Lossy Source Coding of Multiple Gaussian Sources: *m-helper problem*

Ameesh Pandya, Aman Kansal, Greg Pottie and Mani Srivastava
56-147G E-IV, Electrical Engineering Department
University of California, Los Angeles
Box 951594, Los Angeles, CA 90095-1594
Phone: (310) 287-0969, Fax: (310) 206-8495
{ameesh,kansal,pottie,mbs}@ee.ucla.edu

Abstract — **We consider the network information theoretic problem of finding the rate distortion bound when multiple correlated Gaussian sources are present. One of these is the source of interest but some side information from other sources is also transmitted to help reduce the distortion in the reproduction of the first source. The other sources are treated as helpers and are also coded. Special cases of this problem have been solved before, such as when the reproduction is lossless, when the sources are conditionally independent given one of them, or when the number of helpers is limited to one. We consider a generalized version and show that the previously derived expressions fall out as special cases of our bound. Our results can be directly utilized by designers to choose not only how many of the available sources should actually be communicated but also which sources have the highest potential to reduce the distortion.**

I. Introduction

Distributed detection of phenomena is an important problem in sensor networks [16, 8, 15]. It is known that higher reliability and lower probability of detection error can be achieved using multiple observations from a distributed set of sensors and intelligent fusion algorithms [20]. We consider the source coding problem for such a multiterminal scenario.

The multiterminal coding theory problem for two correlated memory-less sources with separate encoders was first addressed by Slepian and Wolf [18]. A related problem of source coding with side information when only one of the sources is reproduced was considered in (Section 14.8, [6]). However, both the above problems considered lossless reproduction. Han and Kobayashi [9], and Csiszar and Korner [7] have also focused on special extensions of Slepian and Wolf. We consider the related problem when multiple correlated sources are available and only one of them is reproduced but instead of lossless coding, the rate-distortion version is considered.

Wyner and Ziv [22, 21] had solved the rate distortion coding problem with uncoded side information, summarized in (Section 14.9, [6]). Related problems have also been considered by Berger [3], Kaspi and Berger [11], Berger et. al. [4], Tung [19] and Berger and Yeung [5]. Berger and Tung [3, 19] had considered the problem when all sources are reproduced instead of one. Oohama [13] solved an extension of the Wyner-Ziv problem when the side information is also coded, for the case of two sources. The extension to more than two sources was considered in [14] when the sources are conditionally independent given one of them. We consider the general problem when sources are correlated and the conditional independence does not hold.

Such a problem is of practical importance in coding when multiple sensors are measuring correlated data. For instance, a network of multiple sensors may be installed to monitor a physical environment, consisting of some resource constrained wireless sensors and some wired sensors. The wireless sensors may be deeply embedded into the environment and hence close to the phenomenon of interest while the wired sensors are farther off. Networked Infomechanical System (NIMS) [10] is an example of such a system. In this case, the wired sensors can provide side information to reduce the rate required at the embedded sensor of interest, and hence save transmission energy at this sensor, for a given distortion. We refer to the sensors providing side information as helpers while the sensor whose measurements are to be reproduced is referred to as the main source. The rate required at the main source is of concern since this source is wireless and resource constrained.

The objective is to find the rate distortion relation between the rate of the main source and the distortion in reproduction, for any arbitrary positive set of rates available for the helpers. The helper rates are assumed to be free (unconstrained) and may be very large in some situations, such as when the helper information is transmitted over high bandwidth wired channels. The exact problem, referred to as the m-helper problem, is specified in the next section and solved in section III. Section IV discusses the significance of helper rates and section V concludes.

II. Problem Statement

Let $X, Y_1, \dots, Y_m$ be correlated random variables such that $\{X_t, Y_{1t}, \dots, Y_{mt}\}_{t=1}^{\infty}$ are jointly normal, stationary and memoryless sources. For each observation time $t = 1, 2, 3, \dots$, the random $(m+1)$-tuplet $(X_t, Y_{1t}, \dots, Y_{mt})$ takes a value in the $(m+1)$-dimensional real space $\mathcal{X} \times \mathcal{Y}_1 \times \dots \times \mathcal{Y}_m$. The probability density function $p_{X,Y_1,\dots,Y_m}(x, y_1, \dots, y_m)$ is $\mathcal{N}\ (0, \Lambda)$ where, the covariance matrix, Λ, is given by

$$\begin{pmatrix} \sigma_X^2 & \rho_{XY_1}\sigma_X\sigma_{Y_1} & \dots & \rho_{XY_m}\sigma_X\sigma_{Y_m} \\ \rho_{XY_1}\sigma_X\sigma_{Y_1} & \sigma_{Y_1}^2 & \dots & \rho_{Y_1Y_m}\sigma_{Y_1}\sigma_{Y_m} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{XY_m}\sigma_X\sigma_{Y_m} & \rho_{Y_1Y_m}\sigma_{Y_1}\sigma_{Y_m} & \dots & \sigma_{Y_m}^2 \end{pmatrix}$$

with $-1 < \rho_{ij} < 1$, $(i,j) \in (X, Y_1, \dots, Y_m)$. Let n independent instances of $\{X_t\}_{t=1}^{\infty}$ be $X^n = \{X_1, X_2, \dots, X_n\}$ and similarly $Y_i^n = \{Y_{i1}, Y_{i2}, \dots, Y_{in}\}$, for $i = 1, 2, \dots, m$. Consider the system depicted in Figure (1). Data sequences X^n and Y_i^n are separately encoded to $\varphi_0(X^n)$ and $\{\varphi_i(Y_i^n)\}_{i=1}^{m}$. The encoder functions φ_0 and φ_i's are defined by

$$\begin{aligned} \varphi_0 &: \mathcal{X}^n \to \mathcal{C}_0 = \{1, 2, \dots, C_0\} \\ \varphi_i &: \mathcal{Y}_i^n \to \mathcal{C}_i = \{1, 2, \dots, C_i\} \end{aligned}$$

The coded (compressed) sequences are sent to a fusion center, and the rates are

$$\frac{1}{n}\log C_i \le R_i + \delta, i = 0, 1, 2, \dots, m$$

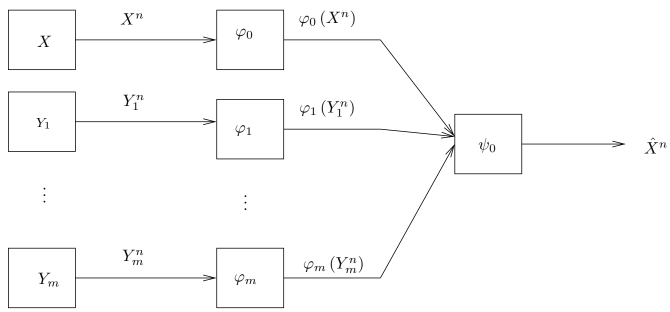

Figure 1: m-helper coding system

where δ is an arbitrary positive number. Note that, all logarithms in this paper are to the base 2. The decoder function observes the $(m+1)$-tuplet $(\varphi_0(X^n), \varphi_1(Y_1^n), \ldots, \varphi_m(Y_m^n))$ to estimate the main source as $\hat{X}^n$. The decoder function ψ_0 is given by

$$\psi_0 : \mathcal{C}_0 \times \mathcal{C}_1 \times \ldots \times \mathcal{C}_m \to \mathcal{X}^n$$

Note that the goal is to reproduce only X, the main source. The other sources, $\{Y_i\}_{i=1}^m$, are used as helpers, and are not reproduced. Hence, there is no distortion constraint on the helpers. Any available rate can be used for coding the helper information. Let

$$d_0 : \mathcal{X}^2 \to [0, \infty)$$

be the squared distortion measure. The average distortion, Δ_0, for $\hat{X}^n = \psi_0(\varphi_0(X^n), \varphi_1(Y_1^n), \ldots, \varphi_m(Y_m^n))$ is defined by,

$$\Delta_0 = E\frac{1}{n}\sum_{t=1}^{n} d_0\left(X_t, \hat{X}_t\right) \le D_0$$

The m-helper problem is to find the rate-distortion relation between R_0 and D_0 for the above coding system.

An attempt to derive the rate-distortion region for the general Gaussian case was made in [2, 1], but that derivation was not entirely correct. Specifically, consider equation (3.5) in [1]:

$$\begin{aligned} n(R_x + \delta) &\ge \log C_1 \\ &\ge H(W_x) \\ &\ge I\left(X^n; \hat{X}^n\right) - \frac{1}{n}\sum_{k=1}^{n} I(X^n; W_k) \end{aligned} \quad (1)$$

where $W_x = \varphi_x(X^n)$ and $W_i = \varphi_i(Y_i^n)$. The second term in (1) i. e. $\frac{1}{n}\sum_{k=1}^{n} I(X^n; W_k)$ is the erroneous term. This term does not account for the correlation between W_i's.

III. Solving the m-helper problem

With the constraints and definitions as described in Section II, we state the following theorem:

Theorem III.1 *For the m-helper coding system, data streams from correlated Gaussian sources can be fused to reduce the data rate, R_0, required for source X. R_0 satisfies the lower bound:*

$$R_0(D_0) \ge \frac{1}{2}\log^+\left[\frac{\sigma_X^2}{D_0}\prod_{i=1}^{m}(1-\rho_i^2)\Gamma_i\right] \quad (2)$$

where $\Gamma_i = 1 - \rho^2_{XY_i|(Y_1,\ldots,Y_{i-1})} + \rho^2_{XY_i|(Y_1,\ldots,Y_{i-1})} \cdot 2^{-2R_i}$,

$\rho_i^2 = 1 - \frac{\sigma^2_{X|Y_1,\ldots,Y_{i-1}}}{\sigma_X^2}$, *and* $\log^+ x = \max\{\log x, 0\}$.

Proof. To simplify the presentation we derive the rate-distortion region for the 2-helper case and then generalize it for the m-helper system.

Set $W_0 = \varphi_0(X^n)$, $W_1 = \varphi_1(Y_1^n)$ and $W_2 = \varphi_2(Y_2^n)$. Then

$$\begin{aligned} n(R_0+\delta) &\ge \log C_0 \\ &\ge h(W_0) \\ &\overset{(a)}{\ge} h(W_0|W_1, W_2) \\ &\overset{(b)}{=} I(X^n; W_0|W_1, W_2) \\ &\overset{(c)}{=} I(X^n; W_0, W_1, W_2) - I(X^n; W_1, W_2) \\ &\ge I(X^n; \hat{X}^n) - I(X^n; W_1) - I(X^n; W_2|W_1) \end{aligned} \quad (3)$$

Here, (a) holds because conditioning reduces entropy, (b) is obtained using the fact that W_0 is a function of X^n, and (c) follows from chain rule of mutual information. Now we express the rate of the second helper, accounting for the correlation among the helpers:

$$\begin{aligned} n(R_2+\delta) &\ge \log C_2 \\ &\ge h(W_2) \\ &\ge h(W_2|W_1, Y_1^n) \\ &= I(Y_2^n; W_2|W_1, Y_1^n) \end{aligned} \quad (4)$$

Observe that for $i = 1, 2$

$$\begin{array}{ccccc} W_i & \to & Y_i^n & \to & X^n \\ (W_1, W_2) & \to & (Y_1^n, Y_2^n) & \to & X^n \end{array} \quad (5)$$

are Markov chains. We use (3) to derive a lower bound on R_0. For this, let

$$\begin{aligned} F_n(D_0) &= \inf_{\hat{X}^n : \Delta_0 \le D_0} \frac{1}{n} I(X^n; \hat{X}^n) \\ G_{n1}(R_1) &= \sup_{\substack{W_1 : \frac{1}{n}I(Y_1^n; W_1) \le R_1 \\ W_1 \to Y_1^n \to X^n}} \frac{1}{n} I(X^n; W_1) \\ G_{n2}(R_2) &= \sup_{\substack{W_2 : \frac{1}{n}I(Y_2^n; W_2) \le R_2 \\ (W_1, W_2) \to (Y_1^n, Y_2^n) \to X^n}} \frac{1}{n} I(X^n; W_2|W_1) \end{aligned} \quad (6)$$

Therefore,

$$R_0 + \delta \ge F_n(D_0+\delta) - G_{n1}(R_1+\delta) - G_{n2}(R_2+\delta) \quad (7)$$

A lower bound on $F_n(D_0)$ can be derived as in [13]:

$$F_n(D_0) \ge \frac{1}{2}\log\frac{\sigma_X^2}{D_0} \quad (8)$$

An upper bound on $G_{n1}(R_1)$ is as derived in [13]:

$$G_{n1}(R_1) \le \frac{1}{2}\log\left(\frac{1}{1-\rho^2_{XY_1}+\rho^2_{XY_1}\cdot 2^{-2R_1}}\right) \quad (9)$$

Now we consider $G_{n2}(R_2)$. This is different from the calculation of $G_{n1}(R_1)$ as the correlation between Y_1 and Y_2 also needs to be considered. To evaluate $G_{n2}(R_2)$, define the random variables $X(y_1) = X|(Y_1 = y_1)$ and $Y_2(y_1) = Y_2|(Y_1 = y_1)$. Now, $E\left[X(y_1)|Y_2(y_1)\right] = aY_2(y_1)$, where $a = \rho_{XY_2|Y_1}\frac{\sigma_{X|Y_1}}{\sigma_{Y_2|Y_1}}$. Hence, we can write

$$X(y_1) = aY_2(y_1) + N$$

where where, N is a zero-mean Gaussian random variable with variance $\sigma_N^2 = \sigma_{X|Y_1}^2\left(1 - \rho_{XY_2|Y_1}^2\right)$ and is independent of $Y_2(y_1)$.

Since the sequences are memoryless, this leads to

$$X^n(y_1^n, w) = aY_2^n(y_1^n, w) + N^n \tag{10}$$

where $w = (w_1, w_2)$, $X^n(y_1^n, w)$ is the conditional random variable X^n conditioned on $Y_1^n = y_1^n$ and $W = w$, and $Y_2^n(y_1^n, w)$ is similarly defined.

Using the entropy power inequality [6] in (10),

$$\begin{aligned} 2^{\frac{2}{n}h(X^n(y_1^n,w))} &\geq 2^{\frac{2}{n}h(aY_2^n(y_1^n,w))} + 2^{\frac{2}{n}h(N^n)} \\ &= 2^{2h(N)} + a^2 2^{\frac{2}{n}h(Y_2^n(y_1^n,w))} \end{aligned} \tag{11}$$

The entropy of N can be substituted in the above expression. This entropy is given by:

$$h(N) = \frac{1}{2}\log\left\{2\pi e\left(\sigma_{X|Y_1}^2(1 - \rho_{XY_2|Y_1}^2)\right)\right\}$$

where $\sigma_{X|Y_1}^2 = \sigma_X^2(1 - \rho_{XY_1}^2)$. With this, (11) becomes

$$\begin{aligned} 2^{\frac{2}{n}h(X^n(y_1^n,w))} \geq\; & a^2 2^{\frac{2}{n}h(Y_2^n(y_1^n,w))} \\ & + 2\pi e\left[\sigma_X^2(1 - \rho_{XY_1}^2)(1 - \rho_{XY_2|Y_1}^2)\right] \end{aligned}$$

Taking the logarithm of the above equation, we get:

$$\frac{1}{n}h(X^n(w, y_1^n)) \geq T\left(\frac{1}{n}h(Y_2^n(y_1^n, w))\right) \tag{12}$$

where:

$$T(x) = \frac{1}{2}\log\left[a^2 2^{2x} + 2\pi e\sigma_X^2\left\{(1 - \rho_{XY_1}^2)(1 - \rho_{XY_2|Y_1}^2)\right\}\right] \tag{13}$$

Next, we take expectations on both sides of (12) with respect to $W = (W_1, W_2)$ and Y_1^n. Note that from the definition of our conditional random variables, it follows that

$$\begin{aligned} E_{W,Y_1^n}[h(X^n(w, y_1^n))] &= h(X^n|W, Y_1^n) \\ E_{W,Y_1^n}[h(Y_2^n(w, y_1^n))] &= h(Y_2^n|W, Y_1^n) \end{aligned}$$

where $E_Z[\cdot]$ denotes expectation w.r.t. Z. Observe that $T(x)$ is a convex function of x. Applying Jensen's inequality, we get

$$\frac{1}{n}h\left(X^n|Y_1^n, W\right) \geq T\left(\frac{1}{n}h(Y_2^n|Y_1^n, W)\right) \tag{14}$$

Since $T(x)$ is monotone increasing with respect to x, the inequality is preserved. From the definition of mutual information, (14) can be rewritten as,

$$\begin{aligned} &\frac{1}{n}h\left(X^n|Y_1^n, W_1, W_2\right) \\ &\quad\geq T\left(\frac{1}{n}h(Y_2^n|Y_1^n) - \frac{1}{n}I(Y_2^n; W_1, W_2|Y_1^n)\right) \end{aligned} \tag{15}$$

By chain rule of mutual information,

$$I(Y_2^n; W_1, W_2|Y_1^n) = I(Y_2^n; W_1|Y_1^n) + I(Y_2^n; W_2|W_1, Y_1^n)$$

Also, from the definition of W_1 it follows that $I(Y_2^n; W_1|Y_1^n) = 0$. Using this in (15), we get

$$\begin{aligned} &\frac{1}{n}h\left(X^n|W_1, W_2, Y_1^n\right) \\ &\quad\geq T\left[\frac{1}{n}h(Y_2^n|Y_1^n) - \frac{1}{n}I(Y_2^n; W_2|W_1, Y_1^n)\right] \end{aligned}$$

Now, using (4) in the above equation:

$$\frac{1}{n}h\left(X^n|W_1, W_2, Y_1^n\right) \geq T\left(\frac{1}{n}h(Y_2^n|Y_1^n) - R_2\right) \tag{16}$$

This can be used to derive $G_{n2}(R_2)$, defined in (6), as follows:

$$\begin{aligned} \frac{1}{n}I(X^n; W_2|W_1) &= \frac{1}{n}h(X^n|W_1) - \frac{1}{n}h(X^n|W_1, W_2) \\ &\overset{(a)}{\leq} \frac{1}{n}h(X^n) - \frac{1}{n}h(X^n|W_1, W_2, Y_1^n) \\ &\overset{(b)}{\leq} \frac{1}{n}h(X^n) - T\left(\frac{1}{n}h(Y_2^n|Y_1^n) - R_2\right) \\ &= \frac{1}{2}\log(2\pi e\sigma_X^2) - T\left(\frac{1}{2}\log(2\pi e\sigma_{Y_2|Y_1}^2) - R_2\right) \end{aligned}$$

where (a) holds because conditioning reduces entropy, and (b) follows from (16).

Now, expressing $T(x)$ using (13) with $x = \frac{1}{2}\log(2\pi e\sigma_{Y_2|Y_1}^2) - R_2$ we obtain

$$G_{n2}(R_2) \leq \frac{1}{2}\log\left(\frac{1}{(1 - \rho_{XY_1}^2)\Gamma_2}\right) \tag{17}$$

$$\text{where,} \qquad \Gamma_2 = 1 - \rho_{XY_2|Y_1}^2 + \rho_{XY_2|Y_1}^2 \cdot 2^{-2R_2}$$

Finally, using (8), (9) and (17) in (7) we get:

$$\begin{aligned} R_0 + \delta \geq\; & \frac{1}{2}\log\frac{\sigma_X^2}{D_0 + \delta} + \frac{1}{2}\log\left(1 - \rho_{XY_1}^2 + \rho_{XY_1}^2 \cdot 2^{-2R_1}\right) \\ & + \frac{1}{2}\log\left((1 - \rho_{XY_1}^2)(1 - \rho_{XY_2|Y_1}^2 + \rho_{XY_2|Y_1}^2 \cdot 2^{-2R_2})\right) \end{aligned}$$

Letting $\delta \to 0$, the outer region for the two helper case becomes:

$$R_0(D_0) \geq \frac{1}{2}\log^+\left[\frac{\sigma_X^2}{D_0}\prod_{i=1}^{2}(1 - \rho_i^2)\Gamma_i\right]$$

where $\Gamma_i = 1 - \rho_{XY_i|(Y_1,\ldots,Y_{i-1})}^2 + \rho_{XY_i|(Y_1,\ldots,Y_{i-1})}^2 \cdot 2^{-2R_i}$,

and $\rho_i^2 = 1 - \frac{\sigma_{X|Y_1,\ldots,Y_{i-1}}^2}{\sigma_X^2}$ leading to $\rho_1^2 = 0$, and $\rho_2^2 = \rho_{XY_1}^2$. Since the joint distribution $p_{X,Y_1,Y_2}(x, y_1, y_2)$ is known, the correlation coefficient $\rho_{XY_2|Y_1}$, required for evaluating Γ_2, can be calculated to be:

$$\rho_{XY_2|Y_1} = \frac{\rho_{XY_2} - \rho_{XY_1}\rho_{Y_1Y_2}}{\sqrt{1 - \rho_{XY_1}^2}\sqrt{1 - \rho_{Y_1Y_2}^2}}$$

Generalizing the two helper case to m-helpers using exactly the same arguments as above, we obtain Theorem III.1. ∎

Note that when there is no helper, (2) collapses to the classic Gaussian rate-distortion expression [6]:

$$R_0(D_0) \geq \frac{1}{2} \log \frac{\sigma_X^2}{D_0} \tag{18}$$

We now consider some examples for which our derived rate-distortion region for the m-helper system collapses to previously known cases.

Example III.2 (One-helper System.) *On substituting $m = 1$ in (2), we obtain:*

$$R_0 \geq \frac{1}{2} \log \left[\frac{\sigma_X^2}{D_0} \left(1 - \rho^2 + \rho^2 \cdot 2^{-2R_1} \right) \right]$$

where ρ is the correlation between the main source X and the helper Y. This is same as the result stated in [13].

Example III.3 (Two-helpers with $R_2 = 0$.) *Consider a main source X and two helpers Y_1 and Y_2. Since $R_2 = 0$, there is no help obtained from Y_2. This is equivalent to the one helper case and thus we should obtain the rate for one-helper. On substituting $m = 2$ and $R_2 = 0$ in (2), we obtain:*

$$R_0 \geq \frac{1}{2} \log \left[\frac{\sigma_X^2}{D_0} \left(1 - \rho_{XY_1}^2 + \rho_{XY_1}^2 \cdot 2^{-2R_1} \right) \right]$$

which is indeed the expected rate.

IV. Significance of Helpers

We now consider the potential benefit that may be derived from using the helper rates to reduce the distortion in reproduction.

Suppose a sensor is able to report on source X. Also, suppose two other sensors, Y_1 and Y_2 (helper sources), are able to sense the source. However they are farther from the source than X and hence their measurements are not worth reproducing. The algorithm for selecting which sensors act as helpers may depend on the quality of measurement at each sensor.

Assume now that the correlation of the helpers with the source X depends on their distances from X. Let the correlation, ρ, follow an inverse power law with distance, d:

$$\rho = \frac{\rho_0}{d^\alpha}$$

where ρ_0 is a constant of proportionality. Let us take $\alpha = 2$ and evaluate $R_0(D_0)$. Other correlation models can also be used, such as exponential in distance [12, 17].

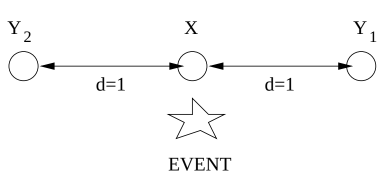

Figure 2: Source Placement

Suppose that each of the helpers is located at distance $d = 1$ from the main source X, and if the three sensors lie along a straight line, the distance between the two helpers becomes $d = 2$ as depicted in Figure 2.

$R_0(D_0)$ is plotted for this scenario in Figure 3. The figure shows that the rate R_0 is reduced when the helpers are used. While the maximum potential benefit is shown by the curve using infinite rates for both the helpers, it can be observed that even with a finite rate, the helpers improve the distortion in the reproduction of X. Also, the graph shows that using more helpers reduces the rate further. However, it may be noted that in practice the distance between the sources will increase as more and more sources are added and hence the correlation will fall. This will make more sources yield diminishing improvement in the rate-distortion performance.

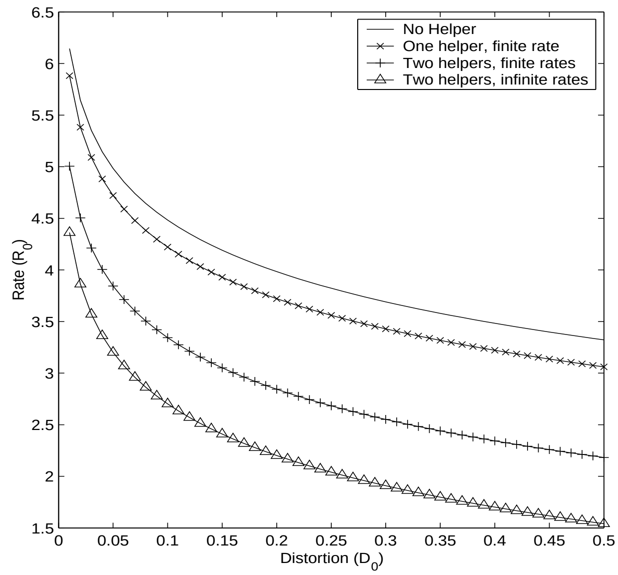

Figure 3: Rate-distortion bounds with multiple helpers. The correlations values are: $\rho_{XY_1} = 0.7, \rho_{XY_2} = 0.7$, and $\rho_{Y_1Y_2} = 0.175$.

V. Conclusions

We considered a multi-terminal network information theory problem when several correlated sources are fused to reproduce a source of interest under a distortion constraint. We presented a generalized solution to this rate-distortion problem with side information and showed that previously known results can be viewed as special cases of the derived expression. The close match for special cases also suggests that the derived lower bound is close to the rate distortion function; however the derivation of the inner region is still an open problem. We also discussed the significance of helper rates and the correlation between them for reducing the distortion in reproduction.

The problem can be extended to the more practical case of non-Gaussian sources. The bound will help definitively compare various data fusion and network coding schemes for wireless networks, with regards to their performance and efficiency.

Acknowledgments

This work was supported in part by the National Science Foundation (NSF) under project #CNS-0306408, the Office of Naval Research (ONR), and the NSF Center for Embedded Networked Sensing under Cooperative Agreement #CCR-0121778 and ITR-Fundamental Limits in Sensor Network #CCR-0217260.

References

[1] M Ahmed. *Decentralized Information Processing in Wireless Peer-to-Peer Networks.* PhD thesis, University of California, Los Angeles, 2003.

[2] M Ahmed and G Pottie. Information theory of wireless sensor networks: the n-helper gaussian case. In *Proceedings of the IEEE International Symposium on Information Theory*, page 436, 2000.

[3] T Berger. *Multiterminal source coding*, pages 171–231. Vienna/New York: Springer-Verlag, 1978.

[4] T Berger, K Houswright, J Omura, S Tung, and J Wolfowitz. An upper bound on the rate distortion function for source coding with partial side information at the decoder. *IEEE Transactions on Information Theory*, IT-25:664–666, November 1979.

[5] T Berger and R Yeung. Multiterminal source encoding with one distortion criterion. *IEEE Transactions on Information Theory*, IT-35:228–236, March 1989.

[6] T Cover and J Thomas. *Elements of Information Theory.* John Wiley & Sons, Inc, 1991.

[7] I Csiszar and J Korner. Towards a general theory of source networks. *IEEE Transactions on Information Theory*, IT-26:155–165, 1980.

[8] D Estrin, R Govindan, and J Heidemann. Embedding the internet: introduction. *Communications of the ACM*, 43(5):38–41, 2000.

[9] T S Han and K Kobayashi. A unified achievable rate region for a general class of multi-terminal source coding systems. *IEEE Transactions on Information Theory*, IT-26:277–288, May 1980.

[10] W J Kaiser, G J Pottie, M B Srivastava, G S Sukhatme, J Villasenor, and D Estrin. Networked infomechanical systems (NIMS) for ambient intelligience. Technical Report 31, Centre For Embedded Networked Sensing, UCLA, 2003. http://www.cens.ucla.edu.

[11] A Kaspi and T Berger. Rate-distortion for correlated sources with partially separated encoders. *IEEE Transactions on Information Theory*, IT-28:828–840, November 1982.

[12] D Marco, E Duarte-Melo, M Liu, and D L Neuhoff. On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data. In *Proceedings of the International Workshop on Information Processing in Sensor Networks (IPSN)*, April 2003.

[13] Y Oohama. Gaussian multiterminal source coding. *IEEE Transactions on Information Theory*, 43(6):1912–1923, November 1997.

[14] Y Oohama. Multiterminal source coding for correlated memoryless gaussian sources with several side information at the decoder. In *Proceedings of IEEE Information Theory and Communications Workshop*, page 100, June 1999.

[15] G J Pottie and W J Kaiser. Wireless integrated network sensors. *Communications of the ACM*, 43(5):51–58, 2000.

[16] A Savvides and M B Srivastava. A distributed computation platform for wireless embedded sensing. In *IEEE International Conference on Computer Design: VLSI in Computers and Processors*, pages 220 – 225, September 2002.

[17] A Scaglione and S D Servetto. On the interdependence of routing and data compression in multi-hop sensor networks. In *Proceedings of the 8th annual international conference on Mobile computing and networking*, pages 140–147. ACM Press, 2002.

[18] D Slepian and J K Wolf. Noiseless coding of correlated information sources. *IEEE Transactions on Information Theory*, IT-19:471–480, July 1973.

[19] S Tung. *Multiterminal source coding.* PhD thesis, Cornell University, Ithaca, NY, May 1978.

[20] P K Varshney. *Distributed Detection and Data Fusion.* New York: Springer-Verlag, 1997.

[21] A Wyner. The rate-distortion function for source coding with side information at the decoder-II: General sources. *Inform. Contr.*, 38:60–80, July 1978.

[22] A Wyner and J Ziv. The rate-distortion function for source coding with side information at the decoder. *IEEE Transactions on Information Theory*, IT-22:1–10, January 1976.