Conference Paper

On the Gaussian Listening-Helper Source-Coding Problem

Author(s):
Bross, Shraga I.; Permuter, Haim H.

Publication Date:
2012

Permanent Link:
https://doi.org/10.3929/ethz-a-007071619

Rights / License:
In Copyright - Non-Commercial Use Permitted
On the Gaussian Listening-Helper Source-Coding Problem

Shraga I. Bross∗, and Haim Permuter†

∗ Engineering Department, Bar-Ilan University, Ramat Gan, 52900, Israel
Email: brosss@biu.ac.il
† Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva
Email: haimp@bgu.ac.il

Abstract— In the Gaussian listening-helper source-coding problem Encoder 1 and Decoder 1 observe respectively a pair of i.i.d. correlated Gaussian sources and both wish to communicate the first source to Decoder 2 subject to a distortion constraint. Encoder 1 sends a message to both decoders and then Decoder 1 – the listening-helper – sends a message just to Decoder 2.

We derive an inner bound on the rate region of this problem by proposing a Gaussian achievable scheme. This inner bound coincides with the solution to the Gaussian optimization problem associated with a single-letter outer bound corresponding to this problem.

Index Terms—multi-terminal source coding, listening-helper problem.

I. INTRODUCTION

The Gaussian listening-helper source-coding problem that we consider is depicted in Fig. 1. Encoder 1 observes a scalar i.i.d. Gaussian source X that is correlated with a scalar i.i.d. source Y observed by Decoder 1. Encoder 1 wishes to communicate the source X to within certain distortion constraint to Decoder 2 by sending a message at rate R to both decoders. Decoder 1 (the listening-helper) is allowed to communicate to Decoder 2 by sending a message at rate R1. Decoder 2 uses both messages to form its estimate of the source X. The goal is to determine the set of all rate pairs (R, R1) that satisfy the distortion constraint.

We derive an inner bound on the rate region for this source-coding problem by proposing a Gaussian scheme in which Encoder 1 vector quantizes its observation using a Gaussian test channel and then compresses the quantized values using proper binning. Decoder 1 quantizes its observation using another test channel and finally, both Encoder 1 and Decoder 1 cooperate to provide Decoder 2 with the enumeration information it needs in order to resolve its ambiguity regarding the compressed information Encoder 1 had generated. Furthermore, we show that the Gaussian solution associated with an upper bound for our problem coincides with the proposed Gaussian achievable scheme.

The paper is organized as follows. In Section II we provide a formal definition of the Gaussian listening-helper source-coding model. In Section III we present our main result, while Section IV is devoted to the description of the main steps of the proof.

The work of S. Bross was supported by the Israel Science Foundation under Grant 49709.

II. PROBLEM FORMULATION

The core problem we consider, the listening-helper source-coding problem, is described as follows. Let \{(X_i, Y_i)\}_{i=1}^n be a pair of zero-mean Gaussian random vectors generated by i.i.d. drawings of a pair of jointly Gaussian random variables with variance \(\sigma_X^2\) and \(\sigma_Y^2\), respectively. With no loss of generality, we can write

\[ Y_i = a_i X_i + N_i^{(y)}, \]

where \(N_i^{(y)} \sim N(0, \sigma_Y^2)\) is independent of \(X_i\). Encoder 1 which observes \(X^n\) sends a message to both decoders using an encoding function

\[ f^{(n)}: X^n \to \{1, 2, \ldots, M^{(n)}\}, \]

where \(X^n\) denotes the set of \(n\)-vectors over \(X\). Based on \(f^{(n)}(X^n)\) and its observation \(Y^n\) Decoder 1 sends a message to Decoder 2 using an encoding function

\[ f_1^{(n)}: \{1, 2, \ldots, M^{(n)}\} \times Y^n \to \{1, 2, \ldots, M_1^{(n)}\}. \]

Decoder 2 then uses both received messages to estimate \(X^n\) using a reproduction function

\[ g^{(n)}: \{1, 2, \ldots, M_1^{(n)}\} \times \{1, 2, \ldots, M_1^{(n)}\} \to \hat{X}_2^n. \]

Definition 1: A rate-distortion triple \((R, R_1, D_2)\) is achievable for the Gaussian listening-helper source-coding problem if there exists a block length \(n\), encoders \(f^{(n)}\) and \(f_1^{(n)}\), and a decoder \(g^{(n)}\) such that,

\[ \begin{align*}
R &\geq \frac{1}{n} \log M^{(n)} \\
R_1 &\geq \frac{1}{n} \log M_1^{(n)} \\
D_2 &\geq \frac{1}{n} \sum_{i=1}^n E \left( (X_i - \hat{X}_2,i)^2 \right).
\end{align*} \]

Let \(\hat{R}(D_2)\) be the closure of the set of achievable rate-distortion triples. The rate region \(\hat{R}(D_2)\) for the Gaussian listening-helper source-coding problem is defined by

\[ \mathcal{R}(D_2) = \left\{ (R, R_1): (R, R_1, D_2) \in \hat{R}(D_2) \right\}. \]
Since we are interested in the Mean-Square-Error (MSE) distortion constraint we can restrict the reproduction function to be the MMSE estimate of $X^n$ given $f^{(n)}(X^n)$ and $f^{(n)}_1(f^{(n)}(X^n), Y^n)$, that is
\[
X^n = \mathbb{E}\left[ X^n | f^{(n)}(X^n), f^{(n)}_1(f^{(n)}(X^n), Y^n) \right].
\]

III. MAIN RESULTS

Our main result is a characterization of an inner bound on the rate-region for the Gaussian listening-helper source-coding problem.

**Proposition 1:** For a positive number $D_2 \leq \sigma_X^2$, the rate-region $\mathcal{R}(D_2)$ contains the rate region $\mathcal{R}_G(D_2)$ defined by
\[
\mathcal{R}_G(D_2) = \bigcup (R, R_1) : \frac{1}{2} \log \sigma^2_{Y_v} \leq R \leq \frac{1}{2} \log \left[ 1 + \sigma^2_{Y_v} \left( \frac{\sigma^2_{Y_v}}{\sigma^2_{W_v}} + \sigma^2_{W_v} \right) \right],
\]
\[
R_1 \geq \frac{1}{2} \log \sigma^2_{Y_v} \left( \frac{\sigma^2_{Y_v}}{\sigma^2_{W_v}} + \sigma^2_{W_v} \right),
\]
\[
R + R_1 \geq \frac{1}{2} \log \frac{\sigma^2_{X} \left( \frac{\sigma^2_{Y_v}}{\sigma^2_{W_v}} + \sigma^2_{W_v} \right)}{\sigma^2_{W_v}} + \frac{\sigma^2_{W_v}}{\sigma^2_{W_v}}
\]
\[
D_2 = \frac{1 + \sigma^2_{W_v} (\frac{1}{\sigma^2_{W_v}} + \frac{\sigma^2_{W_v}}{\sigma^2_{W_v}})}{\sigma^2_{W_v}}.
\]

A. A Gaussian achievable scheme

A natural Gaussian achievable scheme for the listening-helper problem is described as follows. Let $S(U, V)$ be the set of zero-mean jointly Gaussian random variables $U$ and $V$ such that

1. $U \circlearrowright X \circlearrowright Y$ and $X \circlearrowright U Y \circlearrowright V$ are Markov chains.
2. $\sigma^2_{U|U, V} \leq D_2$.

For any $(U, V) \in S(U, V)$ and a large block-length $n$:

- Generate $e^{n\|X(U)\|+\epsilon}$, $\epsilon > 0$ independent codewords $u(1), \ldots, u(e^{n\|X(U)\|+\epsilon})$ randomly according to $p_U$. Assign these codewords uniformly into $e^{nR_1}$ bins. Label them $u(j, l), j \in \{1, \ldots, e^{nR_1}\}, l \in \{1, \ldots, e^{nR_2}\}$.
- Share the codebook and its partition between Encoder 1 and both decoders.
- For each $u(j, l)$ generate $e^{n[R_1+\epsilon+\epsilon]}$, $\epsilon > 0$ independent codewords randomly according to $p_{U|U}$. Label the codewords by $v(j, l, k), k \in \{1, \ldots, e^{n[R_2+\epsilon+\epsilon]}\}$ and share this codebook between Decoder 1 and Decoder 2.

Given a source and a side-information sequence pair $(x, y)$:
1. Encoder 1 looks for the codeword $u(j, l)$ that is jointly typical with $x$, and sends its bin index $l$ to both decoders.
2. Encoder 1 and Decoder 1 send together the enumeration index $j$ to Decoder 2.
3. Decoder 1 looks for the codeword $v(j, l, k)$ that is jointly typical with $(u(j, l), y)$, and sends the index $k$ of that codeword to Decoder 2.

Decoder 1 can recover $u$ with high probability provided that
\[
R \geq I(X; U|Y).
\]

Decoder 1 can comute $v$ with high probability provided that
\[
R_1 \geq I(Y; V|U),
\]
and taking into account that Encoder 1 and Decoder 1 join efforts in sending the index $j$ to Decoder 2 we obtain
\[
R + R_1 \geq I(X; U) + I(Y; V|U).
\]

Decoder 2 computes the MMSE estimate of the source sequence $X^n$ given $(U^n, V^n)$ and condition 2 above guarantees that the estimate meets the desired distortion. The following lemma characterizes an achievable rate-region based on this scheme.

**Lemma 1:** A Gaussian achievable scheme as above which satisfies the long Markov chain $U \circlearrowright X \circlearrowright Y \circlearrowright V$ achieves $\mathcal{R}_G(D_2)$. Moreover, the Gaussian solution associated with an upper bound for our problem coincides with $\mathcal{R}_G(D_2)$.

In Section IV.B we show that for Gaussian $(U, V) \in S(U, V)$ such that $U \circlearrowright X \circlearrowright Y \circlearrowright V$ the rate region associated with (7)-(9) equals $\mathcal{R}_G(D_2)$.

IV. PROOFS

The source-coding problem considered next will be used to show that the Gaussian solution associated with an upper bound for our main problem equals $\mathcal{R}_G(D_2)$.

A. A source-coding sub-problem

Let $\{(X_i, Y_i, Z_i)\}_{i=1}^n$ be a sequence of independent drawings of a triplet of zero-mean jointly Gaussian random variables with covariance $K_{(X,Y,Z)}$. An encoder observes $Y^n$ while both the encoder and the decoder have access to the side-information $Z^n$. The encoder sends a message to the decoder using an encoding function
\[
f^{(n)}_E : Y^n \times Z^n \rightarrow \{1, 2, \ldots, M^{(n)}\}.
\]
Based on $f^{(n)}_E(Y^n, Z^n)$ and its side-information $Z^n$ the decoder forms an estimate for $X^n$ using a reproduction function
\[
g^{(n)}_D : \{1, 2, \ldots, M^{(n)}\} \times Z^n \rightarrow X^n.
\]
This system is shown in Fig. 2.

Definition 2: A rate-distortion pair \((R_E, D)\) is achievable for the Gaussian source-coding sub-problem if there exists a block length \(n\), an encoder \(f_E^n\) and a decoder \(\hat{g}_D^n\) such that,

\[
R_E \geq \frac{1}{n} \log M(n) \quad \text{and} \quad D \geq \frac{1}{n} \sum_{i=1}^{n} \mathbb{E} [(X_i - \hat{X}_i)^2]. \tag{12}
\]

Here also since we are interested in the MSE distortion constraint we can restrict the reproduction function to be the MMSE estimate of \(X^n\) given \(f_E^n(Y^n, Z^n)\) and \(Z^n\), that is

\[
\hat{X}_i = \mathbb{E} \left[ X_i | f_E^n(Y^n, Z^n), Z^n \right].
\]

For a given distortion \(D > 0\), define \(P_0(D)\) as the set of laws \(p_{X^nY^nZ^n}(x, y, z, v)\), \(x, y, z, v \in \mathbb{R}\), such that the marginal law \(f_{X^nY^nZ^n}dv\) identifies with \(p_{X^nY^nZ^n}\), and

\[
\mathbb{E} \left[ (X - \hat{X}(X|V, Z))^2 \right] \leq D.
\]

Then, by a modification of [4, sec. 6.1.1], we have that

\[
R_E \geq R_{Y|Z}(D) = \min_{P_{X^nY^nZ^n} \in P_0(D)} I(Y; V | Z). \tag{13}
\]

Consequently, define an optimization problem \(P\) as follows

\[
\max \; h(Y|V) \quad \text{subject to} \quad D \geq \sigma^2_{Y|VZ}, \tag{14}
\]

then let \(P_G[D, (X, Y, Z), \sigma^2_{Y|VZ}]\) be the corresponding optimal problem assuming \(V\) is Gaussian and denote by \(\varphi\left( P_G[D, (X, Y, Z), \sigma^2_{Y|VZ}] \right)\) the conditional variance \(\sigma^2_{Y|VZ}\) associated with the optimal value of \(P_G\).

Consequently, \(P_G\) may be expressed as

\[
R_{Y|Z}(D) = \min_{p_{X^nY^nZ^n} \in P_0(D)} \frac{1}{2} \log \frac{\sigma^2_{Y|VZ}}{\sigma^2_{Y|VZ}}. \tag{15}
\]

On the other hand, when \(Z^n\) is available just at the decoder, - i.e. \(f_E^n = f_E^n(Y^n)\), by a modification of [2], we have that

\[
\hat{R}_E \geq R_{Y|Z}^{WZ}(D) = \min I(Y; V | Z), \tag{16}
\]

where the minimization is over the set of laws \(P(D)\) defined by

\[
p_{X^nY^nZ^n}(x, y, z, v) = p_{X^nY^nZ^n}(x, y, z)p_G(v|y). \tag{17}
\]

Consequently,

\[
R_{Y|Z}^{WZ}(D) = \min_{p_{X^nY^nZ^n} \in P(D)} \frac{1}{2} \log \frac{\sigma^2_{Y|Z}}{\sigma^2_{Y|VZ}} \quad \text{subject to} \quad D \geq \sigma^2_{Y|VZ},
\]

where the minimization is over the set of laws \(P(D)\) defined by

\[
p_{X^nY^nZ^n}(x, y, z, v) = p_{X^nY^nZ^n}(x, y, z)p_G(v|y).
\]

In [3, Section 3] the author shows that, when the encoder observes the source sequence \(X^n\) - i.e. \(Y = X\), both optimization problems (15) and (18) have the same solution and, in fact, the additional Markov relation \(Z \circ X \circ V^n\) in (18) characterizes the optimal solution of (15). The following lemma asserts that the same is true when the encoder observes a correlated random variable \(Y\) which satisfies a Markov relation.

Lemma 2: For jointly Gaussian random variables \((X, Y, Z)\) satisfying \(Z \circ X \circ Y\), both optimization problems (15) and (18) have the same solution and, in fact, the additional Markov relation in (18) characterizes the optimal solution of (15).

Proof: Omitted.

B. The Gaussian solution associated with an outer bound

Lemma 3: (Single-letter outer bound) If the rate-distortion triple \((R, R_1, D_2)\) is achievable then there exist random variables \((U, V)\) satisfying the Markov relation \(U \circ X \circ Y\), such that

\[
R + R_1 \geq I(X; U) + I(Y; V|U)
\]

\[
R \geq I(X; U|Y), \quad R_1 \geq I(Y; V|U)
\]

\[
D_2 \geq \sigma^2_{Z|UV}. \tag{19}
\]

Proof: Assume that \((R, R_1, D_2)\) is achievable. Let \(T = f^{\circ n}(X^n)\) and \(T_1 = f_1^{\circ n}(f^{\circ n}(X^n), Y^n)\), then the rate \(R\) can be lower bounded as follows

\[
nR \geq \log M(n) \geq H(T) \geq H(T|Y^n)
\]

\[
= I(X^n; T|Y^n) = \sum_{k=1}^{n} I(X_k; T|Y^n, X^{k-1})
\]

\[
\geq \sum_{k=1}^{n} I(X_k; T|Y^n, X^{k-1}, Y_{k-1})
\]

\[
\geq \sum_{k=1}^{n} I(X_k; U_k|Y_k), \tag{20}
\]

where

\[
X^{k-1} \overset{\Delta}{=} (X_1, \ldots, X_{k-1}), \quad U_k \overset{\Delta}{=} (T, Y^{k-1}), \quad Y^n \overset{\Delta}{=} (Y_1, \ldots, Y_{k-1}, Y_{k+1}, \ldots, Y_n),
\]

and the equality (a) follows by the Markov relation \(X_k \circ Y_k \circ Y^n\). Let \(V_k \overset{\Delta}{=} T_1\), then the rate \(R_1\) can be lower bounded as follows

\[
nR_1 \geq \log M_1^{\circ n} \geq H(T_1) \geq H(T_1|T)
\]

\[
= I(Y^n; T_1|T) = \sum_{k=1}^{n} I(Y_k; T_1|T, Y^{k-1})
\]

\[
= \sum_{k=1}^{n} I(Y_k; V_k|U_k). \tag{21}
\]
The sum-rate can be lower bounded as follows

\[
\begin{align*}
n(R + R_1) & \geq \log M^{(n)} + \log M_1^{(n)} \\
& \geq H(T) + H(T_1) \geq H(T; T_1) \\
& = H(T) + H(T_1|T) \\
& = H(T|Y^n) + I(Y; T_1) + H(T_1|T) \\
& = H(T|Y^n) + H(T_1|T) + \sum_{k=1}^{n} I(Y_k; T|Y^{k-1}) \\
& \overset{(b)}{\geq} \sum_{k=1}^{n} [I(X_k; U_k|Y_k) + I(Y_k; V_k|U_k)] \\
& = \sum_{k=1}^{n} [I(X_k; U_k) + I(Y_k; V_k|U_k)] \\
& \overset{(c)}{\geq} \sum_{k=1}^{n} [I(X_k; U_k) + I(Y_k; V_k|U_k)] \\
& \overset{(22)}{=} \sum_{k=1}^{n} [I(X_k; U_k) + I(Y_k; V_k|U_k)].
\end{align*}
\]

Here

(b) follows since \( Y^n \) is a memoryless sequence; and

(c) follows since \( Y_k \perp X_k \perp U_k \) is a Markov chain.

By (2) and the memoryless property of the sequence \((X_k, Y_k), k = 1, \ldots, n\) one can verify the Markov relation \( U_k \perp X_k \perp Y_k \) which implies the Markov relation \( U \perp X \perp Y \).

The combination of (20), (21) and (22) together with the latter Markov relation establish the desired single-letter lower bound.

Define next an optimization problem \( P_G \) over the conditional variance \( \sigma_{X|U|Y}^2 \) and \( V \) as follows

\[
\begin{align*}
\min_{V \sigma_{X|U|Y}^2, \sigma_{X|U|Y}^2} \quad & \frac{1}{2} \log \frac{\sigma_X^2}{\sigma_{X|U|Y}^2} - \frac{\sigma_{X|U|Y}^2}{\sigma_{X|U|Y|V}^2} \\
\text{subject to} \quad & R \geq \frac{1}{2} \log \frac{\sigma_{X|U|Y}^2}{\sigma_{X|U|Y|V}^2}, \quad R_1 \geq \frac{1}{2} \log \frac{\sigma_{X|U|Y}^2}{\sigma_{Y|U|V}^2} \\
& D_2 \geq \sigma_{X|U|V}^2 \quad \text{and} \quad U^G \perp X \perp Y.
\end{align*}
\]

\( P_G \) is the restriction of the outer bound (19) to a Gaussian law and on account of the Markov relation \( U^G \perp X \perp Y \) and (1) it may be expressed as

\[
\begin{align*}
\min_{V \sigma_{X|U|Y}^2, \sigma_{X|U|Y}^2} \quad & \frac{1}{2} \log \frac{\sigma_X^2}{\sigma_{X|U|Y}^2} - \frac{a_0^2 \sigma_{X|U|Y}^2 + a_V^2}{\sigma_{Y|U|V}^2} \\
\text{subject to} \quad & R \geq \frac{1}{2} \log \frac{\sigma_{X|U|Y}^2}{\sigma_{X|U|Y|V}^2}, \quad R_1 \geq \frac{1}{2} \log \frac{a_0^2 \sigma_{X|U|Y}^2 + a_V^2}{\sigma_{Y|U|V}^2} \\
& D_2 \geq \sigma_{X|U|V}^2 \quad \text{and} \quad U^G \perp X \perp Y.
\end{align*}
\]

Lemma 4: In the optimization problem \( P_G \) the conditional variance \( \sigma_{X|U|Y}^2 \) is monotonically increasing with \( \sigma_{X|U|Y}^2 \).

\textbf{Proof:} By the Markov relation \( U^G \perp X \perp Y \) and (1) we may assume that

\[
U^G = X + N^{(u)},
\]

where \( N^{(u)} \sim N(0, \sigma_{W_x}^2) \) is independent of \( X \). Then a straightforward computation yields the relation

\[
\sigma_{X|U|Y}^2 = \sigma_{W_x}^2 \left[ \frac{\sigma_{W_x}^2 + a_y^2}{\sigma_{X|U}^2} \right]^{-1}.
\]

We may express \( P_G \) as a double optimization problem

\[
\min_{\sigma_{X|U|Y}^2, \sigma_{X|U|Y}^2} \quad & \frac{1}{2} \log \frac{\sigma_X^2}{\sigma_{X|U|Y}^2} \left( a_y^2 + \frac{\sigma_{W_x}^2}{\sigma_{X|U}^2} \right) \\
\text{subject to} \quad & R \geq \frac{1}{2} \log \frac{\sigma_{X|U|Y}^2}{\sigma_{X|U|Y|V}^2} + \frac{\sigma_{W_x}^2}{\sigma_{X|U}^2} \\
& R_1 \geq \frac{1}{2} \log \frac{a_0^2 \sigma_{X|U|Y}^2 + \sigma_{W_x}^2}{\sigma_{Y|U|V}^2} \\
& D_2 \geq \sigma_{X|U|V}^2 \quad \text{and} \quad U^G \perp X \perp Y.
\]

Let \( (X, Y, U^G) \) be jointly Gaussian as per (1) and (23) then the choice of \( V^G = Y + N^{(v)} \) where \( N^{(v)} \sim N(0, \sigma_{W_x}^2) \) is independent of \( (X, N^{(u)}), N^{(u)} \) establishes that the optimal solution to \( P_G \) achieves \( \mathcal{R}_G(D_2) \).

\textbf{REFERENCES}