Finite Blocklength Quantized Feedback in Multiuser Channels

Osama Mahmoud

Supervisors

Prof. Gerald Matz

Stefan Farthofer

Institute of Telecommunications

Vienna University of Technology, Vienna, Austria

Diploma Thesis

February 27, 2018

Abstract

Feedback is known to decrease the encoding effort and reduce the decoding error for the single-user channel and it can even enlarge the capacity and the achievable sum rate for the multiple access channel (MAC). Quantization is a major aspect in modern communication systems as it is responsible for the conversion from the analog to the digital domain and it is often modeled as lossy source coding.

We focus on the limited blocklength, as it is very important in the communication nowadays due to complexity and delay constraints especially with the increasing demand of the real-time communication such as audio and video streaming for mobile devices.

We propose a quantized feedback scheme for the single-user AWGN channel and the two-user MAC with additive Gaussian noise. The quantized feedback link is modeled as an information bottleneck subject to a rate constraint.

Furthermore, we study the decoding error probability given a transmit rate and the maximum achievable rate of the quantized feedback scheme and compare it with other schemes like Schalkwjik-Kailath and Polyanskiy. Additionally, we study in detail in what blocklength regions the quantized feedback is beneficial.

Acknowledgments

First, I would like to thank my advisors Prof. Gerald Matz and Stefan Farthofer for their great help and support since the beginning of the thesis. Their doors were always open whenever I needed help or had a question about my research or writing.

I also want to thank all the professors and assistants at the Institute of Telecommunications at Vienna University of Technology for assisting and guiding me a lot during my whole journey in the master studies. I like to thank all my friends and colleagues for all their support, encouragement and standing by my side since I arrived here in Austria and started my Master.

Last but not least, I want to extend my gratitude to my family: my parents, my brother and my sister for believing in me and for supporting me spiritually throughout the whole Master and my life in general.

Contents

List of Figures

Statement of Originality

This is to certify that the work reported in this thesis is my own, and the work done by other authors is appropriately cited.

Osama Mahmoud

Chapter 1

Introduction

1.1 Motivation

In recent years, the demand for real-time data communication such as audio and video streaming is hugely increased. This data communication requires hard delay constraints on its delivery. This leads us to consider a limited blocklength coding in the schemes discussed here. Noiseless feedback in communication systems is known to reduce the coding effort and improves the decoding over a Gaussian channel. It is also known to enlarge the channel capacity and increases the achievable sum rate for the multiple access channel [2].

Quantization is a very important process in modern communication as every digital receiver has to do some kind of quantization in order to do the conversion from analog to digital. So, we apply a quantization to the received signal before being fed back and this feedback link is modeled as an information bottleneck [3]. We will observe how useful this quantized feedback scheme is for a limited blocklength compared with other feedback schemes. In addition, we will examine in what blocklength region the quantized feedback scenario is beneficial and what quantization rates are needed for a certain blocklength.

1.2 Aim of The Work

Perfect feedback without quantization is unrealistic due to the need of the conversion from the analog domain to the digital domain in the modern communication. Thus, we introduce a quantized feedback scheme and study the decoding error probability and the achievable rate.
The main task is to study how valuable the quantized feedback scheme is for a limited blocklength and in what blocklength region is it beneficial compared to other schemes without feedback and with perfect feedback. We applied this scheme in the case of single-user Gaussian channels and 2-user Gaussian multiple access channels.

1.3 Previous Work

In [4] Shannon proved that the channel capacity of a memoryless noisy channel is not increased by noiseless feedback in the single user case. There are still some advantages of the presence of the noiseless feedback such as that it leads to a substantial reduction in the complexity of the coding and decoding required to achieve a given performance over a noisy channel. In 1966 Schalkwijk and Kailath [5] introduced a feedback scheme based on Robbins-Monro procedure [6] and illustrated simplifications when the noisy link is an additive Gaussian noise channel operated under average power constraint with no restrictions on the signal bandwidth. Schalkwjik and Kailath extended their work from the wideband regime to the band-limited channel with signal bandwidth constraint in [7]. In [8] R. G. Gallager developed Schalkwijk and Kailath

scheme and introduced approximated expressions of the decoding error probability as it decreases as a second order exponent in the blocklength for rates below the channel capacity. Ozarow [9] showed that the channel capacity can be enlarged in the presence of feedback when two senders are considered in a multiple access scenario.
In 2009, Polyanskiy [10] developed finite blocklength fundamental limits of the achievable rate for the AWGN channel for a non-feedback system. As finite blocklength is considered here, the transmit rate cannot achieve the channel capacity as the capacity is only achievable when an infinite blocklength is considered. The key parameters of his approximation are the channel dispersion and the channel capacity.

1.4 Background

In this section we will explain important concepts that are necessary to be known before proceeding to the next chapters. [1]

1.4.1 Entropy and Mutual Information

1. Entropy is a measure of the uncertainty of a random variable, in other words, it is a measure of the amount of information that is required on average to describe a random variable. The entropy $H(X)$ of a discrete random variable X is defined by

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log p(x), \tag{1.1}$$

where $p(x)$ is the probability mass function.

2. Relative Entropy is a measure of the distance between two distributions

denoted by $D(p||q)$. The distance between the two probability mass functions $p(x)$ and $q(x)$ can be described by

$$D(p||q) = \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}. \tag{1.2}$$

The relative entropy is always non-negative, and it is equal to zero iff $p = q$.

3. Mutual Information is a measure of the amount of information that one random variable contains about the other random variable. Assume that X and Y are two random variables with joint probability mass function $p(x, y)$. The mutual information $I(X; Y)$ is defined by

$$I(X;Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} = D(p(x,y)||p(x)p(y)). \tag{1.3}$$

4. Differential Entropy is the entropy of a continuous random variable. Let X be a continuous random variable, the differential entropy $h(X)$ is defined as

$$h(x) = -\int_S f(x) \log f(x) dx, \tag{1.4}$$

where $f(x)$ is the probability density function for X and S is called the support set where $f(x) > 0$.

1.4.2 Channel Capacity

Channel capacity is described as the maximum amount of information that can be reliably transmitted over a channel. In the context of information theory,

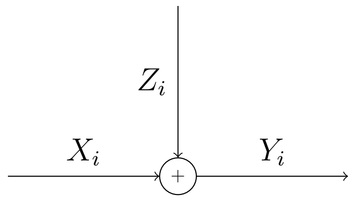


Figure 1.1: Gaussian Channel

it is known as the maximum mutual information. Consider a discrete channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ and the transition probability is given by $p(x|y)$. The information channel capacity is defined by

$$C = \max_{p(x)} I(X;Y). \tag{1.5}$$

1.4.3 Gaussian Channel

One of the most important continuous channels is the Gaussian channel as it is the basis for a large number of situations including radio and satellite links. Figure 1.1 illustrates a time-discrete channel with input X_i, output Y_i and noise $Z_i \sim \mathcal{N}(0, \sigma_z^2)$. The noise is Gaussian with variance σ_z^2 and it is assumed to be independent of X_i. Y_i can be described as follows:

$$Y_i = X_i + Z_i. \tag{1.6}$$

The capacity of the Gaussian channel with signal power P and noise variance σ_z^2 is defined by

$$C = \max_{E\{X^2\} \leq P} I(X;Y) = \frac{1}{2} \log(1 + \frac{P}{\sigma_z^2}). \tag{1.7}$$

And the average transmit power constraint is given by

$$\frac{1}{n} \sum_{i=1}^{n} E\{x_i^2\} \leq P, \tag{1.8}$$

for any transmitted codeword $(x_1, x_2, ..., x_n)$ transmitted over the channel.

1.4.4 Rate-Distortion Theory

To describe an arbitrary real number an infinite number of bits is required, so it is impossible to perfectly represent a continuous random variable with a finite number of bits. Therefore, a distortion measure which is a measure of the distance between the random variable and its representation is necessary. The rate-distortion theory is based on the minimum expected distortion achievable at a particular rate given a source distribution and a distortion measure. Or, equivalently, the minimum rate required to achieve a particular distortion. It can be applied to discrete and continuous random variables.

The rate distortion function for a source X with distribution $p(x)$ and distortion function $d(x,y)$ is [1]

$$R(D) = \min_{p(y|x):\sum_{(x,y)} p(x)p(y|x)d(x,y)\leq D} I(X;Y), \tag{1.9}$$

where the minimization is over all conditional distributions $p(y|x)$ and y is the estimate representation of x.
For a Gaussian source with zero mean and variance of σ^2, the rate distortion function can be described as [1]

$$R(D) = \frac{1}{2}\log\frac{\sigma^2}{D}. \tag{1.10}$$

Figure 1.2 illustrates the relation between the rate and the distortion for a Gaussian source.

1.5 Structure of the Work

In this thesis several coding schemes are discussed starting from Polyanskiy normal approximation in chapter 2 which introduces an approximation of the

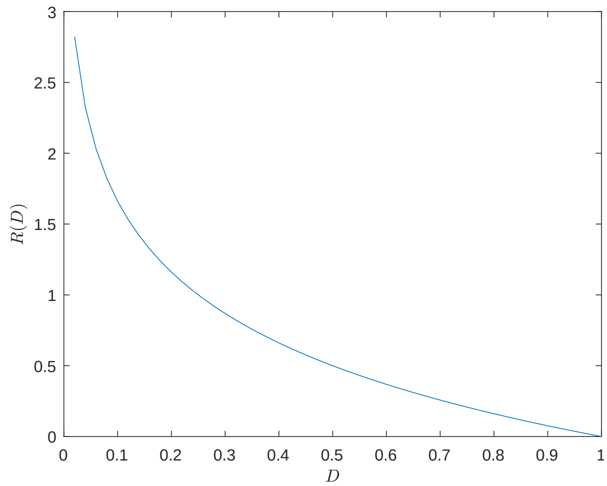


Figure 1.2: Rate-distortion function for Gaussian source

achievable rate and error probability of the direct transmission scenario over a Gaussian channel without feedback. Then in chapter 3, the perfect feedback schemes for single user and multiple users Gaussian channel are examined. In chapter 4, the quantized feedback scheme is introduced and it is compared with the perfect feedback and Polyanskiy schemes. Finally, a conclusion of the overall work is given by chapter 5.

Chapter 2

Polyanskiy Normal Approximation

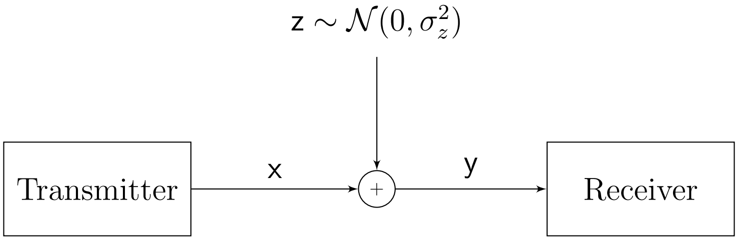


Figure 2.1: Transmission without feedback single-user AWGN channel

In this chapter we start with the very simple scenario which is the transmission without any feedback. So simply we have a transmitter, a Gaussian channel and a receiver as shown in Fig. 2.1. Consider an AWGN channel has an input vector x of length N and a Gaussian noise vector $\mathsf{z} \sim \mathcal{N}(0, \sigma_z^2)$. So the output vector y of length N is given as

$$\mathsf{y} = \mathsf{x} + \mathsf{z} \tag{2.1}$$

The transition probability density function is given by the conditional probability $P(\mathbf{y}|\mathbf{x}) : \mathbf{x} \to \mathbf{y}$. A codebook of M codewords $\{x_1, x_2, ..., x_M\} \subset \mathcal{X}$ are subject to the average power constraint of,

$$\frac{1}{N}\sum_{i=0}^{M} x_i^2 \leq P. \tag{2.2}$$

The minimum blocklength required to reach a given rate and an error probability can be approximately calculated according to Polyanskiy [10] from the channel capacity and the channel dispersion.

The performance limit for the channel in the finite blocklength regime is $M(N, P_e)$ which is the maximum cardinality of a codebook of blocklength N which can be decoded with block error probability not greater than P_e and channel capacity C. The following approximation is asymptotically tight for channel that satisfies the strong converse,

$$\frac{\log M(N, P_e)}{N} \approx C. \tag{2.3}$$

The converse bound is an upper bound on the size of any code with given arbitrary blocklength and error probability.

In practical terms, for many channels, error rates and blocklength ranges the approximation given by (2.3) is too optimistic as an infinite blocklength is needed to achieve the channel capacity. In [11] Polyanskiy, Verdu and Poor showed a much tighter approximation that can be obtained by introducing a second parameter which is the channel dispersion. The channel dispersion V (measured in squared information units per channel use) of a channel with capacity C is given by,

$$V = \lim_{P_e \to 0} \lim_{N \to \infty} \sup \frac{1}{N} \frac{(NC - \log M(N, P_e))^2}{2 \ln \frac{1}{P_e}}. \tag{2.4}$$

For a simple memoryless channels the following approximation is given by [11],

$$\log M(N, P_e) \approx NC - \sqrt{NV} Q^{-1}(P_e) + O(\log N). \tag{2.5}$$

2.1 Single User Channel

Consider a single user AWGN channel with SNR $\gamma = \frac{P}{\sigma_z^2}$ and error probability P_e under the power constraint in (2.2) we have [10]

$$\log M(N, P_e, \gamma) = NC(\gamma) - \sqrt{NV(\gamma)}Q^{-1}(P_e) + \rho_N, \tag{2.6}$$

where

$$\rho_N = O(\log N), \tag{2.7}$$

$$C(\gamma) = \frac{1}{2}\log(1+\gamma), \tag{2.8}$$

$$V(\gamma) = \frac{\gamma(\gamma+2)}{2(\gamma+1)^2}\log^2 e. \tag{2.9}$$

The $O(\log N)$ used in (2.6) is bounded by

$$O(1) \leq \rho_N \leq \frac{1}{2}\log N + O(1). \tag{2.10}$$

An approximation of (2.6) shown in [10] is given by,

$$\log M(N, P_e, \gamma) \approx NC - \sqrt{NV}Q^{-1}(P_e) + \alpha \log N, \tag{2.11}$$

for many practical scenarios for AWGN channels, $\alpha = 1/2$ is found a valid assumption. So we can simply rewrite equation (2.11) as

$$\log M(N, P_e, \gamma) \approx NC - \sqrt{NV}Q^{-1}(P_e) + \frac{1}{2}\log N. \tag{2.12}$$

We are interested in finding an expression of the error probability of this scheme. By using equations (2.3) and (2.12) we can find an expression of P_e with fixed rate as follows,

$$NR = NC - \sqrt{NV}Q^{-1}(P_e) + \frac{1}{2}\log N, \tag{2.13}$$

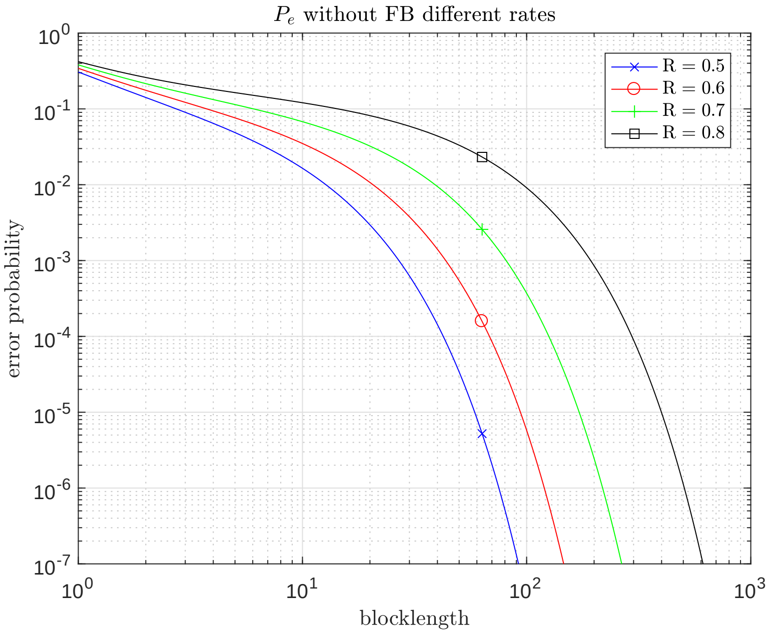


Figure 2.2: P_e without feedback with different rates for Gaussian channel. $C = 1$, $\sigma_z^2 = 1$

$$Q^{-1}(P_e) = \frac{C - R + \frac{1}{2N}\log N}{\sqrt{\frac{V}{N}}}, \tag{2.14}$$

$$P_e = Q\left(\frac{C - R + \frac{1}{2N}\log N}{\sqrt{\frac{V}{N}}}\right), \tag{2.15}$$

$$P_e = \frac{1}{2}\,\mathrm{erfc}\left(\frac{C - R + \frac{1}{2N}\log N}{\sqrt{\frac{2V}{N}}}\right), \tag{2.16}$$

where R here is the transmit rate. So now we found an expression of error probability in (2.16) by reformulating (2.11). Fig. 2.2 shows how the error probabilities are affected by increasing the blocklength while using different rates. It is clearly visible that by increasing blocklength the error probability is decreasing exponentially. Furthermore, we get a better performance when the transmit rate is smaller as we operate at a bigger distance of the ultimate limits.

2.2 Multiple User Channel

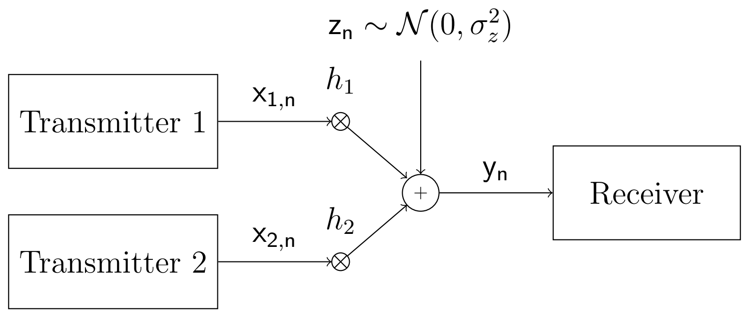


Figure 2.3: Transmission without feedback for 2-user Gaussian MAC

We now consider a symmetric 2-user Gaussian MAC with two senders without feedback as shown in the block diagram in figure 2.3. The transmitters send the independent Gaussian user signals x_1 and x_2 and the received signal is observed as follows

$$\mathsf{y} = h_1 \mathsf{x}_1 + h_2 \mathsf{x}_2 + \mathsf{z}, \tag{2.17}$$

where h_1 and h_2 are the channel gains and they are assumed to be known by the transmitters and the receiver and $\mathsf{z} \sim \mathcal{N}(0, \sigma_z^2)$ is the Gaussian channel noise. We have the average power constraint of

$$\frac{1}{N} \sum_{n=1}^{N} E\{x_{1,n}^2 + x_{2,n}^2\} \leq P_1 + P_2. \tag{2.18}$$

The sum-rate upper bound is given by [12]

$$R_1 + R_2 \leq C(h_1^2 P_1 + h_2^2 P_2) - \sqrt{\frac{V(h_1^2 P_1 + h_2^2 P_2)}{N}} Q^{-1}(P_e) + \frac{\alpha}{N} \log N, \quad (2.19)$$

Thus, we can find an error probability expression for this scheme by solving equation (2.19)

$$P_e = Q\left(\frac{C(h_1^2 P_1 + h_2^2 P_2) - R + \frac{1}{2N} \log N}{\sqrt{\frac{V(h_1^2 P_1 + h_2^2 P_2)}{N}}} \right). \quad (2.20)$$

The capacity and dispersion for the MAC here will be equal to the capacity and dispersion for the single user channel given by (2.8) and (2.9) respectively as this scheme is without feedback and the SNR of the channel will remain the same for the single and the multiple user cases. This will result in having the same error probability for both cases and we can observe that figure 2.2 represents the error probability without feedback for the single and the multiple user Gaussian channels.

Chapter 3

Perfect Feedback

Feedback is known to reduce the coding effort and improves the decoding by decreasing the error probability. For single user channels, feedback cannot increase the capacity of the channel but as we will see later the channel capacity can be enlarged for multiple access channels.
Schalkwijk and Kailath [5] introduced a perfect feedback scheme in 1966. The channel is considered additive noise Gaussian and the feedback is assumed to be noiseless, or in other words, we call it perfect feedback scenario. In 1984 Ozarow [9] presented a feedback scheme for a 2-user Gaussian multiple access channel and proofed that the capacity can be enlarged by such a system.

In 3.1 we will introduce the mathematical description of the Schalkwjik-Kailath (S-K) scheme. The S-K scheme is motivated by Robbins-Monro algorithm which is described in 3.2. Then in 3.3 the perfect feedback scheme is discussed when no bandwidth constraint is considered. In 3.4 we consider the feedback scheme in the band limited regime for single user and multiple user Gaussian channels. Finally, a comparison of the perfect feedback scheme and Polyanskiy scheme will be made in 3.5.

3.1 The Schalkwjik-Kailath Scheme

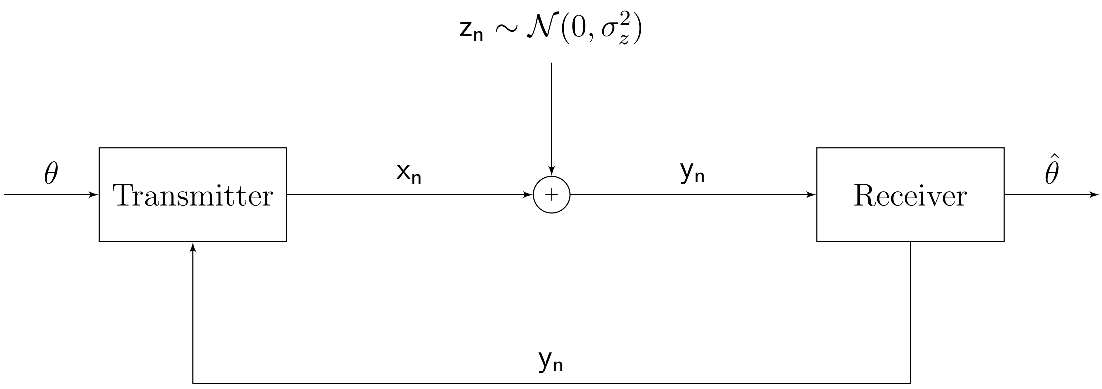


Figure 3.1: Transmission with perfect feedback for single user Gaussian channel

Schalkwjik-Kailath scheme is an iterative scheme based on determining the transmitted message θ from the received message $\hat{\theta}$ after performing N iterations. Figure 3.1 clarifies the block diagram of the transmission over an additive white Gaussian noise channel with noiseless feedback. The received signal is observed as

$$y_n = x_n + z_n, \tag{3.1}$$

where $x_n \in \mathbb{R}$ is the transmitted signal sent across the channel at each channel use $n = 1, 2, ..., N$. And $z_n \in \mathbb{R}$ is assumed to be i.i.d Gaussian noise such that $z_n \sim \mathcal{N}(0, \sigma_z^2)$. The received signal can be written as

$$\mathbf{y} = \mathbf{x} + \mathbf{z}, \tag{3.2}$$

where the input vector $\mathbf{x} = [x_1, x_2, ..., x_N]^T$ and N is the blocklength here. The average power constraints of the transmitted signal is bounded by

$$\frac{1}{N} E\{\mathbf{x}^T \mathbf{x}\} \leq P. \tag{3.3}$$

The transmitter output $\mathbf{x}$ is given by [13]

$$\mathbf{x} = \mathbf{F}\mathbf{z} + \mathbf{g}\theta, \tag{3.4}$$

where $\mathbf{g} \in \mathbb{R}^N$ is a unit vector,

$$\mathbf{g} = [1, 0, ..., 0]^T, \tag{3.5}$$

and the encoding matrix $\mathbf{F} \in \mathbb{R}^{N*N}$ is given as

$$\mathbf{F} = \left\{ \begin{matrix} 0 & 0 & & & \cdots & & 0 \\ -r & 0 & & & & & \\ \frac{-r}{\alpha} & \frac{-r^2}{\alpha} & 0 & & & & \\ \frac{-r}{\alpha^2} & \frac{-r^2}{\alpha^2} & \frac{-r^2}{\alpha} & 0 & & & \vdots \\ \frac{-r}{\alpha^3} & \frac{-r^2}{\alpha^3} & \frac{-r^2}{\alpha^2} & \frac{-r^2}{\alpha} & 0 & & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \\ \frac{-r}{\alpha^{N-2}} & \frac{-r^2}{\alpha^{N-2}} & \frac{-r^2}{\alpha^{N-3}} & & \cdots & \frac{-r}{\alpha} & 0 \end{matrix} \right\}, \tag{3.6}$$

where $\alpha^2 = 1 + P$ and $r = \sqrt{P}$. From (3.4) we observe that each x_n is a linear function of past values of z_n and the message θ.

Now, consider the decoding process at the receiver. The input of the receiver $\mathbf{y}$ is given by (3.2) and can be reformulated using (3.4)

$$\mathbf{y} = \mathbf{F}\mathbf{z} + \mathbf{g}\theta + \mathbf{z} = (\mathbf{I} + \mathbf{F})\mathbf{z} + \mathbf{g}\theta. \tag{3.7}$$

After all the N transmissions, the receiver estimates the original message by combining all the received values as a linear combination. The received message is denoted by $\hat{\theta}$ and is given as

$$\hat{\theta} = \mathbf{q}^T \mathbf{y}, \tag{3.8}$$

where the vector $\mathbf{q} \in \mathbb{R}^N$ is called the combining vector and it is given by

$$\mathbf{q} = [1, \frac{r}{\alpha^2}, \frac{r}{\alpha^3}, ..., \frac{r}{\alpha^N}]^T, \tag{3.9}$$

where $\alpha^2 = 1 + P$ and $r = \sqrt{P}$

3.2 Robbins Monro Procedure

The Schalkwjik-Kailath coding scheme is motivated by Robbins-Monro procedure [6]. Robbins-Monro procedure is a recursive scheme based on determining the transmitted signal from the received signal. The idea is to determine the message θ which is a zero of the function $f(x)$ without knowing the shape of this function. So we receive a noisy version of the transmitted signal x

$$y = f(x) + z, \tag{3.10}$$

where $z \sim \mathcal{N}(0, \sigma_z^2)$ is an additive white Gaussian noise.
To estimate θ we should start with an initial guess $\hat{x}_1$ and make successive guesses according to the following equation,

$$\hat{x}_{n+1} = \hat{x}_n - a_n y_n \quad n = 1, 2, ... \quad , \tag{3.11}$$

where the coefficient $a_n = 1/\alpha n$ according to Sacks' theorem [14].
Assume that the transmitter sends M messages to the receiver and a noiseless feedback link is available. Pick the message point θ which is the midpoint of the message interval. Through this message θ, draw the function $f(\hat{x}_1)$ which is a straight line and can be written as follows

$$f(\hat{x}_1) = \alpha(\hat{x}_1 - \theta), \tag{3.12}$$

with the slope $\alpha > 0$. Start with an initial guess of $\hat{x}_1 = 0.5$ and send $f(\hat{x}_1)$ given by (3.12) to the receiver. The receiver obtains the the following signal

$$y_1 = \alpha(\hat{x}_1 - \theta) + z_1. \tag{3.13}$$

Then the receiver will compute $\hat{x}_2$ according to equation (3.11), $\hat{x}_2 = \hat{x}_1 - ay_1$, and it will retransmit this value to the transmitter. Afterwards, the transmitter sends $f(\hat{x}_2) = \alpha(\hat{x}_2 - \theta)$ and the receiver will receive y_2 and so on. In general, the receiver receives

$$y_n = f(\hat{x}_n) + z_n, \tag{3.14}$$

and computes

$$\hat{x}_{n+1} = \hat{x}_n - \frac{a}{n} y_n. \tag{3.15}$$

Next, $\hat{x}_{n+1}$ is sent back to the transmitter which will send

$$f(\hat{x}_{n+1}) = \alpha(\hat{x}_{n+1} - \theta). \tag{3.16}$$

By solving equations (3.14) and (3.15) we will get

$$\hat{x}_{n+1} = \theta - \frac{1}{\alpha n} \sum_{i=1}^{n} z_i, \tag{3.17}$$

where $z_i \sim \mathcal{N}(0, \sigma_z^2)$ are independent i.i.d additive Gaussian noise and $\hat{x}_{n+1}$ is also Gaussian with mean θ and variance of $\sigma_z^2/\alpha^2 n$.
Suppose that the transmitter has to send one of M possible messages and N iterations are made before the receiver makes its decision. The length of the message interval is $1/M$ and the probability of $x_{N+1} \sim \mathcal{N}(\theta, \sigma_z^2/\alpha^2 N)$ is not located in this interval (the probability of error) is

$$P_e = 2\,\mathrm{erfc}\left(\frac{\frac{1}{2} M^{-1}}{\frac{\sigma_z}{\alpha} \sqrt{N}} \right). \tag{3.18}$$

3.3 Perfect Feedback- No Bandwidth Constraint

In [5] Schalkwijk and Kailath introduced a perfect feedback scheme with no bandwidth limitations. The channel is assumed to be additive noise Gaussian with no bandwidth constraint. Such channels are considered e.g. in space communications. At this point, the feedback link is assumed to be noiseless as the power from the ground to the satellite is much larger than the reverse power from the satellite to the ground. Thus, the first link can be considered to be approximately noiseless.
The channel is assumed to be white Gaussian with power spectral density $N_0/2$ and the average power of the transmitted signal is $P_{average}$ with no constraints imposed on the bandwidth. The channel capacity for this case is given by [5]

$$C = \frac{P_{average}}{N_0 \ln 2} \quad bits/second. \tag{3.19}$$

Equation (3.18) shows that the probability of error P_e that can be driven to zero by increasing N. The signaling rate can be described as follows

$$R = \log \frac{M}{T} \quad bits/second, \tag{3.20}$$

where T is the time interval. If the number of iterations N increased without increasing M the signaling rate will go to zero. A constant rate R can be maintained if M is increased along with T which is monotonically related to N. So the question here is, how rapidly can we increase M with N while still enabling the error probability to vanish for increasing N. Therefore, Schalkwijk-Kailath [5] showed the following relation between M and N

$$M(N) = N^{\frac{1}{2}(1-\epsilon)}. \tag{3.21}$$

Substituting by (3.21) in (3.18), the probability of error can then be written as

$$P_e = 2\,\mathrm{erfc}\left(\frac{\alpha}{2\sigma_z}N^{\epsilon/2}\right). \tag{3.22}$$

The critical rate is determined when $\epsilon = 0$ and it is given by

$$R_{critical} = \left(\frac{\log M(N)}{T}\right)_{\epsilon=0} = \frac{\log N}{2T} = A \quad bits/second, \tag{3.23}$$

where A is constant and it is limited by the average power constraint $P_{average}$. Therefore, A cannot be arbitrarily large and here it is approximately equal to the channel capacity [5]

$$C \approx A = \frac{P_{average}}{N_0 \ln 2}. \tag{3.24}$$

In order to keep the critical rate $R_{critical}$ finite as $T \to \infty$, N must grow exponentially with the time T, thus $N = e^{2AT}$.

The value of the slope α that minimizes the error probability in (3.22) can be determined by equivalently maximizing the square of the argument in (3.22)

$$\frac{\alpha^2}{4\sigma^2}N^{\epsilon}. \tag{3.25}$$

The optimum value of α^2 that maximizes (3.25) can be determined by differentiating (3.25) with respect to α^2 and yields [5]

$$\alpha_0^2 = 6N_0\left(\frac{C}{R}\right). \tag{3.26}$$

With a transmit rate of,

$$R = (1-\epsilon)A, \tag{3.27}$$

for which [5]

$$\epsilon = 1 - \frac{R}{C}\left(\frac{\epsilon^2}{6N_0} + \sum_{i=1}^{N-1}\frac{1}{i}\right)\frac{1}{\ln N}. \tag{3.28}$$

Now, α_0^2 can be substituted in the error probability equation (3.22) and as a result, we will get the following equation

$$P_e = 2\,\mathrm{erfc}\left(\sqrt{3\frac{C}{R}N^{\epsilon}}\right). \tag{3.29}$$

The asymptotic expression of the error probability is similar to the error probability given by (3.22) and by substituting $\sigma_z^2 = N_0/2$, the probability of error is

$$P_e = 2\,\mathrm{erfc}\left(\sqrt{\frac{\alpha^2}{2N_0}N^{\epsilon}}\right). \tag{3.30}$$

Substituting by the optimum α_0^2 given by (3.28) and by using the asymptotic formula of the complementary error function (erfc) shown below

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^{\infty} e^{-t^2}dt, \tag{3.31}$$

we can obtain the following expression for the error probability

$$P_e = \frac{\exp\left(-\frac{3}{2}\frac{C}{R}N^{\epsilon}\right)}{\sqrt{6\pi\left(\frac{C}{R}N^{\epsilon}\right)}}, \tag{3.32}$$

where, $\epsilon = 1 - (1 + \frac{R}{C}\sum_{i=1}^{N-1}\frac{1}{i})$.

Fig. 3.2 illustrates the relation between probability of error and the block length given a fixed capacity and different rates. The capacity is normalized to 1 and the noise power $\sigma^2 = 1$.

As seen from Fig. 3.2 the probability of error is decreasing exponentially with increasing blocklength.

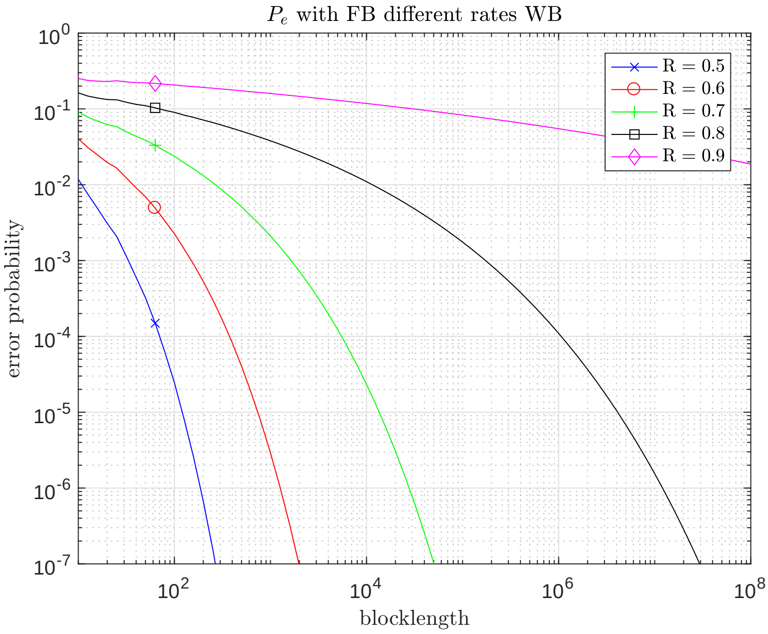


Figure 3.2: P_e for single user Gaussian channel with perfect feedback using different rates in the wide-band regime. $C = 1$, $\sigma_z^2 = 1$

3.4 Perfect Feedback- Band-limited

The previous scheme was in the wide-band regime. Now, we will consider band-limited channel [7] with noiseless feedback. The signal bandwidth is restricted to $(-W, W)$. Accordingly, we have a fixed bandwidth of W which the transmission is not supposed to exceed. The channel capacity is no longer $P_{average}/N_0 \ln 2$ as before, but it will be given by [7]

$$C = W \log(1 + \frac{P_{average}}{N_0 W}) \quad bits/second. \tag{3.33}$$

In order to achieve a constant rate in the case of wide-band, one had to choose $N = e^{2AT}$, which means that the number of transmission had to increase exponentially with time (as discussed before in 3.3). Now we have to meet the bandwidth constraint W. In this case the number of transmissions can only increase linearly with time according to the following equation

$$N = 2WT, \tag{3.34}$$

as the highest number of transmission is approximately equal to $2W$.

3.4.1 Single User Channel

In the beginning, the perfect feedback scheme will be applied to the single user scenario represented by the block diagram in figure 3.1. The received signal $y_n = x_n + z_n$, where z_n is considered additive Gaussian noise with zero mean and variance of σ_z^2. Robert Galleger [8] introduced an approximation of the error probability P_e

$$P_e = 2Q\Bigg(\exp\big(N(C-R)\big)\Bigg), \tag{3.35}$$

where the channel capacity $C = \frac{1}{2}log(1+\gamma)$ and γ is the signal to noise ratio (SNR).

Fig. 3.3 illustrates the probabilities of error for the single user perfect feedback scenario given by (3.35) when using different transmit rates. The channel capacity C is normalized to 1 $bit/second$ and the noise variance $\sigma_z^2 = 1$. As shown the error probability is decreasing doubly exponentially with the blocklength.

3.4.2 Multiple User Channel

The work is extended now to multiple user channel. Therefore, a Gaussian multiple access channel will be considered here. Feedback can enlarge the capacity of the multiple access channels as proven by Gaarder and Wolf [2].

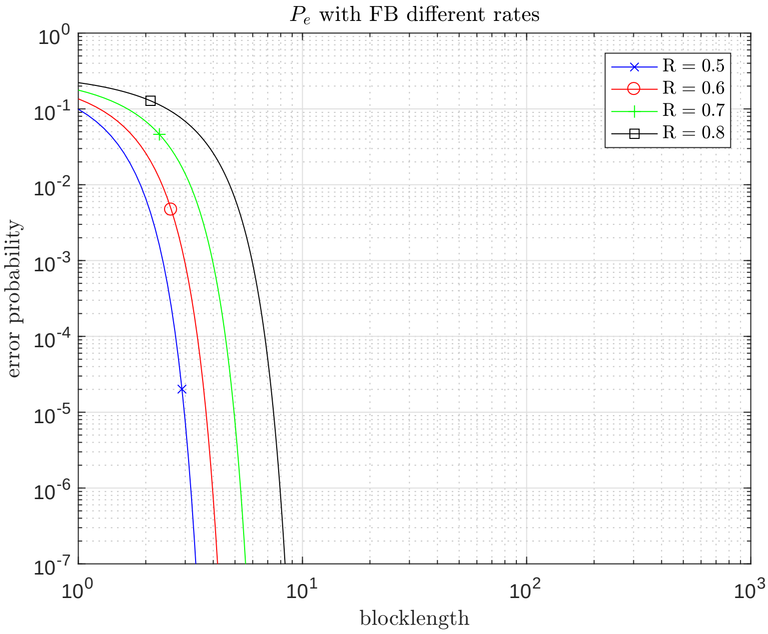


Figure 3.3: P_e with perfect feedback single-user Gaussian channel, $C = 1$, $\sigma_z^2 = 1$

Figure 3.4 illustrates the block diagram of the perfect feedback scheme powered by Ozarow [9] when a Gaussian MAC with two senders is considered. The two transmitters are communicating with a common receiver using the same channel which is considered AWGN with noise $\mathbf{z} \sim \mathcal{N}(0, \sigma_z^2)$. The channel output y is related to the input pair (x_1, x_2) by the conditional probability $p(y|x_1 x_2)$.

Consider a continuous amplitude version of the MAC where the n^{th} output is given by

$$y_n = h_1 x_{1,n} + h_2 x_{2,n} + z_n, \tag{3.36}$$

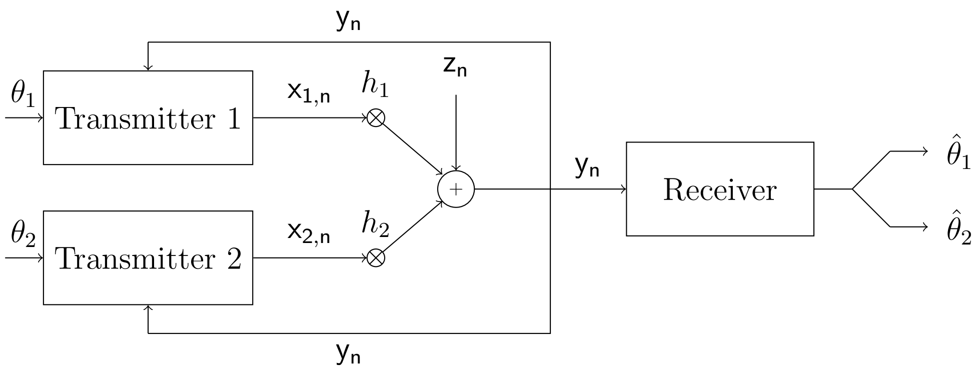


Figure 3.4: Perfect feedback for 2-User Gaussian MAC

where the z_n are a sequence of i.i.d zero mean Gaussian noise with variance σ_z^2. The average transmit power constraint with blocklength N is

$$\frac{1}{N}\sum_{n=1}^{N} E\{x_{1,n}^2 + x_{2,n}^2\} \leq P. \tag{3.37}$$

The capacity region without feedback for a Gaussian MAC given rate pair (R_1, R_2) with transmit powers P_1 and P_2 and channel gains h_1 and h_2 is given by [15][16]

$$\begin{aligned} C = \Big\{(R_1, R_2) : 0 \leq R_1 \leq \frac{1}{2}\log\big(1 + \frac{h_1^2 P_1}{\sigma_z^2}\big), \\ 0 \leq R_2 \leq \frac{1}{2}\log\big(1 + \frac{h_2^2 P_2}{\sigma_z^2}\big), \\ 0 \leq R_1 + R_2 \leq \frac{1}{2}\log\big(1 + \frac{h_1^2 P_1 + h_2^2 P_2}{\sigma_z^2}\big)\Big\}. \end{aligned} \tag{3.38}$$

Ozarow [9] showed that the capacity region of the system when the feedback is applied can be enlarged to

$$C_{fb} = \Big\{(R_1, R_2) : 0 \le R_1 \le \frac{1}{2}\log\big(1 + \frac{h_1^2 P_1}{\sigma_z^2}(1-\rho^2)\big),$$
$$0 \le R_2 \le \frac{1}{2}\log\big(1 + \frac{h_2^2 P_2}{\sigma_z^2}(1-\rho^2)\big), \tag{3.39}$$
$$0 \le R_1 + R_2 \le \frac{1}{2}\log\big(1 + \frac{h_1^2 P_1 + h_2^2 P_2 + 2\sqrt{h_1^2 P_1 h_2^2 P_2}\rho}{\sigma_z^2}\big)\Big\},$$

where $0 \le \rho \le 1$.

The correlation coefficient ρ is the crucial part in (3.39) as the achievable sum capacity can be enlarged according to the value of ρ. If there is no correlation between the two transmit signals, $\rho = 0$, the feedback link will be useless and eq. (3.39) yields eq. (3.38).

For what follows, a symmetric Gaussian MAC is considered for simplicity. Hence, $P_1 = P_2$ and the noise variance σ_z^2 is normalized to 1. To find the optimum value of ρ (ρ^*) for a given channel, all three conditions in (3.39) must be satisfied. Simply, R_1 in the first condition is added to R_2 in the second condition and then it will be equalized with the third condition and thereby we get the optimum ρ^*. We can observe from Fig. 3.5 that the curve of the first condition added to the second condition in (3.39) is strictly decreasing while the curve of the third condition in the same equation is strictly increasing. The intersection point of the two curves leads to ρ^*. The left y-axis of figure 3.5 represents R_1 added to R_2 which are the first and the second conditions respectively: $0 \le R_1 + R_2 \le \log\big(1 + \frac{P_1}{\sigma^2}(1-\rho^2)\big)$ while the right y-axis represents the third condition $R_1 + R_2$. The x-axis shows the correlation coefficient ρ. The channel capacity here for the non-feedback system is normalized to 1 *bit/second.*

At the beginning of a block of N transmissions, each transmitter picks a message θ_i ($i = 1, 2$) from the message alphabet M_i. θ_i is uniformly distributed over M_i equally spaced values in $[-\frac{1}{2}, \frac{1}{2}]$ [9]. For large M_i, θ_i have a vari-

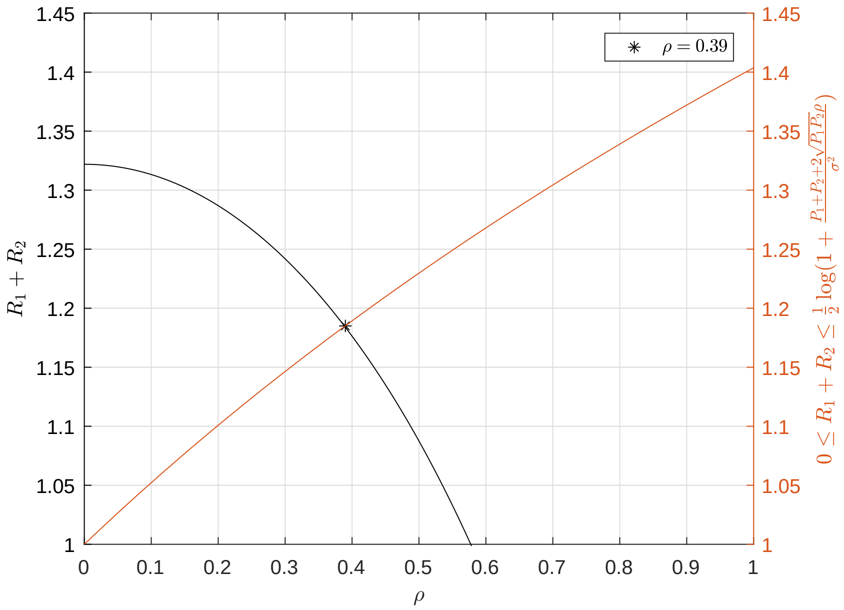


Figure 3.5: Finding the optimum correlation coefficient ρ^*

ance approximately equal to $\frac{1}{12}$. After the nth transmission ($n = 1, 2, ...$) the receiver estimates $\hat{\theta}_i^n$ and computes the estimation error

$$\epsilon_{i,n} = \hat{\theta}_i^n - \theta_i. \tag{3.40}$$

The estimated error $\epsilon_{i,n}$ is a zero mean Gaussian random variable with variance $\alpha_{i,n} = \sigma_z^2/12P_i$.

Now, after finding the optimum correlation coefficient ρ^*, Ozarow [9] achieved an expression of the error probability for a 2-user Gaussian MAC and it is given by

$$P_{e,i} = 2Q\left[\frac{\sigma_z^2}{2\sqrt{\alpha_{i,2}}(\sigma_z^2 + P_i(1-\rho^{*2}))}e^{N\left(C_{fb}-R_i\right)}\right]. \tag{3.41}$$

Substituting by the C_{fb} given by (3.39) in (3.41) we get

$$P_{e,i} = 2Q\left[\frac{\sigma_z^2}{2\sqrt{\alpha_{i,2}}(\sigma_z^2 + P_i(1-\rho^{*2}))}e^{N\left(\frac{1}{2}\log\left(1+\frac{P_1+P_2+2\sqrt{P_1P_2}\rho^*}{\sigma^2}\right)-R_i\right)}\right]. \tag{3.42}$$

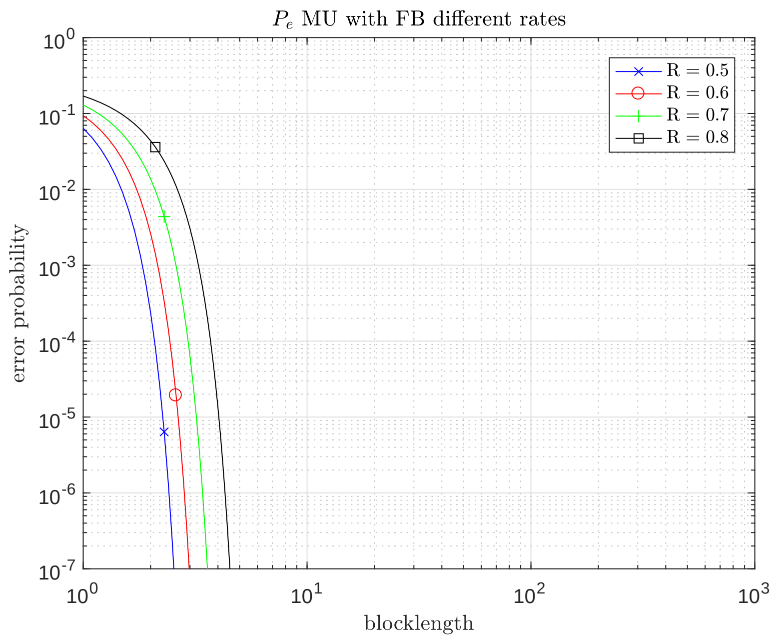


Figure 3.6: P_e perfect feedback for symmetric 2-user Gaussian MAC, $C_{fb} = 1.185$, $\sigma_z^2 = 1$

Fig. 3.6 shows the probability of error of the perfect feedback for symmetric 2-user Gaussian MAC given a various transmit rates. The channel capacity C of the non-feedback is always normalized to 1 and the feedback capacity for the MU system C_{fb} is calculated by Eq. (3.39). The channel is assumed to be symmetric thus, $P_1 = P_2 = 1.5$. Therefore, C_{fb} approximately equals to 1.185 *bits/second*.

3.5 Comparing the Perfect Feedback with Polyanskiy scheme

Next, we compare the perfect feedback scheme with Polyanskiy scheme discussed in the previous chapter. We will observe the advantage of feedback over the non-feedback system by showing the improvement in the error probability that feedback offers compared to Polyanskiy scheme. Figure. 3.7 compares the difference between the two schemes by using different transmit rates for a single user Gaussian channel. As shown the error probability of the transmission without feedback using Polyanskiy normal approximation is exponentially decreasing with increasing blocklength while the error probability of the perfect feedback is decreasing doubly exponentially with increasing blocklength. Therefore, the gap between the two curves is increased when the blocklength is increased. Figure 3.8 shows the difference between the error probability with perfect feedback and the error probability without feedback for single user Gaussian channel by calculating $\Delta P_{e\{norm\}} = P_{e\{Nofb\}}/P_{e\{fb\}}$.

Next, we compare the perfect feedback with polyanskiy scheme for 2-user Gaussian MAC. Figure 3.9 compares the two schemes with different transmit rates. As observed in the single user case, the performance of the perfect feedback scenario is much better than the performance of the non-feedback scenario. And figure 3.10 shows how the gap between the error probability curves is increasing by increasing blocklength.

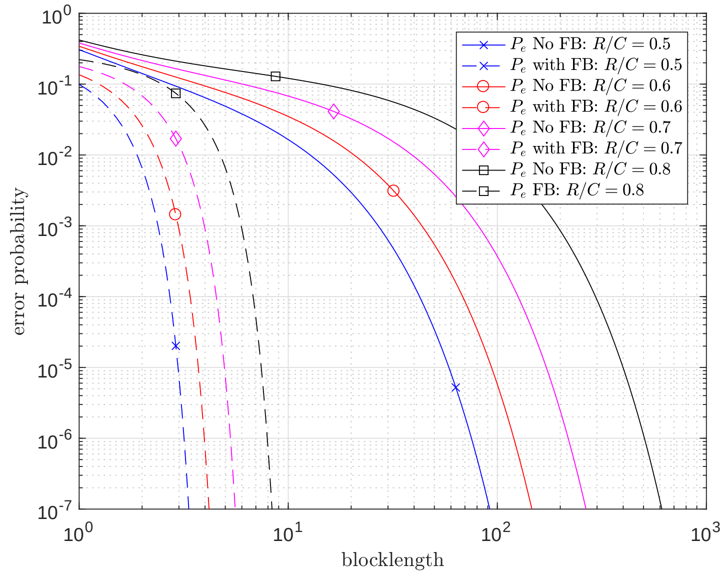


Figure 3.7: P_e perfect feedback Vs P_e no feedback for single user Gaussian channel. $C = 1$, $\sigma_z^2 = 1$

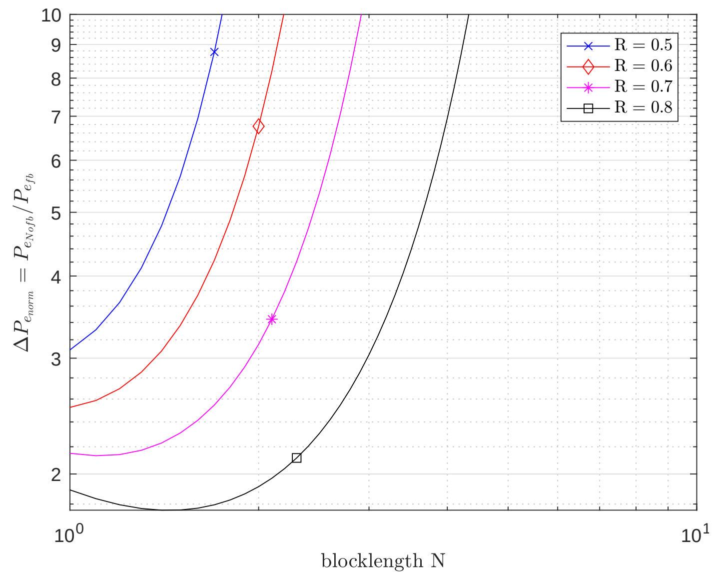


Figure 3.8: $\Delta P_{e\{norm\}}$ for single user Gaussian channel, $C = 1$, $\sigma_z^2 = 1$

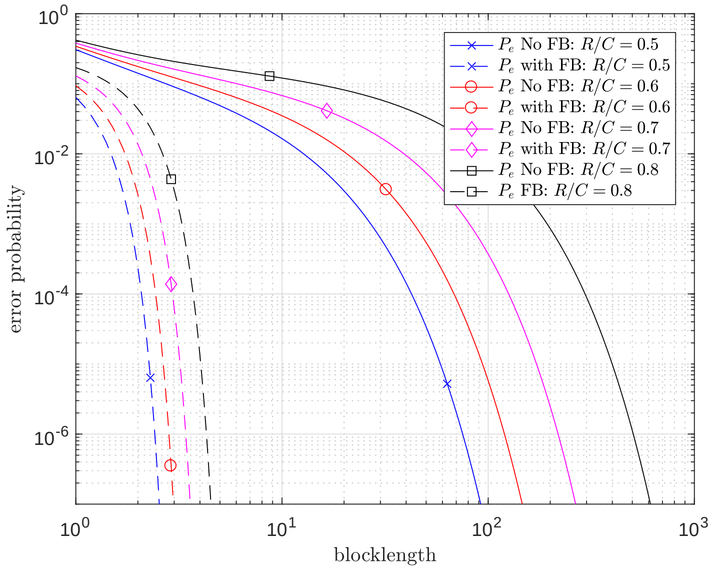


Figure 3.9: P_e perfect feedback Vs P_e no feedback for symmetric 2-user Gaussian MAC. $C_{fb} = 1.185$, $P_1 = P_2 = 1.5$, $\sigma_z^2 = 1$

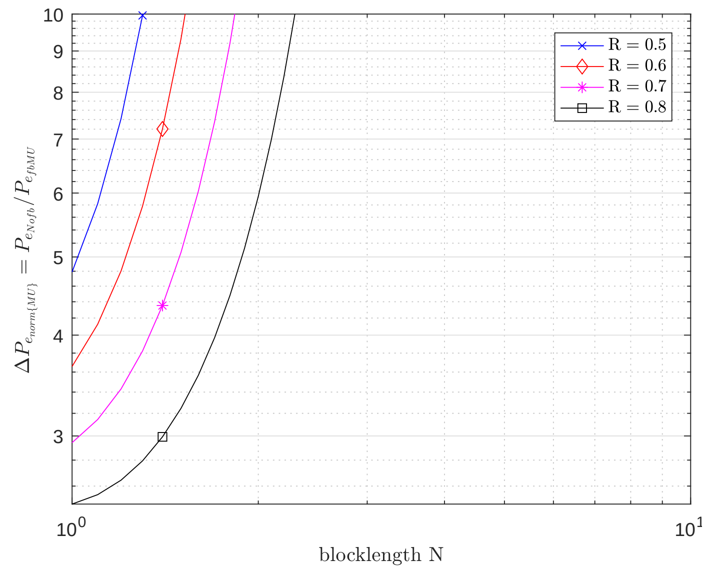


Figure 3.10: $\Delta P_{e\{norm\}}$ for symmetric 2-user Gaussian channel. $C_{fb} = 1.185$, $P_1 = P_2 = 1.5$, $\sigma_z^2 = 1$

Chapter 4

Quantized Feedback

In this chapter we will introduce the main part of the thesis which is the quantized feedback for finite blocklength.
The idea is to apply the quantization to the received signal before being fed back to the transmitter. The quantized feedback link is modeled as an information bottleneck [3].

In section 4.1 the Gaussian information bottleneck method is described. Then we will explain the quantized feedback for the single user Gaussian channel in the wide-band regime in 4.2. After that, the quantized feedback in the band-limited regime will be explained in 4.3 for the single user Gaussian channel and for the 2-user Gaussian MAC and it is compared with the perfect feedback and the non-feedback systems. Finally, we study the achievable rates for the quantized feedback, the perfect feedback and the non-feedback schemes in 4.4.

4.1 Gaussian Information Bottleneck

The information bottleneck method (IBM)[3] [17] was introduced as an information theoretic principle for extracting relevant information that an input random variable $\mathsf{x} \in \mathcal{X}$ contains about an output random variable $\mathsf{y} \in \mathcal{Y}$. The relevant information is defined as the mutual information $I(\mathsf{x};\mathsf{y})$ given the joint distribution $p(\mathsf{x};\mathsf{y})$, where x and y is assumed to be statistically dependent. In this case, y implicitly determines the relevant and the irrelevant features in x. The optimal representation of x would capture the relevant features and compress x by dismissing the irrelevant parts which do not contribute to the prediction of y.

Let $\mathbf{x} - \mathbf{y} - \hat{\mathbf{y}}$ be a Markov chain where $\hat{\mathbf{y}}$ is the quantized version of $\mathbf{y}$ and the joint distribution of $\mathbf{x}$ and $\mathbf{y}$ is assumed to be known. $\mathbf{x}$ and $\mathbf{y}$ are assumed to be jointly Gaussian random vectors with full rank covariance matrices. The information bottleneck method solves the variational problem

$$\min_{p(\hat{y}|y)} I(\mathbf{y};\hat{\mathbf{y}}) - \beta I(\mathbf{x};\hat{\mathbf{y}}), \tag{4.1}$$

where $\mathbf{x}$ is called the relevance variable, $I(\mathbf{x};\hat{\mathbf{y}})$ is the relevant information and $I(\mathbf{y};\hat{\mathbf{y}})$ is the compression rate. β determines the trade-off between the compression rate and the relevant information. It was shown in [17] that the optimal $\hat{\mathbf{y}}$ is jointly Gaussian with $\mathbf{y}$ and can be written as

$$\hat{\mathbf{y}} = \boldsymbol{A}\mathbf{y} + \boldsymbol{\xi}, \tag{4.2}$$

where $\boldsymbol{A}$ is a particular matrix and $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{C}_{\boldsymbol{\xi}})$ is independent on $\mathbf{y}$. Equation (4.1) can be rewritten using (4.2) as

$$\min_{\boldsymbol{A},\boldsymbol{C}_{\boldsymbol{\xi}}} I(\mathbf{y};\boldsymbol{A}\mathbf{y}+\boldsymbol{\xi}) - \beta I(\mathbf{x};\boldsymbol{A}\mathbf{y}+\boldsymbol{\xi}). \tag{4.3}$$

In [17] a solution of (4.3) has been found

$$\boldsymbol{A} = \begin{Bmatrix} [\mathbf{0}; ...; \mathbf{0}] & 0 \leq \beta \leq \beta_1^c \\ [\alpha_1 \boldsymbol{v_1^t}; \mathbf{0}; ...; \mathbf{0}] & \beta_1^c \leq \beta \leq \beta_2^c \\ [\alpha_1 \boldsymbol{v_1^t}; \alpha_2 \boldsymbol{v_2^t}; \mathbf{0}; ...; \mathbf{0}] & \beta_2^c \leq \beta \leq \beta_3^c \\ . & . \\ . & . \\ . & . \end{Bmatrix}, \tag{4.4}$$

where $\{\boldsymbol{v_1^t}, \boldsymbol{v_2^t}, ..., \boldsymbol{v_n^t}\}$ are the the left eigenvectors of $\boldsymbol{C}_{\mathbf{y}|\mathbf{x}} \boldsymbol{C}_{\mathbf{y}}^{-1}$ sorted by their corresponding eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$, $\beta_i^c = \frac{1}{1-\lambda_i}$ are critical β values and the coefficients α_i are given by

$$\alpha_i \widehat{=} \sqrt{\frac{\beta(1-\lambda_i)}{\lambda_i \boldsymbol{v_i^t} \boldsymbol{C}_{\mathbf{y}} \boldsymbol{v_i}}}. \tag{4.5}$$

By using (4.4) and (4.5) the rate information trade-off can be expressed by

$$I(\mathbf{x}; \hat{\mathbf{y}}) = I(\mathbf{y}; \hat{\mathbf{y}}) - \frac{1}{2} \sum_{i=1}^{n} \log(1 - \lambda_i). \tag{4.6}$$

By the data processing inequality, equation (4.6) is bounded by

$$I(\mathbf{x}; \hat{\mathbf{y}}) \leq \min\{I(\mathbf{y}; \hat{\mathbf{y}}), I(\mathbf{x}; \mathbf{y})\}. \tag{4.7}$$

In order to formalize the trade-off between the relevant information and the compression rate, the information-rate function $I(R_q)$ and the rate-information function $R_q(I)$ have to be defined where R_q is the quantization rate.

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

$$I(R_q) \triangleq \max_{p(\hat{y}|y)} I(\mathbf{x}; \hat{\mathbf{y}}) \quad \text{subject to} \quad I(\mathbf{y}; \hat{\mathbf{y}}) \leq R_q. \tag{4.8}$$

The rate-information function $R_q : [0, I(\mathbf{x}; \mathbf{y})] \to \mathbb{R}_+$ is defined as follows:

$$R_q(I) \triangleq \min_{p(\hat{y}|y)} I(\mathbf{y}; \hat{\mathbf{y}}) \quad \text{subject to} \quad I(\mathbf{x}; \hat{\mathbf{y}}) \geq I. \tag{4.9}$$

$I(R_q)$ allows us to quantify the maximum of the relevant information that can be preserved when the compression rate is at most R_q and $R_q(I)$ quantifies the minimum compression rate required when the relevant information is at least I.

The scalar Gaussian channel case is summarized in [8]. The received signal y is given as

$$\mathsf{y} = \mathsf{x} + \mathsf{z}, \tag{4.10}$$

where $\mathsf{z} \sim \mathcal{N}(0, \sigma^2)$ is independent of $\mathsf{x} \sim \mathcal{N}(0, P)$. The quantized version of y is $\hat{\mathsf{y}} = Q(\mathsf{y})$. Let $\mathsf{x} - \mathsf{y} - \hat{\mathsf{y}}$ be a Markov chain, the transition probability density function (pdf) of the overall channel is

$$p(\hat{y}|x) = \int_{-\infty}^{\infty} p(\hat{y}|y)p(y|x)dy, \tag{4.11}$$

where $p(y|x)$ is the transition pdf of the Gaussian channel and $p(\hat{y}|y)$ describes the probabilistic quantization mapping Q.

The capacity of the Gaussian channel $p(y|x)$ with average transmit power constraint P with no quantization is [18]

$$C(\gamma) \triangleq \frac{1}{2}\log(1+\gamma), \tag{4.12}$$

where γ here is the signal-to-noise ratio.

The rate-information function according to [8] is equal to:

$$I(R_q) = C(\gamma) - \frac{1}{2}\log(1 + 2^{-2R_q}\gamma). \tag{4.13}$$

So the rate-information function reaches the channel capacity when the quantization rate R_q goes to infinity.

Since x and y are jointly Gaussian, the overall channel $p(\hat{y}|x)$ is Gaussian too. From [8] and [19] we can write $I(R_q) = C(\gamma_q)$ and

$$\gamma_q = \gamma\frac{1 - 2^{-2R_q}}{1 + 2^{-2R_q}\gamma} \leq \gamma, \tag{4.14}$$

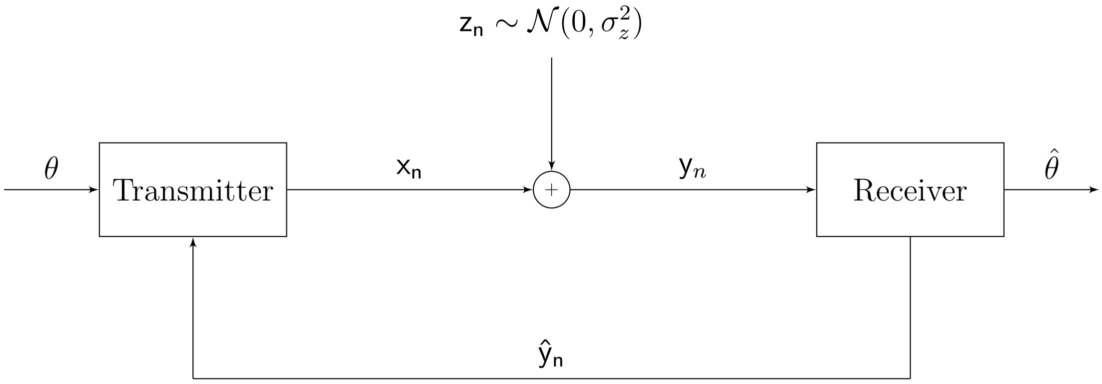


Figure 4.1: Quantized feedback single user AWGN channel

where γ_q is the SNR of the overall channel $p(\hat{y}|x)$ and R_q is the quantization rate. Therefore, we can model the optimal channel output quantization by an additive white Gaussian noise with variance

$$\sigma_q^2 = \sigma_z^2 \frac{1+\gamma}{2^{2R_q} - 1}. \tag{4.15}$$

4.2 Quantized Feedback- No Bandwidth Constraint

Schalkwjik and Kailath introduced a perfect feedback scheme in the wide-band regime [5] and we discussed it in the previous chapter. Now, we will apply the quantization scheme modeled by the information bottleneck method to the S-K scheme.

Figure 4.1 shows the quantized feedback scenario as the quantization is applied to the received signal y before being fed back to the transmitter. $\hat{\mathsf{y}}$ is the quantized version of the received signal y. The channel model is as follows

$$\mathsf{y_n} = \mathsf{x_n} + \mathsf{z_n}, \tag{4.16}$$

where $\mathsf{z_n}$ is an additive Gaussian noise with zero mean and variance of σ_z^2. We impose here the average transmit power constraint $P_{average}$.
The SNR of the overall system $p(\hat{y}|x)$ after applying the quantization to the feedback is given by (4.14). By substituting γ_q in the channel capacity equation given by (4.12) we get the new capacity of the overall channel

$$C_q(\gamma_q) = \frac{1}{2}\log(1 + \gamma_q). \tag{4.17}$$

We can now substitute by this capacity in the error probability given by eq. (3.32) so the probability of error expression for the quantized feedback system becomes as follows

$$P_{e_q} = \frac{\exp\left(-\frac{3}{2}\frac{\frac{1}{2}\log(1+\gamma_q)}{R}N^{\epsilon}\right)}{\sqrt{6\pi\left(\frac{\frac{1}{2}\log(1+\gamma_q)}{R}N^{\epsilon}\right)}}. \tag{4.18}$$

The comparison between the error probability with perfect feedback and the error probability with quantized feedback using quantization rates of 1,3 and 5 are shown in figures 4.2, 4.3 and 4.4 respectively. It is clearly illustrated by the figures that the performance is improved when the quantization rate is increased and the curves of the error probabilities with the same transmit rates become closer to each other. In the example, approximately the same performance is reached when quantization rate of 5 is used as seen in figure 4.4. This is simply because the quantizer introduces a quantization noise and this noise decreases when the quantization rate increases. We can also observe

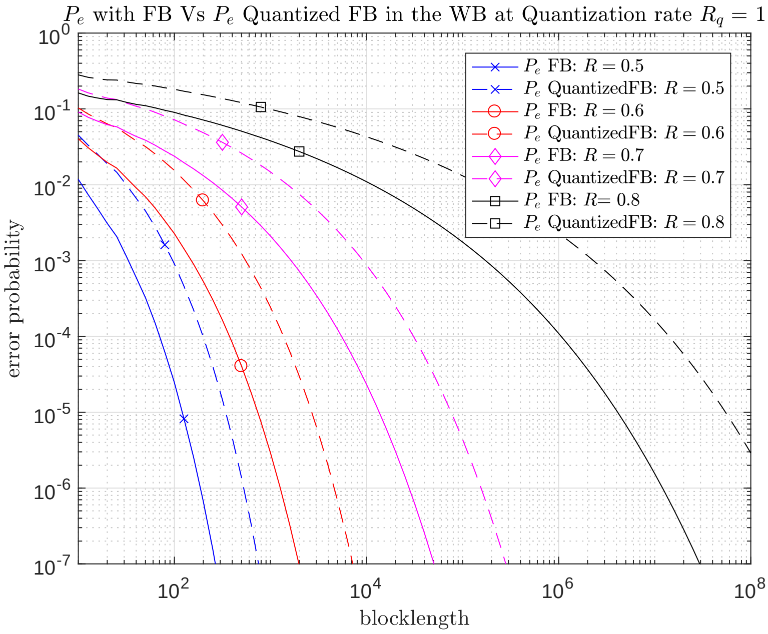


Figure 4.2: P_e perfect feedback Vs P_e quantized feedback at $R_q = 1$ for WB Gaussian Channel, $C = 1$, $\gamma = 3$, $C_q = 0.596$, $\gamma_q = 1.286$

that by increasing blocklength, the difference between the error probability of the quantized feedback P_{e_q} and the error probability of the perfect feedback P_e increases. So we can calculate the normalized difference by obtaining $\Delta P_{e\{norm\}}$

$$\Delta P_{e\{norm\}} = \frac{P_{e_q} - P_e}{P_e}. \tag{4.19}$$

Figures 4.5, 4.6 and 4.7 show how the gap between P_{e_q} and P_e is increasing with the blocklength when quantization rates of 1,3 and 5 are applied respectively. The capacity of the system without quantization is normalized to 1 *bit/second* and the noise variance σ_z^2 is normalized to 1 as well therefore the SNR would be equal to 3. The SNR of the quantized feedback γ_q given by (4.14) is decreased due to the quantization noise introduced by the quantizer. Therefore, the capacity of the overall channel in (4.17) is decreased as well.

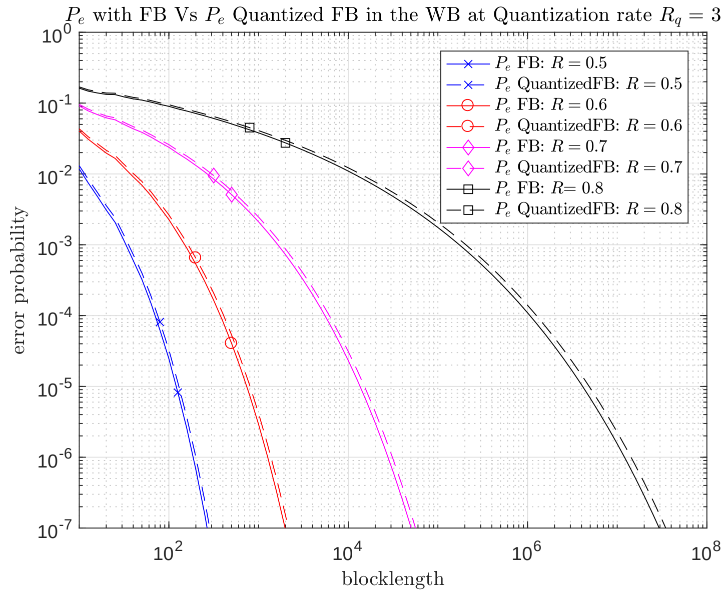


Figure 4.3: P_e perfect feedback Vs P_e quantized feedback at $R_q = 3$ for WB Gaussian Channel, $C = 1$, $\gamma = 3$, $C_q = 0.967$, $\gamma_q = 2.821$

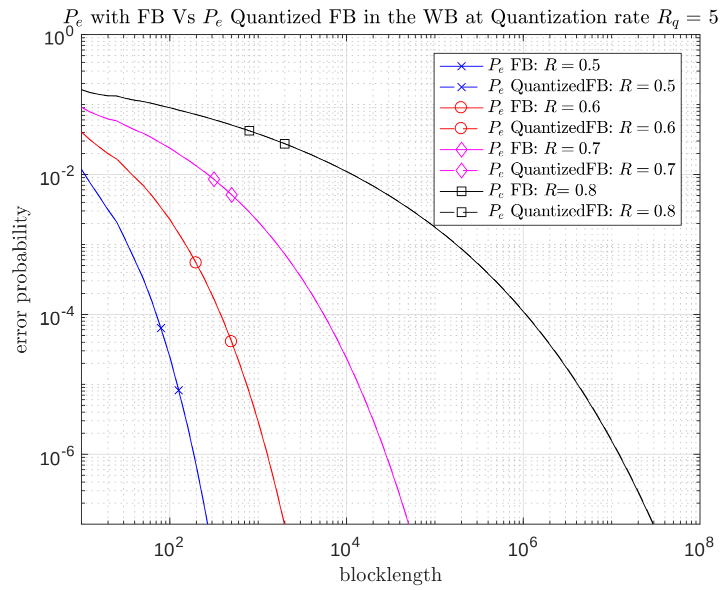


Figure 4.4: P_e perfect feedback Vs P_e quantized feedback at $R_q = 5$ for WB Gaussian Channel, $C = 1$, $\gamma = 3$, $C_q = 0.998$, $\gamma_q = 2.988$

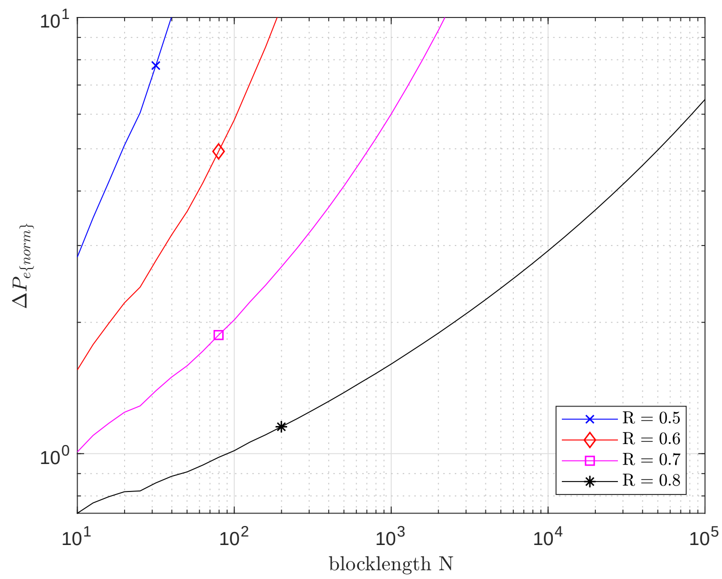


Figure 4.5: $\Delta P_{e\{norm\}}$ at $R_q = 1$ for WB Gaussian Channel, $C = 1$, $\gamma = 3$, $C_q = 0.596$, $\gamma_q = 1.286$

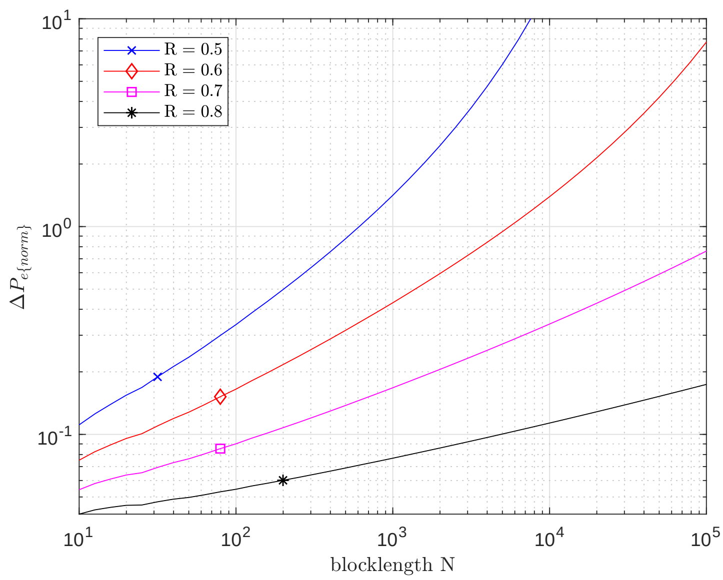


Figure 4.6: $\Delta P_{e\{norm\}}$ at $R_q = 3$ for WB Gaussian Channel, $C = 1$, $\gamma = 3$, $C_q = 0.967$, $\gamma_q = 2.821$

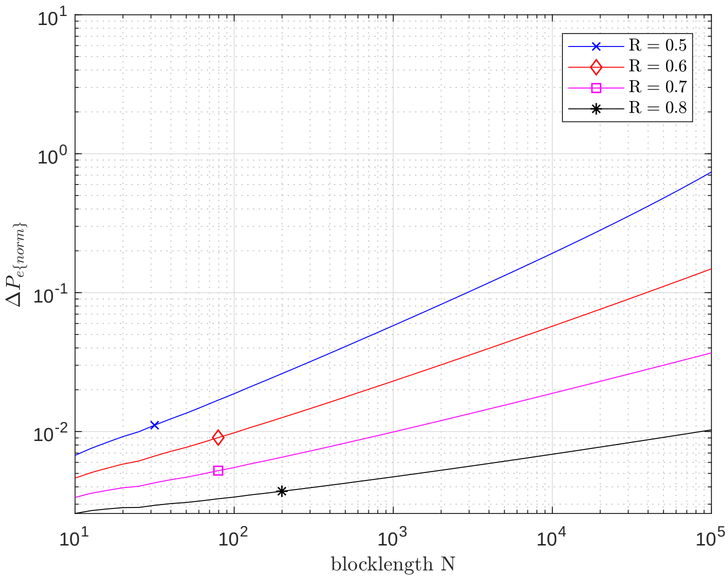


Figure 4.7: $\Delta P_{e\{norm\}}$ at $R_q = 5$ for WB Gaussian Channel, $C = 1$, $\gamma = 3$, $C_q = 0.998$, $\gamma_q = 2.988$

4.3 Quantized Feedback- Band Limited

Here, we will introduce the quantized feedback scheme for the single user and the multiple user Gaussian channel in the band-limited regime.

4.3.1 Single User Channel

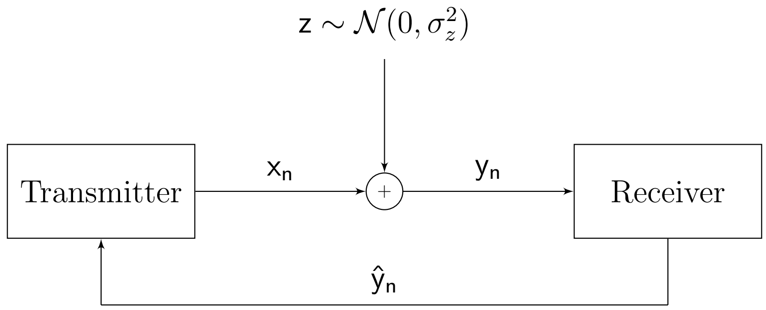


Figure 4.8: Quantized feedback single user AWGN channel

Figure 4.8 shows the block diagram of the quantized feedback system for single user channel. The channel is also considered additive white Gaussian with noise $\mathsf{z} \sim \mathcal{N}(0, \sigma_z^2)$ and the received signal $\mathsf{y} = \mathsf{x} + \mathsf{z}$.

The error probability given by equation (3.35) [8] will be reformulated by substituting the overall capacity C_q in equation (4.17) instead of the channel capacity C of the perfect feedback system. So the error probability equation will be as follows

$$P_{e_q} = 2Q\Bigg(\exp\big(N(C_q - R)\big)\Bigg), \tag{4.20}$$

$$P_{e_q} = 2Q\Bigg(\exp\left[N\left(\frac{1}{2}log(1+\gamma_q) - R\right)\right]\Bigg). \tag{4.21}$$

Figures 4.9, 4.10 and 4.11 show a comparison between the quantized and the perfect feedback schemes using quantization rates of 1, 3 and 5 respectively. In figure 4.9 a low quantization rate is used and the error probability of the quantized feedback is not decreasing anymore when high transmit rates are used (see the transmit rates of $R = 0.8$, $R = 0.7$ and $R = 0.6$ the black, pink and red dashed curves respectively). This happens when the transmit rates become larger than the reduced channel capacity due to the quantization error. But when the quantization rate is increased again as shown in figure 4.10 the error probability of the quantized feedback decreases by increasing blocklength. By increasing the quantization rate, the probability of error of the quantized feedback system becomes more closer to the error probability of the perfect feedback scheme and we reached almost the same performance when a high quantization rate is used as shown in figure 4.11.

In addition, the gap between the error probability of the quantized feedback and the perfect feedback increases by increasing blocklength and this is shown in figures 4.12, 4.13 and 4.14 for the quantization rates of 1, 3, and 5 respectively.

Figure 4.15 combines the probabilities of error without feedback, with perfect feedback and with quantized feedback using different quantization rates. There is a big gap between the blue solid curve and the blue dashed curve representing the error probability without feedback and with perfect feedback respectively. The error probabilities of the quantized feedback scheme are located between the two blue curves and since the quantization rate increases, we get a very close performance to the perfect feedback scenario.

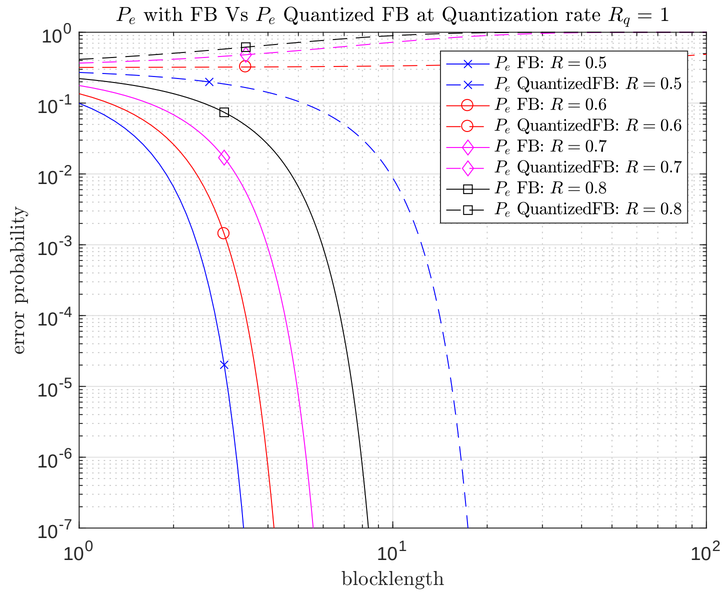


Figure 4.9: P_e perfect feedback Vs P_e quantized feedback at $R_q = 1$ for single user Gaussian channel, $C = 1$, $\gamma = 3$, $C_q = 0.596$, $\gamma_q = 1.286$

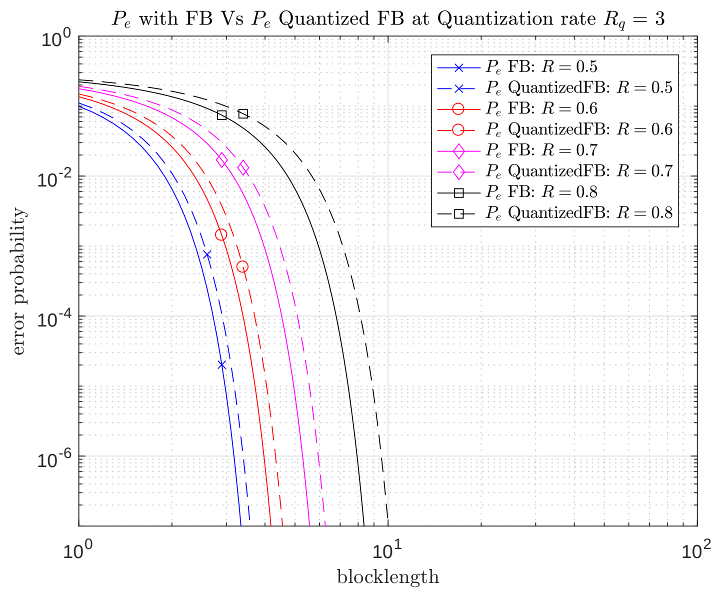


Figure 4.10: P_e perfect feedback Vs P_e quantized feedback at $R_q = 3$ for single user Gaussian channel, $C = 1$, $\gamma = 3$, $C_q = 0.967$, $\gamma_q = 2.821$

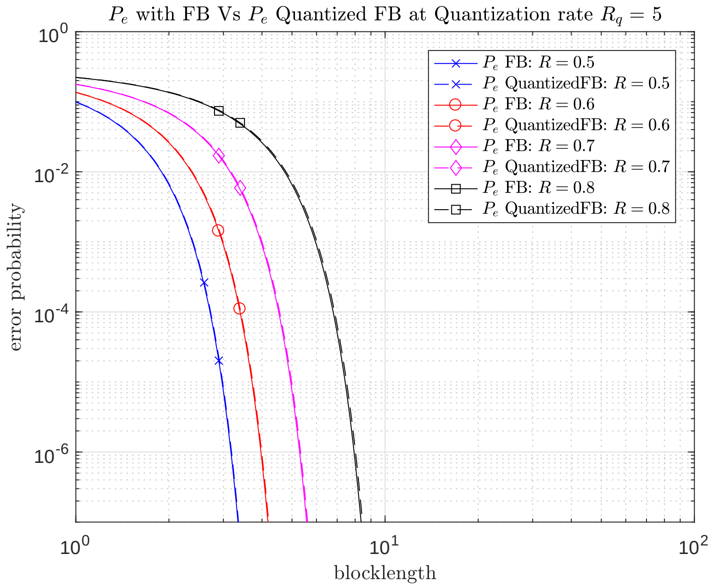


Figure 4.11: P_e perfect feedback Vs P_e quantized feedback at $R_q = 5$ for single user Gaussian channel, $C = 1$, $\gamma = 3$, $C_q = 0.998$, $\gamma_q = 2.988$

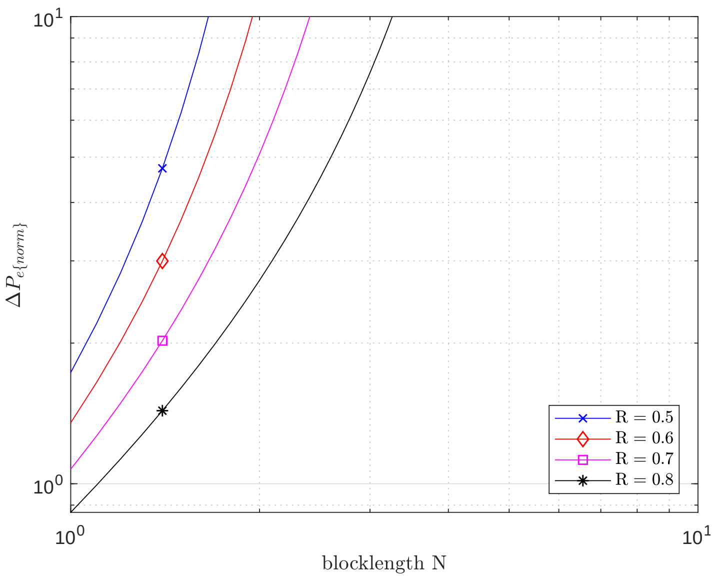


Figure 4.12: $\Delta P_{e\{norm\}}$ at $R_q = 1$ for single user Gaussian channel, $C = 1$, $\gamma = 3$, $C_q = 0.596$, $\gamma_q = 1.286$

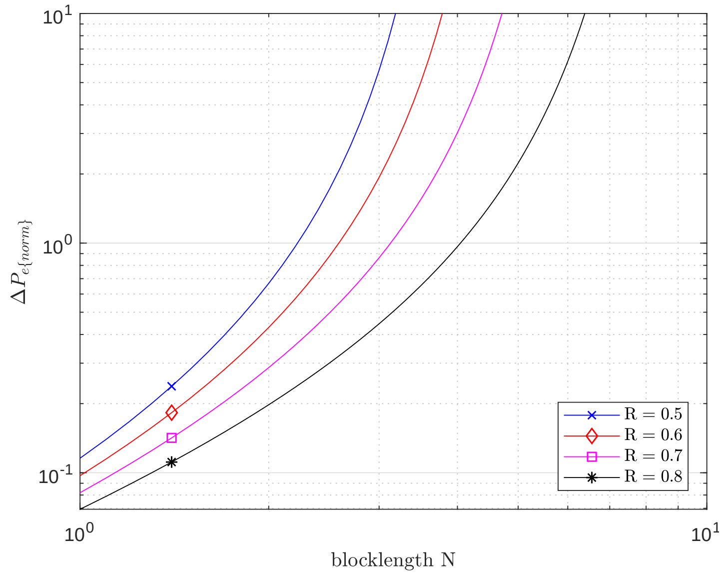


Figure 4.13: $\Delta P_{e\{norm\}}$ at $R_q = 3$ for single user Gaussian channel, $C = 1$, $\gamma = 3$, $C_q = 0.967$, $\gamma_q = 2.821$

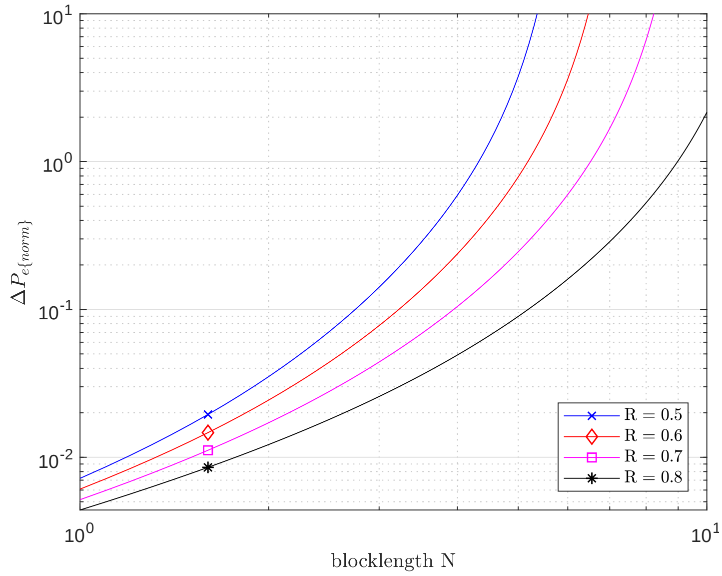


Figure 4.14: $\Delta P_{e\{norm\}}$ at $R_q = 5$ for single user Gaussian channel, $C = 1$, $\gamma = 3$, $C_q = 0.998$, $\gamma_q = 2.988$

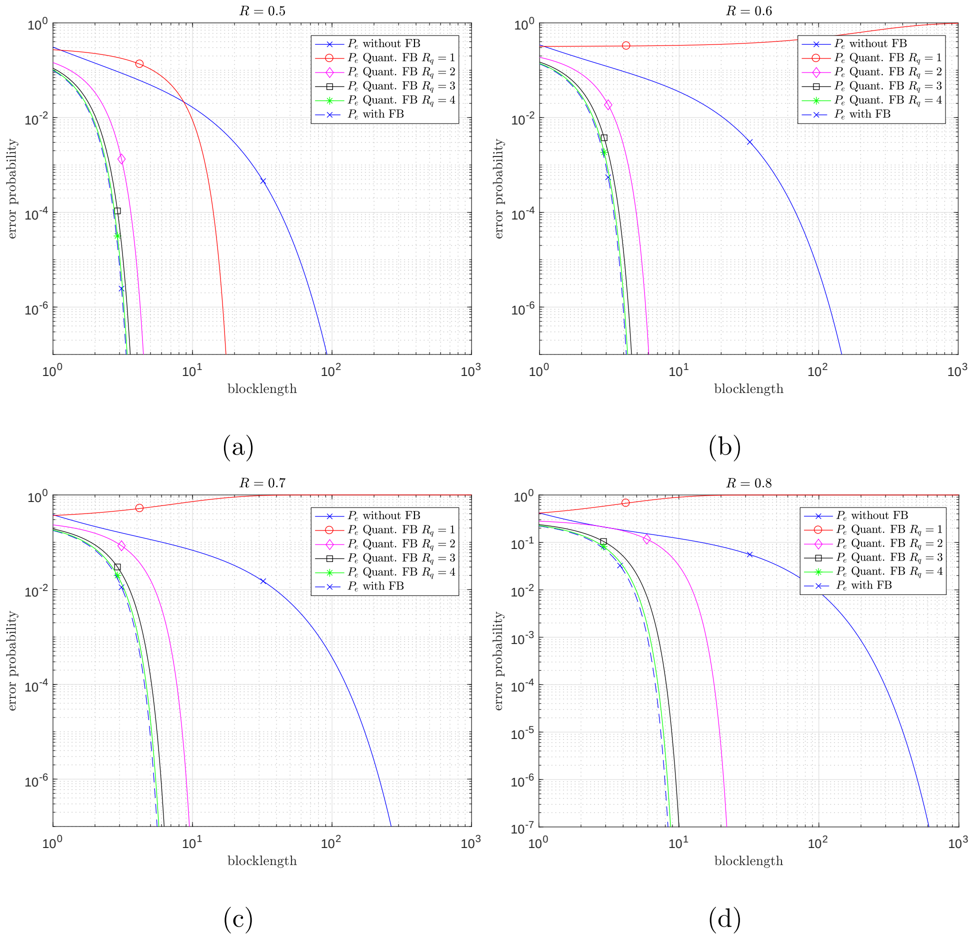
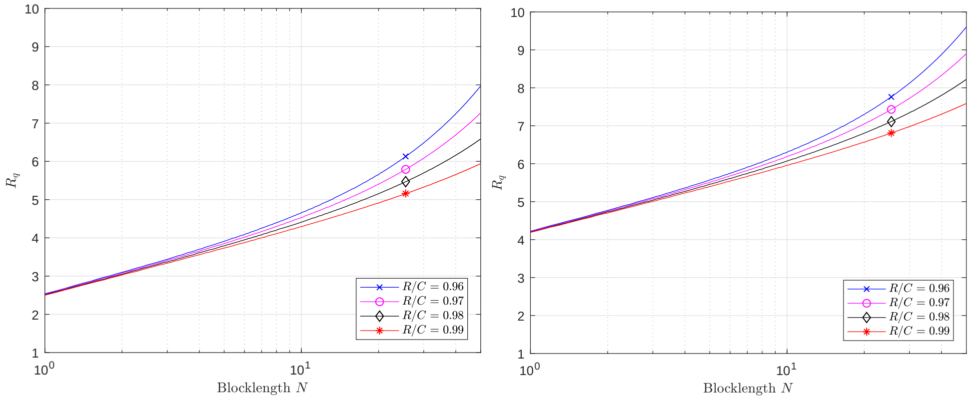


Figure 4.15: Comparing the error probability without feedback, with perfect feedback and with quantized feedback at different quantization rates, $C = 1$, $\sigma_z^2 = 1$. (a) $R = 0.5$. (b) $R = 0.6$. (c) $R = 0.7$. (d) $R = 0.8$.



(a) (b)

Figure 4.16: Quantization rate R_q Vs blocklength N for single user Gaussian channel with $C = 1$ and $\sigma_z^2 = 1$ using different transmit rates for fixed $\Delta P_{e\{norm\}}$. (a) $\Delta P_{e\{norm\}} = 10^{-1}$, (b) $\Delta P_{e\{norm\}} = 10^{-2}$

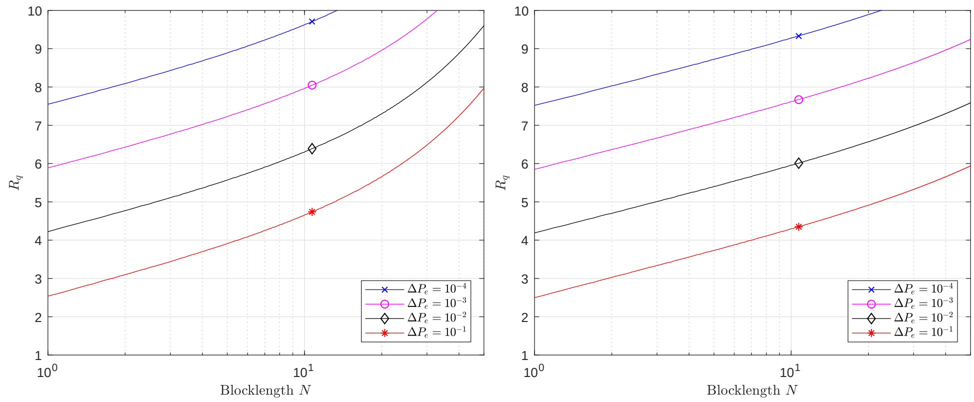
Figure 4.16 shows the quantization rate needed in a certain blocklength region for a specific $\Delta P_{e\{norm\}}$

$$\Delta P_{e\{norm\}} = \frac{P_{e_q} - P_e}{P_e}. \tag{4.22}$$

The previous results demonstrated that, for a higher blocklength we need a higher quantization rate in order to keep a specific difference between the error probability of the quantized feedback and the perfect feedback. We can also observe that a higher quantization rate is needed when a lower transmit rate is used as the $\Delta P_{e\{norm\}}$ is increasing faster than when higher transmit rates are used.

Figure 4.17 illustrates the quantization rate needed for a fixed transmit rate using different $\Delta P_{e\{norm\}}$. We can note that, once we go lower with $\Delta P_{e\{norm\}}$ we will definitely need a higher quantization rate.

Now, we compare the quantized feedback scheme with the no feedback scheme.



(a) (b)

Figure 4.17: Quantization rate R_q Vs blocklength N for single user Gaussian channel with $C = 1$ and $\sigma_z^2 = 1$ using different $\Delta P_{e\{norm\}}$ for fixed transmit rates. (a) $R = 0.96$, (b) $R = 0.99$

The $\Delta P_{e\{norm\}}$ for this case is as follows,

$$\Delta P_{e\{norm\}} = \frac{P_{e_q} - P_{e\{Nofb\}}}{P_{e\{Nofb\}}}. \tag{4.23}$$

Figure 4.18 shows what quantization rate is needed in order to keep the difference of the error probability $\Delta P_{e\{norm\}} = 0$ for increasing blocklength. We need a higher quantization rate when a higher transmit rate is used. We can also can observe that for very high transmit rates the quantized feedback is not beneficial for low blocklength region as we will need infinite quantization rate $R_q \to \infty$ (see the red curve in figure 4.18).

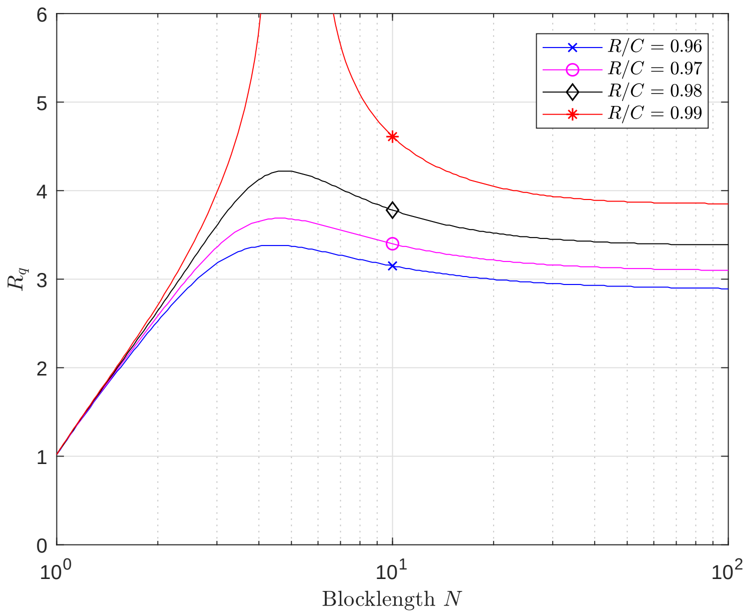


Figure 4.18: Quantization rate R_q Vs blocklength N for single user Gaussian channel with $C = 1$ and $\sigma_z^2 = 1$ using different transmit rates for $\Delta P_{e\{norm\}} = 0$

4.3.2 Multiple User Channel

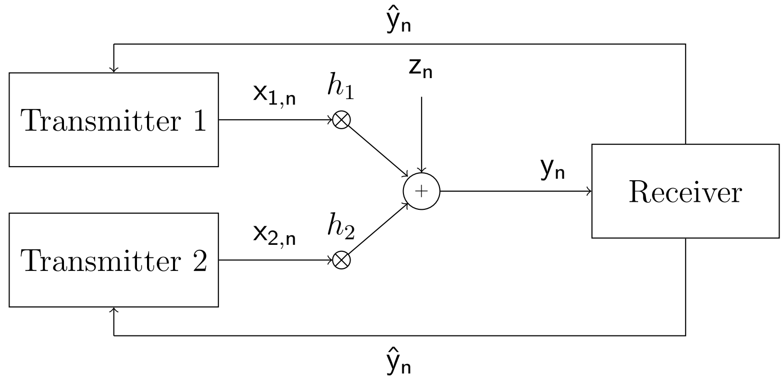


Figure 4.19: Quantized feedback for 2-User Gaussian MAC

In this section the quantized feedback scheme for the Gaussian multiple access channel will be discussed. We study the 2-user symmetric Gaussian MAC with quantized feedback shown in figure 4.19. Here the transmitters send independent Gaussian user signals $x_{1,n}$ and $x_{2,n}$ for $(n = 1, 2, ..., N)$ and the receiver observes the signal

$$y_n = h_1 x_{1,n} + h_2 x_{2,n} + z_n, \tag{4.24}$$

where the channel introduces additive Gaussian noise $\mathbf{z} \sim \mathcal{N}(0, \sigma_z^2)$. The average sum power constraint given by [12] is

$$\frac{1}{N} \sum_{n=1}^{N} E\{x_{1,n}^2 + x_{2,n}^2\} \leq P_1 + P_2. \tag{4.25}$$

Considering the symmetric case, $P_1 = P_2 = P$.

As mentioned in section 3.4.2, Ozarow introduced an error probability expression for the 2-user Gaussian MAC when perfect feedback is applied [9]

$$P_e = 2Q\left[\frac{\sigma_z^2}{2\sqrt{\alpha}(\sigma_z^2 + P(1-\rho^{*2}))} e^{N\left(C_{fb}-R\right)}\right], \tag{4.26}$$

where $\alpha = \sigma_z^2/12P$ is the variance of the estimated error described by equation (3.40) and $R = R_1 = R_2$ is the transmit rate.

The capacity region of the system when the quantized feedback is applied is given by [9]

$$C_{fb_q} = \Big\{(R_1, R_2) : (1) 0 \leq R_1 \leq \frac{1}{2}\log\big(1 + \frac{h_1^2 P_1}{\sigma^2}(1-\rho_q^2)\big),$$
$$(2) 0 \leq R_2 \leq \frac{1}{2}\log\big(1 + \frac{h_2^2 P_2}{\sigma^2}(1-\rho_q^2)\big), \tag{4.27}$$
$$(3) 0 \leq R_1 + R_2 \leq \frac{1}{2}\log\big(1 + \frac{h_1^2 P_1 + h_2^2 P_2 + 2\sqrt{h_1^2 P_1 h_2^2 P_2}\rho_q}{\sigma^2}\big)\Big\},$$

where $\sigma^2 = \sigma_z^2 + \sigma_q^2$ is the noise power of the overall channel $p(\hat{y}|x)$. And σ_q^2 is the variance of the quantization noise given by equation (4.15). To obtain the optimum correlation coefficient ρ_q^*, all the three conditions in equation (4.27) must be satisfied. Thus, R_1 and R_2 in conditions (1) and (2) respectively are added to each other and the result should be equalized to $R_1 + R_2$ in condition (3) (for more details, see section 3.4.2). The capacity of the overall channel after applying the quantizer given ρ_q^*, can be obtained from equations (4.27) and (4.17)

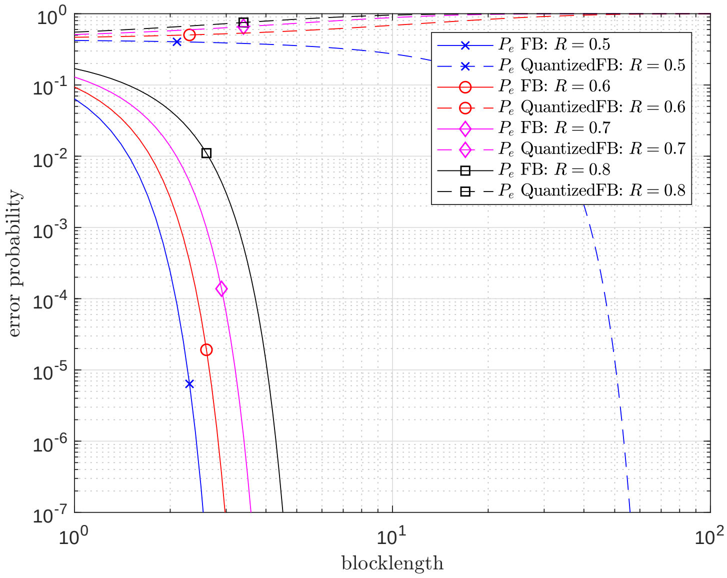
$$C_{fb_q} = \log\Big(1 + \frac{\gamma_q}{2}\big(1 - \rho_q^{*^2}\big)\Big), \tag{4.28}$$

where here γ_q is the signal to noise ration given by (4.14).

Now, we can substitute by C_{fb_q} in equation (4.26) so we get an expression of the error probability of the quantized feedback for 2-user MAC

$$P_{e_q} = 2Q\Bigg[\frac{\sigma^2}{2\sqrt{\alpha}(\sigma^2 + P(1-\rho_q^{*2}))}e^{N\left(C_{fb_q} - R\right)}\Bigg], \tag{4.29}$$

$$P_{e_q} = 2Q\Bigg[\frac{\sigma^2}{2\sqrt{\alpha}(\sigma^2 + P(1-\rho_q^{*2}))}e^{N\left[\log\left(1+\frac{\gamma_q}{2}\left(1-\rho_q^{*^2}\right)\right)-R\right]}\Bigg]. \tag{4.30}$$



[t]

Figure 4.20: P_e perfect feedback Vs P_e quantized feedback for symmetric 2-user Gaussian MAC at R_q=1, $C = 1$, $\sigma_z^2 = 1$, $C_{fb_q} = 0.535$, $\sigma_q^2 = 1.33$, $\rho_q^* = 0.55$

Figures 4.20, 4.21 and 4.22 compare the error probability of the perfect feedback with the error probability of the quantized feedback using quantization rates of 1,3 and 5 respectively for a symmetric 2-user Gaussian MAC. When a low quantization rate is used like in figure 4.20, a big difference in the performance of the two schemes can be clearly seen due to the high quantization error. By increasing the quantization rate, the quantization error decreases. Thus, the performance of the quantized feedback scheme becomes more closer to the perfect feedback scenario as shown in figures 4.21 and 4.22. In figures 4.23, 4.24 and 4.25 we can observe how the difference between the performance of the quantized feedback and the perfect feedback increases by increasing blocklength. The normalized difference $\Delta P_{e\{norm\}}$ between the error probabilities has been calculated according to equation (4.19).

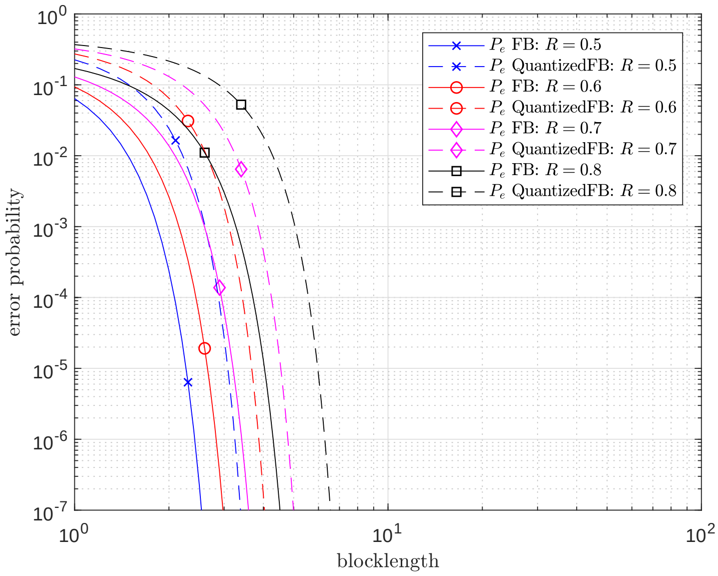


Figure 4.21: P_e perfect feedback Vs P_e quantized feedback for symmetric 2-user Gaussian MAC at R_q=3, $C = 1$, $\sigma_z^2 = 1$, $C_{fb_q} = 1.12$, $\sigma_q^2 = 0.064$, $\rho_q^* = 0.41$

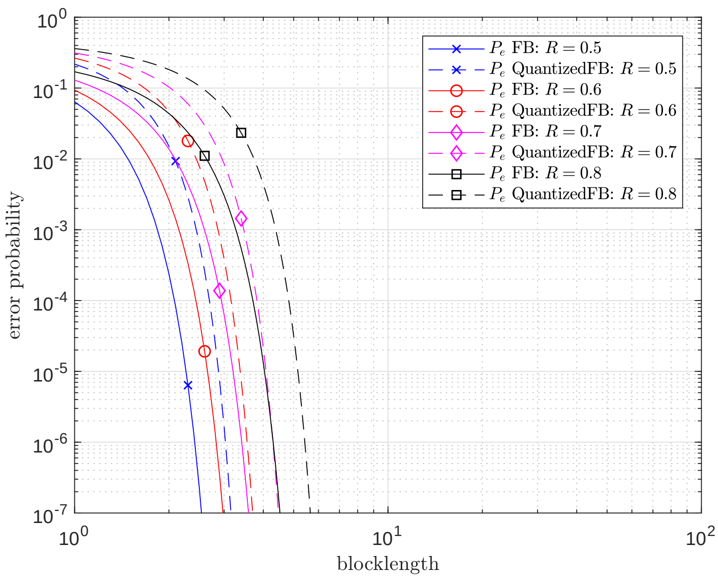


Figure 4.22: P_e perfect feedback Vs P_e quantized feedback for symmetric 2-user Gaussian MAC at R_q=5, $C = 1$, $\sigma_z^2 = 1$, $C_{fb_q} = 1.18$, $\sigma_q^2 = 0.004$, $\rho_q^* = 0.392$

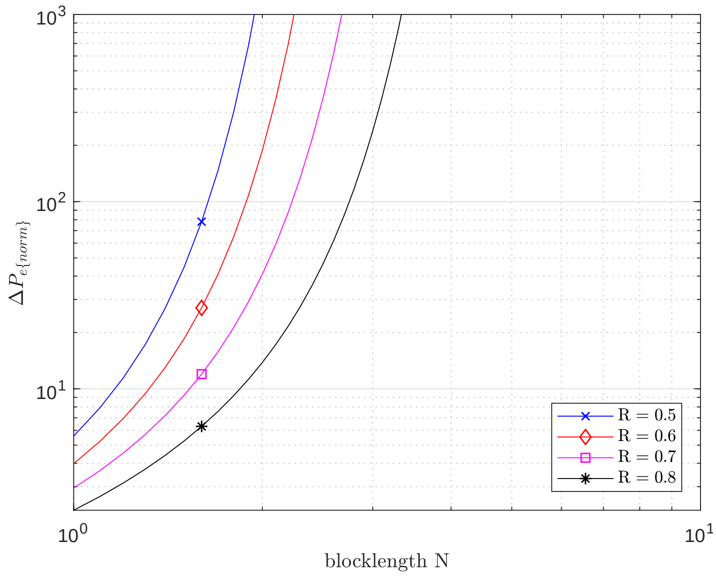


Figure 4.23: $\Delta P_{e\{norm\}}$ at $R_q = 1$ for symmetric 2-user Gaussian channel, $C = 1$, $\sigma_z^2 = 1$, $C_{fb_q} = 0.535$, $\sigma_q^2 = 1.33$, $\rho_q^* = 0.55$

Figure 4.26 combines the probabilities of error without feedback, with perfect feedback and with quantized feedback using different transmit rates. There is a big gap between the blue solid curve and the blue dashed curve representing the error probability without feedback and with perfect feedback respectively. The error probabilities of the quantized feedback scheme are located between the two blue curves and since the quantization rate is increased we get a performance closer to the perfect feedback scenario.

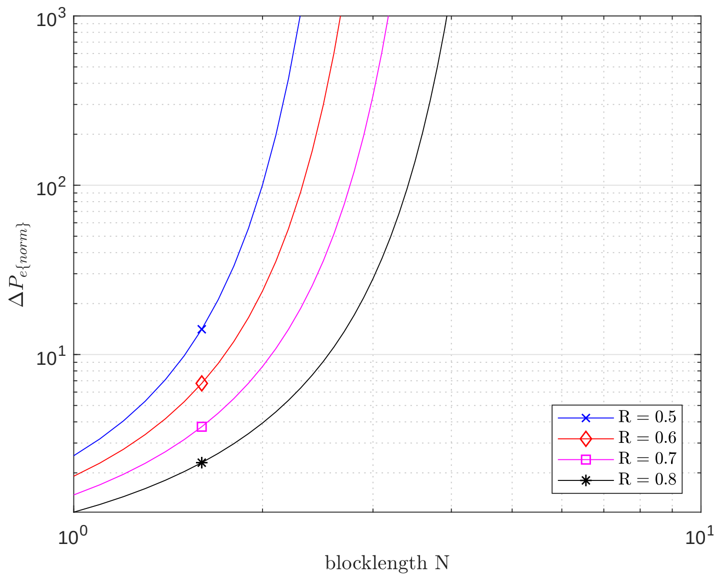


Figure 4.24: $\Delta P_{e\{norm\}}$ at $R_q = 3$ for symmetric 2-user Gaussian channel, $C = 1$, $\sigma_z^2 = 1$, $C_{fb_q} = 1.12$, $\sigma_q^2 = 0.064$, $\rho_q^* = 0.41$

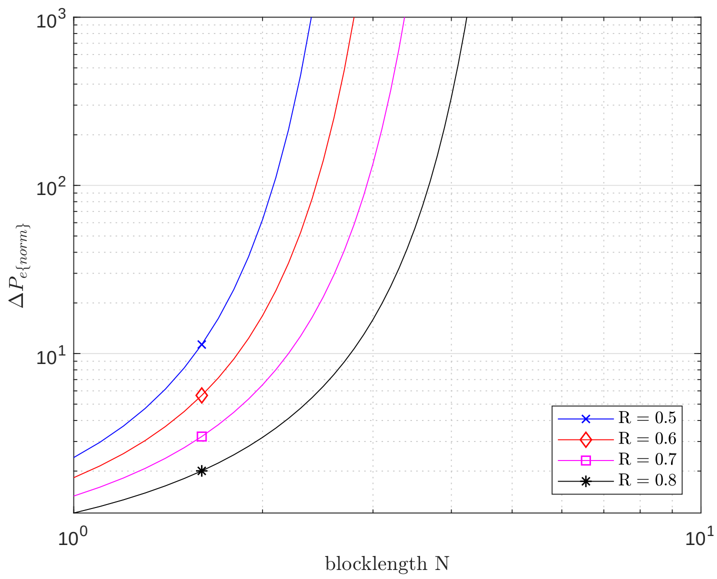


Figure 4.25: $\Delta P_{e\{norm\}}$ at $R_q = 5$ for symmetric 2-user Gaussian channel, $C = 1$, $\sigma_z^2 = 1$, $C_{fb_q} = 1.18$, $\sigma_q^2 = 0.004$, $\rho_q^* = 0.392$

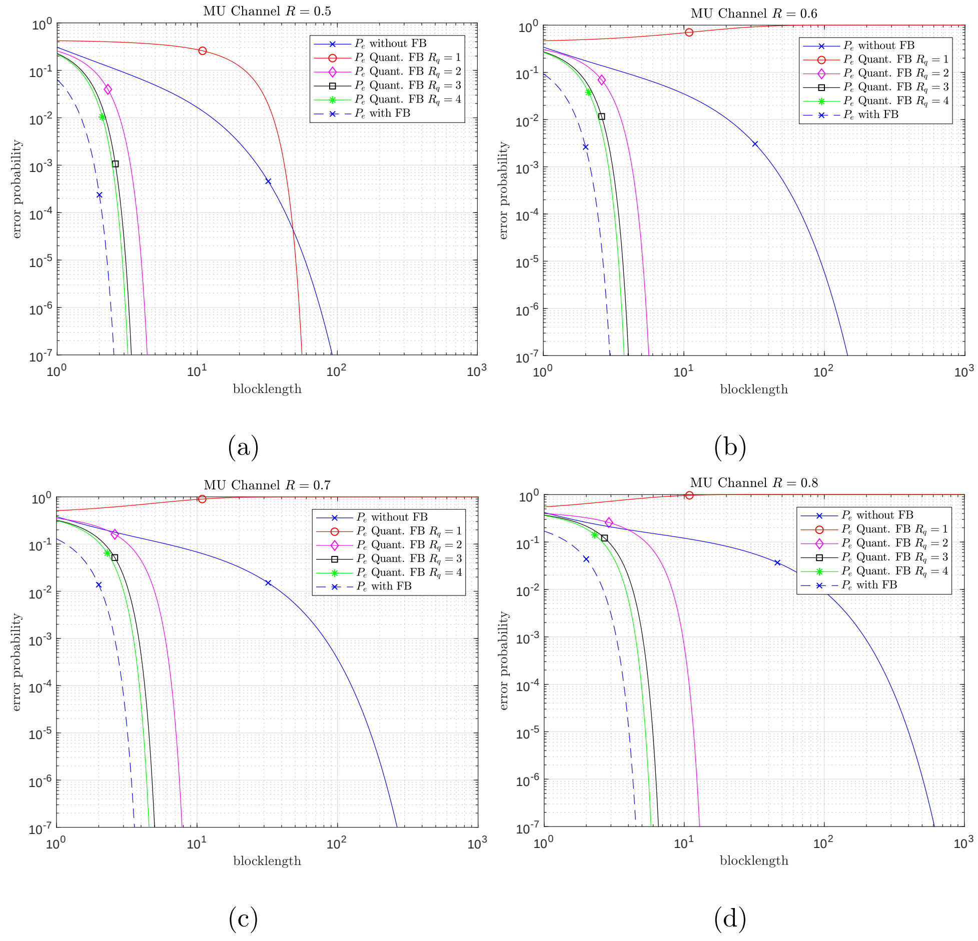
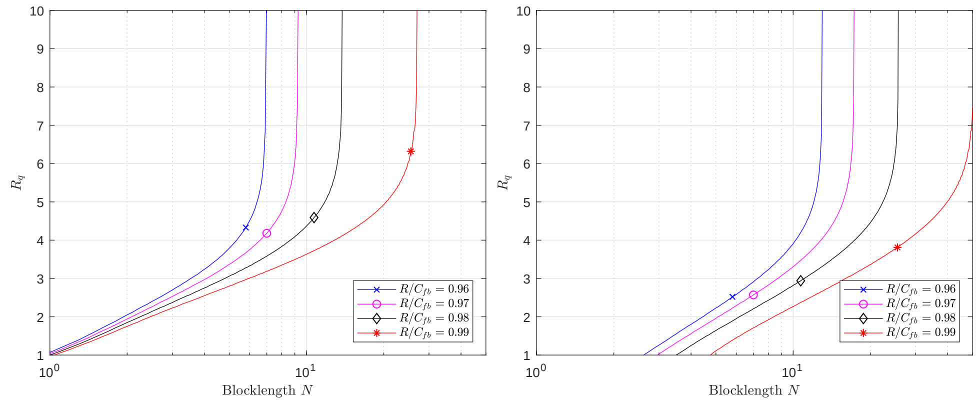


Figure 4.26: Comparing error probability without feedback, with perfect feedback and with quantized feedback for Gaussian MAC. (a) $R = 0.5$. (b) $R = 0.6$. (c) $R = 0.7$. (d) $R = 0.8$.



(a) (b)

Figure 4.27: Quantization rate R_q Vs blocklength N for symmetric 2-user Gaussian channel with $C_{fb} = 1.185$ and $\sigma_z^2 = 1$ using different transmit rates for fixed $\Delta P_{e\{norm\}}$. (a) $\Delta P_{e\{norm\}} = 1$, (b) $\Delta P_{e\{norm\}} = 2$

Figure 4.27 shows the quantization rate needed in a certain blocklength region for a specific $\Delta P_{e\{norm\}}$. We can observe that, for a higher blocklength we need a higher quantization rate in order to keep a constant difference between the error probabilities of the quantized feedback and the perfect feedback until we reach a threshold where an infinite quantization rate $R_q \to \infty$ is needed to keep this constant $\Delta P_{e\{norm\}}$. We can also observe that a higher quantization rate is needed when a lower transmit rate is used as the $\Delta P_{e\{norm\}}$ is increasing faster when higher transmit rates are used.

Continuing now with the comparison between the quantized feedback scheme and the no feedback scheme for the Gaussian MAC. The $\Delta P_{e\{norm\}}$ for this case is as follows,

$$\Delta P_{e\{norm\}} = \frac{P_{e_q} - P_{e\{Nofb\}}}{P_{e\{Nofb\}}}. \tag{4.31}$$

Figure 4.28 shows which quantization rate is needed in order to keep the $\Delta P_{e\{norm\}} = 0$ for increasing blocklength. We need a higher quantization rate

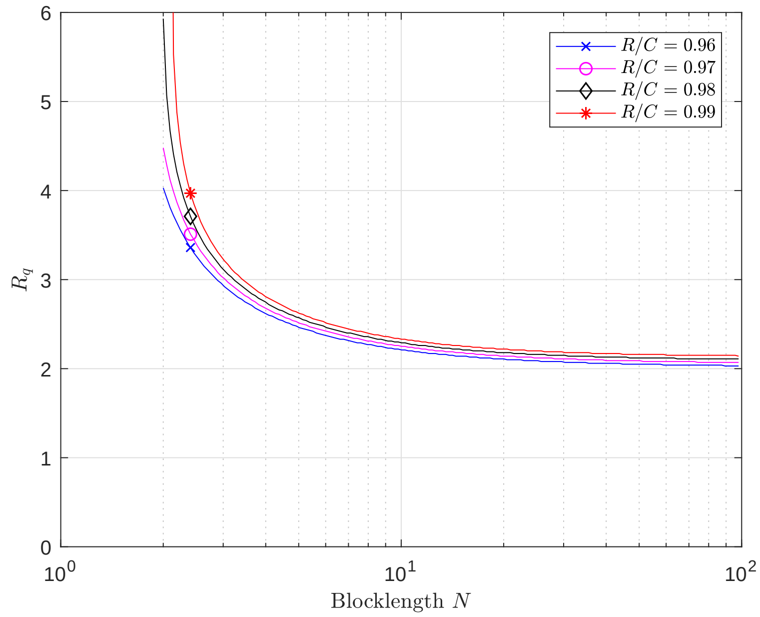


Figure 4.28: Quantization rate R_q Vs blocklength N for symmetric 2-user Gaussian MAC with $C = 1$ and $\sigma_z^2 = 1$ using different transmit rates for $\Delta P_{e\{norm\}} = 0$

when a higher transmit rate is used. We can also observe that for increasing blocklength the quantization rate decreases.

4.4 Achievable Rates

Here we study the maximum achievable rates of the non-feedback scenario, the perfect feedback scenario and the quantized feedback scenario. We will introduce the achievable rate expressions for the three schemes for single and multiple user Gaussian channels. Then a comparison between the achievable rates of the mentioned schemes will be discussed.

4.4.1 Single User Channel

The achievable rate for the single user Gaussian channel was introduced by Polyanskiy [10] and is represented by

$$R \approx C - \sqrt{\frac{V}{N}} Q^{-1}(P_e) + \frac{\alpha}{N} \log N, \tag{4.32}$$

In order to get the achievable rate of the perfect feedback scheme we can just recall the error probability equation (3.35) and reformulate it. The rate equation will be as follows

$$R = C - \frac{\ln\left(Q^{-1}(\frac{P_e}{2})\right)}{N}. \tag{4.33}$$

The achievable rate of the quantized feedback scheme can be calculated by substituting the capacity of the quantized feedback Gaussian channel given by (4.17) in (4.33)

$$R = C_q - \frac{\ln\left(Q^{-1}(\frac{P_e}{2})\right)}{N}. \tag{4.34}$$

Figures (4.29) and (4.30) compare the maximum rates that can be achieved when system without feedback, with perfect feedback and with quantized feedback is considered for single user Gaussian channel given error probability $P_e = 10^{-3}$ for linear and logarithmic x-axis scale respectively. The channel capacity here is normalized to 1 and the noise variance σ_z^2 is assumed to be 1.

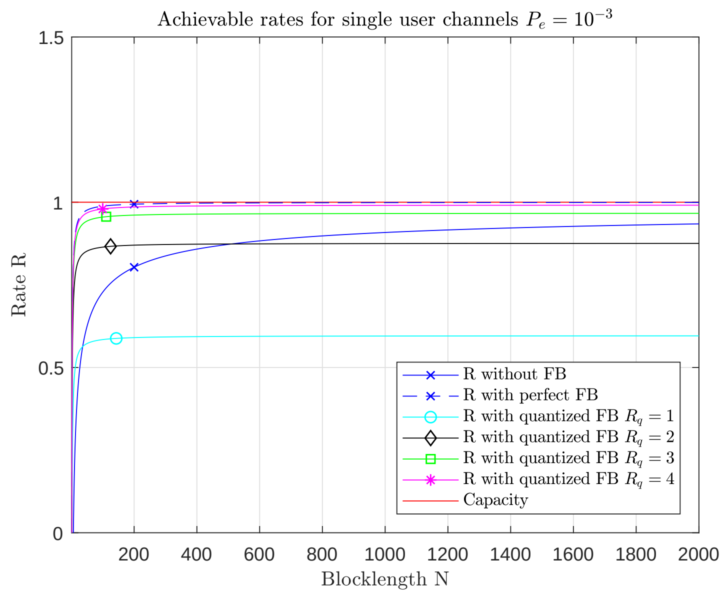


Figure 4.29: Achievable rates for single user Gaussian MAC with error probability $P_e = 10^{-3}$, $C = 1$, $\sigma_z^2 = 1$

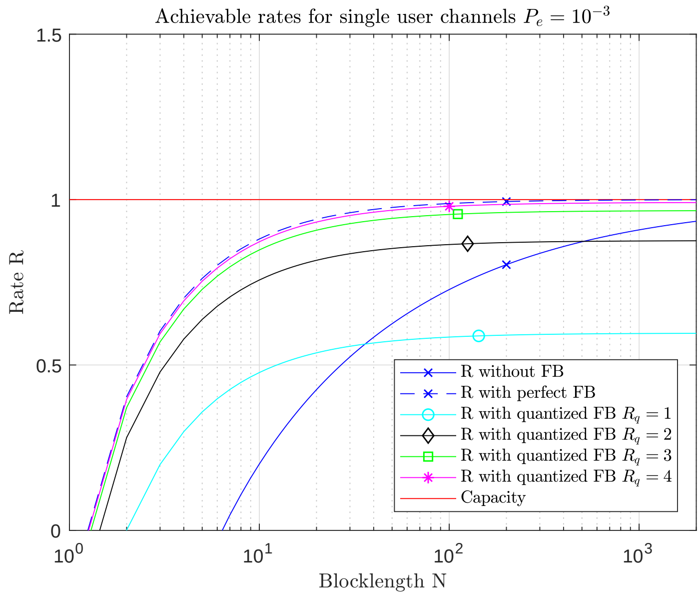


Figure 4.30: Achievable rates for single user Gaussian MAC with error probability $P_e = 10^{-3}$, $C = 1$, $\sigma_z^2 = 1$

4.4.2 Multiple User Channel

Now we will consider the 2-user Gaussian MAC and study the achievable rates of the three mentioned schemes. The achievable sum rate of the channel without feedback when a symmetric 2-user MAC is considered stays the same because the channel capacity will remain the same as in the single user case

$$R_1 + R_2 \approx C - \sqrt{\frac{V}{N}} Q^{-1}(P_e) + \frac{\alpha}{N} \log N. \tag{4.35}$$

Moving on to the achievable sum rate of the perfect feedback scenario, equation (3.41) showed the error probability of the perfect feedback for the 2-user Gaussian MAC. By reformulating this equation we can get an expression of the achievable sum rate

$$R_1 + R_2 = C_{fb} - \frac{1}{N} \ln \left[\frac{2\sqrt{\alpha}\big(\sigma_z^2 + P(1-\rho^{*^2})\big)}{\sigma_z^2} Q^{-1}\Big(\frac{P_e}{2}\Big) \right]. \tag{4.36}$$

The achievable sum rate of the quantized feedback channel can be found by reformulating the error probability equation (4.29)

$$R_1 + R_2 = C_{fb_q} - \frac{1}{N} \ln \left[\frac{2\sqrt{\alpha}\big(\sigma^2 + P(1-\rho_q^{*^2})\big)}{\sigma^2} Q^{-1}\Big(\frac{P_e}{2}\Big) \right], \tag{4.37}$$

where $\sigma^2 = \sigma_z^2 + \sigma_q^2$ and $P = P_1 + P_2$.

Figures 4.31 and 4.32 illustrate a comparison between the achievable sum rates of the three schemes: without feedback, with perfect feedback and with quantized feedback for a symmetric 2-user Gaussian channel. As shown the sum rate of the perfect feedback (the dashed blue curve) exceeds the channel capacity of a system without feedback at a certain blocklength value which means that the capacity is enlarged when we use a blocklength above this value (approximately 200). Almost the same applies to the quantized feedback with a

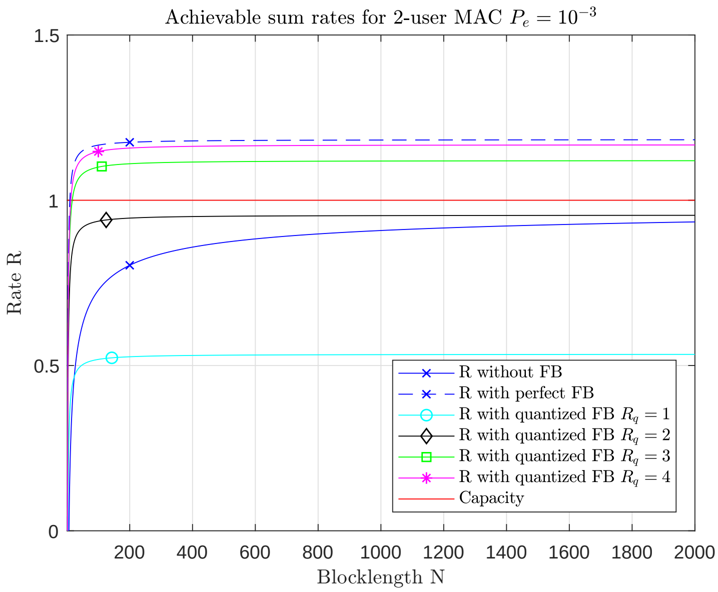


Figure 4.31: Achievable rates for symmetric 2-user Gaussian MAC with error probability $P_e = 10^{-3}$. $\sigma_z^2 = 1$

high quantization rate (the pink curve). In contrast, when very low quantization rates are applied the achievable rate decreases and becomes even worse than the non-feedback scheme (see the cyan curve).

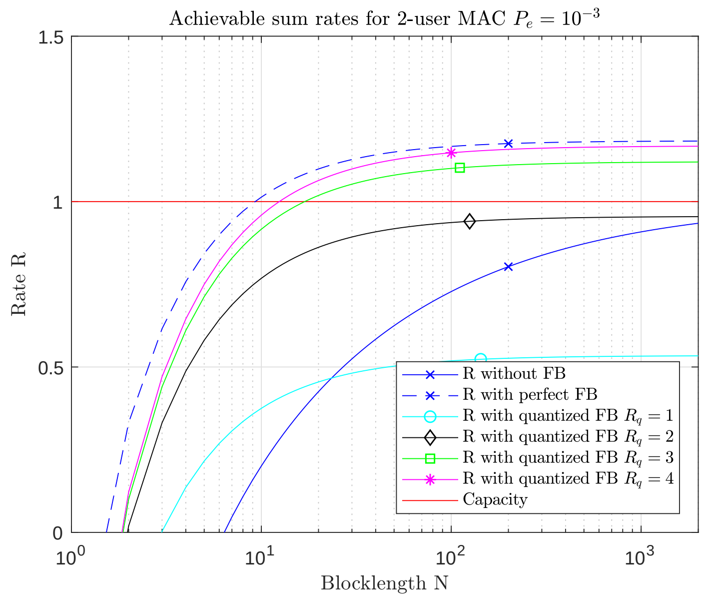


Figure 4.32: Achievable rates for symmetric 2-user Gaussian MAC with error probability $P_e = 10^{-3}$. $\sigma_z^2 = 1$

Chapter 5

Conclusion and Outlook

We studied the decoding error probability and the achievable rates for the system without feedback, with perfect feedback and with quantized feedback considering a finite blocklength for a Gaussian single user and 2-user multiple access channel. We confirmed that the error probability of the non-feedback scheme decreases exponentially with the blocklength while the error probability of the perfect feedback scheme decreases doubly exponentially with increasing blocklength. Furthermore, we confirmed that for the Gaussian MAC, the feedback can enlarge the sum achievable rate compared to the achievable rate of the non-feedback scenario.

Moreover, we introduced a quantized feedback scheme for finite blocklength coding for a Gaussian single user channel and multiple user channel. When the quantization rate increases the quantization error decreases and the performance of the quantized feedback scheme becomes closer to the perfect feedback scenario. We showed in what blocklength regions the quantized feedback is beneficial. For small blocklengths, only low quantization rates are needed to get a performance that is close to the perfect scenario but when the blocklength increases we will need higher quantization rates.

A symmetric 2-user Gaussian MAC is considered for all schemes discussed in the thesis. It is interesting in the future to see the results when asymmetric 2-user MAC is applied. In addition, it is very interesting to see how can we make this quantized feedback scheme practical. An approach will be to apply a practical quantizers and compare them with our quantized feedback scheme which is modeled as an information bottleneck method. Furthermore, we showed that for a higher blocklength, a higher quantization rate is needed to keep the performance very close to the perfect feedback scenario until a certain threshold then an infinite quantization rate is needed. So an important topic is, how can we increase this threshold and still get a very close performance to the perfect feedback for a larger blocklength.

Bibliography

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd edition.* New Jersey: John Wiley & Sons, 2006.

[2] N. T. Gaarder and J. K. Wolf, “The Capacity Region of a Multiple-Access Discrete Memoryless Channel Can Increase with Feedback,” in *IEEE Trans. Inf. Theory*, vol. 21, pp. 100–102, IEEE, Jan. 1975.

[3] N. Tishby and F. Pereira and W. Bialek, “The Information Bottleneck Method,” *Proc. 37th Annu. Allerton Conf. Commun., Control, Comput*, pp. 368–377, Sept. 1999.

[4] C. E. Shannon, “The Zero Error Capacity of a Noisy Channel,” in *IEEE Trans. Inf. Theory*, vol. 2, pp. 8–19, Sept. 1956.

[5] J. Schalkwijk and T. Kailath, “A Coding Scheme for Additive Noise Channels with Feedback-Part I,” in *IEEE Trans. Inf. Theory*, vol. 12, pp. 172–182, IEEE, Apr. 1966.

[6] H. Robbins and S. Monro, “A Stochastic Approximation Method,” in *Ann. Math. Stat.*, vol. 22, pp. 400–407, Sept. 1951.

[7] J. Schalkwijk and T. Kailath, “A Coding Scheme for Additive Noise Channels with Feedback–II: Band-Limited Signals,” in *IEEE Trans. Inf. Theory*, vol. 12, pp. 183–189, IEEE, Apr. 1966.

[8] R. G. Gallager and B. Nakiboglu, “Variations on a Theme by Schalkwijk and Kailath,” in *IEEE Trans. Inf. Theory*, vol. 56, pp. 6–17, Jan. 2010.

[9] L. H. Ozarow, “The Capacity of the White Gaussian Multiple Access Channel with Feedback,” in *IEEE Trans. Inf. Theory*, vol. 30, pp. 623–629, IEEE, July 1984.

[10] Y. Polyanskiy and H. V. Poor and S. Verdu, “Dispersion of Gaussian Channels,” in *IEEE Int. Symp. Inf. Theory (ISIT)*, (Seoul, Korea), July 2009.

[11] Y. Polyanskiy and H. V. Poor and S. Verdu, “Channel Coding Rate in the Finite Blocklength Regime,” *IEEE Trans. Inf. Theory*, vol. 56, pp. 2307–2359, Nov. 2008.

[12] E. MolavianJazi and J. N. Laneman, “A Second-Order Achievable Rate Region for Gaussian Multi-access Channels Via a Central Limit Theorem for Functions,” in *IEEE Trans, Inf. Theory*, vol. 61, pp. 6719–6733, IEEE, Oct. 2015.

[13] Z. Chance and D. J. Love, “Concatenated Coding for the AWGN Channel With Noisy Feedback,” *IEEE Trans. Inf. Theory*, vol. 57, pp. 6633–6649, Oct. 2011.

[14] J. Sacks, “Aymptotic Distribution of Stochastic Approximation Procedures,” in *Ann. Math. Stat.*, vol. 29, pp. 373–405, June 1956.

[15] T. Cover, “Some Advances in Broadcast Channels,” in *Advances in Communication Systems*, vol. 4, (San Francisco: academic press), pp. 229–260, 1975.

[16] A. D. Wyner, "Recent Results and in the Shannon Theory," in *IEEE Trans. Inf. Theory*, vol. 20, pp. 2–10, IEEE, Jan. 1974.

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

[18] C. E. Shannon, "A Mathematical and Theory of Communication," in *Bell Sys. Tech. Journal*, vol. 27, pp. 379–423, 623–656, Oct. 1948.

[19] S. Farthofer and G. Matz, "Linear Superposition Coding for the Asymmetric Gaussian MAC with Quantized Feedback," in *Proc. Asilomar Conference on Signals, Systems and Computers 2016*, (Monterey, CA), pp. 245–249, Nov. 2016.