Lecture 4: Distributed Source Coding

Jun Chen

1 Slepian-Wolf Region:

Give an i.i.d. source $\{X_1(t), \cdots, X_L(t)\}$ with $(X_1(t), \cdots, X_L(t)) \sim p(x_1, \cdots, x_L)$ for all t, the Slepian-Wolf region is

$$\mathcal{R}_{SW} = \left\{ (R_1, \cdots, R_L) : \sum_{i \in \mathcal{A}} R_i \geq H(X_{\mathcal{A}} | X_{\mathcal{A}^c}), \forall \mathcal{A} \subseteq \mathcal{I}_L \right\},$$

where $\mathcal{I}_L = \{1, 2, \cdots, L\}$.

Coding Scheme: For source X_i, there are $2^{nH(X_i)}$ typical sequences of length n. Randomly distributed these typical sequences to 2^{nR_i} bins such that each bin contains $2^{n(H(X_i)-R_i)}$ sequences. Given a source realization $\mathbf{x}_i$, the index of the bin that contains this sequences is conveyed to the decoder. The decoder collects L bin indices and searches for L sequences, one in each bin, that are jointly typical.

We can bound the probability of error event as

$$\begin{aligned}
& \mathbb{P}\Big((\mathbf{x}_1, \cdots, \mathbf{x}_{j-1}, \widetilde{\mathbf{X}}_j, \cdots, \widetilde{\mathbf{X}}_L) \text{ are jointly typical }, \widetilde{\mathbf{X}}_i \text{ and } \mathbf{x}_i \text{ are in the same bin }, i = j, \cdots, L \Big) \\
\leq\ & 2^{n[\sum_{i=j}^{L}(H(X_i)-R_i)]} 2^{n\sum_{i=j}^{L} I(X_i; X_1^{i-1})} \\
=\ & 2^{n[H(X_j^L | X_1^{j-1}) - \sum_{i=j}^{L} R_i]},
\end{aligned}$$

which goes to 0 as $n \to \infty$ if $\sum_{i=j}^{L} R_i > H(X_j^L | X_1^{j-1})$. Other rate constraints can be similarly derived.

It is clear from the proof that the codewords are jointly decoded. So it is natural to ask whether low-complexity successive decoding schemes exist.

2 Wyner-Ziv Rate-Distortion Function:

$\{X(t), Y(t)\}$ is an i.i.d. source and $d(\cdot, \cdot)$ is a distortion measure. $\{Y(t)\}$ is directly available to the decoder but $\{X(t)\}$ is not. $\{Y(t)\}$ is referred to as the decoder side information.

$$R_{WZ}(D) = \min I(X; W | Y)$$

where the minimization is over all W such that $W \to X \to Y$ forma a Markov chain and there exists a function $f : \mathcal{W} \times \mathcal{Y} \to \mathcal{X}$ with $Ed(f(W, Y), X) \leq D$.

The Wyner-Ziv coding consists two steps: 1. Quantization (classic rate-distortion theory); 2. Slepian-Wolf binning.

First independently generate $2^{nI(X;W)}$ codewords $\{\mathbf{W}(j)\}_{j=1}^{2^{nI(X;W)}}$ according to $\prod p(w)$. Randomly distribute these codewords into 2^{nR} bins such that each bin contains $2^{n(H(X;W)-R)}$ codewords. With high probability, for each realization $\mathbf{x}$ we can find a codeword $\mathbf{W}(j^*)$ jointly typical with $\mathbf{x}$. Send the index of the bin that contains $\mathbf{W}(j^*)$ to the decoder. The decoder searches in this bin for a codeword that is jointly typical with $\mathbf{Y}$. The probability of existing other codewords in this bin jointly typical with $\mathbf{Y}$ is upper bounded by

$$2^{n(H(X;W)-R)}2^{-nI(W;Y)},$$

which goes to 0 as $n \to \infty$ if $R > H(X;W) - I(W;Y) = I(X;W|Y)$.

The reason we need the Markov condition $W \to X \to Y$ is that the encoder can only guarantee that $\mathbf{X}$ is typical with $\mathbf{W}$ since $\mathbf{Y}$ is unknown at encoder; with this Markov condition, the fact that X is typical with $\mathbf{W}$ implies $(\mathbf{X}, \mathbf{W}, \mathbf{Y})$ are jointly typical with high probability.

Quadratic Gaussian case: X and Y are zero-mean jointly Gaussian and $d(\cdot,\cdot)$ is the squared distortion measure. Let

$$\Delta X = X - \mathbb{E}(X|Y)$$

and

$$\sigma^2_{X|Y} = E(\Delta X)^2.$$

If $\{Y(t)\}$ is known at both the encoder and decoder, then the rate-distortion function is

$$\frac{1}{2}\log\frac{\sigma^2_{X|Y}}{D}.$$

For the Wyner-Ziv case, let $W = X + Z$ where $Z \sim \mathcal{N}(0, \sigma^2_Z)$ is independent of X and Y. We have

$$\begin{aligned} I(X;W|Y) &= I(X - \mathbb{E}(X|Y); W - \mathbb{E}(X|Y)|Y) \\ &= I(\Delta X; \Delta X + Z) \\ &= \frac{1}{2}\log\frac{\sigma^2_{X|Y}}{D}, \end{aligned}$$

where we let $D = \frac{\sigma^2_{X|Y}\sigma^2_Z}{\sigma^2_{X|Y}+\sigma^2_Z}$. So for the quadratic case, there is no rate loss if the encoder does not have the side information. But it should be emphasized that the underlying coding scheme is very different from the case that the side information is available at encoder. The same performance is achieved at the cost of higher coding complexity.

3 Berger-Tung Region:

The Berger-Tung region can be defined for many different distributed lossy source coding scenarios. Here we shall study the simplest one, called the CEO problem.

CEO problem: $\{X(t), Y_1(t), \cdots, Y_L(t)\}$ is an i.i.d. source with $(X(t), Y_1(t), \cdots, Y_L(t)) \sim p(x, y_1, \cdots, y_L)$ for all t. $\{Y_i(t)\}$ is available at encoder i, $i = 1, \cdots, L$. The decoder reconstructs $\{\hat{X}(t)\}$ such that the distortion constraint is satisfied.

The Berger-Tung coding scheme is a direct generalization of the Wyner-Ziv coding scheme. Roughly speaking, each encoder first quantizes its observation, then apply the Slepian-Wolf coding to the quantized data. This leads to the following achievable region:

$$\mathcal{R}_{BT} = \bigcup \mathcal{R}_{BT}(W_{\mathcal{I}_L}),$$

where

$$\mathcal{R}_{BT}(W_{\mathcal{I}_L}) = \left\{ (R_1, \cdots, R_L) : \sum_{i \in \mathcal{A}} R_i \geq I(Y_{\mathcal{A}}; W_{\mathcal{A}} | W_{\mathcal{A}^c}),\ \forall \mathcal{A} \subseteq \mathcal{I}_L \right\}$$

and the union is over all $(W_1, \cdots, W_L)$ such that $W_i \to Y_i \to (X, Y_{\{i\}^c}, W_{\{i\}^c})$ for all $i \in \mathcal{I}_L$ and there exists a function $f : \prod \mathcal{W}_i \to \mathcal{X}$ with $\mathbb{E}d(X, f(W_1, \cdots, W_L)) \leq D$.

Note: The Berger-Tung region is an achievable rate region. It has been shown that this region is not tight for some cases. But the Berger-Tung region is tight for the important quadratic Gaussian case.

4 Contra-polymatroid Structure:

Definition 4.1. *Let $f : 2^{\mathcal{I}_L} \to \mathcal{R}_+$ be a set function. The polyhedron*

$$\mathcal{G}(f) \triangleq \left\{ (x_1, \cdots, x_L) : \sum_{i \in \mathcal{S}} x_i \geq f(\mathcal{S}), \forall \mathcal{S} \subseteq \mathcal{I}_L, x_i \geq 0, \forall i \right\} \tag{1}$$

is a contra-polymatroid if the set function f satisfies

1. *$f(\emptyset) = 0$ (normalized).*
2. *$f(\mathcal{S}) \leq f(\mathcal{T})$ if $\mathcal{S} \subset \mathcal{T}$ (nondecreasing).*
3. *$f(\mathcal{S}) + f(\mathcal{T}) \leq f(\mathcal{S} \cup \mathcal{T}) + f(\mathcal{S} \cap \mathcal{T})$ (supermodular).*

If f satisfies the three properties, f is called a rank function.

If π is a permutation on the set $\mathcal{I}_L$, define the vector $\mathbf{v}(\pi) \in \mathcal{R}^L$ by $v_{\pi(1)}(\pi) = f(\pi(1))$ and

$$v_{\pi(i)}(\pi) = f(\{\pi(1), \cdots, \pi(i)\}) - f(\{\pi(1), \cdots, \pi(i-1)\})$$

for $i = 2, 3, \cdots, L$.

Theorem 4.1. *Let $\mathcal{G}(f)$ be a contra-polymatroid. Then the points $\mathbf{v}(\pi)$ where π is a permutation on $\mathcal{I}_L$ are precisely the vertices of $\mathcal{G}(f)$. Moreover, if λ is a given vector in $\mathcal{R}_+^L$, then a solution of the optimization problem*

$$\min \lambda \cdot \mathbf{x} \quad \textit{subject to } \mathbf{x} \in \mathcal{G}(f)$$

is attained at a point $\mathbf{v}(\pi^*)$, where the π^* is any permutation such that $\lambda_{\pi^*(1)} \geq \cdots \geq \lambda_{\pi^*(L)}$. Conversely, if f is a set function and $\mathbf{v}(\pi) \in \mathcal{G}(f)$ for every permutation π, then $\mathcal{G}(f)$ is a contra-polymatroid.

Theorem 4.2. *$\mathcal{R}_{SW}$ and $\mathcal{R}_{BT}(W_{\mathcal{I}_L})$ (with $W_i \to Y_i \to (Y_{\{i\}^c}, W_{\{i\}^c})$ for all $i \in \mathcal{I}_L$) are contra-polymatroids.*

The vertices of $\mathcal{R}_{SW}$ are given by

$$\begin{aligned} R_{\pi(i)}(\pi) &= H(X_{\pi(1)}, \cdots, X_{\pi(i)} | X_{\pi(i+1)}, \cdots, X_{\pi(L)}) - H(X_{\pi(1)}, \cdots, X_{\pi(i-1)} | X_{\pi(i)}, \cdots, X_{\pi(L)}) \\ &= H(X_{\pi(i)} | X_{\pi(i+1)}, \cdots, X_{\pi(L)}), \quad i = 1, 2, \cdots, L-1, \\ R_{\pi(L)}(\pi) &= H(X_{\pi(L)}) \end{aligned}$$

for every π. Similarly, the vertices of $\mathcal{R}_{BT}(W_{\mathcal{I}_L})$ are given by

$$\begin{aligned} R_{\pi(i)}(\pi) &= I(Y_{\pi(1)}, \cdots, Y_{\pi(i)}; W_{\pi(1)}, \cdots, W_{\pi(i)} | W_{\pi(i+1)}, \cdots, W_{\pi(L)}) \\ &\quad - I(Y_{\pi(1)}, \cdots, Y_{\pi(i-1)}; W_{\pi(1)}, \cdots, W_{\pi(i-1)} | W_{\pi(i)}, \cdots, W_{\pi(L)}) \\ &= I(Y_{\pi(i)}; W_{\pi(i)} | W_{\pi(i+1)}, \cdots, W_{\pi(L)}), \quad i = 1, 2, \cdots, L-1, \\ R_{\pi(L)}(\pi) &= I(Y_{\pi(L)}; W_{\pi(L)}) \end{aligned}$$

for every π. Clearly, each vertex is associated with a successive decoding scheme.

We can define

$$\mathcal{D}_{SW} = \left\{ (R_1, \cdots, R_L) \in \mathcal{R}_{SW} : \sum_{i=1}^{L} R_i = H(X_{\mathcal{I}_L}) \right\}$$

as the dominant face of $\mathcal{R}_{SW}$ and

$$\mathcal{D}_{BT}(W_{\mathcal{I}_L}) = \left\{ (R_1, \cdots, R_L) \in \mathcal{R}_{BT}(W_{\mathcal{I}_L}) : \sum_{i=1}^{L} R_i = I(Y_{\mathcal{I}_L}; W_{\mathcal{I}_L}) \right\}$$

as the dominant face of $\mathcal{R}(X_{\mathcal{I}_L})$. Why are they non-empty?

The geometric structures of $\mathcal{D}_{SW}$ and $\mathcal{D}_{BT}(W_{\mathcal{I}_L})$ are completely analogous to that of the multiaccess rate region. Again, rate tuples on the boundary of the dominant face can be achieved via successive group decoding. You can easily work it out by yourself.

5 Source Splitting and Quantization Splitting:

Analogous to the rate splitting in multiaccess communications, we have source splitting and quantization splitting in distributed source coding. For simplicity, we only discuss the two-user case. The generalization to the L-user case follows from the same approach in multiaccess channels.

Let's look at the Slepian-Wolf region first. Let (R_1, R_2) be a rate pair on the dominant face. We must have

$$\begin{aligned} & R_1 + R_2 = H(X_1, X_2), \\ & R_1 \in [H(X_1|X_2), H(X_1)], \\ & R_2 \in [H(X_2|X_1), H(X_2)]. \end{aligned}$$

Introduce an auxiliary random variable X_2' such that $X_2' \to X_2 \to X_1$. It is clear that we can find a class of transition probability $p_\epsilon(x_2'|x_2)$ such that $I(X_2; X_2')$ varies from 0 to $H(X_2)$ as ϵ goes from 0 to 1. Hence for X_2' associated with certain ϵ, we have

$$\begin{aligned} R_1 &= H(X_1|X_2'), \\ R_2 &= I(X_2; X_2') + H(X_2|X_1, X_2'). \end{aligned}$$

Correspondingly, decoder first decodes X_2', then X_1, and finally X_2. We can view $\mathbf{X}_2'$ as a block quantization of $\mathbf{X}_2$ in the classic rate-distortion theory. For the special case, we can let X_2' be a deterministic function of X_2. In this case, we have

$$R_2 = H(X_2') + H(X_2|X_1, X_2').$$

Now X_2' is essentially a scalar quantization of X_2.

For a rate tuple (R_1, R_2) on the dominant face of $\mathcal{R}_{BT}(W_{\mathcal{I}_2})$, we can introduce an auxiliary random variable W_2' with $W_2' \to W_2 \to (Y_1, W_1)$ such that

$$\begin{aligned} R_1 &= I(Y_1; W_1 | W_2'), \\ R_2 &= I(Y; W_2') + H(Y_2; W_2 | W_1, W_2'). \end{aligned}$$

Correspondingly, decoder first decodes W_2', then W_1, and finally W_2.

6 Quadratic Gaussian CEO Problem:

Model: $X \sim \mathcal{N}(0, \sigma_X^2)$, $Y = X + N_i$ with $N_i \sim \mathcal{N}(0, \sigma_{N_i}^2)$, $i = 1, \cdots, L$, and $(X, N_1, \cdots, N_L)$ are independent.

Let $W_i = Y_i + T_i$ with $T_i \sim \mathcal{N}(0, \sigma_{T_i}^2)$, $i = 1, \cdots, L$. Here $(X, N_1, \cdots, N_L, T_1, \cdots, T_L)$ are all independent. Let

$$r_i = I(Y_i, W_i | X) = \frac{1}{2} \log \frac{\sigma_{N_i}^2 + \sigma_{T_i}^2}{\sigma_{T_i}^2}, \quad \forall\, i \in \mathcal{I}_L,$$

i.e., $\sigma_{T_i}^2 = \frac{\sigma_{N_i}^2}{\exp(2r_i) - 1}$. We have

$$\begin{aligned} \sigma_{X|W_{\mathcal{A}}}^2 = \mathbb{E}(X - \mathbb{E}(X|W_{\mathcal{A}}))^2 &= \left(\frac{1}{\sigma_X^2} + \sum_{i \in \mathcal{A}} \frac{1}{\sigma_{N_i}^2 + \sigma_{T_i}^2} \right)^{-1} \\ &= \left(\frac{1}{\sigma_X^2} + \sum_{i \in \mathcal{A}} \frac{1 - \exp(-2r_i)}{\sigma_{N_i}^2} \right)^{-1} \end{aligned}$$

and

$$\begin{aligned} \sum_{i \in \mathcal{A}} R_i &\geq I(Y_{\mathcal{A}}; W_{\mathcal{A}} | W_{\mathcal{A}^c}) \\ &= I(X, Y_{\mathcal{A}}; W_{\mathcal{A}} | W_{\mathcal{A}^c}) \\ &= I(X; W_{\mathcal{A}} | W_{\mathcal{A}^c}) + I(Y_{\mathcal{A}}; W_{\mathcal{A}} | X) \\ &= I(X; W_{\mathcal{A}} | W_{\mathcal{A}^c}) + \sum_{i \in \mathcal{A}} r_i \\ &= h(X | W_{\mathcal{A}^c}) - h(X | W_{\mathcal{I}_L}) + \sum_{i \in \mathcal{A}} r_i \\ &= \frac{1}{2} \log \left(2\pi e \sigma_{X|W_{\mathcal{A}^c}}^2 \right) - \frac{1}{2} \log \left(2\pi e \sigma_{X|W_{\mathcal{I}_L}}^2 \right) + \sum_{i \in \mathcal{A}} r_i. \end{aligned}$$

So

$$\begin{aligned} \mathcal{R}_{BT}(W_{\mathcal{I}_L}) &= \left\{ (R_1, \cdots, R_L) : \sum_{i \in \mathcal{A}} R_i \geq \frac{1}{2} \log \left(\frac{\frac{1}{\sigma_X^2} + \sum_{i=1}^{L} \frac{1 - \exp(-2r_i)}{\sigma_{N_i}^2}}{\frac{1}{\sigma_X^2} + \sum_{i \in \mathcal{A}^c} \frac{1 - \exp(-2r_i)}{\sigma_{N_i}^2}} \right) + \sum_{i \in \mathcal{A}} r_i, \forall\, \mathcal{A} \subseteq \mathcal{I}_L \right\} \\ &\triangleq \mathcal{R}_{BT}(r_{\mathcal{I}_L}) \end{aligned}$$

and

$$\mathcal{R}_{BT} = \bigcup_{(r_1,\cdots,r_L)\in\mathcal{F}(D)} \mathcal{R}_{BT}(r_{\mathcal{I}_L}),$$

where

$$\mathcal{F}(D) = \left\{(r_1,\cdots,r_L) \in \mathbb{R}_+^L : \frac{1}{\sigma_X^2} + \sum_{i=1}^{L} \frac{1-\exp(-2r_i)}{\sigma_{N_i}^2} \geq \frac{1}{D}\right\}.$$

The minimum sum-rate distortion function is given by the following optimization problem:

$$\min_{(r_1,\cdots,r_L)\in\mathcal{F}(D)} \frac{1}{2}\log\left(1+\sigma_X^2\sum_{i=1}^{L}\frac{1-\exp(-2r_i)}{\sigma_{N_i}^2}\right) + \sum_{i=1}^{L} r_i$$

which is equivalent to

$$\min_{\frac{1}{\sigma_X^2}+\sum_{i=1}^{L}\frac{1-\exp(-2r_i)}{\sigma_{N_i}^2}=\frac{1}{D}} \frac{1}{2}\log\frac{\sigma_X^2}{D} + \sum_{i=1}^{L} r_i.$$

This optimization problem can be easily solved using Lagrange multiplies. Suppose $\sigma_{N_{\pi(1)}}^2 \leq \cdots \leq \sigma_{N_{\pi(L)}}^2$. Find the largest $\widetilde{L}$ such that

$$\frac{\widetilde{L}}{\sigma_{N_{\pi(\widetilde{L})}}^2} - \left(\frac{1}{D_0(\widetilde{L})} - \frac{1}{D}\right) \geq 0,$$

where $D_0(\widetilde{L}) = (1/\sigma_X^2 + \sum_{i=1}^{\widetilde{L}} 1/\sigma_{N_{\pi(i)}}^2)^{-1}$. We have

$$\begin{aligned}
r_{\pi(i)}^* &= \frac{1}{2}\log\frac{\widetilde{L}}{\sigma_{N_{\pi(i)}}^2\left(\frac{1}{D_0(\widetilde{L})}-\frac{1}{D}\right)}, \quad i=1,\cdots,\widetilde{L}, \\
r_{\pi(i)}^* &= 0, \quad i=\widetilde{L}+1,\cdots,L,
\end{aligned}$$

and the resulting sum-rate distortion function is

$$\frac{1}{2}\log\left\{\frac{\sigma_X^2}{D}\prod_{i=1}^{\widetilde{L}}\left[\frac{\widetilde{L}}{\sigma_{N_{\pi(i)}}^2\left(\frac{1}{D_0(\widetilde{L})}-\frac{1}{D}\right)}\right]\right\}.$$

For the special case when $\sigma_{N_1}^2 = \cdots = \sigma_{N_L}^2 = \sigma_N^2$, the expression can be simplified to

$$\frac{1}{2}\log\left\{\frac{\sigma_X^2}{D}\left(\frac{D\sigma_X^2 L}{D\sigma_X^2 L - \sigma_X^2\sigma_N^2 + D\sigma_N^2}\right)^L\right\}.$$

Note: All the rate tuples on the dominant face of $\mathcal{R}_{BT}(r_{\mathcal{I}_L}^*)$ achieve the mininum sum-rate distortion function. As $L\to\infty$, the above expression converges to

$$\frac{1}{2}\log\frac{\sigma_X^2}{D} + \frac{\sigma_N^2}{2\sigma_X^2}\left(\frac{\sigma_X^2}{D}-1\right).$$

Remark: By exploiting the contra-polymatroid structure of $\mathcal{R}_{BT}(W_{\mathcal{I}_L})$, you can completely characterize the boundary of $\mathcal{R}_{BT}$, not just the minimum sum-rate face.

References

[1] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Info. Theory*, vol.IT-19, pp. 471-480, Jul. 1973.

[2] T. M. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Trans. Inform. Theory*, vol. 21, pp. 226-228, Mar. 1975.

[3] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Info. Theory*, vol. 22, no. 1, pp. 1-10, Jan. 1976.

[4] B. Rimoldi and R. Urbanke, "Asynchronous Slepian-Wolf coding via source-splitting," in *IEEE International Symposium on Information Theory*, Ulm, Germany, June 29 - July 4 1997, p. 271.

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

[6] J. Edmonds, "Submodular functions, matroids and certain polyhedra, in *Combinatorial structures and their applications* (R. Guy, H. Hanani, N. Sauer, and J. Schonheim, eds.), pp. 69-87, Gordon and Breach, New York, 1970. (Proc. Calgary Int. Conf. 1969).

[7] D. N. C. Tse and S. V. Hanly, "Multiaccess fading channels-part I: polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Info. Theory*, vol. 44, pp. 2796-2815, Nov. 1998.

[8] S. I. Gel'fand and M. S. Pinsker, "Coding of sources on the basis of observations with incomplete information". *Problems of Information Transmission*, 15(2):115-125, 1979.

[9] T. J. Flynn and R. M. Gray, "Encoding of correlated observations," *IEEE Trans. Inform. Theory*, vol. 33, pp. 773-787, Nov. 1987.

[10] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Trans. Inform. Theory*, vol. 42, pp. 887-902, May 1996.

[11] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications* (G. Longo, ed.), vol. 229 of *CISM Courses and Lectures*, pp. 171-231, Springer-Verlag, Vienna/New York, 1978.

[12] S. Y. Tung, "Multiterminal source coding," Ph.D. dissertation, School of Electrical Engineering, Cornell Univ., Ithaca, NY, May 1978.

[13] T. Berger, K. Housewright, J. Omura, S. Y. Tung, and J. Wolfowitz, "An upper bound on the rate-distortion function for source coding with partial side information at the decoder," *IEEE Trans. Inform. Theory*, vol. IT-25, pp.664-666, Nov., 1979.

[14] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources (Corresp.)," *IEEE Trans. on Inform. Theory*, vol. 25, pp. 219-221, Mar 1979.

[15] P. Viswanath, "Sum rate of multiterminal gaussian source coding," in *Network Information Theory* (P. Gupta, G. Kramer, and A. Wijngaarden, eds.), DIMACS: Series in Discrete Mathematics and Theoretical Computer Science, American Mathematical Society and DIMACS.

[16] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Trans. on Inform. Theory*, vol. 44, no. 3, pp. 1057-1070, May 1998.

[17] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1549-1559, Sept. 1997.

[18] J. Chen, X. 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. Select. Areas Commun.*, vol. 22, pp. 977-987, Aug. 2004.

[19] Y. Oohama, "Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Trans. on Inform Theory*, submitted for publication.

[20] V. Prabhakaran, D. Tse and K. Ramchandran, "Rate region of the quadratic Gaussian CEO problem," *Proc. International Symposium on Information Theory*, June 27-July 2, 2004, Chicago, USA, pp. 119.

[21] J. Chen and T. Berger, "Successive Wyner-Ziv coding scheme and its application to the quadratic gaussian CEO problem," *IEEE Trans. on Inform Theory*, submitted for publication.