Distributed Coding for Multiple Access Communication with Side Information

Virendra K. Varshneya, Vinod Sharma
Department of Electrical Communication Engg
Indian Institute of Science, Bangalore, India
email: virendra@ece.iisc.ernet.in, vinod@ece.iisc.ernet.in

***Abstract*— In a typical sensor network scenario a goal is to monitor a spatio-temporal process through a number of inexpensive sensing nodes, the key parameter being the fidelity at which the process has to be estimated at distant locations. We study such a scenario in which multiple encoders transmit their correlated data at finite rates to a distant, common decoder over a discrete time multiple access channel under various side information assumptions. In particular, we derive an achievable rate region for this communication problem.**

Keywords: Sensor networks, multiple access channel, distributed joint source-channel coding.

I. Introduction

A commonly used model of sensor networks assumes multiple encoders transmitting at finite rates to a distant, common decoder over a multiple access channel. Since the observations of these sensors are inherently correlated, it is possible to achieve lower rates using astute techniques than would have been required if sensors compressed their respective observations disregarding others'. Moreover, the availability of side information at the encoders or/and at the decoder can also considerably reduce the required rate of transmission. This is true irrespective of the required fidelity at the decoder.

In this paper, we consider the situation depicted in Fig. 1. We consider discrete memoryless dependent finite alphabet sources U_1, U_2 and side information random variables Z_1, Z_2 and Z with known joint distribution $p(u_1, u_2, z_1, z_2, z)$, i.e, $(U_{1i}, U_{2i}, Z_{1i}, Z_{2i}, Z_i), i \geq 1$ form an i.i.d. sequence with distribution $p(u_1, u_2, z_1, z_2, z)$. Our interest lies in the rate region that allows the decoder to satisfy the distortion constraints

$$Ed_1(U_1, \hat{U}_1) \leq D_1, \quad Ed_2(U_2, \hat{U}_2) \leq D_2, \tag{1}$$

where $\hat{U}_i$ is the estimate of U_i produced by the receiver and d_i are arbitrary, bounded distortion measures. We will generalize these results to multiple sources also.

It is well known that source channel separation does not hold in this case ([1], [2]). The best known achievable bound on the rate region for the case of lossless transmission without any side information was obtained by Cover, Gamal and Salehi in [2]. However, they did not prove the converse and hence the capacity region (Since source-channel separation does not hold, it depends on the source statistics also.) for this system is still not available.

We study this problem from lossy perspective under various side information assumptions, as depicted in Fig. 1. The sensors transmit their codewords X_is to a single decoder through a multiple access channel. The receiver obtains Y, which is a superposition of transmitted symbols X_is and the receiver noise. The receiver also has access to side information Z_i available to the encoder i, for all i and some extra side information Z. It uses Y, Z_is and Z to estimate the sensor observations U_i as $\hat{U}_i$. We obtain an inner bound on the rate region (the rates at which the encoders have to transmit) of this problem for a given mean square distortion.

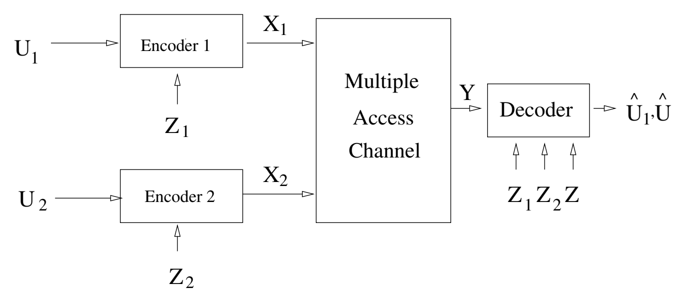

Fig. 1. Lossy communication over a Multiple Access Channel.

The above setup is useful in the following sensor network scenario. N sensor nodes scattered in a random field are sensing the field. The field can be divided in various subregions. The sensors in a subregion transmit their data to a clusterhead in the subregion (See, [3]). Many of the sensors in a subregion may be close to each other and hence their observations can be correlated. These correlated observations need to be transmitted to the clusterhead via the multiple access wireless channel. The correlations among the observations U_i's can be used by the encoders to compress U_i's more efficiently. The clusterhead decodes (estimates) the data from different sensors based on the superposition of symbols it receives and also the data Z it senses itself. The sensors can also communicate with each other (Z_1, Z_2) before sending the data to the clusterhead. Since it is a multiple access channel Z_1, and Z_2 will also be available available to the clusterhead.

The side information Z_1, and Z_2 can also be interpreted as the information transmitted by other sensor nodes in the cluster to the clusterhead which is relayed by the sensors 1, and 2 respectively.

Distributed source coding without channel has also received much attention in literature. For the case when observations are to be decoded losslessly with arbitrarily low error probability, Slepian and Wolf [4] proved the coding theorem for two sensors. Cover [5] extended their results to an arbitrary number of discrete sources with ergodic memory using an important technique now known as "random binning". Inspired by Slepian-Wolf results, Wyner and Ziv

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2006 proceedings.

[6] obtained the rate distortion region for a source when the decoder has access to side information. Their result requires the encoder to communicate at higher rate than it will if it too had access to the same side information. The latter result (when the encoder and the decoder have side information) was first obtained by Gray (See, [7]) and is generally known as conditional rate distortion theorem. The difference in the rates stems from the Markov Chain condition present in the Wyner-Ziv result. The most important contributions to the lossy Slepian-Wolf problem are those of Berger and Tung [7] in the form of an inner and an outer bound on the rate distortion region. Despite numerous attempts (e.g. [8], [9]), exact rate distortion region is still unknown. Recently, Gastpar [10] studied the distributed source coding problem when the decoder has access to side information. He derived an inner and an outer bound on the rate region and proved the tightness of his bounds for the case when the sources are conditionally independent given side information. Finally, in [11] the authors obtain inner and outer bounds on the rate distortion region with side information. The present paper extends the result in [11] by requiring the encoders to communicate over a multiple access channel. All the results mentioned above are special cases of our main result.

Other related work includes that of [12], [13], and [14]. In [12], the authors work with independent channels for the case when sources are to be estimated with arbitrarily small probability of errors. They obtain the joint souce/channel capacity region, and prove the separation theorem for their problem. We pause to mention that our work differs from the work in [12] in two ways. First, we consider a multiple access channel which may be more realistic in some scenario, and secondly we study the more general problem of lossy communication. In [13], the authors study the scaling laws for the transport capacity of a many-to-one data gathering wireless channel. The authors in [14] derive a lower bound on the best achievable end-to-end distortion for the gaussian sensor network, and shows that separate source and channel coding may incur an exponential penalty in terms of communication resources, as a function of the number of sources.

The rest of the paper is organized as follows. After presenting the formal statement of the problem and the main results in Section II for two nodes we prove the results in Section III. In Section IV, we study a few special case of the problem studied by us and relate them to the previous work. We generalize our main results to more than two sources in Section V. The paper is finally summarized and concluded in Section VI.

II. Problem Statement And Main Result

Consider discrete memoryless dependent finite alphabet sources U_1, U_2 and side information random variables Z_1, Z_2 and Z_3 with a known joint distribution $p(u_1, u_2, z_1, z_2, z_3)$. Encoder i after observing and processing the source U_i and side information Z_i communicates with the single decoder over a discrete memoryless multiple access channel $(\mathcal{X}_1 \times \mathcal{X}_2, \mathcal{Y}, p(y|x_1, x_2))$. Let U_i^n denote the sequence $(U_{i,1}, U_{i,2}, \ldots, U_{i,n})$, where $U_{i,j}$ is the j^{th} symbol emitted by source U_i.

Definition 2.1: The source (U_1^n, U_2^n) can be transmitted over the multiple access channel $(\mathcal{X}_1 \times \mathcal{X}_2, \mathcal{Y}, p(y|x_1, x_2))$ with distortions $\mathbf{D} \triangleq (D_1, D_2)$, if for any $\varepsilon > 0$, there is an n_0 such that for all $n > n_0$ there exist encoders $f_{E,i}^n : \mathcal{U}_i^n \times \mathcal{Z}_i^n \rightarrow \mathcal{X}_i^{(n)}$, $i \in \{1, 2\}$, and a decoder $f_D^n : \mathcal{Y}^n \times \mathcal{Z}_1^n \times \mathcal{Z}_2^n \times \mathcal{Z}^n \rightarrow (\hat{\mathcal{U}}_1^n, \hat{\mathcal{U}}_2^n)$, such that $\frac{1}{n} E[\sum_{j=1}^n d(U_{ij}, \hat{U}_{ij})] \leq D_i + \varepsilon$, $i \in \{1, 2\}$, where $(\hat{U}_1^n, \hat{U}_2^n) = f_D(Y^n, Z_1^n, Z_2^n, Z^n)$.

We now state the main theorem of this paper.

Theorem 2.2: A source $(U_1^n, U_2^n) \sim \prod_{i=1}^n p(u_{1i}, u_{2i})$ can be communicated in a distributed fashion over a discrete memoryless multiple access channel $(\mathcal{X}_1 \times \mathcal{X}_2, \mathcal{Y}, p(y|x_1, x_2))$ with distortion (D_1, D_2) if there exist auxiliary random variables (W_1, W_2) satisfying

1) $p(u_1, u_2, z_1, z_2, z, w_1, w_2) = p(u_1, u_2, z_1, z_2, z)$ $.p(w_1|u_1 z_1) p(w_2|u_2 z_2)$;
2) There exists a function $f'_D : \mathcal{W}_1 \times \mathcal{W}_2 \times \mathcal{Z}_1 \times \mathcal{Z}_2 \times \mathcal{Z} \rightarrow (\hat{U}_1, \hat{U}_2)$ such that $\frac{1}{n} E[\sum_{j=1}^n d(U_{ij}, \hat{U}_{ij})] \leq D_i + \varepsilon, i \in \{1, 2\}$

and such that the constraints

$$\begin{aligned} I(U_1; W_1|W_2 Z_1 Z_2 Z) &\leq I(X_1; Y|X_2 W_2 Z_1 Z_2 Z), \\ I(U_2; W_2|W_1 Z_1 Z_2 Z) &\leq I(X_2; Y|X_1 W_1 Z_1 Z_2 Z), \quad (2) \\ I(U_1 U_2; W_1 W_2|Z_1 Z_2 Z) &\leq I(X_1 X_2; Y|Z_1 Z_2 Z). \end{aligned}$$

are satisfied for some random variable X_1 and X_2 satisfying $p(u_1, u_2, z_1, z_2, z, w_1, w_2, x_1, x_2, y) = p(u_1, u_2, z_1, z_2, z) p(w_1|u_1 z_1) p(w_2|u_2 z_2) p(x_1|w_1) p(x_2|w_2)$ $.p(y|x_1 x_2)$.

III. Proof of the Main Result

A. Preliminaries

Let $T_\varepsilon^n(X)$ be the set of strongly ε-typical sequences of length n generated according to the distribution $p_X(x)$ ([1, p. 288,358], [7]). The following lemma, introduced by Berger ([7], [1]), is handy while proving many multiterminal source coding results. It ensures the transitivity of joint typicality. In the following for r.v.'s X, Y, and Z the notation $X - Y - Z$ denotes that {X, Y, Z} forms a Markov sequence.

Lemma 3.1 (Markov Lemma): Suppose $X - Y - Z$. If for a given $(x^n, y^n) \in T_\varepsilon^n(X, Y)$, Z^n is drawn $\sim \prod_{i=1}^n p(z_i|y_i)$, then with high probability $(x^n, y^n, Z^n) \in T_\varepsilon^n(X, Y, Z)$ for n sufficiently large.

The proofs of following Lemmas are simple and omitted for the sake of brevity. The following extension of the Markov lemma is required for proving our results.

Lemma 3.2 (extended Markov Lemma 1): Suppose $W_1 - X_1 Y_1 - X_2 W_2 Y_2 Y$ and $W_2 - X_2 Y_2 - X_1 W_1 Y_1 Y$. If for a given $(x_1^n, x_2^n, y_1^n, y_2^n, y^n) \in T_\varepsilon^n(X_1, X_2, Y_1, Y_2, Y)$, W_1^n and W_2^n are drawn respectively $\sim \prod_{i=1}^n p(w_{1i}|x_{1i} y_{1i})$ and $\prod_{i=1}^n p(w_{2i}|x_{2i} y_{2i})$, then with high probability $(x_1^n, x_2^n, y_1^n, y_2^n, y^n, W_1^n, W_2^n) \in T_\varepsilon^n(X_1 X_2 Y_1 Y_2 Y W_1 W_2)$ for n sufficiently large.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2006 proceedings.

The following lemma is a generalization of Lemma 3.2 to more than two encoders. This is helpful while proving the results for multiple sources.

Lemma 3.3 (extended Markov Lemma 2): Suppose $W_i - X_iY_i - X_{\{i\}^c}W_{\{i\}^c}Y_{\{i\}^c}Y$ for all $i \in \mathcal{E}$, where $\mathcal{E}$ is the set of encoders, and $X_{\{i\}^c} \triangleq (X_k, k \in \mathcal{E} \setminus \{i\})$. If for a given $(x_i^n, y_i^n, y^n, i \in \mathcal{E}) \in T_\varepsilon^n(X_i, Y_i, Y, i \in \mathcal{E})$, W_i^n is drawn $\sim \prod_{j=1}^n p(w_{ij}|x_{ij}y_{ij})$, then with high probability $(x_i^n, y_i^n, y^n, W_i^n, i \in \mathcal{E}) \in T_\varepsilon^n(X_iY_iYW_i, i \in \mathcal{E})$ for n sufficiently large.

We further need the following lemmas during the course of the proof. The proofs given below are established assuming weak typicality for the sake of brevity.

Lemma 3.4: If $(\hat{W}_1^n, W_2^n, Y_1^n, Y_2^n, Y^n) \sim p(w_2y_1y_2y)$ $.p(w_1)$, then $\Pr\Big((\hat{W}_1{}^n, W_2^n, Y_1^n, Y_2^n, Y^n) \in T_\varepsilon^n\Big) \leq 2^{-n\{I(W_1;W_2Y_1Y_2Y)-3\varepsilon\}}$.

Proof :

$$
\begin{aligned}
Pr(\hat{W}_1{}^n, &W_2^n, Y_1^n, Y_2^n, Y^n) \in T_\varepsilon^n) \\
&= \sum_{(w_1^n, w_2^n y_1^n y_2^n y^n) \in T_\varepsilon^n} p(w_2^n y_1^n y_2^n y^n) p(w_1^n) \\
&\leq 2^{n(H(W_1W_2Y_1Y_2Y)+\varepsilon)} 2^{-n(H(W_2Y_1Y_2Y)+H(W_1)-2\varepsilon)} \\
&= 2^{-n\{I(W_1;W_2Y_1Y_2Y)-3\varepsilon\}}
\end{aligned}
$$

■

Lemma 3.5: If $(\hat{W}_1{}^n, \hat{W}_2{}^n, Y_1^n, Y_2^n, Y^n) \sim p(y_1y_2y)$ $.\ p(w_1)p(w_2)$, then $\Pr\Big((\hat{W}_1{}^n, \hat{W}_2{}^n, Y_1^n, Y_2^nY^n) \in T_\varepsilon^n\Big) \leq 2^{-n\{I(W_1;W_2Y_1Y_2Y)+I(W_2;W_1Y_1Y_2Y)-I(W_1;W_2|Y_1Y_2Y)-4\varepsilon\}}$.

Proof :

$$
\begin{aligned}
&Pr(\hat{W}_1{}^n, \hat{W}_2{}^n, Y_1^n, Y_2^nY^n) \in T_\varepsilon^n) \\
&= \sum_{(w_1^n, w_2^n y_1^n y_2^n y^n) \in T_\varepsilon^n} p(y_1^n y_2^n y^n) p(w_1^n) p(w_2^n) \\
&\leq 2^{-n\{H(Y_1Y_2Y)+H(W_1)+H(W_2)-H(W_1W_2Y_1Y_2Y)-4\varepsilon\}} \\
&= 2^{-n\{I(W_1;W_2Y_1Y_2Y)+I(W_2;W_1Y_1Y_2Y)-I(W_1;W_2|Y_1Y_2Y)-4\varepsilon\}}
\end{aligned}
$$

■

B. Proof of main result

We now show the achievability of all points in the rate region (2).

Proof: Fix $p(w_1|u_1z_1)$, $p(w_2|u_2z_2)$, $p(x_1|w_1)$ and $p(x_2|w_2)$ as well as $f'_D(.)$ satisfying the distortion constraints.

Codebook Generation : Let $R'_i = I(U_iZ_i; W_i) + \delta, i \in \{1, 2\}$ for some $\delta > 0$. Generate $2^{nR'_1}$ codewords of length n, sampled iid from the marginal distribution $p(w_1)$. Similary, generate $2^{nR'_2}$ codewords of length n, sampled iid from the marginal distribution $p(w_2)$. For each w_1^n independently generate sequence x_1^n according to $\prod_{j=1}^n p(x_{1j}|w_{1j})$ and for each w_2^n generate sequence x_2^n according to $\prod_{j=1}^n p(x_{2j}|w_{2j})$. Call these sequences $x_1(w_1^n)$ and $x_2(w_2^n)$ respectively. Reveal the codebooks to the encoders and the decoder.

Encoding: For $i \in \{1, 2\}$, given source sequences U_i^n and Z_i^n, the i^{th} encoder looks for a codeword W_i^n such that $(U_i^n, Z_i^n, W_i^n) \in T_\varepsilon^n(U_i, Z_i, W_i)$ and then transmits $x_i(W_i^n)$.

Decoding: Upon receiving y^n, the decoder finds the unique (w_1^n, w_2^n) pair such that$(w_1^n, w_2^n, x_1(W_1^n), x_2(W_2^n), y^n, z_1^n, z_2^n, z^n) \in T_\varepsilon^n$. 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 this encoding-decoding scheme tends to zero as $n \to \infty$. The error can occur because of the following four events **E1** - **E4**. We show that $P(\mathbf{E}_i) \to 0$, for $i = 1, \ldots, 4$.

E1 The encoders do not find codewords. However, from rate distortion theory (See, e.g., [1], pp. 336) $\lim_{n\to\infty} P(\mathbf{E1}) = 0$ if $R'_1 > I(U_1Z_1; W_1)$ and $R'_2 > I(U_2Z_2; W_2)$.

E2 The codewords are not jointly typical with (z_1^n, z_2^n, z^n). Probability of this event goes to zero from Lemma 3.2.

E3 There exists another codeword $\hat{w}_1^n$ such that $(\hat{w}_1^n, W_2^n, x_1(\hat{w}_1^n), x_2(W_2^n), y^n, z_1^n, z_2^n, z^n) \in T_\varepsilon^n$. Define $\alpha \triangleq (\hat{w}_1^n, w_2^n, x_1(\hat{w}_1^n), x_2(w_2^n), y^n, z_1^n, z_2^n, z^n)$. Then,

$$
\begin{aligned}
P(\mathbf{E3}) &= Pr\{\text{ There is } \hat{w}_1^n \neq w_1^n : \alpha \in T_\varepsilon^n\} \\
&\leq \sum_{\substack{\hat{w}_1^n \neq w_1^n: \\ (\hat{w}_1^n, w_2^n, z_1^n, z_2^n, z^n) \in T_\varepsilon^n}} Pr\{\alpha \in T_\varepsilon^n\}. \quad (3)
\end{aligned}
$$

The probability term inside the summation in (3) is

$$
\begin{aligned}
&\leq \sum_{\substack{(x_1(.), x_2(.), \\ y^n): \alpha \in T_\varepsilon^n}} Pr\{x_1(\hat{w}_1^n), x_2(w_2^n), y^n | \hat{w}_1^n, w_2^n, z_1^n, z_2^n, z^n\} \\
&\leq \sum_{-} Pr\{x_1(\hat{w}_1^n)|\hat{w}_1^n\} Pr\{x_2^n(w_2^n), y^n | w_2^n, z_1^n, z_2^n, z^n\} \\
&\leq \sum_{\substack{(x_1(.), x_2(.), y^n): \\ \alpha \in T_\varepsilon^n}} 2^{-n\{H(X_1|W_1)+H(X_2Y|W_2Z_1Z_2Z)-4\varepsilon\}} \\
&\leq 2^{n\{H(X_1X_2Y|W_1W_2Z_1Z_2Z)+2\varepsilon\}}.2^{-n\{H(X_1|W_1)-2\varepsilon\}} \\
&\quad 2^{-n\{H(X_2Y|W_2Z_1Z_2Z)-2\varepsilon\}}.
\end{aligned}
$$

But from hypothesis, we have

$$H(X_1X_2Y|W_1W_2Z_1Z_2Z) - H(X_1|W_1) - H(X_2Y|W_2Z_1Z_2Z)$$

$$
\begin{aligned}
&= H(X_1X_2Y|W_1W_2) - H(X_1|W_1) \\
&\quad -H(X_2Y|W_2Z_1Z_2Z) \\
&= H(X_1|W_1) + H(X_2|W_2) + H(Y|X_1X_2) \\
&\quad -H(X_1|W_1) - H(X_2Y|W_2Z_1Z_2Z) \\
&= H(Y|X_1X_2) - H(Y|X_2W_2Z_1Z_2Z) \\
&= H(Y|X_1X_2W_2Z_1Z_2Z) - H(Y|X_2W_2Z_1Z_2Z) \\
&= -I(X_1; Y|X_2W_2Z_1Z_2Z).
\end{aligned}
$$

Hence,

$$Pr\{(\hat{w}_1^n, W_2^n, x_1(\hat{w}_1^n), x_2(W_2^n), y^n, z_1^n, z_2^n, z^n) \in T_\varepsilon^n\} \leq 2^{-n\{I(X_1;Y|X_2W_2Z_1Z_2Z)-6\varepsilon\}}.$$

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2006 proceedings.

Then from (3)

$$\begin{aligned}P(\mathbf{E3}) &\leq \sum_{\substack{\hat{w}_1^n \neq w_1^n : (\hat{w}_1^n, w_2^n, \\ z_1^n, z_2^n, z^n) \in T_\varepsilon^n}} 2^{-n\{I(X_1;Y|X_2W_2Z_1Z_2Z)-\varepsilon\}} \\ &\leq |\{w_1^n : (w_1^n, w_2^n, z_1^n, z_2^n, z^n) \in T_\varepsilon^n\}| \\ &\quad .2^{-n\{I(X_1;Y|X_2W_2Z_1Z_2Z)-6\varepsilon\}} \\ &\leq |\{w_1^n\}|.Pr\{(\hat{w}_1^n, w_2^n, z_1^n, z_2^n, z^n) \in T_\varepsilon^n\} \\ &\quad .2^{-n\{I(X_1;Y|X_2W_2Z_1Z_2Z)-6\varepsilon\}} \\ &\leq 2^{n\{I(U_1Z_1;W_1)+\varepsilon\}}.2^{-n\{I(W_1;W_2Z_1Z_2Z)-3\varepsilon\}} \\ &\quad .2^{-n\{I(X_1;Y|X_2W_2Z_1Z_2Z)-6\varepsilon\}} \\ &= 2^{n\{I(U_1;W_1|W_2Z_1Z_2Z)\}} \\ &\quad .2^{-n\{I(X_1;Y|X_2W_2Z_1Z_2Z)-10\varepsilon\}}. \end{aligned} \tag{4}$$

The R.H.S. in the above inequality tends to zero if $I(U_1;W_1|W_2Z_1Z_2Z) < I(X_1;Y|X_2W_2Z_1Z_2Z)$. In (4) we have used Lemma 3.4 and the fact that $I(U_1Z_1;W_1) - I(W_1;W_2Z_1Z_2Z)$

$$\begin{aligned} &= H(W_1|W_2Z_1Z_2Z) - H(W_1|U_1Z_1) \\ &= H(W_1|W_2Z_1Z_2Z) - H(W_1|U_1Z_1Z_2ZW_2) \\ &= I(U_1;W_1|W_2Z_1Z_2Z). \end{aligned}$$

Similarly, by symmetry of the problem we require

$$I(U_2;W_2|W_1Z_1Z_2Z) < I(X_2;Y|X_1W_1Z_1Z_2Z).$$

E4 There exist other codewords $\hat{w}_1^n$ and $\hat{w}_2^n$ such that $\alpha \triangleq (\hat{w}_1^n, \hat{w}_2^n, x_1(\hat{w}_1^n), x_2(\hat{w}_2^n), y^n, z_1^n, z_2^n, z^n) \in T_\varepsilon^n$. Then,

$$\begin{aligned}P(\mathbf{E4}) &= Pr\{ \text{ There is } (\hat{w}_1^n, \hat{w}_1^n) \neq (w_1^n, w_2^n) : \alpha \in T_\varepsilon^n\} \\ &\leq \sum_{\substack{\hat{w}_1^n, \hat{w}_1^n) \neq (w_1^n, w_2^n): \\ (\hat{w}_1^n, \hat{w}_2^n, z_1^n, z_2^n, z^n) \in T_\varepsilon^n}} Pr\{\alpha \in T_\varepsilon^n\}.\end{aligned}$$

The probability term inside the summation is

$$\begin{aligned} &\leq \sum_{\substack{(x_1(.), x_2(.), \\ y^n): \alpha \in T_\varepsilon^n}} Pr\{x_1(.), x_2(.), y^n | \hat{w}_1^n, \hat{w}_2^n, z_1^n, z_2^n, z^n\} \\ &\leq \sum_{-} Pr\{x_1(.)|\hat{w}_1^n\} Pr\{x_2^n(.)|\hat{w}_2^n\} Pr\{y^n|z_1^n, z_2^n, z^n\} \\ &\leq \sum_{\substack{(x_1(.), x_2(.), \\ y^n): \alpha \in T_\varepsilon^n}} 2^{-n\{H(X_1|W_1)+H(X_2|W_2)+H(Y|Z_1Z_2Z)-5\varepsilon\}} \\ &\leq 2^{n\{H(X_1X_2Y|W_1W_2Z_1Z_2Z)+2\varepsilon\}} \\ &\quad .2^{-n\{H(X_1|W_1)+H(X_2|W_2)+H(Y|Z_1Z_2Z)-5\varepsilon\}}. \end{aligned}$$

But from hypothesis, we have :

$H(X_1X_2Y|W_1W_2Z_1Z_2Z) - H(X_1|W_1) - H(X_2|W_2) - H(Y|Z_1Z_2Z)$

$$\begin{aligned} &= H(X_1X_2Y|W_1W_2) - H(X_1|W_1) - H(X_2|W_2) - \\ &\quad H(Y|Z_1Z_2Z) \\ &= H(X_1|W_1) + H(X_2|W_2) + H(Y|X_1X_2) - \\ &\quad H(X_1|W_1) - H(X_2|W_2) - H(Y|Z_1Z_2Z) \\ &= H(Y|X_1X_2Z_1Z_2Z) - H(Y|Z_1Z_2Z) \\ &= -I(Y;X_1X_2|Z_1Z_2Z). \end{aligned}$$

Hence,

$$\begin{aligned} &Pr\{(\hat{w}_1^n, \hat{w}_2^n, x_1(\hat{w}_1^n), x_2(\hat{w}_2^n), y^n, z_1^n, z_2^n, z^n) \in T_\varepsilon^n\} \\ &\quad \leq 2^{-n\{I(Y;X_1X_2|Z_1Z_2Z)-7\varepsilon\}}, \end{aligned}$$

and hence

$$\begin{aligned}P(\mathbf{E4}) &\leq \sum_{\substack{(\hat{w}_1^n, \hat{w}_2^n) \neq (w_1^n, w_2^n): \\ (\hat{w}_1^n, \hat{w}_2^n, z_1^n, z_2^n, z^n) \in T_\varepsilon^n}} 2^{-n\{I(Y;X_1X_2|Z_1Z_2Z)-7\varepsilon\}} \\ &\leq |\{(w_1^n, w_2^n) : (w_1^n, w_2^n, z_1^n, z_2^n, z^n) \in T_\varepsilon^n\}| \\ &\quad .2^{-n\{I(X_1X_2;Y|Z_1Z_2Z)-7\varepsilon\}} \\ &\leq |\{w_1^n\}|.|\{w_2^n\}|.Pr\{(\hat{w}_1^n, \hat{w}_2^n, z_1^n, z_2^n, z^n) \in T_\varepsilon^n\} \\ &\quad .2^{-n\{I(X_1X_2;Y|Z_1Z_2Z)-7\varepsilon\}} \\ &\leq 2^{n\{I(U_1Z_1;W_1)+I(U_2Z_2;W_2)+2\varepsilon\}} \\ &\quad .2^{-n\{I(W_1;W_2Z_1Z_2Z)+I(W_2;W_1Z_1Z_2Z)\}} \\ &\quad .2^{-n\{I(W_1;W_2|Z_1Z_2Z)-4\varepsilon\}} \\ &\quad .2^{-n\{I(X_1X_2;Y|Z_1Z_2Z)-7\varepsilon\}} \\ &= 2^{n\{I(U_1U_2;W_1W_2|Z_1Z_2Z)-I(X_1X_2;Y|Z_1Z_2Z)+13\varepsilon\}}.\end{aligned}$$

The R.H.S. tends to 0 if $I(U_1U_2;W_1W_2|Z_1Z_2Z) \leq I(X_1X_2;Y|Z_1Z_2Z)$, where we have used Lemma 3.5 and the independence conditions from the hypothesis.

Thus as $n \to \infty$, with probability tending to 1, the decoder finds the correct sequence (W_1^n, W_2^n) which is jointly strongly typical with (Z_1^n, Z_2^n, Z^n). Since for large n, the empirical distribution of $(W_1^n, W_2^n, Z_1^n, Z_2^n, Z^n)$ is close to its joint distribution which by assumption has distortion (D_1, D_2), our claim follows. ■

IV. SPECIAL CASES

In this section, we study a few special cases of the problem studied, and connect them with the previous work.

1) **Lossy Distributed Source Coding with Side Information ([11])**: Take the multiple access channel to be a dummy channel, i.e., the channel which reproduces it's inputs. In this case we obtain that the sources can be coded with rates R_1 and R_2 to obtain the specified distortions if they satisfy

$$\begin{aligned} I(U_1;W_1|W_2Z_1Z_2Z) &\leq R_1, \\ I(U_2;W_2|W_1Z_1Z_2Z) &\leq R_2, \\ I(U_1U_2;W_1W_2|Z_1Z_2Z) &\leq R_1 + R_2. \end{aligned} \tag{5}$$

2) **Lossy Multiple Access Communication**: Take $(Z_1, Z_2, Z) \perp (U_1, U_2)$. In this case the constraints (2) reduce to

$$\begin{aligned} I(U_1;W_1|W_2) &\leq I(X_1;Y|X_2W_2), \\ I(U_2;W_2|W_1) &\leq I(X_2;Y|X_1W_1), \\ I(U_1U_2;W_1W_2) &\leq I(X_1X_2;Y). \end{aligned} \tag{6}$$

which is an immediate generalization of [2] to the lossy case.

3) **Multiple Access Communication with Correlated Sources (lossless case)**:By taking $(Z_1Z_2Z) \perp (U_1U_2)$,

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2006 proceedings.

$W_1 = U_1$, and $W_2 = U_2$, the constraints (2) reduce to their constraints,

$$\begin{aligned} H(U_1|U_2) &\le I(X_1;Y|X_2U_2), \\ H(U_2|U_1) &\le I(X_2;Y|X_1U_1), \\ H(U_1U_2) &\le I(X_1X_2;Y). \end{aligned} \tag{7}$$

which is the result obtained in [2].

V. Generalization to Several Sources

Let $\mathcal{S} \triangleq \{1, 2, \ldots, n\}$ be the index set of sensors, and $(U_k, Z_k, k \in \mathcal{S}, Z)$ be a discrete memoryless source with joint probability distribution $p(u_k, z_k, k \in \mathcal{S}, z)$ on $\mathcal{Z} \times \prod_{i \in \mathcal{S}}(\mathcal{U}_i \times \mathcal{Z}_i)$ which is observed by $|\mathcal{S}|$ (= the cardinality of $\mathcal{S}$) sensors in a distributed fashion. Sensor k observes (U_k, Z_k) and communicates it's observations over a multiple access channel $(\prod_{i \in \mathcal{S}} \mathcal{X}_i, \mathcal{Y}, p(y|x_i, i \in \mathcal{S}))$ to a decoder which has access to $(Z, Z_k, k \in \mathcal{S})$.

Achievability for n sources can be defined similar to Definition 2.1.

Theorem 5.1: A source $(U_i^n, i \in \mathcal{S}) \sim \prod_{j=1}^n p(u_{ij}, i \in \mathcal{S})$ can be communicated in a distributed fashion over a discrete memoryless multiple access channel $(\prod_{i \in \mathcal{S}} \mathcal{X}_i, \mathcal{Y}, p(y|x_i, i \in \mathcal{S}))$ with distortion $(D_i, i \in \mathcal{S})$ if there exist auxiliary random variables $(W_i, i \in \mathcal{S})$ satisfying

1) $p(u_i, z_i, z, w_i, i \in \mathcal{S}) = p(u_i, z_i, z, i \in \mathcal{S}) \cdot \prod_{i \in \mathcal{S}} p(w_i|u_i z_i)$,
2) There exists a function $f'_D : \prod_{i \in \mathcal{S}}(\mathcal{W}_i \times \mathcal{Z}_i) \times \mathcal{Z} \to (\hat{U}_i, i \in \mathcal{S})$ such that $\frac{1}{n} E[\sum_{j=1}^n d(U_{ij}, \hat{U}_{ij})] \le D_i + \varepsilon$,

and such that the constraints

$$I(U_{\mathcal{A}}; W_{\mathcal{A}}|W_{\mathcal{A}^c} Z_{\mathcal{S}} Z) \le I(X_{\mathcal{A}}; Y|X_{\mathcal{A}^c} W_{\mathcal{A}^c} Z_{\mathcal{S}} Z), \quad \text{for all } \mathcal{A} \subset \mathcal{S}, \tag{8}$$

are satisfied for some $p(u_i, z_i, z, w_i, x_i, y, i \in \mathcal{S}) = p(u_i, z_i, z, i \in \mathcal{S}) \prod_{j \in \mathcal{S}}(p(w_j|u_j z_j) p(x_j|w_j)) p(y|x_i, i \in \mathcal{S})$.

■

The proof of this theorem is similar to that of Theorem 2.2, so we omit it for the sake of brevity.

VI. Conclusions

In this paper, we investigated the distributed lossy compression with various assumptions on the side information at the encoders and the decoder in the scenario of multiple access communication. We derived an inner bound on the rate region for this communication problem. The results are of particular interest to distributed compression and signal processing in low power sensor networks. Our results generalize the results in [2], [4], [7] and [11].

VII. References

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley, N.Y., 1991.

[2] T. M. Cover, A. E. Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inform. Theory, IT-26:648-657*, 1980.

[3] S. J. Baek, G. Veciana, and X. Su, "Minimizing energy consumption in large-scale sensor networks through distributed data compression and hierarchical aggregation," *IEEE JSAC, Vol 22, No 6*, Aug 2004, pp. 1130-1140.

[4] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory, IT-19:471-480*, 1973.

[5] T. M. Cover, "A proof of the data compression theorem of slepian and wolf for ergodic sources," *IEEE Trans. Inform. Theory, IT-22:226-228*, 1975.

[6] 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.

[7] T. Berger, *Multiterminal Source Coding in The Information Theory Approach to Communications*, G. Longo, Ed. Springer-Verlag, N.Y., 1977.

[8] T. Berger and R. W. Yeung, "Multiterminal source coding with one distortion criterion," *IEEE Trans. Inform. Theory, IT-35:228-236*, 1989.

[9] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Trans. Inform. Theory, IT-43:1912-1923*, 1997.

[10] M. Gastpar, "The wyner-ziv problem with multiple sources," *IEEE Trans. Inform. Theory, IT-50:2762-2768, Nov. 2004*.

[11] V. K. Varshneya and V. Sharma, "Lossy distributed source coding with side information." *Submitted.*

[12] J. Barros and S. D. Servetto, "On the capacity of the reachback channel in wireless sensor networks," *Proc. IEEE Workshop on Multimedia Signal Processing*, Dec. 2002.

[13] H. E. Gamal, "On the scaling laws of dense wireless sensor networks: the data gathering channel," *IEEE Tran. Inform. Theory, IT-51:1229-1234*, March 2005.

[14] M. Gastpar and M. Vetterli, "Power, spatio-temporal bandwidth and distortion in large sensor networks," *IEEE JSAC*, April 2005, pp. 745-754.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2006 proceedings.