Linear Superposition Coding for the Asymmetric Gaussian MAC 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 asymmetric two-user multiple access 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. 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]. 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 [2], [3]. Ozarow extended the Schalkwijk-Kailath scheme to the two-user Gaussian MAC with perfect feedback [4] and proved it to achieve the feedback capacity. Since the assumption of perfect feedback is unrealistic, later work focused on noisy feedback. Gastpar extended Ozarow's scheme and applied it to noisy feedback [5], [6] and it was shown by Lapidoth and Wigger that even non-perfect feedback is always beneficial [7]. 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 symmetric Gaussian MAC with quantized feedback [8]. Here, symmetric means identical channel gains. We now extend this work to the asymmetric Gaussian MAC with arbitrary channel gains and quantized feedback.

Notation: We use boldface letters for column vectors and upright sans-serif letters for random variables. The identity matrix is denoted by $\boldsymbol{I}$. Expectation is written as $\mathrm{E}\{\cdot\}$ 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 [9]. All logarithms are to base 2.

Funding by WWTF Grant ICT12-054.

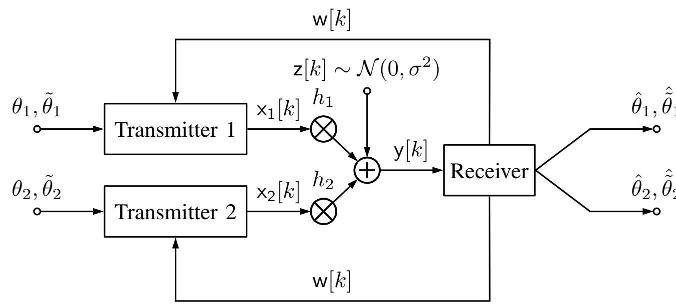

Figure 1: Two-user Gaussian MAC with common feedback link.

II. Contributions

We propose a superposition of feedback-based encoding and conventional (non-feedback) coding for the two-user Gaussian MAC with arbitrary channel gains and quantized feedback. We model the feedback quantization as channel output compression via the information bottleneck principle [10]–[13], which allows the receiver 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 constituent codes. Finally, we show that maximizing the sum rate by optimizing the power (rate) allocation is a difference of convex functions (DC) problem [14] and we solve this problem numerically via the convex-concave procedure (CCP) [15].

A. MAC Model and Proposed Coding Scheme

We study the two-user asymmetric Gaussian MAC with (quantized) feedback (see Fig. 1). Here, the transmitters send the independent length-n Gaussian user signals $x_i[k]$, $i = 1, 2$, and the receiver observes the signal

$$y[k] = h_1 x_1[k] + h_2 x_2[k] + z[k]. \tag{1}$$

The channel gains h_1 and h_2 are assumed to be known by both transmitters and receiver and the channel introduces i.i.d. additive Gaussian noise $\mathsf{z} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{I})$. We impose the average transmit power constraint (here, expectation is with respect to the messages and the channel noise)

$$\frac{1}{n} \sum_{k=1}^{n} \mathrm{E}\big\{x_1^2[k] + x_2^2[k]\big\} \leq P. \tag{2}$$

The transmitters communicate independent messages θ_1, $\tilde{\theta}_1$ and θ_2, $\tilde{\theta}_2$ to a single receiver. 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 $\mathcal{M}_1$, $\widetilde{\mathcal{M}}_1$, $\mathcal{M}_2$, $\widetilde{\mathcal{M}}_2$, and mapped to the transmit signals according to the superposition

$$x_i[k] = \varphi_{i,k}(\theta_i) + \tilde{\varphi}_{i,k}\big(\tilde{\theta}_i, \mathbf{w}^{(k-1)}\big), \tag{3}$$

$k = 1, \ldots, n$. Here, $\varphi_{i,k} : \mathcal{M}_i \to \mathbb{R}$ denotes a conventional encoder with power constraint

$$\frac{1}{n}\sum_{k=1}^{n} \mathrm{E}\big\{\varphi_{i,k}^2(\theta_i)\big\} \leq P_i, \tag{4}$$

with expectation over all possible transmit messages θ_i. The conventional encoder completely ignores the feedback signal. Furthermore, $\mathbf{w}^{(k-1)} = (w[1] \ldots w[k-1])^\top$ denotes the past quantized feedback and $\tilde{\varphi}_{i,k} : \widetilde{\mathcal{M}}_i \times \mathbb{R}^{k-1} \to \mathbb{R}$ is the feedback-based encoder with power constraint

$$\frac{1}{n}\sum_{k=1}^{n} \mathrm{E}\Big\{\tilde{\varphi}_{i,k}^2\big(\tilde{\theta}_i, \mathbf{w}^{(k-1)}\big)\Big\} \leq \tilde{P}_i, \tag{5}$$

with expectation over all possible transmit messages $\tilde{\theta}_i$ and channel realizations. This encoder has causal access to the quantized feedback and works as in the original Ozarow scheme (see Section III-B).

B. Feedback Quantization

In our model, the channel output at the receiver is quantized before being fed back to the transmitters. More specifically, the receiver performs successive cancellation, i.e., it subtracts the estimates $\varphi_{i,k}(\hat{\theta}_i)$ of the conventional codewords from the received signal $\mathsf{y}[k]$ (cf. (1)), and quantizes the residual,

$$\mathsf{w}[k] = \mathcal{Q}\big(\mathsf{y}[k] - h_1\, \varphi_{1,k}(\hat{\theta}_1) - h_2\, \varphi_{2,k}(\hat{\theta}_2)\big).$$

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. It was shown in [12] that rate-information optimal channel output compression amounts to additive Gaussian quantization noise, whose variance in our setup is given by

$$\sigma_q^2 = \sigma^2 \frac{1 + (h_1^2 \tilde{P}_1 + h_2^2 \tilde{P}_2)/\sigma^2}{2^{2R} - 1}. \tag{6}$$

Here, R denotes the quantization rate. For the MAC with noisy feedback the extra noise on the feedback channel is detrimental since it is not known by any node. In our model the transmitters also receive degraded versions of the channel output, but the receiver knows the quantization error of the feedback signal (having itself performed the quantization). With this observation, our scheme can be reduced to an equivalent MAC with perfect feedback in which the quantization error represents additional channel noise. The knowledge of this extra noise is exploited at the receiver.

III. Background

A. Gaussian Information Bottleneck

Feedback quantization causes a penalization of the mutual information underlying capacity. This penalization is captured via the information-rate function $I(R)$ for the compression of Gaussian channel outputs [12]. We thus briefly review the *information bottleneck method* [10], specifically the Gaussian information bottleneck [11].

Let $\mathsf{x} - \mathsf{y} - \mathsf{w}$ be a Markov chain, where w is a compressed representation of y and the joint distribution of x and y is known. The IBM solves the variational problem

$$\min_{p(w|y)} I(\mathsf{y};\mathsf{w}) - \beta I(\mathsf{x};\mathsf{w}). \tag{7}$$

Here, x is called the relevance variable, $I(\mathsf{x};\mathsf{w})$ is the relevant information, and $I(\mathsf{y};\mathsf{w})$ is the compression rate. The parameter $\beta > 0$ determines the trade-off between compression rate and relevant information.

We next consider the case of jointly Gaussian zero-mean random vectors $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^n$, which we assume to have full rank covariance matrices. It was shown in [16] that the optimal $\mathbf{w}$ is jointly Gaussian with $\mathbf{y}$ and can be written as

$$\mathbf{w} = \boldsymbol{A}\mathbf{y} + \boldsymbol{\xi}, \tag{8}$$

where $\boldsymbol{A}$ is a particular matrix and $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{C}_{\boldsymbol{\xi}})$ is independent of $\mathbf{y}$. We next formalize the trade-off between compression rate and relevant information.

Definition 1. *Let $\mathbf{x} - \mathbf{y} - \mathbf{w}$ be a Markov chain. The rate-information function $I : \mathbb{R}_+ \to [0, I(\mathbf{x};\mathbf{y})]$ is defined by*

$$I(R) \triangleq \max_{p(\mathbf{w}|\mathbf{y})} I(\mathbf{x};\mathbf{w}) \quad \text{subject to } I(\mathbf{y};\mathbf{w}) \leq R. \tag{9}$$

The function $I(R)$ quantifies the maximum amount of relevant information that can be preserved when the compression rate is at most R. The definition (9) is similar to rate-distortion theory, only that the minimization of distortion is replaced with a maximization of the relevant information.

Next, we briefly summarize the main results for the scalar case ($n = m = 1$) from [12]. Let $\mathsf{y} = \mathsf{x} + \mathsf{z}$ be a Gaussian channel with signal-to-noise ratio (SNR) $\gamma = P/\sigma^2$. Here, the rate-information function equals [12, Theorem 5]

$$I(R) = C(\gamma) - \frac{1}{2}\log(1 + 2^{-2R}\gamma), \tag{10}$$

with the AWGN capacity

$$C(\gamma) \triangleq \frac{1}{2}\log(1 + \gamma). \tag{11}$$

Thus, the rate-information function approaches channel capacity as the compression rate R goes to infinity. It has been shown in [17] that (10) is the solution to a specific Gaussian rate-distortion problem [18], thereby giving operational meaning to (10). The following Lemma is taken from [12].

Lemma 2. *Let $\mathsf{x} - \mathsf{y} - \mathsf{w}$ be a Markov chain with jointly Gaussian x, y. Optimal compression of y in the sense of the rate-information function yields an equivalent Gaussian channel $p(w|x)$. Therefore, a Gaussian input distribution $p(x)$*

satisfying the power constraint of the channel $p(y|x)$ with equality is capacity-achieving also for the channel $p(w|x)$.

Then, since the overall channel $p(w|x)$ is Gaussian, too, we can write $I(R) = C(\eta)$ where

$$\eta = \gamma \frac{1 - 2^{-2R}}{1 + 2^{-2R}\gamma} \leq \gamma \tag{12}$$

is the equivalent SNR of the channel $p(w|x)$. This means that we can model optimal channel output compression by an additive Gaussian noise term with variance

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

Our two-user MAC model in (1) can be rewritten as a multiple-input, single-output model in [13, Section V.-C.] and results in (13) analogously becoming (6).

B. The Original Ozarow Scheme

Ozarow derived the perfect-feedback capacity of the two-user Gaussian MAC by extending the Schalkwijk-Kailath scheme. The key concept of this scheme is the iterative refinement of the message estimate. In the first two time slots, both transmitters alternately 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 simple linear scheme is capacity-achieving. Ozarow originally assumed that the transmitters have access to perfect channel output feedback and characterized the full capacity region [4].

Theorem 3. *The capacity region of the two-user Gaussian MAC with perfect channel output feedback, transmit powers $\tilde{P}_1$ and $\tilde{P}_2$, and channel gains h_1 and h_2 is the union over $0 \leq \rho \leq 1$ of the rate pairs $(\tilde{R}_1, \tilde{R}_2)$ satisfying*

$$\tilde{R}_1 \leq \frac{1}{2} \log \left(1 + \frac{h_1^2 \tilde{P}_1}{\sigma^2} (1 - \rho^2) \right), \tag{14}$$

$$\tilde{R}_2 \leq \frac{1}{2} \log \left(1 + \frac{h_2^2 \tilde{P}_2}{\sigma^2} (1 - \rho^2) \right), \tag{15}$$

$$\tilde{R}_1 + \tilde{R}_2 \leq \frac{1}{2} \log \left(1 + \frac{h_1^2 \tilde{P}_1 + h_2^2 \tilde{P}_2 + \sqrt{h_1^2 \tilde{P}_1 h_1^2 \tilde{P}_2} \rho}{\sigma^2} \right). \tag{16}$$

IV. Results

A. Achievable Sum Rate

Our proposed scheme is a superposition of a conventional encoder and a feedback-based encoder (cf. (3)). Thus, we can maximize the sum rate by finding the optimum power (equivalently, rate) splitting between the two encodings jointly for both transmitters. The conventionally encoded signals are cancelled at the receiver before quantization such that the quantization only captures the feedback encoding in the forward path. This is possible since the receiver has the quantization noise as side-information (see [8] for details). The sum rate of the conventional coding scheme is thus given by the classical Gaussian MAC capacity C_0 [9] with noise power $\sigma^2 + h_1^2 \tilde{P}_1 + h_2^2 \tilde{P}_2$ (since the feedback-based codewords act as additional interference). Then, the achievable sum rate for fixed feedback transmit powers $\tilde{P}_1$ and $\tilde{P}_2$ and given total transmit power P is

$$\begin{aligned} R_1 + R_2 \leq & \max_{P_1 + P_2 \leq P - \tilde{P}_1 - \tilde{P}_2} C_0(h_1^2 P_1, h_2^2 P_2, \sigma^2 + h_1^2 \tilde{P}_1 + h_2^2 \tilde{P}_2) \\ & \triangleq \mathsf{C}. \end{aligned} \tag{17}$$

The sum rate of the feedback coding scheme is given by the Ozarow capacity C^{FB} [4] with noise power $\sigma^2 + \sigma_q^2$ (due to the additional quantization noise). Then, the achievable sum rate for fixed transmit powers P_1 and P_2 and given total transmit power P is

$$\begin{aligned} \tilde{R}_1 + \tilde{R}_2 \leq & \max_{\tilde{P}_1 + \tilde{P}_2 \leq P - P_1 - P_2} C^{\text{FB}}(h_1^2 \tilde{P}_1, h_2^2 \tilde{P}_2, \sigma^2 + \sigma_q^2) \\ & \triangleq \tilde{\mathsf{C}}. \end{aligned} \tag{18}$$

Thus, the total achievable sum rate of our proposed scheme with given total transmit power P yields

$$\begin{aligned} R_1 + R_2 + \tilde{R}_1 + \tilde{R}_2 \leq & \max_{P_1 + P_2 + \tilde{P}_1 + \tilde{P}_2 \leq P} \mathsf{C} + \tilde{\mathsf{C}} \\ & \triangleq C_S. \end{aligned} \tag{19}$$

Finally, simplifying the nested maximizations (implicitly contained in the expressions for C and $\tilde{\mathsf{C}}$) to one single maximization yields

$$\begin{aligned} C_S = & \max_{P_1 + P_2 + \tilde{P}_1 + \tilde{P}_2 \leq P} C_0(h_1^2 P_1, h_2^2 P_2, \sigma^2 + h_1^2 \tilde{P}_1 + h_2^2 \tilde{P}_2) \\ & + C^{\text{FB}}(h_1^2 \tilde{P}_1, h_2^2 \tilde{P}_2, \sigma^2 + \sigma_q^2). \end{aligned} \tag{20}$$

In our previous work on the symmetric MAC [8] this was the basis a one-dimensional optimization problem, since the capacity expressions directly yielded the convex part and the concave part. In the asymmetric case, (20) results in a multi-dimensional optimization problem where the individual capacity expressions do not directly yield the convex part and the concave part.

Expanding the expressions of the capacities in (20) gives

$$\begin{aligned} C_S = & \max_{P_1 + P_2 + \tilde{P}_1 + \tilde{P}_2 \leq P} C \left(\frac{h_1^2 P_1}{\sigma^2 + h_1^2 \tilde{P}_1 + h_2^2 \tilde{P}_2} \right) \\ & + C \left(\frac{h_2^2 P_2}{\sigma^2 + h_1^2 P_1 + h_1^2 \tilde{P}_1 + h_2^2 \tilde{P}_2} \right) \\ & + C^{\text{FB}}(h_1^2 \tilde{P}_1, h_2^2 \tilde{P}_2, \sigma^2 + \sigma_q^2), \end{aligned} \tag{21}$$

where the signal of transmitter 2 is decoded first and second the signal of transmitter 1. Then, the first two terms reflect the classical Gaussian MAC capacity with successive cancellation. Note that the decoding order of the successive cancellation operation is irrelevant and independent of the actual values of the channel gains h_1 and h_2. Throughout this paper we keep this order of decoding. Changing the order of decoding simply amounts to swapping h_1 and h_2.

B. Concave and Convex Component

Next, we prove that C_S can be split into a concave part $\mathsf{C}^{\cap}(\boldsymbol{p})$ and a convex part $\mathsf{C}^{\cup}(\boldsymbol{p})$ in the power allocation vector $\boldsymbol{p} = (P_1, P_2, \tilde{P}_1, \tilde{P}_2)^\top$,

$$C_S = \max_{\mathbf{1}^\top \boldsymbol{p} \leq P} \mathsf{C}^{\cap}(\boldsymbol{p}) + \mathsf{C}^{\cup}(\boldsymbol{p}). \tag{22}$$

Then, this sum rate maximization problem can be solved by difference of convex functions (DC) programming [15].

Following the lines of the symmetric case [8] and due to the fact that the pointwise minimum of two concave functions (the logarithm is concave and especially the sum (14)+(15) and (16) are concave [4]) preserves concavity [19] we directly identify $C^{\text{FB}}(h_1^2\tilde{P}_1, h_2^2\tilde{P}_2, \sigma^2 + \sigma_q^2)$ as concave in $\boldsymbol{p}$.

Furthermore, we expand $C_0(h_1^2 P_1, h_2^2 P_2, \sigma^2 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2)$ to

$$\begin{aligned} C_0(h_1^2 P_1, h_2^2 P_2, \sigma^2 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2) = \\ 1/2\log(h_1^2 P_1 + \sigma^2 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2) \\ +1/2\log(h_2^2 P_2 + \sigma^2 + h_1^2 P_1 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2) \\ -1/2\log(\sigma^2 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2) \\ -1/2\log(\sigma^2 + h_1^2 P_1 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2). \end{aligned} \tag{23}$$

The concave part $\mathsf{C}'^{\cap}(\boldsymbol{p})$ is given by

$$\begin{aligned} \mathsf{C}'^{\cap}(\boldsymbol{p}) = \;& 1/2\log(h_1^2 P_1 + \sigma^2 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2) \\ &+ 1/2\log(h_2^2 P_2 + \sigma^2 + h_1^2 P_1 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2) \end{aligned} \tag{24}$$

Since the concavity is also preserved when summing concave functions [19] the overall concave component equals

$$\begin{aligned} \mathsf{C}^{\cap}(\boldsymbol{p}) = \;& C^{\text{FB}}(h_1^2\tilde{P}_1, h_2^2\tilde{P}_2, \sigma^2 + \sigma_q^2) + \mathsf{C}'^{\cap}(\boldsymbol{p}) \quad (25) \\ = \;& C^{\text{FB}}(h_1^2\tilde{P}_1, h_2^2\tilde{P}_2, \sigma^2 + \sigma_q^2) \\ &+ 1/2\log(h_1^2 P_1 + \sigma^2 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2) \\ &+ 1/2\log(h_2^2 P_2 + \sigma^2 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2) \quad (26) \end{aligned}$$

The convexity of the overall convex component $\mathsf{C}^{\cup}(\boldsymbol{p})$,

$$\begin{aligned} \mathsf{C}^{\cup}(\boldsymbol{p}) = &-1/2\log(\sigma^2 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2) \\ &-1/2\log(\sigma^2 + h_1^2 P_1 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2), \end{aligned} \tag{27}$$

can be easily shown by direct calculation.

C. DC Programming Solution

Since the sum capacity C_S is the maximum of the sum of a convex and a concave function, or equivalently of a difference of two convex functions, the problem (20) can be solved by difference of convex functions (DC) programming [15], introduced by Yuille and Rangarajan as the convex-concave procedure (CCP) [14]. The DC programming problem in standard form is given by [15]

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) - g_0(x) \\ \text{subject to} \quad & f_i(x) - g_i(x) \leq 0, \quad i = 1, \ldots, m, \end{aligned} \tag{28}$$

where $x \in \mathbb{R}^n$ is the optimization variable and the functions $f_i : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$ have to be convex.

The maximization of (22) can be reformulated in standard form (28), i.e., involving differences of convex functions in the objective and in the constraints and yields the following theorem.

Algorithm 1 Power allocation via CCP

Require: Initial feasible point $\boldsymbol{p}_0$

1: $k := 0$
2: **while** stopping criterion not satisfied **do**
3: Form $\hat{\mathsf{C}}_k(\boldsymbol{p}) = \mathsf{C}^{\cup}(\boldsymbol{p}_k) + (\boldsymbol{p} - \boldsymbol{p}_k)^\top \nabla\mathsf{C}^{\cup}(\boldsymbol{p}_k)$
4: Determine $\boldsymbol{p}_{k+1}$ by solving the convex problem

$$\begin{aligned} \text{minimize} \quad & -\mathsf{C}^{\cap}(\boldsymbol{p}) - \hat{\mathsf{C}}_k(\boldsymbol{p}) \\ \text{subject to} \quad & \mathbf{1}^\top \boldsymbol{p} - P \leq 0, \\ & -\boldsymbol{p} \preceq 0 \end{aligned}$$

5: $k := k + 1$
6: **end while**
7: **return** $\boldsymbol{p}_k$

Theorem 4. *The problem of maximizing the sum capacity of the Gaussian MAC with quantized feedback and a superposition of feedback coding and conventional coding is solved by finding the solution of the DC problem*

$$\begin{aligned} \underset{\boldsymbol{p}}{\textit{minimize}} \quad & -\mathsf{C}^{\cap}(\boldsymbol{p}) - \mathsf{C}^{\cup}(\boldsymbol{p}) \\ \textit{subject to} \quad & \mathbf{1}^\top \boldsymbol{p} - P \leq 0, \\ & -\boldsymbol{p} \preceq 0. \end{aligned} \tag{29}$$

The basic idea of the iterative CCP algorithm [14] 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

$$\nabla\mathsf{C}^{\cup}(\boldsymbol{p}_{k+1}) = -\nabla\mathsf{C}^{\cap}(\boldsymbol{p}_k) \tag{30}$$

which itself is a convex optimization problem. The solution to this auxiliary problem decreases monotonically with increasing k and thus converges to a minimum (or saddle point). Lipp and Boyd [15] give a basic CCP algorithm which requires an initial feasible point $\boldsymbol{p}_0$, which in our case can be any point in the interval $0 \leq \mathbf{1}^\top \boldsymbol{p}_0 \leq P$. Following [15], the CCP approach leads to Algorithm 1 for power allocation.

In step 3 Algorithm 1 uses a linearization $\hat{\mathsf{C}}_k(\boldsymbol{p})$ of the convex component $\mathsf{C}^{\cup}(\boldsymbol{p})$. The gradient $\nabla\mathsf{C}^{\cup}(\boldsymbol{p})$ contained in the linearization can be directly calculated as

$$\nabla\mathsf{C}^{\cup}(\boldsymbol{p}) = -\frac{1}{2}\begin{pmatrix} \frac{h_1^2}{\sigma^2 + h_1^2 P_1 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2} \\ 0 \\ \frac{h_1^2}{\sigma^2 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2} + \frac{h_1^2}{\sigma^2 + h_1^2 P_1 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2} \\ \frac{h_2^2}{\sigma^2 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2} + \frac{h_2^2}{\sigma^2 + h_1^2 P_1 + h_1^2\tilde{P}_1 + h_2^2\tilde{P}_2} \end{pmatrix}. \tag{31}$$

If we were able to find the gradient $\nabla\mathsf{C}^{\cap}(\boldsymbol{p})$ as well and solve the iteration (30) for $\boldsymbol{p}_{k+1}$ we could avoid the auxiliary convex optimization problem. Unfortunately this seems not feasible.

D. Numerical Solution

Figure 2 shows the rates (top) and power allocations (middle and bottom) obtained by solving (29) versus the feedback

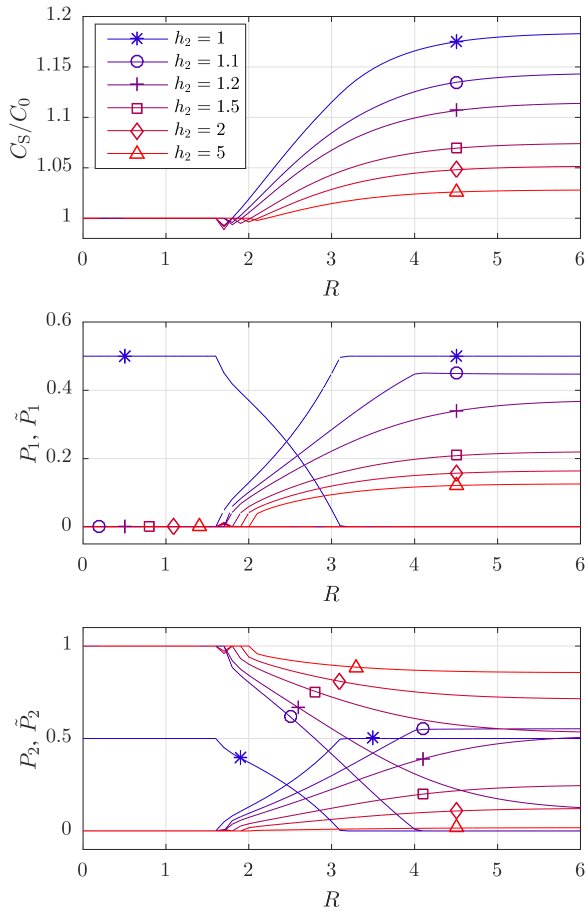

Figure 2: Top: Normalized achievable rates (in bit) with superposition coding. Middle and Bottom: Power splitting (decreasing lines for P_i, increasing lines for $\tilde{P}_i$) for various channel gains h_2. $h_1 = 1$, $P = 1$, $\sigma^2 = 10^{-2}$ fixed. Horizontal axis in all plots is the feedback quantization rate R.

quantization rate R. The linear-feedback capacity of the proposed superposition coding scheme is normalized by the no-feedback MAC capacity C_0. Clearly, when R is too small the feedback is not beneficial at all and the whole transmit power is allocated to the conventional encoding (see middle and bottom part of Fig. 2). Above a certain threshold for R, true superposition is optimal and a capacity gain is achieved. In contrast to the symmetric case where eventually pure feedback coding becomes optimal for very large R (almost perfect feedback), in the asymmetric case (in Fig. 2 for $h_2 > 1.1$) we can observe that true superposition stays optimal even for very large R and converges to a non-trivial power splitting for $R \to \infty$ (perfect feedback). The highest gain in capacity is achieved in the fully symmetric case and strictly decreases with increasing asymmetry of the channel gains.

V. Conclusions

We used the information bottleneck principle to model the quantization of the feedback in a two-user asymmetric Gaussian MAC. Using the results from our previous work on the symmetric Gaussian MAC we generalized this results for the asymmetric case and showed that most results still hold if the asymmetry is not too large. We showed that due to the rate limitation on the common feedback link it is useful to superimpose a conventional (non-feedback) scheme over a feedback coding scheme. We showed that a modified version of the Ozarow scheme and a superimposed conventional encoding scheme which are separated at the receiver 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 this two schemes can be restated as a difference of convex functions program that can be efficiently solved numerically via the concave-convex procedure algorithm. Finally, we showed that the capacity gain decreases with increasing asymmetry of the channel gains.

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] 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.
[3] 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.
[4] 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.
[5] M. Gastpar, "On Noisy Feedback in Gaussian Networks," in *Proc. 43rd Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Sept. 2005.
[6] M. Gastpar and G. Kramer, "On Cooperation Via Noisy Feedback," in *Proc. Int. Zurich Seminar on Commun.*, Zurich, Switzerland, Feb. 2006, pp. 146–149.
[7] 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.
[8] 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.
[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[10] 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.
[11] 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.
[12] 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.
[13] 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.
[14] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Computation*, vol. 15, no. 4, pp. 915–936, 2003.
[15] T. Lipp and S. Boyd, "Variations and extensions of the convex-concave procedure," http://stanford.edu/%7eboyd/papers/cvx%5fccv.html, 2014.
[16] A. Globerson and N. Tishby, "On the optimality of the Gaussian information bottleneck curve," The Hebrew University of Jerusalem, Tech. Rep., Feb. 2004.
[17] M. Meidlinger, A. Winkelbauer, and G. Matz, "On the relation between the Gaussian information bottleneck and MSE-optimal rate-distortion quantization," in *Proc. IEEE Workshop on Statistical Signal Processing (SSP 2014)*, June 2014, pp. 89–92.
[18] T. Berger, *Rate Distortion Theory*. Englewood Cliffs (NJ): Prentice Hall, 1971.
[19] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2006.