Linear Superposition Coding for the Gaussian Broadcast Channel with Quantized Feedback

Stefan Farthofer and Gerald Matz
Institute of Telecommunications, Vienna University of Technology, Vienna (Austria)
Email: stefan.farthofer@nt.tuwien.ac.at, gerald.matz@nt.tuwien.ac.at

***Abstract*—We propose a linear transceiver scheme for the symmetric two-user broadcast channel with additive Gaussian noise and quantized feedback. The quantized feedback link is modeled as an information bottleneck subject to a rate constraint. We introduce a superposition scheme that splits the transmit power between an Ozarow-like linear-feedback code and a conventional code that ignores the feedback. The recently established MAC-BC duality with linear-feedback coding is key as it allows us to extend our previously proposed linear-feedback MAC scheme to the BC. We study the achievable sum rate as a function of the feedback quantization rate and we show that sum rate maximization leads to a difference of convex functions problem that we solve via the convex-concave procedure.**

I. Introduction

Noiseless feedback is known to enhance the capacity of the multiple access channel (MAC) [1] and of the broadcast channel (BC) [2]. For the single-user additive white Gaussian noise channel, Schalkwijk and Kailath proposed a remarkably simple linear feedback scheme that achieves capacity and yields an error probability that decreases doubly exponentially in the block length [3], [4]. Ozarow extended the Schalkwijk-Kailath scheme to the two-user Gaussian MAC with perfect feedback [5] and proved it to achieve the feedback capacity. A similar scheme for the Gaussian BC with perfect feedback was devised by Ozarow and Leung in [6]. However, for the BC with feedback the capacity region is still unknown and the Ozarow-Leung scheme is outperformed by control-theoretic schemes that were recently shown by Amor et al. to achieve the linear-feedback capacity for the BC with perfect feedback [7]. Later work focused on the more realistic assumption of noisy feedback. For the MAC, Gastpar extended Ozarow's scheme and applied it to noisy feedback [8], [9] and it was shown by Lapidoth and Wigger that even non-perfect feedback is always beneficial [10]. For the BC, in [11], the noise in the feedback link is tackled via concatenated coding. Furthermore, [12] studied the Gaussian BC with one-sided noisy feedback. Recently, Wu and Wigger studied a similar scenario [13] as here, but in contrast they allow the receivers to code over the feedback link. Unfortunately, many of the schemes proposed do not have the simplicity of the original Ozarow scheme and the achievable rate regions are very hard to analyze. We previously proposed a simple, Ozarow-like superposition coding scheme for the Gaussian MAC with quantized feedback [14]. We extend this work to the Gaussian BC with quantized feedback using recent MAC-BC duality results [7].

Funding by WWTF Grant ICT12-054.

Specifically, our contributions are as follows. We propose a superposition of feedback-based encoding and conventional (non-feedback) coding for the two-user Gaussian BC with quantized feedback. We model the feedback quantization as channel output compression via the information bottleneck principle [15]–[18], which allows the receivers to use the quantization noise as side-information. We assess the sum rate achievable with our superposition scheme and quantify the impact of the power (equivalently, rate) splitting between the two coding schemes. Finally, we show that maximizing the sum rate by optimizing the the power (rate) allocation is a difference of convex functions (DC) problem [19] and we solve this problem numerically via the convex-concave procedure (CCP) [20].

Notation: We use boldface letters for column vectors and upright sans-serif letters for random variables. Expectation is denoted by $\mathrm{E}\{\cdot\}$ and a Gaussian distribution with mean μ and variance σ^2 is denoted by $\mathcal{N}(\mu,\sigma^2)$. We use the notation $I(\cdot;\cdot)$ for the mutual information [21]. The capacity of a Gaussian channel with signal-to-noise ratio (SNR) γ is denoted by $\mathsf{C}(\gamma) \triangleq \frac{1}{2}\log(1+\gamma)$. All logarithms are to base 2 such that the unit of all information-theoretic quantities is bit.

II. Background

A. Symmetric BC Model

We study the two-user symmetric Gaussian BC with (quantized) feedback (see Figure 1). Here, the transmitter sends the length-n signal $x[k] = x_1[k] + x_2[k]$, where $x_i[k]$, $i = 1,2$, are independent Gaussian user signals. We impose the average power constraints (here, expectation is with respect to the messages and the channel noise) $\frac{1}{n}\sum_{k=1}^{n}\mathrm{E}\{x^2[k]\} \leq P$, such that the average power of $x[k]$ is bounded by P. We represent length-n signals via corresponding length-n vectors, i.e.,

$$\mathbf{x} = (x[1]\dots x[n])^T = \mathbf{x}_1 + \mathbf{x}_2.$$

The channel introduces i.i.d. additive Gaussian noise $\mathbf{z}_1 \sim \mathcal{N}(\mathbf{0},\sigma^2\boldsymbol{I})$ and $\mathbf{z}_2 \sim \mathcal{N}(\mathbf{0},\sigma^2\boldsymbol{I})$ with identical variance at both receivers such that the receive signals read

$$\mathbf{y}_1 = \mathbf{x} + \mathbf{z}_1, \qquad \mathbf{y}_2 = \mathbf{x} + \mathbf{z}_2. \tag{1}$$

The receivers feed back the signals $\mathbf{w}_i$, $i = 1,2$, to the transmitter, which uses the feedback signals in the construction of the transmit signal $\mathbf{x}$.

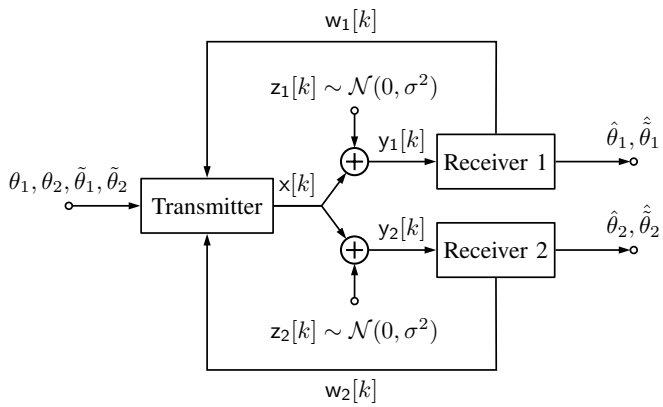

Figure 1: Two-user Gaussian BC with feedback.

B. The Ozarow Scheme

Ozarow characterized the full capacity region [5] of the two-user Gaussian MAC with perfect channel output feedback by extending the Schalkwijk-Kailath scheme [3]. The key idea in this scheme is the iterative refinement of the message estimates. In the first two time slots, both transmitters alternatingly send their raw messages. The remaining time slots are used to transmit updates of the message estimates at the receiver. For infinite block length, this linear scheme is capacity-achieving. For the symmetric case with identical transmit powers $P/2$ for both users, Ozarow obtained the sum capacity as (recall $\gamma = P/\sigma^2$)

$$C_{\text{MAC}} = \mathsf{C}\left(\gamma(1+\rho^*)\right), \tag{2}$$

where the correlation coefficient ρ^* is the solution of

$$\frac{2\rho^*}{(\rho^*-1)^2(\rho^*+1)} = \gamma. \tag{3}$$

C. Linear-Feedback MAC-BC Duality

The duality between the MAC and the BC without feedback is well understood [22]. For the MAC and BC with feedback, duality cannot be obtained in a straightforward manner. In fact, the feedback-capacity of the BC remains unknown to date. The best known achievable rate region turned out to coincide with the linear-feedback capacity region [7]. In [7] it was shown that the linear-feedback capacity region of the dual MAC (which is the capacity region) equals the linear-feedback capacity region of the BC,

$$C_{\text{BC}} = C_{\text{MAC}} = \mathsf{C}\left(\gamma(1+\rho^*)\right). \tag{4}$$

This duality result enables the construction of an Ozarow-like scheme for the BC with feedback that achieves the linear-feedback capacity region.

D. Gaussian Information Bottleneck

The feedback in our scheme is obtained by quantizing the outputs of the Gaussian BC. We thus briefly review Gaussian channel output compression [17] based on the Gaussian information bottleneck (GIB) [15], [16].

Consider the Markov chain $\mathbf{x} - \mathbf{y} - \mathbf{w}$ where $\mathbf{x}$ is the channel input, $\mathbf{y}$ is the channel output, and $\mathbf{w}$ is the feedback signal obtained by compressing $\mathbf{y}$. We assume that $\mathbf{x}$ and $\mathbf{y}$ are jointly Gaussian random vectors with zero mean and full-rank covariance matrix. The trade-off between compression rate and relevant information is captured by the rate-information function $I\colon \mathbb{R}_+ \to [0, I(\mathbf{x};\mathbf{y})]$, which is defined as

$$I(R) \triangleq \max_{p(\mathbf{w}|\mathbf{y})} I(\mathbf{x};\mathbf{w}) \quad \text{subject to } I(\mathbf{y};\mathbf{w}) \leq R, \tag{5}$$

where $p(\mathbf{w}|\mathbf{y})$ denotes the conditional distribution of $\mathbf{w}$ given $\mathbf{y}$. The channel input $\mathbf{x}$ is the relevance variable, $I(\mathbf{x};\mathbf{w})$ is the relevant information in $\mathbf{w}$ about $\mathbf{x}$, and $I(\mathbf{y};\mathbf{w})$ is the compression rate. The rate-information function thus quantifies the maximum amount of relevant information that can be preserved when the compression rate is at most R. The definition (5) is similar to rate-distortion theory, only that the minimization of distortion is replaced with a maximization of the relevant information.

In [17] it was shown that optimal compression of $\mathbf{y}$ in the sense of the rate-information function yields an equivalent channel $p(\mathbf{w}|\mathbf{x})$ which is also Gaussian. Therefore, a Gaussian input distribution $p(\mathbf{x})$ satisfying the power constraint of the channel $p(\mathbf{y}|\mathbf{x})$ with equality is capacity-achieving also for the channel $p(\mathbf{w}|\mathbf{x})$.

Each receiver of our BC can be rewritten as a separate multiple-input, single-output model in [18, Section IV-C], with the point-to-point channel relevant for quantization being the all-ones vector. The resulting rate-information function at SNR $\gamma = P/\sigma^2$ can be shown to equal [17, Theorem 5]

$$I(R) = \mathsf{C}(\gamma) - \frac{1}{2}\log\left(1 + 2^{-2R}\gamma\right). \tag{6}$$

Thus, the rate-information function approaches capacity as the compression rate R goes to infinity. Equivalently, we can write $I(R) = \mathsf{C}(\gamma')$, where

$$\gamma' = \gamma\frac{1-2^{-2R}}{1+2^{-2R}\gamma} \leq \gamma \tag{7}$$

is the equivalent SNR of the channel $p(\mathbf{w}|\mathbf{x})$. Rate-information optimal channel output compression can thus be modeled via additive Gaussian quantization noise with variance

$$\sigma_q^2 = \sigma^2\frac{1+\gamma}{2^{2R}-1}. \tag{8}$$

III. Proposed Scheme

A. Linear Superposition Coding

The transmitter communicates independent messages θ_1, $\tilde{\theta}_1$ and θ_2, $\tilde{\theta}_2$ to the two users (receivers). Here and in what follows, superscript tilde denotes quantities based on exploitation of the channel output feedback. The messages are uniformly drawn from finite sets with cardinalities $\mathcal{M}_1 = 2^{nR_1}$, $\widetilde{\mathcal{M}}_1 = 2^{n\tilde{R}_1}$, $\mathcal{M}_2 = 2^{nR_2}$, $\widetilde{\mathcal{M}}_2 = 2^{n\tilde{R}_2}$ and mapped to the transmit signal components $\mathbf{x}_i$, $i = 1, 2$, according to the superposition

$$x_i[k] = \varphi_{i,k}(\theta_i) + \tilde{\varphi}_{i,k}\big(\tilde{\theta}_i, \mathbf{w}_i^{(k-1)}\big), \tag{9}$$

$k = 1, \ldots, n$. Here, $\varphi_{i,k} : \mathcal{M}_i \to \mathbb{R}$ denotes a conventional encoder that ignores the feedback signal. Furthermore, $\mathbf{w}_i^{(k-1)} = (w_i[1] \ldots w_i[k-1])^T$ denotes the past quantized

feedback from receiver i and $\tilde{\varphi}_{i,k}: \widetilde{\mathcal{M}}_i \times \mathbb{R}^{k-1} \to \mathbb{R}$ is the feedback-based encoder that has causal access to the quantized feedback and works as in the original Ozarow scheme.

The encoder splits the total power P by allocating a fraction $\alpha \in [0,1]$ of that power to the feedback-based codewords and the rest to the non-feedback part, i.e.,

$$\frac{1}{n}\sum_{i=1}^{2}\sum_{k=1}^{n}\mathrm{E}\big\{\varphi_{i,k}^2(\theta_i)\big\} = (1-\alpha)P, \tag{10}$$

$$\frac{1}{n}\sum_{i=1}^{2}\sum_{k=1}^{n}\mathrm{E}\Big\{\tilde{\varphi}_{i,k}^2\big(\tilde{\theta}_i, \mathbf{w}_i^{(k-1)}\big)\Big\} = \alpha P. \tag{11}$$

B. Feedback Quantization

In our model, the channel outputs of each receiver are quantized before being fed back to the transmitter. More specifically, the receivers perform successive cancellation, i.e., subtract the estimates $\varphi_{i,k}(\hat{\theta}_i)$ of the conventional codewords from the received signal, and quantize the residual,

$$\mathsf{w}_i[k] = \mathcal{Q}\big(\mathsf{y}_i[k] - \varphi_{1,k}(\hat{\theta}_1) - \varphi_{2,k}(\hat{\theta}_2)\big), \quad i = 1, 2.$$

This is possible since we consider a fully symmetric BC channel where both users are able to decode both non-feedback messages. The quantization $\mathcal{Q}(\cdot)$ is modeled as an information bottleneck, i.e., the mutual information between compressed received signal and transmit signal is maximized under a rate constraint (cf. Section II-D).

For the BC with noisy feedback the extra noise on the feedback channel is detrimental since it is not known by any node. In our model the transmitter also receives degraded versions of the channel outputs, but the receivers know the quantization errors of their feedback signals (having themselves performed the quantization). With this observation, our scheme can be reduced to an equivalent BC with perfect feedback in which the quantization error represents additional channel noise. The knowledge of this extra noise is exploited at the receivers.

C. Achievable Sum Rate

Our proposed scheme is a superposition of a conventional encoder and a feedback-based encoder (cf. (9)). Pure conventional encoding and pure feedback encoding are special cases obtained with $\alpha = 0$ and $\alpha = 1$, respectively. Thus, we can maximize the sum rate by finding the optimum power (equivalently, rate) splitting between the two encodings.

The conventionally encoded signals are cancelled at the receivers before quantization such that the quantization only captures the feedback encoding in the forward path. This is possible since each receiver has the quantization noise as side-information. The sum rate of the conventional coding scheme is thus given by the classical Gaussian BC capacity [21] with signal power $(1-\alpha)P$ and noise power $\sigma^2 + \alpha P$ (since the feedback-based codewords act as additional interference). The effective SNR thus equals $\frac{(1-\alpha)P}{\sigma^2+\alpha P} = \frac{(1-\alpha)\gamma}{1+\alpha\gamma}$ (recall $\gamma = P/\sigma^2$) and the achievable sum rate is

$$R_1 + R_2 \le \bar{C}(\alpha) \triangleq \mathsf{C}\left(\frac{(1-\alpha)\gamma}{1+\alpha\gamma}\right). \tag{12}$$

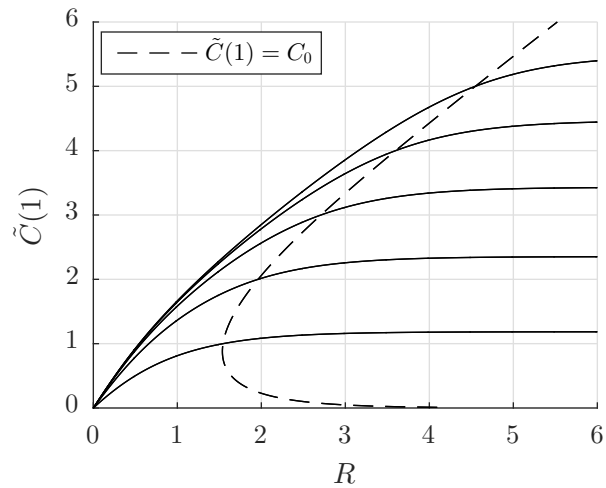

Figure 2: Sum rate $\tilde{C}(1)$ for quantized feedback-based code versus quantization rate R for several SNRs corresponding to no-feedback sum capacities $C_0 = 1 \ldots 5$.

The sum rate achievable with the feedback-based code follows by exploiting the duality between the linear-feedback MAC and the linear-feedback BC (cf. Section II-C) and using the Ozarow scheme for perfect feedback (see Section II-B). The effective SNR in the Ozarow scheme is reduced due to feedback quantization. Specifically, the feedback compression can be seen as additional i.i.d. quantization noise σ_q^2 (cf. Section II-D, specifically (8)). This decreases the SNR for the linear-feedback code according to

$$\tilde{\gamma} = \frac{\alpha P}{\sigma^2 + \sigma_q^2} = \alpha\gamma\frac{1}{1+\sigma_q^2/\sigma^2} = \alpha\gamma\frac{2^{2R}-1}{2^{2R}+\alpha\gamma}. \tag{13}$$

The achievable sum rate for the feedback code is obtained from (4) with γ replaced by $\tilde{\gamma}$, i.e.,

$$\tilde{R}_1 + \tilde{R}_2 \le \tilde{C}(\alpha) \triangleq \mathsf{C}\left(\alpha\gamma\frac{2^{2R}-1}{2^{2R}+\alpha\gamma}(1+\rho^*)\right). \tag{14}$$

Here, ρ^* is determined by (3) with $\tilde{\gamma}$ in place of γ. While this rate is smaller than that achievable with perfect feedback, it can still surpass the BC capacity without feedback for large enough quantization rate R (see Fig. 2).

The overall sum rate for the Gaussian BC with quantized feedback and a superposition of conventional and linear-feedback encoding is finally obtained by optimizing the power allocation (rate splitting) parameter α for the conventional encoding and the feedback-based encoding, i.e., $R_1 + R_2 + \tilde{R}_1 + \tilde{R}_2 \le C_{\mathrm{S}}$ with

$$C_{\mathrm{S}} = \max_{0\le\alpha\le1} \bar{C}(\alpha) + \tilde{C}(\alpha). \tag{15}$$

IV. DC Programming

A. Reformulation of the Sum Rate Maximization

To maximize the sum rate we have to find the optimal power allocation parameter α by solving (15). It can be shown that $\tilde{C}(\alpha)$ is a concave function and $\bar{C}(\alpha)$ is a convex function in

Algorithm 1 Power allocation via CCP

Require: Initial feasible point α_0

1: $k := 0$
2: **while** stopping criterion not satisfied **do**
3: Form $\hat{C}_k(\alpha) = \bar{C}(\alpha_k) + \bar{C}'(\alpha_k)(\alpha - \alpha_k)$
4: Determine α_{k+1} by solving the convex problem

$$\begin{aligned} \text{minimize} \quad & -\tilde{C}(\alpha) - \hat{C}_k(\alpha) \\ \text{subject to} \quad & \alpha - 1 \leq 0, \\ & -\alpha \leq 0 \end{aligned}$$

5: $k := k + 1$
6: **end while**
7: **return** α_k

α. Indeed, by direct calculation, we obtain the second-order derivative of $\bar{C}(\alpha)$ as

$$\frac{\mathrm{d}^2}{\mathrm{d}\alpha^2}\bar{C}(\alpha) = \frac{\log \mathrm{e}}{2}\frac{\gamma^2}{(1+\alpha\gamma)^2} \geq 0, \tag{16}$$

which proves convexity. The concavity of $\tilde{C}(\alpha)$ can be proved similarly by showing that its second-order derivative is non-positive. Since ρ^* is itself the solution of a cubic equation, the derivation is rather lengthy and thus omitted due to space limitations. It follows that the sum capacity C_S is the maximum of the sum of a convex and a concave function, or equivalently, of the difference of two convex functions. The sum rate maximization problem can be solved by difference of convex functions (DC) programming [20], specifically by the convex-concave procedure (CCP) introduced by Yuille and Rangarajan [19]. The maximization problem (15) can be reformulated in standard form [20], i.e., involving differences of convex functions in the objective and in the constraints:

$$\begin{aligned} \underset{\alpha}{\text{minimize}} \quad & -\tilde{C}(\alpha) - \bar{C}(\alpha) \\ \text{subject to} \quad & \alpha - 1 \leq 0, \\ & -\alpha \leq 0. \end{aligned} \tag{17}$$

B. Numerical Solution

The basic idea of the iterative CCP algorithm [19] is to find a point where the gradient of the convex part in the next iteration equals the negative gradient of the concave part of the previous iteration. Intuitively, consider the boundary $\big(\tilde{C}(\alpha), \bar{C}(\alpha)\big)$ of the power splitting rate region; maximizing sum rate is then equivalent to finding the point $\big(\tilde{C}(\alpha), \bar{C}(\alpha)\big)$ whose tangent has slope -1,

$$\frac{\mathrm{d}\bar{C}(\alpha)}{\mathrm{d}\tilde{C}(\alpha)} = \frac{\mathrm{d}\bar{C}(\alpha)/\mathrm{d}\alpha}{\mathrm{d}\tilde{C}(\alpha)/\mathrm{d}\alpha} = -1, \tag{18}$$

Using a superscript prime to denote the first-order derivative, CCP thus aims at

$$\tilde{C}'(\alpha_{k+1}) = -\bar{C}'(\alpha_k), \tag{19}$$

which itself amounts to a convex optimization problem. The solution to this auxiliary problem decreases monotonically with increasing k and thus converges to a minimum (or to saddle point). Lipp and Boyd [20] give a basic CCP algorithm

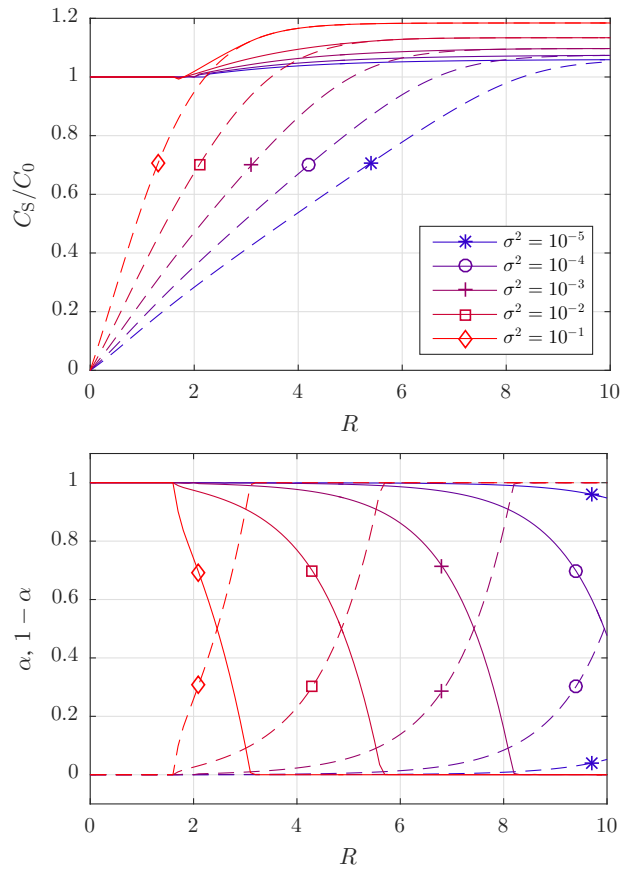

Figure 3: Top: Normalized achievable rates with superposition coding (solid lines) and with pure feedback coding (dashed lines). Bottom: Power splitting (solid lines for $1-\alpha$, dashed lines for α) for various channel noise levels and fixed total transmit power $P = 1$.

which requires an initial feasible point α_0, which in our case can be any point in the interval $[0, 1]$. Following [20], the CCP approach leads to Algorithm 1 for power allocation. The derivative required in step 3 can explicitly be computed as

$$\bar{C}'(\alpha) = -\frac{1}{2}\frac{\gamma}{1+\alpha\gamma}. \tag{20}$$

Fig. 3 shows the rates (top) and power allocations (bottom) obtained by solving (17) versus the feedback quantization rate R. The linear-feedback capacity of the proposed superposition coding scheme is normalized by the no-feedback BC capacity C_0. Clearly, when the feedback quantization rate R is too small the feedback is not beneficial at all and the whole transmit power is allocated to the conventional encoding (see bottom part of Fig. 3). Above a certain threshold for R, true superposition is optimal until eventually pure feedback coding becomes optimal for very large quantization rates R (almost perfect feedback). Fig. 4 shows the achievable power splitting rate region $(\bar{C}(\alpha), \tilde{C}(\alpha))$ and the maximum sum rate C_S.

C. Optimality of Superposition

The bottom plot in Fig. 4 shows the achievable power splitting rate region $\big(\tilde{C}(\alpha), \bar{C}(\alpha)\big)$ for fixed quantization rate

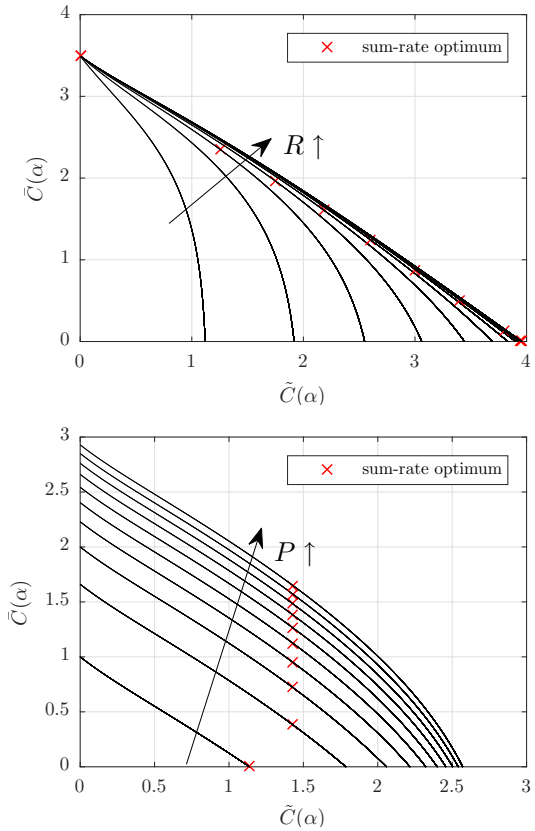

Figure 4: Achievable rate region $(\tilde{C}(\alpha), \bar{C}(\alpha))$ for fixed $C_0 = 3$ and $R = 0.2C_0 \ldots 2.4C_0$ (top) and for fixed $R = 2$ and $P = 0.1P_R \ldots 2P_R$ where P_R is such that $C_0(P_R) = R$ (bottom).

R and variable transmit power P. We see that below a non-trivial power threshold it is optimum to allocate all transmit power to the feedback-coding scheme and above that threshold superposition is optimum. Note that the power allocated to the feedback scheme stays constant while only the power allocated to the non-feedback scheme increases with increasing total transmit power. This can be explained by studying the behavior of the iteration (19), which characterizes a fixed point for the optimal α. In the optimal point the negative gradient of $\tilde{C}(\alpha)$ equals the gradient of $\bar{C}(\alpha)$ (cf. (18)). While the optimal α depends on the total power P, it can be shown that the amount of power αP in the feedback-based code remains constant above a certain power threshold.

V. Conclusions

We used the information bottleneck principle to model the quantization of the feedback in a two-user Gaussian BC. We showed that due to the rate limitation on the feedback links it is useful to superimpose a conventional (non-feedback) scheme and a feedback coding scheme. Using recent linear-feedback MAC-BC duality results we showed that a modified version of the Ozarow scheme and a superimposed conventional encoding scheme, which are separated at the receivers by successive cancellation, generally yields a larger sum capacity than any of the two constituent schemes alone. We demonstrated that the problem of finding the right balance between these two schemes can be restated as a difference of convex functions (DC) program that can be efficiently solved numerically via the concave-convex procedure (CCP) algorithm. We found that true superposition is optimum in regimes with high transmit power or low feedback quantization rate.

References

[1] T. Gaarder and J. K. Wolf, "The Capacity Region of a Multiple-Access Discrete Memoryless Channel Can Increase with Feedback," *IEEE Trans. Inf. Theory*, pp. 100–102, Jan. 1975.

[2] G. Dueck, "Partial Feedback for Two-Way and Broadcast Channels," *Information and Control*, vol. 46, pp. 1–15, 1980.

[3] J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback–I: No bandwidth constraint," *IEEE Trans. Inf. Theory*, vol. 12, no. 2, pp. 172–182, April 1966.

[4] J. Schalkwijk, "A coding scheme for additive noise channels with feedback–II: Band-limited signals," *IEEE Trans. Inf. Theory*, vol. 12, no. 2, pp. 183–189, April 1966.

[5] L. H. Ozarow, "The Capacity of the White Gaussian Multiple Access Channel with Feedback," *IEEE Trans. Inf. Theory*, vol. 30, no. 4, pp. 623–629, July 1984.

[6] L. H. Ozarow and S. K. Leung-Yan-Cheong, "An Achievable Region and Outer Bound for the Gaussian Broadcast Channel with Feedback," *IEEE Trans. Inf. Theory*, vol. 30, no. 4, pp. 667–671, July 1984.

[7] S. B. Amor, Y. Steinberg, and M. Wigger, "MIMO MAC-BC Duality With Linear-Feedback Coding Schemes," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 5976–5998, Nov. 2015.

[8] M. Gastpar, "On Noisy Feedback in Gaussian Networks," in *Proc. 43rd Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Sept. 2005.

[9] M. Gastpar and G. Kramer, "On Cooperation Via Noisy Feedback," in *Proc. Int. Zurich Seminar on Commun.*, Zurich, Switzerland, Feb. 2006, pp. 146–149.

[10] A. Lapidoth and M. Wigger, "On the AWGN MAC With Imperfect Feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5432–5476, Nov. 2010.

[11] Z. Ahmad, Z. Chance, D. J. Love, and C.-C. Wang, "Concatenated Coding Using Linear Schemes for Gaussian Broadcast Channel With Noisy Channel Output Feedback," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4576–4590, Nov. 2015.

[12] S. R. B. Pillai and V. M. Prabhakran, "On the Noisy Feedback Capacity of Gaussian Broadcast Channels," in *Proc. IEEE Inf. Theory Workshop (ITW 2015)*, Jerusalem, Israel, April 2015.

[13] Y. Wu and M. Wigger, "Coding Schemes With Rate-Limited Feedback That Improve Over the No Feedback Capacity for a Large Class of Broadcast Channels," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 2009–2033, April 2016.

[14] S. Farthofer and G. Matz, "Linear Superposition Coding for the Gaussian MAC with Quantized Feedback," in *Proc. 53rd Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Sept. 2015, pp. 1374–1379.

[15] N. Tishby, F. Pereira, and W. Bialek, "The Information Bottleneck Method," in *Proc. 37th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Sept. 1999, pp. 368–377.

[16] G. Chechik, A. Globerson, N. Tishby, and Y. Weiss, "Information Bottleneck for Gaussian Variables," *Journal of Machine Learning Research*, vol. 6, pp. 165–188, Jan. 2005.

[17] A. Winkelbauer and G. Matz, "Rate-information-optimal Gaussian channel output compression," in *Proc. 48th Annu. Conf. on Inf. Sciences and Systems (CISS 2014)*, Princeton, NJ, March 2014.

[18] A. Winkelbauer, S. Farthofer, and G. Matz, "The rate-information trade-off for Gaussian vector channels," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT 2014)*, Honolulu, HI, June 2014, pp. 2849–2853.

[19] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Computation*, vol. 15, no. 4, pp. 915–936, 2003.

[20] T. Lipp and S. Boyd, "Variations and extensions of the convex-concave procedure," http://stanford.edu/%7eboyd/papers/cvx%5fccv.html, 2014.

[21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[22] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, Achievable Rates, and Sum-Rate Capacity of Gaussian MIMO Broadcast Channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.