Achieving Positive Rates over AWGN Channels with Quantized Feedback and Linear Processing

Stefan Farthofer, Andreas Winkelbauer, and Gerald Matz

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

***Abstract*—It is well known that feedback does not increase the capacity of the AWGN channel. Schalkwijk and Kailath presented a feedback scheme that dramatically improves the error probability for finite block lengths. Unfortunately, the Schalkwijk-Kailath scheme requires perfect feedback, which is often not feasible. We study the performance of linear transmission schemes for AWGN channels with separate quantization of the feedback signal and the receiver input signal. The quantizers are designed in a rate-information-optimal manner using the Gaussian information bottleneck. We provide expressions for achievable rates and error probabilities in terms of the so-called information-rate function. It turns out the exploiting knowledge of the quantized feedback signal at the receiver is crucial for achieving strictly positive rates.**

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

I. Introduction

It is well known that channel output feedback does not increase the capacity of memoryless point-to-point channels [1]. However, feedback can dramatically improve the reliability and reduce the coding complexity of capacity-achieving transmission schemes. For the Gaussian channel, Schalkwijk and Kailath presented a remarkably simple scheme which achieves a probability of error that decays doubly exponentially in the block length for all rates below capacity [2], [3]. The Schalkwijk-Kailath scheme has later been extended to make the error probability decay with an exponential *order* that increases linearly in the block length [4]–[6]. Furthermore, Schalkwijk-Kailath coding was shown to achieve the feedback capacity of Gaussian channels with arbitrary autoregressive moving-average noise processes [7].

The key ingredient of the Schalkwijk-Kailath scheme is a stochastic approximation method also known as *iterative refinement* which allows the transmitter to iteratively improve the receiver's estimate of the intended message. A critical requirement for this approach to work is perfect feedback; in case of noisy feedback, the Schalkwijk-Kailath scheme breaks down. In fact, with noisy feedback no linear scheme can achieve positive rates [8]. Noiseless feedback of the analog channel output is unrealistic since in practice the feedback channel would at least have a rate constraint. The rate constraint can be accounted for by quantizing the channel output signal prior to transmitting the feedback. In this context, [9] discusses a quantized feedback scheme for channels with a peak power constraint and [10] augments the linear feedback scheme of [11] with a quantizer.

In this paper, we propose a scheme which is based on [11] but uses separate quantizers for the feedback loop and for the receiver processing. These two quantizers are optimized using the *information bottleneck* (IB) principle [12]. In contrast to [10], we do not assume that the transmitter has access to the quantized channel noise. We show that the proposed scheme achieves positive rates even if the feedback signal is quantized more coarsely than the receiver input signal.

The remainder of this paper is organized as follows. Section II provides the necessary background and definitions. The system model for the proposed quantized feedback scheme is discussed in Section III. In Section IV, we assess the performance of the considered linear feedback scheme in the regime of large block length both for perfect and noisy feedback. Section V proposes a modified linear feedback scheme that always achieves positive rates. Section VI shows numerical results for the finite block length regime that corroborate our findings. Finally, conclusions are provided in Section VII.

Notation: We use boldface letters for column vectors and upright sans-serif letters for random variables. The identity matrix is denoted by $\boldsymbol{I}$ and the superscript $^{\mathrm{T}}$ denotes transposition. The 2-norm and Frobenius norm are denoted by $\|\cdot\|$ and $\|\cdot\|_{\mathsf{F}}$, respectively. Probability and expectation are denoted by $\mathbb{P}\{\cdot\}$ and $\mathbb{E}\{\cdot\}$, respectively. A Gaussian distribution with mean μ and variance σ^2 is denoted by $\mathcal{N}(\mu, \sigma^2)$, and $\mathcal{Q}(\cdot)$ is the Gaussian Q-function. We use the notation of [13] for the mutual information $I(\cdot;\cdot)$. All logarithms are to base 2.

II. Background and Definitions

A. Gaussian Information Bottleneck

The IB method [12] to date has received limited attention outside the machine learning community. We therefore give a brief overview of the IB principle and discuss its solution in the Gaussian case. This allows us to characterize rate-information-optimal quantization.

Let $\mathsf{x} - \mathsf{y} - \mathsf{z}$ be a Markov chain, where z is a compressed representation of y and the joint distribution of x and y is known. The IB considers the following variational problem:

$$\min_{p(z|y)} I(\mathsf{y};\mathsf{z}) - \lambda I(\mathsf{x};\mathsf{z}). \tag{1}$$

Funding by WWTF Grant ICT12-054.

In the context of the IB, x is called the *relevance variable* and the mutual information $I(\mathsf{x};\mathsf{z})$ is called *relevant information*. The trade-off between compression rate $I(\mathsf{y};\mathsf{z})$ and relevant information is determined by the Lagrange parameter λ. We next formalize the trade-off between compression rate and relevant information.

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

$$I(R) \triangleq \max_{p(\mathsf{z}|\mathsf{y})} I(\mathsf{x};\mathsf{z}) \quad \text{subject to } I(\mathsf{y};\mathsf{z}) \le R. \tag{2}$$

It has been shown in [14] that (2) can equivalently be restated as a rate-distortion problem [15] in the Gaussian case. The rate-distortion theorem thus gives operational meaning to (2), i.e., there exists a code which achieves $I(R)$. Hence, the function $I(R)$ allows us to quantify the maximum achievable information rate when the compression rate is R.

If x and y are jointly Gaussian, (1) is referred to as Gaussian information bottleneck (GIB) [16]. In this case, explicit expressions for the information-rate function can be found [17], [18]. Specifically, for an AWGN channel with signal-to-noise ratio (SNR) $\rho \triangleq P/\sigma^2$ and channel output compression we have [17, Theorem 2]

$$I(R) \,=\, C(\rho) - C(2^{-2R}\rho) \,\le\, \min\{C(\rho), R\}. \tag{3}$$

Here,

$$C(\rho) \triangleq \frac{1}{2}\log(1+\rho) \tag{4}$$

is the capacity of the uncompressed channel. The compressed channel is again Gaussian and hence $I(R) = C(\hat{\rho})$ with the equivalent SNR

$$\hat{\rho} = \rho\frac{1-2^{-2R}}{1+2^{-2R}\rho} \le \rho. \tag{5}$$

Equivalently, GIB-optimal channel output compression amounts to additional additive Gaussian noise of variance

$$\sigma^2\frac{1+\rho}{2^{2R}-1}. \tag{6}$$

The relationship between rate-information-optimal compression and rate-distortion-optimal compression in the Gaussian case is studied in [14].

B. The Linear Feedback Scheme of [11]

We consider an AWGN channel with noisy channel output feedback. The channel output at time k is given by

$$\mathsf{y}[k] = \mathsf{x}[k] + \mathsf{u}[k], \tag{7}$$

where $\mathsf{u}[k] \sim \mathcal{N}(0, \sigma_{\mathsf{u}}^2)$ is independent, identically distributed Gaussian noise. The received signal $\mathsf{y}[k]$ is then fed back and in the next time instant the transmitter obtains the noisy version

$$\tilde{\mathsf{y}}[k] = \mathsf{y}[k-1] + \mathsf{v}[k] = \mathsf{x}[k-1] + \mathsf{w}[k], \tag{8}$$

where $\mathsf{w}[k] \triangleq \mathsf{u}[k-1] + \mathsf{v}[k]$. Here, $\mathsf{v}[k] \sim \mathcal{N}(0, \sigma_{\mathsf{v}}^2)$ is independent, identically distributed Gaussian feedback noise. Thus, at time k the transmitter has noisy side information $\{\tilde{\mathsf{y}}[k']\}_{k'<k}$ about all previous channel outputs. Since $\mathsf{x}[k]$ is available at the transmitter, knowing $\tilde{\mathsf{y}}[k]$ is equivalent to knowing the aggregate (channel and feedback) noise $\{\mathsf{w}[k']\}_{k'<k}$ (cf. (8)). The message θ is transmitted using n channel uses and, since this scheme performs iterative refinement, the block length n is equal to the number of iterations.

Using matrix-vector notation, the linear processing at the receiver and the transmitter can be written as follows (all vectors in (9) and (10) are of length n):

$$\mathbf{x} = \boldsymbol{g}\theta + \boldsymbol{F}\mathbf{w}, \tag{9}$$

$$\hat{\theta} = \boldsymbol{q}^{\mathrm{T}}\mathbf{y}, \tag{10}$$

where $\boldsymbol{F} \in \mathbb{R}^{n\times n}$ is a strictly lower triangular matrix (to ensure causality) and $\boldsymbol{g}, \boldsymbol{q} \in \mathbb{R}^n$ are unit-norm vectors. The transmit signal $\mathbf{x}$ consists of the message part $\boldsymbol{g}\theta$ and the noise cancellation part $\boldsymbol{F}\mathbf{w}$ (cf. (9)). Since all quantities in (9) and (10) are Gaussian and the processing is linear, the overall scheme amounts to a Gaussian "superchannel" with capacity

$$C_S = \frac{1}{n}C(\mathrm{snr}), \tag{11}$$

where the corresponding SNR follows from (10) and reads

$$\mathrm{snr} = \frac{\mathbb{E}\{\theta^2\}|\boldsymbol{q}^{\mathrm{T}}\boldsymbol{g}|^2}{\sigma_{\mathsf{u}}^2\left\|\boldsymbol{q}^{\mathrm{T}}(\boldsymbol{I}+\boldsymbol{F})\right\|^2 + \sigma_{\mathsf{v}}^2\left\|\boldsymbol{q}^{\mathrm{T}}\boldsymbol{F}\right\|^2}. \tag{12}$$

The factor $1/n$ in (11) is due to the fact that n iterations are used to transmit the message θ. The linear feedback scheme is optimized by designing $\boldsymbol{F}$, $\boldsymbol{g}$, and $\boldsymbol{q}$ such that the SNR in (12) (and hence the capacity C_S) is maximized.

Let the transmit signal satisfy the average power constraint

$$\mathbb{E}\{\|\mathbf{x}\|^2\} = \mathbb{E}\{\theta^2\} + (\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{v}}^2)\|\boldsymbol{F}\|_{\mathsf{F}}^2 \le nP. \tag{13}$$

Using a power allocation factor $\gamma \in [0,1]$, a fraction of $1-\gamma$ of the total transmit power is allocated to the message part, i.e., we have

$$\mathbb{E}\{\theta^2\} = (1-\gamma)nP, \tag{14}$$

$$(\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{v}}^2)\|\boldsymbol{F}\|_{\mathsf{F}}^2 \le \gamma nP. \tag{15}$$

For fixed γ, [11, Lemma 5] states that the optimal $\boldsymbol{F} = (f_{ij})_{1\le i,j\le n}$, $\boldsymbol{g}$, and $\boldsymbol{q}$ are given by

$$f_{ij} = -\frac{1-\beta_0^2}{\sigma_{\mathsf{u}}^2+\sigma_{\mathsf{v}}^2}\sigma_{\mathsf{u}}^2\beta_0^{i-j-2} \qquad \text{for } i > j, \tag{16}$$

$$\boldsymbol{g} = \boldsymbol{q} = \sqrt{\frac{1-\beta_0^2}{1-\beta_0^{2n}}}\left[1 \;\beta_0\; \beta_0^2 \;\ldots\; \beta_0^{n-1}\right]^{\mathrm{T}}, \tag{17}$$

where $\beta_0 \in (0,1)$ is the smallest positive root of

$$\beta^{2n}\sigma_{\mathsf{u}}^4 - \beta^2 n(\sigma_{\mathsf{u}}^4 + (\sigma_{\mathsf{u}}^2+\sigma_{\mathsf{v}}^2)\gamma P) + (n-1)\sigma_{\mathsf{u}}^4. \tag{18}$$

Using (16)–(18) in (12) yields

$$\mathrm{snr} = \frac{(\sigma_{\mathsf{u}}^2+\sigma_{\mathsf{v}}^2)(1-\gamma)nP}{\sigma_{\mathsf{u}}^2\sigma_{\mathsf{v}}^2 + \sigma_{\mathsf{u}}^4\beta_0^{2(n-1)}} \tag{19}$$

for the SNR. The power allocation factor γ can then be optimized numerically (see [11] for details).

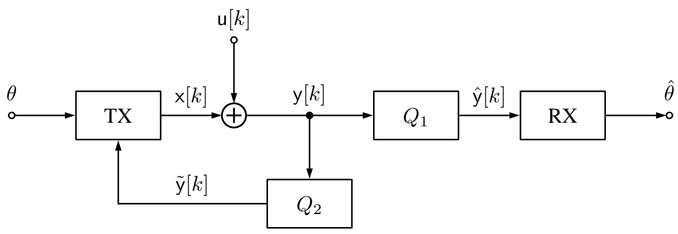

Figure 1: Feedback system with separate quantization of channel output and feedback.

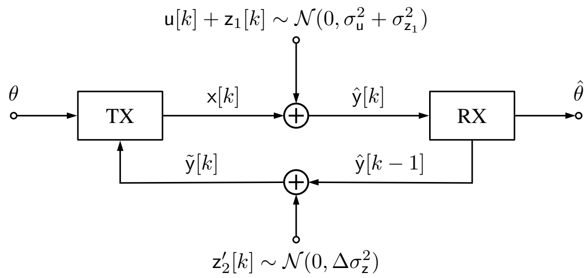

Figure 2: Equivalent system model.

III. System Model

In what follows we apply the linear feedback scheme described above to a scenario with two quantizers Q_1 and Q_2 as depicted in Fig. 1. This setup models the practically relevant case where the receiver observes a quantized version of the channel output and the feedback link is rate-constrained (but otherwise noise-free). We denote the quantized signals by $\hat{\mathsf{y}}[k] = Q_1(\mathsf{y}[k])$ and $\tilde{\mathsf{y}}[k] = Q_2(\mathsf{y}[k-1])$. The compression rates of the two quantizers are denoted by R_{Q_1} and R_{Q_2}. We focus on the case $R_{Q_2} \le R_{Q_1}$. The case $R_{Q_2} > R_{Q_1}$ is briefly discussed at the end of Section IV.

We assume that rate-information-optimal quantizers Q_1 and Q_2 are used that have been designed using the IB principle (cf. Section II-A). Thus, Q_1 and Q_2 amount to additive Gaussian quantization noise,

$$\hat{\mathsf{y}}[k] = \mathsf{y}[k] + \mathsf{z}_1[k], \qquad \tilde{\mathsf{y}}[k] = \mathsf{y}[k-1] + \mathsf{z}_2[k]. \tag{20}$$

The variance of the quantization noise is obtained from (6) with $\sigma^2 = \sigma_{\mathsf{u}}^2$ and $\rho = P/\sigma_{\mathsf{u}}^2$, i.e., $\mathsf{z}_1[k] \sim \mathcal{N}(0, \sigma_{\mathsf{z}_1}^2)$ and $\mathsf{z}_2[k] \sim \mathcal{N}(0, \sigma_{\mathsf{z}_2}^2)$ with $\sigma_{\mathsf{z}_i}^2 = \sigma_{\mathsf{u}}^2(1+\rho)/(2^{2R_{Q_i}} - 1)$. Our assumption on the compression rates allows us to equivalently rewrite $\tilde{\mathsf{y}}[k]$ in (20) as

$$\tilde{\mathsf{y}}[k] = Q_2'\big(Q_1(\mathsf{y}[k-1])\big) = Q_2'(\hat{\mathsf{y}}[k-1]) = \hat{\mathsf{y}}[k-1] + \mathsf{z}_2'[k], \tag{21}$$

where the Gaussian quantization noise $\mathsf{z}_2'[k]$ has variance $\Delta\sigma_{\mathsf{z}}^2 \triangleq \sigma_{\mathsf{z}_2}^2 - \sigma_{\mathsf{z}_1}^2 \ge 0$ and the noise realizations are related as $z_2[k] = z_1[k] + z_2'[k]$. Therefore, $\tilde{\mathsf{y}}[k]$ is a physically degraded version of $\hat{\mathsf{y}}[k-1]$. The additional quantization of $\hat{\mathsf{y}}[k-1]$ with quantizer Q_2' in (21) is performed with rate

$$R_{Q_2}' = \frac{1}{2}\log\frac{2^{2R_{Q_1}} - 1}{2^{2(R_{Q_1} - R_{Q_2})} - 1}. \tag{22}$$

The compression rate in (22) is such that the quantization noise $\mathsf{z}_2'[k]$ has variance $\Delta\sigma_{\mathsf{z}}^2$. Writing the quantized feedback as in (21) allows us to model the system in Fig. 1 by an equivalent AWGN channel with noisy feedback as shown in Fig. 2.

IV. Performance Analysis

A. Preliminaries

We next analyze the performance of the quantized feedback scheme in terms of the error probability $\mathbb{P}\{\hat{\theta} \neq \theta\}$. To this end, we use the following upper bound on the error probability of transmitting equidistant symbols over an AWGN channel:

$$P_e^{(n)} \le 2Q\left(\frac{c_0\sqrt{\mathrm{snr}}}{2^{nR}}\right). \tag{23}$$

Here, c_0 is a constant that depends on the symbol constellation and R is the data rate in bits per channel use. We use the standard definition for achievable rates.

Definition 2. *A data rate R is said to be achievable, if the error probability $P_e^{(n)}$ vanishes as n tends to infinity, i.e., if*

$$\lim_{n\to\infty} P_e^{(n)} = 0. \tag{24}$$

B. Identical Quantization Rates ($R_{Q_1} = R_{Q_2}$)

We first consider the case $R_Q \triangleq R_{Q_1} = R_{Q_2}$ which yields $\tilde{\mathsf{y}}[k] = \hat{\mathsf{y}}[k]$. Here, the feedback is noiseless with respect to the output of quantizer Q_1 and the capacity of the quantized channel is given by

$$I(R_Q) = C\left(\frac{P}{\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{z}_1}^2}\right). \tag{25}$$

Along the lines of Section II-B, the SNR reads

$$\mathrm{snr} = \frac{\mathbb{E}\{\theta^2\}|\boldsymbol{q}^{\mathrm{T}}\boldsymbol{g}|^2}{(\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{z}_1}^2)\left\|\boldsymbol{q}^{\mathrm{T}}(\boldsymbol{I}+\boldsymbol{F})\right\|^2} = \frac{(1-\gamma)nP}{(\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{z}_1}^2)\beta_0^{2(n-1)}}, \tag{26}$$

where $\beta_0 \in (0,1)$ is the smallest positive root of

$$\beta^{2n}(\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{z}_1}^2) - \beta^2 n(\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{z}_1}^2 + \gamma P) + (\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{z}_1}^2)(n-1), \tag{27}$$

and the power constraint is

$$\mathbb{E}\{\|\mathbf{x}\|^2\} = \mathbb{E}\{\theta^2\} + (\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{z}_1}^2)\|\boldsymbol{F}\|_{\mathsf{F}}^2 \le nP. \tag{28}$$

Next, we analyze the asymptotic behavior of the SNR as $n \to \infty$. To this end, we note that for large n (27) yields

$$\beta_0^2 = \frac{\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{z}_1}^2}{\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{z}_1}^2 + \gamma P}. \tag{29}$$

Furthermore, we let

$$\lim_{n\to\infty}\frac{(1-\gamma)nP}{\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{z}_1}^2 + \gamma P} = 1 \quad\Longrightarrow\quad \lim_{n\to\infty}\gamma = 1. \tag{30}$$

Using (29) and (30) in (26) allows us to express the SNR for large n as follows:

$$\mathrm{snr} = \left(1 + \frac{P}{\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{z}_1}^2}\right)^n. \tag{31}$$

Due to (31) the capacity of the superchannel is (cf. (11))

$$C_S = C\left(\frac{P}{\sigma_{\mathsf{u}}^2 + \sigma_{\mathsf{z}_1}^2}\right) = I(R_Q), \tag{32}$$

which is equal to the channel capacity (25).

Proposition 3. *For large block length n and quantization rate $R_Q = R_{Q_1} = R_{Q_2}$, the error probability of the proposed scheme decreases doubly exponentially and is upper bounded as*

$$P_e^{(n)} \le 2\mathcal{Q}\Big(2^{n(I(R_Q)-R)}\Big), \tag{33}$$

for all data rates $R < I(R_Q)$.

Proof: Use (32) to express snr in terms of $I(R_Q)$ and insert the result into (23). ∎

Proposition 3 confirms that the proposed scheme performs asymptotically like the Schalkwijk-Kailath scheme if the compression rates of the two quantizers are equal.

C. Different Quantization Rates ($R_{Q_1} \neq R_{Q_2}$)

We next consider the case $R_{Q_2} < R_{Q_1}$, i.e., the feedback is a degraded version of the output of quantizer Q_1. The SNR is then given by

$$\text{snr} = \frac{(\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_2}^2)(1-\gamma)nP}{(\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_1}^2)\Delta\sigma_\mathsf{z}^2 + (\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_1}^2)^2\beta_0^{2(n-1)}}, \tag{34}$$

where $\beta_0 \in (0,1)$ is the smallest positive root of

$$\beta^{2n}(\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_1}^2)^2 - \beta^2 n\left((\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_1}^2)^2 + (\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_2}^2)\gamma P\right) + (\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_1}^2)^2(n-1). \tag{35}$$

For large n, (35) yields

$$\beta_0^2 = \frac{(\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_1}^2)^2}{(\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_1}^2)^2 + (\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_2}^2)\gamma P}. \tag{36}$$

By inserting (36) into (34) and appropriately letting $\gamma \to 0$ as $n \to \infty$, we obtain

$$\text{snr} = \frac{nP}{\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_1}^2} \tag{37}$$

for the asymptotic behavior of the SNR. Since the SNR in (37) does not grow exponentially in n, the capacity of the superchannel vanishes as $n \to \infty$. This result agrees with the findings in [8].

In the case $R_{Q_1} < R_{Q_2}$, the receiver input $\hat{\mathsf{y}}[k]$ is degraded with respect to the feedback $\tilde{\mathsf{y}}[k]$. Similarly to (21) we can write $\hat{\mathsf{y}}[k-1] = \tilde{\mathsf{y}}[k] + \mathsf{z}_1'[k]$, where $\mathsf{z}_1'[k] \sim \mathcal{N}(0, \sigma_{\mathsf{z}_1}^2 - \sigma_{\mathsf{z}_2}^2)$. From the transmitter's perspective, the feedback signal $\tilde{\mathsf{y}}[k] = \hat{\mathsf{y}}[k-1] - \mathsf{z}_1'[k]$ is a "noisy" version of what the receiver observes. Therefore, also in this case no positive rate can be achieved.

V. Modified Feedback Scheme

We next assume that the receiver knows both $\hat{y}[k]$ and $\tilde{y}[k]$. This is reasonable since typically both quantizers are part of the receiver. With this assumption, a simple modification of the receiver processing allows us to achieve positive rates. We first consider the case $R_{Q_1} > R_{Q_2}$ in what follows.

The transmit signal is given by

$$\mathbf{x} = \boldsymbol{g}\theta + \boldsymbol{F}(\mathbf{u} + \mathbf{z}_2) \tag{38}$$

and the modified receiver processing is

$$\hat{\theta} = \boldsymbol{q}^\mathrm{T}\big(\hat{\mathbf{y}} - \boldsymbol{F}(\tilde{\mathbf{y}} - \hat{\mathbf{y}})\big) = \theta + \boldsymbol{q}^\mathrm{T}(\boldsymbol{I} + \boldsymbol{F})(\mathbf{u} + \mathbf{z}_1). \tag{39}$$

Subtracting $\boldsymbol{F}(\tilde{\mathbf{y}} - \hat{\mathbf{y}})$ cancels the extra feedback quantization noise $\mathbf{z}_2 - \mathbf{z}_1$ (whose variance is $\Delta\sigma_\mathsf{z}^2$). We note that (39) has the same structure as in the case $R_{Q_1} = R_{Q_2}$. However, the noise cancellation at the receiver does not impact the transmit signal (38) which still contains the additional feedback quantization noise. Therefore, designing the modified scheme for the case $R_Q = R_{Q_1}$ would violate the power constraint.

To meet the power constraint, we have to downscale the transmit power by the factor

$$\frac{\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_1}^2}{\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_1}^2 + \Delta\sigma_\mathsf{z}^2} = \frac{\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_1}^2}{\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_2}^2}. \tag{40}$$

Designing $\boldsymbol{F}$, $\boldsymbol{g}$, and $\boldsymbol{q}$ for the case $R_Q = R_{Q_1}$ and taking $\Delta\sigma_\mathsf{z}^2 > 0$ into account by scaling P with (40) yields the asymptotic SNR (cf. (31))

$$\text{snr} = \left(1 + \frac{P}{\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_1}^2}\frac{\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_1}^2}{\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_2}^2}\right)^n. \tag{41}$$

Hence, the capacity of the superchannel is given by

$$C_S = C\left(\frac{P}{\sigma_\mathsf{u}^2 + \sigma_{\mathsf{z}_2}^2}\right) = I(R_{Q_2}). \tag{42}$$

The modified receiver processing (39) thus allows us to achieve positive rates, in contrast to the original scheme discussed in Section IV-C.

Proposition 4. *For large block length n and quantization rates $R_{Q_1} \ge R_{Q_2}$, the error probability of the modified feedback scheme decreases doubly exponentially and is upper bounded as*

$$P_e^{(n)} \le 2\mathcal{Q}\Big(2^{n(I(R_{Q_2})-R)}\Big), \tag{43}$$

for all data rates $R < I(R_{Q_2})$.

Proof: The proof is analogous to that of Proposition 3. ∎

By a similar argument for $R_{Q_2} \ge R_{Q_1}$ it can be shown that the modified scheme achieves any rate $R < I\big(\min\{R_{Q_1}, R_{Q_2}\}\big)$ for arbitrary R_{Q_1} and R_{Q_2}.

VI. Numerical Results

The previous discussion provided asymptotic statements about achievable rates and error probabilities. Especially for small block lengths these results do not give us a complete picture of the performance of the proposed quantized feedback scheme. We thus numerically optimized the power allocation factor γ and evaluated (23). Furthermore, for finite block length the asymptotic expressions (32) and (42) for the capacity of the superchannel are not exact. For Gaussian messages, we have $C_S = \frac{1}{n}I(\hat{\theta};\theta)$.

Fig. 3 shows the results for a data rate of $R = 1\,\text{bit}$ in an AWGN channel with capacity $C = 1.25\,\text{bit}$. With high-rate quantizers $R_{Q_1} = R_{Q_2} = R_Q = 10\,\text{bits}$ (green lines, $\star$ markers), we have $R_Q \gg C$ and thus $\frac{1}{n}I(\hat{\theta};\theta) \approx C$. Note that $R_{Q_1} = R_{Q_2}$ implies $R_{Q_2}' = \infty$ (cf. (22)). The mutual

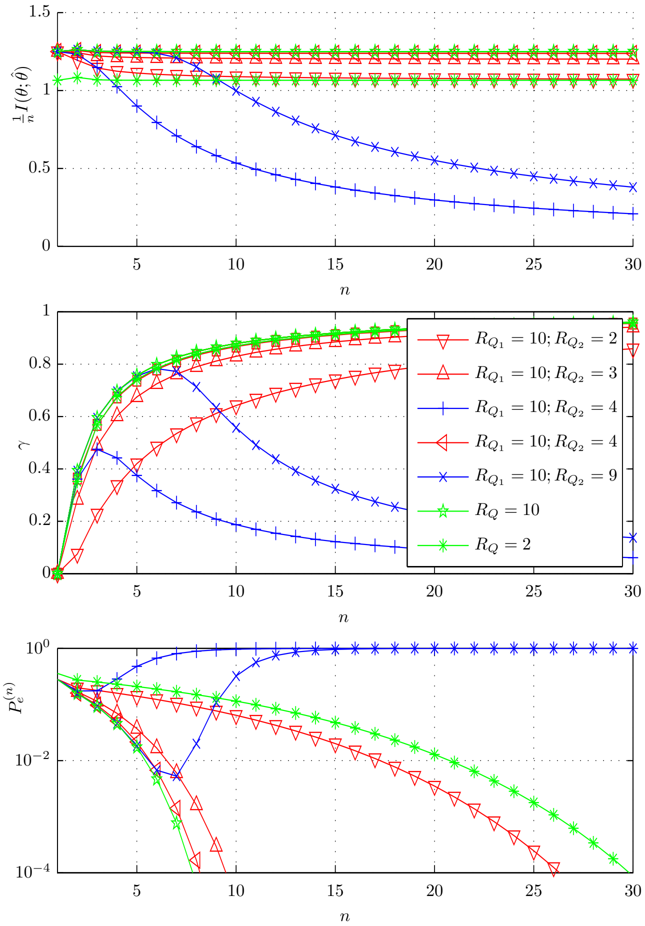

Figure 3: Numerical performance evaluation for finite block lengths: mutual information $\frac{1}{n}I(\hat{\theta};\theta)$ (top); optimal power allocation factor γ (middle); error probability $P_e^{(n)}$ (bottom).

information for the modified feedback scheme (red lines, triangle markers) approaches $I(R_{Q_2})$. In case the quantized feedback signal is not exploited at the receiver (blue lines, cross markers), the mutual information is almost constant up to a certain block length beyond which the performance breaks down rapidly. For $R_{Q_1} \geq R_{Q_2}$, the performance in the asymptotic regime is independent of the actual value of R_{Q_1}. However, this is not true for finite block lengths. We compare the error probabilities for $R_{Q_1} = R_{Q_2} = R_Q = 2$ bits (green lines, $*$ markers) and for the modified scheme with $R_{Q_1} = 10$ bits and $R_{Q_2} = 2$, 3, or 4 bits (red lines, triangle markers). We observe that increasing R_{Q_1} provides a substantial performance improvement for the considered block lengths, although the asymptotic performance is the same for both cases.

VII. Conclusions

We have studied the capacity of an optimal linear feedback scheme for AWGN channels with quantization of the feedback signal and of the receive signal. We have used the GIB, which was recently specialized to optimal output compression of AWGN channels, to design the quantizers in a rate-information-optimal manner. It turned out that the GIB information-rate function is of central importance for the performance of such systems in that it characterizes the achievable data rates as functions of the compression rates. We have shown that the scheme from [11] achieves the same doubly exponential error probability decay as the Schalkwijk-Kailath coding scheme if the same quantizer is used for feedback and receiver processing. Since the scheme from [11] cannot achieve positive rates for different feedback and receiver quantization rates, we have proposed a modified scheme that exploits knowledge of the quantized feedback signal at the receiver. We demonstrated that in this novel linear transceiver processing scheme in combination with GIB quantization results in an equivalent AWGN channel of capacity $I(\min\{R_{Q_1}, R_{Q_2}\})$, in contrast to the results in [8], [11].

References

[1] C. Shannon, “The zero error capacity of a noisy channel,” *IRE Trans. on Inf. Theory*, vol. 2, no. 3, pp. 8–19, Sept. 1956.
[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] A. Kramer, “Improving communication reliability by use of an intermittent feedback channel,” *IEEE Trans. Inf. Theory*, vol. 15, no. 1, pp. 52–60, Jan. 1969.
[5] K. S. Zigangirov, “Upper bounds for the error probability for channels with feedback,” *Probl. Pered. Inform.*, vol. 6, no. 2, pp. 87–92, 1970.
[6] R. Gallager and B. Nakiboglu, “Variations on a theme by Schalkwijk and Kailath,” *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 6–17, Jan. 2010.
[7] Y.-H. Kim, “Feedback capacity of stationary Gaussian channels,” *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 57–85, Jan. 2010.
[8] Y.-H. Kim, A. Lapidoth, and T. Weissman, “The Gaussian channel with noisy feedback,” in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT 2007)*, June 2007, pp. 1416–1420.
[9] N. Martins and T. Weissman, “Coding for additive white noise channels with feedback corrupted by quantization or bounded noise,” *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4274–4282, Sept. 2008.
[10] Z. Chance and D. Love, “On quantization of channel output feedback for the Gaussian channel,” in *Proc. International Waveform Diversity and Design Conference (WDD 2010)*, Aug. 2010, pp. 76–80.
[11] ——, “Concatenated coding for the AWGN channel with noisy feedback,” *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6633–6649, Oct. 2011.
[12] 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.
[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[14] 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.
[15] T. Berger, *Rate Distortion Theory*. Englewood Cliffs (NJ): Prentice Hall, 1971.
[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 Annual Conference on Information Sciences and Systems (CISS 2014)*, March 2014.
[18] A. Winkelbauer, S. Farthofer, and G. Matz, “The rate-information tradeoff for Gaussian vector channels,” in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT 2014)*, 2014.