Title
Broadcasting Gaussian Sources Over Gaussian Channels

Permalink
https://escholarship.org/uc/item/7pz502c8

Author
Gao, Yang

Publication Date
2012-01-01

Peer reviewed|Thesis/dissertation
Broadcasting Gaussian Sources Over Gaussian Channels

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

Yang Gao

June 2012

Dissertation Committee:

Professor Ertem Tuncel, Chairperson
Professor Ilya Dumer
Professor Yingbo Hua
The Dissertation of Yang Gao is approved:

________________________

________________________

________________________

Committee Chairperson

University of California, Riverside
Acknowledgments

I would like to express my gratitude to my advisor, Professor Ertem Tuncel, for leading me into the charming world of information theory. Besides inspiring me with interesting problems, he has also been working closely with me and providing insightful guidance all the time. He sets up a role model for me, in both academics and life.

I am honored and grateful to have Professor Ilya Dumer and Professor Yingbo Hua in my committee.

I would also like to thank Dr. Deniz Gündüz, Dr. Emina Soljanin and Dr. Tian Chao for enlightening discussions.

Last but not least, I want to thank my dear wife Xuemei, who has always been supporting me and made this journey so enjoyable.
To Xuemei.
ABSTRACT OF THE DISSERTATION

Broadcasting Gaussian Sources Over Gaussian Channels

by

Yang Gao

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, June 2012
Professor Ertem Tuncel, Chairperson

This dissertation studies several problems regarding the joint/separate source-channel coding in data transmission over networks. It has been known that separation based transmission schemes achieves optimality when there is only one sender and one receiver. However, other than this point-to-point case, separation is often suboptimal and it is generally unclear how to achieve optimality.

In Chapter 2, several new hybrid digital/analog schemes are proposed for the problem of transmitting a Gaussian source over a Gaussian channel. Previously existing optimal schemes are combined in a way that the output of one encoder is taken as effective channel state information at the other encoder. In each case, the optimum distortion-power tradeoff can be achieved by a continuum of auxiliary random variables for any power allocation between the two encoders.

In Chapter 3, a new scheme is proposed for lossy transmission of a Gaussian source over a Gaussian broadcast channel with source side information available at each receiver. The proposed scheme combines two schemes that were previously shown to achieve optimal point-to-point distortion/power tradeoff simultaneously at all receivers under two distinct conditions stated in terms of channel and side information quality pa-
rameters. For the two-receiver case, the combined scheme is shown to achieve the same kind of optimality for the entire region in the parameter space sandwiched between those two conditions. Crucial to this result is the new degree of freedom discovered in Chapter 2.

In Chapter 4, the problem of broadcasting a pair of correlated Gaussian sources using optimal separate source and channel codes is studied. A universal upper bound of rate (bandwidth compression/expansion ratio) penalty is given and in particular, for a low-distortion scenario, separate coding can be shown to be optimal. For the case with a Gaussian broadcast channel, the power loss of separate coding is also discussed. Although source-channel separation yields suboptimal performance in general, it is shown that the proposed scheme provides competitive performance comparing with best-known schemes.
Contents

List of Figures x

1 Introduction 1
  1.1 Background and Related Work .......................... 1

2 A New Freedom for Point-to-point Transmission 4
  2.1 Proposed Schemes ........................................ 6
    2.1.1 Scheme 1 ............................................ 6
    2.1.2 Scheme 2 ............................................. 10
    2.1.3 Scheme 3 and 4 ....................................... 12
    2.1.4 Discussion ........................................... 16
  2.2 An Application ........................................... 19
  2.3 Two-letter Vector Generalization ......................... 23
    2.3.1 Problem Description and Preliminary Results ........... 23
    2.3.2 Generalization for ρ ≠ 0 ............................. 25
  2.4 Conclusion ................................................ 29

3 Wyner-Ziv Coding over Broadcast Channels: Hybrid Digital/Analog Schemes 31
  3.1 Introduction .............................................. 31
  3.2 Background and notation .................................. 33
  3.3 Results .................................................. 39
    3.3.1 A New Point-to-Point Scheme ......................... 40
    3.3.2 A Basic Scheme: HDA-CDS ............................. 43
    3.3.3 The Main Scheme ...................................... 47
  3.4 Conclusion ............................................... 52
  3.5 Appendix ................................................ 53
    3.5.1 Proof of Theorem 2 .................................. 53
    3.5.2 Proof of Theorem 4 .................................. 54
    3.5.3 Proof of Theorem 6 .................................. 61
    3.5.4 Proof of Lemma 7 .................................... 62
    3.5.5 Proof of Lemma 9 .................................... 63

4 Separate Source-Channel Coding for Broadcasting Gaussian Sources 66
  4.1 Introduction .............................................. 66
  4.2 Separate source-channel coding .......................... 68

viii
## List of Figures

2.1 Proposed scheme 1. A scaled version of the source $X^n$ itself serves as effective CSI. The channel output $V^n$ and the auxiliary codeword $T^n$ jointly serve as effective RSI, and are utilized to reduce the source coding rate as well as to increase the estimation performance. .......................... 7

2.2 Proposed scheme 2. A scaled version of the quantization error $E^n$ serves as effective CSI. The channel output $V^n$ and the auxiliary codeword $T^n$ jointly serve as effective RSI, and are utilized to increase the estimation performance. .......................................................... 7

2.3 Scheme 3 ...................................................... 12

2.4 For $P = 1$ and $\sigma_{W0}^2 = 0.1$, optimal parameters of Scheme 1 are indicated on the $(P_d, \gamma)$-plane for various $(\sigma_{W1}^2, \sigma_N^2)$. The solid lines correspond to $\gamma = \gamma_1$ and $\gamma = \gamma_2$ and the dashed line indicates $\gamma = \gamma_{\text{Costa}}$. ................................. 21

3.1 Illustration of the codebooks of a successful CDS transmission. The cross-hatched codewords are the actually used ones and the double sided arrows denote joint typicality, whereas the arrows with a cross denote atypical pairs. At the receiver side, all the hatched codewords in the source codebook $Z^n$ together form a virtual bin. ......................................................... 38

3.2 Illustration of HDA-WZ. The codebook of auxiliary HDA codeword $T^n$ is shown at both the encoder and the receiver. ................................. 38

3.3 A new hybrid scheme for point-to-point transmission with SSI, featuring the freedom of power allocation between the branches and, for any power allocation, the freedom to choose the auxiliary random codeword from a range. ......................................................... 40

3.4 HDA-CDS for one receiver achieves the optimum distortion without binning and has the same freedom of power allocation and construction of auxiliary random variable. ......................................................... 43

3.5 For WZBC problem, HDA-CDS combines two separately optimal schemes and achieves the trivial outer bound in the whole region between the conditions under which either scheme achieves optimum. ......................................................... 43
3.6 For any given $W_1 > W_2$, each point on the $(N_1, N_2)$-plane corresponds to a system with parameters $(W_1, W_2, N_1, N_2)$. As indicated by the two lines in the figure, the conditions for achieving the trivial outer bound for CDS and the HDA-WZ scheme are $W_1N_1 = W_2N_2$ and $N_1(P + W_1) = N_2(P + W_2)$, respectively. When $N_1 = N_2 = 1$, or the SSI is trivial at each receiver, analog transmission achieves the outer bound, as indicated by the dot. The shaded region, as stated in Theorem 4, indicates where our scheme is achieving the outer bound.

3.7 The encoder of the AHC scheme. An extra analog stream is superimposed with dirty paper coding to incorporate the benefits of analog transmission. The structure of decoders is the same as HDA-CDS.

3.8 Performance comparison between AHC scheme, LDS, analog transmission, and separate coding when $P = 1$. In (a), $N_2 > \frac{W_1}{W_2}N_1$. In (b), $\frac{P+W_1}{P+W_2}N_1 < N_2 < \frac{W_1}{W_2}N_1$, and AHC scheme achieves the trivial outer bound. In (c) and (d), $N_1 < N_2 < \frac{P+W_1}{P+W_2}N_1$. In (e), there is no side information and analog transmission is optimal. In (f), $N_2 < N_1$.

4.1 System model.

4.2 Regions of the $(D_1, D_2)$ plane in [32].

4.3 Feasible ranges of $\nu$ in regions of the $(D_1, D_2)$ plane. The dashed line indicates different regimes of the genie-aided outer bound, as will be seen later.

4.4 Genie-aided system with a degraded channel.

4.5 Source coding rate regions of various schemes.

4.6 The channel-independent rate penalty upper bound $\theta$ over the distortion plane. The color-value mapping is made non-linear intentionally to emphasize the values close to 1.

4.7 The rate penalty $\kappa_{sep}$ over the distortion plane of BSBC.

4.8 The rate penalty $\kappa_{genie}$ over the distortion plane for Gaussian broadcast channel.

4.9 The upper bounds of the power difference in dB. Bound 1 in the figure is the first bound in Theorem 23 and Bound 2 is the second.

4.10 Comparison between the power of genie-aided outer bound, Scheme C and optimal separate coding, $\rho = 0.8$, $N_1 = 1$, $N_2 = 0.5$ and $\kappa = 0.3$.

4.11 Comparison between the genie-aided outer bound, RFZ scheme in [1] and separate coding. $P = 3dB = 1.995$, $N_1 = 0dB = 1$, $N_2 = -5dB = 0.3162$ and $\kappa = 2$.

4.12 Comparison of power between the genie-aided outer bound, RFZ scheme in [1] and separate coding. $P = 3dB = 1.995$, $N_1 = 0dB = 1$, $N_2 = -5dB = 0.3162$ and $\kappa = 2$. The negative regions are shown with white dashed lines.
Chapter 1

Introduction

1.1 Background and Related Work

In his landmark paper [21], Shannon proved that the separation of source and channel coding is optimal for point-to-point data transmission, which is usually referred to as separation theorem. In formal terms, when a memoryless stationary source generating \( W_s \) symbols per second needs to be transmitted over a memoryless stationary channel with \( W_c \) channel uses per second, the minimum achievable distortion \( D \) is given by

\[
D = R^{-1}(\lambda C),
\]

where \( R(D) \) is the rate-distortion function, \( C \) is channel capacity and the rate (bandwidth compression/expansion ratio) \( \lambda = \frac{W_c}{W_s} \).

However, when a network scenario is considered, separation based coding schemes yield suboptimal performance in general. The quest for optimal joint source-channel coding schemes starts with an interesting observation for transmitting a Gaussian source over an additive Gaussian white noise channel. In [12], it was shown that by simply scaling the source and sending uncoded, the same optimal distortion can also be achieved...
without need of coding. The above analog/uncoded scheme is even better than separation based schemes in terms of complexity and delay. The scale-and-send scheme can be extended to broadcasting a single Gaussian source to multiple receivers, but other than that, the simple scheme itself possesses limited ability.

Though this dissertation is mostly concerned with the power-distortion performance of joint or separate schemes, it has to be pointed out that joint source-channel coding schemes are also favored for other reasons, such as graceful performance degradation, robustness, low delay etc.

In [4], a superposition scheme combining the separate coding and analog transmission is shown to achieve optimum distortion with any power allocation between the two streams for transmitting a Gaussian source over a bandwidth-matched Gaussian channel. We generalized the set of optimal schemes in [9]. For bandwidth mismatched cases, a more general scheme was proposed in [27]. Several hybrid digital/analog schemes are proposed for the point-to-point scenario with channel state information or receiver side information in [30].

The scenario with a broadcast or MAC channel is more involved and requires advanced techniques, especially for mismatched bandwidth. In [19], inner and outer distortion bounds are given for broadcasting a single Gaussian source with bandwidth expansion. A variation with receiver side information at each decoder is discussed in [28, 17, 13]. In [11], we proposed a hybrid digital/analog scheme, which, under certain conditions, makes each of the two receivers achieve its optimum distortion simultaneously.

For the case in which two correlated Gaussian sources need to be transmitted, several source coding schemes are analyzed in [22], and the optimal source coding scheme is given in [16]. When the two sources need to be transmitted over a matched Gaussian
broadcast channel and each receiver is only interested in recovering one of the sources, an outer bound is proposed in [22] and various schemes are introduced in [5, 22, 2, 1]. The problem with matched bandwidth is completely solved in [26] with a hybrid scheme. For general mismatched bandwidth, the optimal joint scheme is yet to be found, and we analyzed the performance (loss) of separate coding in [10].

In [25], Tian et al. proved that for a large group of network transmission problems, separate coding is actually optimal, whereas for another large group it is approximately optimal. This result provides insights on whether separate coding is optimal for a specific problem, but it does not indicate how to construct the optimal separate or joint schemes.
Chapter 2

A New Freedom for

Point-to-point Transmission

In point-to-point transmission, minimum distortion can be achieved by using separate source and channel codes, as shown by Shannon in his celebrated work [21]. However, it is well-known that for a Gaussian source transmitted over an additive white Gaussian noise (AWGN) channel with an input power constraint, uncoded transmission, i.e., a simple analog scale-and-transmit scheme, also achieves the minimum mean squared distortion. Recently, Bross et al. [4] proposed a scheme in which the channel-coded quantized source and the uncoded quantization error are superimposed and transmitted. The scheme can achieve the optimum distortion performance for any source quantization rate smaller than the capacity of the AWGN channel by appropriately allocating power between the digital and analog information. Tackling the problem of seamless digital upgrade of analog transmission systems, Puri et al. [18] discussed another hybrid scheme in which a digital upgrade path is added to the old uncoded path and proved its optimality. In the digital path, a Wyner-Ziv encoder is concatenated with a dirty-paper encoder and
the scaled source and the total channel output serve as side information at the encoder and the decoder side, respectively. Schemes of both [4] and [18] cover as special cases i) separation-based digital transmission, and ii) purely analog transmission.

In the presence of source side information (SSI) at the receiver and/or channel state information (CSI) at the transmitter, although optimality of analog transmission breaks down, separate source and channel coding remains optimal [15], [20]. Recently, Wilson et al. [30] proposed several hybrid digital/analog (HDA) schemes that are also optimal in scenarios with SSI and/or CSI. These novel schemes integrate the analog source variable in the auxiliary codeword of a dirty-paper coder [7].

In this chapter, we propose several new schemes which also transmit superimposed coded and uncoded information with or without SSI or CSI, but we fully explore the potential of the dirty-paper coding auxiliary random variable. Specifically, we show that by making better use of the (already) decoded auxiliary codeword, optimality can be achieved by a continuum of auxiliary random variables, even though the effective channel capacity is compromised (even diminished to zero) when a non-Costa variable is chosen. Our schemes specialize to those in [4] and [18] when the auxiliary random variable is appropriately chosen. In the special case of no actual CSI or RSI, the main hybrid digital/analog (HDA) scheme in [30] also becomes an instance of our infinite family of schemes. Unlike in [27], where bandwidth-mismatched transmission of a Gaussian source over a Gaussian channel is considered, we only consider the case with matched bandwidth. The connections and differences between our schemes and those in previous work are further discussed in Sec. 2.1.4.

Although the main goal of this chapter is simply to reveal a previously unnoticed level of freedom in point-to-point transmission, we also briefly look at a multiterminal scenario to demonstrate how this new freedom can be utilized. The result is also extended to a
two letter vector source and the corresponding vector broadcast channel in Sec. 2.3.

2.1 Proposed Schemes

Consider a memoryless zero-mean Gaussian source and an AWGN channel. We assume the source has unit variance without loss of generality. The source emits a source sequence \{X_i\} in an i.i.d. fashion and \(X_i \sim \mathcal{N}(0,1)\). The source is quantized with an optimal vector quantizer, and the quantization can be characterized with a “test channel” \(X = E + Z\), where \(Z\) is the quantized source and \(E\) is the quantization error. For an optimal quantizer, \(E \sim \mathcal{N}(0, \sigma_E^2)\) and \(Z\) is independent of \(E\). The AWGN channel has an input power constraint \(\frac{1}{n} \sum_{i=1}^{n} U_i^2 \leq P\) and the additive noise is i.i.d. Gaussian random variables \(W_i\) drawn from \(\mathcal{N}(0, \sigma_W^2)\).

The optimal distortion can be achieved by two simple schemes. The separation-based approach is to quantize the source to a rate \(R(D) = \frac{1}{2} \log \frac{1}{D} = \frac{1}{2} \log \frac{1}{\sigma_E^2}\) slightly smaller than the channel capacity \(C = \frac{1}{2} \log \frac{P + \sigma_W^2}{\sigma_W^2}\), thereby achieving

\[
D = \frac{\sigma_W^2}{P + \sigma_W^2}.
\]

Using an uncoded transmission approach, we can also scale the source by \(\sqrt{P}\) and send through the channel. After proper estimation, the same optimal distortion can be achieved.

2.1.1 Scheme 1

Consider the system in Fig. 2.1. The channel power is split into two parts \(P_d\) and \(P_u\) for coded and uncoded transmission. The uncoded path just sends a properly scaled version of the source. The coded path sends the source coding index, with dirty-
paper coding. As in [7], the dirty-paper auxiliary codeword $T^n$ is constructed by $T = \sqrt{T_a} \gamma X + U$ with independent $X$ and $U$, and the effective state-dependent channel is $V^n = \sqrt{T_a} X^n + U^n + W^n$. In a successful transmission, the dirty-paper channel decoder can decode $T^n$.

Since both $T^n$ and $V^n$ are available at the decoder, one can use them as side information in Wyner-Ziv coding of the source. Noting that by construction we have the Markov chain $(T, V) \rightarrow X \rightarrow Z$, the minimum source coding rate becomes

\[
I (X; Z | T, V) \\
= I (X; Z) - I (T, V; Z) \\
= h (X) - h (X | Z) - h (T, V) + h (T, V | Z) \\
= h (X) - h (E) - h (T, V) \\
+ h \left( \sqrt{P_a} E + U, \sqrt{P_a} E + U + W \right). \tag{2.1}
\]

If we do not restrict ourselves to the optimal choice of $\gamma$ for the channel capacity, which
is given by Costa in [7] in the form of \( \gamma_{\text{Costa}} = \frac{P_d}{P_d + \sigma_W^2} \), we have the capacity of the new channel for coded transmission as

\[
I(T; V) - I(T; X) = h(V) - h(T, V) + h(U). \tag{2.2}
\]

To guarantee the coded transmission is successful, we need

\[
I(X; Z|T, V) \leq I(T; V) - I(T; X), \tag{2.3}
\]

Expanding (2.3) using (2.1) and (2.2), we obtain

\[
h(X) - h(E) + h\left(\sqrt{P_a \gamma E + U}, \sqrt{P_a E + U + W}\right)
\leq h(V) + h(U)
\]

which becomes

\[
\frac{\det\left(\begin{array}{cc}
 P_a \gamma \sigma_E^2 + P_d & P_a \gamma \sigma_E^2 + P_d \\
 P_a \gamma \sigma_E^2 + P_d & P_a \sigma_E^2 + P_d + \sigma_W^2
\end{array}\right)}{\sigma_E^2} \leq P_d(P + \sigma_W^2)
\]

and finally reduces to

\[
\sigma_E^2 \geq \frac{P_d \sigma_W^2}{P_d(P + \sigma_W^2) - P_a P_d(1 - \gamma)^2 - P_a \sigma_W^2 \gamma^2}. \tag{2.4}
\]

In order to ensure that the capacity in (2.2) is non-negative, or

\[
\frac{1}{2} \log \det\left(\begin{array}{cc}
 P_a \gamma^2 + P_d & P_a \gamma + P_d \\
 P_a \gamma + P_d & P + \sigma_W^2
\end{array}\right) \geq 0,
\]

\( \gamma \) must be confined in a range \( \gamma_1 \leq \gamma \leq \gamma_2 \), where

\[
\gamma_{1,2} = \gamma_{\text{Costa}} \left(1 \pm \sqrt{\frac{P + \sigma_W^2}{P_a}}\right) \tag{2.5}
\]

It is worth noting that in the limit \( P_a \to 0 \), i.e., when the analog path (and hence the CSI) vanishes, the range of \( \gamma \) becomes the entire real line. This is expected, as
the value of $\gamma$ is immaterial in the purely digital regime. At the other extreme, when $P_d \to 0$, it can be shown that $\sigma^2_E \to 1$. To see that, notice that the right-hand side of (2.4) is minimized by $\gamma = \gamma_{\text{Costa}}$ for any $P_d$ and the resultant expression after setting $\gamma = \gamma_{\text{Costa}}$ approaches 1 when $P_d \to 0$. This is also natural, because the source (and channel) coding rate of the digital path vanishes in the purely analog regime.

The reconstruction of $X^n$ is $\hat{X}^n = Z^n + \hat{E}^n$, where $\hat{E}^n$ is the optimal estimate of $E^n$ when $Z^n$, $V^n$, and $T^n$ are known:

$$\hat{E} = k_1 \left( \sqrt{P_a} E + U + W \right) + k_2 \left( \gamma \sqrt{P_a} E \right).$$

If we use an optimal MMSE estimator, $k = [k_1 \ k_2]^T$ is given by the solution of

$$\begin{bmatrix}
P_a \sigma^2_E + P_d + \sigma^2_W & \gamma P_a \sigma^2_E + P_d \\
\gamma P_a \sigma^2_E + P_d & \gamma^2 P_a \sigma^2_E + P_d
\end{bmatrix} k = \begin{bmatrix}
\sqrt{P_a} \sigma^2_E \\
\gamma \sqrt{P_a} \sigma^2_E
\end{bmatrix},$$

which can be computed as

$$k = \alpha \begin{bmatrix}
(1 - \gamma) P_d \\
(\gamma P_d + \sigma^2_W) \gamma - P_d
\end{bmatrix},$$

where

$$\alpha = \frac{\sqrt{P_a} \sigma^2_E}{\sigma^2_E [P_a P_d (1 - \gamma)^2 + P_a \sigma^2_W \gamma^2 + P_d \sigma^2_W]}.$$ 

The square-error distortion then has the form

$$D = E \left\{ (X - \hat{X})^2 \right\} = E \left\{ (E - \hat{E})^2 \right\} = E \left\{ (E - \hat{E}) E \right\} = \sigma^2_E \left( 1 - k_1 \sqrt{P_a} - k_2 \gamma \sqrt{P_a} \right) = \frac{P_d \sigma^2_W}{P_a P_d (1 - \gamma)^2 + P_a \sigma^2_W \gamma^2 + P_d \sigma^2_W \sigma^2_E}.$$
To minimize the distortion, we should obviously choose equality in (2.4), which implies matching source and channel coding rates and

\[ D = \frac{\sigma_W^2}{P + \sigma_W^2}. \]

This means we can achieve the optimal distortion for any power allocation between coded and uncoded paths as in [4] and [18]. Moreover, for any power allocation, we can achieve optimality for a continuous range \( \gamma_1 \leq \gamma \leq \gamma_2 \) rather than with only \( \gamma = \gamma_{\text{Costa}} \). In other words, for any given power allocation, the compromise from the effective channel capacity is fully compensated by the use of the decoded dirty-paper auxiliary codeword, both in estimation and in Wyner-Ziv coding.

### 2.1.2 Scheme 2

As shown in Fig. 2.2, this scheme uses (both as the uncoded component and as the CSI) the quantization error instead of the source itself, where \( c \) is the scaling factor satisfying \( P_a = c^2 \sigma_E^2 \). Thus, \( T^n = c\gamma E^n + U^n \) with underlying \( E \) and \( U \) independent of each other, and \( V^n = cE^n + U^n + W^n \). Since \( Z \) is independent of \( E \), it is easy to conclude that neither \( T^n \) nor \( V^n \) are useful as side information at the receiver to decode \( Z^n \). Therefore Wyner-Ziv coding is not needed. We follow a similar procedure as in Scheme 1 and also consider \( \gamma \) other than \( \gamma_{\text{Costa}} \). To make the coded transmission successful, we need

\[ I(X; Z) \leq I(T; V) - I(cE; T). \]  

Expanding (2.6), we obtain

\[
\frac{1}{2} \log \frac{1}{\sigma_E^2} \leq h(V) + h(U) - h(T, V) = \frac{1}{2} \log \frac{P_t \left(P + \sigma_W^2\right)}{\det(A)}.
\]  

(2.7)
where

\[
A = \begin{bmatrix}
\gamma^2 P_a + P_d & \gamma P_a + P_d \\
\gamma P_a + P_d & P + \sigma_W^2
\end{bmatrix}.
\]

Inequality (2.7) can be re-written as

\[
\sigma^2_E \geq \frac{P_a P_d (1-\gamma)^2 + P_a \sigma_W^2 \gamma^2 + P_d \sigma_W^2}{P_d (P + \sigma_W^2)}. \tag{2.8}
\]

As in Scheme 1, our choice of \(\gamma\) must ensure that the right hand side of (2.6) is non-negative. It can easily be shown that this requirement translates to the same condition \(\gamma_1 \leq \gamma \leq \gamma_2\), where \(\gamma_1\) and \(\gamma_2\) are given as in (2.5). Also, the behavior of Scheme 2 at the two extremes \(P_a \to 0\) and \(P_d \to 0\) is the same as in Scheme 1.

Optimal reconstruction of \(E^n\) is \(\hat{E}^n = k_1 V^n + k_2 T^n\), where \(k = [k_1\ k_2]^T\) satisfies

\[
\begin{bmatrix}
P + \sigma_W^2 & \gamma P_a + P_d \\
\gamma P_a + P_d & \gamma^2 P_a + P_d
\end{bmatrix} \begin{bmatrix}
c \sigma^2_E \\
c \gamma \sigma^2_E
\end{bmatrix} = \begin{bmatrix}
(1-\gamma) P_d \\
(P_d + \sigma_W^2) \gamma - P_d
\end{bmatrix}.
\]

and thus can be found as

\[
k = \frac{c \sigma^2_E}{\det(A)} \begin{bmatrix}
(1-\gamma) P_d \\
(P_d + \sigma_W^2) \gamma - P_d
\end{bmatrix}.
\]

The corresponding minimum square-error distortion is given by

\[
D = E \left\{ (E - \hat{E})^2 \right\} \\
= \sigma^2_E (1 - k_1 c - k_2 \gamma c) \\
= \sigma^2_E \frac{P_d \sigma_W^2}{\det(A)}.
\]

It is again apparent that in order to minimize \(D\), one should choose \(\sigma^2_E\) so as to satisfy equality in (2.8), which results in matched source and channel coding rates and

\[
D = \frac{\sigma^2_W}{P + \sigma_W^2}.
\]
The conclusion is therefore exactly the same as in Scheme 1: For any power allocation, there is a range $\gamma_1 \leq \gamma \leq \gamma_2$ achieving optimality using Scheme 2, and the range is the same as that of Scheme 1.

### 2.1.3 Scheme 3 and 4

With the presence of SSI and/or CSI, analog transmission is no longer an optimal scheme, while separate source and channel coding maintains optimality [15], [20]. Of course, ordinary source and channel coding must be replaced by Wyner-Ziv coding [31] and dirty paper coding [7] when SSI and CSI are present, respectively.

The SSI, if present, is modeled as $Y^n = \rho X^n + N^n$, where $0 \leq \rho < 1$ and $N^n$ is a zero-mean i.i.d. Gaussian vector with variance $\sigma^2_N = 1 - \rho^2$. The CSI $S^n$, if present, is also modeled as zero-mean i.i.d. Gaussian, and affects the channel through the relation $V^n = U^n + W^n + S^n$. Both $N^n$ and $S^n$ are independent of $X^n$, and are also independent of each other. The optimal squared-error distortion with SSI is

$$D_{\text{opt}} = \frac{\sigma^2_N \sigma^2_W}{P + \sigma^2_W},$$

regardless of whether CSI is present or not.

In [30], a new hybrid scheme (HDA) is proposed by embedding the analog source into the auxiliary codeword through the relation $T_h = \kappa X + U + \gamma S$. The encoder finds
a $T^n_h$ in the codebook jointly typical with $X^n$ and $S^n$, and then sends the corresponding $U^n = T^n_h - \kappa X^n - \gamma S^n$ as the channel word. The decoder finds a $T^n_h$ jointly typical with the channel output $V^n$ and side information $Y^n$. It is shown that by choosing

$$\kappa^2 = \frac{P^2}{\sigma_N^2(P + \sigma_W^2)}$$

and

$$\gamma = \frac{P}{P + \sigma_W^2},$$

the effective source coding rate $I(T_h; X, S)$ matches the effective channel capacity $I(T_h; V, Y)$ and the optimal distortion is achieved.

We now propose to combine the HDA scheme in [30] and separate coding. The combination is done the same way as in Scheme 1 and 2, and we shall soon see that a similar new freedom is observed.

The first proposed scheme, termed Scheme 3, is shown in Fig. 2.3. The total channel input power is split into $P_h$ and $P_d$ for hybrid and digital paths, respectively. The HDA auxiliary codeword is constructed by $T^n_h = U^n_h + \kappa E^n + \gamma_h S^n$ and the HDA channel codeword $U^n_h$ is fed to the digital encoder as additional CSI. The dirty-paper auxiliary channel codeword is given by $T^n_d = \gamma_d S^n + \gamma U^n_h + U^n_d$ and the channel output $V^n = U^n_h + U^n_d + W^n$. We choose $\gamma_h$ as

$$\gamma_h = \frac{P_h}{P + \sigma_W^2},$$

and $\gamma_d$ as

$$\gamma_d = \frac{P_h \gamma + P_d}{P + \sigma_W^2}.$$  

(2.10)

At the receiver side, first the digital decoder operates. To ensure that it successfully decodes $Z^n$ (as well as $T^n_d$), we need

$$I(X; Z|Y) \leq I(T_d; V) - I(T_d; S, U_h),$$

(2.11)
where $I(X; Z|Y)$ is the Wyner-Ziv source coding rate and $I(T_d; V) - I(T_d; S, U_h)$ is the effective digital channel capacity. Similarly with Scheme 1, neither $V^n$ and $T^n_d$ is useful in decoding $Z^n$. The only useful side information for that purpose is $Y^n$.

With the choice of $\gamma_h$ and $\gamma_d$ as in (2.9) and (2.10), respectively, (2.11) reduces to

$$\frac{1}{\sigma^2_E} \leq 1 - \frac{1}{\sigma^2_N} + \frac{P_d(P + \sigma^2_W)}{\sigma^2_N(P_h\alpha + \sigma^2_W P_d)} ,$$  

(2.12)

where $\alpha = P_d(1 - \gamma)^2 + \sigma^2_W \gamma^2$. A condition $\gamma_1 \leq \gamma \leq \gamma_2$ is also necessary to ensure that the effective capacity in (2.11) is non-negative, where

$$\gamma_{1,2} = \gamma_{\text{Costa}} \left( 1 \pm \sqrt{\frac{P + \sigma^2_W}{P_h}} \right) .$$  

(2.13)

Comparing (2.13) with (2.5), we observe that the range of feasible $\gamma$ is exactly the same as that in Scheme 1, with $P_a$ playing the role of $P_h$. This also implies that when $P_h \to 0$, the range of $\gamma$ becomes the entire real line and the system turns purely digital. On the other hand, when $P_d \to 0$, $\sigma^2_E \to 1$ and the scheme becomes a pure HDA one.

The HDA decoder has access to $T^n_d$, $V^n$, $Y^n$, and $Z^n$. However, it is not difficult to see that $Z^n$ can be eliminated by defining $\tilde{Y}^n = Y^n - \rho Z^n = \rho E^n + N^n$ as the side information for the source $E^n$. To decode $T^n_h$ successfully, we then need

$$I(T_h; S, E) \leq I(T_h; T_d, V, \tilde{Y}) .$$  

(2.14)

Note that we again make use of the dirty-paper codeword instead of discarding it. Inequality (2.14) boils down to

$$\kappa^2 \leq \frac{P_h^2 \sigma^2_E \alpha}{\sigma^2_E \sigma^2_N(P_h \alpha + P_d \sigma^2_W)} ,$$  

(2.15)

where $\sigma^2_\tilde{Y}$ is the variance of $\tilde{Y}$. It is worth noting that in both (2.12) and (2.15), $\sigma^2_S$ is ultimately canceled in the expression.
The optimal estimate of $X$ is given by the optimal estimate of $E$, which is obtained by

$$
\hat{E}_n = k_1 T_d^n + k_2 T_h^n + k_3 V^n + k_4 \tilde{Y}^n,
$$

where $k = [k_1 \ k_2 \ k_3 \ k_4]^T$ satisfies

$$
A k = b,
$$

in which $A$ is the covariance matrix of $[T_d \ T_h \ V \ \tilde{Y}]^T$ as in (2.16), and $b$ is the correlation between $E$ and $[T_d \ T_h \ V \ \tilde{Y}]^T$:

$$
b = [0 \ \kappa \sigma^2_E \ \rho \sigma^2_E]^T.
$$

The square-error distortion is

$$
D = E \{ (E - \hat{E})^2 \}
= \sigma^2_E - k^T b.
$$

$$
\geq \frac{\sigma^2_E \sigma^2_N \sigma^2_W P_d P_h}{\sigma^2_E \sigma^2_N (P_h \alpha + P_d \sigma^2_W) + \sigma^2_N \sigma^2_W \sigma^2_Y},
$$

where (a) and (b) are satisfied by equality when equalities hold in (2.15) and (2.12), respectively. This implies that for any power allocation between the two streams, proposed Scheme 3 achieves the optimal distortion in the presence of both SSI and CSI.

When $\sigma^2_S = 0$, the CSI is trivial, and so are $\gamma_h$ and $\gamma_d$. Scheme 3 reduces to a scheme with only SSI, and achieves optimal distortion for any feasible $\gamma$ under any
power allocation. When the SSI is trivial, or $\sigma_N^2 = 1$, there is only CSI and Scheme 3 achieves optimal distortion with the same continuum of auxiliary random variables if $\gamma_h$ and $\gamma_d$ satisfy (2.9) and (2.10).

We also propose a scheme, termed Scheme 4, where $X^n$ is input to the HDA encoder instead of $E^n$. The HDA auxiliary codebook is now constructed according to $T_h = U_h + \kappa X + \gamma_h S$ and the digital dirty-paper auxiliary codeword is constructed the same way as in Scheme 3, where $\gamma_h$ and $\gamma_d$ are still chosen according to (2.9) and (2.10), respectively. Unlike Scheme 2, $T_d$ and $V$ do not contain any information about $Z$ and thus are not helpful in source decoding. Therefore, the condition for successful digital transmission is still (2.11). To successfully decode $T_h^n$, the counterpart of (2.14) is now

$$I(T_h;X,S) \leq I(T_h;T_d,V,Y,Z),$$  \hspace{1cm} (2.17)

and it reduces to (2.15) as well. Note since the construction of $T_h$ contains $X$, the right hand side of (2.17) can not be simplified to $I(T_h;T_d,V,\tilde{Y})$ as in Scheme 3. The optimal estimate of $X$ is given by a linear combination of $T_h$, $T_d$, $V$, $Y$, and $Z$. Once again, optimal distortion is achieved when equalities hold in both (2.12) and (2.15), meaning we have the same result in both Schemes 3 and 4.

2.1.4 Discussion

We should point out that although we have a range of feasible $\gamma$, it does not mean we can choose $\gamma$ freely when the quantizer is fixed. Since we need equality in (2.4), (2.8) or (2.12), for any power allocation in each scheme, the quantization error $\sigma_E^2$ (or the source coding rate) has a one-to-one relation with $\gamma$, and the minimum $\sigma_E^2$ (or the maximum rate) is achieved when $\gamma = \gamma_{Costa}$. The moral behind our result is that the maximum effective channel capacity is not necessary to achieve optimality in the hybrid
system, though the source coding rate and compromised effective channel capacity still need to match.

Now we look at some special cases of the proposed schemes.

When \( \gamma = \gamma_{\text{Costa}} \), Scheme 1 reduces to the scheme in [18], in which only \( V \) is used as side information and in estimation. This is because when \( \gamma = \gamma_{\text{Costa}} \) we have

\[
I(T;X|V) = h(T|V) - h(T|V, X)
\]

\[= \quad h(T|V) - h(U|U + W) \quad \text{(a)}
\]

\[= \quad h(T - \gamma V|V) - h(U|U + W) \quad \text{(b)}
\]

\[= \quad h(U|U + W) - h(U|U + W)
\]

\[= \quad 0,
\]

where (a) follows since \( T = \gamma \sqrt{P_a} X + U, \ V = \sqrt{P_a} X + U + W \), and \( X, U, \) and \( W \) are pairwise independent, and (b) follows from the same arguments in [6, Sec. II-D].\(^1\)

The Markov chain \( X - V - T \) implies that \( T \) is useless in estimating \( X \) once we have \( V \) (estimation coefficient \( k_2 = 0 \)). Also, since \( Z \) is a quantized version of \( X \), the Markov chain can be extended to \( Z - X - V - T \), and thus \( I(X;Z|T,V) = I(X;Z|V) \), implying that the source coding rate cannot be reduced by using \( T \) either.

In Scheme 2, we can prove the Markov chain \( E - V - T \) when \( \gamma = \gamma_{\text{Costa}} \) following a similar procedure, which implies \( T \) is not helpful in estimation in that case.

When \( \gamma = 1 \), i.e., when the dirty paper codeword \( T^n \) exactly coincides with what is transmitted through the channel due to perfect cancellation of the CSI, the Markov chains \( Z - X - T - V \) and \( E - T - V \) are obvious in Schemes 1 and 2, respectively. \( V \) is then a corrupted version of \( T \) and thus useless in either scheme, as evidenced by \( k_1 = 0 \).

\(^1\)The arguments are that when \( \gamma = \gamma_{\text{Costa}} \), \( U + W \) and \( T - \gamma V \) are uncorrelated, and thus independent, and that \( (T - \gamma V) - (U + W) - V \) forms a Markov chain.
However, $\gamma = 1$ must be in the range $\gamma_1 \leq \gamma \leq \gamma_2$, where $\gamma_{1,2}$ are given in (2.5). Since $\gamma_1 \leq 0$, it suffices to have $\gamma_2 \geq 1$, or equivalently $\frac{P_a}{P} \geq \Delta$. When $\gamma = 0$, both schemes reduce to simple superpositions without dirty paper coding. Specifically, Scheme 2 reduces to the scheme in [4]. In Scheme 1, $T = U$ and one may initially think that $T$ is not useful in either source coding or estimation because $U$ is chosen to be independent of $X$ and $Z$. However, this is not the case because

$$I(X; Z|T, V) = I(X; Z|U, \sqrt{P_a}X + U + W) = I(X; Z|\sqrt{P_a}X + W) \neq I(X; Z|V),$$

i.e., $T$ does provide information in the presence of $V$.

Finally, if $\gamma \to \gamma_1$ or $\gamma \to \gamma_2$, in both schemes the effective channel capacity approaches zero and the quantization error $\sigma_E^2$ approaches 1, and thus the first two schemes become the same. In this case, in effect, there remain no apparent source and channel coders. However, $T = \gamma \sqrt{P_a}X + U$ or $T = \gamma cE + U$ play the role of forward test channels for source coding, and thus the zero-rate dirty paper codebook essentially becomes a source codebook. The effective quantization error $U^n$ is superimposed with $\gamma \sqrt{P_a}X^n$ or $\gamma cE^n$ and transmitted. The power $P_d$ for the coded transmission is still used even though no digital information is transmitted in the traditional sense. The dirty paper channel encoder and decoder serve as a whole transmission system including a quantizer.

An interesting special case arises when we further let $P_a \to 0$. It can be shown that $\gamma_1 \sqrt{P_a} \to -\kappa$ and $\gamma_2 \sqrt{P_a} \to \kappa$ as $P_a \to 0$, where

$$\kappa = \frac{P}{\sqrt{P + \sigma_W^2}}.$$

This fact, together with the observation that the channel input degenerates into just
$U^n$, implies that our schemes reduce to the main scheme in [30] with trivial CSI and RSI$^2$. 

In [30], several hybrid schemes with actual CSI are proposed by an approach of embedding the analog variable into the dirty paper codeword, as just mentioned. The scheme in Sec. II-D of [30], with trivial CSI, gives a similar result as our Scheme 2 in that for any source coding rate lesser than the Gaussian channel capacity, there is an optimal auxiliary random variable achieving optimum distortion. However, in our Scheme 2, for a source coding rate, both power allocation and the auxiliary random variable can be changed (while maintaining equality in (2.4)) to achieve the optimum distortion, thus providing more freedom.

Another relevant work is [23], where a tradeoff between the digital information rate and the distortion of the CSI is studied. In our Scheme 2, one might think we are doing a similar job by having a coded transmission rate and maintaining the distortion of our artificial CSI. However, as mentioned earlier, our scheme does not have real CSI and it makes a big difference, especially when the split of the channel power is considered. Also, in our scheme, the variance of CSI changes with the message rate.

### 2.2 An Application

Let two nodes in a network communicate via a Gaussian channel with power constraint $P$ and a noise level of $\sigma^2_{W_0}$. The receiving node can recover the Gaussian source with a distortion of $\frac{\sigma^2_{W_0}}{P + \sigma^2_{W_0}}$. Consider the scenario where the channel noise level occasionally rises to $\sigma^2_{W_1}$. When (and only when) this happens, the receiving node gets help from a third node in the form of side information $Y^n$, where unit-variance Gaussian ran-

---

$^2$The scheme referred to is what both schemes in [30, Sec. II-B] and [30, Sec. III-B] reduce to with trivial CSI and RSI, respectively.
dom variable $Y = \rho X + N$ with $\rho \geq 0$, $X \perp N$, and $N \sim \mathcal{N}(0, \sigma_N^2)$. Alternatively, the side information can be expressed as $X = \rho Y + N'$, where $Y \perp N'$ and $N' \sim \mathcal{N}(0, \sigma_N^2)$.

We want to build a scheme that minimizes the resultant distortion in the regime of worse channel and helper side information while maintaining optimality in the original regime of better channel and no side information.

Uncoded transmission is definitely a legitimate choice, and it is well-known to achieve optimal distortion under all channel noise levels in the absence of helper side information ($\sigma_N^2 = 1$). However, in the presence of side information, its distortion becomes

$$D_a = \frac{\sigma_W^2 \sigma_N^2}{P \sigma_N^2 + \sigma_W^2}.$$ 

Unfortunately, a purely digital scheme, designed for $\sigma_W^2$, suffers from the threshold effect and cannot decode the channel codeword under noise level $\sigma_W^2$. The source has to be reconstructed using only the side information. A simple scheme superimposing uncoded and digital paths as in [4] has the same problem: The digital channel codeword cannot be decoded under $\sigma_W^2$ and thus it is meaningless to allocate any power to the digital path.

In a hybrid scheme with dirty-paper coding, $\gamma = \gamma_{\text{Costa}} = \frac{P_d}{P_d + \sigma_W^2}$ could have been a natural choice as well for any power allocation, however, as will be shown below, this choice suffers from the same threshold effect, i.e., $T^n$ cannot be decoded under $\sigma_W^2$.

On the other hand, by utilizing the proposed freedom in our schemes, i.e., considering all feasible $(P_d, \gamma)$, we find for each $(\sigma_W^2, \sigma_N^2)$ a $(P_d, \gamma)$ pair outperforming analog transmission, and of course, any purely digital scheme, and any hybrid scheme with $\gamma = \gamma_{\text{Costa}}$.

In our Scheme 1, to decode the auxiliary codeword when the noise level is $\sigma_W^2$, we
Figure 2.4: For $P = 1$ and $\sigma^2_{W_0} = 0.1$, optimal parameters of Scheme 1 are indicated on the $(P_d, \gamma)$-plane for various $(\sigma^2_{W_1}, \sigma^2_N)$. The solid lines correspond to $\gamma = \gamma_1$ and $\gamma = \gamma_2$ and the dashed line indicates $\gamma = \gamma_{\text{Costa}}$.

need

$$I(X; Z|T, V_0) \leq I(T; V_1, Y) - I(T; X), \quad (2.18)$$

where $V_0 = \sqrt{P_a}X + U + W_0$ and $V_1 = \sqrt{P_a}X + U + W_1$. Note here $Y^n$ is used as another channel output. Since (2.3) has to be satisfied with equality under $\sigma^2_{W_0}$, (2.18) is equivalent to

$$I(T; V_0) \leq I(T; V_1, Y). \quad (2.19)$$

Further, to find out the correct $Z^n$ in a source bin, we need

$$I(Z; T, V_0) \leq I(Z; T, V_1, Y). \quad (2.20)$$

In other words, the combined RSI $(T, V_1, Y)$ must be better than $(T, V_0)$.

Now, when $\gamma = \gamma_{\text{Costa}}$, $Y - X - V_0 - T$ forms a Markov chain, as mentioned earlier, and (2.19) can be written as

$$I(T; V_0, Y) \leq I(T; V_1, Y).$$

On the other hand, since $V_1$ is the output of the channel with the higher noise, we
naturally have

\[ I(T; V_0, Y) \geq I(T; V_1, Y). \]

Thus, successful decoding of \( T^n \) is possible if and only if \( I(T; V_0, Y) = I(T; V_1, Y) \), or equivalently \( I(T; V_0|Y) = I(T; V_1|Y) \), which implies that

\[ I(\gamma_{\text{Costa}} \sqrt{P_a N'} + U; \sqrt{P_a N'} + U + W) \]

remains the same whether \( \sigma^2_W = \sigma^2_{W_0} \) or \( \sigma^2_W = \sigma^2_{W_1} \). It can be easily shown that this happens only when \( P_d = 0 \) (making the mutual information zero). So for any meaningful power allocation, (2.19) is not satisfied and \( T^n \) can not be decoded when \( \gamma = \gamma_{\text{Costa}} \). Therefore, the resultant distortion is not better than that of analog transmission.

Similarly, for Scheme 2, to decode \( T^n \), we need

\[ I(X; Z) \leq I(T; V_1, Y) - I(T; cE), \] (2.21)

which, using (2.6), becomes (2.19) again. Thus, when \( \gamma = \gamma_{\text{Costa}} \), we can also show following similar arguments that (2.21) is satisfied (with equality) only when \( P_d = 0 \), which means \( \gamma = \gamma_{\text{Costa}} \) is not a feasible choice for Scheme 2 either.

Fig. 2.4 shows the numerically-obtained optimal \((P_d, \gamma)\) pairs in Scheme 1 for various \((\sigma^2_{W_1}, \sigma^2_N)\) when \( \sigma^2_{W_0} \) is fixed. As can be seen from the figure, the extra level of freedom introduced here helps reduce the distortion, because the optimal parameters are not those corresponding to analog transmission \((P_d = 0)\), digital transmission \((P_d = 1)\), or hybrid digital/analog coding with \( \gamma = \gamma_{\text{Costa}} \). The exception is, of course, when the quality of the side information is almost trivial, i.e., \( \sigma^2_N \to 1 \), in which case analog transmission is optimal. Following the same procedure, Scheme 2 is observed to have similar results.
2.3 Two-letter Vector Generalization

2.3.1 Problem Description and Preliminary Results

From Scheme 2, we know that for point-to-point transmission, we have the freedom of a range of $\gamma$ with any power allocation and each feasible $\gamma$ corresponds to a $\sigma_E^2$, given by equality of (2.8), which translates to

$$P_a(P_d + \sigma_W^2)\gamma^2 - 2P_aP_d\gamma + P_d(P_a + \sigma_W^2) - P_d(P + \sigma_W^2)\sigma_E^2 = 0. \tag{2.22}$$

From (2.22), we know that $\gamma = \gamma_{\text{Costa}}$ minimizes $\sigma_E^2$ to $\frac{\sigma_W^2}{P_d + \sigma_W^2}$. Denote the two roots of (2.22) with $\gamma_a$ and $\gamma_b$, and the explicit form is

$$\gamma_{a,b} = \frac{P_aP_d \pm \sqrt{P_aP_d(P + \sigma_W^2)[(P_d + \sigma_W^2)\sigma_E^2 - \sigma_W^2]}}{P_a(P_d + \sigma_W^2)}. \tag{2.23}$$

Let

$$\gamma_1 = \frac{\gamma_a + \gamma_b}{2} = \frac{P_d}{P_d + \sigma_W^2}$$

and

$$\gamma_2 = \frac{\gamma_a - \gamma_b}{2} = \pm \frac{\sqrt{P_aP_d(P + \sigma_W^2)[(P_d + \sigma_W^2)\sigma_E^2 - \sigma_W^2]}}{P_a(P_d + \sigma_W^2)}. \tag{2.23}$$

Define

$$f(\gamma) = P_aP_d(\gamma - 1)^2 + P_a\sigma_W^2 \gamma^2 + P_d\sigma_W^2,$$

and we know that

$$f(\gamma_a) = f(\gamma_b) = \sigma_E^2P_d(P + \sigma_W^2).$$

Now consider a two-letter vector Gaussian source

$$\tilde{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix},$$
which can be quantized with an optimal quantizer given by

\[ \tilde{X} = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} + \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}, \]

and \( E\{E_1E_2\} = 0 \). The dirty-paper channel codeword is constructed in a similar way as before, by

\[ \tilde{T} = \tilde{U} + c\Gamma \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}, \]

where

\[ \Gamma = \begin{bmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & \gamma_1 \end{bmatrix}. \]

For successful transmission, we need

\[ I(\tilde{X}; \tilde{Z}) \leq I(\tilde{T}; \tilde{V}) - I(c\tilde{E}; \tilde{T}), \]

which boils down to

\[ \sigma_E^4 \geq \frac{\det (A)}{P_d^2(P + \sigma_W^2)^2} \quad (2.24) \]

and \( A \) is the covariance matrix of \([T_1, T_2, V_1, V_2] \).

\[
A = \begin{bmatrix}
P_d + P_a(\gamma_1^2 + \gamma_2^2) & 2\gamma_1\gamma_2P_a & \gamma_1P_a + P_d & \gamma_2P_a \\
2\gamma_1\gamma_2P_a & P_d + P_a(\gamma_1^2 + \gamma_2^2) & \gamma_2P_a & \gamma_1P_a + P_d \\
\gamma_1P_a + P_d & \gamma_2P_a & P + \sigma_W^2 & 0 \\
\gamma_2P_a & \gamma_1P_a + P_d & 0 & P + \sigma_W^2 \\
\end{bmatrix}
\]

Note (2.24) is satisfied with equality since (2.22) and

\[
\det A
= \left[ P_aP_d(\gamma_a - 1)^2 + P_a\sigma_W^2\gamma_a^2 + P_d\sigma_W^2 \right] \left[ P_aP_d(\gamma_b - 1)^2 + P_a\sigma_W^2\gamma_b^2 + P_d\sigma_W^2 \right]
= f(\gamma_a)f(\gamma_b).
\]
The MMSE estimation of $E_1$ and $E_2$ is given by

$$\hat{E}_i = k_i^T [T_1, T_2, V_1, V_2]^T.$$ 

and the distortion is thus

$$D_{E_i} = \sigma^2_{E_i} - b_i^T A^{-1} b_i ,$$

where $b_1 = c \sigma^2_{E} [\gamma_1, \gamma_2, 1, 0]^T$ and $b_2 = c \sigma^2_{E} [\gamma_2, \gamma_1, 0, 1]^T$ .

The distortion can then be expanded to the following form

$$D_{E_i} = \sigma^2_{W} [f(\gamma_a) + f(\gamma_b)] .$$

From (2.24), we know $\det(A) = \sigma^4_{W} P_d (P + \sigma^2_{W})^2$, and also using the fact that $\gamma_{a,b}$ satisfy (2.22), we have

$$D_{E_1} = D_{E_2} = \frac{\sigma^2_{W}}{P + \sigma^2_{W}} = D_{\text{opt}} .$$

Remark 1 First since we never assume $\gamma_a > \gamma_b$, both possible $\gamma_2$s are valid. Second, from the expression of $\gamma_2$, we know that by changing $\sigma^2_{E}$ ($P_a$ and $P_d$ are still fixed), the absolute value of $\gamma_2$ increases monotonically with $\sigma^2_{E}$. When $\sigma^2_{E} = \frac{\sigma^2_{W}}{P_a + \sigma^2_{W}}$, $\gamma_b = 0$; whereas when $\sigma^2_{E} = 1$, $\gamma_2 = \pm \frac{P_a \sigma^2_{W}}{P_a + \sigma^2_{W}} \sqrt{\frac{P_a + \sigma^2_{W}}{P_a}}$.

2.3.2 Generalization for $\rho \neq 0$

Consider the optimal quantizer $\bar{X} = \bar{Z} + \bar{E}$ where $\bar{E}$ now has the covariance matrix

$$\begin{bmatrix}
\sigma^2_{E} & \rho \sigma^2_{E} \\
\rho \sigma^2_{E} & \sigma^2_{E}
\end{bmatrix} .$$

The condition for successful digital transmission is

$$I(\bar{X}; \bar{Z}) \leq I(\bar{T}; \bar{V}) - I(c \bar{E}; \bar{T}) ,$$

25
which now boils down to

\[(1 - \rho^2)\sigma_E^4 \geq \frac{\text{det} A}{P_d^2(P + \sigma_W^2)^2}\]  

(2.25)

and \(A\) is the covariance matrix of \([T_1, T_2, V_1, V_2]\), which now has the form of (the symmetric lower half is omitted)

\[
\begin{bmatrix}
Pd + Pa(\gamma_1^2 + \gamma_2^2) + 2\rho Pa\gamma_1\gamma_2 & \rho Pa(\gamma_1^2 + \gamma_2^2) + 2\gamma_1\gamma_2 Pa & (\gamma_1 + \rho\gamma_2) Pa + Pd & (\rho\gamma_1 + \gamma_2) Pa \\
Pd + Pa(\gamma_1^2 + \gamma_2^2) + 2\rho Pa\gamma_1\gamma_2 & (\gamma_1 + \rho\gamma_2) Pa & (\gamma_1 + \rho\gamma_2) Pa + Pd & Pa \\
P + \sigma_W^2 & \rho Pa & Pa & P + \sigma_W^2 \\
\end{bmatrix}
\]

The determinant can be written as \(\text{det}(A) = f_1f_2\), where

\[
f_1 = Pa(Pd + \sigma_W^2)(1 + \rho) \left[ \gamma_1^2 + \gamma_2^2 - 2\gamma_{\text{Costa}}\gamma_1 + \gamma_{\text{Costa}} + 2\gamma_2(\gamma_1 - \gamma_{\text{Costa}}) \right] + Pd\sigma_W^2
\]

and

\[
f_2 = Pa(Pd + \sigma_W^2)(1 - \rho) \left[ \gamma_1^2 + \gamma_2^2 - 2\gamma_{\text{Costa}}\gamma_1 + \gamma_{\text{Costa}} - 2\gamma_2(\gamma_1 - \gamma_{\text{Costa}}) \right] + Pd\sigma_W^2
\]

In light of the result of the case with \(\rho = 0\), we notice that if (2.25) is satisfied with equality, we need

\[
\text{det}(A) = [Pd(P + \sigma_W^2)(1 + \rho)\sigma_E^4] \left[ Pd(P + \sigma_W^2)(1 - \rho)\sigma_E^2 \right]
\]

and we will next prove \(f_1\) and \(f_2\) equal the corresponding part. Thus \(f \frac{\rho}{1-\rho} = f \frac{\rho}{1-\rho}\), which translates to

\[(\gamma_1 - \gamma_{\text{Costa}})\gamma_2 = \gamma_{\text{Costa}} \frac{\sigma_W^2}{Pa} \frac{\rho}{1 - \rho^2}; \]  

(2.26)

\[Pd(\gamma_1 - Pd)\gamma_2 = \left[ \frac{Pd}{Pa} \frac{\rho}{1 - \rho^2} - (\gamma_1 - Pd)\gamma_2 \right] \sigma_W^2. \]  

(2.27)

Now we look at the distortion

\[D_{S_i} = \sigma_E^2 - b_i^T A^{-1} b_i, \]
where \( b_1 = c\sigma_E^2[\gamma_1 + \rho\gamma_2, \rho\gamma_1 + \gamma_2, 1, \rho]^T \) and \( b_2 = c\sigma_E^2[\rho\gamma_1 + \gamma_2, \gamma_1 + \rho\gamma_2, \rho, 1]^T \). Then we have

\[
D_{S_1} = D_{S_2} = \frac{P_d\sigma_E^2\sigma_W^2(P_d + \sigma_W^2)}{\det A} \left[ (1 - \rho^2) \left( \gamma_1^2 + \gamma_2^2 - 2\gamma_{\text{Costa}}\gamma_1 + \gamma_{\text{Costa}} \right) + \gamma_{\text{Costa}}\frac{\sigma_W^2}{P_d} \right],
\]

where

\[
\phi = \frac{f_1 - P_d\sigma_W^2}{(1 + \rho)P_a(P_d + \sigma_W^2)} - 2(\gamma_1 - \gamma_{\text{Costa}})\gamma_2
\]

\[
= \frac{f_2 - P_d\sigma_W^2}{(1 - \rho)P_a(P_d + \sigma_W^2)} + 2(\gamma_1 - \gamma_{\text{Costa}})\gamma_2,
\]

Let

\[
\phi = \frac{1}{2} \left[ \frac{f_1 - P_d\sigma_W^2}{(1 + \rho)P_a(P_d + \sigma_W^2)} - 2(\gamma_1 - \gamma_{\text{Costa}})\gamma_2 \right] + \frac{1}{2} \left[ \frac{f_2 - P_d\sigma_W^2}{(1 - \rho)P_a(P_d + \sigma_W^2)} + 2(\gamma_1 - \gamma_{\text{Costa}})\gamma_2 \right]
\]

\[
= \frac{(1 - \rho)f_1 + (1 + \rho)f_2 - 2P_d\sigma_W^2}{2(1 - \rho^2)P_a(P_d + \sigma_W^2)}
\]

and we obtain

\[
D_{S_i} = \frac{P_d\sigma_W^2\sigma_E^2(1 - \rho)f_1 + (1 + \rho)f_2}{f_1f_2} = \frac{P_d\sigma_W^2\sigma_E^2(1 - \rho)f_1}{f_1f_2}
\]

\[
= \frac{P_d\sigma_W^2\sigma_E^2(1 - \rho)}{f_2} = \frac{P_d\sigma_W^2\sigma_E^2(1 - \rho)}{P_d(P + \sigma_W^2)(1 - \rho)\sigma_E^2}
\]

\[
= \frac{\sigma_W^2}{P + \sigma_W^2}.
\]

We are implicitly having a \( \sigma_E^2 \) as follows

\[
\sigma_E^2 = \frac{P_a}{(P + \sigma_W^2)\gamma_{\text{Costa}}} \left( \gamma_1^2 + \gamma_2^2 - 2\gamma_{\text{Costa}}\gamma_1 + \gamma_{\text{Costa}} \right) + \frac{\sigma_W^2}{P + \sigma_W^2} \frac{1}{1 - \rho^2}
\]

and for any given set of parameters \( P_a, P_d, \sigma_W^2, \rho, \sigma_E^2 \leq 1 \) translates to

\[
(\gamma_1 - \gamma_{\text{Costa}})^2 + \gamma_2^2 \leq \gamma_{\text{Costa}} \left[ \frac{P_d + \sigma_W^2}{P_a} + \gamma_{\text{Costa}} - \frac{\sigma_W^2}{P_a(1 - \rho^2)} \right].
\]
The minimal $\sigma_E^2$ occurs when $\gamma - \gamma_{\text{Costa}} = \gamma_2$

$$\sigma_{E\text{min}}^2 = \frac{\sigma_W^2}{P + \sigma_W^2} \left( \frac{1}{1 - \rho} + \frac{P_a}{P_d + \sigma_W^2} \right)$$

When $\sigma_{E\text{min}}^2 > 1$ there is no feasible solution, and to avoid this situation, we need

$$\rho < \frac{P_d(P + \sigma_W^2)}{P_aP_d + (P_d + \sigma_W^2)^2}.$$ 

However, this may not be the only upper bound of $\rho$. Since we have

$$\vec{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} + \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}$$

and $\sigma_X^2 = 1$, the covariance matrix can be written as

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \sigma_{Z_1}^2 & \rho_Z \sigma_{Z_1} \sigma_{Z_2} \\ \rho_Z \sigma_{Z_1} \sigma_{Z_2} & \sigma_{Z_2}^2 \end{bmatrix} + \begin{bmatrix} \sigma_{E_1}^2 & \rho_{E\sigma_{E_1} \sigma_{E_2}} \\ \rho_{E\sigma_{E_1} \sigma_{E_2}} & \sigma_{E_2}^2 \end{bmatrix}.$$ 

Thus we have

$$\rho_{E\sigma_{E_1} \sigma_{E_2}} = -\rho_Z \sigma_{Z_1} \sigma_{Z_2}$$

which translates to

$$\rho_Z = -\rho_{E\sigma_{E_1} \sigma_{E_2}} \sigma_{Z_1} \sigma_{Z_2}.$$

Note $|\rho_Z| \leq 1$ leads to

$$\rho_{Z}^2 \leq \frac{\sigma_{Z_1}^2 \sigma_{Z_2}^2}{\sigma_{E_1}^2 \sigma_{E_2}^2} = \frac{(1 - \sigma_{E_1}^2)(1 - \sigma_{E_2}^2)}{\sigma_{E_1}^2 \sigma_{E_2}^2} = \left( \frac{1}{\sigma_{E_1}^2} - 1 \right) \left( \frac{1}{\sigma_{E_2}^2} - 1 \right),$$

and when $\sigma_{E_1} = \sigma_{E_2}$, it reduces to

$$\rho \leq \frac{1}{\sigma_{E}^2} - 1,$$
which is

$$\sigma_E^2 = \frac{P_a}{(P + \sigma_W^2)\gamma_{\text{Costa}}}(\gamma_1^2 + \gamma_2^2 - 2\gamma_{\text{Costa}}\gamma_1 + \gamma_{\text{Costa}}) + \frac{\sigma_W^2}{P + \sigma_W^2} \frac{1}{1 - \rho^2} \leq \frac{1}{1 + \rho}.$$  

It finally boils down to

$$(\gamma_1 - \gamma_{\text{Costa}})^2 + \gamma_2^2 \leq \gamma_{\text{Costa}} \left( \frac{1}{P_a} (1 - \rho)(P + \sigma_W^2) - \frac{\sigma_W^2}{1 - \rho^2} + \gamma_{\text{Costa}} - 1 \right). \quad (2.29)$$

The RHS of (2.29) needs to be nonnegative, which translates to

$$P_a \sigma_W^2 \rho^2 - (P_d + \sigma_W^2)(P + \sigma_W^2)\rho + P_d(P + \sigma_W^2) \geq 0$$

Since the LHS of (2.29) is positive when $\rho = 0$ and it is negative when $\rho = 1$, $\rho$ has to be smaller than the small root, i.e.

$$\rho \leq \frac{(P_d + \sigma_W^2)(P + \sigma_W^2) - \sqrt{(P_d + \sigma_W^2)^2(P + \sigma_W^2)^2 - 4P_a\sigma_W^2(P + \sigma_W^2)^2}}{2P_a\sigma_W^2}. $$

Comparing (2.28) and (2.29), it is easy to show that the latter bound is tighter. Also notice for a fixed power allocation, when $\rho = 0$, the RHS of the bound reaches its maximum value, which, as expected, equals to (2.23) at $\sigma_E^2 = 1$.

However, this is still not the only bound. Assume we have $xy = c$ and $x^2 + y^2 = k$, then we need $k \geq 2c$ to make sure there indeed exist solutions, and that gives a new bound

$$\rho \leq \frac{P + 3\sigma_W^2 - \sqrt{(P + 3\sigma_W^2)^2 - 4P_a(1 - \gamma_{\text{Costa}})(P_a\gamma_{\text{Costa}} + P_d)}}{2P_a(1 - \gamma_{\text{Costa}})}. $$

2.4 Conclusion

For i.i.d. Gaussian sources and AWGN channels, it is known that various techniques (digital, analog, or hybrid) achieve the optimum distortion. We proposed two new schemes which provide a new degree of freedom, in addition to the classical freedom of
power allocation, in designing hybrid systems. It is interesting that even when we do not use the effective channel at full capacity, we can still achieve optimality by making full use of the decoded auxiliary codeword. With some special choices of the auxiliary dirty paper variable, our schemes reduce to either some known schemes or to simple, yet previously unknown, ones.

To demonstrate how this new freedom can be utilized, we also looked at a multiterminal scenario where the channel quality occasionally worsens, in which case the receiver gets help in the form of side information. We minimized the resultant distortion while maintaining optimality in the original setting. While pure digital coding and hybrid coding with a capacity-maximizing Costa parameter both fail to decode the digital information, thus performing worse than pure analog coding, our schemes, when optimized over their free parameters, outperform analog coding.

The freedom is also generalized to a two-letter vector case. It is shown that when the dimension of the problem grows, there exists larger freedom in constructing the auxiliary dirty-paper codeword.
Chapter 3

Wyner-Ziv Coding over

Broadcast Channels: Hybrid

Digital/Analog Schemes

3.1 Introduction

Consider a sensor network of $K + 1$ nodes taking periodic measurements of a common phenomenon. One node transmits its measurement to the other $K$ nodes over a broadcast channel and each of the $K$ nodes has source side information (SSI) only available to that node. The lossless version of this problem was studied in [28] and termed Slepian-Wolf coding over broadcast channels (SWBC). The more general lossy version was referred to as Wyner-Ziv coding over broadcast channels (WZBC) in [17], and performance of several purely digital schemes were analyzed. In this paper, we consider the bandwidth-matched quadratic Gaussian case of the WZBC problem with emphasis on $K = 2$ receivers. Even in this special case, there is no known scheme that is opti-
mal under all circumstances. However, there are several competitive schemes, some of which, under certain conditions, achieve a trivial outer bound: the minimum distortion that point-to-point transmission would achieve at each individual receiver for the given power level. Of course, achieving this outer bound immediately implies optimality.

The simplest such scheme is analog, i.e., uncoded, transmission of the source, in which the source is scaled to accommodate the power level of the channel. The condition under which the outer bound is achieved in this case is that the side information at each receiver be trivial, i.e., independent of the source.

At the other end of the spectrum, a fully digital joint source-channel coding algorithm, termed the common description scheme (CDS), was proposed in [17] for general sources, distortion measures, and bandwidth expansion factors. CDS is also a simple scheme based on losslessly broadcasting the quantized source by utilizing the binning-free joint decoding technique developed in [28] for the SWBC problem. It was shown in [17] that for the bandwidth-matched quadratic Gaussian case, CDS achieves the trivial outer bound when an appropriately defined “combined” channel and side information quality is constant among the receivers. In [17], CDS was also extended to a dirty-paper setting (termed DPC-CDS), where the channel state information (CSI) is available non-causally at the encoder, and to a layered description scheme (LDS), which was shown to outperform separate source and channel coding. Finally, based on the same techniques, [13] introduced several hybrid digital/analog (HDA) schemes which outperform analog transmission as well as separate coding.

In [14], for point-to-point Wyner-Ziv/dirty-paper coding, a scheme using modulo-lattice modulation was proved to be optimal for the bandwidth-matched quadratic Gaussian case. When the CSI is trivial and the scenario is extended to broadcast channels (and thus the problem becomes WZBC), the scheme is also shown to achieve the trivial
outer bound when another combined channel and side information quality, i.e., defined differently from [17], is constant among the receivers. Later on, [30, Sec. III-B] proposed a closely related scheme with random coding arguments instead of lattices, whereby the analog source is integrated into the auxiliary random codeword. Not surprisingly, when extended to broadcast channels, this scheme achieves the trivial outer bound under the same condition as that in [14], though it is not explicitly mentioned in [30]. In the sequel, we will refer to the scheme in [30] as the HDA-WZ scheme.

In this chapter, we first present a basic WZBC scheme combining the HDA-WZ scheme and DPC-CDS, which we term HDA-CDS. We prove that by making full use of the decoded auxiliary dirty-paper codeword, our scheme achieves the trivial outer bound when the channel and side information quality parameters fall between those yielding constant combined quality with respect to the HDA-WZ scheme and CDS. We also show that HDA-CDS outperforms LDS for any set of system parameters.

To take advantage of analog transmission, especially when the side information is weak, we added an analog stream onto HDA-CDS, also to be potentially used as artificial CSI for both HDA-WZ and CDS components. However, although the analog stream is very useful as a third signal component, we numerically observed that it is useless as CSI, though a rigorous proof seems extremely difficult.

3.2 Background and notation

Let \((X, Y_1, Y_2, \ldots, Y_K)\) be real-valued jointly Gaussian random variables generated in an i.i.d. fashion from \(P_{XY_1Y_2\ldots Y_K}\). The Gaussian source sequence \(X^n\) is to be transmitted over a Gaussian broadcast channel

\[
V_i^n = U^n + W_i^n, \quad i = 1, 2, \ldots, K
\]
where $U^n_i$, $V^n_i$, and $W^n_i$ are the channel input, channel output at receiver $i$, and the corresponding i.i.d. additive white Gaussian channel noise. The channel has an input power constraint
\[ \frac{1}{n} \sum_{j=1}^{n} \mathbb{E}(U_j^2) \leq P. \]
Source side information $Y^n_i$ is available at receiver $i$, where
\[ Y_i = \rho_i X + N_i \]
with $N_i \perp X$ and $\rho_i \in [0, 1]$. Without loss of generality, $X$ and $Y_i$ ($i = 1, 2, \ldots, K$) are assumed to have unit variance and thus the variance of $N_i$ is $N_i = 1 - \rho_i^2$. Hereafter, to ease exposition, a bold font capital letter will denote the variance of the corresponding random variable as in [17]. The reconstruction quality is measured with squared error distortion
\[ d_i(x^n, \hat{x}_i^n) = \frac{1}{n} \sum_{j=1}^{n} (x_j - \hat{x}_{ij})^2 \]
for any source block $x^n$ and reconstruction block $\hat{x}_i^n$.

In the special case of point-to-point transmission, or $K = 1$, the capacity of the Gaussian channel is\(^1\) \(^2\)
\[ C = \frac{1}{2} \log \left( 1 + \frac{P}{W} \right), \]
and the minimum distortion achieved by Wyner-Ziv coding with rate $R$ is given by [31]
\[ D = N 2^{-2R}. \]
Since separate source and channel coding is optimal, as proved in [20], transmission is possible if and only if $R \leq C$, which translates to
\[ D \geq \frac{NW}{P + W} \triangleq D_{\text{opt}}. \]

\(^1\)All logarithms in this chapter are base 2.
\(^2\)The subscripts for the receivers are omitted since there is only one receiver.
For the Gaussian WZBC problem, a trivial outer bound is obtained by letting each receiver achieve its minimum distortion without considering other receivers:

\[ D_i \geq \frac{N_i W_i}{P + W_i}. \]  \hfill (3.1)

There are several schemes for the WZBC problem achieving the trivial outer bound under different conditions, though a scheme optimal under all circumstances is not known.

With the bandwidth-matched case under consideration, the simplest scheme one can use is analog transmission of the source \( X^n \), in which the unit-variance source is scaled with \( \sqrt{P} \) to adapt to the power constraint of the channel, and it achieves the performance

\[ D_i = \frac{N_i W_i}{N_i P + W_i} \]

at each receiver. It is obvious that uncoded transmission achieves the outer bound (3.1) with equality if and only if \( N_i = 1 \), for all \( i = 1, 2, \ldots, K \). This, in turn, corresponds to the extreme situation that \( Y_i \) is independent of \( X \), and hence is of very limited use to us.

The performance of separate source and channel coding is known since both the channel and the side information are degraded. The explicit expressions of the distortion pair for \( K = 2 \) were given in [17, Lemma 3] and are included here for completeness. Without loss of generality, \( W_1 > W_2 \) is assumed. The distortion pair \((D_1, D_2)\) with

\[ \frac{W_1 N_i}{P + W_1} \leq D_1 \leq N_1 \]

is achievable using separate coding if and only if \( D_2 \geq D_{SEP}(D_1) \) where \( D_{SEP}(D_1) \) is the convex hull of

\[ D_{SEP}^2(D_1) = \frac{N_2 N_1^2 W_2 D_1}{[D_1 N_1 + N_2 (N_1 - D_1)] [(W_2 - W_1)N_1 + (P + W_1)D_1]} \]
when $N_1 \geq N_2$ and

$$D^*_S(D_1) = \frac{N_2}{(W_2 - W_1)N_1 + (P + W_1)D_1} \max \left\{ \frac{W_2D_1}{N_1 [N_2W_2 - (P + W_1)D_1 - N_1(W_2 - W_1)]}, \frac{N_2}{N_2 - N_1} \right\}$$

when $N_1 < N_2$.

Another purely digital scheme, the common description scheme (CDS), was proposed in [17], which can be utilized for general sources/channels and bandwidth expansion factors, but we focus here only on the bandwidth-matched case. As illustrated in Figure 3.1, CDS compresses the source sequence $X^n$ to one of $\approx 2^{nI(X;Z)}$ source codewords, say $Z^n(k)$, which in turn is mapped into an independently generated channel word $U^n(k)$ (the number of channel words is still $\approx 2^{nI(X;Z)}$). At receiver $i$, the channel output $V^n_i$ will be jointly typical with $\approx 2^{n[I(X;Z) - I(U;V_i)]}$ channel codewords. When traced back to the source codebook, the corresponding source codewords form a virtual bin (which can overlap with other virtual bins, so does not correspond to binning in the strict sense). Using the side information $Y^n_i$, the actual source codeword $Z^n$ in this bin is then disambiguated with success if and only if

$$I(X;Z|Y_i) \leq I(U;V_i) \quad (3.2)$$

for all $i$. For the quadratic Gaussian case we are interested in here, it was shown in [17] that CDS satisfies (3.1) with equality at all the receivers if and only if

$$N_1W_1 = N_2W_2 = \ldots = N_KW_K.$$

It has to be pointed out that although CDS is governed by the same inequality (3.2) as in separate source and channel coding, it is inherently different since the source

\footnote{The left hand-side of (3.2) coincides with the source coding rate with explicit binning as in Wyner-Ziv coding, and is the same as $I(X;Z) - I(Y_i;Z)$ due to the Markov chain $Y_i \rightarrow X \rightarrow Z$.}
encoder does not perform binning. Intuitively, the benefit of CDS is that, with multiple receivers, it allows a weak receiver with larger channel noise to make up with better side information by avoiding binning at the encoder side. As proposed in [17], CDS can also be used in a dirty-paper setting with a channel \( V_i^n = U^n + S^n + W^n_i \) where the channel state information (CSI) \( S^n \) is also available non-causally at the encoder. By using Costa coding [7], it was shown that transmission is successful if and only if

\[
I(X; Z|Y_i) \leq I(T; V_i) - I(T; S) \tag{3.3}
\]

for all \( i \), where the auxiliary dirty-paper codewords \( T^n \) are generated according to \( T = \gamma S + U \) for some \( \gamma \) with independent \( U \) and \( S \). This extension was termed dirty-paper-coded CDS (DPC-CDS) and it also creates virtual bins at the receivers. DPC-CDS is an important building block of our schemes.

In [17], CDS was also extended to the layered description scheme (LDS) for \( K = 2 \) receivers by adding a refinement layer designated for the receiver with better combined channel and side information quality, i.e., the one with the smaller \( W_i N_i \). At the encoder side, the binning index of the refinement message is mapped to a channel codeword, which, in turn, is used as artificial CSI at the CDS encoder. The receiver first decodes the common information transmitted via CDS and then the additional information is decoded only at the refinement receiver. It was shown that LDS outperforms separate coding. We will compare the performance of our schemes with that of LDS since, to the best of our knowledge, it provides the best known performance in digital schemes.

In [30], a hybrid digital/analog scheme, which will be referred to as the HDA-WZ scheme here, was proposed and proved to be optimal for point-to-point transmission. In the HDA-WZ scheme, an auxiliary random variable is defined as \( T = \kappa X + U \), and a
Figure 3.1: Illustration of the codebooks of a successful CDS transmission. The cross-hatched codewords are the actually used ones and the double sided arrows denote joint typicality, whereas the arrows with a cross denote atypical pairs. At the receiver side, all the hatched codewords in the source codebook $Z^n$ together form a virtual bin.

Figure 3.2: Illustration of HDA-WZ. The codebook of auxiliary HDA codeword $T^n$ is shown at both the encoder and the receiver.

A codebook of size $2^{nR}$ is generated from typical $T^n$, as shown in Figure 3.2. The encoder then finds a codeword $T^n$ that is jointly typical with the source sequence $X^n$, and then sends the corresponding $U^n = T^n - \kappa X^n$. At the receiver side, the unique $T^n$ that is jointly typical with $V^n$ and $Y^n$ is chosen, and $X^n$ is estimated from $T^n$, $V^n$, and $Y^n$.

To guarantee successful decoding of the correct $T^n$,

$$I(T; X) \leq R \leq I(T; V, Y)$$

(3.4)
needs to be satisfied. When
\[ \kappa^2 = \frac{P^2}{N(P + W)} \] (3.5)
the effective source coding rate \( I(T; X) \) and channel coding rate \( I(T; V, Y) \) match, and
the resultant distortion becomes \( D_{opt} \). This also implies that, in the WZBC scenario,
the trivial outer bound (3.1) is achieved by HDA-WZ if and only if
\[ N_1(P + W_1) = N_2(P + W_2) = \ldots = N_K(P + W_K). \]
That is because only in this case, the same choice of \( \kappa \) matches what the receivers need
for optimality, as is apparent from (3.5). Although this condition was not explicitly
mentioned in [30], it was in [14], which uses an equivalent modulo-lattice modulation
instead.

### 3.3 Results

In this section, we will first propose a new scheme for point-to-point transmission
with SSI by combining separate source-channel coding with HDA-WZ, and discuss how
this scheme provides complete freedom in choosing the dirty-paper auxiliary codeword
structure (through the choice of the coefficient \( \gamma \) in DPC) for any power allocation
between the two streams. The point-to-point scheme will then be extended to the
two-receiver broadcast scenario (WZBC) by simply replacing the separate coding block
with DPC-CDS. Utilizing the extra freedom in choosing the DPC coefficient \( \gamma \), we will
then show that this WZBC scheme, termed HDA-CDS, achieves the trivial outer bound
whenever the system parameters fall into the region sandwiched between those for which
CDS and HDA-WZ achieve the same. Finally, we will add an analog layer to HDA-CDS,
and discuss the performance of the resultant scheme, termed AHC.
3.3.1 A New Point-to-Point Scheme

A new scheme for point-to-point transmission with SSI will be introduced in this section, as shown in Figure 3.3.

The scheme is essentially Scheme 3 of Chapter 2 with trivial CSI, and as shown in Sec. 2.1.3, the optimum distortion can be achieved with a range of auxiliary random variables (or equivalently, a range of $\gamma$) for any power allocation between the hybrid stream and the digital stream, by making the hybrid stream serve as artificial CSI for the digital one and fully utilizing the decoded auxiliary dirty-paper codeword.

For completeness, we will show again that the same kind of freedom exists in this system when we make full use of the decoded dirty-paper codeword at the receiver. The source is encoded as usual with an optimal quantizer characterized by a backward test channel $X = Z + E$ with $Z \perp E$, followed by binning. The total channel input power $P$ is split into $P_h = \bar{\nu}P$ and $P_d = \nu P$ for hybrid and digital paths, respectively, where $0 \leq \nu \leq 1$ and $\bar{\nu} = 1 - \nu$. The quantization error $E^n$ is transmitted by HDA-WZ scheme as an analog source, where the HDA auxiliary codeword is constructed by $T_h = U_h + \kappa E$. The HDA channel codeword $U^n_h$ is fed as artificial CSI to the DPC channel encoder, which maps the bin index $i$ into a bin of auxiliary dirty-paper codewords $T^n_d$ characterized by $T_d = \gamma U_h + U_d$, and chooses the unique $T^n_d$ in the bin which is jointly
typical with $U^n_h$. The corresponding digital channel input is $U^n_d = T^n_d - \gamma U^n_h$ and the channel output is $V^n = U^n_h + U^n_d + W^n$. Just as in [9], $\gamma$ is not confined to the “optimal” choice in [7], which is

$$\gamma_{\text{Costa}} = \frac{\nu P}{\nu P + W},$$

but rather, any feasible $\gamma$ is considered.

At the receiver side, the digital decoder operates first. To ensure that it successfully decodes $Z^n$ (as well as $T^n_d$), we need

$$I(X; Z | Y) \leq I(T_d; V) - I(T_d; U_h),$$

where $I(X; Z | Y)$ is the Wyner-Ziv source coding rate and $I(T_d; V) - I(T_d; U_h)$ is the effective digital channel capacity.

The HDA-WZ decoder has access to $T^n_d$, $V^n$, $Y^n$, and $Z^n$. Defining $\tilde{Y} = Y - \rho Z = \rho E + N$ as the effective side information for the “source” $E^n$, we then need

$$I(T^n_h; E) \leq I(T^n_h; T^n_d, V, Y, Z)$$

$$= I(T^n_h; T^n_d, V, \tilde{Y}, Z)$$

$$\overset{(a)}{=} I(T^n_h; T^n_d, V, \tilde{Y})$$

(3.7)

to decode $T^n_h$ successfully, where $(a)$ follows from $Z \perp E$.

Now we are ready to introduce the extra level of freedom of the scheme.

**Theorem 2** $D_{\text{opt}}$ is achievable for any $(\nu, \gamma)$ satisfying $0 \leq \nu \leq 1$ and $\gamma_1 \leq \gamma \leq \gamma_2$ where

$$\gamma_{1,2} = \gamma_{\text{Costa}} \left(1 \pm \sqrt{\frac{P + W}{\nu P}}\right).$$

(3.8)

The proof is given in Sec. 3.5.1.

**Remark 3** Theorem 2 suggests that for any power allocation between the two streams, the proposed scheme achieves the optimal distortion in the presence of SSI, for a range
of auxiliary codeword given by (3.8), instead of only Costa’s construction. Thus even
when the effective channel is not used at its full capacity, one can still achieve the
optimal distortion by fully utilizing the decoded auxiliary codeword. This result implies
the freedom introduced in [9] is not an isolated case. Moreover, as shown in the next
section, since both \( \nu \) and \( \gamma \) can be freely chosen for any receiver, under certain conditions,
we are able to find a \((\nu, \gamma)\) pair making the encoder optimal for both receivers.

It is worth noting that a similar scheme can be found where \( X^n \) is input to the HDA-
WZ encoder instead of \( E^n \). The HDA auxiliary codebook would now be constructed
according to \( T_h = U_h + \kappa X \) and the digital dirty paper auxiliary codeword would be
constructed the same way as in the above scheme. In this alternative scheme, the
condition for successful digital transmission is still given by (3.6). To successfully decode
\( T^n_h \), the counterpart of (3.7) becomes

\[
I(T_h; X) \leq I(T_h; T_d, V, Y, Z), \tag{3.9}
\]

which reduces to (3.23) (in Appendix A) as well. Note that since the construction of \( T_h \)
contains \( X \), the right hand side of (3.9) can not be simplified to \( I(T_h; T_d, V, \tilde{Y}) \) as before,
and the optimal estimate of \( X \) is given by a linear combination of \( T_h, T_d, V, Y, \) and
\( Z \). By rigorous calculation, it can be shown that, once again, optimal distortion \( D_{opt} \)
is achieved when equalities hold in both (3.22) and (3.23). We omit the algebra here as
it is not essential for the purposes of this paper. Finally, this scheme also specializes to
the scheme in [30, Sec. III-C] by letting \( \gamma = 0 \).
3.3.2 A Basic Scheme: HDA-CDS

We can replace the separate digital source and channel encoders inside the dotted box in Figure 3.3 with a DPC-CDS encoder, resulting in a point-to-point scheme shown in Figure 3.4, termed HDA-CDS.

We stick to the previous notation for consistency, though the reader needs to remember that it is an essentially different scheme. As before, the source is quantized with a backward test channel $X = Z + E$, and the quantization error $E^n$ is transmitted by an HDA-WZ stream. The HDA auxiliary random variable is constructed by $T_h = \kappa E + U_h$ and $U_h^n$ is the HDA channel word. What is different is that no source binning is per-
formed and the DPC-CDS encoder directly maps the quantization index into a bin of dirty-paper auxiliary channel codewords. The encoder then uses the CSI $U_h^n$ to choose the right codeword inside the bin, as before.

When there is only one receiver, the governing inequalities remain to be (3.6) and (3.7) and therefore the freedom in Theorem 2 persists.

Inspired by the fact that CDS and HDA-WZ achieve the trivial outer bound under different channel and side information conditions, we apply HDA-CDS on the WZBC problem. Although the number of receivers, $K$, can be arbitrary, we focus here on the two-receiver case, as shown in Figure 3.5. As we shall see, HDA-CDS takes advantage of the new level of freedom revealed for point-to-point transmission, and the combination turns out to be better than the sum of its parts.

Without loss of generality, we assume that the second receiver has lower channel noise, i.e., $W_1 > W_2$, and the channel output at receiver $i \in \{1, 2\}$ is now given by $V_i^n = U_d^n + U_h^n + W_i^n$. At each receiver $i$, with the help of side information $Y_i^n$ and the received channel word $V_i^n$, first $Z^n$ and $T_d^n$ are decoded. From (3.3), this is possible whenever

$$I(X; Z|Y_i) \leq I(T_d; V_i) - I(T_d; U_h).$$

Then the HDA-WZ decoder uses $V_i^n$, $T_d^n$, and $\tilde{Y}_i^n$ to decode $T_h^n$, where $\tilde{Y}_i = Y_i - \rho_i Z = \rho_i E + N_i$. This is possible if

$$I(T_h; E) \leq I(T_h; T_d, V_i, \tilde{Y}_i)$$

as follows from (3.7).

Finally, an MMSE estimate of $X^n$ using $Z^n$, $V_i^n$, $T_d^n$, $T_h^n$, and $Y_i^n$ is performed.

Now, we can optimize over $(\nu, \gamma, \kappa)$ and the quantization rate of the encoder to minimize the distortions. As mentioned before, when $W_1 N_1 = W_2 N_2$, one can simply
Figure 3.6: For any given $W_1 > W_2$, each point on the $(N_1, N_2)$-plane corresponds to a system with parameters $(W_1, W_2, N_1, N_2)$. As indicated by the two lines in the figure, the conditions for achieving the trivial outer bound for CDS and the HDA-WZ scheme are $W_1N_1 = W_2N_2$ and $N_1(P + W_1) = N_2(P + W_2)$, respectively. When $N_1 = N_2 = 1$, or the SSI is trivial at each receiver, analog transmission achieves the outer bound, as indicated by the dot. The shaded region, as stated in Theorem 4, indicates where our scheme is achieving the outer bound.

set $\nu = 1$, with which our scheme reduces to CDS and achieves the outer bound (3.1).

Similarly, when $N_1(P + W_1) = N_2(P + W_2)$, the choice $\nu = 0$ reduces our scheme to pure HDA-WZ, and the outer bound (3.1) is achieved. Since $W_1 > W_2$, these correspond to distinct lines in the $(N_1, N_2)$-plane shown in Figure 3.6. As stated by the following theorem, HDA-CDS achieves the outer bound (3.1) in the entire shaded region sandwiched between these two lines.

**Theorem 4** Given $W_1$ and $W_2$ with $W_1 > W_2$, the hybrid scheme HDA-CDS achieves the outer bound (3.1) for all $(N_1, N_2)$ pairs satisfying

$$\frac{P + W_1}{P + W_2} N_1 < N_2 < \frac{W_1}{W_2} N_1.$$ 

The proof is given in Sec. 3.5.2.

**Remark 5** The intuition of Theorem 4 is that by exploiting the extra level of freedom and benefits of virtual binning, we are able to find an encoder which is optimal for both
receivers when they have “similar” combined channel/SSI quality (so that the system falls in the shaded region in Figure 3.6). In this case, each receiver can achieve its $D_{\text{opt}}$ as if it is the only receiver, although the receivers have different channel noise and SSI quality.

As we show in the next theorem, HDA-CDS performs at least as well as LDS introduced in [5] even outside of the shaded region in Figure 3.6. That is because we can simply achieve the same the performance of the LDS by replacing the separate source and channel coding for the refinement layer of the LDS with HDA-WZ coding, and mimicking LDS with a proper choice of $\gamma$ and corresponding $\kappa$.

**Theorem 6** HDA-CDS can achieve the same $(D_1, D_2)$ pairs as LDS when either $N_2 > \frac{W_1}{W_2} N_1$ or $N_2 < \frac{P + W_1}{P + W_2} N_1$.

The proof is given in Sec. 3.5.3.

Unfortunately, the complete tradeoff between $D_1$ and $D_2$ proved very difficult to derive. However, we numerically observed that outside of the region $\frac{P + W_1}{P + W_2} N_1 \leq N_2 \leq \frac{W_1}{W_2} N_1$, the achievable distortion region (after convexification) is always the convex hull of pure HDA-WZ and LDS. See Figure 3.8(a) and (c)-(f) for examples of this phenomenon.

In summary, HDA-CDS achieves the trivial outer bound in the entire shaded region shown in Figure 3.6, instead of only at the two linear boundaries and in other regions, and HDA-CDS is also shown to outperform the LDS in [17], which, in turn, is the best known digital scheme.
3.3.3 The Main Scheme

Since analog transmission itself is optimal for WZBC with trivial SSI at all the receivers, it will improve the performance if an analog stream is added to the HDA-CDS\(^4\). We propose to add the analog component using dirty paper coding as shown in Figure 3.7. The structure of the decoders remains the same as in HDA-CDS. We call this scheme the Analog-HDA-CDS scheme, or AHC for short. The source vector \(X^n\) is quantized in the same way with the backward test channel \(X = E + Z\) and the quantization error \(E^n\) is transmitted by both HDA-WZ scheme and analog transmission. The analog stream is considered at the dirty-paper digital channel encoder as additional artificial CSI besides \(U^n_h\). The channel power is now split into three parts \(P = P_h + P_a + P_d\) for HDA-WZ, analog, and CDS. The quantization error is scaled by a constant \(c\) so that \(c^2E = P_a\). At the HDA-WZ encoder, the auxiliary codeword is again constructed by \(T_h = U_h + \kappa E\), while the auxiliary dirty paper codeword at the DPC-CDS encoder by \(T_d = U_d + \gamma_h U_h + \gamma_a c E\).

Both receivers first decode \(Z^n\), requiring

\[
I(X; Z|Y_i) \leq I(T_d; V_i) - I(T_d; cE, U_h), \tag{3.12}
\]

\(^4\)Adding the analog stream will improve the performance at least for the case with trivial SSI at all the receivers, and when the analog stream does not help, we can always set \(P_a = 0\) to let AHC reduce to HDA-CDS.

\[
\frac{P_d + W_1}{P_d^2(P + W_1)} \begin{bmatrix}
\gamma_a - \gamma_c & \gamma_h - \gamma_c \\
\gamma_a - \gamma_c & \gamma_h - \gamma_c \\
\end{bmatrix}
\begin{bmatrix}
(P_d + P_h + W_1)P_a & -P_a P_h \\
-P_a P_h & (P_a + P_d + W_1)P_h \\
\end{bmatrix}
\begin{bmatrix}
\gamma_a - \gamma_c \\
\gamma_h - \gamma_c \\
\end{bmatrix}
\leq 1. \tag{3.16}
\]
Figure 3.7: The encoder of the AHC scheme. An extra analog stream is superimposed with dirty paper coding to incorporate the benefits of analog transmission. The structure of decoders is the same as HDA-CDS.

$i = 1, 2$, where the channel output is now $V_i^n = U_h^n + cE^n + U_d^n + W_i^n$. Note that the system parameters should be chosen so that the effective digital channel capacity [the right hand side of (3.12)] is non-negative.

The next lemma translates (3.12) into how $E$, $\gamma_a$, and $\gamma_h$ should be chosen for any $(P_a, P_d, P_h)$ triplet.

**Lemma 7** For any power allocation $(P_a, P_d, P_h)$, both receivers can decode $Z^n$ if and only if

$$\frac{1}{E} \leq 1 + \min_{i=1,2} \left\{ -\frac{1}{N_i} + \frac{P_d(P + W_i)}{\Gamma_i} \right\}, \quad (3.13)$$

where

$$\Gamma_i = P_aP_d\gamma_a^2 + P_a P_h (\gamma_h - \gamma_a)^2 + P_h P_d \bar{\gamma}_h^2 + W_i (P_a \gamma_a^2 + P_h \gamma_h^2 + P_d) \quad (3.14)$$

with $\gamma_a = 1 - \gamma_a$, $\bar{\gamma}_h = 1 - \gamma_h$. Also, $\gamma_a$ and $\gamma_h$ need to satisfy

$$(P_d + P_h + W_1)P_a \gamma_a^2 + (P_d + P_a + W_1)P_h \gamma_h^2 - 2P_a P_h \gamma_a \gamma_h - 2P_a P_d \gamma_a - 2P_h P_d \gamma_h - P_d^2 \leq 0. \quad (3.15)$$

The proof is given in Sec. 3.5.4.
Remark 8 We note here that the ellipse (3.15) has the center \((\gamma_c, \gamma_c)\) with \(\gamma_c = \gamma_{\text{Costa},1} = \frac{P_d}{P_a + W_i}\), and it can be expressed in the standard form as in (3.16) at the bottom of the page.

As for the decoding of \(T^n_h\), it can be accomplished, just as in HDA-CDS, at receiver \(i\) if and only if

\[ I(T_h; E) \leq I(T_h; T_d, V_i, \tilde{Y}_i), \tag{3.17} \]

which is the same as

\[ h(T_h) - h(U_h) \leq h(T_h) + h(T_d, V_i, \tilde{Y}_i) - h(T_h, T_d, V_i, \tilde{Y}_i). \tag{3.18} \]

After some algebra, the condition (3.18) reduces to the quadratic form

\[ c_{i1}\kappa^2 + c_{i2}\kappa + c_{i3} \leq 0, \tag{3.19} \]

where

\[
\begin{align*}
c_{i1} &= P_h E_i(N_i\gamma_h^2 + W_i\gamma_h^2) + P_d E_i W_i, \\
c_{i2} &= -2P_h c E_i [P_d(\gamma_a - \gamma_h) + (P_d + W_i)\gamma_a \gamma_h], \\
c_{i3} &= -P_h^2 \left[ \tilde{Y}_i P_d \gamma_h^2 + W_i \tilde{Y}_i \gamma_h^2 + P_a N_i (\gamma_h - \gamma_a)^2 \right]
\end{align*}
\]

with \(\tilde{Y}_i = \rho_i^2 E + N_i\). The discriminant of (3.19) is given by

\[
\Delta = 4P_h^2 E_i \left( P_d \gamma_h^2 + W_i \gamma_h^2 \right) \left[ P_a P_d N_i \gamma_a^2 + P_a P_h N_i (\gamma_h - \gamma_a)^2 + P_h P_d \tilde{Y}_i \gamma_h^2 \\
+ W_i (P_a N_i \gamma_a^2 + P_h \tilde{Y}_i \gamma_h^2 + P_d \tilde{Y}_i) \right],
\]

which is clearly non-negative. Now since \(c_{i1} \geq 0\) and \(c_{i3} \leq 0\), it immediately follows that (3.19) corresponds to an interval containing \(\kappa = 0\).

At receiver \(i\), the MMSE estimate of \(E^n\) is computed using \(T^n_d, V^n_i, \tilde{Y}_i^n\), and, depending on whether it can be decoded, \(T^n_h\), and the corresponding distortion is given in the following lemma.
Lemma 9 Define

\[ D_i(\kappa) = \frac{P_d P_h E N_i W_i}{c_{i1} \kappa^2 + c_{i2} \kappa + P_a P_h N_i (W_i \gamma_a^2 + P_d \gamma_a^2) + P_h P_d W_i Y_i} \]

for \( i = 1, 2 \), where \( E \) satisfies (3.13) with equality. For any \( \kappa \), the distortion at receiver \( i \) is then given by

\[ D_i = \begin{cases} 
D_i(\kappa) & \text{when } \kappa \text{ satisfies (3.19)} \\
D_i(-\frac{c_{i2}}{2c_{i1}}) & \text{otherwise} 
\end{cases} \]  

(3.20)

The proof is given in Sec. 3.5.5.

We now discuss the performance of the AHC scheme for each aforementioned region on the \((N_1, N_2)\)-plane separately. Since HDA-CDS is a special case of AHC obtained simply by setting \( P_a = 0 \), it is obvious that in the region where \( \frac{P_d W_i}{P_a W_2} N_1 < N_2 < \frac{W_i}{W_2} N_1 \), Theorem 4 still holds. An example of this can be observed in Figure 3.8(b). In the other regions, although analytical dissection is very difficult to obtain, based on extensive numerical simulations we arrived at the following conjectures:

**Conjecture 10** When \( N_2 > \frac{W_i}{W_2} N_1 \), the analog stream is not helpful at all, i.e., optimal distortion performance is achieved when \( P_a = 0 \). Moreover, \( \gamma_h = 1 \) is always the optimal choice, and thus as a consequence of Theorem 6, the performance of AHC coincides with that of LDS.

An example of this phenomenon is shown in Figure 3.8(a).

As a sanity check, this conjecture implies that analog transmission alone can never outperform LDS in this region on the \((N_1, N_2)\)-plane. But this is indeed the case, as was analytically shown in [17, Section V-C].

**Conjecture 11** When \( N_1 < N_2 < \frac{P_d W_i}{P_a W_2} N_1 \), it is observed that \( P_d = 0 \) always, i.e., a simple superposition scheme with only analog and HDA-WZ streams suffices.
Figure 3.8: Performance comparison between AHC scheme, LDS, analog transmission, and separate coding when $P = 1$. In (a), $N_2 > \frac{W_1}{W_2} N_1$. In (b), $\frac{P + W_1}{P + W_2} N_1 < N_2 < \frac{W_1}{W_2} N_1$, and AHC scheme achieves the trivial outer bound. In (c) and (d), $N_1 < N_2 < \frac{P + W_1}{P + W_2} N_1$. In (e), there is no side information and analog transmission is optimal. In (f), $N_2 < N_1$. 

51
As can be seen in Figures 3.8(c) and (d), the resultant performance is strictly better than those of HDA-CDS and analog transmission alone.

**Conjecture 12** When $N_2 \leq N_1$, all three streams make some contribution and the performance is better than those of HDA-CDS and analog alone. In addition, $\gamma_a = \gamma_h = 0$, i.e., dirty paper coding is not necessary and simple superposition is good enough.

The performance is illustrated in Figure 3.8(f), and when $P_a = P$, the scheme reduces to analog transmission, which is indicated by the point on the curve.

Bringing all the conjectures together, one can conclude that the analog stream is useless as CSI, and the AHC scheme can be simplified to the superposition of HDA-CDS and the analog stream. That is because we observe either $P_a = 0$, or $P_d = 0$, or $\gamma_a = \gamma_h = 0$.

### 3.4 Conclusion

For the bandwidth-matched quadratic Gaussian WZBC problem, we proposed a new hybrid digital/analog coding scheme called AHC, and demonstrated that it outperforms all previously known schemes. For the case of two receivers, AHC is analytically shown to achieve (without the help of the analog stream) the trivial outer bound for the entire region in the parameter space sandwiched between the optimality conditions for HDA-WZ and CDS, the two building blocks. This result uses the new level of freedom we discovered for point-to-point transmission, namely, the freedom in choosing the auxiliary dirty-paper codeword in addition to the well known freedom of power allocation. Outside that region, we numerically observed that the AHC scheme reduces to a simple superposition of HDA-CDS and the analog stream, and thus dirty paper coding is not necessary between the two.
3.5 Appendix

3.5.1 Proof of Theorem 2

Inequality (3.6) is the same as

$$I(X; Z) - I(Y; Z) \leq h(V) - h(T_d, V) + h(T_d | U_h),$$

which can be expanded as

$$\frac{1}{2} \log \left(1 - N + \frac{N}{E}\right) \leq \frac{1}{2} \log \frac{(P + W)\nu P}{(P + W)(\gamma^2 \nu P + \nu P) - (\gamma \nu P + \nu P)^2},$$

and finally reduces to

$$\frac{1}{E} \leq 1 + \frac{\nu P - b}{N(b + \nu W)},$$

where $b = \nu [\nu P(1 - \gamma)^2 + W \gamma^2]$. A condition $\gamma_1 \leq \gamma \leq \gamma_2$ is also necessary to ensure that the effective capacity in (3.6) or the right hand side of (3.21) is non-negative, where $\gamma_{1, 2}$ are given in (3.8).

We observe that the range of feasible $\gamma$ in (3.8) is exactly the same as that in [9], with $P_h = \bar{\nu} P$ playing the role of $P_a$ for analog power. Following similar arguments as in [9], this also implies that when $P_h \to 0$, the range of $\gamma$ becomes the entire real line and the system becomes purely digital. On the other hand, when $P_d \to 0$, $E \to 1$ and the scheme becomes equivalent to HDA-WZ. To see that, notice that the right-hand side of (3.22) is maximized by $\gamma = \gamma_{\text{Costa}}$ for any $P_d$ and the resultant expression after setting $\gamma = \gamma_{\text{Costa}}$ approaches 1 as $P_d \to 0$.

Inequality (3.7) is the same as

$$h(T_h) - h(T_h | E) \leq h(T_h) + h(T_d, V, \tilde{Y}) - h(T_h, T_d, V, \tilde{Y}),$$

which reduces to

$$\kappa^2 \leq \frac{\bar{\nu} P \tilde{Y} b}{EN(b + \nu W)} = \frac{\bar{\nu} P b}{b + \nu W} \left(\frac{1}{E} - 1 + \frac{1}{N}\right).$$

(3.23)
The optimal estimate of \( X^n \) is given by the optimal estimate of \( E^n \), which is obtained by \( \hat{E}^n = k_1 T_d^n + k_2 T_h^n + k_3 V^n + k_4 \tilde{Y}^n \), where \( k_1, k_2, k_3, \) and \( k_4 \) satisfy

\[
\begin{bmatrix}
\gamma^2 \tilde{v} P + \nu P & \gamma \tilde{v} P & \gamma \tilde{v} P + \nu P & 0 \\
\gamma \tilde{v} P & \kappa P + \tilde{v} P & \kappa P \rho E & 0 \\
\gamma \tilde{v} P + \nu P & \tilde{v} P & P + W & 0 \\
0 & \kappa P \rho E & 0 & \tilde{Y}
\end{bmatrix}
\begin{bmatrix}
k_1 \\
k_2 \\
k_3 \\
k_4
\end{bmatrix}
= \begin{bmatrix}
k_E \\
k_N \\
k_0 \\
\rho E
\end{bmatrix}
\]

due to the principle of orthogonality. We thus have

\[
\begin{bmatrix}
k_1 \\
k_2 \\
k_3 \\
k_4
\end{bmatrix}
= \frac{E}{EN(b + \nu W)\kappa^2 + \nu \nu P W Y} \cdot \begin{bmatrix}
\nu \kappa N (\gamma_{Costa} - \gamma) (\nu P + W) \\
\nu \kappa N (b + \nu W) \\
\nu \nu P \kappa N (\gamma - 1) \\
\rho \nu \nu P W
\end{bmatrix}
\]

The square-error distortion is given by

\[
D = E \left\{ (E - \hat{E}) E \right\}
= E(1 - k_2 \kappa - k_4 \rho)
= E \frac{\nu \nu P EN W}{EN (b + \nu W) \kappa^2 + \nu \nu P W Y}
\geq \frac{\nu EN W}{(b + \nu W) \tilde{Y}}
\geq \frac{NW}{P + W}
= D_{opt},
\]

where (3.25) and (3.26) are satisfied by equality when equalities hold in (3.23) and (3.22), respectively.

\[\blacksquare\]

### 3.5.2 Proof of Theorem 4

Let us immediately remark that the hypothesis of the theorem implies \( N_1 < N_2 \).
Since (3.10) and (3.11) have the same forms as (3.6) and (3.7) in the point-to-point scenario, by defining

\[ b_i = \bar{\nu}[\nu P(1 - \gamma)^2 + W_i \gamma^2] \]

for \( i = 1, 2 \), they reduce to

\[ \frac{1}{E} \leq 1 + \frac{\nu P - b_i}{N_i(b_i + \nu W_i)} \]  

(3.27)

and

\[ \kappa_i^2 \leq \frac{\bar{\nu} P b_i}{b_i + \nu W_i} \left( \frac{1}{E} - 1 + \frac{1}{N_i} \right) \]  

(3.28)

respectively, as before. Now further define

\[ \alpha_i = \frac{\nu P - b_i}{N_i(b_i + \nu W_i)} \]

and

\[ \kappa_i^2 = \frac{\bar{\nu} P b_i}{b_i + \nu W_i} \left( \frac{1}{E} - 1 + \frac{1}{N_i} \right) \cdot \]

When we consider both receivers, to ensure decoding \( Z^n \) and \( T^n_d \), we need

\[ \frac{1}{E} \leq 1 + \min \{\alpha_1, \alpha_2\} \cdot \]

Further, at each receiver, the HDA auxiliary codeword \( T^n_h \) can be decoded if and only if \( \kappa^2 \leq \kappa_i^2 \).

In addition to (3.27) and (3.28), the range of the free parameter \( \gamma \) has to be confined to \( 0 \leq b_i \leq \nu P \) so that the effective capacity \( I(T_d; V_i) - I(T_d; U_h) \) in (3.10) is non-negative. We refer to pairs \((\nu, \gamma)\) satisfying this requirement as feasible.

In the point-to-point version of our scheme, it is shown that for any feasible pair \((\nu, \gamma)\), the minimum distortion \( D_{\text{opt}} \) can be achieved if \( E \) and \( \kappa \) are chosen so as to satisfy (3.22) and (3.23), i.e., the counterparts of (3.27) and (3.28), with equality. This implies for the WZBC problem at hand that if we can find a feasible pair \((\nu, \gamma)\) so that
(3.27) and (3.28) are satisfied with equality simultaneously for \(i = 1, 2\) with some choice of \(E\) and \(\kappa\), then the trivial outer bound (3.1) can be achieved. This, in turn, requires \(\alpha_1 = \alpha_2\) and \(\kappa_1^2 = \kappa_2^2\) simultaneously.

Now, \(\alpha_1 = \alpha_2\) can be written as

\[
\frac{a_1}{\frac{\nu \beta_1}{\nu} - a_1 N_1} = \frac{a_2}{\frac{\nu \beta_2}{\nu} - a_2 N_2}
\]

(3.29)

where

\[
\beta_i = N_i(P + W_i)
\]

\[
a_i = \nu P - b_i.
\]

Note that \(\beta_2 > \beta_1\) in the region of interest on the \((N_1, N_2)\)-plane.

Similarly, \(\kappa_1^2 = \kappa_2^2\) implies from \(\alpha_1 = \alpha_2\) that

\[
\frac{b_1}{b_1 + \nu W_1} \left( \frac{1}{N_1} + \alpha_1 \right) = \frac{b_2}{b_2 + \nu W_2} \left( \frac{1}{N_2} + \alpha_2 \right),
\]

or equivalently, from (3.29) that

\[
\frac{\beta_1}{a_1^2} \left( \nu P - a_1 \right) = \frac{\beta_2}{a_2^2} \left( \nu P - a_2 \right).
\]

(3.30)

We first ignore the dependency on \(\gamma\), and search for a triplet \((a_1, a_2, \nu)\) in the “interior” feasible set \(0 < \nu < 1\) and \(0 < a_1 < a_2 < \nu P\). Note that \(a_1 \leq a_2\) is guaranteed in the feasible set because of \(W_1 > W_2\). Solving for \(\nu\) in both (3.29) and (3.30) for a fixed \((a_1, a_2)\) pair yields

\[
\nu = \frac{a_1 a_2 (N_2 - N_1)}{a_1 \beta_2 - a_2 \beta_1} = \frac{a_1 a_2 (a_1 \beta_2 - a_2 \beta_1)}{P(a_1^2 \beta_2 - a_2^2 \beta_1)}.
\]

(3.31)

For every fixed \(a_2 > a_1 > 0\), \(0 < \nu < 1\) then implies

\[
\frac{a_1 a_2 (N_2 - N_1)}{a_1 \beta_2 - a_2 \beta_1} < 1
\]

(3.32)
and
\[ a_1\beta_2 - a_2\beta_1 > 0 . \]  \hfill (3.33)

In fact, (3.33) automatically follows from the required consistency between the two solutions of \(\nu\) in (3.31). To see that, rewrite (3.31) as
\[ P(N_2 - N_1)(a_1^2\beta_2 - a_2^2\beta_1) = (a_1\beta_2 - a_2\beta_1)^2 \]  \hfill (3.34)

which, in particular, implies using the non-negativity of the right-hand side and \(\beta_2 > \beta_1\) that
\[ a_2 \leq \sqrt{\frac{\beta_2}{\beta_1}} a_1 < \frac{\beta_2}{\beta_1} a_1 . \]

Note that \(a_2 < \nu P\) is also granted due to the second equality in (3.31) because
\[ \frac{a_1(a_1\beta_2 - a_2\beta_1)}{a_1^2\beta_2 - a_2^2\beta_1} > 1 \]
whenever \(a_1 < a_2\). Thus, it suffices to find \(a_2 > a_1 > 0\) such that (3.32) and (3.34) are simultaneously satisfied, and find the corresponding \(\nu\) using either formula in (3.31).

Now, expanding (3.34), we obtain
\[ \beta_1(PN_2 + N_1W_1)a_2^2 - 2\beta_1\beta_2 a_1 a_2 + \beta_2 a_1^2(PN_1 + N_2W_2) = 0 \]  \hfill (3.35)

which is quadratic in \(a_2\) for every fixed \(a_1\). The discriminant can be computed as
\[ \Delta = 4\beta_1\beta_2 a_1^2 P(N_2 - N_1)(N_1W_1 - N_2W_2) \]
which is strictly positive due to the fact that \(N_1W_1 > N_2W_2\). Thus, (3.35) has two positive roots for any \(a_1 > 0\). Denoting the left-hand side of (3.35) as \(f_{a_1}(a_2)\), it can be shown after some algebra that
\[ f_{a_1}(a_1) = a_1^2(\beta_1 - \beta_2)(N_1W_1 - N_2W_2) < 0 . \]
This implies that setting \( a_2 \) to the larger root of \( f_{a_1}(a_2) = 0 \), we automatically satisfy \( a_2 > a_1 \). What remains to be found is then under what conditions on \( a_1 \) that root also satisfies (3.32). Rewriting (3.32) as

\[
a_2 < \frac{\beta_2 a_1}{\beta_1 + a_1(N_2 - N_1)}
\]

and explicitly computing the larger root of \( f_{a_1}(a_2) = 0 \) as \( a_2 = \phi a_1 \) with

\[
\phi = \frac{\beta_1 \beta_2 + \sqrt{\beta_1 \beta_2 P(N_2 - N_1)(N_1 W_1 - N_2 W_2)}}{\beta_1 (PN_2 + N_1 W_1)}
\]

(3.32) translates to

\[
a_1 < \frac{\beta_2 - \phi \beta_1}{\phi (N_2 - N_1)}.
\]  

(3.36)

Of course, (3.36) is meaningful only if

\[
\beta_2 > \phi \beta_1
\]

which follows after some algebra using \( \beta_2 > \beta_1 \).

In summary, if we pick any

\[
0 < a_1 < \frac{\beta_2 - \phi \beta_1}{\phi (N_2 - N_1)}
\]

together with \( a_2 = \phi a_1 \) and \( \nu \) from (3.31), we simultaneously satisfy \( \alpha_1 = \alpha_2 \) and \( \kappa_1^2 = \kappa_2^2 \) with \( 0 < \nu < 1 \) and \( 0 < a_1 < a_2 < \nu P \).

The only thing that remains is to find a \( \gamma \) that is consistent with \( a_1, a_2, \) and \( \nu \). We have that

\[
\bar{\nu} W_1 \gamma^2 + \bar{\nu} \nu P(1 - \gamma)^2 = \nu P - a_1
\]

\[
\bar{\nu} W_2 \gamma^2 + \bar{\nu} \nu P(1 - \gamma)^2 = \nu P - \phi a_1.
\]

Instead of solving \( \gamma \) directly, let us temporarily treat \( \mu_1 = \gamma^2 \) and \( \mu_2 = (1 - \gamma)^2 \) as free
variables and solve the linear system above:

\[
\begin{align*}
\mu_1 &= \frac{a_1(\phi - 1)}{\nu(W_1 - W_2)} \\
\mu_2 &= \frac{1}{\nu} \left[ 1 - \frac{a_1(\phi W_1 - W_2)}{\nu P(W_1 - W_2)} \right].
\end{align*}
\]

By close inspection, one can actually show that as \(a_1\) varies, we have

\[
\mu_2 = c\mu_1 + d
\]

with

\[
d = 1 - \frac{(\beta_2 - \phi\beta_1)(\phi W_1 - W_2)}{P\phi(W_1 - W_2)(N_2 - N_1)}
\]

and

\[
c = d \cdot \frac{\phi(W_1 - W_2)(N_2 - N_1)}{(\phi - 1)(\beta_2 - \phi\beta_1)}.
\]

Since \(c\) and \(d\) are constants, this results in a line on the \((\mu_1, \mu_2)\)-plane. Also, since \(a_1 \to 0\) and \(a_1 \to \frac{\beta_2 - \phi\beta_1}{\phi(N_2 - N_1)}\) respectively imply \(\mu_1 \to 0\) and \(\mu_1 \to \infty\), this line stretches through the entire interval \(\mu_1 \in (0, \infty)\). In fact, the slope of the line depends only on \(\frac{N_2}{N_1}\) for given \(W_1 > W_2\).

Now, we show that

\[
P\phi(W_1 - W_2)(N_2 - N_1) > (\beta_2 - \phi\beta_1)(\phi W_1 - W_2) \quad (3.37)
\]

and therefore \(0 < d < 1\), implying also that \(c > 0\). Towards that end, rewrite (3.37) as

\[
g(\phi) = \beta_1 W_1 \phi^2 - N_1 W_1(P + W_2)\phi - N_2 W_2(P + W_1)\phi + \beta_2 W_2.
\]

The quadratic form \(g(\phi)\) is minimized at

\[
\phi^* = \frac{N_1 W_1(P + W_2) + N_2 W_2(P + W_1)}{2N_1 W_1(P + W_1)}
= \frac{1}{2} \frac{P + W_2}{P + W_1} + \frac{1}{2} \frac{N_2 W_2}{N_1 W_1}
< 1.
\]
Also,

\[ g(1) = (N_1 W_1 - N_2 W_2)(W_1 - W_2) > 0 \]

implying \( g(\phi) > 0 \) since \( \phi > 1 \).

Going back to \( \mu_1 = \gamma^2 \) and \( \mu_2 = (1 - \gamma)^2 \), a \( \gamma \) value consistent with \( \mu_1 \) and \( \mu_2 \) can be found if and only if

\[ \mu_2 = (1 \pm \sqrt{\mu_1})^2. \]

But \( 0 < d < 1 \) and \( c > 0 \) imply that \( \mu_2 = c \mu_1 + d \) intersects with \( \mu_2 = (1 \pm \sqrt{\mu_1})^2 \). In fact, there are always two intersections (and thus two pairs of \( (\nu, \gamma) \) satisfying \( \alpha_1 = \alpha_2 \) and \( \kappa_1^2 = \kappa_2^2 \)) except when \( c = 1 \), or equivalently when

\[ \frac{N_2}{N_1} = \frac{(P^2 + 2P W_1 + W_1 W_2)^2}{(P^2 + 2P W_2 + W_1 W_2)^2}. \]

\[
\begin{bmatrix}
  P_h + \kappa^2 E & P_h + \kappa \nu E & \kappa \rho_i E & \gamma_h P_h + \gamma \alpha \kappa \nu E \\
  P_h + \kappa \nu E & P + W_i & \rho_i \nu E & P_d + \gamma \alpha c E + \gamma h P_h \\
  \kappa \rho_i E & \rho_i \nu E & \tilde{Y}_i & \gamma \alpha \rho_i \nu E \\
  \gamma_h P_h + \gamma \alpha \kappa \nu E & P_d + \gamma \alpha c E + \gamma h P_h & \gamma \alpha \rho_i \nu E & P_d + \gamma c P_a + \gamma h P_h
\end{bmatrix}
\begin{bmatrix}
k_{i1} \\
k_{i2} \\
k_{i3} \\
k_{i4}
\end{bmatrix}
= \begin{bmatrix}
\kappa \nu E \\
c \nu E \\
\rho_i \nu E \\
\gamma \alpha \nu E
\end{bmatrix}
\]

\[
\frac{k_{i1}}{\Omega_i} = \begin{bmatrix}
-\frac{\gamma_h^2 P_h P_d - \gamma_h^2 P_h W_i - P_d W_i - cP_h(P_d \gamma_h \gamma + W_i \gamma_a \gamma h)}{\kappa \nu E} \\
\frac{P_h P_d(\gamma_h \nu c - \gamma_a c)}{\rho_i P_d P_h} \\
-\rho_i P_d P_h \\
\frac{P_h(P_d + W_i)(\gamma_h - \gamma_{Costa,i}) - c(\gamma_a - \gamma_{Costa,i})}{\Omega_i}
\end{bmatrix}
\]

(3.44)
3.5.3 Proof of Theorem 6

The distortion at each receiver depends on whether $\kappa$ is chosen so that $T_n^u$ can be decoded at that receiver. If $\kappa^2 \leq \kappa_i^2$, $T_n^u$ is decoded at receiver $i$, and hence the distortion is given by (3.24). If, on the other hand, $\kappa^2 > \kappa_i^2$, then $T_n^u$ cannot be decoded, and since both $T_d$ and $V_i$ are independent of $E$ or $\hat{Y}$, the optimal reconstruction of $E^n$ is solely given by $\hat{E}_i^n = k'_i \hat{Y}_i^n$, where

$$k'_i = \frac{\rho_i E}{Y_i}.$$ 

The distortion in this case is given by

$$D_i = \mathbb{E}\left\{ (E - \hat{E})E \right\}$$
$$= \mathbb{E}(1 - k'_i \rho_i)$$
$$= \frac{N_i E}{Y_i}.$$ 

Note that this distortion is the same as in (3.24) with $\kappa = 0$. That is because when $\kappa = 0$, no information about $E^n$ is transmitted and decoding $T_n^u$ does not help at all.

When $N_2 > \frac{W_1}{W_2} N_1$, let $\gamma = 1$, and thus $b_i = \nu W_i$. It is easy to show that when

$$\nu > \frac{1}{N_1} - \frac{1}{N_2},$$

$\alpha_1 > \alpha_2$. It can also be shown that in this case, $0 \leq b_i \leq \nu P$ is automatically satisfied because $W_1 > W_2$. Since now $\kappa_i^2 = \nu^2 P \left( \frac{1}{E} - 1 + \frac{1}{N_i} \right)$ and $N_2 > N_1$, $\kappa_1^2 > \kappa_2^2$ is always satisfied. If we choose $\kappa = \kappa_1$, the first receiver is able to decode both the CDS stream and the HDA stream while the second can only decode the CDS stream. The resultant achievable distortion pairs $(D_1, D_2)$ are given by

$$D_2 = \frac{N_2 W_2}{\nu (P + W_2)}.$$  

(3.38)
and

\begin{align*}
D_1 &= \frac{\nu EN_1}{Y_1} \\
&= \frac{\nu N_1 N_2 W_2}{N_2 W_2 + \nu N_1 P - \nu N_1 W_2} \\
&= \frac{\nu N_1}{1 + N_1 \left( \frac{1}{D_2} - \frac{1}{N_2} \right)}.
\end{align*}

(3.39)

Note that (3.38) and (3.39) are actually the same as (79) and (80) in [17] with \( \gamma = 1 \), where our receiver 2 corresponds to the common receiver in [17] because it has worse combined channel and SSI quality.

When \( N_2 < \frac{P + W_1}{P + W_2} N_1 \), it can be shown that \( \alpha_1 < \alpha_2 \) always. If we set \( \gamma = 0 \), we then have \( b_i = \bar{\nu} P \) and \( \kappa_1^2 < \kappa_2^2 \). When we choose \( \kappa = \kappa_2 \), the first receiver decodes both streams, whereas the second one decodes only the CDS stream, and the resultant distortion pairs \( (D_1, D_2) \) are given by

\begin{equation}
D_1 = \frac{N_1 (\bar{\nu} P + W_1)}{P + W_1}
\end{equation}

(3.40)

and

\begin{align*}
D_2 &= \frac{EN_2 W_2}{(\bar{\nu} P + W_2) Y_2} \\
&= \frac{N_1 N_2 W_2 (\bar{\nu} P + W_1)}{(\bar{\nu} P + W_2) [\nu N_2 P + N_1 (\bar{\nu} P + W_1)]} \\
&= \frac{N_2}{1 + N_2 \left( \frac{1}{D_1} - \frac{1}{N_1} \right)} \cdot \frac{W_2}{\bar{\nu} P}.
\end{align*}

(3.41)

Equations (3.40) and (3.41) are again the same as (79) and (80) in [17], this time with \( \gamma = 0 \) and receiver 1 being the common receiver since \( N_2 W_2 < N_1 W_1 \).

3.5.4 Proof of Lemma 7

The left hand side of (3.12) is

\[ I(X; Z|Y_i) = \frac{1}{2} \log \left( 1 - N_i + \frac{N_i}{E} \right) \]
as before, and the right hand side is

\[ I(T_d; V_i) - I(T_d; cE, U_h) = h(V_i) - h(T_d, V_i) - h(T_d) + h(U_d) = \frac{1}{2} \log \frac{P_d(P + W_i)}{\Gamma_i}, \]

where \( \Gamma_i \) is defined in (3.14).

Taking both receivers into consideration, we thus need

\[ \frac{1}{E} \leq 1 + \min_{i=1,2} \left\{ \frac{1}{N_i} + \frac{P_d(P + W_i)}{\Gamma_i} \right\}. \]

In addition, to make sure the effective channel capacity, i.e., the right hand side of (3.12), is non-negative, \( \gamma_a \) and \( \gamma_h \) need to satisfy

\[ (P_d + P_h + W_i)P_a \gamma_a^2 + (P_d + P_a + W_i)P_h \gamma_h^2 - 2P_aP_h \gamma_a \gamma_h - 2P_aP_d \gamma_a - 2P_hP_d \gamma_h - P_d^2 \leq 0 \quad (3.42) \]

for \( i = 1, 2 \), which corresponds to the intersection of two ellipses on the \( (\gamma_a, \gamma_h) \) plane. By close inspection, it is easy to see that the ellipse corresponding to channel 1 is always contained in that of channel 2, as \( W_1 > W_2 \). It then suffices to consider (3.42) for \( i = 1 \) only, which is the same as (3.15).

\[ 3.5.5 \quad \text{Proof of Lemma 9} \]

When \( T_h^n \) can be decoded, the MMSE estimate of \( E^n \) is \( \hat{E}^n_i = k_{i1}T_h^n + k_{i2}V_i^n + k_{i3}\bar{Y}_i^n + k_{i4}T_d^n \), where the coefficients satisfy (3.43) at the bottom of the page. We thus have (3.44) where

\[ \Omega_i = c_{i1}\kappa^2 + c_{i2}\kappa + P_aP_hN_i (W_i \gamma_a^2 + P_d \gamma_d^2) + P_dP_hW_i\bar{Y}_i \]
\[
\begin{bmatrix}
P + W_i & \rho_i cE & P_d + \gamma_a P_a + \gamma_h P_h \\
\rho_i cE & \tilde{Y}_i & \gamma_a \rho_i cE \\
P_d + \gamma_a P_a + \gamma_h P_h & \gamma_a \rho_i cE & P_d + \gamma_a^2 P_a + \gamma_h^2 P_h
\end{bmatrix}
\begin{bmatrix}
k'_{i1} \\
k'_{i2} \\
k'_{i3}
\end{bmatrix} =
\begin{bmatrix}
cE \\
\rho_i E \\
\gamma_a cE
\end{bmatrix}
\tag{3.46}
\]

\[
\begin{bmatrix}
k'_{i1} \\
k'_{i2} \\
k'_{i3}
\end{bmatrix} =
\frac{E}{\Gamma_i N_i + \mathbf{E} \rho_i^2 (P_h P_d \gamma_h^2 + W_i P_h \gamma_h^2 + W_i P_d)}
\begin{bmatrix}
cN_i [P_d \gamma_a + P_h \gamma_h (\gamma - \gamma_a)] \\
\rho_i [P_h \gamma_h (P_d + W_i)(\gamma - \gamma_{Costa,i})] + P_h P_d \gamma_h + P_d W_i \\
cN_i [P_h (\gamma_a - \gamma_h) - P_d \gamma_a + \gamma_a W_i]
\end{bmatrix}
\tag{3.47}
\]

and \(\gamma_{Costa,i} = \frac{P_d}{P_d + W_i}\). The resultant distortion is

\[
D_i = \mathbb{E} \{(E - \hat{E})^2\}
= E(1 - \kappa k_{i1} - ck_{i2} - \rho_i k_{i3} - \gamma_a ck_{i4})
= \frac{P_d P_h \mathbf{E} N_i W_i}{\Omega_i}.
\tag{3.45}
\]

However, if \(\kappa\) does not satisfy (3.19), since \(T_h^n\) cannot be decoded, the MMSE estimate becomes \(\hat{E}_i^n = k'_{i1} V^n + k'_{i2} \tilde{Y}_i^n + k'_{i3} T_d^n\), where the coefficients satisfy (3.46) on top of the page. We thus have (3.47) and the distortion is then given by

\[
D_i = \mathbb{E} \{(E - \hat{E})^2\}
= E(1 - ck'_{i1} - \rho_i k'_{i2} - \gamma_a ck'_{i3})
= \frac{\mathbf{E} N_i (P_h P_d \gamma_h^2 + W_i P_h \gamma_h^2 + W_i P_d)}{\Gamma_i N_i + \mathbf{E} \rho_i^2 (P_h P_d \gamma_h^2 + W_i P_h \gamma_h^2 + W_i P_d)}.
\tag{3.48}
\]
Now, we observe that substituting (the always feasible) \( \kappa = -\frac{c_i^2}{2c_{i1}} \) in (3.45) yields (3.48). In fact, since \( \kappa = -\frac{c_i^2}{2c_{i1}} \) maximizes \( D_i \) in (3.45), this implies that the worst distortion that can occur when \( T_{ih}^n \) is decodable at receiver \( i \) coincides with the distortion when \( T_{ih}^n \) is not decodable at all, just as noted in Appendix C for HDA-CDS, although the physical reason here is not as obvious.

Comparing (3.45) and (3.48) with (3.20), one can then see that the proof will be complete after showing that \( D_i \) will be minimized when \( E \) satisfies (3.13) with equality. But that easily follows by (i) the fact that the right-hand sides of both (3.45) and (3.48) are decreasing in \( \frac{1}{E} \) (for fixed \( \kappa \)), and (ii) by observing that the interval of \( \kappa \) satisfying (3.19) expands as \( \frac{1}{E} \) increases.
Chapter 4

Separate Source-Channel Coding
for Broadcasting Gaussian Sources

4.1 Introduction

Consider the problem of transmitting two correlated Gaussian sources over a broadcast channel with two receivers, each of which desires only to recover one of the sources. Previous work was mostly focusing on the Gaussian case, i.e., the case with a Gaussian broadcast channel. In [5], it was proven that analog (uncoded) transmission, the simplest possible scheme, is actually optimal when the signal-to-noise ratio (SNR) is below a threshold for the case of matched source and channel bandwidth. To solve the problem for other rates, various hybrid digital/analog (HDA) schemes have been proposed in [22, 2, 1], and [26]. In fact, the HDA scheme in [26] achieves optimal performance for matched bandwidth whenever pure analog transmission does not, thereby
leading to a complete characterization of the achievable power-distortion tradeoff. For
the bandwidth-mismatch case, the HDA schemes proposed in [2] and [1] comprise of
different combinations of previous schemes using either superposition or dirty-paper
coding.

In all the aforementioned work, authors also compared achieved performances with
that of separate source-channel coding. Since the channel is degraded, source coding
boils down to sending a “common” message to both decoders and a “refinement” message
to the decoder at the end of the better channel. In both of the two source coding schemes
proposed in [22], the first source is encoded as the common message, but one scheme
encodes (as the refinement message) the second source independently, and the other
after de-correlating it with the first source. In [26], on the other hand, the second source
is encoded after it is de-correlated with the reconstruction of the first source. Although
this approach provably yields a better performance than the schemes in [22], it is still
not optimal. In [16], it was shown that the optimal rate-distortion (RD) tradeoff in this
source coding scenario is in fact achieved by a scheme called successive coding, whereby
both common and refinement messages are generated by encoding both sources jointly,
instead of using any kind of de-correlation. Although successive coding is a special case
of successive refinement in its general sense, computation of the RD tradeoff, even for
Gaussians, turned out to be non-trivial. A Shannon-type lower bound derived for the
problem was rigorously shown to be tight, yielding an analytical characterization of the
RD tradeoff.

In this chapter, we investigate the performance of separate source and channel coding
for any rate (bandwidth compression/expansion ratio), for any degraded channel. As
discussed in the previous paragraph, the source coding method to be used for optimal
performance is successive coding. Our approach takes two directions: we first analyze
the rate penalty of separate coding over optimal joint coding, which in essence indicates
how many more channel uses we need per source symbol. A channel-independent upper
bound of the rate penalty is proposed and as shown in an example, the upper bound
performs decently for considerably large regions on the distortion plane. Via showing
that the upper bound goes to the trivial lower bound under certain conditions, separate
coding is proved to achieve optimum when either of the sources requires almost lossless
recovery and the other requires a small enough distortion. We then specialize the channel
to binary symmetric broadcast channel (BSBC) and Gaussian broadcast channel, and
show that the rate penalty can be better bounded when we know the channel. The
second direction is to analyze the power loss for the Gaussian case. We compare optimum
separate coding with other separate coding schemes, genie-aided outer bound and the
best known schemes, and show that the separate scheme is competitive in all cases.

In [25], results on the optimality of separate coding are given for general sources and
channels. By applying those results, there is another way to approach this problem,
as suggested by [24]. The idea is to take the Gaussian sources as a vector source and
generalize the individual distortion measure to covariance distortion. But we do not
pursue that line of analysis in this chapter.

4.2 Separate source-channel coding

![System model](image)

Figure 4.1: System model.
As depicted in Figure 4.1, a pair of correlated Gaussian sources \((S^n_1, S^n_2)\) are broadcast to two receivers, and receiver \(i, i \in \{1, 2\}\), is only to reconstruct \(S^n_i\). Without loss of generality, we assume the source sequences are generated in an i.i.d. fashion by 
\[ p_{S_1, S_2} = \mathcal{N}(0, C), \]
where
\[
C = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}
\]
and \(\rho \in [0, 1]\). The transmitter encodes the source sequences to \(X^m\) and thus can be described mathematically as \(X^m = \varphi(S^n_1, S^n_2)\). We define the rate (bandwidth expansion/compression ratio) \(\kappa = \frac{m}{n}\) with the unit of channel uses per source symbol.

Though the channel can be a general broadcast channel described by \(P_{Y_1Y_2|X}\), we put our focus on degraded channel in this chapter. The source coding scheme of a separate source-channel coding approach maps the source sequences to a common message with rate \(R_1\) and a private message for the strong receiver with rate \(R_2\). The channel coder then follows to encode the messages to a channel codeword \(X^m\). Without loss of generality, we assume receiver 2 is the strong receiver and receiver 1 is the weak receiver.

The well-investigated Gaussian case has two receivers, each of which has additive Gaussian white noise \(W^m_i\), given by \(W_i \sim \mathcal{N}(0, N_i)\), where we assume \(N_1 \geq N_2\). The channel output at receiver \(i\) is thus given by \(Y^m_i = X^m + W^m_i\). The Gaussian broadcast channel also has an average input power constraint, given by
\[
\frac{1}{m} \sum_{j=1}^{m} E[(X(j))^2] \leq P.
\]
The capacity region of the Gaussian channel is given in [3] and has the form of
\[
R_1 \leq \frac{1}{2} \log \left( 1 + \frac{\eta P}{\eta P + N_1} \right),
\]
\[
R_2 \leq \frac{1}{2} \log \left( 1 + \frac{\bar{\eta} P}{N_2} \right),
\]
where \(\eta \in [0, 1]\).
Another common example of degraded broadcast channel is the binary symmetric broadcast channel (BSBC), where the two receivers have crossover probabilities \( p_1 \) and \( p_2 \), respectively. We assume \( \frac{1}{2} > p_1 \geq p_2 > 0 \). The capacity region of the BSBC is given in [8] as

\[
R_1 \leq 1 - H(\beta \ast p_1)
\]

\[
R_2 \leq H(\beta \ast p_2) - H(p_2),
\]

where \( H(\cdot) \) is the binary entropy function, \( \beta \in [0, \frac{1}{2}] \), and \( a \ast b = a(1-b) + b(1-a) \).

For any channel, at receiver \( i \), \( Y_{im} \), the corrupted version of \( X_m \), is used to reconstruct source sequence \( S_{ni} \), and can be described as a function \( \hat{S}_{ni} = \phi_1(Y_{im}) \). Analogously, receiver 2 computes \( \hat{S}_{n2} = \phi_2(Y_{2m}) \). Again, in a separate coding scheme, \( \phi_i \) includes a channel decoder and a source decoder. The reconstruction quality is measured with squared-error distortion, i.e.,

\[
d(s^n, \hat{s}^n) = \frac{1}{n} \sum_{j=1}^{n} (s_j - \hat{s}_j)^2,
\]

for any source block \( s^n \) and reconstruction block \( \hat{s}^n \).

Several source coding schemes will be discussed in the following section.

4.2.1 Source Coding Schemes

When we generate the common and private messages at the source encoder with rates \( R_1 \) and \( R_2 \), we can always transfer part of \( R_2 \) to \( R_1 \) since the second receiver can still receive that part in \( R_1 \). The same distortion pair can still be achieved with this rate transfer as long as the cumulative rate \( R_1 + R_2 \) stays the same. Nevertheless, it suffices to consider only the marginal rate region instead of the cumulative rate region, as proved in [29], and thus we will mostly focus on the marginal rate regions of different source coding schemes for its simplicity.
Several separation-based schemes have been previously proposed for the Gaussian case, differing only in their source coding strategy. In the first scheme, termed Scheme A in [22], sources $S_1$ and $S_2$ are encoded as if they are independent, resulting in the source coding rate region for any $(D_1, D_2)$ given by

\[
R_1 \geq \frac{1}{2} \log \frac{1}{D_1}
\]

\[
R_2 \geq \frac{1}{2} \log \frac{1}{D_2}.
\]

Note when we (implicitly) fix $(D_1, D_2)$ and discuss the corresponding source rate regions, the source rate $R_i(D_1, D_2)$ is shorten to $R_i$ for simplicity and the same applies to sequels.

In Scheme B in [22], the second source is written as $S_2 = \rho S_1 + E$, where $S_1 \perp E$, and $S_1$ and $E$ are treated as two new independent sources. Hence we obtain

\[
R_1 \geq \frac{1}{2} \log \frac{1}{D_1}
\]

\[
R_2 \geq \left[ \frac{1}{2} \log \frac{1 - \rho^2}{D_2 - \rho^2 D_1} \right]^+,
\]

where $[x]^+ = \max\{x, 0\}$. Note by construction Scheme B can only achieve the distortion pairs satisfying $D_2 - \rho^2 D_1 \geq 0$.

In the separation-based scheme introduced in [26], which we call Scheme C, $S_1$ is quantized to $\hat{S}_1$ and $S_2$ is then encoded conditioned on $\hat{S}_1$. The resultant source coding rate region becomes

\[
R_1 \geq \frac{1}{2} \log \frac{1}{D_1}
\]

\[
R_2 \geq \left[ \frac{1}{2} \log \frac{1 - \rho^2(1 - D_1)}{D_2} \right]^+.
\]

Of the three, it is obvious that Scheme C achieves the best performance. However, it is still not optimal as we will soon show. The optimal strategy is in fact what is called successive coding in [16], whereby the sources are encoded jointly at both the common
and the refinement layers. The R-D tradeoff for successive coding of Gaussian sources with squared-error distortion was given in [16] parametrically with respect to $\alpha \in [0, 1]$ as

$$R_1(\alpha) = \frac{1}{2} \log \frac{1 - \rho^2}{D_1(1 - \nu^2 \delta_1) - (\rho - \nu \delta_1)^2}$$

$$R_2(\alpha) = \left[ \frac{1}{2} \log \frac{1 - \nu^2 \delta_1}{D_2} \right]^+,$$

where $\delta_1 = 1 - D_1$, and

$$\nu = \begin{cases} 
\nu_0, & \text{if } \nu_0^2 \delta_1 < \delta_2 \\
\nu^*, & \text{if } \rho^2 \delta_1 < \delta_2 \leq \nu_0^2 \delta_1
\end{cases}$$

with $\delta_2 = 1 - D_2$, $\nu^* = \sqrt{\frac{\delta_2}{\delta_1}}$, and $\nu_0$ is the unique root of

$$f_\alpha(\nu) = (1 - \alpha)(\rho - \nu \delta_1)(1 - \nu \rho) - \alpha(\nu - \rho)(1 - \nu^2 \delta_1)$$

in the interval $[\rho, \min(\frac{1}{\rho}, \frac{\rho}{\delta_1})]$.

We first show the RD region of successive coding can be simplified by eliminating both the parameter $\alpha$ and the need to find the roots of the cubic polynomial $f_\alpha(\nu)$.

**Lemma 13** For any distortion pair $(D_1, D_2)$, the boundary of the achievable source coding rate region is given by

$$R_1(\nu) = \frac{1}{2} \log \frac{1 - \rho^2}{D_1(1 - \nu^2 \delta_1) - (\rho - \nu \delta_1)^2}$$

$$R_2(\nu) = \left[ \frac{1}{2} \log \frac{1 - \nu^2 \delta_1}{D_2} \right]^+,$$

where $\nu \in [\rho, \nu_1]$ and $\nu_1 = \min(\frac{1}{\rho}, \frac{\rho}{\delta_1}, \nu^*)$.

**Proof.** The cubic function in [16] is

$$f_\eta(\nu) = (1 - \eta)(\rho - \nu \delta_1)(1 - \nu \rho) - \eta(\nu - \rho)(1 - \nu^2 \delta_1)$$

1When $D_2 > 1 - \rho^2(1 - D_1)$, the optimal strategy degenerates into sending only a common message and estimating $S_2^n$ solely from $\hat{S}_1^n$ for Scheme B, C and successive coding. This trivial case is excluded from the discussion in the sequel.
and when \( f_\eta(\nu) = 0 \), it can be re-written as

\[
\eta = \frac{(\rho - \nu \delta_1)(1 - \nu \rho)}{\nu(1 - \rho^2 - \delta_1(1 - 2\nu \rho + \nu^2))}.
\]

It will be shown that varying \( \eta \) in \([0, 1]\) is equivalent with varying \( \nu \) in \([\rho, \min(\frac{1}{\rho}, \frac{\delta_1}{\delta_1}, \sqrt{\frac{\delta_1}{\delta_1}})]\), by showing \( \eta \) is a monotonically decreasing function of \( \nu \).

When \( \delta_1 > \rho^2 \), \( \rho < \frac{\rho}{\delta_1} < \frac{1}{\rho} \) and also note \( \eta = 1 \) and 0 when \( \nu = \rho \) and \( \frac{\rho}{\delta_1} \), respectively. We examine

\[
h(\nu) = \frac{\rho - \nu \delta_1}{1 - \rho^2 - \delta_1(1 - 2\nu \rho + \nu^2)}
\]

instead of \( \eta \), and

\[
\frac{dh}{d\nu} \propto -1 - \rho^2 + \delta_1 + 2\nu \rho - \nu^2 \delta_1.
\]

The right hand side is a quadratic function of \( \nu \) centered at \( \nu = \frac{\rho}{\delta_1} \) and the maximum value is \(-(1 - \delta_1)(1 - \frac{\rho^2}{\delta_1}) \leq 0. \) (When \( \nu = \rho \), the function value is \(-(1 - \rho^2)(1 - \delta_1)\).)

Similarly, when \( \delta_1 < \rho^2 \), we have \( \rho < \frac{1}{\rho} < \frac{\rho}{\delta_1} \) and in this case, \( \eta = 1 \) and 0 when \( \nu = \rho \) and \( \frac{1}{\rho} \). We examine

\[
h(\nu) = \frac{1 - \nu \rho}{1 - \rho^2 - \delta_1(1 - 2\nu \rho + \nu^2)},
\]

and thus have

\[
\frac{dh}{d\nu} \propto -\rho(1 - \rho^2) - \rho \delta_1 + 2\nu \delta_1 - \nu^2 \rho \delta_1.
\]

The right hand side is centered at \( \nu = \frac{1}{\rho} \) and the maximum value is \(-(1 - \rho^2)(\rho - \frac{\delta_1}{\rho}) \leq 0. \)

\[\blacksquare\]

**Remark 14** We immediately perceive that successive coding specializes to Scheme C with \( \nu = \rho \). A closely related source coding problem is to compress two correlated Gaussian sources for one receiver with individual distortion criteria, and the problem is solved...
in [32]. It was shown in [32] that the rate-distortion function has different forms for different regions of the $(D_1, D_2)$ plane. As shown in Figure 4.2, region $D_1$ is given by $\delta_1 \delta_2 \geq \rho^2$; $D_3$ by $\delta_1 \leq \rho^2 \delta_2$ or $\delta_2 \leq \rho^2 \delta_1$ and the rest is $D_2$. These regions also play a role in our problem setting, since the feasible range of $\nu$ in Lemma 1 changes accordingly.

The resultant feasible ranges over the plane is given in Figure 4.3.

### 4.2.2 The Genie-aided Outer Bound

In [22], a genie-aided outer bound to the distortion region of the Gaussian case is obtained for $\kappa = 1$ by assuming full knowledge of $S_1$ at the second (strong) receiver. In this section, we will prove that for any $\kappa$ and any physically degraded channel, the genie-aided problem, as shown in Figure 4.4, is actually separable.

Let $R_{S_1}(D_1)$ denote the standard rate-distortion function

$$R_{S_1}(D_1) = \min_{P_{\hat{S}_1|S_1}: E[d_1(S_1, \hat{S}_1)] \leq D_1} I(S_1; \hat{S}_1),$$
Figure 4.3: Feasible ranges of $\nu$ in regions of the $(D_1, D_2)$ plane. The dashed line indicates different regimes of the genie-aided outer bound, as will be seen later.

Figure 4.4: Genie-aided system with a degraded channel.
and $R_{S_2|S_1}(D_2)$ is the conditional rate-distortion function

$$R_{S_2|S_1}(D_2) = \min_{p_{\hat{S}_2|S_2} \in \mathcal{E}(d_2(\hat{S}_2,S_2)) \leq D_2} I(S_2;\hat{S}_2|S_1).$$

Thus we have

$$R_{S_1}(D_1) = \frac{1}{2} \log \frac{1}{D_1},$$
$$R_{S_2|S_1}(D_2) = \left[ \frac{1}{2} \log \frac{1 - \rho^2}{D_2} \right]^+. $$

Let $\mathcal{C}$ be the capacity region for the degraded broadcast channel, i.e., the set of rate pairs $(R_1, R_2)$ such that there exists $p_{UX}$ satisfying

$$R_1 \leq I(U;Y_1)$$
$$R_2 \leq I(X;Y_2|U)$$

where

$$p_{UXY_1Y_2} = p_{UX}p_{Y_2|X}p_{Y_1|Y_2}. $$

**Theorem 15** The distortion pair $(D_1, D_2)$ is achievable in the genie-aided setup if and only if

$$(R_{S_1}(D_1), R_{S_2|S_1}(D_2)) \in \kappa \mathcal{C}. $$

**Proof.** The if part is true because one can simply send $R_{S_1}(D_1)$ bits/source sample as common information and $R_{S_2|S_1}(D_2)$ bits/source sample as private information to the second receiver. The issue is whether this is the best one can do, as the common information is thrown away by the second receiver. The answer is affirmative as we will prove by showing that $(D_1, D_2)$ is achievable only if

$$(R_{S_1}(D_1), R_{S_2|S_1}(D_2)) \in \kappa \mathcal{C}. $$
Towards that end, we will use the standard relations that for any $n$,

$$nR_{S_1}(D_1) \leq I(S_1^n; \hat{S}_1^n) \quad (4.5)$$

and

$$nR_{S_2|S_1}(D_2) \leq I(S_2^n; \hat{S}_2^n|S_1^n) \quad (4.6)$$

if $(D_1, D_2)$ is achieved by $(\hat{S}_1^n, \hat{S}_2^n)$.

Now, we have the following chain of inequalities:

\[
I(S_1^n; \hat{S}_1^n) \overset{(a)}{\leq} I(S_1^n; Y_1^m) = H(Y_1^m) - H(Y_1^m|S_1^n) \\
\leq \sum_{i=1}^{m} [H(Y_{1,i}) - H(Y_{1,i}|S_1^n, Y_{1,i-1}^i)] \\
\overset{(b)}{\leq} \sum_{i=1}^{m} [H(Y_{1,i}) - H(Y_{1,i}|S_1^n, Y_{1,i-1}^i, Y_{2,i-1}^i)] \\
= \sum_{i=1}^{m} I(Y_{1,i}; U_i) \quad (4.7)
\]

with $U_i = (S_1^n, Y_{1,i-1}^i, Y_{2,i-1}^i)$, where $(a)$ follows from the Markov chain $S_1^n - Y_1^m - \hat{S}_1^n$, and $(b)$ from the fact that conditioning reduces entropy. Notice that $U_i - X_1 - Y_{2,i} - Y_{1,i}$ forms a Markov chain.
Also,

\[
I(S^n_2; \hat{S}_2^n | S^n_1) \leq I(X^m_2; Y^m_2 | S^n_1) \tag{c}
\]

\[
= H(Y^m_2 | S^n_1) - H(Y^m_2 | X^m, S^n_1)
\]

\[
= \sum_{i=1}^{m} \left[ H(Y_{2,i} | S^n_1, Y_{2,i-1}) - H(Y_{2,i} | X^m, S^n_1, Y_{2,i-1}) \right] \tag{d}
\]

\[
= \sum_{i=1}^{m} \left[ H(Y_{2,i} | S^n_1, Y_{2,i-1}, Y_{i-1}^{i-1}) - H(Y_{2,i} | X^m, U_i) \right] \tag{e}
\]

\[
= \sum_{i=1}^{m} \left[ H(Y_{2,i} | U_i) - H(Y_{2,i} | X_i, U_i, X_{i+1}^{i+1}) \right] \tag{f}
\]

\[
= \sum_{i=1}^{m} I(Y_{2,i}; X_i | U_i) \tag{4.8}
\]

where (c) follows because given \(S^n_1, S^n_2 - X^m - Y^m_2 - \hat{S}^n_2\) forms a Markov chain, (d) because conditioning reduces entropy, (e) because \(Y_{2,i} - (S^n_1, Y_{2,i-1}^{i-1}) - Y_{i-1}^{i-1}\) is a Markov chain, and (f) because \(Y_{2,i} - (X_i, U_i) - (X_{i-1}^{i-1}, X_{i+1}^{i+1})\) is a Markov chain.

Bringing (4.5)-(4.8) together, we have

\[
R_{S_1}(D_1) \leq \kappa \frac{1}{m} \sum_{i=1}^{m} I(Y_{1,i}; U_i)
\]

and

\[
R_{S_2|S_1}(D_2) \leq \kappa  \frac{1}{m} \sum_{i=1}^{m} I(Y_{2,i}; X_i | U_i).
\]

Since \((I(Y_{1,i}; U_i), I(Y_{2,i}; X_i | U_i)) \in \mathcal{C}\) for all \(i\), and since \(\mathcal{C}\) is a convex region, we have the desired result. ■

The above mentioned schemes are shown in Figure 4.5, with \(D_1 = 0.5, D_2 = 0.1,\) and \(\rho = 0.8\). Note Scheme B can not achieve this distortion pair since \(D_2 - \rho^2 D_1 < 0\). The entire boundary of the successive coding source rate region actually consists of 3
segments, and with decreasing $R_2$, they are 1) the line of $R_1 = R_1(\rho)$, $R_2 > R_2(\rho)$; 2) the curve described in Lemma 1; 3) the line connecting $(R_1(\nu_1), R_2(\nu_1))$ to $(R_1(\nu_1) + R_2(\nu_1), 0)$. The first segment is degenerated and the third segment characterizes the rate-transfer region, which uses some rate of the common message to transmit extra private information for the strong receiver. The rate region of the genie-aided problem is similarly constructed except that the middle segment reduces to a point. A Gaussian broadcast channel with $P = 1$, $N_1 = 1$, $N_2 = 0.32$ is also applied to exemplify the difference between the rate regions of the genie-aided scheme and successive coding. We use the effective rate region $\kappa C$ when $\kappa = 0.3$ as the base case and fix either $\kappa$ or $P$ to define the gap between genie-aided scheme and successive coding in terms of $P$ or $\kappa$, respectively. As can be seen from the figure, when the power is fixed, the minimum $\kappa$ needed to make $\kappa C$ intersect with the genie-aided rate region is 1.75, whereas $\kappa = 2.1$ is required for successive coding. On the other hand, when $\kappa$ is fixed, the power gap is
from $P = 240$ to $P = 750$. The same gap seems more significant when it is described with power, and is simply because $\kappa$ serves as a multiplier over the logarithm of power.

In the following sections, we will look at the $\kappa$ penalty and the power loss of separate source channel coding with regard to the genie-aided scheme.

### 4.3 Rate Penalty of Separate Coding

In this section, we first give a channel-independent upper bound of the rate ($\kappa$) penalty of separate coding over optimum joint source-channel coding. Then we will analyze the specific rate penalty for BSBC and Gaussian broadcast channels.

#### 4.3.1 An Upper Bound of the Rate Loss

When we have the freedom to choose the channel bandwidth $m$, or rate $\kappa$, we have the following relation for any given source pairs and desired distortions:

$$\kappa_{\text{genie}} \leq \kappa_{\text{joint}} \leq \kappa_{\text{sep}}$$

where $\kappa_{\text{genie}}$ denotes the minimum $\kappa$ required for the corresponding genie-aided problem, $\kappa_{\text{sep}}$ is the minimum $\kappa$ for the joint source-channel coding scheme, and $\kappa_{\text{sep}}$ is the minimum $\kappa$ for separate coding. For example, in the case shown in Figure 4.5, $\kappa_{\text{genie}} = 1.75$ and $\kappa_{\text{sep}} = 2.1$. $\kappa_{\text{joint}}$ is an unknown value between the two. Therefore, we can define the rate penalty of separate coding as $\frac{\kappa_{\text{sep}}}{\kappa_{\text{joint}}}$, and since $\kappa_{\text{joint}}$ is unknown in general, we can instead use $\kappa_{\text{genie}}$ to attain an upper bound of the rate penalty.

$\kappa_{\text{sep}}$ can be obtained for any degraded channel with known capacity region. Nevertheless, we can further upper bound $\frac{\kappa_{\text{sep}}}{\kappa_{\text{genie}}}$ without knowing the channel. To achieve this upper bound, we first notice that $\kappa_{\text{genie}}$ places the genie-aided source rate pair $(R_{S_1}, R_{S_2} | S_1)$ ($D_i$ will be omitted in the following for simplicity) on the boundary of
\( \kappa_{\text{genie}} \) for any channel, since the slope of the boundary never exceeds \(-1\). Then we connect the origin to the point \((R_{S1}, R_{S2|S1})\) with a line and use the intersection between the line and the successive coding source rate curve as the replacement of the tangent point. We now have a ratio between the length of the intersection point to the origin and the length of the genie-aided source rate point to the origin, and we denote it as \( \theta \).

Obviously it satisfies
\[
1 \leq \frac{\kappa_{\text{sep}}}{\kappa_{\text{joint}}} \leq \frac{\kappa_{\text{sep}}}{\kappa_{\text{genie}}} \leq \theta .
\]

Since \( \theta \) only depends on the sources and desired distortions, it is a universal upper bound of the rate penalty independent of the channel. The value of \( \theta \) is summarized in the following lemma.

**Lemma 16** For any degraded channel, an upper bound \( \theta \) of \( \frac{\kappa_{\text{sep}}}{\kappa_{\text{joint}}} \) is given by

\[
\theta = \begin{cases} 
\theta^* & \text{if } \rho^2 < \delta_2 \text{ and } \frac{R_{S1}}{R_{S2|S1}} \leq \frac{R_1(\nu_1)}{R_2(\nu_1)} \\
\frac{R_1(\nu^*)}{R_{S1}} & \text{if } \rho^2 \delta_1 < \delta_2 \leq \rho^2 \text{ and } \rho^2 \delta_2 \leq \delta_1 \\
\frac{R_1(\frac{1}{\rho}) + R_2(\frac{1}{\rho})}{R_{S1}} & \text{if } \delta_2 \leq \rho^2 \text{ and } \delta_1 < \rho^2 \delta_2 \\
\frac{R_1(\nu_1) + R_2(\nu_1)}{R_{S1} + R_{S2|S1}} & \text{if } \rho^2 < \delta_2 \text{ and } \frac{R_{S1}}{R_{S2|S1}} > \frac{R_1(\nu_1)}{R_2(\nu_1)}
\end{cases}
\]

where \( \theta^* \) is the root of
\[
(1 - \rho^2)^2 D_2^{1-\theta} - (1 - \rho^2) D_1^\theta - \rho^2 - \delta_1 + 2\rho \sqrt{\delta_1} \left[ 1 - (1 - \rho^2)^\theta D_2^{1-\theta} \right] = 0 . \tag{4.9}
\]

**Proof.** If \( \rho^2 < \delta_2 \) and \( \frac{R_{S1}}{R_{S2|S1}} \leq \frac{R_1(\nu_1)}{R_2(\nu_1)} \), the line connecting the origin and \((R_{S1}, R_{S2|S1})\) intersects with the second segment of the successive coding source rate curve. Assume the intersection point corresponds to \( \nu = \nu_2 \) and thus we have
\[
\theta = \frac{R_1(\nu_2)}{R_{S1}} = \frac{R_2(\nu_2)}{R_{S2|S1}} .
\]
Substitute (4.3) and (4.4) in and we have

\[
\frac{1}{D_1^\theta} = \frac{1 - \rho^2}{D_1(1 - \nu_2^2 \delta_1) - (\rho - \nu_2 \delta_1)^2} \\
\left( \frac{1 - \rho^2}{D_2} \right)^\theta = \frac{1 - \nu_2^2 \delta_1}{D_2},
\]

which reduce to

\[
\theta = \log \frac{D_1(1 - \nu_2^2 \delta_1) - (\rho - \nu_2 \delta_1)^2}{D_1} \tag{4.10}
\]

\[
\theta = \log \frac{1 - \nu_2^2 \delta_1}{D_2} \tag{4.11}
\]

and then yield (4.9).

When \(\rho^2 \delta_1 < \delta_2 \leq \rho^2\) and \(\rho^2 \delta_2 \leq \delta_1\), \(R_{S_2|S_1} = 0\). Since \(\nu^*\) is feasible in this region, or \(\nu_1 = \nu^*\), it is clear that \(\nu_2 = \nu^*\) and

\[
\theta = \frac{R_1(\nu^*)}{R_{S_1}},
\]

which boils down to

\[
\theta = \log \frac{1 - \rho^2}{D_1 + D_2 - 1 - \rho^2 + 2\rho \sqrt{\delta_1 \delta_2}}.
\]

When \(\delta_2 \leq \rho^2\) and \(\delta_1 < \rho^2 \delta_2\), \(R_{S_2|S_1} = 0\) but \(\nu^*\) is not feasible and we have

\[
\theta = \frac{R_1(\frac{1}{\rho}) + R_2(\frac{1}{\rho})}{R_{S_1}}.
\]

When \(\delta_2 > \rho^2\) and \(\frac{R_{S_1}}{R_{S_2|S_1}} > \frac{R_1(\nu_1)}{R_2(\nu_1)}\), the intersection point is decided by two lines as follows

\[
R_1 = \frac{R_{S_1}}{R_{S_2|S_1}} R_2 \\
R_1 = -R_2 + R_1(\nu_1) + R_2(\nu_1),
\]

where the value of \(\nu_1\) is from the corresponding regions given in Figure 4.3. The intersection point is then

\[
\left( \frac{R_{S_1} R_1(\nu_1) + R_2(\nu_1)}{R_{S_1} + R_{S_2|S_1}}, \frac{R_{S_2|S_1} R_1(\nu_1) + R_2(\nu_1)}{R_{S_1} + R_{S_2|S_1}} \right),
\]

82
yields
\[ \theta = \frac{R_1(\nu_1) + R_2(\nu_1)}{R_{S_1} + R_{S_2|S_1}}. \]

When \( \delta_2 \leq \rho^2 \delta_1 \), it is obvious that \( \theta = 1 \). \( \blacksquare \)

As shown in the following theorem, in some cases, \( \theta \) approaches 1, implying \( \frac{R_{\text{sep}}}{R_{\text{joint}}} \to 1 \) and separate coding is actually optimal.

**Theorem 17** *Separate source-channel coding achieves the optimum performance when either of the source needs to be recovered losslessly, i.e. \( D_i \to 0 \) \( (i \in \{1, 2\}) \).*

**Proof.** According to the regions in Figure 4.2, we divide the two conditions to 4 cases as follows.

1. \( D_1 \to 0 \) and \( D_2 \leq 1 - \rho^2 \)
2. \( D_2 \to 0 \) and \( D_1 \leq 1 - \rho^2 \)
3. \( D_1 \to 0 \) and \( D_2 > 1 - \rho^2 \)
4. \( D_2 \to 0 \) and \( D_1 > 1 - \rho^2 \)

Both conditions 1) and 2) occur in region \( D_1 \), and it needs to be pointed out that in that region \( \theta = \theta^* \) because \( \nu_1 = \frac{\rho}{\delta_1} \) and we always have \( \frac{R_{S_1}}{R_{S_2|S_1}} \leq \frac{R_1(\frac{\rho}{\delta_1})}{R_2(\frac{\rho}{\delta_1})} \), which is
\[ \frac{\log \frac{1}{D_1}}{\log \frac{1-\rho^2}{D_2}} \leq \frac{\log \frac{1-\rho^2}{D_2}}{\log \frac{1-\rho^2}{D_2}}. \]

From (4.11), notice \( \theta \) can be further upper-bounded by setting \( \nu_2 = \rho \), yielding
\[ \theta \leq \frac{\log \frac{1-\rho^2\delta_1}{D_2}}{\log \frac{1-\rho^2}{D_2}}. \]

Then, for region \( D_1 \), when \( D_1 \to 0 \) or \( D_2 \to 0 \), the right hand side goes to 1 and consequently \( \theta \to 1 \).
Condition 3) is in the trivial region where we know $\theta = 1$.

Condition 4) falls in region $\mathcal{D}_3$ at the lower right side of the distortion plane, where $\nu \in \left[\rho, \frac{1}{\rho}\right]$, and as indicated in Lemma 2, $\theta$ is either $\theta^*$ or $\frac{R_1(\nu_1) + R_2(\nu_1)}{R_{S_1} + R_{S_2|S_1}}$. If $\theta = \theta^*$, the same previous arguments for 2) apply because $\nu = \rho$ is also feasible. If $\theta = \frac{R_1(\nu_1) + R_2(\nu_1)}{R_{S_1} + R_{S_2|S_1}}$, since $\nu_1 = \frac{1}{\rho}$, we have

$$\theta = \frac{R_1\left(\frac{1}{\rho}\right) + R_2\left(\frac{1}{\rho}\right)}{R_{S_1} + R_{S_2|S_1}} = \frac{\log \frac{1}{\rho}}{\log \frac{1}{\rho} - \rho},$$

which also goes to 1 when $D_2 \to 0$. ■

**Remark 18** Note when $D_1 \to 0$, the optimality of separate coding is actually natural from previous results because in this case, the strong receiver can also losslessly recover the first source, and the problem effectively reduce to the genie-aided setup. Since the genie-aided problem is proved separable, separate source-channel coding can certainly achieve optimum.

In Figure 4.6, an example of $\theta$ over the distortion plane is shown with $\rho = 0.8$. Even though the maximum value of $\theta$ is higher than 50, in the entire region of $\mathcal{D}_1$ and most of region $\mathcal{D}_2$, $\theta$ is well less than 1.5, which essentially means separate coding requires no more than 50% excess channel uses in those regions than the optimum. However, since $\theta$ is just an upper bound of the rate penalty, even in other regions, the real rate penalty may be much smaller than shown by $\theta$.

In the following sections, we will evaluate the rate penalty $\frac{\kappa_{sep}}{\kappa_{genie}}$ for BSBC and Gaussian broadcast channel.

84
Figure 4.6: The channel-independent rate penalty upper bound $\theta$ over the distortion plane. The color-value mapping is made non-linear intentionally to emphasize the values close to 1.

Figure 4.7: The rate penalty $\frac{\kappa_{\text{sep}}}{\kappa_{\text{genie}}}$ over the distortion plane of BSBC.
4.3.2 Binary Symmetric Broadcast Channel (BSBC)

For BSBC, to achieve distortion pair \((D_1, D_2)\) with separate coding, we need

\[
R_1 \leq \kappa [1 - H(\beta * p_1)] \\
R_2 \leq \kappa [H(\beta * p_2) - H(p_2)],
\]

which reduce to

\[
\kappa \geq \max \left[ \frac{R_1}{1 - H(\beta * p_1)}, \frac{R_2}{H(\beta * p_2) - H(p_2)} \right]. \tag{4.12}
\]

Since the two parts have opposite monotonicity with respect to \(\beta\), we can solve minimum \(\kappa\) by equating the two parts.

Thus the rate of the genie-aided scheme, \(\kappa_{\text{genie}}\), can be obtained by substituting the source rate pair \((R_{S_1}, R_{S_2}|S_1)\) into (4.12) and solving the \(\beta\) value equating the two parts. On the other hand, \(\kappa_{\text{sep}}\) has to be determined by searching through the \(\kappa\)'s generated by the source rate pairs given in Lemma 1.

As depicted in Figure 4.7, for the given BSBC, the rate penalty \(\frac{\kappa_{\text{sep}}}{\kappa_{\text{genie}}}\) has a maximum value of about 30, which is lower than the more general rate penalty bound \(\theta\) shown in Figure 4.6.

4.3.3 Gaussian Broadcast Channel

Similar to the BSBC case, for a Gaussian broadcast channel characterized by \((P, N_1, N_2)\) with rate \(\kappa\), we have

\[
R_1 \leq \frac{\kappa}{2} \log \left( 1 + \frac{\eta P}{\eta P + N_1} \right) \tag{4.13}
\]
\[
R_2 \leq \frac{\kappa}{2} \log \left( 1 + \frac{\bar{\eta} P}{N_2} \right), \tag{4.14}
\]
The rate penalty $\kappa_{sep}$ of a Gaussian broadcast channel is given in Figure 4.8. Again, the rate penalty of the same distortion pair is lower than the corresponding $\theta$ value as expected.

In the following section, we will analyze the power loss for the case with Gaussian broadcast channel, assuming the rate $\kappa$ fixed.

4.4 Power Loss for the Gaussian Case

In separate coding, the region of all achievable $(P, D_1, D_2)$ triplets can be determined using one of two methods. The conventional method fixes $P$ and searches for the lower envelope of all $(D_1, D_2)$ whose source rate region intersects with the capacity region given in [3]. Alternatively, we can fix $(D_1, D_2)$ and search for the minimum $P$ whose
corresponding capacity region intersects with the source rate region given in Lemma 2. We find this alternative both more convenient and more meaningful. More specifically, it is easier to compare schemes based on the minimum power they need to achieve the same distortion pair, and the ratio of minimum powers yields a single number as a quality measure.

Before we proceed, the explicit form of genie-aided outer bound of the distortion region in [1] is presented for general $\kappa$. The bound is extended from the bound in [22] for $\kappa = 1$, in the form of

$$D_1 \geq \left( 1 + \frac{\eta P}{\bar{\eta} P + N_1} \right)^{-\kappa}$$

$$D_2 \geq (1 - \rho^2) \left( 1 + \frac{\bar{\eta} P}{N_2} \right)^{-\kappa},$$  

(4.15)  

(4.16)

where $\eta \in [0, 1]$ and $\bar{\eta} = 1 - \eta$.

To be able to use the power as the measure, first we need to find out the minimum required power for any given source coding rate pair $(R_1, R_2)$.

**Lemma 19** For any source coding rate pair $(R_1, R_2)$, the minimal required power is given by

$$P(R_1, R_2) = N_1 \left( 2^{2R_1/\kappa} - 1 \right) + N_2 \left( 2^{2R_2/\kappa} - 1 \right) 2^{2R_1/\kappa}.$$  

(4.17)

**Proof.** Inequalities (4.13) and (4.14) imply that $P$ is achievable if and only if there exists $0 < \bar{\eta} < 2^{-2R_1/\kappa}$ such that

$$P \geq \max \left\{ \frac{N_1 \left( 2^{2R_1/\kappa} - 1 \right)}{1 - \bar{\eta} 2^{2R_1/\kappa}}, \frac{N_2 \left( 2^{2R_2/\kappa} - 1 \right)}{\bar{\eta}} \right\}.$$  

Since the terms in the maximum exhibit opposite monotonicity with respect to $\bar{\eta}$ with asymptotes at $\bar{\eta} = 0$ and $\bar{\eta} = 2^{-2R_1/\kappa}$, the minimum power is achieved when the two terms are equal, that is, when

$$\bar{\eta} = \frac{N_2 \left( 2^{2R_2/\kappa} - 1 \right)}{N_1 \left( 2^{2R_1/\kappa} - 1 \right) + N_2 \left( 2^{2R_2/\kappa} - 1 \right) 2^{2R_1/\kappa}}.$$  

88
and has the form in (4.17).

By substituting (4.3) and (4.4) into (4.17), we obtain the minimum power required for the separate coding scheme as a function of $\nu$:

$$P(\nu) = N_1 \left( \left[ \frac{1 - \rho^2}{D_1(1 - \nu^2 \delta_1) - (\rho - \nu \delta_1)^2} \right]^{1/\kappa} - 1 \right)$$

$$+ N_2 \left( \frac{1 - \nu^2 \delta_1}{D_2} \right)^{1/\kappa} \left[ \frac{1 - \rho^2}{D_1(1 - \nu^2 \delta_1) - (\rho - \nu \delta_1)^2} \right]^{1/\kappa}.$$  (4.18)

For bandwidth-matched case, the minimum power of separate coding $P_{\text{sep}} = \min_\nu P(\nu)$ can actually be found analytically for any $(D_1, D_2)$.

Lemma 20 When the bandwidth are matched or $\kappa = 1$, the minimal required power for any $(D_1, D_2)$ is

$$P(\nu_M) = \frac{N_1 \delta_1 D_2 (1 - 2 \nu_M \rho + \nu_M^2) + N_2 (1 - \rho^2) (\delta_2 - \nu_M^2 \delta_1)}{D_2 [(1 - \rho^2) (1 - \delta_1) - \delta_1 (\nu_M - \rho)^2]},$$

where $\nu_M = \min(\nu_L, \nu^*)$, $\nu_L = \frac{N_2 (\rho^2 + \delta_1) + D_2 (N_1 - N_2) - \sqrt{\Delta}}{2 N_2 \delta_1 \rho}$ and $\Delta = [N_2 (\rho^2 - \delta_1) + D_2 (N_1 - N_2)]^2 + 4 N_2 \delta_1 D_2 (N_1 - N_2) (1 - \rho^2)$.

The proof is omitted here.

The following theorem is the counterpart of Theorem 2 for Gaussian broadcast channel in terms of power instead of $\kappa$.

Theorem 21 Separate source-channel coding achieves optimal power-distortion tradeoff when $(D_1, D_2)$ satisfies either of the following conditions

1. $D_1 \to 0$ and $D_2 \leq 1 - \rho^2$ ,

2. $D_2 \to 0$ and $D_1 \leq 1 - \rho^2$.

Proof. We first find the minimum power the genie-aided outer bound (4.15) and (4.16) requires. Note that when $D_2 > 1 - \rho^2$, (4.16) will hold for any $\eta \in [0, 1]$, and hence the
minimum power is obtained solely from (4.15), whereas when \( D_2 \leq 1 - \rho^2 \), the minimum power satisfies equality in both (4.15) and (4.16). Combining the two cases, we obtain the concise expression

\[
P_{\text{genie}} = N_2 D_1^{-1/\kappa} \left[ \left( \frac{D_2}{1 - \rho^2} \right)^{-1/\kappa} - 1 \right]^+ + N_1 (D_1^{-1/\kappa} - 1) .
\]

On the other hand, from (4.18), we have

\[
P(\rho) = N_2 D_1^{-1/\kappa} \left[ \left( \frac{D_2}{1 - \rho^2 \delta_1} \right)^{-1/\kappa} - 1 \right] + N_1 (D_1^{-1/\kappa} - 1) .
\]

Since \( \delta_1 = 1 - D_1 \), it is easy to see that when \( D_2 \leq 1 - \rho^2 \)

\[
\frac{P(\rho)}{P_{\text{outer bound}}} \to 1 \quad \text{as} \quad D_1 \to 0 ,
\]

and since \( \nu = \rho \) is feasible, the minimum power of separate coding satisfies \( P_{\text{sep}} \leq P(\rho) \).

Therefore the performance of separate coding scheme approaches the outer bound, or

\[
\frac{P_{\text{sep}}}{P_{\text{outer bound}}} \to 1, \quad \text{when} \quad D_2 \leq 1 - \rho^2 \quad \text{and} \quad D_1 \to 0 .
\]

Similarly, by setting \( \nu = \frac{\rho}{\delta_1} \), we have

\[
P \left( \frac{\rho}{\delta_1} \right) = N_1 \left[ \left( \frac{1 - \rho^2}{D_1 (1 - \rho^2 \delta_1)} \right)^{1/\kappa} - 1 \right] + N_2 \left[ \left( \frac{1 - \rho^2}{D_2} \right)^{1/\kappa} - 1 \right] \left[ \frac{1 - \rho^2}{D_1 (1 - \rho^2 \delta_1)} \right]^{1/\kappa}
\]

and when \( D_2 \to 0 \), \( \frac{P(\frac{\rho}{\delta_1})}{P_{\text{genie}}} \to 1 \). Note when \( \delta_1 \delta_2 \geq \rho^2 \), \( \min \left( \frac{1}{\rho}, \frac{\rho}{\delta_1}, \nu^* \right) = \frac{\rho}{\delta_1} \), which again implies \( P_{\text{sep}} \leq P(\frac{\rho}{\delta_1}) \), thus proving the second part of the theorem. ■

**Remark 22** Here we proved that the genie-aided outer bound is tight in the region of \( (1 - D_1)(1 - D_2) \geq \rho^2 \) when either \( D_1 \) or \( D_2 \) goes to 0, and the performance of separate coding approaches the outer bound. The condition that either \( D_1 \) or \( D_2 \) goes to 0 translates to infinite channel SNR.

Again, the first condition is a natural result just as in Theorem 17.
In fact, as we show in the following theorem, separate coding is approximately optimal for the entire region \( \delta_1 \delta_2 \geq \rho^2 \), in the sense that the power ratio \( \frac{P_{\text{sep}}}{P_{\text{genie}}} \) can be upper-bounded universally in \((N_1, N_2, D_1, D_2)\).

**Theorem 23** When \( \delta_1 \delta_2 \geq \rho^2 \),

\[
\frac{P_{\text{sep}}}{P_{\text{genie}}} \leq \min \left\{ 1 + \frac{(1 + \rho)^{1/\kappa} [1 + \rho^2]^{1/\kappa} - 1}{(1 + \rho)^{1/\kappa} - (1 - \rho)^{1/\kappa}}, \frac{(1 + \rho)^{2/\kappa} - 1}{1 + 2\rho} \right\}.
\]

**Proof.** The first bound is true because

\[
\frac{P_{\text{sep}}}{P_{\text{genie}}} \leq \frac{P(\rho)}{P_{\text{genie}}} \leq \frac{N_2}{N_1} \left[ (1 - \rho^2) \delta_1^{1/\kappa} - D_2^{1/\kappa} \right] + D_2^{1/\kappa} \left( 1 - D_1^{1/\kappa} \right)
\]

\[
= \frac{N_2}{N_1} \left[ (1 - \rho^2) \delta_1^{1/\kappa} - D_2^{1/\kappa} \right] + D_2^{1/\kappa} \left( 1 - D_1^{1/\kappa} \right)
\]

\[
\leq \frac{(1 - \rho^2) \delta_1^{1/\kappa} - (D_1 D_2)^{1/\kappa}}{(1 - \rho^2)^{1/\kappa} - (D_1 D_2)^{1/\kappa}}
\]

\[
\leq \frac{(1 - \rho^2) \delta_1^{1/\kappa} - (1 - \rho)^{2/\kappa}}{(1 - \rho^2)^{1/\kappa} - (1 - \rho)^{2/\kappa}}
\]

\[
\leq 1 + \frac{(1 + \rho)^{1/\kappa} [(1 + \rho^2)^{1/\kappa} - 1]}{(1 + \rho)^{1/\kappa} - (1 - \rho)^{1/\kappa}}
\]

where (a) follows since \( \nu = \rho \) is feasible, (b) by \( \frac{N_2}{N_1} \leq 1 \), (c) follows since \( D_1 D_2 \leq (1 - \rho)^2 \), and (d) since \( \delta_1 \geq \rho^2 \).

Since \( \frac{\rho}{\sqrt{\delta_1}} \) is feasible as

\[
\rho \leq \frac{\rho}{\sqrt{\delta_1}} \leq \frac{\rho}{\delta_1}
\]

by relaxing \( P_{\text{sep}} \) to \( P \left( \frac{\rho}{\sqrt{\delta_1}} \right) \), the second bound can be obtained in a similar way as follows. First, we have

\[
P \left( \frac{\rho}{\sqrt{\delta_1}} \right) = N_1 \left[ D_1^{-1/\kappa} \left( \frac{1 - \rho^2}{1 - 2\rho^2} \right)^{1/\kappa} - 1 \right] + N_2 \left[ \left( \frac{1 - \rho^2}{D_2} \right)^{1/\kappa} - 1 \right] \left( \frac{1 - \rho^2}{1 - 2\rho^2} \right)^{1/\kappa}.
\]
Thus
\[
\frac{P_{\text{sep}}}{P_{\text{genie}}} \leq \frac{D_1^{-1/\kappa} \left( \frac{1 - \rho^2}{1 - \frac{2\rho^2}{1 + \sqrt{\delta_1}}} \right)^{1/\kappa} - 1 + \frac{N_2}{N_1} D_1^{-1/\kappa} \left[ \left( \frac{1 - \rho^2}{\rho_2} \right)^{1/\kappa} - 1 \right] \left( \frac{1 - \rho^2}{1 - \frac{2\rho^2}{1 + \sqrt{\delta_1}}} \right)^{1/\kappa}}{D_1^{-1/\kappa} - 1 + \frac{N_2}{N_1} D_1^{-1/\kappa} \left[ \left( \frac{1 - \rho^2}{\rho_2} \right)^{1/\kappa} - 1 \right]}
\]
\[
= \left( \frac{1 - \rho^2}{1 - \frac{2\rho^2}{1 + \sqrt{\delta_1}}} \right)^{1/\kappa} + \frac{\left( \frac{1 - \rho^2}{1 - \frac{2\rho^2}{1 + \sqrt{\delta_1}}} \right)^{1/\kappa} - 1}{D_1^{-1/\kappa} - 1}
\]
\[
\leq \left( \frac{1 - \rho^2}{1 - \frac{2\rho^2}{1 + \sqrt{\delta_1}}} \right)^{1/\kappa} + \frac{\left( \frac{1 - \rho^2}{1 - \frac{2\rho^2}{1 + \sqrt{\delta_1}}} \right)^{1/\kappa} - 1}{(1 - \rho^2)^{-1/\kappa} - 1}
\]
\[
= \left[ \frac{(1 + \rho)^2}{1 + 2\rho} \right]^{1/\kappa} + \frac{\left[ \frac{(1 + \rho)^2}{1 + 2\rho} \right]^{1/\kappa} - 1}{(1 - \rho^2)^{-1/\kappa} - 1}
\]
where the last inequality comes from the fact that \(D_1 = 1 - \rho^2\) maximizes the right-hand side.

Combining the two bounds, we have the result in the theorem. ■

Figure 4.9: The upper bounds of the power difference in dB. Bound 1 in the figure is the first bound in Theorem 23 and Bound 2 is the second.
A typical example of the performance of the bounds is given in Figure 4.9. As can be seen in the figure, there is about less than 1dB gap between the bounds and the maximum power difference. In general, the first bound is tighter for smaller $\rho$ values, whereas the second is tighter for larger $\rho$ values.

In the following section, we compare our separate coding scheme with various other schemes over the entire distortion plane.

### 4.4.1 Performance comparison

#### 4.4.1.1 Separation-based schemes

![Comparison between the power of genie-aided outer bound, Scheme C and optimal separate coding. \( \rho = 0.8, N_1 = 1, N_2 = 0.5, \kappa = 0.3 \).](image)

(a) Power difference between separate coding and the outer bound.

(b) Power difference between Scheme C and optimal separate coding.

Figure 4.10: Comparison between the power of genie-aided outer bound, Scheme C and optimal separate coding. $\rho = 0.8, N_1 = 1, N_2 = 0.5$ and $\kappa = 0.3$.

As illustrated in [26] for $\kappa = 1$, the outer bound in (4.15) and (4.16) is not always tight. Nevertheless, we can still compare $P_{\text{sep}}$ and $P_{\text{genie}}$ for any $(D_1, D_2)$, which provides an upper bound to the ratio of the minimum separate and joint coding power levels. To show the optimality of our separate coding scheme, we also compare our scheme
with Scheme C, which provides the best performance among the three separation-based schemes mentioned earlier. The minimum required power of Scheme C can be obtained from (4.1) and (4.2) as

\[ P_C = N_2 D_1^{-1/\kappa} \left( \left( \frac{D_2}{1 - \rho^2 \delta_1} \right)^{-1/\kappa} - 1 \right) + N_1 (D_1^{-1/\kappa} - 1) . \]

Note that \( P_C = P(\rho) \) in the non-trivial distortion regions.

As an example with bandwidth compression, we show the power ratio between our separate coding scheme and the outer bound in Figure 4.10(a), and that between Scheme C and our separate coding scheme in Figure 4.10(b), both in dB. For reference, the black curves illustrate the different distortion regions in [32]. The lower left corner region is actually \( \delta_1 \delta_2 \geq \rho^2 \), where, in general, small dB differences are observed, as implied by the two theorems above. As can be seen from the figure, even for highly correlated sources, the optimum separate coding scheme does not require too much extra power in most of the \((D_1, D_2)\) plane. Again, since the outer bound is not always tight, the large power difference in some regions may be dramatically reduced when the outer bound is replaced by the optimum performance. For smaller \( \rho \) values, we observe that the power difference is very small in the entire plane, as an example illustrates in the next section. This is a natural result because separate coding is optimal for independent sources and small \( \rho \) value means the sources are not highly dependent.

There is also noticeable power difference between our scheme and Scheme C, and we numerically observe large dB values near the point \( D_1 = 1 - \rho^2 \) and \( D_2 = 0 \), for which we can obtain analytically the power ratio.

**Theorem 24** When \( D_1 = 1 - \rho^2 \) and \( D_2 \to 0 \),

\[ \frac{P_C}{P_{sep}} \to (1 + \rho^2)^{1/\kappa} . \]
The proof uses the fact that the power of our separate coding scheme goes to that of the outer bound when approaching this point.

When the two receivers have the same noise level, i.e., \( N_1 = N_2 \), it can be shown that our separate coding scheme achieves the corresponding rate-distortion function \( R(D_1, D_2) \) in [32], whereas none of the three separate coding schemes mentioned earlier has the same performance.

### 4.4.1.2 Hybrid Digital/Analog (HDA) schemes

![Figure 4.11: Comparison between the genie-aided outer bound, RFZ scheme in [1] and separate coding. \( P = 3dB = 1.995, N_1 = 0dB = 1, N_2 = -5dB = 0.3162 \) and \( \kappa = 2 \).](image)

In [2] and [1], a group of hybrid digital/analog (HDA) schemes were proposed for bandwidth-mismatched case, where analog, digital, and hybrid schemes are layered with superposition or dirty-paper-coding. The achievable distortion region can be found by varying power allocation and scaling coefficients. In [1], an HDA scheme from [19] for broadcasting a common source with bandwidth expansion was adapted for the problem of broadcasting correlated sources and is termed the RFZ scheme. In addition, a scheme, termed the HWZ scheme, containing an analog layer and two digital layers each with a
Wyner-Ziv coder and a channel coder, was also proposed. It is argued by an example in [1] that the HWZ scheme performs similar to the RFZ scheme. Here we compare our separate coding scheme with the outer bound and the RFZ/HWZ scheme in Figure 4.11. For this comparison, we revert to the more familiar $(D_1, D_2)$ plot for the exact same $(P, \rho, \kappa, N_1, N_2)$ as those used in the examples in [1].

As seen in Figure 4.11(a), when $\rho$ is small, the separate coding scheme almost coincides with the outer bound and outperforms RFZ/HWZ schemes. When the sources are highly correlated as in Figure 4.11(b), the separate coding scheme is still better than the RFZ/HWZ schemes when $D_2$ is lower than a certain value, and also provides competitive performance when it is higher. We observed similar performance behavior when we compared the separate coding scheme to the layered schemes in [2] for bandwidth compression.

We also conduct power comparison for the entire distortion plane for the same set of parameters between RFZ scheme and separate coding. As can be seen from the figures, when $\rho = 0.2$, separate coding is very close to the genie-aided outer bound and is better than RFZ scheme in most of the distortion regions. Even in the region where RFZ scheme performs better (marked with white dashed lines), the difference is rather little. When $\rho = 0.8$, again, RFZ scheme does not outperform separate coding much and in most of the regions, separate coding performs closely to the genie-aided outer bound.

4.5 Conclusion

The performance of optimum separate source-channel coding scheme for broadcasting two correlated Gaussians is analyzed. We propose a channel-independent upper bound

---

2. The performance of RFZ is plotted to represent both schemes as their curves almost coincide at least for this set of parameters in [1].
of the rate penalty of separate coding over optimum joint coding, and prove that separate coding actually achieves optimum under two conditions. The minimum power required for a given distortion pair is used as a tool to compare performances between separate coding and other previous schemes. In a certain low distortion region, the power loss is analytically bounded. It is illustrated that this separate coding scheme outperforms other known separate schemes, and is competitive with best-known HDA schemes.

Figure 4.12: Comparison of power between the genie-aided outer bound, RFZ scheme in [1] and separate coding. $P = 3dB = 1.995$, $N_1 = 0dB = 1$, $N_2 = -5dB = 0.3162$ and $\kappa = 2$. The negative regions are shown with white dashed lines.
Chapter 5

Conclusions

In this dissertation, first a new level of freedom in designing transmission schemes for Gaussian sources over Gaussian channels is revealed. It is shown that the optimum distortion can be achieved with various schemes with a continuum of dirty-paper codeword by making good use of the decoded dirty-paper codeword at the receiver. This freedom plays an important role in the HDA-CDS scheme proposed for the Wyner-Ziv coding over broadcast channel problem. HDA-CDS scheme achieves the trivial outer bound for all the systems in the region of the parameter space sandwiched between the optimality conditions for HDA-WZ and CDS schemes, and the upgraded AHC scheme outperforms all known schemes.

The optimal separation based scheme is analyzed for broadcasting two correlated Gaussian sources over a Gaussian channel. It is shown that the performance of separate coding is competitive to the best known hybrid schemes for bandwidth-mismatch cases, and in fact achieves optimality under certain conditions. The performance loss of separate coding is analytically upper bounded in terms of power and rate loss, though the optimal joint scheme remains unknown.
There are still many open problems to be worked on, even for the problems discussed in this dissertation. In this dissertation, it is proved that for WZBC problem, when the channel quality and side information quality are matched in some sense, the two receivers can achieve simultaneous optimum. A natural question would be whether this is possible for all channel quality and side information quality combinations. Also, when there are more than two receivers, it is unclear whether the receivers can achieve their own optimum at the same time with a single scheme. Since all the discussions so far are under the assumption of matched bandwidth, the same questions would raise again when the bandwidth are not required to match.

For the problem of broadcasting correlated Gaussian sources, the special case with matched bandwidth has been solved with a hybrid joint scheme, but the optimal scheme in general is still open.
Bibliography


