

The Dispersion of Superposition Coding for Gaussian Broadcast Channels

Ayse Unsal, Jean-Marie Gorce

► **To cite this version:**

Ayse Unsal, Jean-Marie Gorce. The Dispersion of Superposition Coding for Gaussian Broadcast Channels. IEEE Information Theory Workshop 2017, Nov 2017, Kaohsiung, Taiwan. hal-01643260

HAL Id: hal-01643260

<https://hal.archives-ouvertes.fr/hal-01643260>

Submitted on 21 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Dispersion of Superposition Coding for Gaussian Broadcast Channels

Ayşe Ünsal

INSA Lyon, France

ayse.unsal@insa-lyon.fr

Jean-Marie Gorce

INSA Lyon, France

jean-marie.gorce@insa-lyon.fr

Abstract—In this paper, we analyze the performance of superposition coding for Gaussian broadcast channels with finite blocklength. To this end, we adapt two different achievability bounds, the dependence testing and the $\kappa\beta$ bounds introduced by Polyanskiy *et al.* in 2010 to the broadcast setting. The distinction between these bounds lies in fixing either the input or the output distributions of the channel. For the first case of the dependence testing bound, an upper bound on the average error probability of the system is derived whereas for the latter, lower bounds on the maximal code sizes of each user are presented.¹

I. INTRODUCTION

It is well known that, superposition coding is optimum for degraded Gaussian broadcast channels (BC) in the asymptotic regime which is guaranteed by the vanishing error probability as the blocklength is allowed to grow boundlessly [1], [2]. In general, classical information theory studies and analyzes fundamental limits of communication systems subject to interference and channel noise using coding schemes that require asymptotically large blocklengths. In this case, the channel coding rate is obtained through the expectation of the *mutual information random variable* or the *information density* (to be defined later) based on the input distribution. The information density is a theoretic notion, some measure, of the channel depending on the input signal and the channel noise.

There is a practical requirement for coding with finite blocklength to which the classical information theory results are no longer applicable. The loss in the channel capacity due to coding with finite blocklength is explained with a second order statistic of the information density, namely the *channel dispersion term* denoted V for a single user channel. In [3], the author has shown that, the following equality holds for the single-user discrete-memoryless channels

$$\log M^*(n, \epsilon) = nC - Q^{-1}(\epsilon)\sqrt{nV} + O(\log n) \quad (1)$$

where n , ϵ , M and C denote the blocklength, the error probability, the maximum code size and the channel capacity, respectively. In [4], the above result is generalized to various point-to-point channel models including the AWGN channel where the channel dispersion term is defined as a function of the signal-to-noise ratio.

In [5], the authors extended the use of the channel dispersion term to multi-user channels and studied three network information problems as the Slepian–Wolf problem, the multiple-access (MAC) and *asymmetric broadcast* (ABC) for discrete channels only. In the multiuser settings of [5] the scalar dispersion term V , is shown to be a positive semi-definite matrix and inner bounds for the (n, ϵ) -capacity regions of the discrete memoryless MAC and ABC are presented. In [6], maximum achievable rate regions are presented for the Gaussian MAC. The result of [6] depends only on the first and second order statistic of the relevant mutual information random variables

represented by a scalar distribution where the authors firstly present the multi-access adaptation of the dependence testing bound by Polyanskiy *et al.* which constitutes the inspiration of the results presented in the first part of this paper.

In this paper, we consider a two–receiver degraded Gaussian broadcast channel (BC) in the setting of superposition coding with finite blocklength and we adapt the so-called *dependence testing* (DT) bound to the broadcast setting through random coding and threshold decoding to provide an upper bound on the average error probability of the system. To this end, we derive the channel dispersion terms for each error event that is defined through the information density functions based on the asymptotic capacity region. The overall bound consists of three different error events as the mis-detection, confusion and the input signals being confined to a pre-defined set. We derive the probability distributions of the channel dispersion terms for the mis-detection and confusion errors along with the corresponding parameters in order to derive the error probabilities. In addition, $\kappa\beta$ achievability bounds based on a binary hypothesis test to choose between two possible output distributions are derived on the maximal code sizes of both users. Section II describes the channel model for the addressed problem that is followed by the broadcast adaptation of the DT bound in section III. Lastly, in Section IV we give a short reminder of the original $\kappa\beta$ bound in an AWGN P2P channel which is followed by its generalization to a degraded Gaussian BC in the setting of superposition coding.

II. SYSTEM MODEL

We consider the following real-valued channel model

$$Y_{j,i} = X_i + Z_{j,i}, \quad (2)$$

for $j = 1, 2$ and $i = 1, \dots, n$ where X_i corresponds to the encoded messages that are $m_1 \in [1, 2, \dots, M_1]$ and $m_2 \in [1, 2, \dots, M_2]$, respectively subject to the channel noise terms $Z_{1,i} \sim \mathcal{N}(0, N_1)$ and $Z_{2,i} \sim \mathcal{N}(0, N_2)$ and the following channel transition probability density

$$P_{Y_j^n | X^n}(y_j^n | x^n) = (2\pi N_j)^{-n/2} e^{-\frac{\|y_j^n - x^n\|^2}{2N_j}}. \quad (3)$$

User j decodes message m_j using the observation Y_j^n . We assume that the channel of user 2, the weaker user, has the larger of the two noise variances, i.e. $N_2 > N_1$. For the achievable schemes considered here, the stronger user decodes the message m_2 prior to decoding its own message. We define the average error probability based on either user to be in error

$$\epsilon = \Pr[\hat{m}_1 \neq m_1 \cup \hat{m}_2 \neq m_2] \quad (4)$$

which is further bounded by the union bound as follows

$$\epsilon \leq \epsilon_1 + \epsilon_2 \quad (5)$$

¹This work was supported by Orange Labs under Grant no:F05151 and by the French National Research Agency project ARBURST.

where $\epsilon_j = \Pr[\hat{m}_j \neq m_j]$ with \hat{m}_j being the estimated m_j on the corresponding user. The error probability in decoding message 1 is expanded as

$$\begin{aligned} \epsilon_1 &= \Pr\left[\{\hat{m}_2 \neq m_2\}\right] \Pr\left[\{\hat{m}_1 \neq m_1\}|\{\hat{m}_2 \neq m_2\}\right] \\ &\quad + \Pr\left[\{\hat{m}_2 = m_2\}\right] \Pr\left[\{\hat{m}_1 \neq m_1\}|\{\hat{m}_2 = m_2\}\right] \\ &\stackrel{(a)}{\leq} \epsilon_{11} + \epsilon_{21} \end{aligned} \quad (6)$$

where \hat{m}_2 denotes the estimate of message 2 on user 1. In step (a), we denote $\epsilon_{11} = \Pr\left[\{\hat{m}_1 \neq m_1\}|\{\hat{m}_2 = m_2\}\right]$ and $\epsilon_{21} = \Pr\left[\{\hat{m}_2 \neq m_2\}\right]$. This expansion is based on successive decoding.

We consider two different cases for this channel model depending on the bounding technique either fixing the input distribution as in the case of DT bound or with fixed output distribution as the $\kappa\beta$ bound. These two cases are studied in the following subsections starting with the DT bound.

A. Random codebooks

This subsection presents the detailed model considered for the broadcast adaptation of the DT bound on the error probability. Here we have an input and two outputs with the n -letter channel model given as in (2) where

$$X^n = X_1^n + X_2^n \quad (7)$$

with $X_j^n \sim \mathcal{N}(0, P_j \mathbf{I}_n)$ for $P_1 = \alpha P$ and $P_2 = \bar{\alpha} P$ for $\alpha \in (0, 1)$ and $\bar{\alpha} = 1 - \alpha$.

The decoding rule is set for the threshold decoding as

$$i(x^n; y^n) > \gamma \quad (8)$$

where γ is some threshold and the information density or the mutual information random variable denoted $i(\cdot; \cdot)$ is defined as

$$i(x^n; y^n) = \log \frac{dP_{Y^n|X^n}(y^n|x^n)}{dP_{Y^n}(y^n)} \quad (9)$$

Here $P(\cdot|\cdot)$ and $P(\cdot)$ respectively denote conditional and marginal distributions.

Definition II.1. An $(n, M_1, M_2, \epsilon, P)$ code is composed of the messages $m_1 \in [1, 2, \dots, M_1]$ and $m_2 \in [1, 2, \dots, M_2]$ encoded into n i.i.d. sequences X_1 and X_2 that are chosen randomly according to $N(0, \alpha P)$ and $N(0, \bar{\alpha} P)$, respectively with the following average power constraints

$$\frac{1}{M_1} \sum_{i=1}^{M_1} \mathbb{E}\|x_{1,i}\|^2 \leq n\alpha P, \quad \frac{1}{M_2} \sum_{i=1}^{M_2} \mathbb{E}\|x_{2,i}\|^2 \leq n\bar{\alpha} P \quad (10)$$

and the average probability of error defined by (4).

B. Peak power

For the broadcast adaptation of the $\kappa\beta$ bound [4], we assume codebooks constructed as in (7) subject to the power constraint defined as

$$\|x^n(m_1, m_2)\|^2 = \sum_{i=1}^n |x_{1,i}(m_1, m_2) + x_{2,i}(m_2)|^2 \leq nP, \quad (11)$$

$\forall m_1, m_2$ where we assume equal power per codeword for the cloud center $X_{2,i}(m_2)$ for $i = 1, \dots, n$, $m_2 = 1, \dots, M_2$ and $m_1 = 1, \dots, M_1$ s.t.

$$\sum_{i=1}^n |x_{2,i}(m_2)|^2 = nP_2, \quad \forall m_2. \quad (12)$$

For $X_{1,i}(m_1, m_2)$, we have

$$\begin{aligned} \|x^n(m_1, m_2)\|^2 &= \|x_1(m_1, m_2)\|^2 + \|x_2(m_2)\|^2 \\ &\quad + 2\langle x_1(m_1, m_2), x_2(m_2) \rangle \\ &\stackrel{(a)}{=} nP_1(m_1, m_2) + nP_2 + 2n\rho(m_1, m_2)\sqrt{P_1(m_1, m_2)P_2} \\ &\leq nP \end{aligned} \quad (13)$$

with $\rho(m_1, m_2) \in [-1, 1]$ through the following definition

$$\sum_{i=1}^n x_{1,i}^2(m_1, m_2) = nP_1(m_1, m_2), \quad \forall m_1, m_2. \quad (14)$$

If we choose $X_1^n(m_1, m_2)$ in the null space of $X_2^n(m_2)$ for each m_2 then we can assign a constant power $P_1(m_1, m_2) = P - P_2$. Note that this is not orthogonal multiplexing but the effective dimensionality of the codebook $X_1^n(m_1, m_2)$ is $n - 1$ for each m_2 . This model is considered for the $\kappa\beta$ bound covered in Section IV. The feasible set of permissible inputs F_n is defined as

$$F_n := \{x^n : \|x^n\|^2 \leq nP\} \subset \mathbb{R}^n \quad (15)$$

III. DEPENDENCE TESTING BOUND –SUPERPOSITION CODING AND THRESHOLD DECODING

The dependence testing bound on the error probability of a P2P channel presented in [4, Theorem 20] is adapted to the model defined in Section II-A for a two-receiver Gaussian BC. The next theorem reminds the reader of the classical result by [1], [2], [7] for this channel in the asymptotic regime.

Theorem III.1. The capacity region of the Gaussian BC is given as

$$R_1 \leq C\left(\frac{\alpha P}{N_1}\right), \quad R_2 \leq C\left(\frac{\bar{\alpha} P}{\alpha P + N_2}\right) \quad (16)$$

where α is a constant confined in $[0, 1]$ and $\bar{\alpha} = 1 - \alpha$. The Shannon capacity $C(\cdot)$ is defined for a signal-to-noise ratio of P as $C(P) = \frac{1}{2} \log(1 + P)$.

The aim of this part is to provide a finite blocklength adaptation of this result and to introduce an achievability bound on the system error probability. We state our first result with the following theorem which provides the broadcast adaptation of [4, Theorem 20] using the mutual information functions in (16) achieving the capacity region of Theorem III.1. The average error probability defined in (4) with (6) is upper bounded based on the decoding rule (8).

Theorem III.2. An $(n, M_1, M_2, \epsilon, P)$ code exists for the channel as described in Definition II.1 with the average error probability satisfying

$$\epsilon \leq \Pr[i(X_2^n; Y_2^n) < \gamma_2] + \eta_2 \Pr[i(X_2^n; \bar{Y}_2^n) > \gamma_2] \quad (17)$$

$$+ \Pr[i(X^n; Y_1^n|X_2^n) < \gamma_1] + \eta_1 \Pr[i(X^n; \bar{Y}_1^n|X_2^n) > \gamma_1] \quad (18)$$

$$+ \Pr[i(X_2^n; Y_1^n) < \gamma_2] + \eta_2 \Pr[i(X_2^n; \bar{Y}_1^n) > \gamma_2] \quad (19)$$

$$+ \Pr[X^n \notin F] \quad (20)$$

where \bar{Y}_j^n follows the same distribution as the output signal Y_j^n but is independent of the input X^n , the thresholds are set as $\gamma_j = \log \eta_j$ and $\eta_j = (M_j - 1)/2$ for $j = 1, 2$ and F denotes the set of permissible inputs.

Proof. The capacity region of the two-receiver Gaussian BC given by Theorem III.1 is achieved through the information densities $i(X^n; Y_1^n|X_2^n)$ and $i(X_2^n; Y_2^n)$ for user 1 and 2, respectively. $i(X^n; Y_1^n|X_2^n)$ and $i(X_2^n; Y_2^n)$ are defined using

(9). Expectation of these information densities bring out the mutual information functions that compose the asymptotic capacity region in Theorem III.1. Unlike the asymptotic case, the error event for decoding message 2 using the observation of the strong user Y_1^n cannot be ignored with the corresponding error probability ϵ_{21} as shown by (6). Applying the threshold decoding rule defined by (8) to these three functions, we obtain

$$i(X_2^n; Y_2^n) > \gamma_2, \quad (21)$$

$$i(X^n; Y_1^n | X_2^n) > \gamma_1, \quad (22)$$

$$i(X_2^n; Y_1^n) > \gamma_2. \quad (23)$$

Combining the dependence testing bound proposed in [4, Theorem 20] with (21), (22) and (23) yields (17), (18) and (19), respectively. The proposed bound consists of three types of errors. The first case of *mis-detection*, which is shown by the first terms in (17)-(19), corresponds to the information density of a correct pair of input-output to remain below the threshold. In the second case, the *confusion* error, shown by the second terms in (17)-(19), occurs when a pair exceeds the threshold where the output does not depend on the input signal that was transmitted. The probabilities (17), (18) and (19) are $\epsilon_2 = \Pr[\hat{m}_j \neq m_j]$, $\epsilon_{11} = \Pr[\{\hat{m}_1 \neq m_1\} | \{\hat{m}_2 = m_2\}]$ and $\epsilon_{21} = \Pr[\{\hat{m}_2 \neq m_2\}]$, respectively. Lastly, due to the input signal defined by (7), we have $\Pr[X^n \notin F]$ given by (20) as the last error event of the proposed bound that represents the probability of the channel input X^n not being chosen from the feasible set denoted F . \square

Hereafter, starting with decoding message 2 on the weaker user, we evaluate the probabilities of each error event considered in the proposed bound in detail.

1) *Decoding m_2* : This scheme uses successive decoding where initially message 2 is decoded using the observation of Y_2^n . This event is represented by the information density $i(X_2^n; Y_2^n)$. The terms in the first line of Theorem III.2 are evaluated in the following starting with the mis-detection error event $i(X_2^n; Y_2^n) < \gamma_2$. We derive $i(X_2^n; Y_2^n)$ as

$$\begin{aligned} \frac{1}{n} i(X_2^n; Y_2^n) &= \frac{1}{n} \log \frac{P_{Y_2^n | X_2^n}(Y_2^n | X_2^n)}{P_{Y_2^n}(Y_2^n)} \\ &= \frac{1}{n} \log \frac{(2\pi)^{-n/2} (\alpha P + N_2)^{-n/2} e^{-\frac{\|Y_2^n - X_2^n\|^2}{2(\alpha P + N_2)}}}{(2\pi)^{-n/2} (P + N_2)^{-n/2} e^{-\frac{\|Y_2^n\|^2}{2(P + N_2)}}} \\ &= \frac{1}{n} \log \left[\left(\frac{P + N_2}{\alpha P + N_2} \right)^{n/2} e^{\frac{\|Y_2^n\|^2}{2(P + N_2)} - \frac{\|Y_2^n - X_2^n\|^2}{2(\alpha P + N_2)}} \right] \\ &= C \left(\frac{\bar{\alpha} P}{\alpha P + N_2} \right) + \frac{1}{n} \log e \left(\frac{\|Y_2^n\|^2}{2(P + N_2)} - \frac{\|Y_2^n - X_2^n\|^2}{2(\alpha P + N_2)} \right) \end{aligned} \quad (24)$$

Let us denote the additional term to the Shannon capacity $C \left(\frac{\bar{\alpha} P}{\alpha P + N_2} \right)$ by v_{22} . (24) is rewritten as

$$\frac{1}{n} i(X_2^n; Y_2^n) = C \left(\bar{\alpha} P / (\alpha P + N_2) \right) + v_{22} \quad (25)$$

As for v_{22} , we get

$$v_{22} \stackrel{(a)}{=} \frac{1}{n} \log e \left[f_1 \|X_1^n + X_2^n + Z_2^n\|^2 + f_2 \|X_1^n + Z_2^n\|^2 \right]$$

$$\begin{aligned} &\stackrel{(b)}{=} \frac{1}{n} \log e \left[f_1 \|X_2^n + \tilde{Z}_2^n\|^2 + f_2 \|\tilde{Z}_2^n\|^2 \right] \\ &= \frac{1}{n} \log e \left[f_1 \|X_2^n\|^2 + (f_1 + f_2) \|\tilde{Z}_2^n\|^2 + 2f_1 \langle X_2^n, \tilde{Z}_2^n \rangle \right] \\ &= \frac{1}{n} \log e \sum_{i=1}^n \begin{bmatrix} X_{2,i} & \tilde{Z}_{2,i} \end{bmatrix} \mathbf{A} \begin{bmatrix} X_{2,i} & \tilde{Z}_{2,i} \end{bmatrix}^T \\ &\stackrel{(c)}{=} \frac{1}{n} \log e \sum_{i=1}^n [\mu_{1,i} \ \mu_{2,i}] \mathbf{P}_2 \mathbf{A} \mathbf{P}_2 [\mu_{1,i} \ \mu_{2,i}]^T \end{aligned} \quad (26)$$

In step (a) we substituted $f_1 = [2(P + N_2)]^{-1}$, $f_2 = -[2(\alpha P + N_2)]^{-1}$, in step (b) we defined $\tilde{Z}_2^n = X_1^n + Z_2^n$ and in step (c), the standard Gaussian random variables $\mu_{1,i}, \mu_{2,i}$ are introduced where \mathbf{A} and \mathbf{P}_2 are 2×2 matrices given by

$$\mathbf{A} = \begin{bmatrix} f_1 & f_1 \\ f_1 & f_1 + f_2 \end{bmatrix}, \quad (27)$$

and $\mathbf{P}_2 = \text{diag}(\sqrt{\bar{\alpha} P}, \sqrt{\alpha P} + \sqrt{N_2})$. The eigenvalues of the matrix product $\mathbf{A}' = \mathbf{P}_2 \mathbf{A} \mathbf{P}_2$ yield $\lambda_1 = \frac{\sqrt{\bar{\alpha} P}}{2\sqrt{(P + N_2)}}$, $\lambda_2 = -\lambda_1$. Rewriting \mathbf{A}' through its eigen decomposition $\mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$ where $\mathbf{\Lambda}$ is a diagonal matrix with the eigenvalues on the diagonal elements and \mathbf{U} is the corresponding eigenvector, we define $[\tilde{\mu}_{1,i}^2, \tilde{\mu}_{2,i}^2]^T$ which is the product $\mathbf{U}^T [\mu_{1,i}, \mu_{2,i}]^T$. Using λ_1 and λ_2 , the dispersion term v_{22} yields

$$v_{22} = \frac{1}{n} \log e \times \lambda_1 \sum_{i=1}^n (\tilde{\mu}_{1,i}^2 - \tilde{\mu}_{2,i}^2). \quad (28)$$

The weighted difference of chi-squares with n degrees of freedom follows the *variance-gamma distribution*² with the location parameter m and asymmetry parameter denoted by b are both 0 and $\delta = \theta = \sqrt{\frac{P + N_2}{\bar{\alpha} P}}$. Substituting v_{22} into (25), the probability of mis-detection error in decoding m_2 on user 2 yields

$$\Pr[i(X_2^n; Y_2^n) < \gamma_2] = \Pr \left[v_{22} < \frac{\gamma_2 - nC \left(\frac{\bar{\alpha} P}{\alpha P + N_2} \right)}{\log e} \right] \quad (29)$$

The confusion error event represents the case where the information density exceeds the threshold according to the decoding rule (8) where the channel output does not depend on the input signal that is transmitted over the channel. More precisely, for a codeword C_k^n from a given codebook $(C_0^n, C_1^n, \dots, C_{M_j-1}^n)$ for user j , the confusion error event represents the information density of interest to exceed the threshold with C_l^n given that the input is C_k^n for all $l < k$. In decoding m_2 , we have for the confusion error event

$$\begin{aligned} \frac{1}{n} i(X_2^n; \bar{Y}_2^n) &= \frac{1}{n} \log \frac{P_{Y_2^n | X_2^n = C_k^n}(Y_2^n | X_2^n = C_k^n)}{P_{Y_2^n}(Y_2^n)} \\ &= \frac{1}{n} \log \frac{(2\pi)^{-n/2} (2(\alpha P + N_2))^{-n/2} e^{-\frac{\|Y_2^n - C_k^n\|^2}{2(\alpha P + N_2)}}}{(2\pi)^{-n/2} (2(P + N_2))^{-n/2} e^{-\frac{\|Y_2^n\|^2}{2(P + N_2)}}} \\ &= C \left(\frac{\bar{\alpha} P}{(\alpha P + N_2)} \right) + v_{c,2}. \end{aligned} \quad (30)$$

²The variance-gamma distribution with $\lambda > 0$ degrees of freedom, the real-valued location parameter m and the asymmetry parameter b , is defined as

$$f(t) = \frac{\theta^{2\lambda} |t - m|^{\lambda - 1/2} K_{\lambda - 1/2}(\delta |t - m|)}{\sqrt{\pi} \Gamma(\lambda) (2\delta)^{\lambda - 1/2}} e^{b(t - m)}, \quad -\infty < t < \infty$$

where K_λ and Γ denote the modified Bessel function of the second kind and the gamma function, respectively for real δ , $\theta = \sqrt{\delta^2 - b^2} > 0$.

The additional term $v_{c,2}$ is further derived as

$$\begin{aligned} v_{c,2} &= \frac{1}{n} \log e \left(g_1 \|C_k^n + \tilde{Z}_2^n\|^2 + g_2 \|C_k^n - C_l^n + \tilde{Z}_2^n\|^2 \right) \\ &= \frac{1}{n} \log e \sum_{i=1}^n [C_{k,i} C_{l,i} \tilde{Z}_{2,i}] \mathbf{A}_c [C_{k,i} C_{l,i} \tilde{Z}_{2,i}]^T \\ &= \frac{1}{n} \log e \sum_{i=1}^n [\mu_{1,i} \mu_{2,i} \mu_{3,i}] \mathbf{P}_{c2} \mathbf{A}_c \mathbf{P}_{c2} [\mu_{1,i} \mu_{2,i} \mu_{3,i}]^T \end{aligned}$$

where $g_1 = [2(P + N_2)]^{-1}$ and $g_2 = -[2(\alpha P + N_2)]^{-1}$ with $\mathbf{P}_{c2} = \text{diag}(\sqrt{\alpha P}, \sqrt{\alpha P}, \sqrt{\alpha P} + \sqrt{N_2})$

$$\mathbf{A}_c = \begin{bmatrix} (g_1 + g_2) & -g_2 & (g_1 + g_2) \\ -g_2 & g_2 & -g_2 \\ (g_1 + g_2) & -g_2 & (g_1 + g_2) \end{bmatrix} \quad (31)$$

and $\mu_{j,i} \sim N(0, 1)$ for $j = 1, \dots, 4$. We remind the reader that $\tilde{Z}_2^n = X_1^n + Z_2^n$. The non-zero eigenvalues of the matrix product $\mathbf{P}_{c2} \mathbf{A}_c \mathbf{P}_{c2}$ are $\lambda_{c1} = \frac{-(\alpha P + \sqrt{\alpha P(P + N_2)})}{2(\alpha P + N_2)}$ and $\lambda_{c2} = \frac{-\alpha P + \sqrt{\alpha P(P + N_2)}}{2(\alpha P + N_2)}$. The dispersion term $v_{c,2}$ is obtained in the following form which constitutes a weighted sum of chi-squared variables with n degrees of freedom, a special case of the gamma distribution.

$$v_{c,2} = \frac{1}{n} \log e \sum_{i=1}^n [\lambda_{c1} \mu_{1,i}^2 + \lambda_{c2} \mu_{2,i}^2] \quad (32)$$

The confusion error probability for decoding m_2 on user 2 is

$$\Pr [i(X_2^n; \bar{Y}_2^n) > \gamma_2] = \Pr \left[v_{c,2} > \frac{\gamma_2 - nC \left(\frac{\alpha P}{\alpha P + N_2} \right)}{\log e} \right] \quad (33)$$

Let us denote the threshold functions in the right hand sides (r.h.s.) of (29) and (33) by ζ_2 . Combining (29) and (33), we obtain the total error probability in decoding message 2 given by (17), which in fact is $\epsilon_2 = \Pr [\hat{m}_2 \neq m_2]$.

$$\epsilon_2 = \Pr [v_{22} < \zeta_2] + \eta_2 \Pr [v_{c,2} > \zeta_2] \quad (34)$$

2) *Decoding m_1* : In order to reconstruct m_1 , in the setting of superposition coding, the strong user that is subject to a channel with a smaller noise variance, firstly decodes m_2 with $i(X_2^n; Y_1^n)$. Note that in the asymptotic case, the error probability for this event is insignificant. In a similar fashion to $i(X_2^n; Y_2^n)$, we have for $i(X_2^n; Y_1^n)$

$$\frac{1}{n} i(X_2^n; Y_1^n) = C \left(\frac{\alpha P}{\alpha P + N_1} \right) + v_{21} \quad (35)$$

where v_{21} follows a variance-gamma distribution with n degrees of freedom, $m = b = 0$ and $\delta = \theta = \sqrt{\frac{P + N_1}{\alpha P}}$. Substituting v_{21} into (35), the probability of mis-detection error in decoding m_2 on user 1 yields

$$\Pr [i(X_2^n; Y_1^n) < \gamma_2] = \Pr [v_{21} < \zeta_{21}] \quad (36)$$

with $\zeta_{21} = \frac{\gamma_2 - nC \left(\frac{\alpha P}{\alpha P + N_1} \right)}{\log e}$. As for the confusion error probability of decoding message 2 on the stronger user, we have

$$\Pr [i(X_2^n; \bar{Y}_1^n) > \gamma_2] = \Pr [v_{c,21} > \zeta_{21}] \quad (37)$$

where $v_{c,21} = \frac{1}{n} \log e \sum_{i=1}^n [\lambda_{c5} \mu_{1,i}^2 + \lambda_{c6} \mu_{2,i}^2]$ for $\mu_{1,i}^2$ and $\mu_{2,i}^2$ being chi-squared variables with n degrees of freedom, $\lambda_{c5} = \frac{-(\alpha P + \sqrt{\alpha P(P + N_1)})}{2(\alpha P + N_1)}$ and $\lambda_{c6} = \frac{-\alpha P + \sqrt{\alpha P(P + N_1)}}{2(\alpha P + N_1)}$. The

total error probability in decoding m_2 on the first user given by (19) becomes

$$\epsilon_{21} = \Pr [v_{21} < \zeta_{21}] + \eta_2 \Pr [v_{c,21} > \zeta_{21}] \quad (38)$$

which is equivalent to $\epsilon_{21} = \Pr [\hat{m}_2 \neq m_2]$. The derivation follows the same way as in (30). Therefore it is omitted due to space limitations.

For the stronger user to decode its intended message m_1 using the observation of Y_1^n given the information of X_2^n , we derive the relevant information density to evaluate the probability of mis-detecting m_1 as

$$\begin{aligned} \frac{1}{n} i(X^n; Y_1^n | X_2^n) &= \frac{1}{n} \log \frac{P_{Y_1^n | X_1^n, X_2^n}(Y_1^n | X_1^n, X_2^n)}{P_{Y_1^n | X_2^n}(Y_1^n | X_2^n)} \\ &= C(\alpha P / N_1) + v_{11} \end{aligned} \quad (39)$$

The additional term v_{11} yields

$$\begin{aligned} v_{11} &= \frac{1}{n} \log e [f_3 \|X_1^n + Z_1^n\|^2 + f_4 \|Z_1^n\|^2] \\ &= \frac{1}{n} \log e \sum_{i=1}^n [X_{1,i} Z_{1,i}] \mathbf{B} [X_{1,i} Z_{1,i}]^T \\ &\stackrel{(d)}{=} \frac{1}{n} \log e \sum_{i=1}^n [\mu_{1,i} \mu_{4,i}] \mathbf{P}_1 \mathbf{B} \mathbf{P}_1 [\mu_{1,i} \mu_{4,i}]^T \end{aligned} \quad (40)$$

for $f_3 = \frac{1}{2(\alpha P + N_1)}$ and $f_4 = -\frac{1}{2N_1}$ and in step (d) we define the standard Gaussian random variables $\mu_{1,i}, \mu_{4,i}$ with 2×2 matrices \mathbf{B} and \mathbf{P}_1 given by

$$\mathbf{B} = \begin{bmatrix} f_3 & f_3 \\ f_3 & f_3 + f_4 \end{bmatrix}, \quad (41)$$

and $\mathbf{P}_1 = \text{diag}(\sqrt{\alpha P}, \sqrt{N_1})$. The eigenvalues of $\mathbf{B}' = \mathbf{P}_1 \mathbf{B} \mathbf{P}_1$ are $\lambda_4 = \frac{1}{2} \sqrt{\frac{\alpha P}{\alpha P + N_1}}$ and $\lambda_5 = -\frac{1}{2} \sqrt{\frac{\alpha P}{\alpha P + N_1}}$. Finally we obtain the channel dispersion term v_{11} in the following form

$$v_{11} = \frac{1}{n} \log e \times \lambda_4 \sum_{i=1}^n (\tilde{\mu}_{1,i}^2 - \tilde{\mu}_{4,i}^2) \quad (42)$$

since $\lambda_4 = -\lambda_5$ and $\tilde{\mu}_{l,i}$'s for $l = 1, 4$ are obtained through substitution of the eigen decomposition of \mathbf{B}' . v_{11} follows the variance gamma distribution with n degrees of freedom, $m = b = 0$ and $\delta = \theta = \sqrt{\frac{\alpha P + N_1}{\alpha P}}$.

In the evaluation of the achievability bound proposed in Theorem III.2 in order to derive the confusion error probability in detecting m_1 , we have

$$\frac{1}{n} i(X^n; \bar{Y}_1^n | X_2^n) = C(\alpha P / N_1) + v_{c,1} \quad (43)$$

The dispersion term $v_{c,1}$ is obtained in the form of $\frac{1}{n} \log e \sum_{i=1}^n [\lambda_{c3} \mu_{1,i}^2 + \lambda_{c4} \mu_{2,i}^2]$ with $\lambda_{c3} = \frac{-\alpha P + \sqrt{\alpha P(\alpha P + N_1)}}{2N_1}$ and $\lambda_{c4} = \frac{-\alpha P + \sqrt{\alpha P(\alpha P + N_1)}}{2N_1}$. Finally, summing up the probabilities of mis-detection and confusion for user 1 to detect its intended message m_1 , we obtain

$$\epsilon_{11} = \Pr [v_{11} < \zeta_{11}] + \eta_1 \Pr [v_{c,1} > \zeta_{11}]. \quad (44)$$

for $\zeta_{11} = \frac{\gamma_1 - nC \left(\frac{\alpha P}{N_1} \right)}{\log e}$.

Given the power constraint (10), the probability of the input being confined to the set F_n yields $\Pr \{X^n \notin F_n\} = 0$. In summary, combining ϵ_2 , ϵ_{21} and ϵ_{11} respectively given by (34), (38) and (44), the derivation of the achievability bound of Theorem III.2 is complete.

IV. $\kappa\beta$ BOUND-ACHIEVABILITY

The idea behind the $\kappa\beta$ bound is the optimality of the binary hypothesis testing problem and the Neyman-Pearson lemma. In this part, we adapt the $\kappa\beta$ bound proposed in [4, Theorem 41] to the Gaussian BC. The degraded channel can be considered as two AWGN P2P channels since the channel between the transmitter and the weak receiver is the cascade of the channel from the transmitter to the strong receiver and the one from the strong receiver to the weak receiver. The input distribution is no longer limited to the Gaussian distribution. To apply superposition coding and obtain the two-user equivalent of the $\kappa\beta$ bound, we consider the mutual information functions from Section III.

Let $P_{Y_j^n}$ denote $\mathcal{N}(0, \sigma_{Y_j}^2 \mathbf{I}_n)$ with $\sigma_{Y_j}^2 = P + N_j$ for $j = 1, 2$. Using the definition of (9), $i(X_2^n; Y_2^n)$ under $P_{Y_2^n}$ is given as

$$G_{n_2} = \frac{n}{2} \log \sigma_{Y_2}^2 - \frac{nP_2}{P_1 + N_2} \frac{\log e}{2} + \frac{1}{2} \log e \sum_{i=1}^n \left[(1 - \sigma_{Y_2}^2) S_i^2 + 2\sqrt{P_2/(P_1 + N_2)} \sigma_{Y_2} S_i \right] \quad (45)$$

where $S_i \sim \mathcal{N}(0, 1)$ for $i = 1, \dots, n$. Under the conditional distribution $P_{Y_2^n | X_2^n}$ the same function yields

$$H_{n_2} = \frac{n}{2} \log(\sigma_{Y_2}^2) + \frac{nP_2}{P_1 + N_2} \frac{\log e}{2\sigma_{Y_2}^2} + \frac{\log e}{2\sigma_{Y_2}^2} \sum_{i=1}^n \left[(1 - \sigma_{Y_2}^2) S_i^2 + 2\sqrt{P_2/(P_1 + N_2)} S_i \right] \quad (46)$$

In a similar manner, the mutual information random variable $i(X^n; Y_1^n | X_2^n)$ under $P_{Y_1^n}$ becomes

$$G_{n_1} = \frac{n}{2} \log \sigma_{Y_1}^2 - \frac{nP_1}{N_1} \frac{\log e}{2} + \frac{1}{2} \log e \sum_{i=1}^n \left[(1 - \sigma_{Y_1}^2) S_i^2 + 2\sqrt{P_1/N_1} \sigma_{Y_1} S_i \right] \quad (47)$$

where we define $Y_1'^n = Y_1^n - X_2^n$ with $\sigma_{Y_1'}^2 = P_1 + N_1$ since P_2 is subtracted off the sum power P . Under the conditional distribution $P_{Y_1'^n | X_1^n}$ we obtain the information density as

$$H_{n_1} = \frac{n}{2} \log(\sigma_{Y_1'}^2) + \frac{P_1}{N_1} \frac{\log e}{2\sigma_{Y_1'}^2} + \frac{\log e}{2\sigma_{Y_1'}^2} \sum_{i=1}^n \left[(1 - \sigma_{Y_1'}^2) S_i^2 + 2\sqrt{P_1/N_1} S_i \right] \quad (48)$$

Lastly, $i(X_2^n; Y_1^n)$ is respectively denoted by G_{n_3} and H_{n_3} under $P_{Y_1^n}$ and $P_{Y_1^n | X_2^n}$, which will not be given here explicitly.

Using Theorem 40 of [4] for the degraded Gaussian BC, we define β functions as follows

$$\beta_{a,k} = \Pr[G_{n_k} \geq \gamma_k] \quad (49)$$

where $\Pr[H_{n_k} \geq \gamma_k] = a_k$ for $a_k = 1 - \epsilon_k$ as defined by (4) and with any positive γ_k and $k = 1, 2, 3$. The resulting achievability bound for the Gaussian BC is stated in the next theorem.

Theorem IV.1. For any ϵ_k , $n \geq 1$, $\tau_k \in [0, 1]$, $k = 1, 2, 3$, and the chosen P_{Y_j} for $j = 1, 2$ with F_n as defined by (15), the maximal code sizes denoted M_j^* of the j^{th} user in a two

receiver Gaussian degraded BC are bounded by

$$M_1^* \geq \frac{\kappa_{\tau_1,1}(F_n, P_{Y_1^n})}{\beta_{1-\epsilon_1+\tau_1,1}(x, P_{Y_1^n})} \quad (50)$$

$$M_2^* \geq \max \left\{ \frac{\kappa_{\tau_2,2}(F_n, P_{Y_2^n})}{\beta_{1-\epsilon_2+\tau_2,2}(x, P_{Y_2^n})}, \frac{\kappa_{\tau_3,3}(F_n, P_{Y_2^n})}{\beta_{1-\epsilon_3+\tau_3,3}(x, P_{Y_2^n})} \right\} \quad (51)$$

where

$$\kappa_{\tau_k,k}(F_n, P_{Y_j^n}) = P_{0,k} \left[\frac{p_{1,k}(r)}{p_{0,k}(r)} \geq \psi_k \right] \quad (52)$$

with ψ_k satisfying $P_{1,k} \left[\frac{p_{1,k}(r)}{p_{0,k}(r)} \geq \psi_k \right] = \tau_k$. The probability distributions $p_{0,k}(r)$ and $p_{1,k}(r)$ are defined as $p_{0,k}(r) = \frac{1}{\Gamma(n/2)\omega_k^{n/2}} r^{n/2-1} e^{-r/\omega_k}$ and $p_{1,k}(r) = \frac{1}{2} e^{(r+v_k)/2} \left(\frac{r}{v_k}\right)^{n/4-1/2} I_{n/2-1}(\sqrt{v_k r})$ with the modified Bessel function of the first kind $I_b(y) = (y/2)^b \sum_{j=0}^{\infty} \frac{(y^2/4)^j}{j! \Gamma(b+j+1)}$ and the following parameters for $k = 1$, $\omega_1 = 2(N_1 + P_1)$ and $v_1 = \frac{nP_1}{N_1}$, for $k = 2$, $\omega_2 = 2(N_2 + P)$ and $v_2 = \frac{nP_2}{P_1 + N_2}$ whereas for $k = 3$, $\omega_3 = 2(N_1 + P)$ with $v_3 = \frac{nP_2}{P_1 + N_1}$.

Proof. Detailed proofs of the general case for a P2P-AWGN channel can be found in the original paper [4, Theorems 25, 40 and 42]. $\beta_{1-\epsilon_k+\tau_k,k}(x, P_{Y_j^n})$ for $a = 1 - \epsilon_k + \tau_k$ given by (49) is derived using the functions (45)-(48). As for evaluating $\kappa_{\tau_k,k}$'s the following definitions on the distribution of the output signals $P_{0,k}$ and the distribution of the output given the input $P_{1,k}$ are set using superposition coding as follows

$$P_{0,1} \sim \sum_{i=1}^n (P_1 + N_1) S_i^2, \quad P_{1,1} \sim \sum_{i=1}^n \left(\sqrt{P_1} + \sqrt{N_1} S_i \right)^2$$

$$P_{0,2} \sim \sum_{i=1}^n (P + N_2) S_i^2, \quad P_{1,2} \sim \sum_{i=1}^n \left(\sqrt{P_2} + \sqrt{(P_1 + N_2)} S_i \right)^2$$

$$P_{0,3} \sim \sum_{i=1}^n (P + N_1) S_i^2, \quad P_{1,3} \sim \sum_{i=1}^n \left(\sqrt{P_2} + \sqrt{(P_1 + N_1)} S_i \right)^2$$

□

V. DISCUSSION

We introduced the dependence testing and the $\kappa - \beta$ bounds for the two-receiver Gaussian BC. Both types of bounds can be easily extended to scenarios with k -users for $k \geq 2$. It is clear that unlike the well-known optimality of superposition coding for Gaussian BC, k user adaptation of the proposed bounds in Theorems III.2 and IV.1 would perform poorly due to the additional error event using the successive cancellation.

REFERENCES

- [1] P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Transactions on Information Theory*, vol. 19, pp. 197-207, December 1973.
- [2] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Probl. Peredachi Inf.*, vol. 10, pp. 3-14, 1974.
- [3] V. Strassen, "Asymptotische Abschätzungen in Shannon's Informationstheorie," in *Transactions of the Third Prague Conference on Information Theory*, 1962, pp. 689-723.
- [4] Y. Polyanskiy, H. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, pp. 2307-2359, December 2010.
- [5] V. Tan and O. Kosut, "On the dispersions of three network information theory problems," *IEEE Transactions on Information Theory*, vol. 60, pp. 881-903, February 2014.
- [6] E. MolavianJazi and J. Laneman, "A second-order achievable rate region for Gaussian multi-access channels via a central limit theorem for functions," *IEEE Transactions on Information Theory*, vol. 61, pp. 6719-6733, December 2015.
- [7] T. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. 18, pp. 2-14, January 1972.