$\kappa\beta$ Bounds for Gaussian Broadcast Channels with Finite Blocklength

Ayse UNSAL, Jean-Marie GORCE

Univ Lyon, INSA Lyon, Inria, CITI, France
6 Avenue des Arts, 69621, Villeurbanne, France
ayse.unsal@insa-lyon.fr, jean-marie.gorce@insa-lyon.fr

Résumé – Dans cet article, nous étudions les performances atteignables d'une transmission par superposition de codes dans un canal Gaussien à diffusion avec deux récepteurs et en régime à longueur de codes finie. Dans ce but, nous adaptons la borne atteignable introduite par Polyanskiy *et al.* en 2010 pour un canal point à point au canal à diffusion. De plus, une nouvelle borne supérieure (converse) est proposée et caractérise les débits conjoints non atteignables en fonction de la longueur des codes et pour une probabilité d'erreur donnée.

Abstract – We analyze the achievable performance of superposition coding in a two–receiver Gaussian broadcast channel (BC) with finite blocklength. To this end, we adapt the achievability bound on maximal code size of a point-to-point (P2P) channel introduced by Polyanskiy *et al.* in 2010 to the broadcast setting. Additionally, a new converse bound on maximal code sizes of each user in a two-user Gaussian BC is introduced for a given probability of error.

1 Introduction

The strongest results from classical information theory hold for vanishing error probability and asymptotically large blocklengths or equivalently signal dimensionality in analyzing the performance limits of communication systems. In the finite-dimensional case imposed by small payload transmission, many of the classical error-probability bounds stemming from information theoretic results are no longer applicable, especially for multiuser channels due to delay and complexity limitations. In the asymptotic regime, the channel coding rate approaches the mutual information which is the expectation of *the information density* or *the mutual information random variable* based on the joint probability density function of the input and output. The mutual information random variable is a theoretic notion, some measure, of the channel depending on the input signal and the channel noise to be defined later. It is shown by Strassen in 1962 that the variance of the mutual information random variable, namely *the channel dispersion term* denoted V, plays a crucial role for a discrete memoryless channel (DMC) in the finite blocklength regime and that the logarithm of the maximal code size $M^*(n,\epsilon)$ is given as a function of the error probability ϵ and the blocklength n as follows

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

where C denotes the channel capacity. In [1], Polyanskiy *et al.* analyzed the performance of coding with finite blocklength for various types of point-to-point (P2P) channels.

In [2], the results of [1] is adapted to the Slepian-Wolf problem, multiple-access channel (MAC) and asymmetric broadcast channel (ABC). In [3], maximum achievable rate regions are presented for the DM–MAC and the Gaussian MAC.

In this paper, we consider a two–receiver degraded Gaussian broadcast channel (BC) using superposition coding with finite blocklength and adapt the so-called $\kappa\beta$ bound to the two-user broadcast setting. $\kappa\beta$ achievability bounds, which are based on a binary hypothesis test to choose between two possible output distributions, are derived on the maximal code sizes for each user as functions of the error probability and the blocklength. In addition, we adapt the converse bound on the maximal code size of the P2P AWGN channel by Polyanskiy *et al.* to the two–receiver Gaussian BC.

The next section describes the considered model for the addressed problem.Section 3 provides a short reminder of the original $\kappa\beta$ bound for the AWGN P2P channel which is followed by its generalization to the degraded Gaussian BC in the setting of superposition coding. New converse bounds on the code sizes of each user and on their product are presented in Section 4 in addition to the numerical comparison results of the obtained bounds.

2 System Model

We consider the real-valued channel model

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

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

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

User j decodes message $\hat{m}_j$ from observation Y_j^n. We assume that the channel of user 2, the weaker user, has the larger of the two noise variances. For the achievable scheme considered here, the stronger user decodes the message $\hat{\hat{m}}_2$ prior to decoding its own message. We have the overall average error probability defined as

$$\epsilon = \epsilon_1 + \epsilon_2 \tag{4}$$

where $\epsilon_j = \Pr[\hat{m}_j \neq m_j]$. For the achievability scheme based on successive decoding using $\hat{\hat{m}}_2$ we have for ϵ_1

$$\begin{aligned}\epsilon_1 &= \Pr\left[\{\hat{\hat{m}}_2 \neq m_2\}\right]\Pr\left[\{\hat{m}_1 \neq m_1\}|\{\hat{\hat{m}}_2 \neq m_2\}\right] \\ &\quad + \Pr\left[\{\hat{\hat{m}}_2 = m_2\}\right]\Pr\left[\{\hat{m}_1 \neq m_1\}|\{\hat{\hat{m}}_2 = m_2\}\right] \\ &\leq \epsilon_{11} + \epsilon_{21}\end{aligned} \tag{5}$$

where $\epsilon_{11} = \Pr\left[\{\hat{m}_1 \neq m_1\}|\{\hat{\hat{m}}_2 = m_2\}\right]$ and $\epsilon_{21} = \Pr\left[\{\hat{\hat{m}}_2 \neq m_2\}\right]$. The decoding rule is set for the threshold decoding as

$$i(x^n; y^n) > \log \eta \tag{6}$$

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

$$i(x^n; y^n) :== \log \frac{dP_{Y^n|X^n}(y^n|x^n)}{dP_{Y^n}(y^n)} \tag{7}$$

Here $P(.,.)$, $P(.|.)$ and $P(.)$ respectively denote joint, conditional and marginal distributions.

For the broadcast adaptation of the $\kappa\beta$ bound [1], we also assume codebooks constructed as $X^n = X_1^n + X_2^n$ and we define the power constraint on the channel input

$$\|x^n(m_1, m_2)\|^2 = \sum_{i=1}^{n} |x_{1,i}(m_1, m_2) + x_{2,i}(m_2)|^2 \leq nP, \tag{8}$$

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

$$\sum_{i=1}^{n} |x_{2,i}(m_2)|^2 = nP_2,\ \forall m_2. \tag{9}$$

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

$$\begin{aligned}&\|x^n(m_1, m_2)\|^2 = \|x_1(m_1, m_2)\|^2 + \|x_2(m_2)\|^2 \\ &+ 2\langle x_1(m_1, m_2), x_2(m_2)\rangle \\ &\overset{(a)}{=} nP_1(m_1, m_2) + nP_2 + 2n\rho(m_1, m_2)\sqrt{P_1(m_1, m_2)P_2} \leq nP\end{aligned} \tag{10}$$

for $\rho(m_1, m_2)$ confined in $[-1, 1]$ through defining

$$\sum_{i=1}^{n} x_{1,i}^2(m_1, m_2) = nP_1(m_1, m_2),\ \forall m_1, m_2. \tag{11}$$

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 given in Section 3.The feasible set of permissible inputs F_n is defined as

$$F_n := \{x^n : \|x^n\|^2 \leq nP\} \subset \mathbb{R}^n \tag{12}$$

3 $\kappa\beta$ bound–Achievability

The idea behind the $\kappa\beta$ bound introduced in [1] both for discrete and Gaussian channels is the optimality of the binary hypothesis testing problem and the Neyman-Pearson lemma. Imagine the binary hypothesis test between the two possible distributions, for instance P and Q, a random variable $W \in$ W can take on. A randomized test, denoted $\beta_a(P, Q)$ is set through the transformation $P_{Z|W} : \mathrm{W} \to \{0, 1\}$ where the test chooses the distribution Q and P for 0 and 1, respectively. As for the case of codewords with cost constraints where all of the codewords belong to the feasible set $F \subset A$, A being the input set, the corresponding measure of performance denoted $\kappa_\tau(F, Q_Y)$ for the hypothesis test is to choose between Q_Y and $\{P_{Y|X=x}\}_{x\in F}$ for $0 < \tau < \epsilon$. As stated in [1, Theorem 25], for a discrete P2P channel, there exists an (M, ϵ) code with codewords chosen from the feasible set F that satisfies for any $\epsilon \in (0, 1)$

$$M \geq \sup_{0<\tau<\epsilon} \sup_{Q_Y} \frac{\kappa_\tau(F, Q_Y)}{\sup_{x\in F} \beta_{1-\epsilon+\tau}(x, Q_Y)} \tag{13}$$

This bound is particularized to the Gaussian channel yielding

$$M^*(n, \epsilon, P) \geq \kappa_\tau^n / \beta_{1-\epsilon+\tau}^n \tag{14}$$

where M^* denotes the maximal code size and $\forall x \in F$, $\beta_{1-\epsilon+\tau}^n(x, Q_Y) = \beta_{1-\epsilon+\tau}^n$.

3.1 Two –User Degraded Gaussian BC

In this part, we adapt the $\kappa\beta$ bound given above by (14) to a two receiver degraded Gaussian BC. The degraded channel can be observed 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. Note that this bounding technique does not constrain the distribution of the channel input to any particular distribution.

The next theorem reminds the reader of the classical result from [4, 5] for the capacity region of the two-user Gaussian BC in the asymptotic regime.

Theorem 1 (Cover 1972, Bergmans 1974)**.** *The capacity region of the Gaussian BC is given as*

$$R_1 \leq C\left(\frac{\alpha P}{N_1}\right),\ \ R_2 \leq C\left(\frac{\bar{\alpha}P}{\alpha P + N_2}\right) \tag{15}$$

where α is a constant confined in $[0, 1]$ and $\bar{\alpha} = 1 - \alpha$. The Shannon capacity $C(.)$ is defined for SNR P as

$$C(P) = \frac{1}{2}\log(1 + P) \tag{16}$$

Using Theorem 1, we define the information densities of the Gaussian BC $i(X_2^n; Y_2^n)$ and $i(X^n; Y_1^n|X_2^n)$ using (7). Expectation of these information densities bring out the two mutual information functions that compose the capacity region of the Gaussian BC given in Theorem 1. Note that unlike the asymptotic case, the error event that corresponds to decoding the cloud center using the observation of the strong user cannot be ignored as shown in (4)-(5) with the corresponding error probability ϵ_{21}. Therefore we also take $i(X_2^n;Y_1^n)$ into account in our derivation.

Let us set $P_{Y_j^n} = \mathcal{N}(0, \sigma^2_{Y_j}\mathbf{I}_n)$ with $\sigma^2_{Y_j} = P + N_j$ for $j = 1,2$. The mutual information random variable $i(X_2^n;Y_2^n)$ under $P_{Y_2^n}$ is given as

$$G_{n_2} = \frac{n}{2}\log\sigma^2_{Y_2} - \frac{nP_2}{P_1+N_2}\frac{\log \mathrm{e}}{2} + \frac{1}{2}\log \mathrm{e}\sum_{i=1}^{n}\left[(1-\sigma^2_{Y_2})S_i^2 + 2\sqrt{P_2/(P_1+N_2)}\sigma_{Y_2}S_i\right] \tag{17}$$

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

$$H_{n_2} = \frac{n}{2}\log(\sigma^2_{Y_2}) + \frac{nP_2}{P_1+N_2}\frac{\log \mathrm{e}}{2\sigma^2_{Y_2}} + \frac{\log \mathrm{e}}{2\sigma^2_{Y_2}}\sum_{i=1}^{n}\left[\left(1-\sigma^2_{Y_2}\right)S_i^2 + 2\sqrt{P_2/(P_1+N_2)}S_i\right] \tag{18}$$

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

$$G_{n_1} = \frac{n}{2}\log\sigma^2_{Y_1'} - \frac{nP_1}{N_1}\frac{\log \mathrm{e}}{2} + \frac{1}{2}\log \mathrm{e}\sum_{i=1}^{n}\left[(1-\sigma^2_{Y_1'})S_i^2 + 2\sqrt{P_1/N_1}\sigma_{Y_1'}S_i\right] \tag{19}$$

for $\sigma^2_{Y_1'} = P_1 + N_1$ since P_2 is subtracted off the sum power P. As for the conditional distribution $P_{Y_1'^n|X_1^n}$ we get the information density as

$$H_{n_1} = \frac{n}{2}\log(\sigma^2_{Y_1'}) + \frac{P_1}{N_1}\frac{\log \mathrm{e}}{2\sigma^2_{Y_1'}} + \frac{\log \mathrm{e}}{2\sigma^2_{Y_1'}}\sum_{i=1}^{n}\left[\left(1-\sigma^2_{Y_1'}\right)S_i^2 + 2\sqrt{P_1/N_1}S_i\right] \tag{20}$$

Lastly, we denote $i(X_2^n;Y_1^n)$ by G_{n_3} and H_{n_3} under the distributions of $P_{Y_1^n}$ and $P_{Y_1^n|X_2^n}$, respectively. We adapt [1, Theorem 40] to the setting of superposition coding for the degraded Gaussian BC and define β functions as follows

$$\beta_{a,k} = \Pr[G_{n_k} \ge \gamma_k] \tag{21}$$

where $\Pr[H_{n_k} \ge \gamma_k] = a_k$ for $a_k = 1 - \epsilon_k$ as defined by (4) and with any positive γ_k and $k = 1,2,3$.

Theorem 2. *For any ϵ_k, $n \ge 1$, $\tau_k \in [0,1]$, $k = 1,2,3$, and the chosen P_{Y_j} for $j = 1,2$ with F_n as defined by (12), the maximal code sizes denoted M_j^* of the j^{th} user in a two receiver Gaussian degraded BC are bounded by*

$$M_1^* \ge \frac{\kappa_{\tau_1,1}(F_n, P_{Y_1^n})}{\beta_{1-\epsilon_1+\tau_1,1}(x, P_{Y_1^n})} \tag{22}$$

$$M_2^* \ge \max\left\{\frac{\kappa_{\tau_2,2}(F_n, P_{Y_2^n})}{\beta_{1-\epsilon_2+\tau_2,2}(x, P_{Y_2^n})}, \frac{\kappa_{\tau_3,3}(F_n, P_{Y_2^n})}{\beta_{1-\epsilon_3+\tau_3,3}(x, P_{Y_2^n})}\right\} \tag{23}$$

where $\kappa_{\tau_k,k}(F_n, P_{Y_j^n}) = P_{0,k}\left[\frac{p_{1,k}(r)}{p_{0,k}(r)} \ge \psi_k\right]$ with ψ_k satisfying $P_{1,k}\left[\frac{p_{1,k}(r)}{p_{0,k}(r)} \ge \psi_k\right] = \tau_k$. The probability distributions $p_{0,k}(r)$ and $p_{1,k}(r)$ are defined as

$$p_{0,k}(r) = \frac{1}{\Gamma(n/2)\omega_k^{n/2}}r^{n/2-1}\mathrm{e}^{r/\omega_k}$$

$$p_{1,k}(r) = \frac{1}{2}\mathrm{e}^{(r+\upsilon_k)/2}\left(\frac{r}{\upsilon_k}\right)^{n/4-1/2}I_{n/2-1}(\sqrt{\upsilon_k r})$$

with the modified Bessel function of the first kind $I_b(y) = (y/2)^b\sum_{j=0}^{\infty}\frac{(y^2/4)^l}{l!\Gamma(b+l+1)}$ and the following parameters for $k = 1$ $\omega_1 = 2(N_1+P_1)$ and $\upsilon_1 = \frac{nP_1}{N_1}$, for $k = 2$ $\omega_2 = 2(N_2+P)$ and $\upsilon_2 = \frac{nP_2}{P_1+N_2}$ whereas for $k = 3$ $\omega_3 = 2(N_1+P)$ with $\upsilon_3 = \frac{nP_2}{P_1+N_1}$.

Proof. Detailed proofs of the general case for a P2P-AWGN channel can be found in the original paper [1, 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 (21) is derived using the functions (17)-(20). As for evaluating $\kappa_{\tau_k,k}$'s the following definitions on $P_{0,k}$ and $P_{1,k}$ are set using superposition coding

$$P_{Y_1'^n} = P_{0,1} \sim \sum_{i=1}^{n}(P_1+N_1)S_i^2$$

$$P_{Y_2^n} = P_{0,2} \sim \sum_{i=1}^{n}(P+N_2)S_i^2$$

$$P_{Y_1^n} = P_{0,3} \sim \sum_{i=1}^{n}(P+N_1)S_i^2$$

$$P_{Y_1'^n|X_1^n} = P_{1,1} \sim \sum_{i=1}^{n}\left(\sqrt{P_1}+\sqrt{N_1}S_i\right)^2$$

$$P_{Y_2^n|X_2^n} = P_{1,2} \sim \sum_{i=1}^{n}\left(\sqrt{P_2}+\sqrt{(P_1+N_2)}S_i\right)^2$$

$$P_{Y_1^n|X_2^n} = P_{1,3} \sim \sum_{i=1}^{n}\left(\sqrt{P_2}+\sqrt{(P_1+N_1)}S_i\right)^2$$

for $S_i \sim N(0,1)$. □

4 A Converse Bound

In [6, Theorem 2], it was shown using the classical results of Sato [7] that in a physically degraded two–user BC the error probability of the system is upper bounded by the maximum of the two individual error probabilities per user and lower bounded by the error probability of the cooperative BC, which is equivalent to a P2P channel. Here, we have

$$2\max\{\epsilon_1, \epsilon_2\} \ge \epsilon \ge \epsilon_{\mathrm{c}} \tag{24}$$

where ϵ_1 and ϵ_2 and ϵ_{c} represent the first and second terms in (4) and the error probability of the cooperative BC, respectively.

[1, Theorem 41] presented a converse bound on the maximal code size of the P2P-AWGN channel. This bound is derived as a function of two random variables derived based on

$i(X^n; Y_c^n)$ which is denoted by H_n and G_n under $P(Y_c^n)$ and $P(Y_c^n|X^n)$, respectively. The corresponding functions for the cooperative channel are defined as

$$H_n = \frac{n}{2}\log_2(1+P_c) + \frac{1}{2}\frac{P_c \log e}{1+P_c}\sum_{i=1}^{n}\left(1 - S_i^2 + \frac{2S_i}{\sqrt{P_c}}\right), \tag{25}$$

$$G_n = \frac{n}{2}\log_2(1+P_c) - \frac{P_c \log e}{2}\sum_{i=1}^{n}\left(1 + S_i^2 - 2\sqrt{1+\frac{S_i}{P_c}}\right) \tag{26}$$

with $S_i \sim N(0,1)$ and $P_c = P/N_c$ where $N_c = \min\{N_1, N_2\}$. The maximal code size of the AWGN channel for any n and the average error probability ϵ is upper bounded by $(\Pr[G_n \geq \gamma_n])^{-1}$ given that $\Pr[H_n \geq \gamma_n] = 1 - \epsilon$ is satisfied [1, Theorem 41]. We also define the functions $H_{n_{c,j}}$ and $G_{n_{c,j}}$, respectively through (25) and (26) where P_c is replaced by P/N_j for $j = 1, 2$.

Theorem 3. *For a Gaussian broadcast channel with two receivers with a power constraint as given by (8), the following inequalities on the maximal code sizes of user 1 and 2 denoted M_1^* and M_2^*, respectively hold for any n*

$$M_1^* \leq \frac{1}{\Pr[G_{n_{c,1}} \geq \gamma_{c,1}]} \tag{27}$$

$$M_2^* \leq \frac{1}{\Pr[G_{n_{c,2}} \geq \gamma_{c,2}]} \tag{28}$$

$$M_1^* M_2^* \leq \frac{1}{\Pr[G_n \geq \gamma_c]} \tag{29}$$

where the positive thresholds in each inequality $\gamma_{c,1}$, $\gamma_{c,2}$ and γ_c satisfy $\Pr[H_{n_{c,1}} \geq \gamma_{c,1}] = 1 - \epsilon_1$, $\Pr[H_{n_{c,2}} \geq \gamma_{c,2}] = 1 - \epsilon_2$, $\Pr[H_n \geq \gamma_c] = 1 - \epsilon_c$ respectively for $2\max\{\epsilon_1, \epsilon_2\} \geq \epsilon_c$.

Proof. First two inequalities (27) and (28) corresponding respectively to user 1 and 2 are a direct application of the AWGN bound of [1, Theorem 41] using (25) and (26). For the third inequality (29) in order to apply the AWGN bound, we use the channel where the two receivers are allowed to cooperate with an error probability that cannot exceed the one of the broadcast channel as shown in [7, 6]. This last inequality corresponds to the cooperative channel with the product code sizes of the two users and the corresponding condition on γ_c depends on the error probability satisfying (24). □

Figure 1 shows the numerical comparison of the converse bound from Theorem 3 labeled as the outer bound with the asymptotic capacity region given by Theorem 1 and the superposition coding applied to the $\kappa\beta$ bound where κ is derived through [1, Lemma 43]. Sato's bounding technique yields a very loose bound since the converse is partially outside of the infinite blocklength capacity. The dotted line represents the time-sharing between two single-user achievable rates $R_j = \frac{1}{n}\log M_j$ where M_j is given by (14) as functions of P_1 and P_2, respectively for user 1 and 2. A more fair comparison can be made by limiting time-sharing to within 1000 dimensions where $n_1 + n_2 = 1000$ and R_j is (14) as a function of n_j. In the case of an asymmetric channel, the advantage of superposition coding over orthogonal multiplexing with short blocklength is significant as in the asymptotic regime.

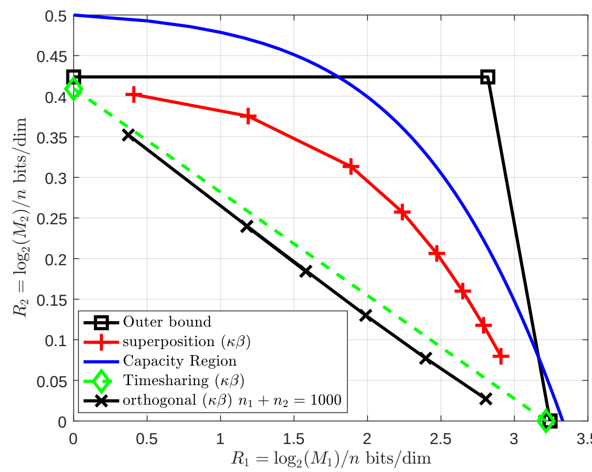

Figure 1: Numerical comparison of the obtained bounds for $P = 100$, $\epsilon = 10^{-3}$, $N_1 = 1$ and $N_2 = 100$.

5 Acknowledgement

This work has been partly supported by the INSA Lyon - SPIE ICS chair on the Internet of Things, by Orange Labs under Grant no F05151 and by the French National Research Agency project ARBURST.

References

[1] 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.

[2] 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.

[3] 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.

[4] T. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. 18, pp. 2–14, January 1972.

[5] P. Bergmans, "A simple converse for broadcast channels with additive white gaussian noise," *IEEE Transactions on Information Theory*, pp. 279–280, March 1974.

[6] L. Weng, S. Pradhan, and A. Anastasopoulos, "Error exponent regions for Gaussian broadcast and multiple-access channels," *IEEE Transactions on Information Theory*, vol. 54, pp. 2919–2942, July 2008.

[7] H. Sato, "An outer bound to the capacity region of broadcast channels," *IEEE Transactions on Information Theory*, vol. 24, pp. 374–377, May 1978.