ON THE ACHIEVABLE SUM RATE OF MULTITERMINAL SOURCE CODING FOR A CORRELATED GAUSSIAN VECTOR SOURCE

Jie Chen, A. Lee Swindlehurst

Center for Pervasive Communications and Computing (CPCC) Department of EECS, University of California, Irvine Irvine, CA 92697, USA {jie.chen, swindle}@uci.edu

ABSTRACT

In wireless sensor networks, many monitoring problems can be cast in the form of distributed estimation. If the data links from the sensor nodes to the fusion center have limited capacity, there is a tradeoff between estimation precision and transmission rate. This kind of decentralized estimation system is equivalent to the so-called indirect multiterminal source coding problem, and the Berger-Tung inner bound is the best known achievable rate region boundary. In this paper, we attempt to evaluate the Berger-Tung sum rate for a vector source with correlated components. First we formulate the sum rate as a nonconvex optimization problem with a distortion constraint. Then we propose a method to find an approximate solution. Numerical experiments show the approximate solution is accurate if the required distortion level is relatively small. Its appropriateness is also verified by simulation results based on practical quantizer design.

Index Terms— Indirect multiterminal source coding, distributed estimation, CEO problems, rate distortion theory

1. INTRODUCTION

Wireless sensor networks have been the subject of active research for the past decade, and they find many uses in civil, industrial, commercial and military applications. Such networks are often used for distributed estimation, in which geographically distributed sensors make measurements or local estimates and forward them to a fusion center, which conducts the processing necessary to extract useful information from the reported data. In practice, the local measurements have to be quantized prior to transmission, and there is clearly a tradeoff between the level of quantization (or equivalently the sensor's transmission rate) and the final estimation accuracy. With knowledge of the required estimation accuracy and the statistical characteristics of the source and noise, the fusion center can optimally determine the sensors' individual transmission rates and feed this information back to the sensor nodes in order to efficiently use the available computing and communication resources.

This kind of system is equivalent to the indirect multiterminal source coding problem, which was first studied in [1] and referred to as the central estimation officer (CEO) problem. For memoryless Gaussian scalar sources, the rate region of the CEO problem has been completely characterized in [2, 3, 4]. The authors of [5] gave a new proof for the lower bound of the CEO problem by utilizing the semidefinite partial order of the distortion covariance matrices. In [6], the CEO model was extended to include the vector source case. Xiao and Luo in [7] proposed a numerical method to evaluate the performance for Gaussian vector sources, but they assume the components of the source vector are uncorrelated. In this paper, we attempt to evaluate the Berger-Tung sum rate for a vector source with *correlated* components. First, we formulate the sum rate calculation as a nonconvex optimization problem with a distortion constraint. Using Schur's complement, we propose a method to find an approximate solution. Numerical experiments show the approximate solution is nearly optimal if the target distortion level is relatively small, and thus it is useful for system designs with a highaccuracy requirement. We then implement the multiterminal source coding model using trellis coded vector quantization, and the subsequent empirical simulation results also verify the validity of the approximate solution.

The rest of the paper is organized as follows. In Section 2, we describe the system model and formulate the sum-rate optimization problem for the two-terminal source coding case. Section 3 then proposes a method to find an approximate solution. Simulation results are presented in Section 4 for cases with either symmetric or asymmetric sensor noise. Section 5 summarizes the paper and discusses potential future work.

2. SYSTEM MODEL

Without loss of generality, we assume there are two sensor nodes in the system. The analysis in this paper can be natu-

This work is supported by the Air Force Office of Scientific Research under grant FA9550-10-1-0310, and by the National Science Foundation under grant CCF-0916073.



Fig. 1. Indirect two-terminal source coding

rally extended to the case of more than two nodes. A block diagram model is illustrated in Fig. 1.

The fusion center is interested in recovering the realvalued source vector sequence $\mathbf{x}(t), t = 1, \ldots, \infty$, where every $\mathbf{x}(t)$ is an i.i.d. Gaussian random vector with zero mean. The corresponding covariance matrix $\mathbf{C}_{\mathbf{x}}$ is assumed in general to be non-diagonal, *i.e.*, the components of $\mathbf{x}(t)$ are correlated. We use $\mathbf{v}_i(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{\mathbf{v}_i})$ to indicate that the measurement noise (local estimation error) is real-valued zero-mean Gaussian with covariance $\mathbf{C}_{\mathbf{v}_i}$, and we assume $\mathbf{v}_i(t)$ is also i.i.d. in time. The observation at sensor *i* is given by $\mathbf{y}_i(t) = \mathbf{x}(t) + \mathbf{v}_i(t)$, and $\hat{\mathbf{x}}(t)$ is the estimate obtained at the fusion center. The distortion is defined as $\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^n \|\mathbf{x}(t) - \hat{\mathbf{x}}(t)\|^2$.

Based on [4, 6, 7], if there exist auxiliary random variables \mathbf{w}_1 and \mathbf{w}_2 such that $\mathbf{w}_i \to \mathbf{y}_i \to (\mathbf{x}, \mathbf{y}_{\{1,2\}\setminus i}, \mathbf{w}_{\{1,2\}\setminus i})$ forms a Markov chain, and $E\{\|\mathbf{x} - \hat{\mathbf{x}}(\mathbf{w}_1, \mathbf{w}_2)\|^2\}$ is no greater than the distortion level D, the Berger-Tung achievable rate region is the convex hull of

$$R(\mathbf{w}_{1}, \mathbf{w}_{2}) = \{ (R_{1}, R_{2}) \mid R_{1} \ge I(\mathbf{y}_{1}; \mathbf{w}_{1} | \mathbf{w}_{2}), R_{2} \ge I(\mathbf{y}_{2}; \mathbf{w}_{2} | \mathbf{w}_{1}), R_{1} + R_{2} \ge I(\mathbf{y}_{1}, \mathbf{y}_{2}; \mathbf{w}_{1}, \mathbf{w}_{2}) \}.$$
(1)

In other words, the minimum achievable sum rate can be obtained through minimizing $I(\mathbf{y}_1, \mathbf{y}_2; \mathbf{w}_1, \mathbf{w}_2)$. As in [4, 6, 7], we assume the auxiliary random variables are Gaussian and can be further expressed through forward test channels:

$$\mathbf{w}_1 = \mathbf{A}_1 \mathbf{y}_1 + \mathbf{z}_1 \tag{2}$$

$$\mathbf{w}_2 = \mathbf{A}_2 \mathbf{y}_2 + \mathbf{z}_2. \tag{3}$$

Here \mathbf{A}_i is a matrix and $\mathbf{z}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{\mathbf{z}_i})$ is independent of \mathbf{y}_i . In terms of minimum distortion, the best estimate of \mathbf{x} from \mathbf{w}_i can be obtained using the MMSE estimator $\hat{\mathbf{x}} = E(\mathbf{x}|\mathbf{w}_1, \mathbf{w}_2)$. We use $\tilde{\mathbf{x}}$ to denote the estimation error and $\mathbf{C}_{\tilde{\mathbf{x}}}$ its covariance matrix, so that the distortion requirement $E\{\|\mathbf{x} - \hat{\mathbf{x}}(\mathbf{w}_1, \mathbf{w}_2)\|^2\} \leq D$ is equivalent to $\operatorname{tr}(\mathbf{C}_{\tilde{\mathbf{x}}}) \leq D$. If we let

$$\mathbf{Q}_{i} = \left(\mathbf{C}_{\mathbf{v}_{i}} + \mathbf{A}_{i}^{-1}\mathbf{C}_{\mathbf{z}_{i}}(\mathbf{A}_{i}^{-1})^{T}\right)^{-1}, \qquad (4)$$

then the covariance of $\tilde{\mathbf{x}}$ is

$$\mathbf{C}_{\tilde{\mathbf{x}}} = (\mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_1 + \mathbf{Q}_2)^{-1}.$$
 (5)

The Gaussianity of y_i and w_i results in

$$I(\mathbf{y}_1, \mathbf{y}_2; \mathbf{w}_1, \mathbf{w}_2) = \frac{1}{2} \log \frac{\det(\mathbf{C}_{\mathbf{y}}) \det(\mathbf{C}_{\mathbf{w}})}{\det(\mathbf{C}_{\mathbf{yw}})}, \quad (6)$$

where $\mathbf{C}_{\mathbf{y}}$, $\mathbf{C}_{\mathbf{w}}$, and $\mathbf{C}_{\mathbf{yw}}$ are the covariance matrices of $(\mathbf{y}_1^T \mathbf{y}_2^T)^T$, $(\mathbf{w}_1^T \mathbf{w}_2^T)^T$ and $(\mathbf{y}_1^T \mathbf{y}_2^T \mathbf{w}_1^T \mathbf{w}_2^T)^T$, respectively. After a series of mathematical manipulations, we end up with

$$I(\mathbf{y}_1, \mathbf{y}_2; \mathbf{w}_1, \mathbf{w}_2) = \frac{1}{2} \log \frac{\det(\mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_1 + \mathbf{Q}_2) \det(\mathbf{C}_{\mathbf{x}})}{\det(\mathbf{I} - \mathbf{Q}_1 \mathbf{C}_{\mathbf{v}_1}) \det(\mathbf{I} - \mathbf{Q}_2 \mathbf{C}_{\mathbf{v}_2})}.$$
 (7)

So the minimal Berger-Tung sum rate is the result of the following minimization:

$$\min_{\mathbf{Q}_{1},\mathbf{Q}_{2}} \frac{1}{2} \log \frac{\det(\mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_{1} + \mathbf{Q}_{2}) \det(\mathbf{C}_{\mathbf{x}})}{\det(\mathbf{I} - \mathbf{Q}_{1}\mathbf{C}_{\mathbf{v}_{1}}) \det(\mathbf{I} - \mathbf{Q}_{2}\mathbf{C}_{\mathbf{v}_{2}})} \quad (8)$$
s.t. tr $\left((\mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_{1} + \mathbf{Q}_{2})^{-1}\right) \leq D$
 $\mathbf{Q}_{1} \geq 0, \mathbf{Q}_{2} \geq 0.$

3. APPROXIMATE SOLUTION

The optimization problem (8) is nonconvex and obtaining an exact solution is difficult. To find an approximate solution, we first divide the objective function into two parts,

$$I(\mathbf{y}_{1}, \mathbf{y}_{2}; \mathbf{w}_{1}, \mathbf{w}_{2}) = \frac{1}{2} \log \frac{1}{\det(\mathbf{I} - \mathbf{Q}_{1}\mathbf{C}_{\mathbf{v}_{1}}) \det(\mathbf{I} - \mathbf{Q}_{2}\mathbf{C}_{\mathbf{v}_{2}})} + \frac{1}{2} \log \frac{\det(\mathbf{C}_{\mathbf{x}})}{\det(\mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_{1} + \mathbf{Q}_{2})^{-1}}.$$
(9)

The minimization of the first term can be expressed as

$$\min_{\mathbf{Q}_{1},\mathbf{Q}_{2}} \frac{1}{2} \log \frac{1}{\det(\mathbf{I} - \mathbf{Q}_{1}\mathbf{C}_{\mathbf{v}_{1}})\det(\mathbf{I} - \mathbf{Q}_{2}\mathbf{C}_{\mathbf{v}_{2}})} \quad (10)$$
s.t. $\operatorname{tr}\left((\mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_{1} + \mathbf{Q}_{2})^{-1}\right) \leq D$
 $\mathbf{Q}_{1} \succeq 0, \mathbf{Q}_{2} \succeq 0,$

which is still not convex in the constraints. However, as we will see later in this section, we can find the exact solution to this subproblem, which can be plugged into (7) to compute a sum rate.

In addition, we will see that the second term has a lower bound, which in turn bounds the sum rate in (7). Numerical experiments show that the computed sum rate is extremely close to the sum rate bound for small target distortions. So in this range the solution to problem in (10) can be used as a good approximation for the Berger-Tung sum rate problem. To solve (10), we replace the distortion constraint with a stricter matrix constraint **D** that bounds the error covariance: $(\mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_1 + \mathbf{Q}_2)^{-1} \leq \mathbf{D}$. Using Schur's complement, it is equivalent to require

$$\begin{pmatrix} \mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_1 + \mathbf{Q}_2 & \mathbf{I} \\ \mathbf{I} & \mathbf{D} \end{pmatrix} \succeq \mathbf{0}.$$
 (11)

The optimization problem (10) can then be recast as

$$\min_{\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{D}} \frac{1}{2} \log \frac{1}{\det(\mathbf{I} - \mathbf{Q}_1 \mathbf{C}_{\mathbf{v}_1}) \det(\mathbf{I} - \mathbf{Q}_2 \mathbf{C}_{\mathbf{v}_2})}$$
(12)

s.t.
$$\operatorname{tr}(\mathbf{D}) \le D$$
 (13)

$$\begin{pmatrix} \mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_1 + \mathbf{Q}_2 & \mathbf{I} \\ \mathbf{I} & \mathbf{D} \end{pmatrix} \succeq 0$$
(14)

$$\mathbf{D} \succeq \mathbf{0} \tag{15}$$

$$\mathbf{Q}_1 \succeq 0 \tag{16}$$

$$\mathbf{Q}_2 \succeq \mathbf{0}. \tag{17}$$

With this modification, the problem is now a convex problem and can be efficiently solved. Generally speaking, the solutions to (10) and (12) are equal only when the constraints (13) and (14) are simultaneously fulfilled with equality. Otherwise the two solutions are not necessarily the same. In the proof of the following theorem, we will see that in fact the solution to (12) also minimizes (10).

Theorem 1. The optimal \mathbf{Q}_1 and \mathbf{Q}_2 of problem (12) also minimize (10).

Proof. The first order derivatives of the Lagrangian L with respect to the optimization variables are

$$\frac{\partial L}{\partial \mathbf{Q}_1} = \left(\mathbf{C}_{\mathbf{v}_1} (\mathbf{I} - \mathbf{Q}_1 \mathbf{C}_{\mathbf{v}_1})^{-1} \right) - \mathbf{\Lambda}_1 - \mathbf{\Lambda}_4^{TL}$$
(18)

$$\frac{\partial L}{\partial \mathbf{Q}_2} = \left(\mathbf{C}_{\mathbf{v}_2} (\mathbf{I} - \mathbf{Q}_2 \mathbf{C}_{\mathbf{v}_2})^{-1} \right) - \mathbf{\Lambda}_2 - \mathbf{\Lambda}_4^{TL}$$
(19)

$$\frac{\partial L}{\partial \mathbf{D}} = \lambda \mathbf{I} - \mathbf{\Lambda}_3 - \mathbf{\Lambda}_4^{BR},\tag{20}$$

where $\lambda \geq 0$, $\Lambda_i \succeq 0$ are the Lagrange multipliers, and

$$\mathbf{\Lambda}_4 = \begin{pmatrix} \mathbf{\Lambda}_4^{TL} & \mathbf{\Lambda}_4^{TR} \\ \mathbf{\Lambda}_4^{BL} & \mathbf{\Lambda}_4^{BR} \end{pmatrix}.$$
(21)

The KKT conditions are

$$\operatorname{tr}(\mathbf{Q}_1 \boldsymbol{\Lambda}_1) = 0 \tag{22}$$

$$\operatorname{tr}(\mathbf{Q}_2 \mathbf{\Lambda}_2) = 0 \tag{23}$$

$$tr(\mathbf{D}\mathbf{\Lambda}_3) = 0 \tag{24}$$

$$\lambda(\operatorname{tr}(\mathbf{D}) - D) = 0 \tag{25}$$

tr
$$\left(\begin{pmatrix} \mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_1 + \mathbf{Q}_2 & \mathbf{I} \\ \mathbf{I} & \mathbf{D} \end{pmatrix} \mathbf{\Lambda}_4 \right) = 0.$$
 (26)

From (24) it can be shown that $\Lambda_3 = 0$ and therefore

$$\lambda \mathbf{I} - \mathbf{\Lambda}_4^{BR} = \mathbf{0}.$$
 (27)



Fig. 2. Approximate solution and the lower bound

If $\lambda > 0$, the above equation results in a nonsingular Λ_4^{BR} . From (26) we have

$$(\mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_1 + \mathbf{Q}_2)\mathbf{D}\mathbf{\Lambda}_4^{BR} = \mathbf{\Lambda}_4^{BR}, \qquad (28)$$

so in this case $\mathbf{D} = (\mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_1 + \mathbf{Q}_2)^{-1}$. Since $\lambda > 0$, equation (25) requires $\operatorname{tr}(\mathbf{D}) = D$. In other words the constraints (13) and (14) are simultaneously fulfilled with equality, so the solutions to (10) and (12) are equivalent.

If $\lambda = 0$, from (27) we have $\Lambda_4^{BR} = \mathbf{0}$. Further by using (18), (19), (22) and (23), we obtain $\mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{0}$. This solution makes the minimal value of (12) equal zero. Due to the positive semidefiniteness of \mathbf{Q}_1 and \mathbf{Q}_2 , we have $(\mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_1 + \mathbf{Q}_2)^{-1} \leq \mathbf{C}_{\mathbf{x}}$. Therefore

$$D \ge \operatorname{tr}(\mathbf{D}) \ge \operatorname{tr}(\mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{Q}_1 + \mathbf{Q}_2)^{-1} = \operatorname{tr}(\mathbf{C}_{\mathbf{x}}).$$
(29)

For the $D \geq \operatorname{tr}(\mathbf{C}_{\mathbf{x}})$ case, it is easy to verify the minimal value of (10) is also zero. In fact, this corresponds to the trivial case where there is no data transmission from the sensor nodes to the fusion center.

In summary, no matter what value λ has, the solutions to (10) and (12) are equivalent.

Plugging the \mathbf{Q}_1 and \mathbf{Q}_2 of Theorem 1 into (7), we obtain an approximation of the Berger-Tung sum rate. The second term in (9) is lower bounded by $\frac{1}{2} \log \frac{\det(\mathbf{C}_{\mathbf{x}})}{\det(\mathbf{D})}$. From the property $\det(\mathbf{D}) \leq (\frac{1}{n} \operatorname{tr}(\mathbf{D}))^n$ and the fact that $\operatorname{tr}(\mathbf{D}) \leq D$, the sum rate is lower bounded by

$$\frac{1}{2}\log\frac{\det(\mathbf{C}_{\mathbf{x}})}{\left(\frac{D}{n}\right)^{n}\det(\mathbf{I}-\mathbf{Q}_{1}\mathbf{C}_{\mathbf{v}_{1}})\det(\mathbf{I}-\mathbf{Q}_{2}\mathbf{C}_{\mathbf{v}_{2}})}.$$
(30)

Our numerical experiments consistently show that for small distortions D, the approximate sum rate and the bound are very close. So in this range the solution in Theorem 1 can



Fig. 3. Coding chain

be regarded as a good approximation to the original problem. An example is given in Fig. 2, where

$$\mathbf{C}_{\mathbf{x}} = \begin{pmatrix} 0.99 & 0.9\\ 0.9 & 0.99 \end{pmatrix} \text{ and } \mathbf{C}_{\mathbf{v}_1} = \mathbf{C}_{\mathbf{v}_2} = \begin{pmatrix} 0.01 & 0\\ 0 & 0.01 \end{pmatrix}.$$
(31)

4. NUMERICAL RESULTS

In this section, we use practical quantizers to verify the proposed solution in Section 3. As reported in the literature, the Berger-Tung achievable sum rate can be approached by using powerful quantizers and Slepian-Wolf encoders. Fig. 3 shows the coding chain. In the figure, Q, EC, and SW stand for quantizer, entropy encoder, and Slepian-Wolf encoder respectively. We use a trellis coded vector quantizer (TCVQ) in our simulation, and ideal entropy and Slepian-Wolf encoding are assumed. The optimal Q_1 and Q_2 in Theorem 1 are used in the MMSE calculation.

Fig. 4 shows the result of a case where the noise power is the same at the two sensors. The dimension of the source vector is two. The correlation between the two components and the noise power is characterized in (31). The target distortion level is 0.0102. The dashed line corresponds the sum-rate lower bound obtained from (30). The solid curve describes the boundary of (1) when the proposed approximate solution is adopted. The cross point is the rate pair resulting from the empirical simulations. Fig. 5 shows the result of a case with asymmetric noise. The parameters are the same as before, except the noise power at node 2 is 0.04. The target distortion level is 0.0166. Both of the simulation results indicate that the approximate solution is very near the optimal sum rate.

5. CONCLUSION

In this paper, we found an approximate solution to the vector Berger-Tung achievable sum rate problem for a correlated Gaussian source. Numerical experiments show it is near optimal if high estimation accuracy is required. Furthermore, we used TCVQ to implement the multiterminal source coding model, and our empirical simulation results also indicate



the near optimality of the approximate solution. Possible future work includes examining the performance gap that results from using an optimal LDPC encoder for the Slepian-Wolf encoder. In addition, the use of linear MMSE estimators in the coding chain is only optimal if the quantization noise is i.i.d. Gaussian. Since this assumption cannot be guaranteed in practice, finding non-linear estimators to improve the estimation performance is also of interest.

6. REFERENCES

- T. Berger, Zhen Zhang, and H. Viswanathan, "The CEO problem," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 887–902, 1996.
- [2] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1057–1070, 1998.
- [3] V. Prabhakaran, D. Tse, and K. Ramachandran, "Rate region of the quadratic Gaussian CEO problem," in *Proc. Int. Symp. Information Theory ISIT 2004*, 2004.
- [4] Jun Chen, Xin Zhang, T. Berger, and S. B. Wicker, "An upper bound on the sum-rate distortion function and its corresponding rate allocation schemes for the CEO problem," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 977–987, 2004.
- [5] Jia Wang, Jun Chen, and Xiaolin Wu, "On the sum rate of Gaussian multiterminal source coding: New proofs and results," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3946–3960, 2010.
- [6] S. Tavildar and P. Viswanath, "On the sum-rate of the vector Gaussian CEO problem," in *Proc. Conf Signals, Systems and Computers Record of the Thirty-Ninth Asilomar Conf*, 2005, pp. 3–7.
- [7] Jin-Jun Xiao and Zhi-Quan Luo, "Optimal rate allocation for the vector Gaussian CEO problem," in *Proc. 1st IEEE Int Computational Advances in Multi-Sensor Adaptive Processing Workshop*, 2005, pp. 56–59.