

Distributed Joint Source Channel Coding on a Multiple Access Channel with Side Information

R Rajesh
Dept. of ECE
Indian Institute of Science
Bangalore, India
Email: rajesh@pal.ece.iisc.ernet.in

V K Varshneya
IBM India Research Lab
Bangalore, India
Email: virendra@pal.ece.iisc.ernet.in

Vinod Sharma
Dept. of ECE
Indian Institute of Science
Bangalore, India
Email: vinod@ece.iisc.ernet.in

Abstract— We consider the problem of transmission of several discrete sources over a multiple access channel (MAC) with side information at the sources and the decoder. Source-channel separation does not hold for this channel. Sufficient conditions are provided for transmission of sources with a given distortion. The channel could have continuous alphabets (Gaussian MAC is a special case). Various previous results are obtained as special cases.

Keywords: Multiple access channel, Side information, Distributed joint source-channel coding.

I. INTRODUCTION AND SURVEY

The use of sensor networks to monitor spatio-temporal processes is gaining interest due to the availability of using inexpensive, distributed sensor nodes. These nodes have limited computational and storage capabilities and very limited energy [1]. These sensor nodes need to transmit their observations to a fusion center which uses this data to estimate the sensed random field. Since transmission is very energy intensive, it is important to minimize transmission.

The proximity of the sensing nodes to each other induces high correlations between the observations of adjacent sensors. One can exploit these correlations to compress the transmitted data significantly. Furthermore, some of the nodes can be more powerful and can act as cluster heads ([2]). Neighboring nodes can first transmit their data to a cluster head which can further compress information before transmission to the fusion center. The transmission of data from sensor nodes to their cluster-head is usually through a multiple access channel (MAC). At the fusion center the underlying physical process is estimated. The main trade-off possible is between the rates at which the sensors send their observations and the distortion incurred in the estimation at the fusion center. The availability of side information at the encoders and/or the decoder can reduce the rate of transmission ([9], [21]).

The above considerations open up new interesting problems in multi-user information theory and the quest for finding the optimal performance for various models of sources, channels and side information have made this an active area of research. The optimal solution is unknown except in a few simple cases. In this paper a joint source channel coding approach

This work was partially supported by the DRDO-IISc program on Advanced Research in Mathematical Engineering.

is discussed under various assumptions of side information and distortion criteria. The paper gives sufficient conditions for transmission of finite alphabet sources over a discrete alphabet MAC. Suitable examples are also given. These conditions are further extended to continuous alphabet channels. Our results generalize the previous results available on this problem.

In the following we survey the related literature. The seminal paper, Cover, El Gamal and Salehi [5] provided sufficient conditions for transmitting losslessly correlated observations over a multiple access channel (MAC). However single letter characterization of the capacity region is still unknown. Indeed Duek [7] proved that the conditions given in [5] are only sufficient and may not be necessary. Joint source channel coding with discrete sources, channels and side information is discussed in [18]. In this paper we extend the side information scenario and also consider continuous alphabet channels. This covers the important special case of the Gaussian MAC.

The classical results on distributed source coding are in [3], [17] and [21]. More recent results are [9] and [19] (see also references therein). Exact rate region for the distributed Gaussian source coding problem is provided in [20]. The scaling laws for the problem without side information are discussed in [10] and it is shown that the standard practice of separating source from channel coding may require exponential growth in communication bandwidth, as a function of number of sensors. One necessary and two sufficient conditions for transmission of bi-variate Gaussian sources over a Gaussian MAC (GMAC) are given in [12]. The joint source-channel coding problem also bears relationship to the CEO problem [4]. The Gaussian version of the CEO problem is studied in [13].

The organization of the paper is as follows. The discrete alphabet sources on a discrete alphabet MAC are discussed in Section II and previous results are shown as special cases. An example is provided to show the effectiveness of side information in Section III. In Section IV the results of Section II are generalized to multiple sources. Extensions to continuous channels are given in Section V. Section VI gives the proof of the main theorem. Section VII summarizes the results.

II. DISCRETE ALPHABET SOURCES OVER A DISCRETE CHANNEL

In this section we consider the transmission of discrete memoryless dependent sources, through a discrete memoryless multiple access channel with side information and fidelity requirements at the decoder (see Fig 1).

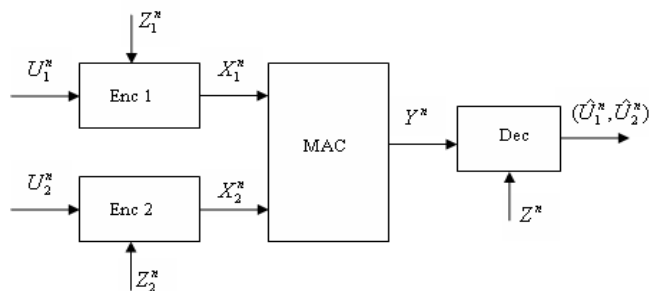


Fig. 1. Transmission of correlated sources over a MAC with side information

We consider discrete alphabet sources (U_1, U_2) and side information random variables Z_1, Z_2, Z with a known joint distribution $F(u_1, u_2, z_1, z_2, z)$. Side information Z_i is available to source i , $i = \{1, 2\}$ and the decoder has side information Z . The random vector sequence $\{(U_{1n}, U_{2n}, Z_{1n}, Z_{2n}, Z_n), n \geq 1\}$ formed from the source outputs and the side information with distribution F is independent identically distributed in time. The sources transmit their codewords X_i 's to a single decoder through a memoryless multiple access channel. The channel output Y has distribution $P(y|x_1, x_2)$ if x_1 and x_2 are transmitted at that time. The decoder receives Y and also has access to the side information Z . It uses Y and Z to estimate the sensor observations U_i as \hat{U}_i , $i = \{1, 2\}$. It is of interest to find encoders and a decoder such that the encoded $\{U_{1n}, U_{2n}, n \geq 1\}$ can be transmitted over the given MAC such that $E[d_1(U_1, \hat{U}_1)] \leq D_1$ and $E[d_2(U_2, \hat{U}_2)] \leq D_2$ where d_i are bounded distortion measures upper bounded by d_{max} and D_i are the given distortion constraints. The encoders at the two users do not communicate with each other except via the side information. Source channel separation does not hold in this case.

We will denote $\{U_{ij}\}, j = 1, 2, \dots, n$ by $U_i^n, i = 1, 2$.

Definition : The source (U_1^n, U_2^n) can be transmitted over the multiple access channel with distortions $\mathbf{D} \triangleq (D_1, D_2)$ if for any $\epsilon > 0$ there is an n_0 such that for all $n > n_0$ there exist encoders $f_{E,i}^n : U_i^n \times Z_i^n \rightarrow X_i^n, i \in \{1, 2\}$ and a decoder $f_D^n : Y^n \times Z^n \rightarrow (\hat{U}_1^n, \hat{U}_2^n)$ such that $\frac{1}{n} E \left[\sum_{j=1}^n d(U_{ij}, \hat{U}_{ij}) \right] \leq D_i + \epsilon, i \in \{1, 2\}$ where $(\hat{U}_1^n, \hat{U}_2^n) = f_D(Y^n, Z^n)$.

We denote the joint distribution of (U_1, U_2) by $p(u_1, u_2)$ and let $p(y|x_1, x_2)$ be the transition probabilities of the MAC. Since the MAC is memoryless, $p(y^n|x_1^n, x_2^n) = \prod_{j=1}^n p(y_j|x_{1j}, x_{2j})$. $X \leftrightarrow Y \leftrightarrow Z$ will indicate that $\{X, Y, Z\}$ form a Markov chain.

Now we state the main Theorem.

Theorem 1: A source can be transmitted over the multiple

access channel with distortions (D_1, D_2) if there exist random variables (W_1, W_2, X_1, X_2) such that

$$1. p(u_1, u_2, z_1, z_2, z, w_1, w_2, x_1, x_2, y) = p(u_1, u_2, z_1, z_2, z)p(w_1|u_1, z_1)p(w_2|u_2, z_2)p(x_1|w_1)p(x_2|w_2)p(y|x_1, x_2).$$

2. There exists a function $f_D : \mathcal{W}_1 \times \mathcal{W}_2 \times \mathcal{Z} \rightarrow (\hat{\mathcal{U}}_1 \times \hat{\mathcal{U}}_2)$ such that $E[d(U_i, \hat{U}_i)] \leq D_i + \epsilon, i = 1, 2$ and the constraints

$$\begin{aligned} I(U_1, Z_1; W_1|W_2, Z) &< I(X_1; Y|X_2, W_2, Z), \\ I(U_2, Z_2; W_2|W_1, Z) &< I(X_2; Y|X_1, W_1, Z), \\ I(U_1, U_2, Z_1, Z_2; W_1, W_2|Z) &< I(X_1, X_2; Y|Z) \end{aligned} \quad (1)$$

are satisfied where $U_i, Z_i, \mathcal{Z}, W_i, X_i, Y, \hat{U}_i$ are the sets in which $U_i, Z_i, Z, W_i, X_i, Y, \hat{U}_i$ take values and $\epsilon > 0$ is any small constant. ■

The proof of the theorem is given in section VI. Due to strong typicality, the proof of this theorem works if in addition to estimating (U_1, U_2) we were also estimating at the receiver a function $g(U_1, U_2)$ and the distortion measures were jointly over $(U_1, U_2, \hat{U}_1, \hat{U}_2)$. We need to pick f_D and W_1, W_2 such that in addition to $E[d(U_i, \hat{U}_i)] \leq D_i + \epsilon, i = 1, 2$ they also satisfy the corresponding extra conditions.

This theorem is generalized to the continuous channel alphabet case in Section V. This will cover the practically important case of a Gaussian MAC.

A. Special Cases

In the following, to obtain the results for the lossless case, we will take $\hat{U}_i = U_i, i = 1, 2$ and the distortion measure as Hamming distance. Our results generalize the system in [18].

1) *Correlated sources with lossless transmission over multiuser channels with receiver side information [11]*: If we consider $(Z_1, Z_2) \perp (U_1, U_2), W_1 = U_1$ and $W_2 = U_2$ then we recover the conditions

$$\begin{aligned} H(U_1|U_2, Z) &< I(X_1; Y|X_2, U_2, Z), \\ H(U_2|U_1, Z) &< I(X_2; Y|X_1, U_1, Z), \\ H(U_1, U_2|Z) &< I(X_1, X_2; Y|Z) \end{aligned} \quad (2)$$

in Theorem 2.1 in [11].

2) *Lossy multiple access communication*: Take $(Z_1, Z_2, Z) \perp (U_1, U_2)$. In this case the constraints in (1) reduce to

$$\begin{aligned} I(U_1; W_1|W_2) &< I(X_1; Y|X_2, W_2), \\ I(U_2; W_2|W_1) &< I(X_2; Y|X_1, W_1), \\ I(U_1, U_2; W_1, W_2) &< I(X_1, X_2; Y). \end{aligned} \quad (3)$$

This is an immediate generalization of [5] to the lossy case.

3) *Lossless multiple access communication with correlated sources*: Take $(Z_1, Z_2, Z) \perp (U_1, U_2)$ and $W_1 = U_1$ and $W_2 = U_2$ then the constraints of (1) reduce to the conditions in [5].

4) *Lossy distributed source coding with side information:*

The multiple access channel is taken as a dummy channel which reproduces its inputs. In this case we obtain that the sources can be coded with rates R_1 and R_2 to obtain the specified distortions at the decoder if

$$\begin{aligned} R_1 &> I(U_1, Z_1; W_1 | W_2, Z), \\ R_2 &> I(U_2, Z_2; W_2 | W_1, Z), \\ R_1 + R_2 &> I(U_1, U_2, Z_1, Z_2; W_1, W_2 | Z). \end{aligned} \quad (4)$$

This recovers the result in [19], and generalizes the results in [9], [17] and [21].

5) *Mixed Side Information [8]:* Consider the system in Fig 1(c) of [8]. This has the system in Fig 1(a) and (b) of [8] as special cases. The results of Fig 1(c) can be recovered from our Theorem if we take X, Y, Z, W in [8] as $U_1 = X, Z = (Z, Y), Z_1 = Y, W_1 = W$. U_2 and Z_2 are taken to be constants.

III. EXAMPLE

We provide an example to show the reduction possible in transmission rates by exploiting the correlation between the sources, the side information and the permissible distortions.

Consider (U_1, U_2) with the joint distribution: $P(U_1 = 0; U_2 = 0) = P(U_1 = 1; U_2 = 1) = 1/3; P(U_1 = 1; U_2 = 0) = P(U_1 = 0; U_2 = 1) = 1/6$. If we use independent encoders which do not exploit the correlation among the sources then we need $R_1 \geq H(U_1) = 1$ bit and $R_2 \geq H(U_2) = 1$ bit for lossless coding of the sources. If we use the coding scheme in [23], then $R_1 \geq H(U_1|U_2) = 0.918$ bits, $R_2 \geq H(U_2|U_1) = 0.918$ bits and $R_1 + R_2 \geq H(U_1, U_2) = 1.918$ bits.

Next consider a multiple access channel such that $Y = X_1 + X_2$ where X_1 and X_2 take values from the alphabet $\{0, 1\}$ and Y takes values from the alphabet $\{0, 1, 2\}$. This does not satisfy the separation conditions ([16]). The sum capacity C of such a channel with independent X_1 and X_2 is 1.5 bits and if we use source-channel separation, the given sources are not admissible because $H(U_1, U_2) > C$. Now we use a joint-source channel code to improve the capacity of the channel. Take $U_1 = X_1$ and $U_2 = X_2$. Then the capacity of the channel is improved to 1.585 bits. This is still not enough to transmit the sources over the given MAC. Next we exploit the side information.

The side-information random variables are generated as follows. Z_1 is generated from U_2 by using a binary symmetric channel (BSC) with cross over probability $p = 0.3$. Similarly Z_2 is generated from U_1 by using the same BSC. Let $Z = (Z_1, Z_2, V)$, V is generated by $V = U_1 \cdot U_2 \cdot N$, where N is a binary random variable with $P(N = 0) = P(N = 1) = 0.5$ independent of U_1 and U_2 and \cdot denotes the logical AND operation. Then from (1) if we use just the side information Z_1 the sum rate for sources can be 1.8 bits and by symmetry the same holds for Z_2 . If we use Z_1 and Z_2 then we get 1.683 bits as the sum rate and 1.606 bits if only V is used. If all the information in Z_1, Z_2, Z is used then we need

$R_1 + R_2 \geq 1.4120$ bits. Thus with the aid of Z_1, Z_2, Z we can transmit (U_1, U_2) losslessly over the MAC even with independent X_1 and X_2 .

Next we consider the distortion criterion to be the Hamming distance and the allowable distortion as 4%. Then for compressing the individual sources without side information $R_1 \geq H(p) - H(d) = 0.758$ bits, where $H(x) = -x \log_2(x) - (1-x) \log_2(1-x)$. Thus we still cannot transmit (U_1, U_2) with this distortion when (X_1, X_2) are independent. If U_1 and U_2 are encoded, exploiting their correlations, (X_1, X_2) can be correlated. Assume side information $Z = (Z_1, Z_2)$ to be available at the decoder only. Thus we need $R_1 \geq I(U_1; W_1) - I(Z_1; W_1)$ where W_1 is an auxiliary random variable generated from U_1 . W_1 and Z_1 are related by a cascade of a BSC with crossover probability 0.3 with a BSC with crossover probability 0.04. This implies that $R_1 \geq 0.6577$ bits and $R_2 \geq 0.6577$ bits.

IV. EXTENSION TO MULTIPLE SOURCES

The above results can be generalized to the multiple source (≥ 2) case. Let $\mathcal{S} = \{1, 2, \dots, m\}$ be the set of sensors. Let $p(u_1, u_2, \dots, u_m)$ denote the joint distribution of the sources.

Theorem 2: Sources $(U_i^n, i \in \mathcal{S})$ can be communicated in a distributed fashion over the discrete memoryless multiple access channel $p(y|x_i, i \in \mathcal{S})$ with distortions $(D_i, i \in \mathcal{S})$ if there exist auxiliary random variables $(W_i, X_i, i \in \mathcal{S})$ satisfying

$$\begin{aligned} 1. \quad &p(u_i, z_i, z, w_i, x_i, y, i \in \mathcal{S}) = \\ &p(u_i, z_i, z, i \in \mathcal{S}) p(y|x_i, i \in \mathcal{S}) \\ &\prod_{j \in \mathcal{S}} p(w_j|u_j, z_j) p(x_j|w_j) \end{aligned}$$

2. There exists a function $f_D : \prod_{j \in \mathcal{S}} (W_j) \times \mathcal{Z} \rightarrow (\hat{U}_i, i \in \mathcal{S})$ such that $E[d(U_i, \hat{U}_i)] \leq D_i + \epsilon$, $i \in \mathcal{S}$ and the constraints

$$I(U_A, Z_A; W_A | W_{A^c}, Z) < I(X_A; Y | X_{A^c}, W_{A^c}, Z) \\ \text{for all } A \subset \mathcal{S}$$

are satisfied where A^c is the compliment of set A . ■

The proof is along the lines of the proof of Theorem 1 and is omitted for brevity.

V. DISCRETE SOURCES OVER A CONTINUOUS ALPHABET CHANNEL

Theorem 1 can be extended to discrete alphabet sources and continuous channels. We provide the changes required in the proof of Theorem 1 when the channel alphabets are uncountable (we will take them real-valued). An important practical example of this is Gaussian MAC.

Now (U_1, U_2, Z_1, Z_2, Z) will take values in a discrete alphabet but (X_1, X_2, Y) can be real valued. For example, we show in [14], [15] that if the channel is a Gaussian MAC then one should choose (X_1, X_2) to be jointly Gaussian. Indeed in [14], [15] we also provide a distributed algorithm to map discrete (U_1, U_2, Z_1, Z_2) into jointly Gaussian (X_1, X_2) which satisfy the Markovian assumptions in Theorem 1.

The main difference in the proof for this scenario is that for real valued r.v.s we cannot use strong typicality. One

could use weak typicality but then one needs to deal with distortion typical sets as in [6], Chapter 13. But to include this will need substantial changes in the proofs. However, in the current setup, since $(U_1, U_2, Z_1, Z_2, Z, W_1, W_2)$ are discrete, we can encode (U_i^n, Z_i^n) into W_i^n , $i = 1, 2$ such that these form *strongly* typical sequences as in the original proof. By Markov Lemma $(U_1^n, Z_1^n, W_1^n, W_2^n, Z_2^n, U_2^n)$ form a strongly typical sequence. At the decoder we find the unique (W_1^n, W_2^n) which are *weakly* typical with the received sequence $(x_1(W_1^n), x_2(W_2^n), Y^n, Z^n)$ (although the proof uses strongly typical sequences, if we use weak typicality at the decoder this part of the proof goes through). With a high probability we obtain the original sequences (W_1^n, W_2^n) which were coded and transmitted and hence is *strongly* typical with (U_1^n, U_2^n, Z^n) . Thus in this set the distortion constraints are satisfied.

Similar extension to the multi-user Theorem 2 also holds.

The above proof does not work for continuous alphabet sources. But recently we have worked throughout with weak typicality and distortion typical sequences to generalize Theorem 1 to continuous alphabet sources and channels (see [15]).

VI. PROOF OF THEOREM 1

The scheme involves distributed quantization of the sources and the side information followed by correlation preserving mapping to the channel codewords. The decoding approach involves first decoding the quantized version (W_1, W_2) of (U_1, U_2) , and then obtaining estimate (\hat{U}_1, \hat{U}_2) as a function of (W_1, W_2) and the decoder side information.

We show the achievability of all points in the rate region (1).

Proof: Fix $p(w_1|u_1, z_1), p(w_2|u_2, z_2), p(x_1|w_1), p(x_2|w_2)$ as well as $f_D^n(\cdot)$ satisfying the distortion constraints.

Codebook Generation: Let $R'_i = I(U_i, Z_i; W_i) + \delta$, $i = 1, 2$ for some $\delta > 0$. Generate $2^{nR'_i}$ codewords of length n , sampled iid from the marginal distribution $p(w_i)$, $i = 1, 2$. For each w_i^n independently generate sequence X_i^n according to $\prod_{j=1}^n p(x_{ij}|w_{ij})$, $i = 1, 2$. Call these sequences $x_i(w_i^n)$, $i = 1, 2$. Reveal the codebooks to the encoders and the decoder.

Encoding: For $i = 1, 2$, given the source sequence U_i^n and Z_i^n , the i^{th} encoder looks for a codeword W_i^n such that $(U_i^n, Z_i^n, W_i^n) \in T_\epsilon^n(U_i, Z_i, W_i)$ and then transmits $X_i(W_i^n)$ where $T_\epsilon^n(\cdot)$ is the set of ϵ strongly typical sequences ([6]) of length n .

Decoding: Upon receiving Y^n , the decoder finds the unique (W_1^n, W_2^n) pair such that $(W_1^n, W_2^n, x_1(W_1^n), x_2(W_2^n), Y^n, Z^n) \in T_\epsilon^n$. If it fails to find such a unique pair, the decoder declares an error and incurs a maximum distortion of d_{max} .

In the following we show that the probability of error for the encoding decoding scheme tends to zero as $n \rightarrow \infty$. The error can occur because of the following four events **E1-E4**. We show that $P(\mathbf{E}i) \rightarrow 0$, for $i = 1, 2, 3, 4$.

E1 The encoders do not find the codewords. However from rate distortion theory [6] $\lim_{n \rightarrow \infty} P(E_1) = 0$ if $R'_i > I(U_i, Z_i; W_i)$, $i = 1, 2$.

E2 The codewords are not jointly typical with Z^n . Probability of this event goes to zero from the extended Markov Lemma [18].

E3 There exists another codeword \hat{w}_1^n such that $(\hat{w}_1^n, W_2^n, x_1(\hat{w}_1^n), x_2(W_2^n), Y^n, Z^n) \in T_\epsilon^n$. Define $\alpha \triangleq (\hat{w}_1^n, W_2^n, x_1(\hat{w}_1^n), x_2(W_2^n), Y^n, Z^n)$. Then,

$$\begin{aligned} P(\mathbf{E3}) &= Pr\{\text{There is } \hat{w}_1^n \neq w_1^n : \alpha \in T_\epsilon^n\} \\ &\leq \sum_{\hat{w}_1^n \neq w_1^n : (\hat{w}_1^n, W_2^n, Z^n) \in T_\epsilon^n} Pr\{\alpha \in T_\epsilon^n\} \end{aligned} \quad (5)$$

The probability term inside the summation in (5) is

$$\begin{aligned} &\leq \sum_{(x_1(\cdot), x_2(\cdot), y^n) : \alpha \in T_\epsilon^n} Pr\{x_1(\hat{w}_1^n), x_2(w_2^n), y^n | \hat{w}_1^n, w_2^n, z^n\} \\ &\leq \sum_{(x_1(\cdot), x_2(\cdot), y^n) : \alpha \in T_\epsilon^n} Pr\{x_1(\hat{w}_1^n) | \hat{w}_1^n\} Pr\{x_2(w_2^n), y^n | w_2^n, z^n\} \\ &\leq \sum_{(x_1(\cdot), x_2(\cdot), y^n) : \alpha \in T_\epsilon^n} 2^{-n\{H(X_1|W_1) + H(X_2, Y|W_2, Z) - 4\epsilon\}} \\ &\leq 2^{nH(X_1, X_2, Y|W_1, W_2, Z)} 2^{-n\{H(X_1|W_1) + H(X_2, Y|W_2, Z) - 4\epsilon\}} \end{aligned}$$

But from hypothesis, we have

$$\begin{aligned} &H(X_1, X_2, Y|W_1, W_2, Z) - H(X_1|W_1) - H(X_2, Y|W_2, Z) \\ &= H(X_1|W_1) + H(X_2|W_2) + H(Y|X_1, X_2) \\ &\quad - H(X_1|W_1) - H(X_2, Y|W_2, Z) \\ &= H(Y|X_1, X_2) - H(Y|X_2, W_2, Z) \\ &= H(Y|X_1, X_2, W_2, Z) - H(Y|X_2, W_2, Z) \\ &= -I(X_1; Y|X_2, W_2, Z). \end{aligned}$$

Hence,

$$\begin{aligned} &Pr\{(\hat{w}_1^n, W_2^n, x_1(\hat{w}_1^n), x_2(W_2^n), Y^n, Z^n) \in T_\epsilon^n\} \\ &\leq 2^{-n\{I(X_1; Y|X_2, W_2, Z) - 6\epsilon\}}. \end{aligned} \quad (6)$$

Then from (5)

$$\begin{aligned} P(\mathbf{E3}) &\leq \sum_{\hat{w}_1^n \neq w_1^n : (\hat{w}_1^n, w_2^n, z^n) \in T_\epsilon^n} 2^{-n\{I(X_1; Y|X_2, W_2, Z) - 6\epsilon\}} \\ &\leq |\{\hat{w}_1^n : (\hat{w}_1^n, w_2^n, z^n) \in T_\epsilon^n\}| 2^{-n\{I(X_1; Y|X_2, W_2, Z) - 6\epsilon\}} \\ &\leq |\{\hat{w}_1^n\}| \cdot Pr\{\hat{w}_1^n, w_2^n, z^n \in T_\epsilon^n\} \cdot 2^{-n\{I(X_1; Y|X_2, W_2, Z) - 6\epsilon\}} \\ &\leq 2^{n\{I(U_1, Z_1; W_1) + \epsilon\}} \cdot 2^{-n\{I(W_1; W_2, Z) - 3\epsilon\}} \\ &\quad \cdot 2^{-n\{I(X_1; Y|X_2, W_2, Z) - 6\epsilon\}} \\ &= 2^{n\{I(U_1, Z_1; W_1|W_2, Z)\}} \cdot 2^{-n\{I(X_1; Y|X_2, W_2, Z) - 10\epsilon\}}. \end{aligned} \quad (7)$$

The R.H.S of the above inequality tends to zero if $I(U_1, Z_1; W_1|W_2, Z) < I(X_1; Y|X_2, W_2, Z)$. In (7) we have used the fact that

$$\begin{aligned} &I(U_1, Z_1; W_1) - I(W_1; W_2, Z) \\ &= H(W_1|W_2, Z) - H(W_1|U_1, Z_1) \\ &= H(W_1|W_2, Z) - H(W_1|U_1, Z_1, W_2, Z) \\ &= I(U_1, Z_1; W_1|W_2, Z). \end{aligned}$$

Similarly, by symmetry of the problem we require $I(U_2, Z_2; W_2|W_1, Z) < I(X_2; Y|X_1, W_1, Z)$.

E4 There exist other codewords \hat{w}_1^n and \hat{w}_2^n such that $\alpha \triangleq (\hat{w}_1^n, \hat{w}_2^n, x_1(\hat{w}_1^n), x_2(\hat{w}_2^n), y^n, z^n) \in T_\epsilon^n$. Then,

$$\begin{aligned} P(\mathbf{E4}) &= Pr\{\text{There is } (\hat{w}_1^n, \hat{w}_2^n) \neq (w_1^n, w_2^n) : \alpha \in T_\epsilon^n\} \\ &\leq \sum_{(\hat{w}_1^n, \hat{w}_2^n) \neq (w_1^n, w_2^n) : (\hat{w}_1^n, \hat{w}_2^n, z^n) \in T_\epsilon^n} Pr\{\alpha \in T_\epsilon^n\}. \quad (8) \end{aligned}$$

The probability term inside the summation in (8) is

$$\begin{aligned} &\leq \sum_{(x_1(\cdot), x_2(\cdot), y^n) : \alpha \in T_\epsilon^n} Pr\{x_1(\hat{w}_1^n), x_2(\hat{w}_2^n), y^n | \hat{w}_1^n, \hat{w}_2^n, z^n\} \\ &\leq \sum \dots Pr\{x_1(\hat{w}_1^n) | \hat{w}_1^n\} Pr\{x_2(\hat{w}_2^n) | \hat{w}_2^n\} Pr\{y^n | z^n\} \\ &\leq \sum_{(x_1(\cdot), x_2(\cdot), y^n) : \alpha \in T_\epsilon^n} 2^{-n\{H(X_1|W_1) + H(X_2|W_2) + H(Y|Z) - 5\epsilon\}} \\ &\leq 2^{nH(X_1, X_2, Y|W_1, W_2, Z)} \\ &\quad 2^{-n\{H(X_1|W_1) + H(X_2|W_2) + H(Y|Z) - 7\epsilon\}}. \end{aligned}$$

But from hypothesis, we have

$$\begin{aligned} &H(X_1, X_2, Y|W_1, W_2, Z) - H(X_1|W_1) - H(X_2|W_2) \\ &- H(Y|Z) = H(Y|X_1, X_2) - H(Y|Z) \\ &= H(Y|X_1, X_2, Z) - H(Y|Z) \\ &= -I(X_1, X_2; Y|Z). \end{aligned}$$

Hence,

$$\begin{aligned} &Pr\{(\hat{w}_1^n, \hat{w}_2^n, x_1(\hat{w}_1^n), x_2(\hat{w}_2^n), y^n, z^n) \in T_\epsilon^n\} \\ &\leq 2^{-n\{I(X_1, X_2; Y|Z) - 7\epsilon\}}. \quad (9) \end{aligned}$$

Then from (8)

$$\begin{aligned} P(\mathbf{E4}) &\leq \sum_{\substack{(\hat{w}_1^n, \hat{w}_2^n) \neq (w_1^n, w_2^n) \\ (\hat{w}_1^n, \hat{w}_2^n, z^n) \in T_\epsilon^n}} 2^{-n\{I(X_1, X_2; Y|Z) - 7\epsilon\}} \\ &\leq |\{(\hat{w}_1^n, \hat{w}_2^n) : (\hat{w}_1^n, \hat{w}_2^n, z^n) \in T_\epsilon^n\}| 2^{-n\{I(X_1, X_2; Y|Z) - 7\epsilon\}} \\ &\leq |\{\hat{w}_1^n\}| \cdot |\{\hat{w}_2^n\}| \cdot Pr\{(\hat{w}_1^n, \hat{w}_2^n, z^n) \in T_\epsilon^n\} \\ &\quad 2^{-n\{I(X_1, X_2; Y|Z) - 7\epsilon\}} \\ &\leq 2^{n\{I(U_1, Z_1; W_1) + I(U_2, Z_2; W_2) + 2\epsilon\}} \\ &\quad 2^{-n\{I(W_1; W_2, Z) + I(W_2; W_1, Z) + I(W_1; W_2|Z) - 4\epsilon\}} \\ &\quad 2^{-n\{I(X_1, X_2; Y|Z) - 7\epsilon\}} \\ &= 2^{n\{I(U_1, U_2, Z_1, Z_2; W_1, W_2|Z)\}} 2^{-n\{I(X_1, X_2; Y|Z) - 13\epsilon\}}. \end{aligned}$$

The RHS of the above inequality tends to zero if $I(U_1, U_2, Z_1, Z_2; W_1, W_2|Z) < I(X_1, X_2; Y|Z)$.

Thus as $n \rightarrow \infty$, with probability tending to 1, the decoder finds the correct sequence (W_1^n, W_2^n) which is jointly strongly typical with (U_1^n, U_2^n, Z^n) . Since for large n , the empirical distribution of (W_1^n, W_2^n) is close to its joint distribution which by assumption has distortion (D_1, D_2) , our claim follows.

VII. CONCLUSIONS

We have provided sufficient conditions for transmission of several correlated sources over a multiple access channel. There may be side information available at the encoders and the decoder. Lossless and lossy transmission are considered. The channels could have discrete or continuous alphabets. Previous results are shown as special cases.

REFERENCES

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A survey on sensor networks. *IEEE Communications Magazine*, pages 1–13, august 2002.
- [2] S. J. Back, G. Veciana, and X. Su. Minimizing energy consumption in large-scale sensor networks through distributed data compression and hierarchical aggregation. *IEEE JSAC*, 22(6):1130–1140, Aug 2004.
- [3] T. Berger. *Multiterminal source coding in the information theory approach to communication*. Springer-Verlag, N.Y., 1977.
- [4] T. Berger, Z. Zhang, and H. Viswanathan. The CEO problem. *IEEE Trans. Inform. Theory*, IT-42:887–902, 1996.
- [5] T. M. Cover, A. E. Gamel, and M. Salehi. Multiple access channels with arbitrarily correlated sources. *IEEE Trans. Inform. Theory*, IT - 26:648–657, 1980.
- [6] T. M. Cover and J. A. Thomas. *Elements of Information theory*. Wiley Series in Telecommunication, N.Y., 2004.
- [7] G. Dueck. A note on the multiple access channel with correlated sources. *IEEE Trans. Inform. Theory*, IT -27(2):232–235, 1981.
- [8] M. Fleming and M. Effros. On rate distortion with mixed types of side information. *IEEE Trans. Inform. Theory*, 52(4):1698–1705, 2006.
- [9] M. Gastpar. Wyner-ziv problem with multiple sources. *IEEE Trans. Inform. Theory*, 50(11):2762–2768, 2004.
- [10] M. Gastpar and M. Vetterli. Power spatio-temporal bandwidth and distortion in large sensor networks. *IEEE JSAC*, 23(4):745–754, 2005.
- [11] D. Gunduz and E. Erkip. Transmission of correlated sources over multiuser channels with receiver side information. *UCSD ITA Workshop, San Diego, CA*, Jan 2007.
- [12] A. Lapidoth and S. Tinguely. Sending a bi-variate Gaussian source over a Gaussian MAC. *IEEE International Symp. on Inform. Theory 06*, 2006.
- [13] Y. Oohama. The rate distortion function for quadratic Gaussian CEO problem. *IEEE Trans. Inform. Theory*, IT -44:1057–1070, 1998.
- [14] R. Rajesh and V. Sharma. A joint source-channel coding scheme for transmission of correlated discrete sources over a Gaussian multiple access channel. *submitted*.
- [15] R. Rajesh and V. Sharma. Distributed joint source-channel coding on a multiple access channel with side information. *Technical Report, Available: Arxiv, http://arxiv.org/PS_cache/arxiv/pdf/0803/0803.1445v1.pdf*, 2008.
- [16] S. Ray, M. Medard, M. Effros, and R. Kotter. On separation for multiple access channels. *Proc. IEEE Inform. Theory Workshop*, 2006.
- [17] D. Slepian and J. K. Wolf. Noiseless coding of correlated information sources. *IEEE Trans. Inform. Theory*, IT-19:471–480, 1973.
- [18] V. K. Varsheneya and V. Sharma. Distributed coding for multiple access communication with side information. *Proc. IEEE Wireless Communication and Networking Conference (WCNC)*, April 2006.
- [19] V. K. Varsheneya and V. Sharma. Lossy distributed source coding with side information. *Proc National Conference on Communication (NCC)*, Jan 2006.
- [20] A. B. Wagner, S. Tavildar, and P. Viswanath. The rate region of the quadratic Gaussian two terminal source coding problem. *Arxiv*, 2005.
- [21] A. Wyner and J. Ziv. The rate distortion function for source coding with side information at the receiver. *IEEE Trans. Inform. Theory*, IT-22:1–11, 1976.