# Coding Strategies for Multiple-Access Channels With Feedback and Correlated Sources

Lawrence Ong, *Student Member, IEEE*, and Mehul Motani, *Member, IEEE*

*Abstract*—The multiple-access channel with feedback and correlated sources (MACFCS) models a sensor network in which sensors collect and transmit correlated data to a common sink. We present four achievable rate regions and a capacity outer bound for the MACFCS. For the first achievable region, we construct a decode–forward based coding strategy. The sources first exchange their data, and then cooperate to send full information to the destination. We term this strategy full decoding at sources with decode-forward (FDS-DF). For two of the other achievable regions, we first perform Slepian–Wolf coding to remove the correlation among the source data. This is followed by either i) a compress–forward based coding strategy for the multiple-access channel with feedback, or ii) an existing coding strategy for the multiple-access channel. We also find another achievable region using a multihop coding strategy, which only uses point-to-point coding (no cooperation). From numerical computations, we see that different strategies perform better under certain source correlation structures and network topologies. More specifically, FDS-DF approaches the capacity when i) the inter-source distance decreases, or ii) the correlation among the sources gets higher. Furthermore, the cooperative coding strategies considered support larger achievable rate regions than the noncooperative multihop strategy.

*Index Terms*—Achievable rates, capacity, correlated sources, generalized feedback, multiple-access channel, multiterminal networks.

## I. INTRODUCTION

**W**E investigate the multiple-access channel with feedback and correlated sources (MACFCS) [1], [2]. This channel is a combination of the multiple-access channel with correlated sources (MACCS) and the multiple-access channel with feedback (MACF). The MACFCS serves as a model for the wireless sensor network in which multiple sources send possibly correlated data to a single destination. At the same time, each source receives feedback from the channel and we allow each node to receive different feedback.

### A. Related Work

The MACCS (with a common part) was studied by Slepian and Wolf [3], who derived an achievable region. In their paper, separate source coding and channel coding are used, where

the source coding is first performed to remove the correlation among the sources. The channel coding for the multiple-access channel (MAC) with independent sources is then employed. The MACCS (with possibly no common part) was considered by Cover *et al.* [4]. They showed, by using a simple example, that separating source and channel coding is not optimal. They derived an achievable region for the MACCS using a combined source–channel coding strategy to preserve the correlation among the channel inputs. Outer bounds on the capacity of the MACCS were derived with infinite letter characterization by Cover *et al.* [4] and later improved by Kang and Ulukus [5] to finite-letter expressions. While [3] assumes a certain structure for the correlation among the sources, we study arbitrarily correlated sources in this paper.

The MACF (with independent sources) was investigated by Cover and Leung [6], who derived an achievable region assuming all nodes receive common feedback. Ozarow [7] found the capacity of the Gaussian MACF with common feedback and derived a capacity outer bound for the discrete memoryless MACF with common feedback. King [8] investigated the MACF with all sources receiving common feedback, which is possibly different from what the destination receives, and derived an achievable region for the channel. Willems [9] and Carleial [10] further generalized the MACF with common channel feedback to the case where each node receives possibly different channel feedback, and derived achievable regions of the channel. Sendonaris *et al.* [11], [12] considered the Gaussian MACF with different feedback to different nodes. They derived an information-theoretic achievable region based on cooperation among the source nodes, and showed how the cooperation scheme can be implemented in a practical code-division multiple-access system.

Combining the MACF and the MACCS, we arrive at the MACFCS. One practical system modeled by the MACFCS is a sensor network in which every sensor is capable of transmitting as well as receiving, and each sensor collects data and aims to send them to a single destination. We note that the data collected by the sensor nodes might be correlated, e.g., if they are located close to one another.

Applying coding strategies designed for the MACF or the MACCS might be suboptimal for the MACFCS. Coding strategies for the MACF ignore the correlation among the sources, while coding strategies for the MACCS ignore the feedback from the channel to the sources. Taking both these extra pieces of information into account can help to enlarge the achievable region.

Murugan *et al.* [13] investigated the Gaussian MACFCS with a total average power constraint on the sources. Their coding approach is based on time-division multiple access (TDMA) with

the nodes operating in half-duplex. Our work differs from [13] in that we consider a general MACFCS (including both discrete memoryless channels and Gaussian channels) with full-duplex nodes, in which the source nodes can transmit and receive simultaneously. For the Gaussian case, we impose average power constraints on individual sources, rather than a total average power constraint. King [8] considered the MACFCS with each source observing an independent private message, all sources observing a common message, and all nodes (all sources and the destination) receiving the same feedback. In this paper, we consider arbitrary source correlation and possibly different feedback to all nodes.

Now, we briefly describe the different coding strategies investigated in this paper.

### B. Coding Strategies for the MACFCS

There are numerous coding strategies which we can apply to the MACFCS. The aim of this paper is not to list all of them, but to compare different strategies and to study their strengths and weaknesses. In this paper, we study the following coding strategies for the MACFCS.

i) *Full Decoding at Sources with Decode–Forward Channel Coding (FDS-DF):* In FDS-DF, the general idea is for the sources to communicate so that every source has the complete data of the other sources. They then cooperate to send the combined data to the destination. Since the data of different nodes are correlated, a node does not need to send all its data to other nodes for them to fully decode the data.

ii) *Source Coding for Correlated Sources and Compress–Forward Channel Coding for the MACF (SC-CF):* Source coding for correlated sources [3] is first performed at every source node to remove the correlation among the sources. At this point, we have turned the problem into that of channel coding for the MACF with independent sources. We then construct a coding strategy for the MACF based on the compress–forward technique to transport the independent data to the destination.

iii) *Source Coding for Correlated Sources and MAC Channel Coding (SC-MAC):* Source coding with correlated sources is performed at individual nodes. Then we use a channel coding for the MAC [14], [15] to send the independent data to the destination. In this case, we disregard the feedback from the channel to the source nodes.

iv) *Multihop Coding with Data Aggregation (MH-DA):* The nodes are sequenced (with the last node being the destination) to form a route. Each node (except the first node) decodes the data from the previous node in the route, combines it with its own data, and forwards all data (data that it decodes from the previous node, plus its own data, less the correlated part of the data which the next node already has) to the next node in the route. This continues until the second last node sends all aggregated data to the destination.

*Remark 1:* The first two strategies, i.e., FDS-DF and SC-CF, use coding ideas for the relay channel, in which the relay helps the source to send data to the destination. These two strategies exploit the fact that there is an embedded relay channel in the MACFCS.

*Remark 2:* In SC-MAC, the sources ignore feedback from the channel. Feedback certainly has the potential to increase rates, but taking it into account carries with it a certain amount of complexity, both from a hardware and processing viewpoint. This is the motivation for SC-MAC and we find that this simple strategy can actually be better under certain topologies.

*Remark 3:* The first three strategies mentioned above involve multiuser coding (e.g., multipoint-to-multipoint), which requires a certain amount of coordination for synchronization and cooperation. In MH-DA, all transmissions are single-point-to-single-point, i.e., a node only decodes from a node behind it, treating all other transmissions as noise. We note that there are many practical coding schemes available for single-point-to-single-point communication. Through MH-DA, we can study the loss in performance of single-point-to-single-point coding in a multiterminal network.

*Remark 4:* Barros and Servetto [16] consider the problem of communicating correlated sources over a network of independent point-to-point links. The strategy in [16] includes MH-DA as a special case and can be used for the MACFCS with appropriate modifications.

### C. The Value of Cooperation in the MACFCS

In the wireless channel, which is broadcast in nature, every node hears the transmissions of other nodes. It can treat the transmissions as pure noise, or make use of the received transmissions for cooperation. In the coding strategies described above, the nodes *cooperate* in the encoding and decoding of the data. Across the strategies, we find different levels of cooperation.

In all the strategies, we see nodes cooperate in the source coding, i.e., a node takes into account of other nodes (their data or correlation structure) during its data encoding. In FDS-DF, all the nodes send cooperative data (of all sources) to the destination. In SC-CF and SC-MAC, source coding for correlated sources is performed prior to channel coding. We can view this as a form of cooperation in the encoding. In MH-DA, a node receives data from the previous node, combines them with its own data, and sends the aggregated data to the next node. Again, we see cooperation in the encoding of data.

Now, we see how the nodes cooperate in the channel coding, e.g., multiple nodes decode the transmission of a node, and a node decodes the transmissions of multiple nodes. In FDS-DF, when the sources are exchanging data, the destination, overhearing these transmissions, makes use of the transmissions to aid its decoding of the data. In SC-CF, each source hears the transmissions of other sources, quantizes them, bins them, and sends them to the destination. In SC-MAC, though, the sources ignore the transmissions of other nodes, the destination listens to all the source nodes. The coding strategies above involve channel coding for multiple users. In contrast, MH-DA only considers node pairs, i.e., point-to-point coding. Encoding and

decoding are done only for two nodes. Each node only transmits to one node (down the route) and each node only decodes from one node, ignoring all other transmissions. Hence, we see minimum cooperation in MH-DA.

### D. Contributions

In this paper, we address the following questions.
  i) What rates are achievable on the MACFCS and what coding strategies achieve these rates?
 ii) What are the characteristics of different coding strategies for the MACFCS?
iii) How does the study of the MACFCS help us to understand better coding and cooperation in sensor networks?

Our main contributions are as follows.
  i) We derive an outer bound on the capacity of the MACFCS, which turns out to be the cut-set bound [17], [18].
 ii) We construct a new coding strategy for the MACFCS, where the source nodes first exchange information and then cooperate to send full information to the destination. We derive an achievable region using this strategy.
iii) We construct a compress–forward based coding strategy for the MACF, with each node receiving possibly different channel feedback. King [8] derived an achievable region for the MACF with all sources receiving common feedback using combined decode–forward and compress–forward strategies.
 iv) We combine source coding for correlated sources [3] and our newly constructed compress–forward strategy for the MACF, and arrive at a new achievable region for the MACFCS.
  v) We combine existing schemes, i.e., source coding for correlated sources [3] with the MAC channel coding [14], [15] to arrive at another achievable region for the MACFCS.
 vi) We find another achievable region of the MACFCS using a multihop coding strategy.
vii) We compute achievable regions of the different strategies on the additive white Gaussian noise (AWGN)-MACFCS.
viii) We show that certain strategies perform better under certain source correlation structures and channel topologies. More specifically, we observe the following for the symmetrical MACFCS (where the sources are equidistant from the destination, and they have the same amount of private information to send).
    a) When the inter-source links get better than the source–destination links, FDS-DF approaches the capacity outer bound.
    b) When the correlation among the sources gets higher, FDS-DF approaches the capacity outer bound.
  When one source is far away from the destination and another source is closer to the destination, SC-CF gives a better performance compared to FDS-DF and SC-MAC.
 ix) By comparing different coding strategies for the MACFCS, we show the value of cooperation in the multiple-source single-sink sensor network.
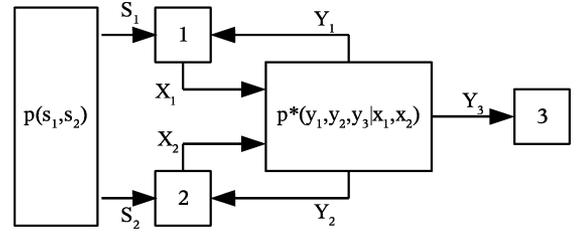


Fig. 1. The three-node MACFCS.

### E. Organization

The remainder of the paper is organized as follows. In Section II, we describe the MACFCS channel model that we will be using throughout the paper. Also, we define the AWGN-MACFCS. In Section III, we derive a capacity outer bound for the MACFCS. These will serve as benchmarks for the coding strategies constructed in Section IV. We compare the performance of different coding strategies on the AWGN-MACFCS in Section V. This is followed by a discussion of the results in Section VI and conclusions in Section VII.

## II. CHANNEL MODEL

Fig. 1 depicts the three-node MACFCS. The three-node discrete memoryless MACFCS is denoted by $\big(\mathcal{S}_1 \times \mathcal{S}_2, p(s_1, s_2), \mathcal{X}_1 \times \mathcal{X}_2, p^*(y_1, y_2, y_3 | x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_3\big)$. $s_1 \in \mathcal{S}_1$, and $s_2 \in \mathcal{S}_2$ are the source messages collected by nodes 1 and 2, respectively, and they are drawn from some discrete bivariate distribution $p(s_1, s_2)$. Here, $\mathcal{S}_1, \mathcal{S}_2, \mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2,$, and $\mathcal{Y}_3$ are seven finite sets. $p^*(y_1, y_2, y_3 | x_1, x_2)$ defines the channel transition probability on $\mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_3$ for each $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$. $x_1$ and $x_2$ are the inputs to the channel from nodes 1 and 2, respectively. $y_1$, $y_2$, and $y_3$ are the channel outputs to nodes 1, 2, and 3 (the destination), respectively.

The channel is memoryless because the current outputs $(y_{1i}, y_{2i}, y_{3i})$ depend on the past inputs $(x_1^i, x_2^i)$ only through the current transmitted symbols $(x_{1i}, x_{2i})$.

*Definition 1:* A block code for the MACFCS consists of an integer $n$, two sets of encoding functions $\{f_{1i}, f_{2i}\}_{i=1}^n$ at nodes 1 and 2, where

$$X_{1i} = f_{1i}(S_1^n, Y_{11}, Y_{12}, \ldots, Y_{1i-1}) \tag{1a}$$
$$X_{2i} = f_{2i}(S_2^n, Y_{21}, Y_{22}, \ldots, Y_{2i-1}) \tag{1b}$$

and a decoding function at node 3, $g : \mathcal{Y}_3^n \to \mathcal{S}_1^n \times \mathcal{S}_2^n$, such that

$$g(Y_3^n) = (\hat{S}_1^n, \hat{S}_2^n) \tag{2}$$

where $\hat{S}_1^n$ and $\hat{S}_2^n$ are estimates of $S_1^n$ and $S_2^n$, respectively.

Without loss of generality, we assume that each encoder knows all $n$ source messages before the encoding of each block. We can also think of the $n$ source messages as one combined message from the source at the beginning of each block of the encoding. Hence, for any choice of codes, the joint

probability mass function on $(\mathcal{S}_1^n, \mathcal{S}_2^n, \mathcal{X}_1^n, \mathcal{X}_2^n, \mathcal{Y}_1^n, \mathcal{Y}_2^n, \mathcal{Y}_3^n)$ is given by

$$
\begin{aligned}
&p(s_1^n, s_2^n, x_1^n, x_2^n, y_1^n, y_2^n, y_3^n) \\
&= \prod_{i=1}^{n} p(s_{1i}, s_{2i}) \prod_{i=1}^{n} p(x_{1i} | s_1^n, y_{11}, y_{12}, \ldots, y_{1i-1}) \\
&\quad \cdot p(x_{2i} | s_2^n, y_{21}, y_{22}, \ldots, y_{2i-1}) \cdot p^*(y_{1i}, y_{2i}, y_{3i} | x_{1i}, x_{2i}). \quad \text{(3a)}
\end{aligned}
$$

*Definition 2:* The error probability is defined as

$$
\begin{aligned}
P_e &= \Pr\{(\hat{S}_1^n, \hat{S}_2^n) \neq (S_1^n, S_2^n)\} \quad &\text{(4a)} \\
&= \sum_{(s_1^n, s_2^n) \in \mathcal{S}_1^n \times \mathcal{S}_2^n} p(s_1^n, s_2^n) \\
&\quad \cdot \Pr\{(\hat{S}_1^n, \hat{S}_2^n) \neq (s_1^n, s_2^n) | (S_1^n, S_2^n) = (s_1^n, s_2^n)\}. \quad &\text{(4b)}
\end{aligned}
$$

*Definition 3:* We say that $(S_1, S_2)$ can be reliably transmitted to the destination per channel use if for any $\epsilon > 0$, there exists a sequence of block codes $\{\{f_{1i}, f_{2i}\}_{i=1}^n, g\}$ such that $P_e < \epsilon$.

We define an *achievable region* of the MACFCS as a set of triplets $[H(S_1|S_2), H(S_2|S_1), H(S_1, S_2)]$ for which we can reliably send $(S_1, S_2)$ to the destination per channel use, for some $p^*(y_1, y_2, y_3|x_1, x_2)$. The *capacity region* is the closure of the set of all achievable regions.

### A. The Gaussian MACFCS

In the $T$-node AWGN-MACFCS, the channel output to node $t$ is given by

$$
Y_t = \sum_{\substack{i=1 \\ i \neq t}}^{T-1} \sqrt{\kappa d_{it}^{-\eta}} X_i + Z_t, \qquad t = 1, 2, \ldots, T \quad \text{(5)}
$$

where $Y_t$ is the received signal at node $t$, $X_i$ is the signal transmitted by node $i$, with power constraint $E[X_i^2] \leq P_i$, and $Z_t$ is independent, zero-mean, Gaussian random variable at the receiver of node $t$ with variance $N_t$. We assume perfect echo cancellation, meaning that a node can cancel its own transmission perfectly. We assume the standard path loss model for signal propagation, in which $\eta$ is the path-loss exponent ($\eta \geq 2$ with equality for free space transmission) and $\kappa$ is a positive constant. $d_{it}$ is the Euclidean distance between nodes $i$ and $t$.

In Section V, where we compare different strategies for the AWGN-MACFCS, we use an alternative but useful definition of achievable region. The reason is that regions in the three-dimensional space are difficult to plot and compare. Hence, we use the following version of achievable region for the AWGN-MACFCS. With a fixed correlation structure $H(S_1|S_2)$, $H(S_2|S_1)$, and $H(S_1, S_2)$, and node positions, an achievable region is the set of average transmit power pairs $[P_1, P_2]$ for which we can reliably send $(S_1, S_2)$ to the destination per channel use. Similarly, the capacity is defined as the closure of the set of all achievable regions. Note that we can convert an achievable region for the discrete memoryless MACFCS to that for the AWGN-MACFCS.

## III. CAPACITY OUTER BOUND

In this section, we derive an outer bound on the capacity of the MACFCS.

*Theorem 1 (Cut-Set Outer Bound):* Let $(\mathcal{S}_1 \times \mathcal{S}_2, p(s_1, s_2), \mathcal{X}_1 \times \mathcal{X}_2, p^*(y_1, y_2, y_3|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_3)$ be a discrete memoryless three-node MACFCS. The source symbols $(S_1, S_2)$ can be reliably transmitted to the destination per channel use only if

$$
[H(S_1|S_2), H(S_2|S_1), H(S_1, S_2)] \in \boldsymbol{R} \quad \text{(6)}
$$

where

$$
\boldsymbol{R} = \left\{ [R_1, R_2, R_3] : \left\{ \begin{array}{l} R_1 \leq I(X_1; Y_2, Y_3|X_2) \\ R_2 \leq I(X_2; Y_1, Y_3|X_1) \\ R_3 \leq I(X_1, X_2; Y_3) \end{array} \right\} \right\} \quad \text{(7)}
$$

for some

$$
p(x_1, x_2). \quad \text{(8)}
$$

In other words, an outer bound on the capacity of the MACFCS is given by $\boldsymbol{R}_{\text{OB}} = \bigcup_{\mathcal{P}} \boldsymbol{R}$, where $\mathcal{P}$ is the set of all distributions satisfying (8), and $\bigcup$ is the union of sets operator.

*Remark 5:* We call the above outer bound the cut-set outer bound (CS-OB) as it turns out to be a special case of the cut-set argument by Gastpar [17], [18]. Now, we start with the cut-set argument and see how it simplifies to the CS-OB. We partition the network into two sets, with a *cut* separating the sets. We assume that all nodes in each set can fully cooperate. We obtain bounds by associating each cut with a corresponding point-to-point system. Consider the cut separating the sets $\{1\}$ and $\{2, 3\}$. The transmission rate from node 1 to nodes 2 and 3 is bounded by the corresponding point-to-point system $X_1 \rightarrow (Y_2, Y_3)|X_2$ (using the notation in [17]). In this point-to-point channel, node 3 receives side information $S_2$ from node 2. For node 3 to reliably decode $S_1$, node 1 needs to transmit at least $H(S_1|S_2)$ bits across the cut, to node 3. Hence, we get $H(S_1|S_2) \leq \max_{p(x_1,x_2)} I(X_1; Y_2, Y_3|X_2)$ (cf. [17, eq. (3.9)]). Applying this argument to the cut separating $\{2\}$ and $\{1, 3\}$, we obtain the second inequality in (7). Consider the cut separating $\{1, 2\}$ and $\{3\}$. We need to transmit $(S_1, S_2)$ across the cut, and the transmission rate is bounded by the corresponding point-to-point system $(X_1, X_2) \rightarrow (Y_3)$. Hence, we get $H(S_1, S_2) \leq \max_{p(x_1,x_2)} I(X_1, X_2; Y_3)$ (cf. [17, eq. (3.2)]). Note that for the point-to-point system, feedback does not increase the capacity, and can be ignored.

*Remark 6:* In the AWGN-MACFCS, $\boldsymbol{R}_{\text{OB}}$ can be found by considering only jointly Gaussian input distributions. We can show that choosing Gaussian input distributions maximizes every mutual information expression in (7) [19, Proposition 2]. Hence, in the AWGN-MACFCS

$$
\bigcup_{p(x_1, x_2)} \boldsymbol{R} = \bigcup_{\text{jointly Gaussian } x_1, x_2} \boldsymbol{R}.
$$

*Proof of Theorem 1:* Given any code $\{\{f_{1i}, f_{2i}\}_{i=1}^n, g\}$ for the MACFCS, the probability function on the joint ensemble $S_1^n, S_2^n, X_1^n, X_2^n, Y_1^n, Y_2^n, Y_3^n$ is given by

$$p(s_1^n, s_2^n, x_1^n, x_2^n, y_1^n, y_2^n, y_3^n)$$
$$= \prod_{i=1}^n p(s_{1i}, s_{2i}) \prod_{i=1}^n p(x_{1i}|s_1^n, y_{11}, y_{12}, \ldots y_{1i-1})$$
$$\cdot p(x_{2i}|s_1^n, y_{21}, y_{22}, \ldots, y_{2i-1}) \cdot p^*(y_{1i}, y_{2i}, y_{3i}|x_{1i}, x_{2i}). \tag{9a}$$

By Fano's inequality [20]

$$H(S_1^n, S_2^n|Y_3^n) \le n \log_2 |\mathcal{S}_1 \times \mathcal{S}_2| P_e + 1 \triangleq n\delta_n. \tag{10}$$

Now, we consider $H(S_1|S_2)$.

$$nH(S_1|S_2) = H(S_1^n|S_2^n) \tag{11a}$$
$$= I(S_1^n; Y_3^n|S_2^n) + H(S_1^n|Y_3^n, S_2^n) \tag{11b}$$
$$\le I(S_1^n; Y_3^n|S_2^n) + H(S_1^n, S_2^n|Y_3^n) \tag{11c}$$
$$\le I(S_1^n; Y_3^n|S_2^n) + n\delta_n. \tag{11d}$$

Now

$$I(S_1^n; Y_3^n|S_2^n)$$
$$= H(Y_3^n|S_2^n) - H(Y_3^n|S_1^n, S_2^n) \tag{12a}$$
$$= \sum_{i=1}^n \left[ H(Y_{3i}|S_2^n, Y_3^{i-1}) - H(Y_{3i}|S_1^n, S_2^n, Y_3^{i-1}) \right] \tag{12b}$$
$$\le \sum_{i=1}^n \left[ H(Y_{3i}) - H(Y_{3i}|X_{1i}, X_{2i}, S_1^n, S_2^n, Y_3^{i-1}) \right] \tag{12c}$$
$$= \sum_{i=1}^n \left[ H(Y_{3i}) - H(Y_{3i}|X_{1i}, X_{2i}) \right] \tag{12d}$$
$$= \sum_{i=1}^n I(X_{1i}, X_{2i}; Y_{3i}), \tag{12e}$$

where
- (12b) and (12e) are by the chain rule;
- (12c) is because conditioning reduces entropy;
- (12d) is because of the memoryless channel, $Y_{3i}$ and $(Y_3^{i-1}, S_1^n, S_2^n)$ are independent given $(X_{1i}, X_{2i})$.

Also

$$I(S_1^n; Y_3^n|S_2^n) \le I(S_1^n; Y_2^n, Y_3^n|S_2^n) \tag{13a}$$
$$= \sum_{i=1}^n I(S_1^n; Y_{2i}, Y_{3i}|S_2^n, Y_2^{i-1}, Y_3^{i-1}) \tag{13b}$$
$$= \sum_{i=1}^n I(Y_{2i}, Y_{3i}; S_1^n|X_{2i}, S_2^n, Y_2^{i-1}, Y_3^{i-1}) \tag{13c}$$
$$= \sum_{i=1}^n \big[ H(Y_{2i}, Y_{3i}|X_{2i}, S_2^n, Y_2^{i-1}, Y_3^{i-1})$$
$$\qquad - H(Y_{2i}, Y_{3i}|X_{2i}, S_1^n, S_2^n, Y_2^{i-1}, Y_3^{i-1}) \big] \tag{13d}$$

$$\le \sum_{i=1}^n \big[ H(Y_{2i}, Y_{3i}|X_{2i})$$
$$\qquad - H(Y_{2i}, Y_{3i}|X_{1i}, X_{2i}, S_1^n, S_2^n, Y_2^{i-1}, Y_3^{i-1}) \big] \tag{13e}$$
$$= \sum_{i=1}^n \left[ H(Y_{2i}, Y_{3i}|X_{2i}) - H(Y_{2i}, Y_{3i}|X_{1i}, X_{2i}) \right] \tag{13f}$$
$$= \sum_{i=1}^n I(X_{1i}; Y_{2i}, Y_{3i}|X_{2i}) \tag{13g}$$

where
- (13b) is by the chain rule and the memoryless nature of the channel;
- (13c) is because $X_{2i}$ is a function of $(S_2^n, Y_2^{i-1})$;
- (13e) is because conditioning reduces entropy;
- (13f) is because $(Y_{2i}, Y_{3i})$ and $(S_1^n, S_2^n, Y_2^{i-1}, Y_3^{i-1})$ are independent given $(X_{1i}, X_{2i})$.

In addition,

$$nH(S_1, S_2) = H(S_1^n, S_2^n) \tag{14a}$$
$$= I(S_1^n, S_2^n; Y_3^n) + H(S_1^n, S_2^n|Y_3^n) \tag{14b}$$
$$\le I(S_1^n, S_2^n; Y_3^n) + n\delta_n. \tag{14c}$$

Consider $I(S_1^n, S_2^n; Y_3^n)$:

$$I(S_1^n, S_2^n; Y_3^n)$$
$$= \sum_{i=1}^n I(S_1^n, S_2^n; Y_{3i}|Y_3^{i-1}) \tag{15a}$$
$$= \sum_{i=1}^n \left[ H(Y_{3i}|Y_3^{i-1}) - H(Y_{3i}|S_1^n, S_2^n, Y_3^{i-1}) \right] \tag{15b}$$
$$\le \sum_{i=1}^n \left[ H(Y_{3i}) - H(Y_{3i}|X_{1i}, X_{2i}, S_1^n, S_2^n, Y_3^{i-1}) \right] \tag{15c}$$
$$= \sum_{i=1}^n \left[ H(Y_{3i}) - H(Y_{3i}|X_{1i}, X_{2i}) \right] \tag{15d}$$
$$= \sum_{i=1}^n I(X_{1i}, X_{2i}; Y_{3i}). \tag{15e}$$

Now we introduce a new variable $Q$ independent of $S_1^n, S_2^n, X_1, X_2, Y_1, Y_2, Y_3$ [21] that takes values in the set $\{1, 2, \ldots, n\}$ with probability

$$\Pr\{Q = i\} = \frac{1}{n}, \quad i \in \{1, 2, \ldots, n\} \tag{16}$$

and such that

$$Q \to (S_1^n, S_2^n) \to (X_{1Q}, X_{2Q}) \to (Y_{1Q}, Y_{2Q}, Y_{3Q}) \tag{17}$$

forms a Markov chain. We set

$$X_1 \triangleq X_{1Q}, X_2 \triangleq X_{2Q}, Y_1 \triangleq Y_{1Q}, Y_2 \triangleq Y_{2Q}, Y_3 \triangleq Y_{3Q}. \tag{18}$$

Now, (12e) becomes

$$\sum_{i=1}^n I(X_{1i}, X_{2i}; Y_{3i}) = \sum_{i=1}^n I(X_{1Q}, X_{2Q}; Y_{3Q}|, Q=i) \tag{19a}$$

$$= nI(X_1, X_2; Y_3|Q) \tag{19b}$$
$$\leq nI(X_1, X_2; Y_3). \tag{19c}$$

Similarly

$$\sum_{i=1}^{n} I(X_{1i}; Y_{2i}, Y_{3i}|X_{2i}) \leq nI(X_1; Y_2, Y_3|X_2). \tag{20}$$

Taking the limit as $n \to \infty$ and $P_e \to 0$, and combining (11d), (12e), (13g), (19c), and (20), we have

$$H(S_1|S_2) \leq \min\{I(X_1, X_2; Y_3), I(X_1; Y_2, Y_3|X_2)\}. \tag{21}$$

By symmetry, we can show that

$$H(S_2|S_1) \leq \min\{I(X_2, X_1; Y_3), I(X_2; Y_1, Y_3|X_1)\}. \tag{22}$$

Combining (14c), (15e), and (19c), we have

$$H(S_1, S_2) \leq I(X_1, X_2; Y_3). \tag{23}$$

Equation (23) guarantees

$$H(S_1|S_2) \leq I(X_1, X_2; Y_3) \tag{24a}$$
$$H(S_2|S_1) \leq I(X_1, X_2; Y_3). \tag{24b}$$

Hence, we have Theorem 1. $\qquad\square$

## IV. ACHIEVABILITY

Now, we present four achievable regions for the three-node MACFCS using four different coding strategies.

### A. Full Decoding at Sources With Decode–Forward Channel Coding (FDS-DF)

In this strategy, every node decodes the data from all other nodes, and all nodes cooperate to send combined data to the destination. We note that for the nodes to cooperate, they must first agree on the data to be sent. In order to do this, each of them must first decode the data from all other nodes.

In brief, this strategy does the following. Since $S_1$ and $S_2$ are correlated, using [22, Theorem 2], node 1 only needs to send $H(S_1|S_2)$ compressed bits to node 2 for it to decode $S_1$. Node 2 does the same. Now, both nodes have $S_1$ and $S_2$. They then cooperate to transmit the full information, i.e., $(S_1, S_2)$, to the destination. At the same time, nodes 1 and 2 send the next (new) message to each other.

Murugan *et al.* [13] proposed a similar coding scheme where the transmissions are split into two phases. In the first phase, the source nodes communicate with each other using TDMA. At the end of the first phase, each source has the data of all nodes. In the second phase, all sources cooperate to transmit to the destination. In this paper, we offer a more general coding scheme. Each source node transmits cooperative information of the previous block (data that it decodes from other nodes together with its own data) and new information (which is to be decoded by other sources and the destination) simultaneously. Since all nodes agree on the same fully decoded information of

the previous block, *coherent combining* can be achieved in the AWGN channel. We show that the coding strategy proposed by [13] is a special case of ours.

Using FDS-DF, we can show that the region given in the following theorem is achievable.

*Theorem 2 (FDS-DF):* Let $\big(S_1 \times S_2, p(s_1, s_2), \mathcal{X}_1 \times \mathcal{X}_2, p^*(y_1, y_2, y_3|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_3\big)$ be a discrete memoryless three-node MACFCS. $(S_1, S_2)$ can be reliably transmitted to the destination per channel use if the following conditions hold:

$$H(S_1|S_2) < \min[I(X_1; Y_2|Q, W_0, W_1, W_2, X_2),$$
$$I(W_1; Y_3|Q, W_0, W_2)$$
$$+ I(X_1; Y_3|Q, W_0, W_1, W_2, X_2)] \tag{25a}$$
$$H(S_2|S_1) < \min[I(X_2; Y_1|Q, W_0, W_1, W_2, X_1),$$
$$I(W_2; Y_3|Q, W_0, W_1)$$
$$+ I(X_2; Y_3|Q, W_0, W_1, W_2, X_1)] \tag{25b}$$
$$I(S_1; S_2) < I(W_0; Y_3|Q, W_1, W_2) \tag{25c}$$
$$H(S_1) < I(W_0, W_1; Y_3|Q, W_2)$$
$$+ I(X_1; Y_3|Q, W_0, W_1, W_2, X_2) \tag{25d}$$
$$H(S_2) < I(W_0, W_2; Y_3|Q, W_1)$$
$$+ I(X_2; Y_3|Q, W_0, W_1, W_2, X_1) \tag{25e}$$
$$H(S_1|S_2) + H(S_2|S_1) < I(W_1, W_2; Y_3|Q, W_0)$$
$$+ I(X_1, X_2; Y_3|Q, W_0, W_1, W_2) \tag{25f}$$
$$H(S_1, S_2) < I(X_1, X_2; Y_3|Q) \tag{25g}$$

where

$$p(q, x_1, x_2, y_1, y_2, y_3, w_0, w_1, w_2) = p(q)p(w_0|q)p(w_1|q)$$
$$\cdot p(w_2|q)p(x_1|q, w_0, w_1, w_2)p(x_2|q, w_0, w_1, w_2)$$
$$\cdot p^*(y_1, y_2, y_3|x_1, x_2, x_3). \tag{26a}$$

$W_0 \in \mathcal{W}_0$, $W_1 \in \mathcal{W}_1$, and $W_2 \in \mathcal{W}_2$ are auxiliary random variables with cardinalities $|\mathcal{W}_0| \times |\mathcal{W}_1| \times |\mathcal{W}_2| \leq \min\{|\mathcal{X}_1| \times |\mathcal{X}_2|, |\mathcal{Y}_1|, |\mathcal{Y}_2|, |\mathcal{Y}_3|\}$. $Q \in \mathcal{Q}$ is a time-sharing variable which determines the portion of time we use a particular distribution $p_1(x_1, x_2, y_1, y_2, y_3, w_0, w_1, w_2)$, $p_2(x_1, x_2, y_1, y_2, y_3, w_0, w_1, w_2)$, and so on. Here, $|\mathcal{Q}| \leq 3$.

*Remark 7:* We note that by setting

$$Q = 0 : W_0 = 0, W_1 = 0, W_2 = 0, X_2 = 0, \tag{27a}$$
$$Q = 1 : W_0 = 0, W_1 = 0, W_2 = 0, X_1 = 0, \tag{27b}$$
$$Q = 2 : X_1 = f(W_0, W_1, W_2), X_2 = f(W_0, W_1, W_2) \tag{27c}$$

for some deterministic function $f(\cdot)$, we end up with the half-duplex coding scheme proposed in [13]. At time $Q = 0$, node 1 transmits, and at time $Q = 1$, node 2 transmits. After both nodes fully decode the messages from each other, they coherently transmit at time $Q = 2$. However, in [13], the destination only decodes at time $Q = 2$. In other words, the terms $I(\cdots; Y_3|Q = i, \ldots)$ for $i = 1, 2$ are excluded. In Theorem 2,
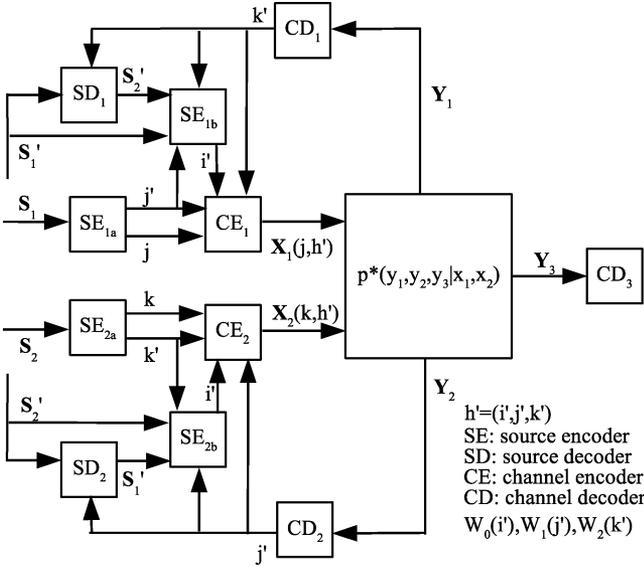
Fig. 2. Coding for the MAC with feedback and correlated sources using the decode–forward strategy.

the destination decodes at all $Q = 0, 1, 2$ and hence the achievable region can be larger.

*Outline of the Proof For Theorem 2:* Now, we present an outline of the proof for Theorem 2. The complete proof can be found in Appendix A. We ignore $Q$ in the following discussion to simplify the expressions.

The codebook generation is as follows.

i) Fix the probability mass functions $p(w_0), p(w_1), p(w_2)$, $p(x_1|w_0, w_1, w_2)$, and $p(x_2|w_0, w_1, w_2)$.

ii) Generate $2^{n[I(S_1;S_2)+\epsilon]}$ independent and identically distributed (i.i.d.) sequences $\boldsymbol{w}_0$ according to $\prod_{i=1}^{n} p(w_{0i})$. Index them $\boldsymbol{w}_0(i)$, $i \in \{1, 2, \ldots, 2^{n[I(S_1;S_2)+\epsilon]}\}$.

iii) Generate $2^{n[H(S_1|S_2)+\epsilon]}$ i.i.d. sequences $\boldsymbol{w}_1$ according to $\prod_{i=1}^{n} p(w_{1i})$. Index them $\boldsymbol{w}_1(j)$, $j \in \{1, 2, \ldots, 2^{n[H(S_1|S_2)+\epsilon]}\}$.

iv) Generate $2^{n[H(S_2|S_1)+\epsilon]}$ i.i.d. sequences $\boldsymbol{w}_2$ according to $\prod_{i=1}^{n} p(w_{2i})$. Index them $\boldsymbol{w}_2(k)$, $k \in \{1, 2, \ldots, 2^{n[H(S_2|S_1)+\epsilon]}\}$.

v) Define $h' = (i', j', k')$. For each $(\boldsymbol{w}_0(i'), \boldsymbol{w}_1(j'), \boldsymbol{w}_2(k'))$, generate $2^{n[H(S_1|S_2)+\epsilon]}$ sequences $\boldsymbol{x}_1$ according to $\prod_{i=1}^{n} p(x_{1i}|w_{0i}(i'), w_{1i}(j'), w_{2i}(k'))$. Index them $\boldsymbol{x}_1(j, h')$, $j \in \{1, 2, \ldots, 2^{n[H(S_1|S_2)+\epsilon]}\}$.

vi) Again for each $(\boldsymbol{w}_0(i'), \boldsymbol{w}_1(j'), \boldsymbol{w}_2(k'))$, independently generate $2^{n[H(S_2|S_1)+\epsilon]}$ sequences $\boldsymbol{x}_2$ according to $\prod_{i=1}^{n} p(x_{2i}|w_{0i}(i'), w_{1i}(j'), w_{2i}(k'))$. Index them $\boldsymbol{x}_2(k, h')$, $k \in \{1, 2, \ldots, 2^{n[H(S_2|S_1)+\epsilon]}\}$.

The encoding steps (refer to Fig. 2) are as follows.

i) Slepian and Wolf [3, Theorem 2] showed that when node 1 only knows $\boldsymbol{s}_1$ and node 2 knows $\boldsymbol{s}_2$, node 1 can encode $\boldsymbol{s}_1$ using $n[H(S_1|S_2) + \epsilon]$ bits (indexed by $j$) and it can be decoded by node 2. Similarly, node 2 can use $n[H(S_2|S_1)+\epsilon]$ bits (indexed by $k$) to encode $\boldsymbol{s}_2$. Node 1 transmits $\boldsymbol{x}_1(j, h')$, and node 2 transmits $\boldsymbol{x}_2(k, h')$, where $h'$ is the cooperative information from the previous block. We use prime to indicate the index from the previous block.

ii) At the beginning of the new block, assume that node 1 correctly estimates $k'$ sent by node 2. Using $\boldsymbol{s}'_1$, it can decode $\boldsymbol{s}'_2$. Node 2 does likewise to decode $\boldsymbol{s}'_1$.

iii) Both sources now compress $(\boldsymbol{s}'_1, \boldsymbol{s}'_2)$ down to $n[H(S_1, S_2) + 3\epsilon]$ bits and index it by $h' \in \{1, \ldots, 2^{n[H(S_1, S_2)+3\epsilon]}\}$. Now, create $2^{n[H(S_1|S_2)+H(S_2|S_1)+2\epsilon]}$ bins and index each bin by a unique $(j', k')$. Assign $h'$ to the bins so that each bin contains $2^{n[I(S_1;S_2)+\epsilon]}$ entries. Index the entries $i' \in \{1, \ldots, 2^{n[I(S_1;S_2)+\epsilon]}\}$. Hence, each $h'$ can be represented by a unique triplet $(i', j', k')$.

iv) In the new block, node 1 sends $\boldsymbol{x}_1(j, i', j', k')$ and node 2 sends $\boldsymbol{x}_2(k, i', j', k')$.

The decoding steps are as follows.

i) Upon observing the sequence $\boldsymbol{y}_1$, node 1 declares $\hat{k}$ has been sent by node 2 if there exists a unique $\hat{k}$ such that $\left(\boldsymbol{x}_1(j, h'), \boldsymbol{w}_0(i'), \boldsymbol{w}_1(j'), \boldsymbol{w}_2(k'), \boldsymbol{x}_2(\hat{k}, h'), \boldsymbol{y}_1\right) \in \mathcal{A}_\epsilon$. We use hat to indicate the estimate. Here, $\mathcal{A}_\epsilon$ is the set of jointly typical sequences [20, p. 195]. We note that node 1 knows $h' = (i', j', k')$, which is the full information from the previous block, and its own information $j$. It can determine the correct $k$ with diminishing error probability if

$$H(S_2|S_1) < I(X_2; Y_1|W_0, W_1, W_2, X_1). \tag{28}$$

ii) Similarly, observing the sequence $\boldsymbol{y}_2$, node 2 declares $\hat{j}$ has been sent by node 1 if there exists a unique $\hat{j}$ such that $\left(\boldsymbol{x}_1(\hat{j}, h'), \boldsymbol{w}_0(i'), \boldsymbol{w}_1(j'), \boldsymbol{w}_2(k')\boldsymbol{x}_2(k, h'), \boldsymbol{y}_2\right) \in \mathcal{A}_\epsilon$. Node 2 can determine the correct $j$ with diminishing error probability if

$$H(S_1|S_2) < I(X_1; Y_2|W_0, W_1, W_2, X_2). \tag{29}$$

iii) Node 3 decodes $(\hat{i}, \hat{j}, \hat{k})$ over two blocks. In the first block, assuming that it has already correctly decoded $h' = (i', j', k')$ from the previous block, it finds a set of $(\hat{j}, \hat{k}) \in \mathcal{L}_1$ where $\left(\boldsymbol{x}_1(\hat{j}, h'), \boldsymbol{x}_2(\hat{k}, h'), \boldsymbol{w}_0(i'), \boldsymbol{w}_1(j'), \boldsymbol{w}_2(k'), \boldsymbol{y}_3\right) \in \mathcal{A}_\epsilon$. In the second block, it then finds another set of $(\hat{j}, \hat{k}) \in \mathcal{L}_2$ and a unique $\hat{i}$ where $\left(\boldsymbol{w}_0(\hat{i}), \boldsymbol{w}_1(\hat{j}), \boldsymbol{w}_2(\hat{k}), \boldsymbol{y}_3\right) \in \mathcal{A}_\epsilon$. It declares $(\hat{i}, \hat{j}, \hat{k})$ has been sent if there is a unique $\hat{i}$ and a unique pair of $(\hat{j}, \hat{k})$ in $\mathcal{L}_1 \cap \mathcal{L}_2$. This can be done with diminishing error probability if

$$I(S_1; S_2) < I(W_0; Y_3|W_1, W_2) \tag{30a}$$

$$H(S_1|S_2) < I(W_1; Y_3|W_0, W_2)$$
$$+ I(X_1; Y_3|W_0, W_1, W_2, X_2) \tag{30b}$$

$$H(S_2|S_1) < I(W_2; Y_3|W_0, W_1)$$
$$+ I(X_2; Y_3|W_0, W_1, W_2, X_1) \tag{30c}$$

$$H(S_1) < I(W_0, W_1; Y_3|W_2)$$
$$+ I(X_1; Y_3|W_0, W_1, W_2, X_2) \tag{30d}$$

$$H(S_2) < I(W_0, W_2; Y_3|W_1)$$
$$+ I(X_2; Y_3|W_0, W_1, W_2, X_1) \tag{30e}$$

$$H(S_1|S_2) + H(S_2|S_1) < I(W_1, W_2; Y_3|W_0)$$
$$+ I(X_1, X_2; Y_3|W_0, W_1, W_2) \tag{30f}$$

$$H(S_1, S_2) < I(X_1, X_2; Y_3). \tag{30g}$$

We consider all possible error combinations. Assuming that $(i, j, k)$ were sent, (30a) guarantees that the $\Pr(\hat{i} \neq i, \hat{j} = j, \hat{k} = k) < \epsilon$ for any $\epsilon > 0$. Equation (30b) guarantees that $\Pr(\hat{i} = i, \hat{j} \neq j, \hat{k} = k) < \epsilon$, (30c) guarantees that $\Pr(\hat{i} = i, \hat{j} = j, \hat{k} \neq k) < \epsilon$, (30d) guarantees that $\Pr(\hat{i} \neq i, \hat{j} \neq j, \hat{k} = k) < \epsilon$, (30e) guarantees that $\Pr(\hat{i} = i, \hat{j} = j, \hat{k} \neq k) < \epsilon$, (30f) guarantees that $\Pr(\hat{i} = i, \hat{j} \neq j, \hat{k} \neq k) < \epsilon$, and (30g) guarantees that $\Pr(\hat{i} \neq i, \hat{j} = j, \hat{k} \neq k) < \epsilon$.

iv) With $(\hat{i}, \hat{j}, \hat{k})$, node 3 can determine $\hat{h}$ and decode $(\hat{s}_1, \hat{s}_2)$.

The total probability of error can be bounded, for large $n$, if (28), (29), and (30a)–(30g) hold.

For the cardinality of the auxiliary random variables, using the method in [23], we can show that $|\mathcal{W}_0| \times |\mathcal{W}_1| \times |\mathcal{W}_1| \leq \min\{|\mathcal{X}_1| \times |\mathcal{X}_2|, |\mathcal{Y}_1|, |\mathcal{Y}_2|, |\mathcal{Y}_3|\}$. Since the achievable region of FDS-DF can be plotted in the three-dimensional space, $|\mathcal{Q}| \leq 3$ is sufficient.

Hence, we have Theorem 2. □

The probability error analysis can be found in Appendix A. The achievable region of FDS-DF on the AWGN-MACFCS can be found in Appendix B.

*Remark 8:* In FDS-DF, the sources only need to exchange $nH(S_1|S_2) + nH(S_2|S_1)$ bits in order for them to know the full information $(s_1, s_2)$. When both sources know the full information, they then cooperate (achieving coherent combining in the AWGN channel) to send the full information to the destination.

Under certain channel conditions, that all nodes fully decode the data of all other nodes might not be desirable. One example is when node 1 is far from the destination and node 2 is close to the destination. In this case, it is not necessary for node 1 to decode all of node 2's data. We note that if the sources only exchange partial information, they are not able to cooperate to send the full information to the destination. They can only cooperate to send the data that they exchange (in contrast with FDS-DF in which the sources can cooperate to send more information than what they exchange). Without full decoding at the sources, we study a few other types of coding strategies where full decoding of all messages $(s_1, s_2)$ only occurs at the destination. We use the following method. First, source coding is performed at each individual source node to remove the correlation among the sources (see Section IV-B). At this point, we have turned the problem into that of channel coding for the MACF with independent sources. Then we apply a channel coding strategy for the MACF to transmit independent information to the destination.

### B. Source Coding for Correlated Sources

Source coding for correlated sources is first performed at every source node. This removes correlation between the sources. This does not require physical communication among the sources. Each source node forms independent inputs to its channel encoder.

Recall that nodes 1 and 2 receive $s_1$ and $s_2$ from their respective sources. The data are correlated and drawn according to $p(s_1, s_2)$. First, we consider a noiseless channel. With node 1 knowing only $s_1$ and node 2 knowing only $s_2$, the destination
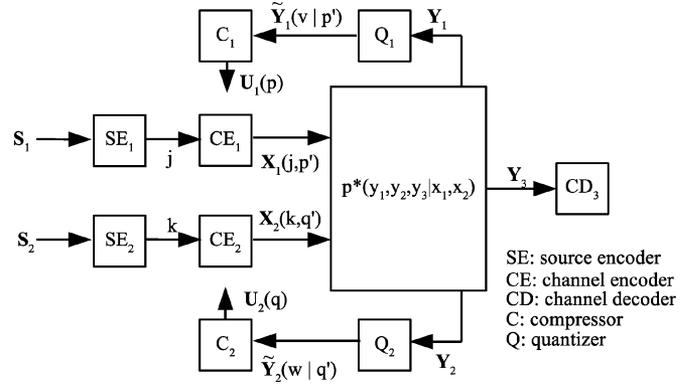


Fig. 3. Coding for the MAC with feedback (with independent sources) using the compress–forward strategy.

can reconstruct $(s_1, s_2)$ reliably if node 1 encodes $s_1$ with rate $R_1$ and node 2 encodes $s_2$ with rate $R_2$ [3], where

$$R_1 \geq H(S_1|S_2) \tag{31a}$$

$$R_2 \geq H(S_2|S_1) \tag{31b}$$

$$R_1 + R_2 \geq H(S_1, S_2). \tag{31c}$$

Fig. 3 shows independent data $(j, k)$ after source coding. After receiving $n$ source messages, $s_1$, encoder 1 encodes the data to $j \in \{1, 2, \ldots, 2^{nR_1}\}$. Encoder 2 receives $s_2$ and encodes the data to $k \in \{1, 2, \ldots, 2^{nR_2}\}$. $R_1$ and $R_2$ are within the constraints (31a)–(31c).

Now, we consider an unreliable channel and explore how channel coding can help the destination to recover $j$ and $k$. With these, it can recover $s_1$ and $s_2$.

### C. Source Coding for Correlated Sources and Compress–Forward Channel Coding for the MACF (SC-CF)

In this subsection, we derive an achievable region for the MACF based on the compress–forward strategy. Combining this with the source coding rate constraints in Section IV-B, we derive another achievable region for the MACFCS. We term this coding strategy source coding for correlated sources and compress–forward channel coding for the MACF (SC-CF). To the best of our knowledge, the compress–forward strategy has not been studied on the MACF where each node receives possibly different channel feedback. The compress–forward strategy was first introduced by Cover and El Gamal [21] for the single-relay channel. It was subsequently extended to the multiple-relay channel by Kramer *et al.* [19] in which the strategy is termed the compress-and-forward strategy. King [8] derived an achievable region for the MACF, with all sources receiving common feedback, using combined decode–forward and compress–forward strategies. In this paper, we construct a compress–forward strategy for the MACF with possibly different feedback to every node. Here, we do not combine the compress–forward strategy with the decode–forward strategy as we want to compare the performance of different strategies. With the different strategies described in this paper, we can easily pick and combine different strategies to get another achievable region.

*Remark 9:* It has been shown [19] that in relay channels with different topologies, the decode–forward strategy (which is also known as the decode-and-forward strategy) or the compress–forward strategy can achieve higher rates. For the Gaussian relay channel, a rough guide is that when the relay is closer to the source, the decode–forward strategy achieves higher rates; while when the relay is closer to the destination, the compress–forward strategy achieves higher rates. This suggests that SC-CF might give larger achievable regions on the MACFCS compared to FDS-DF under different topologies.

Using the compress–forward strategy, each node transmits independent information as well as a quantized and binned version of its received signal. Referring to Fig. 3, $j$ and $k$ are independent information after performing source coding on a block of $n$ correlated source messages $(\boldsymbol{s}_1, \boldsymbol{s}_2)$. Consider node 1 as an example first. From the received signal $\boldsymbol{y}_1$, it produces a quantized version $\tilde{\boldsymbol{y}}_1$. It then bins $\tilde{\boldsymbol{y}}_1$ to $\boldsymbol{u}_1$. In the next block, it sends new information $j$ as well as $\boldsymbol{u}_1$. We can view this as node 1 helping node 2 to send a noisy, quantized, and binned version of node 2's signal, $k$, without needing to fully decode $k$. Node 2 does likewise.

Using SC-CF, we show that the following region is achievable.

*Theorem 3 (SC-CF):* Let $\big(\mathcal{S}_1 \times \mathcal{S}_2, p(s_1, s_2), \mathcal{X}_1 \times \mathcal{X}_2, p^*(y_1, y_2, y_3 | x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_3\big)$ be a discrete memoryless three-node MACFCS. The source messages $(S_1, S_2)$ can be reliably transmitted to the destination per channel use if

$$H(S_1|S_2) < I(X_1; \tilde{Y}_1, \tilde{Y}_2, Y_3 | Q, U_1, U_2, X_2) \quad (32a)$$

$$H(S_2|S_1) < I(X_2; \tilde{Y}_1, \tilde{Y}_2, Y_3 | Q, U_1, U_2, X_1) \quad (32b)$$

$$H(S_1, S_2) < I(X_1, X_2; \tilde{Y}_1, \tilde{Y}_2, Y_3 | Q, U_1, U_2) \quad (32c)$$

where the mutual information is taken over all joint probability mass functions

$$p(q, u_1, u_2, x_1, x_2, \tilde{y}_1, \tilde{y}_2, y_1, y_2, y_3) = p(q)p(u_1|q)p(x_1|q, u_1)$$
$$\cdot p(u_2|q)p(x_2|q, u_2)p(\tilde{y}_1|q, y_1, x_1, u_1)p(\tilde{y}_2|q, y_2, x_2, u_2)$$
$$\cdot p^*(y_1, y_2, y_3 | x_1, x_2) \quad (33a)$$

subject to the following constraints:

$$I(U_1; Y_3 | Q, U_2) > I(\tilde{Y}_1; Y_1 | Q, X_1, U_1)$$
$$- I(\tilde{Y}_1; Y_3 | Q, \tilde{Y}_2, U_1, U_2) \quad (34a)$$

$$I(U_2; Y_3 | Q, U_1) > I(\tilde{Y}_2; Y_2 | Q, X_2, U_2)$$
$$- I(\tilde{Y}_2; Y_3 | Q, \tilde{Y}_1, U_1, U_2) \quad (34b)$$

$$I(U_1, U_2; Y_3 | Q) > I(\tilde{Y}_1; Y_1 | Q, X_1, U_1) + I(\tilde{Y}_2; Y_2 | Q, X_2, U_2)$$
$$- I(\tilde{Y}_1, \tilde{Y}_2; Y_3 | Q, U_1, U_2). \quad (34c)$$

Here, $U_1 \in \mathcal{U}_1, U_2 \in \mathcal{U}_2, \tilde{Y}_1 \in \tilde{\mathcal{Y}}_1$, and $\tilde{Y}_2 \in \tilde{\mathcal{Y}}_2$ are auxiliary random variables. $|\mathcal{U}_1| \times |\mathcal{U}_2| \leq \min\{|\mathcal{X}_1| \times |\mathcal{X}_2|, |\mathcal{Y}_3|\}, |\tilde{\mathcal{Y}}_1|$, and $|\tilde{\mathcal{Y}}_2|$ are finite. $Q \in \mathcal{Q}$ is the time sharing variable, and $|\mathcal{Q}| \leq 3$.

*Outline of the Proof for Theorem 3:* Now, we give a brief outline of the proof for Theorem 3. The error probability analysis can be found in Appendix C. We ignore $Q$ in the following discussion to simplify the expressions.

Fig. 3 shows independent data $(j, k)$ after source coding. Channel encoder 1 receives $j \in \{1, 2, \ldots, 2^{nR_1}\}$ and channel encoder 2 receives $k \in \{1, 2, \ldots, 2^{nR_2}\}$ for every $n$ source messages. Now, we study a channel coding scheme to ensure that the independent data after source coding can be reliably transmitted to the destination. The codebook generation is as follows.

i) Fix $p(u_1), p(x_1|u_1), p(u_2), p(x_2|u_2), p(\tilde{y}_1|y_1, x_1, u_1)$, and $p(\tilde{y}_2|y_2, x_2, u_2)$.

ii) Generate $2^{nR'_1}$ i.i.d. sequences $\boldsymbol{u}_1$ according to $\prod_{i=1}^n p(u_{1i})$. Index them $\boldsymbol{u}_1(p), p \in \{1, \ldots, 2^{nR'_1}\}$. Generate $2^{nR'_2}$ i.i.d. sequences $\boldsymbol{u}_2$ according to $\prod_{i=1}^n p(u_{2i})$. Index them $\boldsymbol{u}_2(q), q \in \{1, \ldots, 2^{nR'_2}\}$.

iii) For each $\boldsymbol{u}_1(p')$, generate $2^{nR_1}$ sequences $\boldsymbol{x}_1$ according to $\prod_{i=1}^n p(x_{1i}|u_{1i}(p'))$. Index them $\boldsymbol{x}_1(j, p')$, $j \in \{1, \ldots, 2^{nR_1}\}$. For each $\boldsymbol{u}_2(q')$, generate $2^{nR_2}$ sequences $\boldsymbol{x}_2$ according to $\prod_{i=1}^n p(x_{2i}|u_{2i}(q'))$. Index them $\boldsymbol{x}_2(k, q'), k \in \{1, \ldots, 2^{nR_2}\}$.

iv) For each $\boldsymbol{u}_1(p')$, generate $2^{n\tilde{R}_1}$ sequences $\tilde{\boldsymbol{y}}_1$ according to $\prod_{i=1}^n p(\tilde{y}_{1i}|u_{1i}(p'))$. We define

$$p(\tilde{y}_1|u_1) =$$
$$\frac{\sum_{u_2, x_1, x_2, \tilde{y}_2, y_1, y_2, y_3} p(u_1, u_2, x_1, x_2, \tilde{y}_1, \tilde{y}_2, y_1, y_2, y_3)}{\sum_{u_2, x_1, x_2, \tilde{y}_1, \tilde{y}_2, y_1, y_2, y_3} p(u_1, u_2, x_1, x_2, \tilde{y}_1, \tilde{y}_2, y_1, y_2, y_3)}$$
$$(35a)$$

where $p(u_1, u_2, x_1, x_2, \tilde{y}_1, \tilde{y}_2, y_1, y_2, y_3)$ is defined in (33). Index them $\tilde{\boldsymbol{y}}_1(v|p'), v \in \{1, \ldots, 2^{n\tilde{R}_1}\}$.

v) Similarly, for each $\boldsymbol{u}_2(q')$, generate $2^{n\tilde{R}_2}$ sequences $\tilde{\boldsymbol{y}}_2$ according to $\prod_{i=1}^n p(\tilde{y}_{2i}|u_{2i}(q'))$. We define

$$p(\tilde{y}_2|u_2) =$$
$$\frac{\sum_{u_1, x_1, x_2, \tilde{y}_1, y_1, y_2, y_3} p(u_1, u_2, x_1, x_2, \tilde{y}_1, \tilde{y}_2, y_1, y_2, y_3)}{\sum_{u_1, x_1, x_2, \tilde{y}_1, \tilde{y}_2, y_1, y_2, y_3} p(u_1, u_2, x_1, x_2, \tilde{y}_1, \tilde{y}_2, y_1, y_2, y_3)} \cdot$$
$$(36a)$$

Index them $\tilde{\boldsymbol{y}}_2(w|q'), w \in \{1, \ldots, 2^{n\tilde{R}_2}\}$.

vi) Randomly partition the set $\{1, 2, \ldots, 2^{n\tilde{R}_1}\}$ into $2^{nR'_1}$ cells $S_p, p \in \{1, \ldots, 2^{nR'_1}\}$; and partition the set $\{1, \ldots, 2^{n\tilde{R}_2}\}$ into $2^{nR'_2}$ cells $S_q, q \in \{1, \ldots, 2^{nR'_2}\}$.

The encoding steps are as follows. Basically, node 1 quantizes its received signal from the previous block and bins it. It sends the binned information together with new information from the source in the new block. Node 2 does likewise.

i) In the beginning of block $t+1$, remembering its previous transmission in block $t$, $\boldsymbol{x}_1(j^t, q^{t-1})$, and $\boldsymbol{u}_1(q^{t-1})$, and observing its received signal in block $t$, $\boldsymbol{y}_1(t)$, it finds a unique $v^t$ for which $(\boldsymbol{x}_1(j^t, p^{t-1}), \boldsymbol{u}_1(p^{t-1}), \boldsymbol{y}_1(t), \tilde{\boldsymbol{y}}_1(v^t|p^{t-1})) \in \mathcal{A}_\epsilon$. Berger [24, Lemma 2.1.3] showed that node 1 can find such a $v^{t-1}$ with probability tending to 1, with a large enough $n$, if

$$\tilde{R}_1 > I(\tilde{Y}_1; Y_1 | X_1, U_1). \quad (37)$$

Here, $v^t$ is the quantized version of $\boldsymbol{y}_1(t)$.

ii) Now, node 1 bins $v^t$ to $p^t$. It finds $p^t$ for which $v^t \in S_{p^t}$. It then sends $\boldsymbol{x}_1(j^{t+1}, p^t)$ in block $t+1$, where $j^{t+1}$ is the new message from the source. Here, $p^t$ is to be decoded

and used by the destination to estimate $v^t$. We see here that node 1 helps node 2 to send a noisy, quantized, and binned version of node 2's signal to the destination.

iii) In block $t + 1$, node 2 quantizes $\boldsymbol{y}_2(t)$ to $w^t$. It can find a unique $w^t$ with probability tending to 1 if

$$\tilde{R}_2 > I(\tilde{Y}_2; Y_2 | X_2, U_2). \tag{38}$$

It bins $w^t$ to $q^t$, where $w^t \in S_{q^t}$. It then sends $\boldsymbol{x}_2(k^{t+1}, q^t)$ in block $t$, where $k^{t+1}$ is the new information.

The decoding steps are as follows. The destination first decodes the quantized and binned information from nodes 1 and 2. It then estimates the quantized information. Using its received signal and the estimated quantized information, it decodes the messages from nodes 1 and 2.

i) At the end of block $t + 1$, the destination receives $\boldsymbol{y}_3(t + 1)$. It declares $(\hat{p}^t, \hat{q}^t)$ were sent by nodes 1 and 2 if it can find a unique pair of $(\hat{p}^t, \hat{q}^t)$ for which $(\boldsymbol{u}_1(\hat{p}^t), \boldsymbol{u}_2(\hat{q}^t), \boldsymbol{y}_3(t + 1)) \in \mathcal{A}_\epsilon$. This can be done with an arbitrarily small error probability if the following inequalities hold:

$$R_1' < I(U_1; Y_3 | U_2) \tag{39a}$$
$$R_2' < I(U_2; Y_3 | U_1) \tag{39b}$$
$$R_1' + R_2' < I(U_1, U_2; Y_3). \tag{39c}$$

ii) At the end of block $t + 1$, assume that the destination has correctly decoded $(p^{t-1}, q^{t-1})$ and $(p^t, q^t)$. It uses its received signal in block $t$ to find a set $\mathcal{L}(t)$ of $(v^t, w^t)$ such that $(\tilde{\boldsymbol{y}}_1(v^t | p^{t-1}), \tilde{\boldsymbol{y}}_2(w^t | q^{t-1}), \boldsymbol{u}_1(p^{t-1}), \boldsymbol{u}_2(q^{t-1}), \boldsymbol{y}_3(t)) \in \mathcal{A}_\epsilon$. It declares that $(\hat{v}^t, \hat{w}^t)$ were sent if it can find a unique $(\hat{v}^t, \hat{w}^t) \in \{(\hat{v}^t, \hat{w}^t) : \hat{v}^t \in S_{p^t} \text{ and } \hat{w}^t \in S_{q^t}\} \cap \mathcal{L}(t)$. This can be done reliably if

$$\tilde{R}_1 < I(\tilde{Y}_1; Y_3 | \tilde{Y}_2, U_1, U_2) + R_1' \tag{40a}$$
$$\tilde{R}_2 < I(\tilde{Y}_2; Y_3 | \tilde{Y}_1, U_1, U_2) + R_2' \tag{40b}$$
$$\tilde{R}_1 + \tilde{R}_2 < I(\tilde{Y}_1, \tilde{Y}_2; Y_3 | U_1, U_2) + R_1' + R_2'. \tag{40c}$$

iii) At the end of block $t + 1$, assume that the destination has correctly decoded $(v^t, w^t)$ and $(p^{t-1}, q^{t-1})$. It uses $\tilde{\boldsymbol{y}}_1(v^t | p^{t-1})$, $\tilde{\boldsymbol{y}}_2(w^t | q^{t-1})$, and $\boldsymbol{y}_3(t)$. It declares $(\hat{j}^t, \hat{k}^t)$ were sent if $(\boldsymbol{x}_1(\hat{j}^t, p^{t-1}), \boldsymbol{x}_2(\hat{k}^t, q^{t-1}), \boldsymbol{u}_1(p^{t-1}), \boldsymbol{u}_2(q^{t-1}), \tilde{\boldsymbol{y}}_1(v^t | p^{t-1}), \tilde{\boldsymbol{y}}_2(\hat{w}^t | q^{t-1}), \boldsymbol{y}_3(t)) \in \mathcal{A}_\epsilon$. This can be done with diminishing error probability if

$$R_1 < I(X_1; \tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2, X_2) \tag{41a}$$
$$R_2 < I(X_2; \tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2, X_1) \tag{41b}$$
$$R_1 + R_2 < I(X_1, X_2; \tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2). \tag{41c}$$

We see that node 3 decodes $(j^t, k^t)$ at the end of block $t + 1$.

Combining these rate constraints for the MACF using the compress–forward strategy and the constraints for the source coding, (31a)–(31c), we get Theorem 3. $\qquad\square$

The probability error analysis can be found in Appendix C. The achievable region of the AWGN-MACFCS using SC-CF can be found in Appendix D.

## D. Source Coding for Correlated Sources and the MAC Channel Coding (SC-MAC)

Now, we consider a coding strategy for the MACF that ignores the feedback from the channel to the source nodes. Each source now simply sends independent messages as it would in the MAC. We call this strategy SC-MAC, and we will see later that it actually does well in certain network topologies. A coding strategy that achieves the capacity of the MAC was found by Liao [14] and Ahlswede [15]. Combining source coding for correlated sources and this channel coding for the MAC, we have the following theorem.

*Theorem 4 (SC-MAC):* Let $(\mathcal{S}_1 \times \mathcal{S}_2, p(s_1, s_2), \mathcal{X}_1 \times \mathcal{X}_2, p^*(y_1, y_2, y_3 | x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_3)$ be a discrete memoryless three-node MACFCS. The source messages $(S_1, S_2)$ can be reliably transmitted to the destination per channel use if the following inequalities hold:

$$H(S_1 | S_2) \leq I(X_1; Y_3 | X_2) \tag{42a}$$
$$H(S_2 | S_1) \leq I(X_2; Y_3 | X_1) \tag{42b}$$
$$H(S_1, S_2) \leq I(X_1, X_2; Y_3) \tag{42c}$$

where $p(x_1, x_2) = p(x_1) p(x_2)$.

## E. Combination of Other Strategies

There is a multitude of ways which we can combine a coding strategy for the MACCS with that for the MACF to arrive at a coding strategy for the MACFCS. The aim of this paper is not to list all of them. In this section, we briefly mention a few combinations.

i) Combining source coding for correlated sources and the *partial decode–forward* channel coding for the MACF by Carleial [10, Theorem 1]: After source coding, each source node has independent data. Each source now exchanges part of their data with other source nodes. They then cooperate to send the exchanged data to the destination. We call that partial decode–forward strategy as every source only decodes part of the data of other sources. An achievable region for the MACFCS can be derived by combining the source-coding constraints for correlated sources (constraint (31a)–(31c) in Section IV-B) and the channel coding constraints of the partial decode–forward strategy for the MACF [10, Constraints (3a), (3b), (7a)–(7q)].

ii) Combining source coding for correlated sources and the partial decode–forward channel coding for the MACF by Willems [9, Theorem 7.1]: Similar to [10], the sources exchange part of their data through the channel feedback link. They then cooperate to send the exchanged data to the destination. An achievable region for the MACFCS can be derived by combining the source-coding constraints for correlated sources (constraint (31a)–(31c) in Section IV-B) and the channel-coding constraints of the partial decode–forward strategy for the MACF [9, Theorem 7.1].

iii) Combining coding strategy by Cover *et al.* for MACCS without common part [4, Constraints (3)] and the compress–forward strategy for the MACF (that we derived in Section IV-C): Each node sends information encoded directly from the source (so that correlation is preserved among the transmitted signals) as well as the received (via the feedback links), quantized, and binned signals from other nodes.

iv) Combining coding strategy by Cover *et al.* for MACCS without common part [4, Constraints (3)] and the partial decode–forward strategy by Carleial [10, Theorem 1] or Willems [9, Theorem 7.1]: Each node sends information encoded directly from the source (so that correlation is preserved among the transmitted signals). At the same time, the source nodes partially decode the data from other nodes, and cooperate to send the exchanged data to the destination.

*Remark 10:* The strategy mentioned in i) above (in Section IV-E) is different from FDS-DF (in Theorem 2). In the former, the channel encoders at nodes 1 and 2 receive independent data stream, after performing source coding for correlated sources. Then, Carleial's technique for MACF is applied directly. Hence, if we want the nodes to cooperatively send the full information $(S_1^n, S_2^n)$, they must exchange at least $nH(S_1, S_2)$ bits. In FDS-DF, however, the channel encoders receives correlated data from the sources (so, we do not apply Carleial's technique directly here), and so they only need to exchange $nH(S_1|S_2) + nH(S_2|S_1)$ bits to be able to cooperatively send the full information.

*Remark 11:* We note that FDS-DF, SC-CF, SC-MAC, and strategies i) and ii) in Section IV-E are based on separate source and channel coding. Strategies iii) and iv) in Section IV-E are based on combined source and channel coding. Evaluating the performance of the combined source and channel coding strategies in the Gaussian channel is difficult as it involves discrete and continuous variables.

*Remark 12:* The achievable regions for FDS-DF, SC-CF, and SC-MAC are derived assuming that the number of source symbols received per unit time equals the number of channel transmissions per unit time. However, using these separate source- and channel-coding strategies, we can easily match the source symbol rate to the channel usage rate, without rederiving the coding strategies. Considering a general case when we wish to send $k$ pairs of source symbols using $n$ channel transmissions, the achievable regions can be found by simply replacing the mutual information expressions by $kH(\cdot) \leq nI(\cdot)$. In this way, the achievability question for a particular MACFCS is no longer just "whether we can reliably transmit a pair of $(S_1, S_2)$ per channel use," but more generally, "at what rate $k/n$ we can reliably transmit $k$ pairs of $(S_1, S_2)$ per $n$ channel uses." However, using combined source and channel coding strategies, we need to modify the coding strategies such that the probability distributions involve $k$ source symbols and $n$ channel input symbols, e.g., [4, eq. (87)]. Doing so, the achievable region will no longer be a single-letter characterization.

In Sections IV-A–IV-E, we investigated coding strategies for the MACFCS where the nodes exploit the broadcast/multiple-access nature of the channel. They cooperate in the sense that either the transmission from a node is decoded/processed by more than one node (broadcast nature) or a node decodes/processes the transmissions from more than one node (multiple-access nature). In the following subsection, we study a strategy in which the network is abstracted to a set of point-to-points links. A collection of links forms a route, and data are then passed down the route from the source to the destination using point-to-point coding.

*F. Multihop Coding With Data Aggregation (MH-DA)*

In the multihop coding with data aggregation strategy (MH-DA), data are passed from one node to another, until they reach the destination. First, we number the nodes in a sequence, which we call a route. The last node in the route is the destination. We consider a *combine–forward* multihop coding, where each node decodes the data from the previous node in the route, combines that with its own data, and forwards the aggregated data (those that it decodes from the previous node, plus its own data, less the correlated part with the data at the next node) to the next node in the route. In the three-node MACFCS, assuming that node 1 receives $\boldsymbol{s}_1$ from its source and node 2 receives $\boldsymbol{s}_2$ from its source, they do the following.

i) Node 1 compresses $\boldsymbol{s}_1$ down to $nH(S_1|S_2)$ bits, indexes it by $j \in \{1, \ldots, 2^{nH(S_1|S_2)}\}$, and sends it to node 2.

ii) We know that upon receiving $j$, node 2 can decode $\boldsymbol{s}_1$.

iii) Node 2 compresses $(\boldsymbol{s}_1, \boldsymbol{s}_2)$ to $k \in \{1, \ldots, 2^{nH(S_1, S_2)}\}$, using $nH(S_1, S_2)$ bits, and sends it to the destination.

In this multihop coding scheme, a node only decodes from the node behind in the route. The achievable region of the MACFCS using MH-DA is given in the following theorem.

*Theorem 5 (MH-DA):* Let $\big(\mathcal{S}_1 \times \mathcal{S}_2, p(s_1, s_2), \mathcal{X}_1 \times \mathcal{X}_2, p^*(y_1, y_2, y_3|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_3\big)$ be a discrete memoryless three-node MACFCS. $(S_1, S_2)$ can be reliably transmitted to the destination per channel use if the following holds:

$$H(S_1|S_2) < I(X_1; Y_2|X_2, Q) \tag{43a}$$
$$H(S_1, S_2) < I(X_2; Y_3|Q) \tag{43b}$$

where

$$p(q, x_1, x_2, y_1, y_2, y_3)$$
$$= p(q)p(x_2|q)p(x_1|x_2, q)p^*(y_1, y_2, y_3|x_1, x_2). \tag{44}$$

$Q \in \mathcal{Q}$ is the time-sharing variable and $|\mathcal{Q}| \leq 2$.

The proof for Theorem 5 is straightforward and will be omitted.

Now we consider a time-division MH-DA for the three-node AWGN-MACFCS. By time division, we mean that only one source transmits at a time, i.e., for fraction $(1 - f)$ of the time $(0 < f < 1)$, node 1 transmits and node 2 does not transmit; for fraction $f$ of the time, node 2 transmits and node 1 does not transmit. This might be done to reduce interference among the nodes.

The achievable region of the AWGN-MACFCS using the time division MH-DA is

$$H(S_1|S_2) < \frac{1-f}{2} \log\left(1 + \frac{kd_{12}^{-\eta}P_1}{(1-f)N_2}\right) \quad (45a)$$

$$H(S_1, S_2) < \frac{f}{2} \log\left(1 + \frac{kd_{23}^{-\eta}P_2}{fN_3}\right). \quad (45b)$$

We have presented four achievable regions for the MACFCS using different coding strategies, and suggested a few coding strategies for the MACFCS by combining coding strategies for the MACCS and the MACF. In the next section, we compare the four achievable regions to the CS-OB.

## V. COMPARISON OF CODING STRATEGIES

In this section, we plot and compare achievable regions for the different strategies in the three-node AWGN-MACFCS. We consider *symmetrical topologies*, i.e., both sources are of equidistant from the destination, and also linear asymmetrical topologies, where the three nodes form a straight line, with node 2 placed between node 1 and the destination. Although Gaussian input distributions may not be optimal, we choose $X_1$, $X_2$, and the auxiliary random variables to be Gaussian for the sake of comparison.

### A. Design Methodology

We perform numerical calculations to compare the achievable regions of different coding strategies and the CS-OB to gain insights into how node position and data correlation affect the performance. First, we study the effect of node position. For this analysis, we assume symmetrical source data, meaning $H(S_1|S_2) = H(S_2|S_1)$. This is a reasonable assumption for sensor networks when homogeneous sensors are deployed, and each sensor is sensing the environment at the same rate. For the computations, we fix $H(S_1|S_2) = 0.5$, $H(S_2|S_1) = 0.5$, and $I(S_1; S_2) = 0.5$. Although there are many combinations of node positions that one can study, we group them into three main categories.

  i) Symmetrical topology with the sources closer to the destination than they are to one another.
  ii) Symmetrical topology with the sources closer to one another than they are to the destination.
  iii) Asymmetrical topology. Without loss of generality, we assume that node 1 is further away from the destination than node 2 is from the destination. For simplicity, we study linear topologies where the three nodes form a straight line.

Taking a closer look at MH-DA, we note that this strategy is more suitable for asymmetrical topologies. This is because in symmetrical topologies, there is no reason why we would arrange the nodes in a route and "load" the node at the end of route. This strategy makes sense in the asymmetrical topology where some nodes are nearer to the destination. Hence, we analyze the performance of MH-DA only in the asymmetrical topology.

After investigating the effect of node position, we study the effect of varying the correlation between the sources on the performance of the various coding strategies. As rationalized above, we still keep the source data symmetrical, i.e., $H(S_1|S_2) = H(S_2|S_1)$. We vary $I(S_1; S_2)$ while keeping one of the following constant.



(a) Network topology

$N_j=1$, $d_{12}=1$, $d_{13}=d_{23}=0.5$, $H(S_1|S_2)=I(S_1;S_2)=H(S_2|S_1)=0.5$

(b) $P_1$ versus $P_2$.
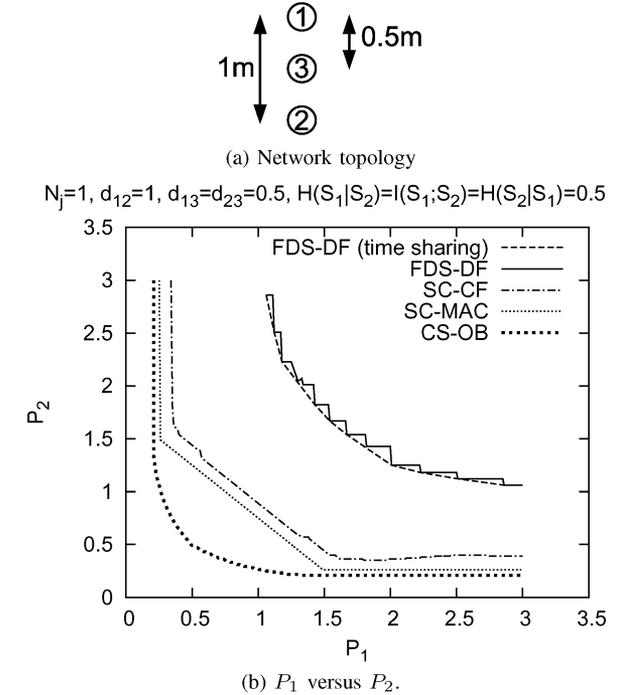
Fig. 4. Minimum power required to transmit $(S_1, S_2)$ to the destination per channel use, with weak inter-source link.



(a) Network topology

$N_j=1$, $d_{12}=1$, $d_{13}=d_{23}=5$, $H(S_1|S_2)=I(S_1;S_2)=H(S_2|S_1)=0.5$

(b) $P_1$ versus $P_2$.
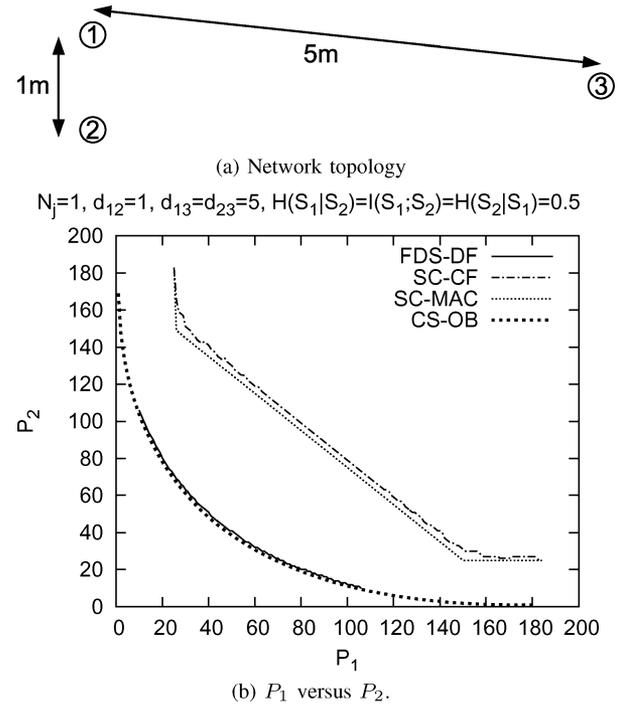
Fig. 5. Minimum power required to transmit $(S_1, S_2)$ to the destination per channel use, with weak source–destination links.

  i) The information of each source $H(S_1)$ and $H(S_2)$ is constant.
  ii) The total information $H(S_1, S_2)$ is constant.

### B. The Effect of Node Position

Figs. 4 and 5 show the minimum average transmit powers (energy per channel use) required for nodes 1 and 2 to reliably

transmit a pair of $(S_1, S_2)$ to the destination per channel use. The achievable region is the region above the line. Note that we plot average transmit powers on both axes. So, if the nodes transmit with an average power pair in the achievable region, the nodes can reliably send $(S_1, S_1)$ to the destination per channel use. We denote the average power of nodes 1 and 2 by $P_1$ and $P_2$, respectively.

We consider symmetrical source data with the following values: $H(S_1|S_2) = 0.5$, $H(S_2|S_1) = 0.5$, and $I(S_1; S_2) = 0.5$. First, we compare the two symmetrical topologies: 1) when the sources are further away from each other than they are from the destination, and 2) when the sources are further away from the destination than they are from each other. The first setup studies the case where the source–destination links are better than the inter-source link while the second setup studies the case where the inter-source link is better than the source–destination links.

When the inter-source link is weak, Fig. 4 shows that SC-CF and SC-MAC perform better than FDS-DF, i.e., the achievable regions for the SC-CF and SC-MAC contain that of FDS-DF. FDS-DF performs worst among the three strategies as the strategy requires each source node to get all the data from other nodes. This imposes an extra constraint on the average transmit power of the source nodes. When the source–destination link is stronger, a better strategy is to send the signals directly to the destination than to seek help from other sources.

On the other hand, when the inter-source link is strong, Fig. 5 shows that FDS-DF performs better than SC-CF and SC-MAC. The transmission bottleneck is now at the source–destination link. A good inter-source link lets each source node fully decode the messages from other nodes using little transmit power. In FDS-DF, the sources then use most of the transmit power to send the full information coherently to the destination. Coherent combining makes a significant gain in transmission rate on the source–destination link. Also, we see that the achievable region of FDS-DF comes very close to the CS-OB when the inter-source link is much better than the source–destination link.

*Remark 13:* In these two scenarios, we consider symmetrical topologies $(d_{13} = d_{23})$ and symmetrical source data $(H(S_1|S_2) = H(S_2|S_1))$. We term the channel with symmetrical topology and symmetrical source data *symmetrical MACFCS*. In the symmetrical MACFCS, using FDS-DF, the total average transmit power is minimized when the nodes transmit at the same average power. In other words, it is more efficient for the nodes to share the load in transmitting data than for one to transmit at higher power. We can see this from the nonlinearity in the coherent combining term in (69), or from Figs. 4 and 5, that the power curves are convex. However, for SC-CF and SC-MAC, there is a range for which individual source node can vary their transmit power while maintaining the minimum total average transmit power. We can see this from the mutual information expression in (87), (88), and (92a) that the relationship between $P_1$ and $P_2$ is linear, keeping $P_1 + P_2$ constant, or from Figs. 4 and 5, that there are portions of the SC-CF and SC-MAC curves where the slope is $-1$.

*Remark 14:* The staircase behavior of the FDS-DF curve in Fig. 4 is caused by the optimization involving different inequalities in Theorem 2 and the finite step size of $\{\alpha_{ij}\}$ in (63a) and



(a) Network topology

$d_{12}=2$, $d_{23}=0.5$, $d_{13}=2.5$, $H(S_1|S_2)=I(S_1;S_2)=H(S_2|S_1)=0.5$

(b) $P_1$ versus $P_2$.
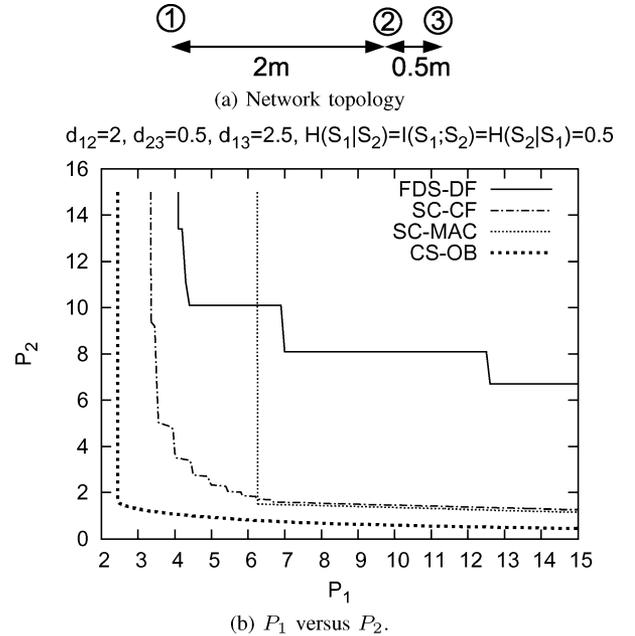
Fig. 6. Minimum power required to transmit $(S_1, S_2)$ to the destination per channel use, in a linear topology.

(63b). $\{\alpha_{ij}\}$ are the power splits of node $i$ used to carry different messages. The definition can be found in Appendix B. With time sharing, reliable transmission can be achieved using an average transmit power above the dotted line. Hence, time sharing enlarges the achievable region of FDS-DF. This explains why the time-sharing random variable $Q$ is included in Theorem 2.

*Remark 15:* In the symmetrical MACFCS, SC-MAC performs better than SC-CF. This means after we remove the correlation among the sources, using the feedback of the channel via the compress–forward strategy is worse than not using the feedback at all. This can be explained as follows. When the nodes are of same distance from the destination and have same amount of information to send, it is better for each of them to send their own message to the destination directly. It does not help when they try to help other nodes by sending a noisy, compressed, and binned version of what they received. The power can be better used to send their own uncorrupted data.

That SC-MAC always outperforms SC-CF is no longer true in the asymmetrical topology. Fig. 6 shows the minimum power curves without time sharing. For illustration, we choose $d_{12} = 2$, $d_{23} = 0.5$, and $d_{13} = 2.5$. We note that choosing node 2 close to node 1 resembles a symmetrical topology. From the graph, we see that using SC-MAC, the minimum power required at node 1 is 6.25 W. Using FDS-DF, the minimum power required at node 1 is 4.1 W. We can further reduce the power at node 1 to 3.35 W by increasing the power at node 2 by using SC-CF.

Using SC-MAC, we ignore the feedback in the channel. Hence, node 1 needs to transmit at least $H(S_1|S_2)$ bits to the destination, which is situated 2.5 m away. Note that without using the feedback, there is no way node 2 can help in sending this portion of the message. However, we can reduce node 1's transmit power by using FDS-DF. Now, node 1 needs to send at least $H(S_1|S_2)$ bits to node 2, which is 2 m away. Node
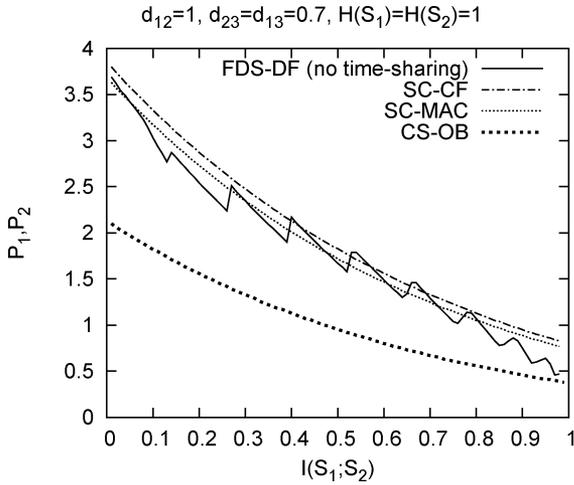
Fig. 7. Minimum power required to transmit $(S_1, s_2)$ to the destination per channel use, with different message correlation but constant $H(S_1)$ and $H(S_2)$.
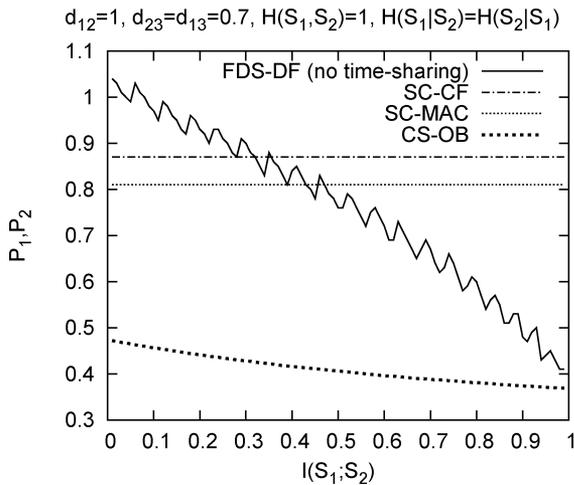


Fig. 8. Minimum power required to transmit $(S_1, S_2)$ to the destination per channel use, with different message correlation but constant $H(S_1, S_2)$.

2, which is nearer to the destination can help node 1 to relay the message to the destination. If we wish to further reduce the transmission power of node 1, we use SC-CF. Using this strategy, node 2 does not need to fully decode $S_1$. Node 2 acts as an additional (but noisy) antenna for the destination. Hence, this further enhances the "reception" of node 1's message. So node 1 needs to send at least $H(S_1|S_2)$ bits to the destination, equipped with an additional (noisy) antenna at node 2.

*Remark 16:* From Fig. 6, we see the staircase behavior of the SC-CF curve. It shows that time sharing increases the achievability region of SC-CF. This accounts for the use of the time-sharing auxiliary random variable $Q$ in Theorem 3.

### C. The Effect of Source Correlation

Now, we study how correlation among the source data affects the different coding strategies for the MACFCS. We consider symmetrical topologies. Figs. 7 and 8 depict achievable regions of the different coding strategies. We plot the equal power point for different correlation values.

From both graphs, we see that all three strategies perform either better or do not change when the data are more correlated. This make sense since if each node knows a larger portion of other nodes' data, it is easier for the nodes to cooperate.

When the nodes transmit at equal power, the achievable regions of SC-CF and SC-MAC do not vary with the correlation as long as the total information $H(S_1, S_2)$ remains constant and the correlation is symmetrical (Fig. 8). This is because using these two strategies, source coding is first performed. After that, the nodes send independent data to the destination. We know that the minimum total rate ($R_1 + R_2$ in (31c)) for which the nodes must transmit remains constant if $H(S_1, S_2)$ is constant.

In the same graph, although the total information $H(S_1, S_2)$ stays constant, increasing the correlation of the data enlarges the achievable region of FDS-DF. The reason is that when the correlation is higher, more power can be used for coherent transmission. The nodes need less power for inter-source communication.

When the sources are fully correlated, i.e., $H(S_1|S_2)$ and $H(S_2|S_1)$ approach zero, the achievable region of FDS-DF approaches the CS-OB. This does not come as a surprise as when $H(S_1|S_2) = H(S_2|S_1) = 0$, every source node knows other nodes' data. They can cooperate to form a multiple-transmit antenna without wasting any power to exchange data. Hence, it achieves the CS-OB.

The achievable regions of all the three strategies are far from the CS-OB when the inter-source distance is large compared to the source–destination distance and the correlation between the sources are low. To achieve the CS-OB, all sources need to cooperate to send full information. When the correlation is low and the inter-source link is weak, the sources "waste" a larger portion of the transmit power to communicate among themselves in FDS-DF. For SC-CF and SC-MAC, as no coherent combining is possible, the achievable regions are far from the CS-OB. This highlights the value of cooperation in the MACFCS.

*Remark 17:* We notice that in Figs. 7 and 8, the FDS-DF curves are zig-zag. This is because we plot the equal power point ($P_1 = P_2$) for the non-time-sharing FDS-DF. As can be seen from Fig. 4, time sharing might improve the FDS-DF region at the equal power point. The non-time-sharing FDS-DF curve in Fig. 4 coincides with the time-sharing curve at equal power point only at certain correlation levels. Hence, the deviation from the time-sharing line for different correlation levels accounts for the zig-zag behavior of the FDS-DF curve in Figs. 7 and 8 when we change the correlation level.

### D. Comparing MH-DA With Other Strategies

Figs. 9 and 10 compare MH-DA with other strategies in a three-node AWGN-MACFCS. As explained in Section V-A, we will only consider the linear topology when comparing MH-DA with other strategies. We consider the cases when node 2 is closer to node 1 and when node 2 is closer to the destination. We show that in both cases, we can always find a strategy with multiuser coding (FDS-DF, SC-CF, or SC-MAC) that outperforms MH-DA.

Using MH-DA, we penalize the nodes toward the end of the route as they need to send more information. In this example, node 2 needs to send full information, which is at least

TABLE I
NODE POSITIONING, CORRELATION, AND CODING STRATEGIES FOR SYMMETRICAL AWGN-MACFCS

| | Low correlation | High correlation |
|---|---|---|
| Good inter-source link[♯] | FDS-DF approaches CS-OB | FDS-DF approaches |
| Poor inter-source link[♯] | FDS-DF, SC-CF, SC-MAC far from CS-OB | CS-OB |

[♯]Relative to the source-destination links.

$d_{12}$=0.5, $d_{23}$=2, $d_{13}$=2.5, $H(S_1|S_2)$=$I(S_1;S_2)$=$H(S_2|S_1)$=0.5
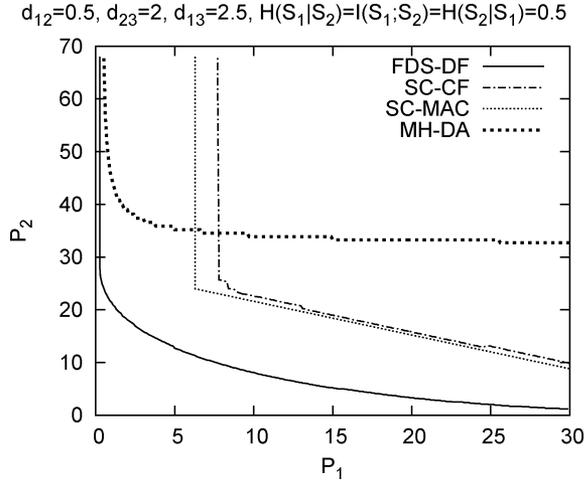


Fig. 9. Minimum power required to transmit $(S_1, S_2)$ to the destination per channel use, with node 2 closer to node 1.

$d_{12}$=2, $d_{23}$=0.5, $d_{13}$=2.5, $H(S_1|S_2)$=$I(S_1;S_2)$=$H(S_2|S_1)$=0.5
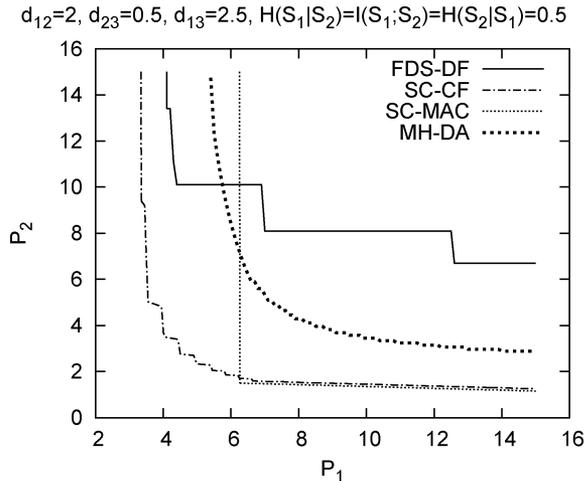


Fig. 10. Minimum power required to transmit $(S_1, S_2)$ to the destination per channel use, with node 2 closer to the destination.

$H(S_1, S_2)$ bits, "alone" to the destination as the destination only decodes from node 2. Hence, the minimum required $P_2$ is high. In other strategies, node 1 helps node 2 to transmit to the destination and hence a lower $P_2$ is possible. An exception is FDS-DF when node 2 is closer to the destination. Here, node 2 needs to transmit at high power to ensure that node 1 (which is situated further away) can fully decode its transmission.

Nodes in the beginning of the route benefit from MH-DA, but only when the "next hop" is near. From Fig. 9, we see that $P_1$ can be low when node 2 is closer to node 1. When node 2 is further away (Fig. 10), node 1 suffers. FDS-DF and SC-CF help node 1 to lower its transmit power as both node 2 and the destination are listening.

## VI. REFLECTIONS

The analyses of the different coding strategies for the AWGN-MACFCS help us to understand better how sensor nodes can cooperate in a network given node positions and correlation structures. We summarize the results from the numerical computations for the symmetrical AWGN-MACFCS in Table I.

*Remark 18:* When every source node can fully decode from other sources using little power (i.e., when the inter-source link is good or when the data are highly correlated), FDS-DF is a good choice of coding strategy. The sources can coherently transmit to the destination with little inter-node communication.

*Remark 19:* For the symmetrical topology when the inter-sources links are poor (e.g., when sources surround the destination), SC-MAC proves to be useful. A complicated scheme like SC-CF does not improve the achievable region.

*Remark 20:* For the asymmetrical topology, we note that SC-CF gives a better performance compared to FDS-DF and SC-MAC. SC-CF allows the furthest node to transmit at lower power as other source nodes now act as additional antennas for the destination.

*Remark 21:* For the linear asymmetrical topology with symmetrical source data, Figs. 9 and 10 show that we can always find a multiuser coding strategy that outperforms MH-DA. The problem with MH-DA is that it uses point-to-point coding and unfairly loads nodes nearer the end of the route. Multiuser coding strategies mitigate this by allowing richer forms of cooperation between nodes. This highlights the value of cooperative coding in the multiple-source network.

*Remark 22:* We investigated the three-node MACFCS in this paper. This simple example enabled us to demonstrate the characteristics of different coding strategies. Consider a sensor network. The correlation between the measured data often depends on the inter-sensor positions. A shorter inter-sensor distance usually results in a higher correlation between the data of the two sensors. Hence, the upper right cell and the lower left cell in Table I are of greater interest. If the sensors are closer to one another than they are from the sink, which normally results in a high correlation among the data, they should fully decode the data from all sensors and transmit coherently to the sink. This can be done by using FDS-DF based coding schemes. If the sensors are scattered around the sink, which normally results in a lower correlation among the data, a simple coding strategy like SC-MAC might be sufficient.

*Remark 23:* In a network with more nodes, mixed coding strategies can be used. Here, we give an example of how the results in this paper could help us to design a coding scheme for sensor network with more nodes. If there is a group of sensors situated further away from the destination and another group

closer to the destination, we suggest that sensors that are further from the destination form a group and fully decode the data from each other. They, as a group, then cooperate with sensors nearer to the destination via SC-CF.

## VII. CONCLUSION

In this paper, we presented four achievable regions for the MACFCS. In addition, we derived an outer bound on the capacity of the MACFCS, which turned out to be the cut-set bound. Using Gaussian channels as examples, we compared the achievable regions of different strategies to the cut-set outer bound. We showed that FDS-DF, SC-CF, and SC-MAC can each give superior performance in certain channel settings. From the comparison, we found that when the sources are closer to each other than they are to the destination, the achievable region of FDS-DF approaches the cut-set outer bound as the inter-source distance decreases. The same strategy also approaches the cut-set outer bound when the correlation between the sources gets higher. In symmetrical topologies, when the inter-source links are weak but the source–destination links are good, SC-MAC proves to be useful. In asymmetrical topologies, SC-CF can give better performance compared to the other strategies. Our study of the three-node MACFCS sheds light on how one could effectively operate a wireless sensor network where each sensor is tasked to collect and send correlated data measurements to a common sink.

## APPENDIX A
### PROOF OF THEOREM 2

In this appendix, we prove Theorem 2. We calculate the error probabilities and show that they diminish when certain conditions are satisfied.

*Proof of Theorem 2:* In each block, node 1 encodes $\boldsymbol{s}_1$ to $j$. Similarly, node 2 encodes $\boldsymbol{s}_2$ to $k$. Assuming noiseless channel, node 1 receives $k$ correctly and node 2 receives $j$ correctly. We define the following source coding error events:

$$\mathcal{E}_{0a} \triangleq \{\text{node 2 wrongly decodes } \boldsymbol{s}_1\} \tag{46a}$$
$$\mathcal{E}_{0b} \triangleq \{\text{node 1 wrongly decodes } \boldsymbol{s}_2\} \tag{46b}$$
$$\mathcal{E}_0 \triangleq \mathcal{E}_{0a} \cup \mathcal{E}_{0b}. \tag{46c}$$

Using the results by Slepian and Wolf [3, Theorem 2], $\Pr(\mathcal{E}_{0a})$ and $\Pr(\mathcal{E}_{0b})$ can be bounded by $\epsilon$ if $j$ is encoded with no less than $n[H(S_1|S_2) + \epsilon]$ bits and $k$ is encoded in no less than $n[H(S_2|S_1) + \epsilon]$ bits. Hence

$$\Pr(\mathcal{E}_0) \leq \Pr(\mathcal{E}_{0a}) + \Pr(\mathcal{E}_{0b}) < 2\epsilon. \tag{47}$$

Now, both sources have $(j, k, \boldsymbol{s}_1, \boldsymbol{s}_2)$. They compress $(\boldsymbol{s}_1, \boldsymbol{s}_2)$ to $h \in \{1, \ldots, 2^{n[H(S_1, S_2) + 3\epsilon]}\}$. We know that the destination can correctly decode $(\boldsymbol{s}_1, \boldsymbol{s}_2)$ from $h$ if $h$ is at least $n[H(S_1, S_2) + 3\epsilon]$ bits. Now, we create $2^{n[H(S_1|S_2) + H(S_2|S_1) + 2\epsilon]}$ bins and index each bin by a unique $(j, k)$. Assign $h$ to the bins so that each bin contains $2^{n[I(S_1; S_2) + \epsilon]}$ entries. Index the entries $i \in \{1, \ldots, 2^{n[I(S_1; S_2) + \epsilon]}\}$. Hence, each $h$ can be represented by a unique triplet $(i, j, k)$.

Assume that in the beginning of block $t$, nodes 1 and 2 have correctly received $(j^{t-1}, k^{t-1})$ and determined $i^{t-1}$. They send $\boldsymbol{x}_1(j^t|h^{t-1})$ and $\boldsymbol{x}_2(k^t|h^{t-1})$, respectively, where $h^{t-1} = (i^{t-1}, j^{t-1}, k^{t-1})$. At the end of block $t$, nodes 1 and 2 received $\boldsymbol{y}_1(t)$ and $\boldsymbol{y}_2(t)$, respectively. We define the following error events at nodes 1 and 2:

$$\mathcal{E}_1 \triangleq \left\{ \boldsymbol{w}_0(i^{t-1}), \boldsymbol{w}_1(j^{t-1}), \boldsymbol{w}_2(k^{t-1}), \boldsymbol{x}_1(j^t|h^{t-1}), \right. \\ \left. \boldsymbol{x}_2(k^t|h^{t-1}), \boldsymbol{y}_1(t) \right\} \notin \mathcal{A}_\epsilon \tag{48a}$$

$$\mathcal{E}_2 \triangleq \left\{ \boldsymbol{w}_0(i^{t-1}), \boldsymbol{w}_1(j^{t-1}), \boldsymbol{w}_2(k^{t-1}), \boldsymbol{x}_1(j^t|h^{t-1}), \right. \\ \left. \boldsymbol{x}_2(k|h^{t-1}), \boldsymbol{y}_1(t) \right\} \in \mathcal{A}_\epsilon \tag{48b}$$

$$\mathcal{E}_3 \triangleq \left\{ \boldsymbol{w}_0(i^{t-1}), \boldsymbol{w}_1(j^{t-1}), \boldsymbol{w}_2(k^{t-1}), \boldsymbol{x}_1(j^t|h^{t-1}), \right. \\ \left. \boldsymbol{x}_2(k^t|h^{t-1}), \boldsymbol{y}_2(t) \right\} \notin \mathcal{A}_\epsilon \tag{48c}$$

$$\mathcal{E}_4 \triangleq \left\{ \boldsymbol{w}_0(i^{t-1}), \boldsymbol{w}_1(j^{t-1}), \boldsymbol{w}_2(k^{t-1}), \boldsymbol{x}_1(j|h^{t-1}), \right. \\ \left. \boldsymbol{x}_2(k^t|h^{t-1}), \boldsymbol{y}_2(t) \right\} \in \mathcal{A}_\epsilon \tag{48d}$$

for all $j \in \{1, 2, \ldots, 2^{n[H(S_1|S_2) + \epsilon]}\} \setminus \{j^t\}$ and $k \in \{1, 2, \ldots, 2^{n[H(S_2|S_1) + \epsilon]}\} \setminus \{k^t\}$. By the asymptotic equipartition property (AEP), for sufficiently large $n$

$$\Pr(\mathcal{E}_1) < \epsilon \tag{49a}$$
$$\Pr(\mathcal{E}_3) < \epsilon. \tag{49b}$$

The probability that error event $\mathcal{E}_2$ occurs for all $k \neq k^t$ is given by

$$\Pr(\mathcal{E}_2)$$
$$= \sum_{k \neq k^t} \sum_{\{\boldsymbol{w}_0, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}_1\} \in \mathcal{A}_\epsilon} p(\boldsymbol{x}_2|\boldsymbol{w}_0, \boldsymbol{w}_1, \boldsymbol{w}_2)$$
$$\cdot p(\boldsymbol{w}_0, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{x}_1, \boldsymbol{y}_1) \tag{50a}$$
$$< 2^{n[H(S_1|S_2) + \epsilon]} 2^{n[H(W_0, W_1, W_2, X_1, X_2, Y_1) + \epsilon]}$$
$$\times 2^{-n[H(X_2|W_0, W_1, W_2) - \epsilon]} 2^{-n[H(W_0, W_1, W_2, X_1, Y_1) - \epsilon]} \tag{50b}$$
$$= 2^{n[H(S_2|S_1) + H(X_2|W_0, W_1, W_2, X_1, Y_1) - H(X_2|W_0, W_1, W_2) + 4\epsilon]} \tag{50c}$$
$$= 2^{n[H(S_2|S_1) + H(X_2|W_0, W_1, W_2, X_1, Y_1) - H(X_2|W_0, W_1, W_2, X_1) + 4\epsilon]} \tag{50d}$$
$$= 2^{n[H(S_2|S_1) - I(X_2; Y_1|W_0, W_1, W_2, X_1) + 4\epsilon]}. \tag{50e}$$

This can be made arbitrarily small if

$$H(S_2|S_1) < I(X_2; Y_1|W_0, W_1, W_2, X_1) - 4\epsilon \tag{51}$$

holds and $n$ is sufficiently large. Similarly, we can show that $\Pr(\mathcal{E}_4)$ can be made small if

$$H(S_1|S_2) < I(X_1; Y_2|W_0, W_1, W_2, X_2) - 4\epsilon. \tag{52}$$

Now we look at the error probability at the destination. Assume that nodes 1 and 2 send $\boldsymbol{x}_1(j^t|h^{t-1})$ and $\boldsymbol{x}_2(k^t|h^{t-1})$, respectively, in block $t$; and $\boldsymbol{x}_1(j^{t+1}|h^t)$ and $\boldsymbol{x}_2(k^{t+1}|h^t)$, respectively, in block $t + 1$. Assume that the destination has correctly

decoded $h^{t-1} = (i^{t-1}, j^{t-1}, k^{t-1})$. We define the following error events at the destination:

$$\mathcal{E}_{5a} \triangleq \Big\{ \boldsymbol{w}_0(i^{t-1}), \boldsymbol{w}_1(j^{t-1}), \boldsymbol{w}_2(k^{t-1}), \boldsymbol{x}_1(j^t | h^{t-1}),$$
$$\boldsymbol{x}_2(k^t | h^{t-1}), \boldsymbol{y}_3(t) \Big\} \notin \mathcal{A}_\epsilon \quad (53a)$$

$$\mathcal{E}_{5b} \triangleq \Big\{ \boldsymbol{w}_0(i^t), \boldsymbol{w}_1(j^t), \boldsymbol{w}_2(k^t), \boldsymbol{y}_3(t+1) \Big\} \notin \mathcal{A}_\epsilon \quad (53b)$$

$$\mathcal{E}_5 \triangleq \mathcal{E}_{5a} \cap \mathcal{E}_{5b} \quad (53c)$$

$$\mathcal{E}_6 \triangleq \Big\{ \boldsymbol{w}_0(i), \boldsymbol{w}_1(j^t), \boldsymbol{w}_2(k^t), \boldsymbol{y}_3(t+1) \Big\} \in \mathcal{A}_\epsilon \quad (53d)$$

$$\mathcal{E}_{7a} \triangleq \Big\{ \boldsymbol{w}_0(i^{t-1}), \boldsymbol{w}_1(j^{t-1}), \boldsymbol{w}_2(k^{t-1}), \boldsymbol{x}_1(j | h^{t-1}),$$
$$\boldsymbol{x}_2(k^t | h^{t-1}), \boldsymbol{y}_3(t) \Big\} \in \mathcal{A}_\epsilon \quad (53e)$$

$$\mathcal{E}_{7b} \triangleq \Big\{ \boldsymbol{w}_0(i^t), \boldsymbol{w}_1(j), \boldsymbol{w}_2(k^t), \boldsymbol{y}_3(t+1) \Big\} \in \mathcal{A}_\epsilon \quad (53f)$$

$$\mathcal{E}_7 \triangleq \mathcal{E}_{7a} \cap \mathcal{E}_{7b} \quad (53g)$$

$$\mathcal{E}_{8a} \triangleq \Big\{ \boldsymbol{w}_0(i^{t-1}), \boldsymbol{w}_1(j^{t-1}), \boldsymbol{w}_2(k^{t-1}), \boldsymbol{x}_1(j^t | h^{t-1}),$$
$$\boldsymbol{x}_2(k | h^{t-1}), \boldsymbol{y}_3(t) \Big\} \in \mathcal{A}_\epsilon \quad (53h)$$

$$\mathcal{E}_{8b} \triangleq \Big\{ \boldsymbol{w}_0(i^t), \boldsymbol{w}_1(j^t), \boldsymbol{w}_2(k), \boldsymbol{y}_3(t+1) \Big\} \in \mathcal{A}_\epsilon \quad (53i)$$

$$\mathcal{E}_8 \triangleq \mathcal{E}_{8a} \cap \mathcal{E}_{8b} \quad (53j)$$

$$\mathcal{E}_{9a} \triangleq \Big\{ \boldsymbol{w}_0(i^{t-1}), \boldsymbol{w}_1(j^{t-1}), \boldsymbol{w}_2(k^{t-1}), \boldsymbol{x}_1(j | h^{t-1}),$$
$$\boldsymbol{x}_2(k^t | h^{t-1}), \boldsymbol{y}_3(t) \Big\} \in \mathcal{A}_\epsilon \quad (53k)$$

$$\mathcal{E}_{9b} \triangleq \Big\{ \boldsymbol{w}_0(i), \boldsymbol{w}_1(j), \boldsymbol{w}_2(k^t), \boldsymbol{y}_3(t+1) \Big\} \in \mathcal{A}_\epsilon \quad (53l)$$

$$\mathcal{E}_9 \triangleq \mathcal{E}_{9a} \cap \mathcal{E}_{9b} \quad (53m)$$

$$\mathcal{E}_{10a} \triangleq \Big\{ \boldsymbol{w}_0(i^{t-1}), \boldsymbol{w}_1(j^{t-1}), \boldsymbol{w}_2(k^{t-1}), \boldsymbol{x}_1(j^t | h^{t-1}),$$
$$\boldsymbol{x}_2(k | h^{t-1}), \boldsymbol{y}_3(t) \Big\} \in \mathcal{A}_\epsilon \quad (53n)$$

$$\mathcal{E}_{10b} \triangleq \Big\{ \boldsymbol{w}_0(i), \boldsymbol{w}_1(j^t), \boldsymbol{w}_2(k), \boldsymbol{y}_3(t+1) \Big\} \in \mathcal{A}_\epsilon \quad (53o)$$

$$\mathcal{E}_{10} \triangleq \mathcal{E}_{10a} \cap \mathcal{E}_{10b} \quad (53p)$$

$$\mathcal{E}_{11a} \triangleq \Big\{ \boldsymbol{w}_0(i^{t-1}), \boldsymbol{w}_1(j^{t-1}), \boldsymbol{w}_2(k^{t-1}), \boldsymbol{x}_1(j | h^{t-1}),$$
$$\boldsymbol{x}_2(k | h^{t-1}), \boldsymbol{y}_3(t) \Big\} \in \mathcal{A}_\epsilon \quad (53q)$$

$$\mathcal{E}_{11b} \triangleq \Big\{ \boldsymbol{w}_0(i^t), \boldsymbol{w}_1(j), \boldsymbol{w}_2(k), \boldsymbol{y}_3(t+1) \Big\} \in \mathcal{A}_\epsilon \quad (53r)$$

$$\mathcal{E}_{11} \triangleq \mathcal{E}_{11a} \cap \mathcal{E}_{11b} \quad (53s)$$

$$\mathcal{E}_{12a} \triangleq \Big\{ \boldsymbol{w}_0(i^{t-1}), \boldsymbol{w}_1(j^{t-1}), \boldsymbol{w}_2(k^{t-1}), \boldsymbol{x}_1(j | h^{t-1}),$$
$$\boldsymbol{x}_2(k | h^{t-1}), \boldsymbol{y}_3(t) \Big\} \in \mathcal{A}_\epsilon \quad (53t)$$

$$\mathcal{E}_{12b} \triangleq \Big\{ \boldsymbol{w}_0(i), \boldsymbol{w}_1(j), \boldsymbol{w}_2(k), \boldsymbol{y}_3(t+1) \Big\} \in \mathcal{A}_\epsilon \quad (53u)$$

$$\mathcal{E}_{12} \triangleq \mathcal{E}_{12a} \cap \mathcal{E}_{12b} \quad (53v)$$

for all $i \neq i^{t+1}$, $j \neq j^{t+1}$, and $k \neq k^{t+1}$.

By AEP and for sufficiently large $n$, $\Pr(\mathcal{E}_{5a}) < \epsilon$ and $\Pr(\mathcal{E}_{5b}) < \epsilon$. Hence, $\Pr(\mathcal{E}_5) \leq \Pr(\mathcal{E}_{5a}) \Pr(\mathcal{E}_{5b}) < \epsilon^2$

$$\Pr(\mathcal{E}_6)$$
$$= \sum_{i \neq i^t} \sum_{\{\boldsymbol{w}_0, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{y}_3\} \in \mathcal{A}_\epsilon} p(\boldsymbol{w}_0) p(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{y}_3) \quad (54a)$$
$$< 2^{n[I(S_1;S_2)+\epsilon]} 2^{n[H(W_0,W_1,W_2,Y_3)+\epsilon]} 2^{-n[H(W_0)-\epsilon]}$$
$$2^{-n[H(W_1,W_2,Y_3)-\epsilon]} \quad (54b)$$

$$\leq 2^{n[I(S_1;S_2)+\epsilon]} 2^{-n[H(W_0|W_1,W_2)-\epsilon]} 2^{n[H(W_0|W_1,W_2,Y_3)-2\epsilon]}$$
$$\quad (54c)$$
$$= 2^{n[I(S_1;S_2)-I(W_0;Y_3|W_1,W_2)+4\epsilon]}. \quad (54d)$$

Hence, $\Pr(\mathcal{E}_6)$ can be made small if

$$I(S_1; S_2) < I(W_0; Y_3 | W_1, W_2) - 4\epsilon \quad (55)$$

holds and $n$ is sufficiently large.

$$\Pr(\mathcal{E}_7)$$
$$= \sum_{j \neq j^t} \Bigg[ \sum_{\{\boldsymbol{w}_0, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}_3\} \in \mathcal{A}_\epsilon} p(\boldsymbol{x}_1 | \boldsymbol{w}_0, \boldsymbol{w}_1, \boldsymbol{w}_2)$$
$$\cdot p(\boldsymbol{w}_0, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{x}_2, \boldsymbol{y}_3)$$
$$\times \sum_{\{\boldsymbol{w}_0, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{y}_3\} \in \mathcal{A}_\epsilon} p(\boldsymbol{w}_1) p(\boldsymbol{w}_0, \boldsymbol{w}_2, \boldsymbol{y}_3) \Bigg] \quad (56a)$$
$$< 2^{n[H(S_1|S_2)+\epsilon]} 2^{-n[I(X_1;Y_3|W_0,W_1,W_2,X_2)-3\epsilon]}$$
$$\times 2^{-n[I(W_1;Y_3|W_0,W_2)-3\epsilon]} \quad (56b)$$
$$= 2^{n[H(S_1|S_2)-I(X_1;Y_3|W_0,W_1,W_2,X_2)-I(W_1;Y_3|W_0,W_2)+7\epsilon]}. \quad (56c)$$

Hence $\Pr(\mathcal{E}_7)$ can be made small if

$$H(S_1|S_2) < I(W_1; Y_3 | W_0, W_2)$$
$$+ I(X_1; Y_3 | W_0, W_1, W_2, X_2) - 7\epsilon \quad (57)$$

holds and $n$ is sufficiently large.

Similarly, $Pr(\mathcal{E}_8)$ be made arbitrarily small if

$$H(S_2|S_1) < I(W_2; Y_3 | W_0, W_1)$$
$$+ I(X_2; Y_3 | W_0, W_1, W_2, X_1) - 7\epsilon. \quad (58)$$

$$\Pr(\mathcal{E}_9)$$
$$= \sum_{(i,j) \neq (i^t, j^t)} \Bigg[ \sum_{\{\boldsymbol{w}_0, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}_3\} \in \mathcal{A}_\epsilon} p(\boldsymbol{x}_1 | \boldsymbol{w}_0, \boldsymbol{w}_1, \boldsymbol{w}_2)$$
$$\cdot p(\boldsymbol{w}_0, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{x}_2, \boldsymbol{y}_3)$$
$$\times \sum_{\{\boldsymbol{w}_0, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{y}_3\} \in \mathcal{A}_\epsilon} p(\boldsymbol{w}_0, \boldsymbol{w}_1) p(\boldsymbol{w}_2, \boldsymbol{y}_3) \Bigg] \quad (59a)$$
$$< 2^{n[I(S_1;S_2)+\epsilon+H(S_1|S_2)+\epsilon]} 2^{-n[I(X_1;Y_3|W_0,W_1,W_2,X_2)-3\epsilon]}$$
$$\times 2^{-n[I(W_0,W_1;Y_3|W_2)-3\epsilon]} \quad (59b)$$
$$= 2^{n[H(S_1)-I(X_1;Y_3|W_0,W_1,W_2,X_2)-I(W_0,W_1;Y_3|W_2)+8\epsilon]}. \quad (59c)$$

Hence, $\Pr(\mathcal{E}_9)$ can be made small if

$$H(S_1) < I(W_0, W_1; Y_3 | W_2) + I(X_1; Y_3 | W_0, W_1, W_2, X_2) - 8\epsilon \quad (60)$$

holds and $n$ is sufficiently large.

Similarly, $Pr(\mathcal{E}_{10})$, $Pr(\mathcal{E}_{11})$, and $Pr(\mathcal{E}_{12})$ can be made arbitrarily small if

$$H(S_2) < I(W_0, W_2; Y_3 | W_1)$$
$$+ I(X_2; Y_3 | W_0, W_1, W_2, X_1) - 8\epsilon \quad (61a)$$

$$H(S_1|S_2) + H(S_2|S_1) < I(W_1, W_2; Y_3|W_0)$$
$$+ I(X_1, X_2; Y_3|W_0, W_1, W_2) - 8\epsilon \tag{61b}$$

$$H(S_1, S_2) < I(X_1, X_2; Y_3) - 9\epsilon \tag{61c}$$

hold, respectively.

If all these constraints are satisfied and if $n$ is large enough, the total probability of error can be bounded by

$$P_e = \bigcup_{i=0}^{12} \Pr(\mathcal{E}_i) < \sum_{i=0}^{12} \Pr(\mathcal{E}_i) < 12\epsilon + \epsilon^2 \tag{62}$$

for any $\epsilon > 0$.

Combining these rate constraints and adding the time-sharing random variable $Q$, we get Theorem 2. $\square$

## APPENDIX B
## ACHIEVABILITY OF FDS-DF ON AWGN-MACFCS

On the AWGN channel, using FDS-DF, nodes 1 and 2 send the following, respectively:

$$X_1 = \sqrt{\alpha_{10}P_1}W_0 + \sqrt{\alpha_{11}P_1}W_1 + \sqrt{\alpha_{12}P_1}W_2 + \sqrt{\alpha_{13}P_1}U_1 \tag{63a}$$
$$X_2 = \sqrt{\alpha_{20}P_2}W_0 + \sqrt{\alpha_{21}P_2}W_1 + \sqrt{\alpha_{22}P_2}W_2 + \sqrt{\alpha_{23}P_2}U_2 \tag{63b}$$

where $W_i$ and $U_j$ are independent Gaussian random variables with unit power $E[W_i^2] = [U_j^2] = 1, \forall i = 0, 1, 2$ and $\forall j = 1, 2$. $0 \le \sum_{k=0}^3 \alpha_{jk} \le 1$ for $j = 1, 2$. Recall that the channel outputs are

$$Y_1 = \sqrt{\kappa d_{21}^{-\eta}}X_2 + Z_1 \tag{64a}$$
$$= \sqrt{\kappa d_{21}^{-\eta}P_1}(\sqrt{\alpha_{10}}W_0 + \sqrt{\alpha_{11}}W_1 + \sqrt{\alpha_{12}}W_2$$
$$+ \sqrt{\alpha_{13}}U_1) + Z_1, \tag{64b}$$
$$Y_2 = \sqrt{\kappa d_{12}^{-\eta}}X_1 + Z_2 \tag{64c}$$
$$= \sqrt{\kappa d_{12}^{-\eta}P_2}(\sqrt{\alpha_{20}}W_0 + \sqrt{\alpha_{21}}W_1 + \sqrt{\alpha_{22}}W_2$$
$$+ \sqrt{\alpha_{23}}U_2) + Z_2, \tag{64d}$$
$$Y_3 = \sqrt{\kappa d_{13}^{-\eta}}X_1 + \sqrt{\kappa d_{23}^{-\eta}}X_2 + Z_3 \tag{64e}$$
$$= \left[\sqrt{\kappa d_{13}^{-\eta}\alpha_{10}P_1} + \sqrt{\kappa d_{23}^{-\eta}\alpha_{20}P_2}\right] W_0$$
$$+ \left[\sqrt{\kappa d_{13}^{-\eta}\alpha_{11}P_1} + \sqrt{\kappa d_{23}^{-\eta}\alpha_{21}P_2}\right] W_1$$
$$+ \left[\sqrt{\kappa d_{13}^{-\eta}\alpha_{12}P_1} + \sqrt{\kappa d_{23}^{-\eta}\alpha_{22}P_2}\right] W_2$$
$$+ \sqrt{\kappa d_{13}^{-\eta}\alpha_{13}P_1}U_1 + \sqrt{\kappa d_{23}^{-\eta}\alpha_{23}P_2}U_2 + Z_3. \tag{64f}$$

Now, we calculate the mutual information terms in Theorem 2.

$$I(X_1; Y_2|W_0, W_1, W_2, X_2) = \frac{1}{2}\log\left(1 + \frac{\kappa d_{12}^{-\eta}\alpha_{13}P_1}{N_2}\right) \tag{65a}$$

$$I(X_2; Y_1|W_0, W_1, W_2, X_1) = \frac{1}{2}\log\left(1 + \frac{\kappa d_{21}^{-\eta}\alpha_{23}P_2}{N_1}\right). \tag{65b}$$

For $(a, b, c) \in \left\{\{0, 1, 2\}^3 : a \ne b \ne c\right\}$

$$I(W_a; Y_3|W_b, W_c)$$
$$= \frac{1}{2}\log\left(1 + \frac{\kappa\left[\sqrt{d_{13}^{-\eta}\alpha_{1a}P_1} + \sqrt{d_{23}^{-\eta}\alpha_{2a}P_2}\right]^2}{\kappa d_{13}^{-\eta}\alpha_{13}P_1 + \kappa d_{23}^{-\eta}\alpha_{23}P_2 + N_3}\right) \tag{66a}$$

$$I(W_a, W_b; Y_3|W_c)$$
$$= \frac{1}{2}\log\left(1 + \frac{\kappa\left[\sqrt{d_{13}^{-\eta}\alpha_{1a}P_1} + \sqrt{d_{23}^{-\eta}\alpha_{2a}P_2}\right]^2}{\kappa d_{13}^{-\eta}\alpha_{13}P_1 + \kappa d_{23}^{-\eta}\alpha_{23}P_2 + N_3}\right.$$
$$\left. + \frac{\kappa\left[\sqrt{d_{13}^{-\eta}\alpha_{1b}P_1} + \sqrt{d_{23}^{-\eta}\alpha_{2b}P_2}\right]^2}{\kappa d_{13}^{-\eta}\alpha_{13}P_1 + \kappa d_{23}^{-\eta}\alpha_{23}P_2 + N_3}\right). \tag{66b}$$

Also, for $(a, b) \in \left\{\{1, 2\}^2 : a \ne b\right\}$

$$I(X_a; Y_3|W_0, W_1, W_2, X_b) = \frac{1}{2}\log\left(1 + \frac{\kappa d_{a3}^{-\eta}\alpha_{a3}P_a}{N_3}\right) \tag{67}$$

and

$$I(X_1, X_2; Y_3|W_0, W_1, W_2)$$
$$= \frac{1}{2}\log\left(1 + \frac{\kappa d_{13}^{-\eta}\alpha_{13}P_1 + \kappa d_{23}^{-\eta}\alpha_{23}P_2}{N_3}\right). \tag{68}$$

Finally

$$I(X_1, X_2; Y_3)$$
$$= \frac{1}{2}\log\left(1 + \frac{\sum_{i=0}^2 \kappa\left[\sqrt{d_{13}^{-\eta}\alpha_{1i}P_1} + \sqrt{d_{23}^{-\eta}\alpha_{2i}P_2}\right]^2}{N_3}\right.$$
$$\left. + \frac{\kappa d_{13}^{-\eta}\alpha_{13}P_1 + \kappa d_{23}^{-\eta}\alpha_{23}P_2}{N_3}\right). \tag{69a}$$

## APPENDIX C
## PROOF OF THEOREM 3

In this section, we prove Theorem 3.

*Proof of Theorem 3:* Node 1 receives $\boldsymbol{y}_1(t)$ in block $t$. It knows $\boldsymbol{x}_1(j^t, p^{t-1})$ and $\boldsymbol{u}_1(p^{t-1})$. It finds $v^t$ such that

$\left(\tilde{\boldsymbol{y}}_1(v^t|p^{t-1}), \boldsymbol{y}_1(t), \boldsymbol{x}_1(j^t, p^{t-1}), \boldsymbol{u}_1(p^{t-1})\right) \in \mathcal{A}_\epsilon$. Using [24, Lemma 2.1.3], node 1 can find such a $v^t$ with probability that tends to 1 as $n \to \infty$ if

$$\tilde{R}_1 > I(\tilde{Y}_1; Y_1|X_1, U_1). \tag{70}$$

By similar argument, node 2 can find $w^t$ with probability that tends to 1 as $n \to \infty$ such that $\left(\tilde{\boldsymbol{y}}_2(w^t|q^{t-1}), \boldsymbol{y}_2(t), \boldsymbol{x}_2(k^t, q^{t-1}), \boldsymbol{u}_2(q^{t-1})\right) \in \mathcal{A}_\epsilon$ if

$$\tilde{R}_2 > I(\tilde{Y}_2; Y_2|X_2, U_2). \tag{71}$$

Supposed that nodes 1 and 2 send $\boldsymbol{x}_1(j^{t+1}, p^t)$ and $\boldsymbol{x}_2(k^{t+1}, q^t)$, respectively, in block $t+1$. Define the following event where the destination wrongly decodes the quantized and binned signal $p^t$ or $q^t$:

$$\mathcal{E}_1 \triangleq \left(\boldsymbol{u}_1(p^t), \boldsymbol{u}_2(q^t), \boldsymbol{y}_3(t+1)\right) \notin \mathcal{A}_\epsilon \tag{72a}$$
$$\mathcal{E}_2 \triangleq \left(\boldsymbol{u}_1(p), \boldsymbol{u}_2(q^t), \boldsymbol{y}_3(t+1)\right) \in \mathcal{A}_\epsilon \tag{72b}$$
$$\mathcal{E}_3 \triangleq \left(\boldsymbol{u}_1(p^t), \boldsymbol{u}_2(q), \boldsymbol{y}_3(t+1)\right) \in \mathcal{A}_\epsilon \tag{72c}$$
$$\mathcal{E}_4 \triangleq \left(\boldsymbol{u}_1(p), \boldsymbol{u}_2(q), \boldsymbol{y}_3(t+1)\right) \in \mathcal{A}_\epsilon \tag{72d}$$

for all $p \in \{1, 2, \ldots, 2^{nR_1'}\} \setminus \{p^t\}$ and $q \in \{1, 2, \ldots, 2^{nR_2'}\} \setminus \{q^t\}$.

By AEP, $\Pr(\mathcal{E}_1) < \epsilon$ for large $n$. We can show that $\Pr(\mathcal{E}_2), \Pr(\mathcal{E}_3)$, and $\Pr(\mathcal{E}_4)$ can be bounded by $\epsilon$ for large $n$ if the following holds:

$$R_1' < I(U_1; Y_3|U_2) - 3\epsilon \tag{73a}$$
$$R_2' < I(U_2; Y_3|U_1) - 3\epsilon \tag{73b}$$
$$R_1' + R_2' < I(U_1, U_2; Y_3) - 3\epsilon. \tag{73c}$$

At the end of block $t$, assume that the destination has already correctly decoded the quantized and binned signals $p^t, q^t, p^{t-1}$, and $q^{t-1}$. Suppose that $v^t$ and $w^t$ are the quantized values of nodes 1 and 2, respectively. We define the following events where the destination decodes the estimates wrongly, for all $v \in \{1, 2, \ldots, 2^{n\tilde{R}_1}\} \setminus \{v^t\}$ and $w \in \{1, 2, \ldots, 2^{n\tilde{R}_2}\} \setminus \{w^t\}$:

$$\mathcal{E}_5 \triangleq \left(\tilde{\boldsymbol{y}}_1(v^t|p^{t-1}), \tilde{\boldsymbol{y}}_2(w^t|q^{t-1}), \boldsymbol{u}_1(p^{t-1}), \boldsymbol{u}_2(q^{t-1}), \boldsymbol{y}_3(t)\right)$$
$$\notin \mathcal{A}_\epsilon, \tag{74a}$$
$$\mathcal{E}_{6a} \triangleq \left(\tilde{\boldsymbol{y}}_1(v|p^{t-1}), \tilde{\boldsymbol{y}}_2(w^t|q^{t-1}), \boldsymbol{u}_1(p^{t-1}), \boldsymbol{u}_2(q^{t-1}), \boldsymbol{y}_3(t)\right)$$
$$\in \mathcal{A}_\epsilon, \tag{74b}$$
$$\mathcal{E}_6 \triangleq \mathcal{E}_{6a} \cap \{v \in S_{p^t}\}, \tag{74c}$$
$$\mathcal{E}_{7a} \triangleq \left(\tilde{\boldsymbol{y}}_1(v^t|p^{t-1}), \tilde{\boldsymbol{y}}_2(w|q^{t-1}), \boldsymbol{u}_1(p^{t-1}), \boldsymbol{u}_2(q^{t-1}), \boldsymbol{y}_3(t)\right)$$
$$\in \mathcal{A}_\epsilon, \tag{74d}$$
$$\mathcal{E}_7 \triangleq \mathcal{E}_{7a} \cap \{w \in S_{q^t}\}, \tag{74e}$$
$$\mathcal{E}_{8a} \triangleq \left(\tilde{\boldsymbol{y}}_1(v|p^{t-1}), \tilde{\boldsymbol{y}}_2(w|q^{t-1}), \boldsymbol{u}_1(p^{t-1}), \boldsymbol{u}_2(q^{t-1}), \boldsymbol{y}_3(t)\right)$$
$$\in \mathcal{A}_\epsilon, \tag{74f}$$
$$\mathcal{E}_8 \triangleq \mathcal{E}_{8a} \cap \{v \in S_{p^t}\} \cap \{w \in S_{q^t}\}. \tag{74g}$$

By AEP, $\Pr(\mathcal{E}_5) < \epsilon$ for large $n$. The probability of the event $\mathcal{E}_6$ is as follows:

$$\Pr(\mathcal{E}_6)$$
$$= \Pr\left(\mathcal{E}_{6a} \cap \{v \in S_{p^t}\}\right) \tag{75a}$$

$$= \sum_{\substack{v \neq v^t \\ v \in S_{p^t}}} \sum_{(\tilde{\boldsymbol{y}}_1, \tilde{\boldsymbol{y}}_2, \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{y}_3) \in \mathcal{A}_\epsilon} p(\tilde{\boldsymbol{y}}_1|\tilde{\boldsymbol{y}}_2, \boldsymbol{u}_1, \boldsymbol{u}_2) p(\tilde{\boldsymbol{y}}_2, \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{y}_3)$$
$$\tag{75b}$$
$$< 2^{n(\tilde{R}_1 - R_1')} \times 2^{n[H(\tilde{Y}_1, \tilde{Y}_2, U_1, U_2, Y_3) + \epsilon]}$$
$$\times 2^{-n[H(\tilde{Y}_1|\tilde{Y}_2, U_1, U_2) - \epsilon]} \times 2^{-n[H(\tilde{Y}_2, U_1, U_2, Y_3) - \epsilon]} \tag{75c}$$
$$= 2^{n(\tilde{R}_1 - R_1')} \times 2^{-n(I(\tilde{Y}_1; Y_3|\tilde{Y}_2, U_1, U_2) - 3\epsilon)}. \tag{75d}$$

This can be made small, for a large $n$, if

$$\tilde{R}_1 < I(\tilde{Y}_1; Y_3|\tilde{Y}_2, U_1, U_2) + R_1' - 3\epsilon. \tag{76}$$

Similarly, $\Pr(\mathcal{E}_7) < \epsilon$ and $\Pr(\mathcal{E}_8) < \epsilon$ for large $n$ if

$$\tilde{R}_2 < I(\tilde{Y}_2; Y_3|\tilde{Y}_1, U_1, U_2) + R_2' - 3\epsilon \tag{77a}$$
$$\tilde{R}_1 + \tilde{R}_2 < I(\tilde{Y}_1, \tilde{Y}_2; Y_3|U_1, U_2) + R_1' + R_2' - 3\epsilon. \tag{77b}$$

Now, supposed that nodes 1 and 2 send $\boldsymbol{x}_1(j^t, p^{t-1})$ and $\boldsymbol{x}_2(k^t, q^{t-1})$, respectively, in block $t$. Assume that the destination has correctly estimated $v^t, w^t, p^{t-1}$, and $q^{t-1}$. It decodes $(j^t, k^t)$ using $\tilde{\boldsymbol{y}}_1, \tilde{\boldsymbol{y}}_2$, as well as its received symbol $\boldsymbol{y}_3(t)$. The error events, where the destination wrongly decodes the source signal(s), are as follows:

$$\mathcal{E}_9 \triangleq \left(\boldsymbol{x}_1(j^t, p^{t-1}), \boldsymbol{x}_2(k^t, q^{t-1}), \boldsymbol{u}_1(p^{t-1}), \boldsymbol{u}_2(q^{t-1}),\right.$$
$$\left.\tilde{\boldsymbol{y}}_1(v^t|p^{t-1}), \tilde{\boldsymbol{y}}_2(w^t|q^{t-1}), \boldsymbol{y}_3(t)\right) \notin \mathcal{A}_\epsilon \tag{78a}$$
$$\mathcal{E}_{10} \triangleq \left(\boldsymbol{x}_1(j, p^{t-1}), \boldsymbol{x}_2(k^t, q^{t-1}), \boldsymbol{u}_1(p^{t-1}), \boldsymbol{u}_2(q^{t-1}),\right.$$
$$\left.\tilde{\boldsymbol{y}}_1(v^t|p^{t-1}), \tilde{\boldsymbol{y}}_2(w^t|q^{t-1}), \boldsymbol{y}_3(t)\right) \in \mathcal{A}_\epsilon \tag{78b}$$
$$\mathcal{E}_{11} \triangleq \left(\boldsymbol{x}_1(j^t, p^{t-1}), \boldsymbol{x}_2(k, q^{t-1}), \boldsymbol{u}_1(p^{t-1}), \boldsymbol{u}_2(q^{t-1}),\right.$$
$$\left.\tilde{\boldsymbol{y}}_1(v^t|p^{t-1}), \tilde{\boldsymbol{y}}_2(w^t|q^{t-1}), \boldsymbol{y}_3(t)\right) \in \mathcal{A}_\epsilon \tag{78c}$$
$$\mathcal{E}_{12} \triangleq \left(\boldsymbol{x}_1(j, p^{t-1}), \boldsymbol{x}_2(k, q^{t-1}), \boldsymbol{u}_1(p^{t-1}), \boldsymbol{u}_2(q^{t-1}),\right.$$
$$\left.\tilde{\boldsymbol{y}}_1(v^t|p^{t-1}), \tilde{\boldsymbol{y}}_2(w^t|q^{t-1}), \boldsymbol{y}_3(t)\right) \in \mathcal{A}_\epsilon. \tag{78d}$$

By AEP, $\Pr(\mathcal{E}_9) < \epsilon$ for large $n$. Now

$$\Pr(\mathcal{E}_{10})$$
$$= \sum_{j \neq j^t} \sum_{(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{u}_1, \boldsymbol{u}_2, \tilde{\boldsymbol{y}}_1, \tilde{\boldsymbol{y}}_2, \boldsymbol{y}_3) \in \mathcal{A}_\epsilon} p(\boldsymbol{x}_1|\boldsymbol{u}_1)$$
$$\cdot p(\boldsymbol{x}_2, \boldsymbol{u}_1, \boldsymbol{u}_2, \tilde{\boldsymbol{y}}_1, \tilde{\boldsymbol{y}}_2, \boldsymbol{y}_3) \tag{79a}$$
$$< 2^{nR_1} \sum_{(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{u}_1, \boldsymbol{u}_2, \tilde{\boldsymbol{y}}_1, \tilde{\boldsymbol{y}}_2, \boldsymbol{y}_3) \in \mathcal{A}_\epsilon} p(\boldsymbol{x}_1|\boldsymbol{u}_1) p(\boldsymbol{x}_2, \boldsymbol{u}_1, \boldsymbol{u}_2)$$
$$\cdot p(\tilde{\boldsymbol{y}}_1, \tilde{\boldsymbol{y}}_2, \boldsymbol{y}_3|\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{x}_2) \tag{79b}$$
$$\leq 2^{nR_1} 2^{n[H(U_1, U_2, X_1, X_2, \tilde{Y}_1, \tilde{Y}_2, Y_3) + \epsilon]} 2^{-n[H(X_1|U_1) - \epsilon]}$$
$$\times 2^{-n[H(U_1, U_2, X_2) - \epsilon]} 2^{-n[H(\tilde{Y}_1, \tilde{Y}_2, Y_3|U_1, U_2, X_2) - \epsilon]} \tag{79c}$$
$$\leq 2^{n[R_1 + H(X_1, \tilde{Y}_1, \tilde{Y}_2, Y_3|U_1, U_2, X_2) - H(X_1|U_1, U_2, X_2)]}$$
$$\times 2^{n[-H(\tilde{Y}_1, \tilde{Y}_2, Y_3|U_1, U_2, X_2) + 4\epsilon]} \tag{79d}$$
$$= 2^{n[R_1 - I(X_1; \tilde{Y}_1, \tilde{Y}_2, Y_3|U_1, U_2, X_2) + 4\epsilon]}. \tag{79e}$$

$\Pr(\mathcal{E}_{10})$ can be made small if

$$R_1 < I(X_1; \tilde{Y}_1, \tilde{Y}_2, Y_3|U_1, U_2, X_2) - 4\epsilon. \tag{80}$$

Similarly, $\Pr(\{\mathcal{E}_{11})$ and $\Pr(\mathcal{E}_{12})$ can be bounded if

$$R_2 < I(X_2; \tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2, X_1) - 4\epsilon \quad (81a)$$

$$R_1 + R_2 < I(X_1, X_2; \tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2) - 4\epsilon \quad (81b)$$

hold, respectively.

Combining these rate constraints for the MACF using the compress–forward strategy and the constraints for the source coding, (31a)–(31c), and adding the time-sharing random variable $Q$, we get Theorem 3. $\qquad\square$

## APPENDIX D
## ACHIEVABILITY OF SC-CF ON AWGN-MACFCS

On the AWGN-MACFCS, using SC-CF, nodes 1 and 2 send $X_1 = U_1 + V_1$ and $X_2 = U_2 + V_2$, respectively. Here $U_1$ (quantized and binned information of the previous block from $Y_1$), $V_1$ (new information from source 1), $U_2$ (old quantized and binned information of the previous block from $Y_2$), and $V_2$ (new information from source 2) are independent Gaussian random variables with power constraints $E[U_1^2] \leq P_{U1}$, $E[V_1^2] \leq P_{V1}$ $E[U_2^2] \leq P_{U2}$, and $E[V_2^2] \leq P_{V2}$, respectively. We note that $P_1 = P_{U1} + P_{V1}$ and $P_2 = P_{U2} + P_{V2}$.

The nodes receive

$$Y_1 = \sqrt{\kappa d_{21}^{-\eta}} X_2 + Z_1 = \sqrt{\kappa d_{21}^{-\eta}}(U_2 + V_2) + Z_1 \quad (82a)$$

$$Y_2 = \sqrt{\kappa d_{12}^{-\eta}} X_1 + Z_2 = \sqrt{\kappa d_{12}^{-\eta}}(U_1 + V_1) + Z_2 \quad (82b)$$

$$Y_3 = \sqrt{\kappa d_{13}^{-\eta}}(U_1 + V_1) + \sqrt{\kappa d_{23}^{-\eta}}(U_2 + V_2) + Z_3 \quad (82c)$$

where $Z_1 \sim \mathcal{N}(0, N_1)$, $Z_2 \sim \mathcal{N}(0, N_2)$, and $Z_3 \sim \mathcal{N}(0, N_3)$ are independent noise. The quantized signals are

$$\tilde{Y}_1 = Y_1 + \tilde{Z}_1 = \sqrt{\kappa d_{21}^{-\eta}} X_2 + Z_1 + \tilde{Z}_1$$
$$= \sqrt{\kappa d_{21}^{-\eta}}(U_2 + V_2) + Z_1 + \tilde{Z}_1 \quad (83a)$$

$$\tilde{Y}_2 = Y_2 + \tilde{Z}_2 = \sqrt{\kappa d_{12}^{-\eta}} X_1 + Z_2 + \tilde{Z}_2$$
$$= \sqrt{\kappa d_{12}^{-\eta}}(U_1 + V_1) + Z_2 + \tilde{Z}_2 \quad (83b)$$

where $\tilde{Z}_1 \sim \mathcal{N}(0, \tilde{N}_1)$ and $\tilde{Z}_2 \sim \mathcal{N}(0, \tilde{N}_2)$ are independent quantization noise.

Now

$$I(X_1; \tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2, X_2)$$
$$= H(\tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2, X_2) - H(\tilde{Y}_2, Y_3 | U_1, U_2, X_1, X_2)$$
$$\qquad (84a)$$

$$= H\Big(\sqrt{\kappa d_{21}^{-\eta}}(U_2 + V_2) + Z_1 + \tilde{Z}_1, \sqrt{\kappa d_{12}^{-\eta}}(U_1 + V_1) + Z_2$$
$$+ \tilde{Z}_2, \sqrt{\kappa d_{13}^{-\eta}}(U_1 + V_1) + \sqrt{\kappa d_{23}^{-\eta}}(U_2 + V_2) + Z_3 \Big| U_1, U_2,$$
$$U_2 + V_2 \Big) - H\Big(\sqrt{\kappa d_{21}^{-\eta}}(U_2 + V_2) + Z_1 + \tilde{Z}_1,$$
$$\sqrt{\kappa d_{12}^{-\eta}}(U_1 + V_1) + Z_2 + \tilde{Z}_2, \sqrt{\kappa d_{13}^{-\eta}}(U_1 + V_1)$$
$$+ \sqrt{\kappa d_{23}^{-\eta}}(U_2 + V_2) + Z_3 | U_1, U_2, U_1 + V_1, U_2 + V_2 \Big) \quad (84b)$$

$$= H\Big(Z_1 + \tilde{Z}_1, \sqrt{\kappa d_{12}^{-\eta}} V_1 + Z_2 + \tilde{Z}_2, \sqrt{\kappa d_{13}^{-\eta}} V_1 + Z_3\Big)$$
$$- H\Big(Z_1 + \tilde{Z}_1, Z_2 + \tilde{Z}_2, Z_3\Big). \quad (84c)$$

The first term is

$$H\Big(Z_1 + \tilde{Z}_1, \sqrt{\kappa d_{12}^{-\eta}} V_1 + Z_2 + \tilde{Z}_2, \sqrt{\kappa d_{13}^{-\eta}} V_1 + Z_3\Big)$$
$$= \frac{1}{2}\log(2\pi e)^3$$
$$\times \begin{vmatrix} N_1 + \tilde{N}_1 & 0 & 0 \\ 0 & \kappa d_{12}^{-\eta} P_{V1} + N_2 + \tilde{N}_2 & \kappa\sqrt{d_{12}^{-\eta} d_{13}^{-\eta}} P_{V1} \\ 0 & \kappa\sqrt{d_{12}^{-\eta} d_{13}^{-\eta}} P_{V1} & \kappa d_{13}^{-\eta} P_{V1} + N_3 \end{vmatrix}$$
$$\qquad (85a)$$
$$= \frac{1}{2}\log(2\pi e)^3 \Big[N_1 + \tilde{N}_1\Big]\Big[N_3(N_2 + \tilde{N}_2) + (\kappa d_{12}^{-\eta} N_3$$
$$+ \kappa d_{13}^{-\eta}(N_2 + \tilde{N}_2))P_{V1}\Big]. \quad (85b)$$

The second term is

$$H\Big(Z_1 + \tilde{Z}_1, Z_2 + \tilde{Z}_2, Z_3\Big)$$
$$= \frac{1}{2}\log(2\pi e)^3 \Big[(N_1 + \tilde{N}_1)(N_2 + \tilde{N}_2)N_3\Big]. \quad (86)$$

Thus

$$I(X_1; \tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2, X_2)$$
$$= \frac{1}{2}\log\left[1 + \frac{\kappa d_{12}^{-\eta} P_{V1}}{N_2 + \tilde{N}_2} + \frac{\kappa d_{13}^{-\eta} P_{V1}}{N_3}\right]. \quad (87)$$

Similarly, we can show that

$$I(X_2; \tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2, X_1)$$
$$= \frac{1}{2}\log\left[1 + \frac{\kappa d_{21}^{-\eta} P_{V2}}{N_1 + \tilde{N}_1} + \frac{\kappa d_{23}^{-\eta} P_{V2}}{N_3}\right]. \quad (88)$$

Now, we evaluate $I(X_1, X_2; \tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2)$ and $H(\tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2)$ as shown in (89)–(90c) at the top of the following page.

The first term is given in (90) and the second term is

$$H(\tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2, X_1, X_2)$$
$$= \frac{1}{2}\log(2\pi e)^3 (N_1 + \tilde{N}_1)(N_2 + \tilde{N}_2)N_3. \quad (91)$$

Hence

$$I(X_1, X_2; \tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2)$$
$$= \frac{1}{2}\log\Big[1 + \frac{\kappa d_{12}^{-\eta} P_{V1}}{N_2 + \tilde{N}_2} + \frac{\kappa d_{21}^{-\eta} P_{V2}}{N_1 + \tilde{N}_1} + \frac{\kappa d_{13}^{-\eta} P_{V1}}{N_3}$$
$$+ \frac{\kappa d_{23}^{-\eta} P_{V2}}{N_3} + \frac{\kappa^2 d_{12}^{-\eta} d_{21}^{-\eta} P_{V1} P_{V2}}{(N_1 + \tilde{N}_1)(N_2 + \tilde{N}_2)}$$
$$+ \frac{\kappa^2 d_{21}^{-\eta} d_{13}^{-\eta} P_{V1} P_{V2}}{(N_1 + \tilde{N}_1)N_3} + \frac{\kappa^2 d_{12}^{-\eta} d_{23}^{-\eta} P_{V1} P_{V2}}{(N_2 + \tilde{N}_2)N_3}\Big]. \quad (92a)$$

$$I(X_1, X_2; \tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2)$$
$$= H(\tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2) - H(\tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2, X_1, X_2) \tag{89}$$

$$H(\tilde{Y}_1, \tilde{Y}_2, Y_3 | U_1, U_2)$$
$$= \frac{1}{2} \log(2\pi e)^3 \begin{vmatrix} \kappa d_{21}^{-\eta} P_{V2} + N_1 + \tilde{N}_1 & 0 & \kappa\sqrt{d_{21}^{\eta} d_{23}^{-\eta}} P_{V2} \\ 0 & \kappa d_{12}^{-\eta} P_{V1} + N_2 + \tilde{N}_2 & \kappa\sqrt{d_{12}^{-\eta} d_{13}^{-\eta}} P_{V1} \\ \kappa\sqrt{d_{21}^{\eta} d_{23}^{-\eta}} P_{V2} & \kappa\sqrt{d_{12}^{-\eta} d_{13}^{-\eta}} P_{V1} & \kappa d_{13}^{-\eta} P_{V1} + \kappa d_{23}^{-\eta} P_{V2} + N_3 \end{vmatrix} \tag{90a}$$

$$= \frac{1}{2} \log(2\pi e)^3 \Big[ \kappa d_{12}^{-\eta} P_{V1} N_3 (N_1 + \tilde{N}_1) + \kappa d_{21}^{-\eta} P_{V2} N_3 (N_2 + \tilde{N}_2)$$
$$+ \kappa d_{13}^{-\eta} P_{V1}(N_1 + \tilde{N}_1)(N_2 + \tilde{N}_2) + \kappa d_{23}^{-\eta} P_{V2}(N_1 + \tilde{N}_1)(N_2 + \tilde{N}_2) + \kappa^2 d_{12}^{-\eta} d_{21}^{-\eta} P_{V1} P_{V2} N_3$$
$$\kappa^2 d_{21}^{-\eta} d_{13}^{-\eta} P_{V1} P_{V2}(N_2 + \tilde{N}_2) + \kappa^2 d_{12}^{-\eta} d_{23}^{-\eta} P_{V1} P_{V2}(N_1 + \tilde{N}_1) + (N_1 + \tilde{N}_1)(N_2 + \tilde{N}_2) N_3 \Big] \tag{90b}$$

$$\triangleq \frac{1}{2} \log(2\pi e)^3 \mathcal{B}_1. \tag{90c}$$

---

We can show that

$$I(U_1, Y_3 | U_2) = \frac{1}{2} \log\left[1 + \frac{\mathcal{C}_1}{\mathcal{B}_2}\right] \tag{93a}$$

$$I(U_2, Y_3 | U_1) = \frac{1}{2} \log\left[1 + \frac{\mathcal{C}_2}{\mathcal{B}_2}\right] \tag{93b}$$

$$I(U_1, U_2, Y_3) = \frac{1}{2} \log\left[1 + \frac{\mathcal{C}_1 + \mathcal{C}_2}{\mathcal{B}_2}\right] \tag{93c}$$

where $\mathcal{B}_2 \triangleq \kappa d_{13}^{-\eta} P_{V1} + \kappa d_{23}^{-\eta} P_{V2} + N_3$, $\mathcal{C}_1 \triangleq \kappa d_{13}^{-\eta} P_{U1} = \kappa d_{13}^{-\eta}(P_1 - P_{V1})$, and $\mathcal{C}_2 \triangleq \kappa d_{23}^{-\eta} P_{U2} = \kappa d_{23}^{-\eta}(P_2 - P_{V2})$. Also

$$I(\tilde{Y}_1; Y_1 | X_1, U_1) = \frac{1}{2} \log\left[1 + \frac{\kappa d_{21}^{-\eta} P_2 + N_1}{\tilde{N}_1}\right] \tag{94a}$$

$$\triangleq \frac{1}{2} \log\left[1 + \frac{\mathcal{D}_1}{\tilde{N}_1}\right] \tag{94b}$$

$$I(\tilde{Y}_2; Y_2 | X_2, U_2) = \frac{1}{2} \log\left[1 + \frac{\kappa d_{12}^{-\eta} P_1 + N_2}{\tilde{N}_2}\right] \tag{94c}$$

$$\triangleq \frac{1}{2} \log\left[1 + \frac{\mathcal{D}_2}{\tilde{N}_2}\right]. \tag{94d}$$

We write

$$I(\tilde{Y}_1; Y_3 | \tilde{Y}_2, U_1, U_2) = H(Y_3 | \tilde{Y}_2, U_1, U_2) - H(Y_3 | \tilde{Y}_1, \tilde{Y}_2, U_1, U_2).$$

Evaluating and simplifying, we get

$$I(\tilde{Y}_1; Y_3 | \tilde{Y}_2, U_1, U_2)$$
$$= \frac{1}{2} \log\left[1 + \frac{\kappa^2 d_{23}^{-\eta} d_{21}^{-\eta} P_{V2}^2(\kappa d_{12} P_{V1} + N_2 + \tilde{N}_2)}{\mathcal{B}_1}\right]. \tag{95}$$

So, constraint (34a) becomes

$$\left(1 + \frac{\mathcal{D}_1}{\tilde{N}_1}\right) < \left(1 + \frac{\mathcal{C}_3}{\mathcal{B}_1}\right)\left(1 + \frac{\mathcal{C}_1}{\mathcal{B}_2}\right) \tag{96}$$

where $\mathcal{C}_3 \triangleq \kappa^2 d_{23}^{-\eta} d_{21}^{-\eta} P_{V2}^2(\kappa d_{12}^{-\eta} P_{V1} + N_2 + \tilde{N}_2)$. Similarly, constraint (34b) becomes

$$\left(1 + \frac{\mathcal{D}_2}{\tilde{N}_2}\right) < \left(1 + \frac{\mathcal{C}_4}{\mathcal{B}_1}\right)\left(1 + \frac{\mathcal{C}_2}{\mathcal{B}_2}\right) \tag{97}$$

where $\mathcal{C}_4 \triangleq \kappa^2 d_{13}^{-\eta} d_{12}^{-\eta} P_{V1}^2(\kappa d_{21}^{-\eta} P_{V2} + N_1 + \tilde{N}_1)$. Finally, constraint (34c) becomes

$$\left(1 + \frac{\mathcal{D}_1}{\tilde{N}_1}\right)\left(1 + \frac{\mathcal{D}_2}{\tilde{N}_2}\right) < \left(1 + \frac{\mathcal{C}_3 + \mathcal{C}_4}{\mathcal{B}_1}\right)\left(1 + \frac{\mathcal{C}_1 + \mathcal{C}_2}{\mathcal{B}_2}\right). \tag{98}$$

We note that the achievability derived in Theorem 3 makes use of the Markov lemma [24, Lemma 4.1], which requires strong typicality. Though strong typicality does not extend to continuous random variables, we can generalize the Markov lemma for Gaussian inputs and thus show that the rate governed by (87), (88), and (92a) is achievable [19].

## ACKNOWLEDGMENT

## REFERENCES

[1] L. Ong and M. Motani, "Achievable rates for the multiple-access channels with feedback and correlated sources," in *Proc. 43rd Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Sep. 2005.

[2] L. Ong and M. Motani, "The multiple-access channel with feedback and correlated sources," in *Proc. IEEE Int. Symp. Information Theory*, Seattle, WA, Jul. 2006, pp. 2129–2133.

[3] D. Slepian and J. K. Wolf, "A coding theorem for multiple-access channels with correlated sources," *Bell Syst. Tech. J.*, vol. 52, no. 7, pp. 1037–1076, Sep. 1973.

[4] T. M. Cover, A. A. El-Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 6, pp. 648–657, Nov. 1980.

[5] W. Kang and S. Ulukus, "An outer bound for multiple-access channels with correlated sources," in *Proc. Conf. Information Science and Systems*, Princeton, NJ, Mar. 2006, pp. 7–12.

[6] T. M. Cover and C. S. K. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 3, pp. 292–298, May 1981.

[7] L. H. Ozarow, "The capacity of the white Gaussian multiple-access channel with feedback," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 4, pp. 623–629, Jul. 1984.

[8] R. C. King, "Multiple Access Channels with Generalized Feedback," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1978.

[9] F. M. J. Willems, "Information Theoretical Results for the Discrete Memoryless Multiple Access Channel," Ph.D. dissertation, Katholieke Universiteit, Leuven, Belgium, 1982.

[10] A. B. Carleial, "Multiple-access channels with different generalized feedback signals," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 6, pp. 841–850, Nov. 1982.

[11] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[12] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-part II: Implementation aspects and performance analysis," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[13] A. D. Murugan, P. K. Gopala, and H. El-Gamal, "Correlated sources over wireless channels: Cooperative source-channel coding," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 988–998, Aug. 2004.

[14] H. Liao, "Multiple Access Channel," Ph.D. dissertation, Univ. Hawaii, Honolulu, HI, 1972.

[15] R. Ahlswede, "The capacity of a channel with two senders and two receivers," *Ann. Probab.*, vol. 2, pp. 805–814, Oct. 1974.

[16] J. Barros and S. D. Servetto, "Network information flow with correlated sources," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 155–170, Jan. 2006.

[17] M. Gastpar, A. J. v. W. P. Gupta, and G. Kramer, Eds., "On source-channel communication in networks," in *Proc. DIMACS Workshop on Network Information Theory*, Piscataway, NJ, Mar. 2003, pp. 217–238.

[18] M. Gastpar, "Cut-set arguments for source-channel networks," in *Proc. IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 34.

[19] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[21] T. M. Cover and A. A. El-Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.

[22] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.

[23] M. Salehi, "Improved computational efficiency in derivation of capacity regions in network information theory," *Computers and Electr. Eng. (Elsevier)*, vol. 22, no. 4, pp. 247–256, 1999.

[24] T. Berger, "Multiterminal source coding," in *Lecture Notes Presented at the 1977 CISM Summer School*, Udine, Italy, Jul. 1977, pp. 171–231.