Linear Superposition Coding for the Gaussian MAC with Quantized Feedback

Stefan Farthofer and Gerald Matz

Institute of Telecommunications, Vienna University of Technology
Gusshausstrasse 25/389, 1040 Vienna, Austria
email: {stefan.farthofer, gerald.matz}@nt.tuwien.ac.at

***Abstract*—We propose a linear transceiver scheme for the symmetric two-user multiple access channel with additive white 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 of each user between an Ozarow-like code (designed for perfect rate-limited feedback) 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.**

***Index Terms*—Channel capacity, feedback, quantization, linear superposition coding, iterative refinement, information bottleneck**

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 (AWGN) 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]. 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]. It was shown by Lapidoth and Wigger that even non-perfect feedback is always beneficial [7]. Shaviv and Steinberg derived achievable rate regions for the Gaussian MAC with common rate-limited feedback using block Markov coding [8]. Unfortunately many of these schemes do not have the algorithmic simplicity of the original Ozarow scheme and the achievable rate regions are very hard to analyze. We therefore propose a simple, Ozarow-like superposition coding scheme. Our contributions are as follows.

- We study the two-user Gaussian MAC with quantized feedback. Building on the assumption that the receiver has the quantization noise as side-information, we propose a superposition of feedback-based encoding and conventional (non-feedback) coding.
- We model the quantization process as channel output compression via the information bottleneck principle [9]–[12].
- We assess the sum rate achievable with our superposition scheme and quantify the impact of the power (equivalently, rate) allocation between the two coding schemes. We show that the resulting sum rate in general is larger than the sum rates achievable with either of the two constituent schemes alone.
- We show that maximizing the sum rate by optimizing the the power (rate) allocation is a difference of convex functions (DC) problem [13] and we solve this problem numerically via the convex-concave procedure (CCP) [14].
- Eventually we show that there is a non-trivial transmit power regime where the proposed superposition coding is optimal.

The remainder of this paper is organized as follows. Section II provides the necessary background and definitions including the system model for the proposed superposition coding scheme. In Section III we asses the superposition coding scheme with quantized feedback in detail and study the tradeoff between feedback coding and non-feedback coding. Section IV studies the achievable rate region. Finally, conclusions are provided in Section V.

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 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 [15]. All logarithms are to base 2.

II. Background and Definitions

A. System Model

This paper studies the two-user Gaussian MAC with one common receiver. The users communicate independent messages θ_1, θ_1^{FB} and θ_2, θ_2^{FB} to the receiver. The messages are uniformly drawn from finite sets with cardinalities $\mathcal{M}_1 = 2^{nR_1}$, $\mathcal{M}_1^{\mathrm{FB}} = 2^{nR_1^{\mathrm{FB}}}$, $\mathcal{M}_2 = 2^{nR_2}$, $\mathcal{M}_2^{\mathrm{FB}} = 2^{nR_2^{\mathrm{FB}}}$ and mapped to the length-n transmit signals x_i, $i = 1, 2$,

Funding by WWTF Grant ICT12-054.

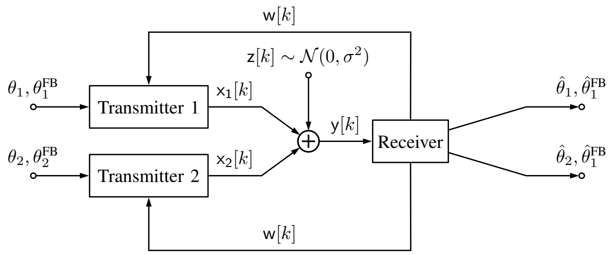

Figure 1: Two-user Gaussian MAC with common feedback link.

according to the superposition

$$x_i[k] = \varphi_{i,k}(\theta_i) + \varphi_{i,k}^{\mathrm{FB}}(\theta_i^{\mathrm{FB}}, w[1], \ldots, w[k-1]), \quad (1)$$

where $k = 1, \ldots, n$. Here, $w[k]$ is the quantized feedback from the receiver (see below) and $\varphi_{i,k} : \mathcal{M}_i \to \mathbb{R}$ denotes a conventional encoder that ignores the feedback signal. Furthermore, $\varphi_{i,k}^{\mathrm{FB}} : \mathcal{M}_i^{\mathrm{FB}} \times \mathbb{R}^{k-1} \to \mathbb{R}$ denotes a feedback-based encoder that has causal access to the quantized feedback and works similar to the original Ozarow scheme. Both encoders produce Gaussian codewords that fulfill the average power constraints

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

$$\frac{1}{n}\sum_{k=1}^{n} \mathrm{E}\{\varphi_{i,k}^{\mathrm{FB}}(\theta_i^{\mathrm{FB}}, \mathsf{w}[1], \ldots, \mathsf{w}[k-1])^2\} \leq P_i^{\mathrm{FB}}, \quad (3)$$

where the expectation is with respect to the messages and the channel noise. We assume a symmetric total transmit power constraint,

$$P_i + P_i^{\mathrm{FB}} \leq P. \quad (4)$$

The channel (see Fig. 1) introduces i.i.d. additive Gaussian noise $\mathbf{z}$. The receive signal reads

$$\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z}, \qquad \mathbf{z} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{I}). \quad (5)$$

The receiver forms the feedback signal by successive cancellation (subtracting the estimate $\hat{\mathbf{x}}^{\mathrm{NOFB}}$ of the conventional codeword from the receive signal) followed by quantization, $\mathbf{w} = \mathcal{Q}(\mathbf{y} - \hat{\mathbf{x}}^{\mathrm{NOFB}})$.

B. 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 [11]. We thus briefly review the *information bottleneck method* (IBM) [9], specifically the Gaussian information bottleneck (GIB) [10].

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

$$\min_{p(w|y)} I(\mathsf{y}; \mathsf{w}) - \beta I(\mathsf{x}; \mathsf{w}). \quad (6)$$

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

$$\mathbf{w} = \boldsymbol{A}\mathbf{y} + \boldsymbol{\xi}, \quad (7)$$

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*

$$I(R) \triangleq \max_{p(\mathbf{w}|\mathbf{y})} I(\mathbf{x}; \mathbf{w}) \quad \text{subject to } I(\mathbf{y}; \mathbf{w}) \leq R. \quad (8)$$

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 (8) 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 [11]. 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 [11, Theorem 5]

$$I(R) = C(\gamma) - \frac{1}{2}\log(1 + 2^{-2R}\gamma), \quad (9)$$

with the AWGN capacity

$$C(\gamma) \triangleq \frac{1}{2}\log(1 + \gamma). \quad (10)$$

Thus, the rate-information function approaches channel capacity as the compression rate R goes to infinity. The following Lemma is taken from [11].

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

$$\eta = \gamma\frac{1 - 2^{-2R}}{1 + 2^{-2R}\gamma} \leq \gamma \quad (11)$$

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

$$\sigma_q^2 = \sigma^2\frac{1 + \gamma}{2^{2R} - 1}. \quad (12)$$

Our two-user MAC model in (5) can be rewritten as a multiple-input, single-output model in [12, Sec. IV. C.], with the channel being the all-ones vector. The resulting rate-information function is the same as in the scalar case (9) with the difference that the total transmit power equals $2P$, i.e.,

$$I(R) = C(2\gamma) - \frac{1}{2}\log(1 + 2^{-2R}2\gamma). \tag{13}$$

It has been shown in [17] that (13) is the solution to a specific Gaussian rate-distortion problem [18], thereby giving operational meaning to (13).

C. 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 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 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]. The symmetry of the problem implies that the sum capacity is achieved with identical transmit powers P and is given by

$$C^{\mathrm{FB}} = \frac{1}{2}\log\left(1 + \frac{2P}{\sigma^2}(1+\rho^*)\right), \tag{14}$$

where the correlation coefficient ρ^* is the solution of

$$\frac{P}{\sigma^2} = \frac{2\rho^*}{(\rho^*-1)^2(\rho^*+1)}. \tag{15}$$

D. Achievable Rate Region

Our proposed scheme is a superposition of a conventional encoder and a feedback-based encoder (cf. (1)). Pure conventional encoding and pure feedback encoding are special cases obtained with $P^{\mathrm{FB}} = 0$ and $P^{\mathrm{FB}} = P$, respectively. Thus, we have to find the optimum power (equivalently, rate) allocation between the two encodings. As we argued, we assume that the conventionally encoded signals are canceled at the receiver before quantization such that the quantization only captures the feedback encoding in the forward path. This is possible since we assume that the receiver has the quantization noise as side-information. The sum capacity of the conventional coding scheme is given by the classical Gaussian MAC capacity with signal power $P–P^{\mathrm{FB}}$ and noise power $\sigma^2 + 2P^{\mathrm{FB}}$ (since the feedback-based codewords act as additional interference). The capacity of the feedback-based coding scheme is given by the Ozarow result with an SNR that is reduced due to feedback quantization. The sum capacity of the superposition scheme is then defined as follows.

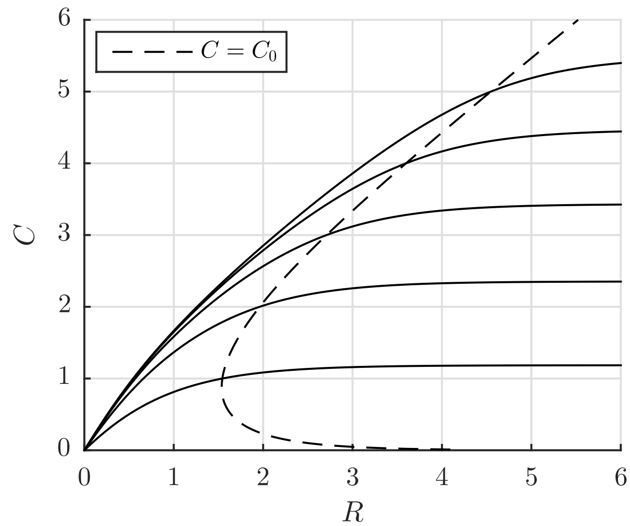

Figure 2: Sum rate for quantized feedback versus quantization rate R, parameterized by various perfect feedback sum capacities $C_0 = 1 \ldots 5$ (all quantities in bit).

Definition 3. *The sum-capacity of the Gaussian MAC with quantized feedback and a superposition of conventional and feedback-based encoding is given by the constrained maximization problem*

$$\begin{aligned} C_S = \max_{P^{FB}} \quad & \underbrace{C\left(\frac{2(P-P^{FB})}{\sigma^2+2P^{FB}}\right)}_{\textit{conventional}} + \\ & \underbrace{C\left(\frac{2P^{FB}}{\sigma^2}\frac{2^{2R}-1}{2^{2R}+P^{FB}/\sigma^2}(1+\rho^*)\right)}_{\textit{feedback-based}} \\ \textit{s.t.} \quad & 0 \le P^{FB} \le P. \end{aligned} \tag{16}$$

III. The Superposition Coding Scheme with Quantized Feedback

In our model the channel output is quantized before it is fed back to the transmitters. Thus, the transmitters receive a degraded version of the channel output. However, in our model we assume that the receiver knows the quantization noise, since usually the quantizer is located at the receiver. Therefore, we study a modified scheme with direct quantization of the received signal before message estimation. The scheme then reduces to the perfect feedback case with additional channel noise due to the quantization. The knowledge of this additional noise may be exploited at the receiver. However we only use the side-information in order to guarantee the equivalence of degraded quantized feedback and channel output quantization.

A. Feedback-Capacity Reduction due to Quantization

The quantization of the received signal is modeled as an information bottleneck, i.e., the mutual information of

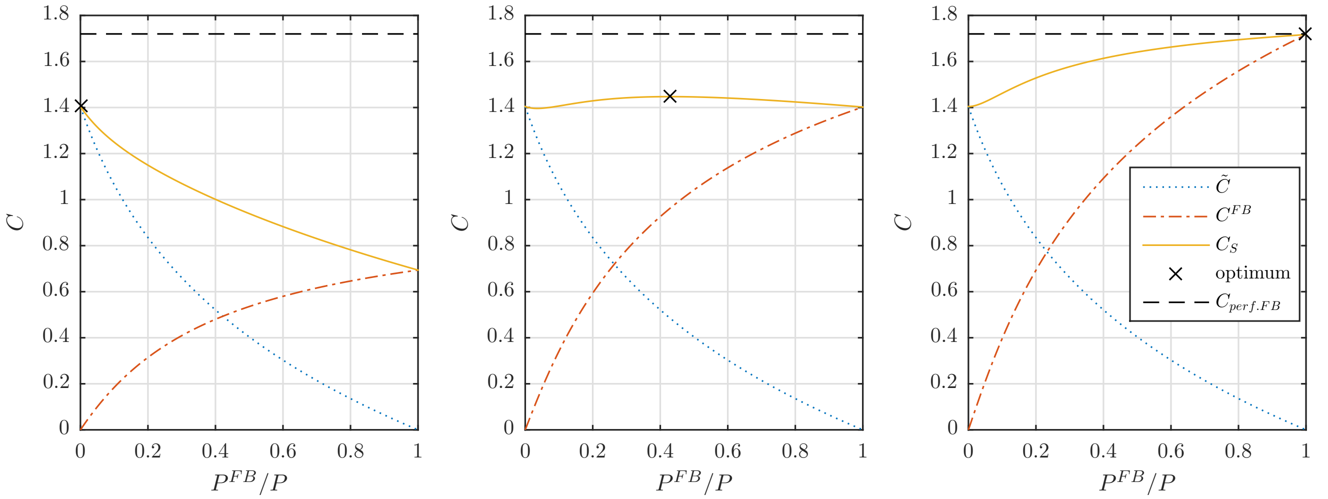

Figure 3: Sum capacity C_S versus power allocation P^FB/P for $C_0 = 1$ and $R = 0.5C_0$ (left), $R = 1.5C_0$ (middle), and $R = 5C_0$ (right); all quantities in bit.

compressed received signal and transmit signals is maximized under a rate constraint. The compression can be equivalently be seen as additional i.i.d. quantization noise σ_q^2 (12) as illustrated in Section II-B. This decreases the SNR to

$$\frac{P}{\sigma^2+\sigma_q^2} = \frac{P}{\sigma^2}\frac{1}{1+\sigma_q^2/\sigma^2} = \frac{P}{\sigma^2}\frac{1-2^{-2R}}{1+P/\sigma^2 2^{-2R}}, \qquad (17)$$

which translates into a reduced feedback capacity

$$C^\mathrm{FB} = \frac{1}{2}\log\left(1+\frac{2P}{\sigma^2}\frac{1-2^{-2R}}{1+P/\sigma^2 2^{-2R}}(1+\rho^*)\right). \qquad (18)$$

The benefit is due to the fact that we may get a higher capacity, although there is a quantization process involved, than in the case of no quantization, but also no feedback.

Fig. 2 shows the feedback-capacity C^FB versus the quantization rate R. We see that asymptotically the feedback is beneficial if the quantization rate is at least $R = C_0 - 0.5$. The point at which feedback becomes useful (corresponding to the dashed line in Fig. 2) is characterized by

$$C\left(\frac{2P}{\sigma^2}\frac{(1+\rho^*)}{1+\sigma_q^2/\sigma^2}\right) = C\left(\frac{2P}{\sigma^2}\right) \implies \rho^* = \frac{\sigma_q^2}{\sigma^2}. \qquad (19)$$

Since ρ^* must simultaneously fulfil (15) (with modified noise) we can rewrite (15) as

$$\frac{P}{\sigma^2}\frac{1}{1+\rho^*} = \frac{2\rho^*}{(\rho^*-1)^2(\rho^*+1)} \implies \frac{P}{\sigma^2} = \frac{2\rho^*}{(\rho^*-1)^2}. \qquad (20)$$

The only relevant solution of this quadratic equation in ρ^* is given by

$$\rho^* = \frac{\sigma_q^2}{\sigma^2} = \frac{P/\sigma^2 - \sqrt{2P/\sigma^2+1}}{P/\sigma^2}. \qquad (21)$$

B. Tradeoff between Feedback Coding and Non-Feedback Coding

The rates of the conventional and feedback-based encoding scheme are respectively given by

$$\tilde{C}(P^\mathrm{FB} = C\left(\frac{2(P-P^\mathrm{FB})}{\sigma^2+2P^\mathrm{FB}}\right), \qquad (22)$$

$$C^\mathrm{FB}(P^\mathrm{FB}) = C\left(\frac{2P^\mathrm{FB}}{\sigma^2}\frac{2^{2R}-1}{2^{2R}+P^\mathrm{FB}/\sigma^2}(1+\rho^*)\right). \qquad (23)$$

Clearly, $\tilde{C}(P^\mathrm{FB})$ is a strictly decreasing function in P^FB and C^FB is strictly increasing in P^FB. As a consequence there is a tradeoff between feedback-based coding and conventional coding. The optimum is either the extreme case where all transmit power is allocated to the feedback coding scheme ($P^\mathrm{FB} = P$) or to the conventional coding ($P^\mathrm{FB} = 0$), or the optimum is in fact a true superposition of both coding schemes depending on the quantization rate R. The tradeoff for this three cases is illustrated in Fig. 3.

IV. The Achievable Rate Region

A. Formulation of the Capacity Optimization Problem as a DC problem

To maximize the sum rate we have to find the ideal power allocation (rate splitting) between the conventional encoding and the feedback-based encoding. To this end, we have to solve (16), i.e.,

$$C_\mathrm{S} = \max_{P^\mathrm{FB}} \tilde{C}(P^\mathrm{FB}) + C^\mathrm{FB}(P^\mathrm{FB}) \qquad \text{s.t.} \quad 0 \le P^\mathrm{FB} \le P. \qquad (24)$$

It can be shown that $C^\mathrm{FB}(P^\mathrm{FB})$ is a concave function and $\tilde{C}(P^\mathrm{FB})$ is a convex function in P^FB. Indeed, by direct

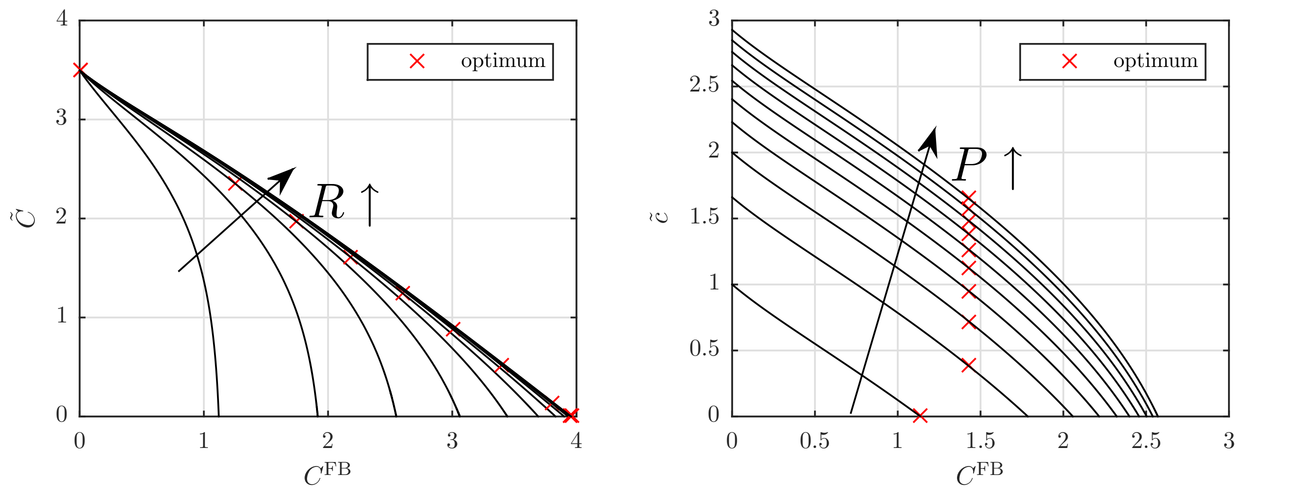

Figure 4: Achievable rate region for fixed $C_0 = 3$ and $R = 0.2C_0 \ldots 2.4C_0$ (left) and for fixed $R = 2$ and $P = 0.1P_R \ldots 2P_R$ (right); P_R is such that $C_0(P_R) = R$; all quantities in bit.

calculation we obtain the second-order derivative of $\tilde{C}(P^{\text{FB}})$ as

$$\tilde{C}''(P^{\text{FB}}) = \frac{2}{(2P_{FB} + \sigma^2)^2} \geq 0, \tag{25}$$

which proves convexity. The concavity of C^{FB} could also be proved by showing that the second derivative is non-positive. Since ρ^* is itself the solution of a cubic equation, the proof is rather lengthy and hard to show analytically. However we numerically proved that it is indeed concave. In summary, the sum capacity C_{S} is the sum of a convex and a concave function, or equivalently the difference of two convex functions. This problem can be solved by difference of convex functions (DC) programming [14], introduced by Yuille and Rangarajan as the convex-concave procedure (CCP) [13]. The DC programming problem in standard form is given by [14]

$$\begin{aligned} &\underset{x}{\text{minimize}} && f_0(x) - g_0(x) \\ &\text{subject to} && f_i(x) - g_i(x) \leq 0, \quad i = 1, \ldots, m, \end{aligned} \tag{26}$$

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. Rewriting our maximization problem (24) in standard form (26) yields the following result.

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*

$$\begin{aligned} &\underset{P^{FB}}{\textit{minimize}} && -C^{FB}(P^{FB}) - \tilde{C}(P^{FB}) \\ &\textit{subject to} && P^{FB} - P \leq 0 \\ & && -P^{FB} \leq 0. \end{aligned} \tag{27}$$

B. The Solution of the DC problem

The basic idea of the iterative CCP algorithm [13] 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

$$\frac{\partial}{\partial P^{\text{FB}}} C^{\text{FB}}\left(P^{\text{FB}}_{(k+1)}\right) = -\frac{\partial}{\partial P^{\text{FB}}} \tilde{C}\left(P^{\text{FB}}_{(k)}\right), \tag{28}$$

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 [14] give a basic CCP algorithm which requires an initial feasible point $P^{\text{FB}}_{(0)}$, which in our case can be any point in the interval $[0, P]$. Following [14], the CCP approach leads to Algorithm 1 for power allocation.

Algorithm 1 Power allocation via CCP

Require: Initial feasible point $P^{\text{FB}}_{(0)}$

1: $k := 0$
2: **while** stopping criterion not satisfied **do**
3: *Convexify.* Form $\hat{C}_{(k)}\left(P^{\text{FB}}\right) = \tilde{C}\left(P^{\text{FB}}_{(k)}\right) + \tilde{C}'\left(P^{\text{FB}}_{(k)}\right)\left(P^{\text{FB}} - P^{\text{FB}}_{(k)}\right)$
4: *Solve.* Determine $P^{\text{FB}}_{(k+1)}$ as solution of the convex problem

$$\begin{aligned} &\text{minimize} && -C^{\text{FB}}(P^{\text{FB}}) - \hat{C}_{(k)}(P^{\text{FB}}) \\ &\text{subject to} && P^{\text{FB}} - P \leq 0 \\ & && -P^{\text{FB}} \leq 0 \end{aligned}$$

5: $k := k + 1$
6: **end while**
7: **return** $P^{\text{FB}}_{(k)}$

Fig. 4 illustrates the optima which are the maxima of the sum capacities in the various achievable rate regions.

C. The Non-Trivial Region where a Superposition is Optimal

The right panel in Fig. 4 shows the achievable rate region for fixed quantization rate R and variable transmit power P. We see that up to a specific transmit power it is optimum to allocate all transmit power to the feedback-coding scheme and from a non-trivial threshold a 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. Intuitively this can be explained by studying the behaviour of the iteration (28) which characterizes a fixed point for the optimal $P^{\mathrm{FB}}_{(k)}$. In the optimum point the negative gradient of C^{FB} equals the gradient of $\tilde{C}$, which can be calculated as

$$\frac{\partial}{\partial P^{\mathrm{FB}}}\tilde{C}\left(P^{\mathrm{FB}}\right)=\frac{\partial}{\partial P^{\mathrm{FB}}}C\left(\frac{2(P-P^{\mathrm{FB}})}{\sigma^2+2P^{\mathrm{FB}}}\right)=-\frac{1}{2P^{\mathrm{FB}}+\sigma^2}. \tag{29}$$

Obviously, this gradient is independent of the total available transmit power and the equation for C^{FB} does not contain the total transmit power at all. Therefore, the optimum must also be independent of the actual value of the total transmit power if the total transmit power is at least $P \geq P^{\mathrm{FB}}$ and is limited by the quantization rate R.

V. Conclusions

We used the information bottleneck principle to model the quantization of the feedback in a two-user Gaussian MAC. 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 (DC) program that can be efficiently solved numerically via the concave-convex procedure (CCP) algorithm. Finally, we showed that true superposition is optimum when a certain transmit power threshold is exceeded.

References

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

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

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

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

[5] M. Gastpar, "On Noisy Feedback in Gaussian Networks," in *43rd Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, 2005.

[6] M. Gastpar and G. Kramer, "On Cooperation Via Noisy Feedback," in *International Zurich Seminar on Communications*, 2006, pp. 146–149.

[7] A. Lapidoth and M. Wigger, "On the AWGN MAC With Imperfect Feedback," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5432–5476, 2010.

[8] D. Shaviv and Y. Steinberg, "On the Multiple-Access Channel With Common Rate-Limited Feedback," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3780–3795, 2013.

[9] N. Tishby, F. Pereira, and W. Bialek, "The Information Bottleneck Method," in *Proc. 37th Allerton Conf. on Communication, Control, and Computing*, Sept. 1999, pp. 368–377.

[10] 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.

[11] A. Winkelbauer and G. Matz, "Rate-Information-Optimal Gaussian Channel Output Compression," in *Proc. 48th Annual Conference on Information Sciences and Systems (CISS 2014)*, March 2014.

[12] 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)*, 2014.

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

[14] T. Lipp and S. Boyd, "Variations and extensions of the convex-concave procedure," 2014.

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

[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.

[18] T. Berger, *Rate Distortion Theory*. Englewood Cliffs (NJ): Prentice Hall, 1971.