

Approximate Characterizations for the Gaussian Source Broadcast Distortion Region

Chao Tian, *Member, IEEE*, Suhas Diggavi, *Member, IEEE*,
and Shlomo Shamai (Shitz), *Fellow, IEEE*

Abstract—We consider the joint source-channel coding problem of sending a Gaussian source on a K -user Gaussian broadcast channel with bandwidth mismatch. A new outer bound to the achievable distortion region is derived using the technique of introducing more than one additional auxiliary random variable, which was previously used to derive sum-rate lower bound for the symmetric Gaussian multiple description problem. By combining this outer bound with the achievability result based on source-channel separation, we provide approximate characterizations of the achievable distortion region within constant multiplicative factors. Furthermore, we show that the results can be extended to general broadcast channels, and the performance of the source-channel separation based approach is also within the same constant multiplicative factors of the optimum.

Index Terms—Gaussian source, joint source-channel coding, squared error distortion.

I. INTRODUCTION

Shannon's source-channel separation theorem essentially states that asymptotically there is no loss from optimum by decoupling the source coding component and channel coding component in a point-to-point communication system [1]. This separation result tremendously simplifies the concept and design of communication systems, and it is also the main reason for the division between research in source coding and channel coding. However, it is also well known that in many multi-user settings, such a separation indeed incurs certain performance loss; see, *e.g.*, [2]–[4]. For this reason, joint source-channel coding has attracted an increasing amount of attention as the communication systems become more and more complex.

One of the most intriguing problems in this area is joint source-channel coding of a Gaussian source on a Gaussian broadcast channel with K users under an average power constraint. It was observed by Gobbleck [2] that when the source bandwidth and the channel bandwidth are matched, *i.e.*, one channel use per source sample, directly sending the source samples on the channel after a simple scaling is in fact optimal,

but the separation-based scheme suffers a performance loss [4]. However, when the source bandwidth and the channel bandwidth are not matched, such a simple scheme is no longer optimal. Many researchers have considered this problem, and significant progress has been made toward finding better coding schemes based on hybrid digital and analog signaling; see, *e.g.*, [5]–[10] and the references therein.

In spite of the progress on the achievability schemes, our overall understanding on this problem is still quite limited. As pointed out by Caire [11], the key difficulty appears to be finding meaningful outer bounds. Such outer bounds not only can provide a concrete basis to evaluate various achievability schemes, but also may provide insights into the structure of good or even optimal codes, and may further suggest simplification of the possibly quite complex optimal schemes in certain distortion regimes. In this regard, the result by Reznic *et al.* [8] is particularly important, where they derived a non-trivial outer bound for the achievable distortion region for the two-user system. This outer bound relies on a technique previously used in the multiple description problem by Ozarow [12], where one additional random variable beyond those in the original problem is introduced. The bound given in [8] is however rather complicated, and was only shown to be asymptotically tight for certain high signal to noise ratio regime.

In this work, we derive an outer bound for the K -user problem using a similar technique as that used in [8], however, more than one additional random variable is introduced. The technique used here also bears some similarity to that used in [13]. The outer bound has a more concise form than the one given in [8], but for the $K = 2$ case, it can be shown that they are equivalent. This outer bound is in fact a set of outer bounds parametrized by $K - 1$ non-negative variables. Though one can optimize over these variables to find the tightest one, this optimization problem appears difficult. Thus we take an approach similar to the one taken in [13], and choose some specific values for the $K - 1$ variables which gives specific outer bounds. Moreover, by combining these specific outer bounds with the simple achievability scheme based on source-channel separation, we provide approximate characterizations¹ of the achievable distortion region within some universal constant multiplicative factors, independent of the signal to noise ratio and the bandwidth mismatch factor. In one of the approximations, the multiplicative factor is roughly

The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Seoul, Korea, June-July 2009.

The work of S. Shamai was supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless Communications, NEWCOM++.

C. Tian is with AT&T Labs-Research, Florham Park, NJ 07932, USA. (email: tian@research.att.com)

S. N. Diggavi is with the Department of Electrical Engineering, University of California, Los Angeles, CA 90095, USA. (email: suhas@ee.ucla.edu)

S. Shamai is with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel. (email: sshlomo@ee.technion.ac.il)

¹We would like to thank David Tse for discussions in ITA 2008 on the formulation of the question, where he was motivated by his solution to a deterministic version of this problem.

of form 2^k for the distortion at the k -th user, while in the other, the factor is K for all the distortions. Thus although Shannon's source-channel separation result does not hold strictly in this problem, it indeed holds in an approximate manner. In fact, this set of results is extremely flexible, and can be applied in the case with an infinite number of users but the minimum achievable distortion is bounded away from zero, for which we can conclude that the source-channel separation based approach is also within certain finite constant multiplicative factors of the optimum. In this case, these constants can be upper bounded by factors related to the disparity between the best and worse distortions, which is not influenced by the number of users being infinite.

Though the outer bound is derived using techniques that have some precedents in the information theory literature, the difficulty lies in determining which terms to bound. In contrast to pure source coding problems or pure channel coding problems, where we can usually meaningfully bound a linear combination of rates, in a joint source-channel coding problem the notion of rates does not exist. In [8], the lower bound on one distortion is given in terms of a function of the other distortion in the two-user problem. It is clear that such a proof approach becomes unwieldy for the general K -user case. In this work, we instead derive bounds for some quantity which at the first sight may seem even unrelated to the problem, but eventually serves as an interface between the source and channel coding components, thus replacing the role of "rates" in traditional Shannon theory proofs.

Inspired by a recent work of Avestimehr, Caire and Tse [14], where source-channel separation in more general networks is considered, we further show that our technique can be conveniently extended to general broadcast channels, and the source-channel separation based scheme is within the same multiplicative constants of the optimum as for the Gaussian channel case.

The rest of the paper is organized as follows. Section II gives the necessary notation and reviews an important lemma useful in deriving the outer bound. The main results are presented in Section III, and the proofs for these results are given in Section IV. The extension to general broadcast channels is given in Section V, and Section VI concludes the paper.

II. PROBLEM DEFINITION AND REVIEW

In this section, we give a formal definition of the Gaussian source broadcast problem in the context of Gaussian broadcast channels; the notation will be generalized in Section V when other broadcast channels are considered.

Let $\{S(i)\}_{i=1,2,\dots}$ be a stationary and memoryless Gaussian source with zero-mean and unit-variance. The vector $(S(1), S(2), \dots, S(m))$ will be denoted as S^m . We use \mathbb{R} to denote the domain of reals, and \mathbb{R}_+ to denote the domain of non-negative reals. The Gaussian memoryless broadcast channel is given by the model

$$Y_k = X + Z_k, \quad k = 1, 2, \dots, K, \quad (1)$$

where Y_k is the channel output observed by the k -th receiver, and Z_k is the zero-mean additive Gaussian noise on the

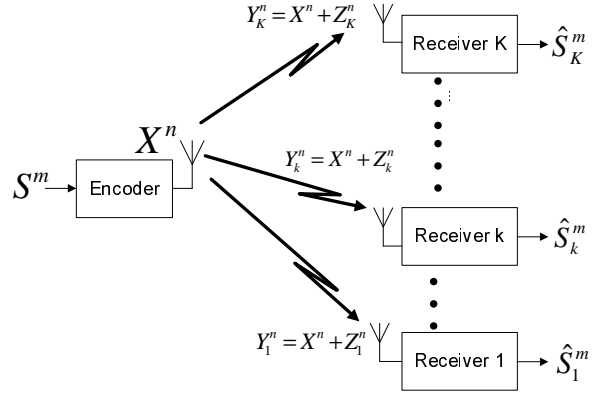


Fig. 1. Joint source-channel coding with bandwidth mismatch.

channel input X . Thus the channel is memoryless in the sense that $(Z_1(i), Z_2(i), \dots, Z_K(i))_{i=1,2,\dots}$ is a stationary and memoryless process. The variance of Z_k is denoted as N_k , and without loss of generality, we shall assume

$$N_1 \geq N_2 \geq \dots \geq N_K. \quad (2)$$

The mean squared error distortion measure is used, which is given by $d(s^m, \hat{s}^m) = \frac{1}{m} \sum_{i=1}^m (s(i) - \hat{s}(i))^2$. The encoder maps a source sample block of length m into a channel input block of length n , and each decoder maps the corresponding channel output block of length n into a source reconstruction block of length m . The bandwidth mismatch factor is thus defined as

$$b = \frac{n}{m}, \quad (3)$$

which is essentially the (possibly fractional) channel uses per source sample; see Fig. 1. The channel input is subject to an average power constraint.

We can make the codes in consideration more precise by introducing the following definition.

Definition 1: An $(m, n, P, d_1, d_2, \dots, d_K)$ Gaussian source-channel broadcast code is given by an encoding function

$$f : \mathbb{R}^m \rightarrow \mathbb{R}^n, \quad (4)$$

such that

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}(X(i))^2 \leq P, \quad (5)$$

and K decoding functions

$$g_k : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad k = 1, 2, \dots, K, \quad (6)$$

and their induced distortions

$$d_k = \mathbb{E}d(S^m, g_k(f(S^m) + Z_k^n)), \quad k = 1, 2, \dots, K, \quad (7)$$

where $\mathbb{E}(\cdot)$ is the expectation operation.

Note that there are two kinds of independent randomness in the system, the first of which is by the source, and the second is by the channel noises; the expectation operation in (7) is taken over both of them. In the definition, $+$ in the expression $f(S^m) + Z_k^n$ is understood as the length- n vector addition.

From the above definition, it is clear that the performance of any Gaussian joint source-channel code depends only on the marginal distribution of (S^m, X^n, Y_k^n) , but not the joint distribution $(S^m, X^n, Y_1^n, Y_2^n, \dots, Y_K^n)$. This implies that physical degradedness does not differ from statistical degradedness in terms of the system performance. Since the Gaussian broadcast channel is always statistically degraded, we shall assume physical degradedness from here on without loss of generality. The channel noises can thus be written as

$$Z_k = Z_{k+1} + \Delta Z_k, \quad k = 1, 2, 3, \dots, K, \quad (8)$$

where ΔZ_k is a zero-mean Gaussian random variable with variance $\Delta N_k = N_k - N_{k+1}$, which is independent of everything else; for convenience, we define $Z_{K+1} \triangleq 0$, and it follows that $\Delta N_K = N_K$ and $Y_{K+1} = X$.

Definition 2: A distortion vector $(D_1, D_2, \dots, D_K) \in \mathbb{R}_+^K$, where $1 \geq D_1 \geq D_2 \geq \dots \geq 0$ is achievable under power constraint P and bandwidth mismatch factor b , if for any $\epsilon > 0$ and sufficiently large m , there exist an integer $n \leq bm$ and an $(m, n, P, d_1, d_2, \dots, d_K)$ Gaussian source-channel broadcast code such that

$$D_i + \epsilon \geq d_i, \quad i = 1, 2, \dots, K. \quad (9)$$

Note that the constraint $1 \geq D_1 \geq D_2 \geq \dots \geq 0$ is without loss of generality, because otherwise the problem can be reduced to an alternative one with fewer users due to the assumed physical degradedness. The collection of all the achievable distortion vectors under power constraint P and bandwidth mismatch factor b is denoted by $\mathcal{D}(P, b)$, and this is the region that we are interested in.

One important result we need in this work is the following lemma, which is a slightly different version of the one given in [13].

Lemma 1: Let W be a random variable jointly distributed with the Gaussian source vector S^m in the alphabet \mathcal{W} , such that there exists a deterministic mapping $g : \mathcal{W} \rightarrow \mathbb{R}^m$ satisfying

$$\mathbb{E}d(S^m, g(W)) \leq D. \quad (10)$$

Let $U = S + V$ and $U' = S + V + V'$, where V and V' are mutually independent Gaussian random variables independent of the Gaussian source S and the random variable W , with variance σ^2 and σ'^2 , respectively. Then with $\sigma^2 \triangleq \tau$ and $\sigma^2 + \sigma'^2 \triangleq \tau'$, we have

1) Mutual information bound

$$I(W; U'^m) \geq \frac{m}{2} \log \frac{1 + \tau'}{D + \tau'}, \quad (11)$$

2) Bound on mutual information difference

$$I(W; U^m) - I(W; U'^m) \geq \frac{m}{2} \log \frac{(1 + \tau)(D + \tau')}{(1 + \tau')(D + \tau)}. \quad (12)$$

The proof of this lemma is almost identical to the one given in [13]. The only difference between the two versions is that in [13] the random variable W is in fact a deterministic function of S^m , however it is rather straightforward to verify that this condition was never used in the proof given in [13]; we include

the proof of this lemma in the Appendix for completeness.

III. MAIN RESULTS FOR GAUSSIAN BROADCAST CHANNELS

Our main results for Gaussian source broadcast on Gaussian broadcast channels are summarized in Theorem 1, Corollary 1, Proposition 1, Corollary 2 and Corollary 3, the proofs of which are given in the next section; extensions of these results to general broadcast channels are given in Section V.

Define the region in (13) on the top of next page, which is in fact the inner bound via source-channel separation. Next define the regions in (14) and (15) also on the top of next page, which are in fact outer bounds to the achievable distortion region. We have the the following theorem.

Theorem 1:

$$\hat{\mathcal{D}}(P, b) \subseteq \mathcal{D}(P, b) \subseteq \underline{\mathcal{D}}^*(P, b) \cap \underline{\mathcal{D}}(P, b). \quad (16)$$

Theorem 1 is stated as inner and outer bounds to the achievable distortion region, however it can be observed that the bounds have similar forms, and their difference, in terms of distortions, can be bounded by some multiplicative constants. The following corollary follows directly from Theorem 1, by comparing (13) and (14).

Corollary 1: If $(D_1, D_2, \dots, D_K) \in \mathcal{D}(P, b)$, and if $D_k \geq 2D_{k+1}$ for $k = 1, 2, \dots, K - 1$, then $(2D_1, 2^2D_2, \dots, 2^K D_K) \in \hat{\mathcal{D}}(P, b)$.

The condition $D_k \geq 2D_{k+1}$ in Corollary 1 is to ensure that the distortion vector $(2D_1, 2^2D_2, \dots, 2^K D_K)$ satisfies the monotonicity requirement in Definition 2 and (13). This result has the following intuitive interpretation if the condition indeed holds that $D_k \geq 2D_{k+1}$ for all $k = 1, 2, \dots, K - 1$: if a genie helps the separation-based scheme by giving each individual user half a bit information per source sample, and at the same time all the better users also receive this half a bit information for free, then the separation-based scheme is as good as the optimal scheme.

This approximation can in fact be refined, and for this purpose, the following additional definition is needed. For any $1 \geq D_1 \geq D_2 \geq \dots \geq D_K \geq 0$, we associate with it a *relaxed distortion vector* $(D_1^*, D_2^*, \dots, D_K^*)$ and a binary labeling vector (B_1, B_2, \dots, B_K) in a recursive manner

$$(D_k^*, B_k) = \begin{cases} (D_{k-1}^*, 0) & \text{if } 2^{1+\sum_{j=1}^{k-1} B_j} \frac{D_k}{D_{k-1}^*} \geq 1 \\ (2^{1+\sum_{j=1}^{k-1} B_j} D_k, 1) & \text{otherwise} \end{cases} \quad (17)$$

for $k = 1, 2, \dots, K$, and we have defined $D_0^* = 1$ for convenience. It is easily verified that $D_k^* \geq D_{k+1}^*$ for $k = 1, 2, \dots, K - 1$, and moreover $D_k^* \leq 2^{1+\sum_{j=1}^{k-1} B_j} D_k$.

Proposition 1: Let $(D_1^*, D_2^*, \dots, D_K^*)$ be the relaxed distortion vector of (D_1, D_2, \dots, D_K) . If $(D_1, D_2, \dots, D_K) \in \mathcal{D}(P, b)$, then $(D_1^*, D_2^*, \dots, D_K^*) \in \hat{\mathcal{D}}(P, b)$.

The notion of relaxed distortion vector essentially removes the rather artificial condition $D_k \geq 2D_{k+1}$ in Corollary 1. When this condition does not hold for some k , the relaxed distortion vector is introduced to replace $(2D_1, 2^2D_2, \dots, 2^K D_K)$, which in this case does not satisfy the monotonicity requirement in Definition 2 and thus is not a

$$\hat{\mathcal{D}}(P, b) \triangleq \left\{ (D_1, D_2, \dots, D_K) : \sum_{k=1}^K \Delta N_k D_k^{-\frac{1}{b}} \leq P + N_1, \quad 1 \geq D_1 \geq D_2 \geq \dots \geq D_K \geq 0 \right\}. \quad (13)$$

$$\underline{\mathcal{D}}^*(P, b) \triangleq \left\{ (D_1, D_2, \dots, D_K) : \sum_{k=1}^K \Delta N_k (2^k D_k)^{-\frac{1}{b}} \leq P + N_1, \quad 1 \geq D_1 \geq D_2 \geq \dots \geq D_K \geq 0 \right\}. \quad (14)$$

$$\underline{\mathcal{D}}(P, b) \triangleq \left\{ (D_1, D_2, \dots, D_K) : \sum_{k=1}^K \Delta N_k (K D_k)^{-\frac{1}{b}} \leq P + N_1, \quad 1 \geq D_1 \geq D_2 \geq \dots \geq D_K \geq 0 \right\}. \quad (15)$$

valid choice of a distortion vector; nevertheless, in this case, the difference between the original distortion vector and its relaxed version is in fact smaller, being $2^{1+\sum_{j=1}^{k-1} B_j}$, instead of 2^k for D_k as in the case already considered in Corollary 1.

Proposition 1 can be used in the situation where there are an infinite number of users such as in a fading channel. Let the set of users indexed by x and their associated distortions be denoted as D_x , since there may be an uncountably infinite many of them. If we apply the construction given in (17), with B_i replaced by B_x , $\sup\{D_x\}$ taking the role of D_1 and $\inf\{D_x\}$ taking the role of D_K , then the following lemma is straightforward by observing that $\inf\{D_x^*\} \leq \sup\{D_x^*\} \leq 2 \sup\{D_x\}$ and $\inf\{D_x^*\} = 2^{\sum_x B_x} \inf\{D_x\}$.

Lemma 2: The sequence B_x specified by (17) satisfies $\sum_x B_x \leq \log_2(\sup\{D_x\}) - \log_2(\inf\{D_x\}) + 1$.

It is clear that the maximum multiplicative constant is less than $2^{1+\sum_x B_x}$ in the statement of Proposition 1. If there exists a lower bound on the achievable distortion for the best user, denoted as d_{\min} , which is strictly positive, *i.e.*, $\inf\{D_x\} \geq d_{\min} > 0$, then since $\sup\{D_x\} \leq 1$, the multiplicative factor can be bounded as

$$2^{1+\sum_x B_x} \leq \frac{4}{d_{\min}}.$$

Thus even when the number of users is infinite, as long as the lower bound d_{\min} is bounded away from zero, the multiplicative factors are in fact finite. More formally, we have the following corollary².

Corollary 2: For an infinite number of users indexed by x with $\inf_{\{D_x\} \in \mathcal{D}(P, b)} \inf\{D_x\} \geq d_{\min}$, let $\{D_x^*\}$ be the relaxed distortion vector of $\{D_x\}$. If $\{D_x\} \in \mathcal{D}(P, b)$, then $\{D_x^*\} \in \hat{\mathcal{D}}(P, b)$, and furthermore, $\sup\{\frac{D_x^*}{D_x}\} \leq \frac{4}{d_{\min}}$.

The next corollary gives another version of the approximation, essentially stating that for any achievable distortion vector, its K -fold multiple is achievable using the separation approach. In terms of the genie-aided interpretation, the genie only needs to provide $\frac{1}{2} \log K$ bits common information to the users in the separation-based scheme, then it is as good as the optimal scheme. More formally, the following corollary follows directly from Theorem 1.

Corollary 3: If $(D_1, D_2, \dots, D_K) \in \mathcal{D}(P, b)$, then $(K D_1, K D_2, \dots, K D_K) \in \hat{\mathcal{D}}(P, b)$, where

²Here we directly take the number of users to infinity in Proposition 1, however a more rigorous approach is to derive the outer bounds for this case and show the result holds. This can indeed be done either along the line of the proof given in Section IV with careful replacement of summation by integral, or more straightforwardly along the line of proof given in Section V.

$(x_1, x_2, \dots, x_K) \stackrel{+}{K} = (\min(1, x_1), \min(1, x_2), \dots, \min(1, x_K))$.

Theorem 1, Proposition 1 and the corollaries provide approximate characterizations of the achievable distortion region, essentially stating that the loss of the source-channel separation approach is bounded by constants. The bound on the gap is chosen to be (largely) independent of a specific distortion tuple on the boundary of $\mathcal{D}(P, b)$, but it will become clear in the next section that such a choice is not necessary.

The proofs of Theorem 1 and Proposition 1 rely heavily on the following outer bound, which is one of the main contributions of this work.

Theorem 2: Let $\tau_1 \geq \tau_2 \geq \dots \geq \tau_{K-1}$ be any $K-1$ non-negative real values, and $\tau_K = 0$. If $(D_1, D_2, \dots, D_K) \in \mathcal{D}(P, b)$, then

$$\sum_{k=1}^K \Delta N_k \left[\frac{(1 + \tau_k) \prod_{j=2}^k (D_j + \tau_{j-1})}{\prod_{j=1}^k (D_j + \tau_j)} \right]^{\frac{1}{b}} \leq P + N_1. \quad (18)$$

With the above theorem in mind, let us denote the set of distortion vectors satisfying (18) for a specific choice of $\tau_1 \geq \tau_2 \geq \dots \geq \tau_{K-1}$ as $\underline{\mathcal{D}}(P, b, \tau_1, \dots, \tau_{K-1})$, *i.e.*, (19) as given on the top of next page. Thus Theorem 2 essentially states that $\mathcal{D}(P, b) \subseteq \underline{\mathcal{D}}(P, b, \tau_1, \dots, \tau_{K-1})$ for any valid choice of $\tau_1, \tau_2, \dots, \tau_{K-1}$. The following corollary is then immediate.

Corollary 4:

$$\mathcal{D}(P, b) \subseteq \bigcap_{\tau_1 \geq \tau_2 \geq \dots \geq \tau_{K-1} \geq 0} \underline{\mathcal{D}}(P, b, \tau_1, \dots, \tau_{K-1}). \quad (20)$$

To illustrate Corollary 4, let us consider the case $K = 2$ for which the bound involves only one parameter τ_1 . For this case, it can be shown through some algebra that this outer bound is equivalent to the one given in [8]. In Fig. 2, we illustrate the outer bounds for several specific choices of $\tau_1 = \tau$. For comparison, the achievable region using the proposed scheme in [8] is also given. Note that although the inner bound given by this scheme is extremely close to the outer bound, it appears they do not match exactly.

It is worth emphasizing that we view this outer bound differently from the authors in [8]: for each possible value of τ_1 we view the condition (18) as specifying an outer bound for the distortion region (D_1, D_2) ; in contrast, the authors of [8] viewed the distortion D_2 as being lower bounded by a function of D_1 , and the parameter τ_1 was viewed as an additional variable that is subject to optimization, and consequently only the optimal choice of τ_1 value was of interest. These two views are complementary, however the former view appears to be more natural for the K -user problem, which also readily

$$\underline{\mathcal{D}}(P, b, \tau_1, \dots, \tau_{K-1}) \triangleq \left\{ (D_1, D_2, \dots, D_K) : \sum_{k=1}^K \Delta N_k \left[\frac{(1 + \tau_k) \prod_{j=2}^k (D_j + \tau_{j-1})}{\prod_{j=1}^k (D_j + \tau_j)} \right]^{\frac{1}{b}} \leq P + N_1, \right. \\ \left. 1 \geq D_1 \geq D_2 \geq \dots \geq D_K \geq 0 \right\}. \quad (19)$$

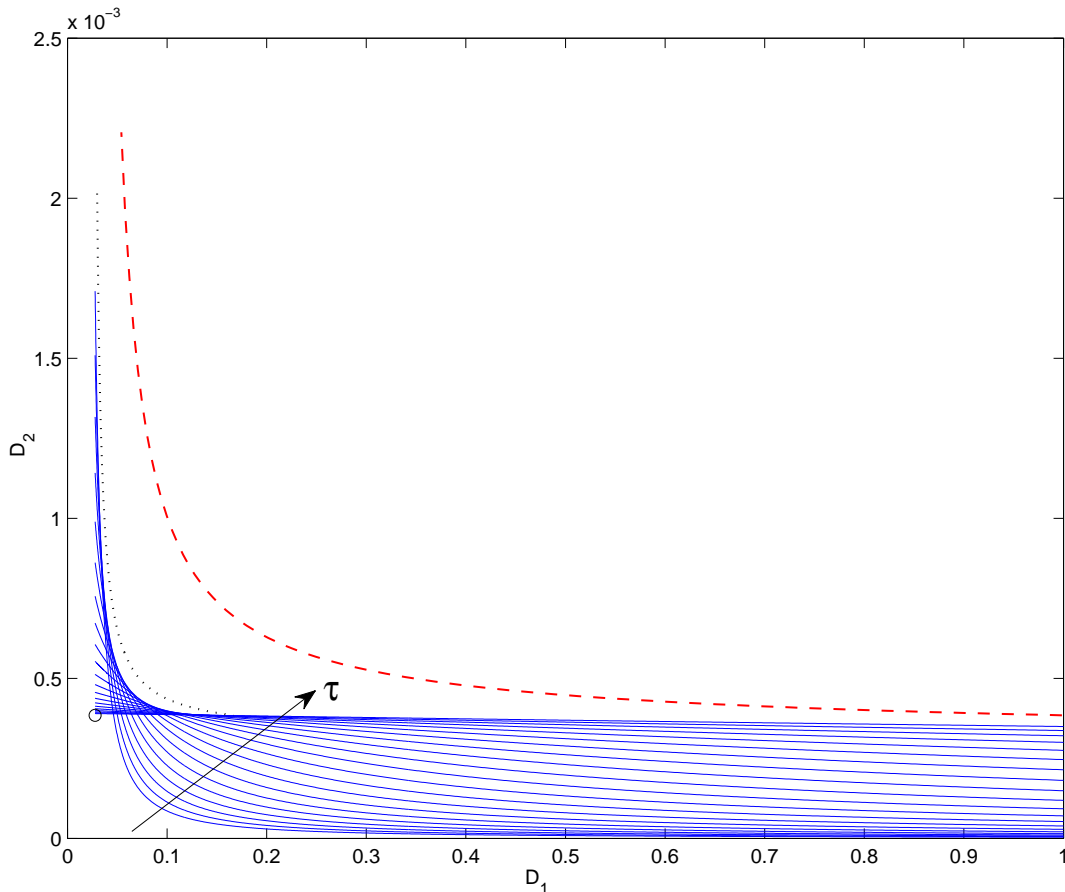


Fig. 2. Illustration of the outer bounds as $\tau_1 = \tau$ varies, for a channel where $N_1 = 10$, $N_2 = 1$, $P = 50$, and $b = 2$. The solid blue lines are the outer bounds, and the dashed red line is the inner bound based on source-channel separation; the black circle gives the trivial outer bound with both users at the respective optimum in the point-to-point setting. For comparison, the achievable region using the proposed scheme in [8] is also given as the dotted black line.

leads to the approximate characterizations. In certain cases, the second view may be more convenient, such as when we are given a specific achievable distortion tuple, and wish to determine how much further improvement is possible or impossible.

For $K = 2$, the properties of the outer bound were thoroughly investigated in [8]. In certain regimes, this outer bound in fact degenerates for the case of bandwidth compression, and it is looser than the trivial outer bound with each user being optimal in the point-to-point setting³. Due to its non-

linear form, the optimization of this bound is rather difficult, and it also appears difficult to determine whether it is always looser than the trivial outer bound in all distortion regimes with bandwidth compression. Nevertheless, it is clear that this outer bound always holds whether the bandwidth is expanded or compressed, and the approximate characterizations are valid in either case.

A different and simpler approximate characterization may in fact be more useful for the bandwidth compression case⁴. Consider a different genie who helps the separation-based

³We would like to thank Dr. Zvi Reznic for clarifying this point in a private communication.

⁴We would again like to thank David Tse as well as one anonymous reviewer for pointing out this different approximate characterization.

scheme by giving each individual user half a bit information *per channel use*, and at the same time all the better users also receive this half a bit information for free, then the genie-aided separation-based scheme is as good as the optimal scheme, and moreover each user can in fact achieve the optimal point-to-point distortion. To see this approximation holds, first observe that the following broadcast channel rates are achievable by using the Gaussian broadcast channel capacity region characterization [17] (it is particularly easy by using the alternative Gaussian broadcast channel capacity characterization given in (22))

$$R_k = \max \left(\frac{1}{2} \log_2 \left(1 + \frac{P}{N_k} \right) - \frac{1}{2} \log_2 \left(1 + \frac{P}{N_{k-1}} \right) - \frac{1}{2}, 0 \right), \quad k = 1, 2, \dots, K. \quad (21)$$

The k -th user can thus utilize a total rate of $\sum_{i=1}^k R_i$ per channel use on this broadcast channel; together with the genie-provided rates, it will have at least a total rate of $\frac{1}{2} \log_2 \left(1 + \frac{P}{N_k} \right)$ per channel use, *i.e.*, the optimal point to point channel rate. Since the Gaussian source is successively refinable [15], it is now clear that each user can achieve the optimal point-to-point distortion with this genie-aided separation-based scheme. Note that though this approximation is good for bandwidth compression, it can be rather loose when the bandwidth expansion factor is large. In contrast, the approximations given in Theorem 1 and Proposition 1 are independent of the bandwidth mismatch factor (the genie provides information in terms of *per source sample*); another difference is that the approximations given in Theorem 1 and Proposition 1 rely on the new outer bound, instead of the simple point-to-point distortion outer bound.

It is clear from the above discussion that the outer bound in Theorem 1 may be further improved by taking its intersection with the trivial point-to-point outer bound. In the remainder of this paper, we do not pursue such possible improvements, but instead focus on the proofs for the results stated in Theorem 1 and Proposition 1.

IV. PROOF OF THE MAIN RESULTS FOR GAUSSIAN BROADCAST CHANNELS

The proofs of the main results for Gaussian source broadcast on Gaussian broadcast channels are given in this section. We start by establishing a simple inner bound for the distortion region $\mathcal{D}(P, b)$ based on source-channel separation, and then focus on deriving an outer bound, or more precisely a set of outer bounds. The approximate characterizations are then rather straightforward by combining these two bounds. From here on, we shall use natural logarithm for concreteness, though choosing logarithm of a different base does not make any essential difference.

A. A Simple Inner Bound

The source-channel separation based coding scheme we consider is extremely simple, which is the combination of a Gaussian successive refinement source code and a Gaussian broadcast channel code; this scheme was thoroughly investigated in [16], and a solution for the optimal power allocation

was given to minimize the expected end-user distortion. Since Gaussian broadcast channel is degraded, a better user can always decode completely the messages sent to the worse users, and thus a successive refinement source code is a perfect match for this channel. Note that such a source-channel separation approach is not optimal in general for this joint source-channel coding problem; see for example [4].

The Gaussian broadcast channel capacity region is well known [17], which is usually given in a parametric form in terms of the power allocation. In this work, we will use an alternative representation, which first appeared in [18] and was instrumental for deriving the optimal power allocation solution in [16]. The Gaussian broadcast channel capacity region (per channel use) can be written in the form in (22) as given on the top of next page.

The rate R_k is the individual message rate intended only to the k -th user, however due to the degradedness, all the better users can also decode this message. Since the Gaussian source is successively refinable [15], by combining an optimal Gaussian successive refinement source code with a Gaussian broadcast code that (asymptotically) achieves (22), we have the following theorem.

Theorem 3:

$$\hat{\mathcal{D}}(P, b) \subseteq \mathcal{D}(P, b). \quad (23)$$

Proof: We wish to show that any $(D_1, D_2, \dots, D_K) \in \hat{\mathcal{D}}(P, b)$ is indeed achievable. Using the separation scheme, we only need to show the channel rates (R_1, R_2, \dots, R_K) specified by

$$D_k = \exp(-2b \sum_{j=1}^k R_j), \quad k = 1, 2, \dots, K, \quad (24)$$

are achievable on this Gaussian broadcast channel. The non-negative vector (R_1, R_2, \dots, R_K) is uniquely determined by (D_1, D_2, \dots, D_K) , and it is straightforwardly seen that it indeed satisfies the inequality in (22). The proof is thus complete. ■

B. An Outer Bound

Next we derive a set of conditions that any achievable distortion vector (D_1, D_2, \dots, D_K) has to satisfy, *i.e.*, Theorem 2.

Proof of Theorem 2: Let us first introduce a set of auxiliary random variables, defined as

$$U_k = S + V_k, \quad k = 1, 2, \dots, K-1, \quad (25)$$

where V_k 's are zero Gaussian random variables, with variance τ_k , and furthermore

$$V_k = V_{k+1} + \Delta V_k, \quad k = 1, 2, \dots, K-1, \quad (26)$$

where ΔV_k is a zero-mean Gaussian random variable, independent of everything else, with variance $\Delta \tau_k = \tau_k - \tau_{k+1}$. For convenience, we define $U_K = S$, which implies $\tau_K \triangleq 0$; furthermore, define $U_0 \triangleq 0$, *i.e.*, being a constant. This technique of introducing auxiliary random variables beyond those in the original problem was previously used in [8], [12],

$$\mathcal{C} = \left\{ (R_1, R_2, \dots, R_M) : R_k \geq 0, k = 1, 2, \dots, K, \sum_{k=1}^K \Delta N_k \exp \left(2 \sum_{j=1}^k R_j \right) \leq P + N_1 \right\}. \quad (22)$$

[13] to derive outer bounds, and specifically in [13] more than one random variable was introduced, whereas in [8], [12] only one was introduced.

For any encoding and decoding functions, we consider a quantity which bears some similarity to the expression for the Gaussian broadcast channel capacity (22), and we denote this quantity as $E_{f,g}(\tau_1, \tau_2, \dots, \tau_{K-1})$ due to its sum exponential form

$$E_{f,g}(\tau_1, \tau_2, \dots, \tau_{K-1}) \triangleq \sum_{k=1}^K \Delta N_k \exp \left[\frac{2}{n} \sum_{j=1}^k I(U_j^m; Y_j^n | U_1^m, U_2^m, \dots, U_{j-1}^m) \right]. \quad (27)$$

The subscript (f, g) makes it clear that this quantity depends on the specific encoding and decoding functions. Next we shall derive universal upper and lower bounds for this quantity regardless the specific choice of functions (f, g) , which eventually yield an outer bound for $\mathcal{D}(P, b)$.

Let (f, g) be any encoding and decoding functions that (asymptotically) achieve the distortions (D_1, D_2, \dots, D_K) . We first derive a lower bound for $E_{f,g}(\tau_1, \tau_2, \dots, \tau_{K-1})$. Observe that for $j = 2, 3, \dots, K$,

$$\begin{aligned} I(U_j^m; Y_j^n | U_1^m, U_2^m, \dots, U_{j-1}^m) &= I(U_j^m; Y_j^n) - I(U_{j-1}^m; Y_j^n) \\ &\geq \frac{m}{2} \log \frac{(1 + \tau_j)(D_j + \tau_{j-1})}{(1 + \tau_{j-1})(D_j + \tau_j)}, \end{aligned} \quad (28)$$

where the equality is due to the Markov string $U_1^m \leftrightarrow U_2^m \leftrightarrow \dots \leftrightarrow U_{K-1}^m \leftrightarrow S^m \leftrightarrow X^n \leftrightarrow Y_j^n$, and the inequality is by Lemma 1. Moreover, also by Lemma 1, we have

$$I(U_1^m; Y_1^n) \geq \frac{m}{2} \log \frac{1 + \tau_1}{D_1 + \tau_1}. \quad (29)$$

It follows that

$$\begin{aligned} &\sum_{j=1}^k I(U_j^m; Y_j^n | U_1^m, U_2^m, \dots, U_{j-1}^m) \\ &\geq \frac{m}{2} \log \frac{1 + \tau_1}{D_1 + \tau_1} + \frac{m}{2} \sum_{j=2}^k \log \frac{(1 + \tau_j)(D_j + \tau_{j-1})}{(1 + \tau_{j-1})(D_j + \tau_j)} \\ &= \frac{m}{2} \log \frac{1 + \tau_k}{D_1 + \tau_1} + \frac{m}{2} \sum_{j=2}^k \log \frac{D_j + \tau_{j-1}}{D_j + \tau_j} \end{aligned} \quad (30)$$

Summarizing the above bounds, we have

$$E_{f,g}(\tau_1, \tau_2, \dots, \tau_{K-1}) \geq \sum_{k=1}^K \Delta N_k \exp \left[\frac{1}{b} \log \frac{1 + \tau_k}{D_1 + \tau_1} + \frac{1}{b} \sum_{j=2}^k \log \frac{D_j + \tau_{j-1}}{D_j + \tau_j} \right]. \quad (31)$$

Next we turn to upper-bounding $E_{f,g}(\tau_1, \tau_2, \dots, \tau_{K-1})$, and first write the following.

$$\begin{aligned} &\frac{2}{n} \sum_{j=1}^k I(U_j^m; Y_j^n | U_1^m, U_2^m, \dots, U_{j-1}^m) \\ &= \frac{2}{n} \sum_{j=1}^k [I(U_j^m; Y_j^n) - I(U_{j-1}^m; Y_j^n)] \\ &= \frac{2}{n} \sum_{j=1}^k [h(Y_j^n | U_{j-1}^m) - h(Y_j^n | U_j^m)] \\ &= \frac{2}{n} \sum_{j=1}^k h(Y_j^n | U_{j-1}^m) - \frac{2}{n} \sum_{j=1}^k h(Y_j^n | U_j^m). \end{aligned} \quad (32)$$

Applying the entropy power inequality [19] for $j = 1, 2, \dots, K-1$, we have

$$\begin{aligned} &\exp \left[\frac{2}{n} h(Y_j^n | U_{j-1}^m) \right] \\ &\geq \exp \left[\frac{2}{n} h(Y_{j+1}^n | U_j^m) \right] + \exp [\log(2\pi e \Delta N_j)] \\ &= \exp \left[\frac{2}{n} h(Y_{j+1}^n | U_j^m) \right] + 2\pi e \Delta N_j. \end{aligned} \quad (33)$$

For $j = K$, it is clear that

$$\begin{aligned} \exp \left[\frac{2}{n} h(Y_K^n | U_K^m) \right] &= \exp \left[\frac{2}{n} h(Y_K^n | S^m) \right] \\ &= 2\pi e N_K = 2\pi e \Delta N_K. \end{aligned} \quad (34)$$

By defining $\exp \left[\frac{2}{n} h(X^n | S^m) \right] \triangleq 0$, it now follows that

$$\begin{aligned} &E_{f,g}(\tau_1, \tau_2, \dots, \tau_{K-1}) \\ &= \sum_{k=1}^K \Delta N_k \exp \left[\frac{2}{n} \sum_{j=1}^k I(U_j^m; Y_j^n | U_1^m, U_2^m, \dots, U_{j-1}^m) \right] \\ &\leq \sum_{k=1}^K \Delta N_k \frac{\exp \left[\frac{2}{n} \sum_{j=1}^k h(Y_j^n | U_{j-1}^m) \right]}{\prod_{j=1}^k [\exp(\frac{2}{n} h(Y_{j+1}^n | U_j^m)) + 2\pi e \Delta N_j]}. \end{aligned} \quad (35)$$

We bound this summation, by considering the summands in the reversed order, *i.e.*, $k = K, K-1, \dots, 1$. Starting with the summands when $k = K-1$ and $k = K$, we have (36) as given on the top of next page. Continuing this line of reduction, we finally arrive at (37) when $k = 1$ where the last inequality is by the concavity of the $\log(\cdot)$ function and the given power constraint.

Combining (31) and (37), it is clear that for any encoding

$$\begin{aligned}
& \Delta N_{K-1} \frac{\exp\left[\frac{2}{n} \sum_{j=1}^{K-1} h(Y_j^n | U_{j-1}^m)\right]}{\prod_{j=1}^{K-1} \left[\exp\left(\frac{2}{n} h(Y_{j+1}^n | U_j^m)\right) + 2\pi e \Delta N_j\right]} + \Delta N_K \frac{\exp\left[\frac{2}{n} \sum_{j=1}^K h(Y_j^n | U_{j-1}^m)\right]}{\prod_{j=1}^K \left[\exp\left(\frac{2}{n} h(Y_{j+1}^n | U_j^m)\right) + 2\pi e \Delta N_j\right]} \\
&= \frac{\exp\left[\frac{2}{n} \sum_{j=1}^{K-1} h(Y_j^n | U_{j-1}^m)\right]}{\prod_{j=1}^{K-1} \left[\exp\left(\frac{2}{n} h(Y_{j+1}^n | U_j^m)\right) + 2\pi e \Delta N_j\right]} \left[\Delta N_{K-1} + \Delta N_K \frac{\exp\left[\frac{2}{n} h(Y_K^n | U_{K-1}^m)\right]}{2\pi e \Delta N_K} \right] \\
&= \frac{1}{2\pi e} \frac{\exp\left[\frac{2}{n} \sum_{j=1}^{K-1} h(Y_j^n | U_{j-1}^m)\right]}{\prod_{j=1}^{K-2} \left[\exp\left(\frac{2}{n} h(Y_{j+1}^n | U_j^m)\right) + 2\pi e \Delta N_j\right]}. \tag{36}
\end{aligned}$$

$$\begin{aligned}
E_{f,g}(\tau_1, \tau_2, \dots, \tau_{K-1}) &\leq \Delta N_1 \frac{\exp\left[\frac{2}{n} h(Y_1^n)\right]}{\exp\left[\frac{2}{n} h(Y_2^n | U_1^m)\right] + 2\pi e \Delta N_1} + \frac{1}{2\pi e} \frac{\exp\left[\frac{2}{n} \sum_{j=1}^2 h(Y_j^n | U_{j-1}^m)\right]}{\left[\exp\left(\frac{2}{n} h(Y_2^n | U_1^m)\right) + 2\pi e \Delta N_1\right]} \\
&= \frac{\exp\left[\frac{2}{n} h(Y_1^n)\right]}{\exp\left[\frac{2}{n} h(Y_2^n | U_1^m)\right] + 2\pi e \Delta N_1} \left[\Delta N_1 + \frac{\exp\left[\frac{2}{n} h(Y_2^n | U_1^m)\right]}{2\pi e} \right] \\
&= \frac{\exp\left[\frac{2}{n} h(Y_1^n)\right]}{2\pi e} \leq P + N_1, \tag{37}
\end{aligned}$$

and decoding functions (f, g)

$$\begin{aligned}
P + N_1 &\geq E_{f,g}(\tau_1, \tau_2, \dots, \tau_{K-1}) \\
&\geq \sum_{k=1}^K \Delta N_k \exp\left[\frac{1}{b} \log \frac{1 + \tau_k}{D_1 + \tau_1} + \frac{1}{b} \sum_{j=2}^k \log \frac{D_j + \tau_{j-1}}{D_j + \tau_j} \right], \tag{38}
\end{aligned}$$

which completes the proof. \blacksquare

The meaning of the newly introduced random variable U_k can be roughly understood as the message meant for the k -th user. Under this interpretation, the term $\frac{1}{n} I(U_k^m; Y_k^n | U_1^m, U_2^m, \dots, U_{k-1}^m)$ in the quantity $E_{f,g}(\tau_1, \tau_2, \dots, \tau_{K-1})$ essentially represents the individual rate intended for the k -th user in the Gaussian broadcast channel; this informal understanding provides the rationale for bounding $E_{f,g}(\tau_1, \tau_2, \dots, \tau_{K-1})$. This interpretation is nevertheless not completely accurate, and thus the outer bound is likely to be not tight in general, but suffices to provide approximate characterizations.

C. The Approximate Characterizations

Now we are ready to prove Theorem 1 and Proposition 1.

Proof of Theorem 1 and Proposition 1: The first inclusion $\mathcal{D}(P, b) \subseteq \mathcal{D}(P, b)$ in Theorem 1 is simply Theorem 3, and thus we focus on the other inclusion $\mathcal{D}(P, b) \subseteq \underline{\mathcal{D}}^*(P, b) \cap \underline{\mathcal{D}}(P, b)$, for which we prove $\mathcal{D}(P, b) \subseteq \underline{\mathcal{D}}^*(P, b)$ and $\mathcal{D}(P, b) \subseteq \underline{\mathcal{D}}(P, b)$ separately. From Theorem 2, it is clear that if $(D_1, D_2, \dots, D_K) \in \mathcal{D}(P, b)$, then (18) holds with any $\tau_1 \geq \tau_2 \geq \dots \geq \tau_{K-1} \geq 0$, and thus (18) holds when we choose $\tau_k = D_k$ for $k = 1, 2, \dots, K-1$. It follows that the following condition has to be satisfied by any achievable

distortion vector

$$\sum_{k=1}^K \Delta N_k \left[\frac{(1 + D_k) \prod_{j=2}^k (D_j + D_{j-1})}{\prod_{j=1}^k (D_j + D_j)} \right]^{\frac{1}{b}} \leq P + N_1. \tag{39}$$

However, notice that

$$\begin{aligned}
& \sum_{k=1}^K \Delta N_k \left[\frac{(1 + D_k) \prod_{j=2}^k (D_j + D_{j-1})}{\prod_{j=1}^k (D_j + D_j)} \right]^{\frac{1}{b}} \\
&\geq \sum_{k=1}^K \Delta N_k \left[\frac{\prod_{j=2}^k D_{j-1}}{\prod_{j=1}^k 2D_j} \right]^{\frac{1}{b}} = \sum_{k=1}^K \Delta N_k (2^k D_k)^{-\frac{1}{b}}. \tag{40}
\end{aligned}$$

It now follows straightforwardly that any achievable distortion vector has to satisfy

$$\sum_{k=1}^K \Delta N_k (2^k D_k)^{-\frac{1}{b}} \leq P + N_1, \tag{41}$$

and $\mathcal{D}(P, b) \subseteq \underline{\mathcal{D}}^*(P, b)$ is proved.

To prove $\mathcal{D}(P, b) \subseteq \underline{\mathcal{D}}(P, b)$, note again that any achievable distortion vector has to satisfy Theorem 2, and because of the similarity between the forms given in (15) and (18), we only need to prove

$$\frac{1}{KD_k} \leq \frac{(1 + \tau_k) \prod_{j=2}^k (D_j + \tau_{j-1})}{\prod_{j=1}^k (D_j + \tau_j)}, \quad k = 1, 2, \dots, K, \tag{42}$$

for some specific choice of $\tau_1 \geq \tau_2 \geq \dots \geq \tau_{K-1} \geq 0$. We first consider the case that $D_1 \leq 1/K$; the case that $D_1 > 1/K$ needs to be treated in a slightly different manner, as we shall discuss later. We take an induction approach, for which

the following auxiliary quantities are needed

$$\alpha_k = \frac{D_k(1 + \tau_k)}{D_{k+1} + \tau_k}, \quad k = 1, 2, \dots, K-1. \quad (43)$$

We claim that there exist $\tau_1 \geq \tau_2 \geq \dots \geq \tau_{K-1} \geq 0$ such that (42) holds with equality for $k = 1, 2, \dots, K-1$, and holds for $k = K$ with equality or inequality; moreover with these τ_k 's we have $\alpha_k \leq (K-k)^{-1}$. First consider the case $k = 1$, since the function

$$\Phi_k(\tau) = \frac{1 + \tau}{D_k + \tau} \quad (44)$$

is monotonically decreasing and continuous in the range $[0, \infty]$, as long as

$$\Phi_1(0) = \frac{1}{D_1} \geq \frac{1}{KD_1} \geq 1 = \Phi_1(\infty), \quad (45)$$

there exists a unique solution of τ_1 such that (42) holds with equality. This is indeed true for $D_1 \leq 1/K$, which gives

$$\tau_1 = \frac{(K-1)D_1}{1 - KD_1}. \quad (46)$$

It follows that

$$\begin{aligned} \alpha_1 &= \frac{D_1(1 + \tau_1)}{D_2 + \tau_1} = \frac{D_1(1 - D_1)}{D_2(1 - KD_1) + (K-1)D_1} \\ &\leq \frac{D_1}{(K-1)D_1} = \frac{1}{K-1}, \end{aligned} \quad (47)$$

and thus our claim holds for $k = 1$. Next suppose the claim is true for $k = k^*$ and we prove it is also true $k = k^* + 1$, for which we wish to find τ_{k^*+1} such that

$$\begin{aligned} \frac{1}{KD_{k^*+1}} &= \frac{(1 + \tau_{k^*+1}) \prod_{j=2}^{k^*+1} (D_j + \tau_{j-1})}{\prod_{j=1}^{k^*+1} (D_j + \tau_j)} \\ &= \frac{(1 + \tau_{k^*}) \prod_{j=2}^{k^*} (D_j + \tau_{j-1}) (1 + \tau_{k^*+1})(D_{k^*+1} + \tau_{k^*})}{\prod_{j=1}^{k^*} (D_j + \tau_j) (1 + \tau_{k^*})(D_{k^*+1} + \tau_{k^*+1})} \\ &= \frac{1}{KD_{k^*}} \frac{(1 + \tau_{k^*+1})(D_{k^*+1} + \tau_{k^*})}{(1 + \tau_{k^*})(D_{k^*+1} + \tau_{k^*+1})}, \end{aligned} \quad (48)$$

where the last equality is by the supposition that the claim holds true for k^* . Again by the monotonicity and continuity of $\Phi_k(\tau)$, as long as the choice of $(\tau_1, \tau_2, \dots, \tau_{k^*})$ satisfies

$$\begin{aligned} \Phi_{k^*+1}(0) &= \frac{1}{D_{k^*+1}} \geq \frac{D_{k^*}(1 + \tau_{k^*})}{D_{k^*+1}(D_{k^*+1} + \tau_{k^*})} \\ &\geq \frac{1 + \tau_{k^*}}{D_{k^*+1} + \tau_{k^*}} = \Phi_{k^*+1}(\tau_{k^*}), \end{aligned} \quad (49)$$

there exists a valid solution τ_{k^*+1} in $[0, \tau_{k^*}]$ for (48) to hold. The second inequality in (49) is clearly true, and thus we only need to consider the first inequality. However, notice that

$$\begin{aligned} \frac{D_{k^*}(1 + \tau_{k^*})}{D_{k^*+1}(D_{k^*+1} + \tau_{k^*})} &= \frac{\alpha_{k^*}}{D_{k^*+1}} \\ &\leq \frac{1}{(K - k^*)D_{k^*+1}} \leq \frac{1}{D_{k^*+1}}, \end{aligned} \quad (50)$$

and we thus conclude that there indeed exists a valid solution

of τ_{k^*+1} for (48), or more precisely

$$\tau_{k^*+1} = \frac{D_{k^*+1}(1 - \alpha_{k^*})}{\alpha_{k^*} - D_{k^*+1}}. \quad (51)$$

To bound α_{k^*+1} , we write

$$\begin{aligned} \alpha_{k^*+1} &= \frac{D_{k^*+1}(1 + \tau_{k^*+1})}{D_{k^*+2} + \tau_{k^*+1}} \leq \frac{D_{k^*+1}(1 + \tau_{k^*+1})}{\tau_{k^*+1}} \\ &= \frac{\alpha_{k^*}(1 - D_{k^*+1})}{1 - \alpha_{k^*}} \leq \frac{\alpha_{k^*}}{1 - \alpha_{k^*}} \\ &\leq \frac{(K - k^*)^{-1}}{1 - (K - k^*)^{-1}} = \frac{1}{K - k^* - 1}, \end{aligned} \quad (52)$$

where the last inequality is by the monotonicity of $x/(1-x)$ in $[0, 1]$, and the fact $\alpha_{k^*} \leq (K - k^*)^{-1}$. The induction is thus complete. It only remains to check that when $k = K$ the inequality (42) still holds, for which we have

$$\begin{aligned} &\frac{\prod_{j=2}^K (D_j + \tau_{j-1})}{DK \prod_{j=1}^{K-1} (D_j + \tau_j)} \\ &= \frac{(1 + \tau_{K-1}) \prod_{j=2}^{K-1} (D_j + \tau_{j-1})}{\prod_{j=1}^{K-1} (D_j + \tau_j)} \frac{D_K + \tau_{K-1}}{(1 + \tau_{K-1})DK} \\ &= \frac{1}{KD_{K-1}} \frac{D_K + \tau_{K-1}}{(1 + \tau_{K-1})DK} = \frac{1}{\alpha_{K-1}KD_K} \geq \frac{1}{KD_K}, \end{aligned} \quad (53)$$

where the last inequality is by the fact $\alpha_{K-1} \leq 1$. We have thus proved that $\mathcal{D}(P, b) \subseteq \underline{\mathcal{D}}(P, b)$ for the case $D_1 \leq 1/K$.

Next we briefly discuss the case $D_1 \geq D_2 \geq \dots \geq D_r > 1/K \geq D_{r+1} \geq \dots \geq D_K$, and we shall prove that (42) holds for some $\tau_1 \geq \tau_2 \geq \dots \geq \tau_{K-1} \geq 0$. Notice that by choosing $\tau_1, \tau_2, \dots, \tau_r$ sufficiently large, we can make

$$\frac{(1 + \tau_k) \prod_{j=2}^k (D_j + \tau_{j-1})}{\prod_{j=1}^k (D_j + \tau_j)} \geq \frac{1}{KD_k}, \quad k = 1, 2, \dots, r. \quad (54)$$

because of the strict inequality given in $D_1 \geq D_2 \geq \dots \geq D_r > 1/K$, and the fact that the left-hand-side of (54) goes to 1 when we send $\tau_1, \tau_2, \dots, \tau_r$ to infinity. We have also $\frac{1}{K-r} \geq \frac{1}{K} \geq D_{r+1} \geq \dots \geq D_K$, and thus there exist $\tau_{r+1} \geq \dots \geq \tau_{K-1}$ such that

$$\frac{1}{(K-r)D_k} \leq \frac{(1 + \tau_k) \prod_{j=r+2}^k (D_j + \tau_{j-1})}{\prod_{j=r+1}^k (D_j + \tau_j)} \quad (55)$$

holds with equality for $k = r+1, r+2, \dots, K-1$ and with either equality or inequality for $k = K$, by applying the result for the previously discussed case in a system with $K-r$ users. Since we can choose $\tau_1, \tau_2, \dots, \tau_r$ sufficiently large, it is clear that indeed this set of τ_k 's makes (42) hold for $k = r+1, r+2, \dots, K$. This completes the proof for $\mathcal{D}(P, b) \subseteq \underline{\mathcal{D}}(P, b)$.

To prove Proposition 1, we need to choose different values for $(\tau_1, \tau_2, \dots, \tau_{K-1})$. Essentially, when the condition $2D_k \leq D_{k-1}$ is not satisfied for certain k , we will choose to ignore the contribution of D_k in the outer bound of Theorem 2 by taking an appropriate value of τ_k . For convenience, define the set $\mathcal{B} = \{k : B_k = 1\}$, where B_k is the labeling function given

before Proposition 1; denote the member of \mathcal{B} in an increasing order as $k_1, k_2, \dots, k_{|\mathcal{B}|}$, where $|\mathcal{B}|$ is the cardinality of the set \mathcal{B} . The value of τ_k 's are set by the following recursive formula

$$\tau_k = \begin{cases} D_k & \text{if } B_k = 1 \\ \tau_{k-1} & \text{if } B_k = 0 \end{cases}, \quad k = 1, 2, \dots, K-1, \quad (56)$$

where we have defined $\tau_0 \triangleq 1$ for convenience.

Note that $\tau_{k+1} \leq \tau_k$ for $k = 1, 2, \dots, K-2$. It follows from Theorem 2 that this achievable distortion vector has to satisfy (57) on the top of next page, where for convenience we define $k_{|\mathcal{B}|+1} = K+1$; to see the equality holds, note that the terms for $k \notin \mathcal{B}$ are combined with the terms of $k \in \mathcal{B}$, because this choice of $(\tau_1, \tau_2, \dots, \tau_{K-1})$ cancels out some of terms in the numerator and the denominator. This implies that the distortion vector has to satisfy

$$\begin{aligned} P + N_1 &\geq N_1 - N_{k_1} + \sum_{i=1}^{|\mathcal{B}|} (N_{k_i} - N_{k_{i+1}}) (2^i D_{k_i})^{-\frac{1}{b}} \\ &= N_1 - N_{k_1} + \sum_{i=1}^{|\mathcal{B}|} (N_{k_i} - N_{k_{i+1}}) (D_{k_i}^*)^{-\frac{1}{b}} \\ &= \sum_{k=1}^K \Delta N_k (D_k^*)^{-\frac{1}{b}}, \end{aligned} \quad (58)$$

where the last two equalities are by the definition of D_k^* . The proof can now be completed by applying Theorem 3. ■

V. GAUSSIAN SOURCES ON GENERAL BROADCAST CHANNELS

In this section, we show that the results for sending Gaussian sources on Gaussian broadcast channels can be conveniently extended to general broadcast channels, which was inspired by a recent work by Avestimehr, Caire and Tse [14].

The broadcast channel is now given by an arbitrary conditional distribution $P(Y_1, Y_2, \dots, Y_K | X)$, in the alphabets $(\mathcal{Y}_1, \mathcal{Y}_2, \dots, \mathcal{Y}_K, \mathcal{X})$. For brevity, we omit repeating the definition of the codes here. To distinguish from the Gaussian channel case, we denote the achievable distortion region as $\mathcal{D}_g(b)$, with a bandwidth mismatch factor b . For the separation-based scheme, we shall consider combining successive refinement source codes with broadcast codes with degraded message set [20]. Particularly, for an arbitrary permutation $\pi : \{1, 2, \dots, K\} \rightarrow \{1, 2, \dots, K\}$, the degraded message set requirement implies that there are a total of K independent messages (W_1, W_2, \dots, W_K) , such that the user $\pi(k)$ should decode the messages W_1, W_2, \dots, W_k . For an arbitrary permutation π , let us denote the achievable distortion by the separation-based approach of combining successive refinement source code with the broadcast code with degraded message set as $\hat{\mathcal{D}}_g^\pi(b)$, and the overall achievable distortion region using this separation-based approach is given by

$$\hat{\mathcal{D}}_g(b) = \bigcup_{\pi} \hat{\mathcal{D}}_g^\pi(b). \quad (59)$$

Clearly the convex closure of the above region is also achievable, however such generality is not required.

It is worth noting that since the characterization of the broadcast channel capacity region with degraded message set

is still an open problem for $K > 2$, we do not have a characterization for the region $\hat{\mathcal{D}}_g(b)$. However, if the broadcast channel is degraded, then only one permutation needs to be considered; moreover, if the capacity region of the broadcast channel (with degraded message set or it is a degraded broadcast channel) is known, such as for the Gaussian broadcast channel case, the region $\hat{\mathcal{D}}_g(b)$ can then be completely characterized.

Now we present the counterpart of Corollary 1, Proposition 1, Corollary 2 and Corollary 3 for general broadcast channels.

Corollary 5: If $(D_1, D_2, \dots, D_K) \in \mathcal{D}_g(P, b)$, and if $D_k \geq 2D_{k+1}$ for $k = 1, 2, \dots, K-1$, then $(2D_1, 2^2D_2, \dots, 2^KD_K) \in \hat{\mathcal{D}}_g(P, b)$.

Proposition 2: Let $(D_1^*, D_2^*, \dots, D_K^*)$ be the relaxed distortion vector of (D_1, D_2, \dots, D_K) as defined in (17). If $(D_1, D_2, \dots, D_K) \in \mathcal{D}_g(b)$, then $(D_1^*, D_2^*, \dots, D_K^*) \in \hat{\mathcal{D}}_g(b)$.

Corollary 6: For an infinite number of users indexed by x with $\inf_{\{D_x\} \in \mathcal{D}_g(b)} \inf\{D_x\} \geq d_{\min} > 0$, let $\{D_x^*\}$ be the relaxed distortion vector of $\{D_x\}$ as defined in (17). If $\{D_x\} \in \mathcal{D}_g(b)$, then $\{D_x^*\} \in \hat{\mathcal{D}}_g(b)$, and furthermore, $\sup\{\frac{D_x}{D_x^*}\} \leq \frac{4}{d_{\min}}$.

Corollary 7: If $(D_1, D_2, \dots, D_K) \in \mathcal{D}_g(P, b)$, then $(KD_1, KD_2, \dots, KD_K) \in \hat{\mathcal{D}}_g(P, b)$.

We only prove Corollary 7 here, since the proofs of Corollary 5, Proposition 2 and Corollary 6 are quite similar.

Proof: Assume a distortion vector (D_1, D_2, \dots, D_K) is indeed achievable with certain joint source-channel coding scheme. Let us view the induced random mapping from \mathcal{S}^m to \mathcal{Y}_i^n , $i = 1, 2, \dots, K$, by this joint source-channel code as a super-broadcast-channel; without loss of generality, let us assume $D_1 \geq D_2 \geq \dots \geq D_K > 0$.

We pick up the story from (28) and (29), and claim that there exists a degraded message set broadcast code on the super-broadcast-channel with asymptotic rate per each length- m block as follows

$$\begin{aligned} R_1^m &= \frac{m}{2} \log \frac{1 + \tau_1}{D_1 + \tau_1}, \\ R_j^m &= \frac{m}{2} \log \frac{(1 + \tau_j)(D_j + \tau_{j-1})}{(1 + \tau_{j-1})(D_j + \tau_j)}, \quad j = 2, 3, \dots, K. \end{aligned} \quad (60)$$

It is not difficult to see that the distribution $U_1^m, U_2^m, \dots, U_K^m, X^m$ can be used to construct a well-known super-position code with the above rates [19] by (28) and (29). This code needs to span over n' blocks, and note that the induced super-broadcast-channel is block-wise memoryless. We only need to confirm that receiver k can decode all the messages up to the k -th layer by successive decoding. Observe that since

$$I(U_1^m; Y_k^n) \geq \frac{m}{2} \frac{1 + \tau_1}{D_k + \tau_1} \geq \frac{m}{2} \frac{1 + \tau_1}{D_1 + \tau_1} = R_1^m, \quad (61)$$

indeed receiver k can decoder the first layer code, and thus recover the codewords based on U_1 . The above inequality essentially shows that the channel to Y_k^n is more powerful than that to Y_1^n , with the given channel input distribution, although the broadcast channel itself is not necessarily degraded. For

$$\begin{aligned}
& \sum_{k=1}^K \Delta N_k \left[\frac{(1 + \tau_k) \prod_{j=2}^k (D_j + \tau_{j-1})}{\prod_{j=1}^k (D_j + \tau_j)} \right]^{\frac{1}{b}} \\
& = N_1 - N_{k_1} + \sum_{i=1}^{|\mathcal{B}|} (N_{k_i} - N_{k_{i+1}}) \left[\frac{(1 + D_{k_i}) \prod_{j=2}^i (D_{k_j} + D_{k_{j-1}})}{\prod_{j=1}^i (D_{k_j} + D_{k_j})} \right]^{\frac{1}{b}} \leq P + N_1, \tag{57}
\end{aligned}$$

the j -th layer where $j \leq k$, we have

$$\begin{aligned}
I(U_j^m; Y_k^n | U_{j-1}^m) & \geq \frac{m}{2} \log \frac{(1 + \tau_j)(D_k + \tau_{j-1})}{(1 + \tau_{j-1})(D_k + \tau_j)} \\
& \geq \frac{m}{2} \log \frac{(1 + \tau_j)(D_j + \tau_{j-1})}{(1 + \tau_{j-1})(D_j + \tau_j)} = R_j^m, \tag{62}
\end{aligned}$$

because of the monotonicity of the function

$$f(D) = \frac{D + \tau_{j-1}}{D + \tau_j}, \tag{63}$$

when $\tau_{j-1} \geq \tau_j$, and the fact $D_k \leq D_j$. Thus our claim is indeed true.

Using this set of degraded message set broadcast channel codes, we can achieve (or more precisely, approach arbitrarily close to) the following distortion D_k^* for any $\tau_1 \geq \tau_2 \geq \dots \geq \tau_{K-1} \geq 0$, and $\tau_K \triangleq 0$,

$$D_k^* = \exp\left(-2 \sum_{j=1}^k \frac{R_j^m}{m}\right) = \frac{\prod_{j=1}^k (D_j + \tau_j)}{(1 + \tau_k) \prod_{j=2}^k (D_j + \tau_{j-1})}. \tag{64}$$

However, in the proof of Theorem 1, we have already showed that there exist τ 's such that (42) holds, *i.e.*,

$$D_k^* \leq K D_k, \quad k = 1, 2, \dots, K. \tag{65}$$

The proof is thus complete. \blacksquare

VI. CONCLUDING REMARKS

We derived a new outer bound to the achievable distortion region for the joint source-channel coding problem of sending a Gaussian source on Gaussian broadcast channels with bandwidth mismatch. When combined with a simple separation-based achievability result, this new bound leads to approximate characterizations of the achievable distortion region within some universal constant multiplicative factors. These results are further extended to the case of Gaussian source broadcast on general broadcast channels.

The outer bound was not fully optimized, which seems to be a difficult problem by itself. It may be beneficial to investigate the outer bound more thoroughly when more powerful achievability schemes become available. In the current work, we only considered the separation-based scheme that yields approximate characterizations.

The technique used in this work can be applied to the problem of multi-source broadcast on more complex communication networks under certain conditions. We believe that similar results hold for many classes of suitably well-behaved networks. In a follow-up work to this paper, we have shown

approximate separation for a class of such networks along with other results on source-channel separation [21].

The outer bounding technique of introducing auxiliary random variables used in this work is inspired by those used in [8], [12], [13]. We believe this technique is also promising for other multi-user information theoretic problems. The role of the auxiliary random variable introduced in [8], [12] was not quite well understood or interpreted previously, and may even appear mysterious to many researchers less familiar with the specific problems being treated. The current work (see also [13]) has made the meaning of the introduced random variables explicit. More specifically, they are introduced to either substitute the messages in the channel coding problem, or to substitute the source reconstructions in the source coding problem. By this substitution, the quantities representing rates are replaced by the corresponding information quantities. With the interpretation made clear in a general manner, it is our hope that this technique can inspire other meaningful results in the future.

APPENDIX

Proof of Lemma 1: Define $Z' = V + V'$ and $Z = V$, and thus $U' = S + Z'$ and $U = S + Z$. To prove the first statement, we consider the following chain of inequalities

$$\begin{aligned}
& I(W; U'^m) \\
& = mh(U') - h(U'^m | W) \\
& = mh(U') - h(S^m + Z'^m | W) \\
& = mh(U') - h(S^m + Z'^m - \hat{S}^m | W) \\
& \stackrel{(a)}{\geq} mh(U') - h(S^m + Z'^m - \hat{S}^m) \\
& \stackrel{(b)}{\geq} mh(U') - \sum_{i=1}^m h[S(i) + Z'(i) - \hat{S}(i)] \\
& \stackrel{(c)}{\geq} mh(U') \\
& \quad - \sum_{i=1}^m \frac{1}{2} \log \left\{ (2\pi e) \mathbb{E}[(S(i) + Z'(i) - \hat{S}(i))^2] \right\} \\
& = mh(U') - \sum_{i=1}^m \frac{1}{2} \log \left[(2\pi e) (\mathbb{E}d(S(i), \hat{S}(i)) + \tau') \right],
\end{aligned}$$

where \hat{S}^m is the reconstruction with W , and its i -th position is denoted as $\hat{S}(i)$. The inequality (a) is because conditioning reduces entropy, (b) is because of the chain rule for differential entropy and the fact that conditioning reduces entropy, and (c) is because Gaussian distribution maximizes the differential entropy for a given second moment. Since $\log(\cdot)$ is a concave

function, we have

$$\begin{aligned} & \sum_{i=1}^m \frac{1}{2} \log \left[(2\pi e) (\mathbb{E}d(S(i), \hat{S}(i)) + \tau') \right] \\ & \leq \frac{m}{2} \log \left[2\pi e \left(\mathbb{E}d(S^n, \hat{S}^n) + \tau' \right) \right]. \end{aligned}$$

It follows

$$\begin{aligned} I(W; U'^m) & \geq mh(U') - \frac{m}{2} \log \left[2\pi e \left(\mathbb{E}d(S^m, \hat{S}^m) + \tau' \right) \right] \\ & \geq mh(U') - \frac{m}{2} \log [2\pi e(D + \tau')] \\ & = \frac{m}{2} \log \frac{1 + \tau'}{D + \tau'}, \end{aligned}$$

which is the first claim in the lemma.

To prove the second claim, we write the following

$$\begin{aligned} I(W; U^m) - I(W; U'^m) & = mh(U) - mh(U') + h(U'^m|W) - h(U^m|W). \end{aligned}$$

For the latter two terms, we have

$$\begin{aligned} h(U'^m|W) - h(U^m|W) & \stackrel{(a)}{=} h(U'^m|W) - h(U^m|V'^m, W) \\ & \stackrel{(b)}{=} h(U'^m|W) - h(U'^m|V'^m, W) \\ & = I(U'^m; V'^m|W), \end{aligned}$$

where (a) is because V'^m is independent of U^m and W ; (b) is by the definition of U' . Continuing the chain of inequalities, we have

$$\begin{aligned} & I(U'^m; V'^m|W) \\ & \stackrel{(a)}{=} h(V'^m) - h(V'^m|S^m + V^m + V'^m, W) \\ & = h(V'^m) - h(V'^m|S^m + V^m + V'^m, \hat{S}^m, W) \\ & \stackrel{(b)}{\geq} h(V'^m) - h(V'^m|S^m - \hat{S}^m + V^m + V'^m) \\ & \stackrel{(c)}{\geq} \sum_{i=1}^m \left[h(V'(i)) - h(V'(i)|S(i) - \hat{S}(i) + V(i) + V'(i)) \right] \\ & = \sum_{i=1}^m I(V'(i); S(i) - \hat{S}(i) + V(i) + V'(i)) \\ & \stackrel{(d)}{\geq} \sum_{i=1}^m \frac{1}{2} \log \frac{\mathbb{E}d(S(i), \hat{S}(i)) + \tau'}{\mathbb{E}d(S(i), \hat{S}(i)) + \tau} \\ & \stackrel{(e)}{\geq} \frac{m}{2} \log \frac{D + \tau'}{D + \tau}, \end{aligned}$$

where (a) is because V' is independent of W ; (b) is because conditioning reduces entropy; (c) is by applying the chain rule, and the facts that V'^m is an i.i.d. sequence and conditioning reduces entropy; (d) is by applying the mutual information game result that Gaussian noise is the worst additive noise under a variance constraint [19] (pg. 263, ex. 1), and taking $V'(i)$ as channel input; finally (e) is due to the convexity and monotonicity of $\log \frac{x+\tau'}{x+\tau}$ in $x \in (0, \infty)$ when $\tau' \geq \tau \geq 0$. This completes the proof for the second claim. ■

ACKNOWLEDGMENTS

The authors would like to thank David Tse for the stimulating discussions at several occasions as well as his insightful

comments. The authors are also grateful to the anonymous reviewers for their comments.

REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, pp. 623–656, Jul., Oct. 1948.
- [2] T. J. Goblick, "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Information Theory*, vol. 11, no. 4, pp. 558–567, Oct. 1965.
- [3] T. Cover, A. E. Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Information Theory*, vol. 26, no. 6, pp. 648–657, Nov. 1980.
- [4] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: lossy source-channel communication revisited," *IEEE Trans. Information Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.
- [5] S. Shamai, S. Verdú and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Information Theory*, vol. 44, no. 3, pp. 564–579, Mar. 1998.
- [6] U. Mittal and N. Phamdo, "Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communications," *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1082–1102, May 2002.
- [7] M. Skoglund, N. Phamdo, and F. Alajaji, "Hybrid digital-analog source-channel coding for bandwidth compression/expansion," *IEEE Transactions on Information Theory*, vol. 52, no. 8, pp. 3757–3763, Aug. 2006.
- [8] Z. Reznic, M. Feder and R. Zamir, "Distortion bounds for broadcasting with bandwidth expansion," *IEEE Trans. Information Theory*, vol. 52, no. 8, pp. 3778–3788, Aug. 2006.
- [9] V. M. Prabhakaran, R. Puri and K. Ramachandran, "Hybrid analog-digital strategies for source-channel broadcast," *Proc. 43rd Allerton Conference on Communication, Control and Computing*, Allerton, IL, Sep. 2005, pp. 94–113.
- [10] K. Narayanan, G. Caire and M. Wilson, "Duality between broadcasting with bandwidth expansion and bandwidth compression," *Proc. IEEE International Symposium on Information Theory*, Nice, France, Jul. 2007, pp. 1161–1165.
- [11] G. Caire, "Distortion region in common source broadcasting," Open question session, *Information Theory and Applications Inaugural Workshop*, San Diego, CA, Feb. 2006.
- [12] L. Ozarow, "On a source-coding problem with two channels and three receivers," *Bell Syst. Tech. Journal*, vol. 59, pp. 1909–1921, Dec. 1980.
- [13] C. Tian, S. Mohajer, and S. Diggavi, "Approximating the Gaussian multiple description rate region under symmetric distortion constraints," *IEEE Trans. Information Theory*, vol. 55, no. 8, pp. 3869–3891, Aug. 2009.
- [14] S. Avestimehr, G. Caire and D.N.C. Tse "On source-channel separation in networks," arXiv:0901.2082v1.
- [15] W. H. R. Equitz and T. M. Cover, "Successive refinement of information," *IEEE Trans. Information Theory*, vol. 37, no. 2, pp. 269–275, Mar. 1991.
- [16] C. Tian, A. Steiner, S. Shamai, and S. Diggavi, "Successive refinement via broadcast: optimizing expected distortion of a Gaussian source over a Gaussian fading channel," *IEEE Trans. Information Theory*, vol. 54, no. 7, pp. 2903–2918, Jul. 2008.
- [17] P. Bergmans, "A simple converse proof for the broadcast channels with additive white Gaussian noise," *IEEE Trans. Information Theory*, vol. 20, no. 2, pp. 279–280, Mar. 1974.
- [18] D. N. C. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *U.C. Berkeley technical report UCB/ERL M99/7, 1999*; available at <http://www.eecs.berkeley.edu/Pubs/TechRpts/1999/3578.html>
- [19] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York: Wiley, 1991.
- [20] J. Korner, K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Information Theory*, vol. 23, no. 1, pp. 60–64, Mar. 1977.
- [21] C. Tian, J. Chen, S. Diggavi and S. Shamai, "Optimality and approximate optimality of source-channel separation in networks," *Proc. IEEE International Symposium on Information Theory*, Austin, TX, USA, Jun. 2010, pp. 495–499. see also <http://arxiv.org/abs/1004.2648>

Chao Tian (S'00, M'05) received the B.E. degree in Electronic Engineering from Tsinghua University, Beijing, China, in 2000 and the M.S. and Ph. D. degrees in Electrical and Computer Engineering from Cornell University, Ithaca, NY in 2003 and 2005, respectively.

Dr. Tian was a postdoctoral researcher at Ecole Polytechnique Federale de Lausanne (EPFL) from 2005 to 2007. He joined AT&T Labs—Research, Florham Park, New Jersey in 2007, where he is now a Senior Member of Technical Staff. His research interests include multi-user information theory, joint source-channel coding, quantization design and analysis, as well as image/video coding and processing.

Suhas N. Diggavi (S'93, M'99) received the B. Tech. degree in electrical engineering from the Indian Institute of Technology, Delhi, India, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1998.

After completing his Ph.D., he was a Principal Member Technical Staff in the Information Sciences Center, AT&T Shannon Laboratories, Florham Park, NJ. Since then he had been in the faculty of the School of Computer and Communication Sciences, EPFL, where he directed the Laboratory for Information and Communication Systems (LICOS). He is currently a Professor, in the Department of Electrical Engineering, at the University of California, Los Angeles. His research interests include wireless communications networks, information theory, network data compression and network algorithms.

He is a recipient of the 2006 IEEE Donald Fink prize paper award, 2005 IEEE Vehicular Technology Conference best paper award and the Okawa foundation research award. He is currently an editor for ACM/IEEE Transactions on Networking and IEEE Transactions on Information Theory. He has 8 issued patents.

Shlomo Shamai (Shitz) (S'80, M'82, SM'89, F'94) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1975, 1981 and 1986 respectively.

During 1975-1985 he was with the Communications Research Labs in the capacity of a Senior Research Engineer. Since 1986 he is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, where he is now the William Fondiller Professor of Telecommunications. His research interests encompasses a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai (Shitz) is an IEEE Fellow, and the recipient of the 2011 Claude E. Shannon Award. He is the recipient of the 1999 van der Pol Gold Medal of the Union Radio Scientifique Internationale (URSI), and a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IT/COM societies paper award, the 2007 IEEE Information Theory Society Paper Award, the 2009 European Commission FP7, Network of Excellence in Wireless COMMunications (NEWCOM++) Best Paper Award, and the 2010 Thomson Reuters Award for International Excellence in Scientific Research. He is also the recipient of 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for the SHANNON THEORY OF THE IEEE TRANSACTIONS ON INFORMATION THEORY, and has also served on the Board of Governors of the Information Theory Society.