2-1-2015

Resource allocation for layered broadcast over relay-assisted channels

Mohamed Adel Attia

AMERICAN UNIVERSITY IN CAIRO

RESOURCE ALLOCATION FOR LAYERED BROADCAST OVER RELAY-ASSISTED CHANNELS

by

Mohamed Adel Attia

A thesis submitted in partial fulfillment for the degree of
Master of Science

in the

SCHOOL OF SCIENCES AND ENGINEERING
ELECTRONICS AND COMMUNICATIONS ENGINEERING

July 2015

RESOURCE ALLOCATION FOR LAYERED BROADCAST OVER RELAY-ASSISTED CHANNELS

Presented by

Mohamed Adel Attia

For The Degree of

Master of Science

In Electronics and Communications Engineering

By

Mohamed Adel Attia

Examiners' Committee: Approved

Dr. Yasser Gadallah

Dr. Karim Seddik

Dr. Amr Elsherif

Advisors' Committee:

Dr. Karim Seddik

AMERICAN UNIVERSITY IN CAIRO

Abstract

SCHOOL OF SCIENCES AND ENGINEERING
ELECTRONICS AND COMMUNICATIONS ENGINEERING

Master of Science

by Mohamed Adel Attia

The topic of this thesis is on the application of multilayer transmission using the broadcast approach on a relay-assisted channel. Unlike single layer transmission, where all transmitted information bits have the same protection level by the channel coding scheme, multilayer transmission schemes combine successive refinement layered source coding with ordered protection levels of the source layers. Consequently, the receiver will be able to decode some information when the channel is faded and all information when the channel is good. The multilayer transmission schemes have gained a lot of interest in the information theory and the communication theory literature, where most researchers are interested in the broadcast approach since it is the optimal transmission strategy.

In this thesis, we consider a fading relay channel where the source uses layered source coding with successive refinement. The source layers are transmitted using superposition coding at the source with optimal resource allocation. The destination applies successive interference cancellation after optimally combining the direct and relayed signals. The resource allocation for the layers is subject to optimization in order to maximize the expected user satisfaction that is usually defined by a differentiable concave increasing utility function of the total decoded rate at the destination. As special cases, we consider two utility functions; namely, the expected total decoded rate at the receiver and the expected rate distortion of a Gaussian source. We also assume that only the channel statistics are known at the receiver. The relay is half-duplex and applies different relaying strategies, and we have investigated the Amplify-and-Forward, and Decode-and-Forward strategies in particular.

First, we consider the case of Decode-and-Forward relays where we consider two layers only with predetermined rates for simplicity, and we solve the problem of optimal power allocation among the two layers at the source and the relay using random search methods. After that, we solve the optimal power allocation problem for any number of layers with fixed rates over an Amplify-and-Forward relays. An approximation for the end-to-end channel quality is presented in terms of the statistics of the three links of the channel model. Furthermore, we obtain that for some conditions, it is optimal to send only one layer. Finally, we solve the joint optimal power and rate allocation problem for any number of layers over an Amplify-and-Forward relays. We also consider the theoretical case of infinite number of layers representing an upper bound for the performance. Moreover, we show that with a small number of layers, we can approach the performance upper bound. We provide many numerical examples for the three cases above to show the prospected gains of using the relays on the expected utility for different channel conditions.

Acknowledgements

All praise be to Allah, the lord of the worlds, most Gracious, most merciful.

I would like to express my deep sense of gratitude to my advisor Dr. Karim Seddik for his guidance, patience, motivation, knowledge, and continuous support throughout my graduate studies over the past two years. It is not often that one can find an advisor who treats him as his younger brother. He has taught me innumerable lessons and insights on the workings of academic research in general.

I am also very grateful to Dr. Mohammad Shaqfeh for his scientific advice, knowledge, and many insightful discussions and suggestions through the course of this work.

My sincere regards to my friends, and colleagues who have directly or indirectly helped me in this thesis.

Last but not least; this work is dedicated to my dear parents for their encouragement and love during my whole life.

Contents

List of Figures

List of Tables

Abbreviations

AF	**A**mplify and **F**orward
CDF	**C**umulative **D**istribution **F**unction
CSI	**C**hannel **S**tate **I**nformation
DF	**D**ecode and **F**orward
ISI	**I**nter **S**ymbol **I**nterference
KKT	**K**arush **K**uhn **T**ucker
LOS	**L**ine **O**f **S**ight
MRC	**M**aximum **R**atio **C**ombining
PDF	**P**robability **D**ensity **F**unction
QoS	**Q**uality **o**f **S**ervice
SR	**S**uccessive **R**efinement
SNR	**S**ignal to **N**oise **R**atio
SIR	**S**uccessive **I**nterference **C**ancellation

Chapter 1

Introduction

Large demands for high *Quality of Service QoS* for new multimedia applications has increased the attention towards replacing the conventional single-layer transmission with new transmission schemes, in order to overcome some technical challenges related to the impairments of the wireless medium. Wireless networks are suffering from fading effects caused by multi-path propagation and shadowing over the wireless medium. Therefore, it is infeasible to maintain a constant QoS for all users in every instantaneous moment.

1.1 Wireless Fading Channels

Wireless medium is suffering from many impairments due to the fading nature of the channels. Fading is the random fluctuations in the amplitude and the phase of the received signals due to the reflections of the transmitted signals [1], [2]. These reflections are caused by many obstacles in the *Line Of Sight LOS* between the transmitter and the receiver as shown in Figure 1.1. If the tansmitted signal is $x(t)$, then the received signal $y(t)$ can be expressed as

$$y(t) = \sum_{\ell=1}^{L} h_{\ell}(t)x(t-\tau_{\ell}) + n(t), \tag{1.1}$$

where L is the number of paths, $n(t)$ is the additive noise, and τ_{ℓ} and $h_{\ell}(t)$ are the delay and the channel coefficient for ℓ-th path, respectively. The fading of the wireless channel can be classified to flat-fading or multipath fading according to the value of the delay spread of the channel $\Delta\tau$, which is the difference between the maximum and the minimum values of the delay spread $\Delta\tau = \max_{\ell} \tau_l - \min_{\ell} \tau_l$.

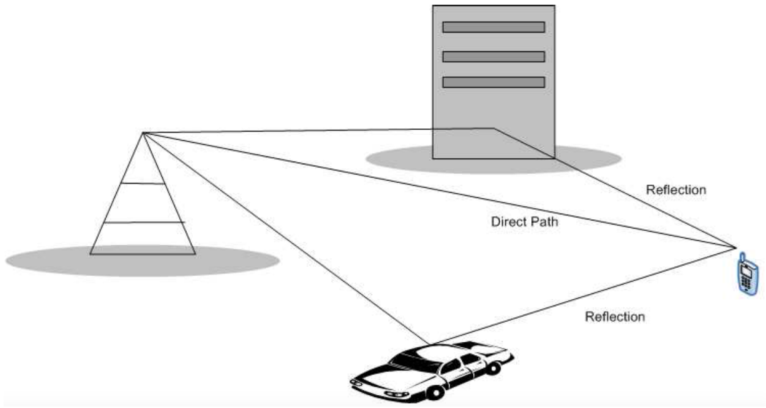


FIGURE 1.1: The Wireless Fading Channel

If the delay spread is small compared to the symbol duration then is channel is known as flat (frequency non-selective) fading. Otherwise, the fading is known as multipath (frequency selective) fading, where different fading values are assigned for different frequency components, and this will result in *Inter Symbol Interference ISI*. Due to this impairments of the wireless channel, it is infeasible to maintain a constant QoS for all users.

1.2 Single Layer Transmission

Single-layer transmission has been widely used in the last few decades. From an information theoretic perspective, the information in the codeword is either decoded successfully or unsuccessfully as a whole, depending on the channel condition, usually described as the "On-Off" nature for the codewords. If the channel capacity is lower than the rate of the codeword (i.e., channel is in deep fade), an error occurs and the codeword as a whole will not be decoded reliably. On the other hand, the codeword is decoded successfully and reliably if the channel capacity exceeds the rate. Consequently, due to the dispersive nature of the channel, rate adaptation is usually sought to compensate for the variation in the channel capacity over the time, in order for the receiver to decode the data reliably.

However, rate adaptation in order to avoid the events of outage requires a complete knowledge of the channel at the transmitter (i.e., *Channel State Information CSI*). For multimedia systems, this approach may be impractical. The CSI may be not available at the transmitter, especially for delay-constrained applications where feedback latency is intolerable. Moreover, for multicast systems and for all users to decode the information successfully, it is needed to maintain the rate lower than the capacity of the worst channel in the system. This implies that all users will be receiving the worst QoS, which is related to the source information rate, even if their channel state is good.

Therefore, for the multimedia multicast applications using single layer transmission, a trade-off exists between maintaining high QoS, and reducing the outage probability at the receivers. Restricting the rate lower than the capacity of the worst channel in the system will decrease the outage events while maintaining the worst QoS for all the users. Conversely, transmitting with high source rates will enhance the opportunity of receiving high quality in good channel conditions but with high probability of outage for users suffering from bad channel conditions.

1.3 Multilayer Transmission

Unlike the "On-Off" nature for the single layer transmission, "multilayer" transmission schemes combine *Successive Refinement SR* layered source coding [3], [4] with ordered protection levels at the physical layer. Therefore, the "base" source layer is given higher priority and protected more than the "enhancement" source layers even by allocating more power or reducing the allocated rate so that they can be decoded at relatively low *Signal to Noise Ratio SNR* thresholds, while the enhancement source layers are decoded at higher SNR thresholds. Consequently, Multilayer transmission makes it possible to partially decode the information when the channel is not good. Thus, the receiver will be able to get an "acceptable but not perfect" QoS when the channel is faded and "perfect" QoS when the channel is good. Of course, this is better than the "all or nothing" scenario that is inherent in the conventional single-layer transmission schemes.

In addition to its practical merits in multimedia broadcasting and multi-casting applications, the multilayer transmission schemes have gained a lot of interest in the information theory and the communication theory literature, where most researchers are interested in the "broadcast approach" since it is the optimal transmission strategy. In this case, the source layers are protected independently using different channel codewords and transmitted jointly using superposition coding at the physical layer [5], [6], [7] with *Successive Interference Cancellation SIC* at the receiver. Therefore, the first layer is decoded first, and the upper layers in this case are treated as interference. Once the layer is decoded, it is subtracted from the total received signal, and this process is repeated up to the supported layer by the channel quality.

Other possible approaches for the multilayer transmission including coupling SR source coding with "orthogonal transmission" of the layers was considered in [8] with predetermined rates for the source layer. In [9] a comparison was done assuming equal rates of the source layers between the approach of combining SR coding with broadcast strategy and other cross layer approaches including progressive (i.e. orthogonal) transmission of the source layers and hybrid analog/digital schemes [10].

The design parameters that are subject to optimization in the broadcast approach are the allocated rates and power ratios of the different source layers. By changing the allocated rate and power ratio for the layer, the SNR threshold for this layer changes. Thus, the adaptation of the design parameters produces a lot of trade-offs in the system performance. In order to decrease the SNR threshold for the layer so as to be decoded more reliably, it required to decrease the rate for the layer which affects the over-all QoS, or allocate more of the total power budget for this layer which will cause, as a consequence, an increase in the SNR threshold for the remaining layers. Therefore, an optimization over the design parameters (rates and power ratios) is required in order to optimally enhance the long-term performance. The performance can be described by a utility function of the decoded rates at the receiver. There are several measures for the system performance. Two of the common objectives are maximizing the expected rate e.g. [11], [12]. and minimizing the expected distortion of the Gaussian source, e.g. [13], [14].

1.4 Thesis Contribution and Organization

Our main interest in this work is in the application of multilayer transmission using the broadcast approach in the context of relay-assisted networks [15], [16]. Our initial contribution in this topic was by considering *Decode and Forward DF* relays [17]. We have also investigated the *Amplify and Forward AF* scenario for the multilayer transmission later in [18], where we extended the solution algorithm presented in [19] for fixed and predetermined rates of the source layers. Then, we examined recently the extension of the optimization framework for jointly optimal rate and power allocation presented in [20] to the AF relay channel case. As well known, relaying and cooperative communication schemes are involved in LTE-Advanced systems and they have been a very active research area recently. For brevity, we refer here to only few examples of important contributions in this interesting topic [18, 21–26].

The remainder of this thesis is organized as follows; we consider a simple case of two layer transmission and solve the optimal power allocation problem assuming DF relays in Chapter 2. In Chapter 3, we examine the case of the AF relays and solve the optimal power allocation problem for any number of layers. Next, we solve the optimal jointly rate and power allocation for the AF relaying case and for any number of layers in Chapter 4. Finally, we summarize the main conclusion and some directions for future research in Chapter 5.

1.4.1 Power Optimization for Layered Transmission Over DF Relay Channels (Chapter 2)

In Chapter 2, we consider layered source transmission over a single-node half-duplex relay channel. Therefore, the transmission is done over two consecutive time slots with equal duration and bandwidth. The layers are transmitted at the source with optimal power allocation to both the relay and the destination in the first time slot. The relay is transmitting using DF relaying strategy, where the relay decodes the information received from the source and only forwards the layers decoded successfully with new optimal power allocation. We characterize the expected utility function describing the average user satisfaction for two cases; namely, the expected total decoded rate at the receiver and the expected rate distortion of a Gaussian source. The optimal power allocation problem is solved so that the average utility function is maximized.

Obtaining the closed form expression for the end-to-end channel condition is not feasible in this case, then we consider a simple case of two layer transmission with fixed rates. Since the optimal power allocation at the relay may not necessarily follow the same power allocation at the source. Hence, the optimal power allocation at the relay should be considered as well, and we need to find the optimal power allocation for the two layers at both the source and the relay. As a result, the number of optimization variables increases considerably (since the power allocation at the relay will be conditional on the number of layers decoded at the relay). Hence, the number of optimization variables becomes $\frac{M(M+3)}{2} - 1$, where M is the total number of layers. Furthermore, the solution presented in [19] cannot be applied. So, numerical random search methods should be applied which becomes inefficient and expensive in terms of the computation load as the number of optimization variables increases. We provide several numerical results describing the gain for using the relay with different positions.

1.4.2 Optimal Power Allocation for Layered Broadcast Over AF Relay Channels (Chapter 3)

In Chapter 3, we describe a layered source transmission aided with AF single node relay. The transmission is done over two consecutive time slots with equal duration and bandwidth. In the first time slot, the source transmits the layers with optimal power allocation to both the relay and the destination. The relay forwards the information received in the first time slot from the source to the destination with the same power ratios in the second time slot. Note that the relay here only forwards the data after amplifying without the need to decode the date. The advantage of AF relays is that there is no error propagation as in the DF strategies, because the relays do not perform any hard-decision operation on the received signal. Also, the relays is transmitting the layers with the same power ratios as the source. Therefore, it is feasible in this case to obtain the end-to-end channel quality in order to use the solution algorithm presented in [19] for any number of layers with linear complexity. However, in the AF protocol noise accumulates with the desired signal along the transmission path.

We solve the optimal power allocation problem for M number of layers with predetermined rates, so that the utility function is maximized. We also consider two special cases of maximizing the average rate, or minimizing the expected distortion of a Gaussian source. An approximation is proposed for the end-to-end channel quality given that all three links in our channel model are Rayleigh faded. Consequently, we can use the algorithm proposed in [19] to solve the problem, and we provide several numerical results describing the gain of using the relay with different relay position. We also show that it may be optimal not to transmit all the layers depending on the channel condition. Furthermore, we show that using the AF relays may achieve lower values for the maximized utility function than the no-relay case because of the multiplexing loss of transmitting over two time slots.

1.4.3 Jointly Optimal Power and Rate Allocation for Layered Broadcast Over AF Relay Channels (Chapter 4)

The jointly optimal rate and power allocation for multi-layer transmission over AF relays is considered in Chapter 4. It is true that the predetermined rate allocation is considered more practical, where the joint source-channel coding is not feasible. However, the fixed rate scenarios considered in Chapter 3 are considered sub-optimal compared to the flexible rate allocation case considered in this chapter. We use the same channel approximation proposed in Chapter 3, so that we can apply the solution algorithms presented in [20]. Furthermore, we consider the theoretical case of infinite number of layers, which gives an upper bound for the performance and is considered the global optimal solution. However, it is shown that this upper bound can be achieved even with small number of layers. Numerical results are provided showing the gain of using the AF relay for different relay positions, and the gain of using optimally joint rate and power allocation over other suboptimal schemes.

Chapter 2

Power Optimization for Layered Transmission Over DF Relay Channels

2.1 Introduction

Our contribution of this chapter is on the investigation of multilayer transmission on a relay channel, where we consider a layered transmission scheme with optimal power allocation at the source and the relay. This work is already published in [17] where we assume that the relay applies the selection-relaying *Decode-and-Forward DF* strategy [21]. We consider a utility function as a measure for the user satisfaction. This function can be defined flexibly for many cases; such as maximizing the expected rate or minimizing the expected distortion. The optimization problem is formulated as utility maximization with known layers rates. The expected utility function in our problem is a function of the channel statistics of the three links in the channel (i.e., source-destination, source-relay and relay-destination).

In this chapter, we show that finding the end-to-end channel quality as in [19] is infeasible. Therefore, we only find the solution in a special case of two-layer transmission with predetermined rates. The extension into more than two layers in the DF relay case becomes prohibitively more complex. We characterize the expected utility function and use it in the power optimization problem. Also, we apply the "random search" method [27, Chapter 14] to solve this optimization problem, and we provide several numerical examples to show the gains in the maximum expected utility when relaying is applied. The random search method is not considered an efficient way in solving the problem, as it is time consuming and needs high number of iterations to find an accurate solution. However, it is simple and suits the objective function under optimization, which is found to be extremely complex. Other possible methods in finding the optimal solution is found in [27], such as Gradient and Steepest Ascent methods where we move closer and faster to the maximum. However, it needs extra mathematical manipulations and approximations in order to find a closed form for the objective function.

2.2 System Model and Problem Formulation

2.2.1 System Model and Transmission Scheme

We consider a system that consists of three nodes; source, destination and relay. We assume that the source is Gaussian and it is encoded into two layers, L_1 and L_2, with rates R_1 and R_2 respectively. L_1 is the base layer, and L_2 is the enhancement layer that refines the description in L_1. The relay is half-duplex and applies selection-relaying DF strategy. Therefore, the transmission is carried over two consecutive time slots of equal duration and bandwidth. In the first time slot, the source broadcasts L_1 and L_2 to the relay and the destination using superposition coding, where L_1 should be decoded first then L_2 using SIC. If the relay is able to decode one or both layers, it forwards the decoded layer(s) to the destination using new complementary[1] codewords in the second time slot. Otherwise, the source re-transmits L_1 and L_2 in the second time slot using new complementary codewords. In both cases, the destination tries to decode L_1 first and then L_2 based on the received codewords in the two time slots of the transmission.

We assume that the three nodes are equipped with a single antenna. Also, we assume that the source and the relay transmit using constant power P_s and P_r respectively. We use the notation $\frac{|h|^2P}{\sigma^2}$ to represent the Signal-to-Noise-Ratio (SNR) of the channel, where $|h|$ is the magnitude of the channel gain and σ^2 is the received noise variance. Therefore, the SNR over the three links of the relay channel can be denoted using γ_{sr}, γ_{sd} and γ_{rd} for the source-relay, source-destination, and relay-destination links, respectively, and they are given as

$$\gamma_{\mathsf{sr}} = \frac{|h_{\mathsf{sr}}|^2 P_s}{\sigma^2_{\mathsf{sr}}}, \quad \gamma_{\mathsf{sd}} = \frac{|h_{\mathsf{sd}}|^2 P_s}{\sigma^2_{\mathsf{sd}}}, \quad \gamma_{\mathsf{rd}} = \frac{|h_{\mathsf{rd}}|^2 P_r}{\sigma^2_{\mathsf{rd}}}. \tag{2.1}$$

Furthermore, we assume that the channel gains, and consequently the SNRs, stay constant for the duration of one transmission block, which consists of two consecutive time slots. However, γ_{sr}, γ_{sd} and γ_{rd} vary from one channel block to another randomly. Furthermore, we assume that the source and the relay do not know the instantaneous values of the SNRs while transmitting.

[1] We assume that the codewords achieves the information-theoretic maximum achievable rates of decode-and-forward over the relay channel.

In the numerical results in Section 2.4, we assume that the variation (i.e., fading) of the channels' gain is Rayleigh distributed[2]. Hence, the *Probability Density Function PDF* of the channels follow an exponential distribution, and they are given as

$$f_{\mathsf{sd}}\left(\gamma_{\mathsf{sd}}\right) = \frac{1}{\bar{\gamma}} \exp\left(\frac{-\gamma_{\mathsf{sd}}}{\bar{\gamma}}\right), \tag{2.2a}$$

$$f_{\mathsf{sr}}\left(\gamma_{\mathsf{sr}}\right) = \frac{1}{m\bar{\gamma}} \exp\left(\frac{-\gamma_{\mathsf{sr}}}{m\bar{\gamma}}\right), \tag{2.2b}$$

$$f_{\mathsf{rd}}\left(\gamma_{\mathsf{rd}}\right) = \frac{1}{m\bar{\gamma}} \exp\left(\frac{-\gamma_{\mathsf{rd}}}{m\bar{\gamma}}\right), \tag{2.2c}$$

for the source-destination, source-relay and relay-destination links, respectively. In (2.2), $\bar{\gamma}$ is the average SNR for the direct source-destination link and m is the ratio between the average SNR of the source-relay and the relay-destination links to the source-destination link. We assume that $\bar{\gamma}$ and m are known at the source and the relay and they are used in the optimization of the power allocation over the two layers L_1 and L_2 at these two nodes to maximize the expected utility function, denoted U, of the total decoded rate, denoted $\bar{R}$, at the destination.

The optimization variables are denoted α_1 and α_2 for the ratios of the total power at the source that are allocated to L_1 and L_2, respectively. Additionally, we have β_1 and β_2 for the ratios of the total power at the relay that are allocated to L_1 and L_2, respectively, when the relay transmits both layers. However, when the relay can decode only L_1, it forwards only this layer, and hence it allocates all of its power to it.

2.2.2 Mathematical Notation and Problem Formulation

First, we define the following functions because we will use them frequently in the sequel.

$$\mathcal{C}_1(\gamma, \epsilon) = \log\left(1 + \frac{(1-\epsilon)\gamma}{1+\epsilon\gamma}\right), \tag{2.3a}$$

$$\mathcal{C}_2(\gamma, \epsilon) = \log\left(1 + \epsilon\gamma\right). \tag{2.3b}$$

[2]The extension of the results of this thesis into other channel fading models is straightforward.

The functions in (2.3) define the information-theoretic maximum achievable rates for the transmission of two layers over a Gaussian channel with SNR γ, and power ratio of the enhancement layer equals ϵ.

$$\Gamma_1(x,\epsilon) = \frac{x-1}{1-\epsilon x}, \tag{2.4a}$$
$$\Gamma_2(x,\epsilon) = \frac{x-1}{\epsilon}. \tag{2.4b}$$

Also, we use the functions in (2.4) to denote the information-theoretic minimum SNR threshold that is required in order to be able to decode two layers over a Gaussian channel with x a function of the layer rate as will be explained in the sequel, and ϵ is the power ratio allocated to the enhancement layer.

In this thesis, our objective is to maximize the expected user satisfaction determined by the utility function $U(\bar{R})$. The utility function $U(\bar{R})$ can be flexibly defined to employ many special cases such as minimizing the expected distortion of a Gaussian source or maximizing the expected rate. The optimization problem is to optimally allocate the power among the two layers such that the expectation of the utility function $E[U(\bar{R})]$ is maximized. The expectation of the utility function can be described as

$$E\left[U(\bar{R})\right] = U(R_1).P_{\mathsf{d1}} + U(R_1 + R_2).P_{\mathsf{d2}}, \tag{2.5}$$

where P_{d1} and P_{d2} are defined as

$$P_{\mathsf{d1}} \equiv \Pr\left(\text{Destination can decode } \mathrm{L}_1 \text{ only}\right), \tag{2.6a}$$
$$P_{\mathsf{d2}} \equiv \Pr\left(\text{Destination can decode both } \mathrm{L}_1 \text{ and } \mathrm{L}_2\right). \tag{2.6b}$$

These probabilities can be characterized as

$$P_{\mathsf{d1}} = P_{\mathsf{d1|r0}}.P_{\mathsf{r0}} + P_{\mathsf{d1|r1}}.P_{\mathsf{r1}} + P_{\mathsf{d1|r2}}.P_{\mathsf{r2}}. \tag{2.7}$$

Similarly, we have

$$P_{\mathsf{d2}} = P_{\mathsf{d2|r0}}.P_{\mathsf{r0}} + P_{\mathsf{d2|r1}}.P_{\mathsf{r1}} + P_{\mathsf{d2|r2}}.P_{\mathsf{r2}}, \tag{2.8}$$

where the following notations are used

$$P_{\mathsf{r}0} \equiv \Pr\left(\text{Relay cannot decode } \mathrm{L}_1 \text{ and } \mathrm{L}_2\right), \tag{2.9a}$$
$$P_{\mathsf{r}1} \equiv \Pr\left(\text{Relay can decode } \mathrm{L}_1 \text{ only}\right), \tag{2.9b}$$
$$P_{\mathsf{r}2} \equiv \Pr\left(\text{Relay can decode both } \mathrm{L}_1 \text{ and } \mathrm{L}_2\right), \tag{2.9c}$$

$$P_{\mathsf{d}1|\mathsf{r}0} \equiv \Pr\left(\text{Destination can decode } \mathrm{L}_1 \text{ only}|\text{Relay cannot decode } \mathrm{L}_1 \text{ and } \mathrm{L}_2\right), \tag{2.10a}$$
$$P_{\mathsf{d}1|\mathsf{r}1} \equiv \Pr\left(\text{Destination can decode } \mathrm{L}_1 \text{ only}|\text{Relay can decode } \mathrm{L}_1 \text{ only}\right), \tag{2.10b}$$
$$P_{\mathsf{d}1|\mathsf{r}2} \equiv \Pr\left(\text{Destination can decode } \mathrm{L}_1 \text{ only}|\text{Relay can decode both } \mathrm{L}_1 \text{ and } \mathrm{L}_2\right), \tag{2.10c}$$

$$P_{\mathsf{d}2|\mathsf{r}0} \equiv \Pr\left(\text{Destination can decode both } \mathrm{L}_1 \text{ and } \mathrm{L}_2|\text{Relay cannot decode } \mathrm{L}_1 \text{ and } \mathrm{L}_2\right), \tag{2.11a}$$
$$P_{\mathsf{d}2|\mathsf{r}1} \equiv \Pr\left(\text{Destination can decode both } \mathrm{L}_1 \text{ and } \mathrm{L}_2|\text{Relay can decode } \mathrm{L}_1 \text{ only}\right), \tag{2.11b}$$
$$P_{\mathsf{d}2|\mathsf{r}2} \equiv \Pr\left(\text{Destination can decode both } \mathrm{L}_1 \text{ and } \mathrm{L}_2|\text{Relay can decode both } \mathrm{L}_1 \text{ and } \mathrm{L}_2\right). \tag{2.11c}$$

Notice that $P_{\mathsf{r}i}$ depends on γ_{sr} while $P_{\mathsf{d}i|\mathsf{r}j}$ depends on γ_{sd} and γ_{rd}. We assume that the channels are fading independently.

The main optimization problem is

$$\max_{\alpha_2,\beta_2} \quad E[U(\bar{R})], \tag{2.12a}$$
$$\text{subject to} \quad 0 \leq \alpha_2, \beta_2 \leq 1, \tag{2.12b}$$

where α_1 and β_1 are equal to

$$\alpha_1 = 1 - \alpha_2, \qquad \beta_1 = 1 - \beta_2 \tag{2.13}$$

The first step to solve (2.12) is to find the probability that the destination is able to decode only layer L_1 correctly, then find the probability that the destination is able to decode both layers L_1 and L_2 correctly. These two probabilities will be substituted in (2.5), and then it is required to find optimal power ratios α_2 and β_2 to solve (2.12) using "random search" numerical method.

2.3 Characterizing the Successful Decoding Probabilities

2.3.1 Successful Decoding Probabilities at the Relay

$$
\begin{aligned}
P_{\mathsf{r}0} &= \Pr\left(R_1 > \frac{1}{2}\mathcal{C}_1(\gamma_{\mathsf{sr}}, \alpha_2)\right) && (2.14a)\\
&= \Pr\left(\gamma_{\mathsf{sr}} < \Gamma_1(2^{2R_1}, \alpha_2)\right) && (2.14b)\\
&= F_{\mathsf{sr}}\left(\Gamma_1(2^{2R_1}, \alpha_2)\right) && (2.14c)
\end{aligned}
$$

where F_{sr} denoted the cumulative distribution function (CDF) of γ_{sr}.

$$
\begin{aligned}
P_{\mathsf{r}1} &= \Pr\left(R_1 \le \frac{1}{2}\mathcal{C}_1(\gamma_{\mathsf{sr}}, \alpha_2) \text{ and } R_2 > \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{sr}}, \alpha_2)\right) && (2.15a)\\
&= \Pr\left(\Gamma_1(2^{2R_1}, \alpha_2) \le \gamma_{\mathsf{sr}} < \Gamma_2(2^{2R_2}, \alpha_2)\right) && (2.15b)\\
&= F_{\mathsf{sr}}\left(\max\left(\Gamma_1(2^{2R_1}, \alpha_2), \Gamma_2(2^{2R_2}, \alpha_2)\right)\right) \\
&\qquad - F_{\mathsf{sr}}\left(\Gamma_1(2^{2R_1}, \alpha_2)\right) && (2.15c)
\end{aligned}
$$

The maximum of $\Gamma_1(2^{2R_1}, \alpha_2)$ and $\Gamma_2(2^{2R_2}, \alpha_2)$ is taken in order to take into consideration the cases when $\Gamma_2(2^{2R_2}, \alpha_2) < \Gamma_1(2^{2R_1}, \alpha_2)$. In this case, the probability of decoding L_1 only is zero.

$$
\begin{aligned}
P_{\mathsf{r}2} &= \Pr\left(R_1 \le \frac{1}{2}\mathcal{C}_1(\gamma_{\mathsf{sr}}, \alpha_2) \text{ and } R_2 \le \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{sr}}, \alpha_2)\right) && (2.16a)\\
&= \Pr\left(\gamma_{\mathsf{sr}} \ge \max\left(\Gamma_1(2^{2R_1}, \alpha_2), \Gamma_2(2^{2R_2}, \alpha_2)\right)\right) && (2.16b)\\
&= 1 - F_{\mathsf{sr}}\left(\max\left(\Gamma_1(2^{2R_1}, \alpha_2), \Gamma_2(2^{2R_2}, \alpha_2)\right)\right) && (2.16c)
\end{aligned}
$$

2.3.2 Case: The Relay cannot decode any Layer

Next, we characterize the conditional probabilities starting with the case when the relay is not able to decode any layer. In this case, the source re-transmits L_1 and L_2 using new codewords and with the same power ratios α_1 and α_2.

$$
\begin{aligned}
P_{\mathsf{d1|r0}} &= \Pr\left(R_1 \le \mathcal{C}_1(\gamma_{\mathsf{sd}}, \alpha_2) \text{ and } R_2 > \mathcal{C}_2(\gamma_{\mathsf{sd}}, \alpha_2)\right) && \text{(2.17a)}\\
&= \Pr\left(\Gamma_1(2^{R_1}, \alpha_2) \le \gamma_{\mathsf{sd}} < \Gamma_2(2^{R_2}, \alpha_2)\right) && \text{(2.17b)}\\
&= F_{\mathsf{sd}}\left(\max\left(\Gamma_1(2^{R_1}, \alpha_2), \Gamma_2(2^{R_2}, \alpha_2)\right)\right)\\
&\qquad - F_{\mathsf{sd}}\left(\Gamma_1(2^{R_1}, \alpha_2)\right) && \text{(2.17c)}
\end{aligned}
$$

$$
\begin{aligned}
P_{\mathsf{d2|r0}} &= \Pr\left(R_1 \le \mathcal{C}_1(\gamma_{\mathsf{sd}}, \alpha_2) \text{ and } R_2 \le \mathcal{C}_2(\gamma_{\mathsf{sd}}, \alpha_2)\right) && \text{(2.18a)}\\
&= \Pr\left(\gamma_{\mathsf{sd}} \ge \max\left(\Gamma_1(2^{R_1}, \alpha_2), \Gamma_2(2^{R_2}, \alpha_2)\right)\right) && \text{(2.18b)}\\
&= 1 - F_{\mathsf{sd}}\left(\max\left(\Gamma_1(2^{R_1}, \alpha_2), \Gamma_2(2^{R_2}, \alpha_2)\right)\right) && \text{(2.18c)}
\end{aligned}
$$

2.3.3 Case: The Relay can decode Only One Layer

In this case when the relay can decode only L_1, it transmits only this layer using a new codeword in the second time slot. Therefore, L_1 is allocated the full power of the relay in this case.

$$
P_{\mathsf{d1|r1}} = \Pr\left(R_1 \le \frac{1}{2}\mathcal{C}_1(\gamma_{\mathsf{sd}}, \alpha_2) + \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{rd}}, 1) \text{ and } R_2 > \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{sd}}, \alpha_2)\right) \tag{2.19}
$$

The region defined by the condition $R_1 \le \frac{1}{2}\mathcal{C}_1(\gamma_{\mathsf{sd}}, \alpha_2) + \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{rd}}, 1)$ can be characterized in terms of γ_{sd} and γ_{rd} as

$$
\gamma_{\mathsf{rd}} \ge \max\left(0, \Gamma_2\left(\frac{2^{2R_1}(1+\alpha_2\gamma_{\mathsf{sd}})}{1+\gamma_{\mathsf{sd}}}, 1\right)\right) \tag{2.20}
$$

Equivalently, this region can be characterized as

$$\gamma_{\mathsf{sd}} \geq \max\left(0, \Gamma_1\left(\frac{2^{2R_1}}{1+\gamma_{\mathsf{rd}}}, \alpha_2\right)\right) \tag{2.21}$$

Furthermore, the condition $R_2 > \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{sd}}, \alpha_2)$ is equivalent to

$$\gamma_{\mathsf{sd}} < \Gamma_2\left(2^{2R_2}, \alpha_2\right) \tag{2.22}$$

Therefore, based on (2.20) and (2.22), we can write (2.19) as

$$\begin{aligned} P_{\mathsf{d1|r1}} &= \int_0^{\Gamma_2\left(2^{2R_2},\alpha_2\right)} f_{\mathsf{sd}}(\gamma_{\mathsf{sd}}) \int_{\omega(\gamma_{\mathsf{sd}})}^{\infty} f_{\mathsf{rd}}(\gamma_{\mathsf{rd}}) d\gamma_{\mathsf{rd}} d\gamma_{\mathsf{sd}} && \text{(2.23a)}\\ &= \int_0^{\Gamma_2\left(2^{2R_2},\alpha_2\right)} f_{\mathsf{sd}}(\gamma_{\mathsf{sd}}) \left(1 - F_{\mathsf{rd}}\left(\omega(\gamma_{\mathsf{sd}})\right)\right) d\gamma_{\mathsf{sd}}, && \text{(2.23b)} \end{aligned}$$

where $\omega(\gamma_{\mathsf{sd}})$ is defined as

$$\omega(\gamma_{\mathsf{sd}}) = \max\left(0, \Gamma_2\left(\frac{2^{2R_1}(1+\alpha_2\gamma_{\mathsf{sd}})}{1+\gamma_{\mathsf{sd}}}, 1\right)\right) \tag{2.24}$$

Equivalently, based on (2.21) and (2.22), we can write (2.19) as

$$P_{\mathsf{d1|r1}} = \int_0^{\infty} f_{\mathsf{rd}}(\gamma_{\mathsf{rd}}) \int_{\zeta(\gamma_{\mathsf{rd}})}^{\Gamma_2\left(2^{2R_2},\alpha_2\right)} f_{\mathsf{sd}}(\gamma_{\mathsf{sd}}) d\gamma_{\mathsf{sd}} d\gamma_{\mathsf{rd}}, \tag{2.25}$$

where $\zeta(\gamma_{\mathsf{rd}})$ is defined as

$$\zeta(\gamma_{\mathsf{rd}}) = \min\left(\Gamma_2\left(2^{2R_2}, \alpha_2\right), \max\left(0, \Gamma_1\left(\frac{2^{2R_1}}{1+\gamma_{\mathsf{rd}}}, \alpha_2\right)\right)\right) \tag{2.26}$$

In a similar way, we can obtain $P_{\mathsf{d2|r1}}$

$$\begin{aligned} P_{\mathsf{d2|r1}} &= \Pr\left(R_1 \leq \frac{1}{2}\mathcal{C}_1(\gamma_{\mathsf{sd}}, \alpha_2) + \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{rd}}, 1) \text{ and } R_2 \leq \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{sd}}, \alpha_2)\right) && \text{(2.27a)}\\ &= \int_{\Gamma_2\left(2^{2R_2},\alpha_2\right)}^{\infty} f_{\mathsf{sd}}(\gamma_{\mathsf{sd}}) \left(1 - F_{\mathsf{rd}}\left(\omega(\gamma_{\mathsf{sd}})\right)\right) d\gamma_{\mathsf{sd}}, && \text{(2.27b)} \end{aligned}$$

where $\omega(\gamma_{\mathsf{sd}})$ is defined in (2.24). Equivalently, we can characterize $P_{\mathsf{d2|r1}}$ as

$$P_{\mathsf{d2|r1}} = \int_0^\infty f_{\mathsf{rd}}(\gamma_{\mathsf{rd}}) \int_{\eta(\gamma_{\mathsf{rd}})}^\infty f_{\mathsf{sd}}(\gamma_{\mathsf{sd}}) d\gamma_{\mathsf{sd}} d\gamma_{\mathsf{rd}}, \tag{2.28}$$

where $\eta(\gamma_{\mathsf{rd}})$ is defined as

$$\eta(\gamma_{\mathsf{rd}}) = \max\left(0, \Gamma_1\left(\frac{2^{2R_1}}{1+\gamma_{\mathsf{rd}}}, \alpha_2\right), \Gamma_2\left(2^{2R_2}, \alpha_2\right)\right) \tag{2.29}$$

2.3.4 Case: The Relay can decode Both Layers

When the relay can decode both layers, it forwards both layers using power ratios β_1 and β_2, respectively. The conditional probabilities in this case are characterized as follows.

$$\begin{aligned} P_{\mathsf{d1|r2}} = \Pr\Big(R_1 \leq \frac{1}{2}\mathcal{C}_1(\gamma_{\mathsf{sd}}, \alpha_2) + \frac{1}{2}\mathcal{C}_1(\gamma_{\mathsf{rd}}, \beta_2) \text{ and} \\ R_2 > \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{sd}}, \alpha_2) + \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{rd}}, \beta_2)\Big) \end{aligned} \tag{2.30}$$

The region defined by the condition $R_1 \leq \frac{1}{2}\mathcal{C}_1(\gamma_{\mathsf{sd}}, \alpha_2) + \frac{1}{2}\mathcal{C}_1(\gamma_{\mathsf{rd}}, \beta_2)$ can be characterized in terms of γ_{sd} and γ_{rd} as

$$\gamma_{\mathsf{rd}} \geq \max\left(0, \Gamma_1\left(\frac{2^{2R_1}(1+\alpha_2\gamma_{\mathsf{sd}})}{1+\gamma_{\mathsf{sd}}}, \beta_2\right)\right) \tag{2.31}$$

Equivalently, this region can be characterized as

$$\gamma_{\mathsf{sd}} \geq \max\left(0, \Gamma_1\left(\frac{2^{2R_1}(1+\beta_2\gamma_{\mathsf{rd}})}{1+\gamma_{\mathsf{rd}}}, \alpha_2\right)\right) \tag{2.32}$$

Similarly, the region defined by the condition $R_2 > \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{sd}}, \alpha_2) + \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{rd}}, \beta_2)$ can be characterized in terms of γ_{sd} and γ_{rd} as

$$\gamma_{\mathsf{rd}} < \max\left(0, \Gamma_2\left(\frac{2^{2R_2}}{1+\alpha_2\gamma_{\mathsf{sd}}}, \beta_2\right)\right) \tag{2.33}$$

Equivalently, this region can be characterized as

$$\gamma_{\mathsf{sd}} < \max\left(0, \Gamma_2\left(\frac{2^{2R_2}}{1+\beta_2\gamma_{\mathsf{rd}}}, \alpha_2\right)\right) \tag{2.34}$$

Therefore, based on (2.31) and (2.33), we can write (2.30) as

$$\begin{aligned} P_{\mathsf{d1|r2}} = \int_0^{\Gamma_2\left(2^{2R_2},\alpha_2\right)} & f_{\mathsf{sd}}(\gamma_{\mathsf{sd}})\Big(F_{\mathsf{rd}}\left(\max\left(\sigma(\gamma_{\mathsf{sd}}), \xi(\gamma_{\mathsf{sd}})\right)\right) \\ & - F_{\mathsf{rd}}\left(\sigma(\gamma_{\mathsf{sd}})\right)\Big) d\gamma_{\mathsf{sd}}, \end{aligned} \tag{2.35a}$$

where $\sigma(\gamma_{\mathsf{sd}})$ is defined as

$$\sigma(\gamma_{\mathsf{sd}}) = \max\left(0, \Gamma_1\left(\frac{2^{2R_1}(1+\alpha_2\gamma_{\mathsf{sd}})}{1+\gamma_{\mathsf{sd}}}, \beta_2\right)\right), \tag{2.36}$$

and $\xi(\gamma_{\mathsf{sd}})$ is defined as

$$\xi(\gamma_{\mathsf{sd}}) = \max\left(0, \Gamma_2\left(\frac{2^{2R_2}}{1+\alpha_2\gamma_{\mathsf{sd}}}, \beta_2\right)\right) \tag{2.37}$$

Equivalently, based on (2.32) and (2.34), we can write (2.30) as

$$\begin{aligned} P_{\mathsf{d1|r2}} = \int_0^{\Gamma_2\left(2^{2R_2},\beta_2\right)} & f_{\mathsf{rd}}(\gamma_{\mathsf{rd}})\Big(F_{\mathsf{sd}}\left(\max\left(\theta(\gamma_{\mathsf{rd}}), \phi(\gamma_{\mathsf{rd}})\right)\right) \\ & - F_{\mathsf{sd}}\left(\theta(\gamma_{\mathsf{rd}})\right)\Big) d\gamma_{\mathsf{rd}}, \end{aligned} \tag{2.38a}$$

where $\theta(\gamma_{\mathsf{rd}})$ is defined as

$$\theta(\gamma_{\mathsf{rd}}) = \max\left(0, \Gamma_1\left(\frac{2^{2R_1}(1+\beta_2\gamma_{\mathsf{rd}})}{1+\gamma_{\mathsf{rd}}}, \alpha_2\right)\right), \tag{2.39}$$

and $\phi(\gamma_{\mathsf{rd}})$ is defined as

$$\phi(\gamma_{\mathsf{rd}}) = \max\left(0, \Gamma_2\left(\frac{2^{2R_2}}{1+\beta_2\gamma_{\mathsf{rd}}}, \alpha_2\right)\right) \tag{2.40}$$

In a similar way, we can obtain $P_{\mathsf{d2|r2}}$

$$P_{\mathsf{d2|r2}} = \Pr\Big(R_1 \le \frac{1}{2}\mathcal{C}_1(\gamma_{\mathsf{sd}}, \alpha_2) + \frac{1}{2}\mathcal{C}_1(\gamma_{\mathsf{rd}}, \beta_2) \text{ and } R_2 \le \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{sd}}, \alpha_2) + \frac{1}{2}\mathcal{C}_2(\gamma_{\mathsf{rd}}, \beta_2)\Big) \tag{2.41a}$$

$$= \int_0^\infty f_{\mathsf{sd}}(\gamma_{\mathsf{sd}})\Big(1 - F_{\mathsf{rd}}\Big(\max\left(\sigma(\gamma_{\mathsf{sd}}), \xi(\gamma_{\mathsf{sd}})\right)\Big)\Big)d\gamma_{\mathsf{sd}} \tag{2.41b}$$

where $\sigma(\gamma_{\mathsf{sd}})$ and $\xi(\gamma_{\mathsf{sd}})$ are defined in (2.36) and (2.37). Equivalently, we can characterize $P_{\mathsf{d2|r2}}$ as

$$P_{\mathsf{d2|r2}} = \int_0^\infty f_{\mathsf{rd}}(\gamma_{\mathsf{rd}})\Big(1 - F_{\mathsf{sd}}\left(\max\left(\theta(\gamma_{\mathsf{rd}}), \phi(\gamma_{\mathsf{rd}})\right)\right)\Big)d\gamma_{\mathsf{rd}}, \tag{2.42a}$$

where $\theta(\gamma_{\mathsf{rd}})$ and $\phi(\gamma_{\mathsf{rd}})$ are defined in (2.39) and (2.40).

2.4 Numerical Results

We present several numerical results in this section with the assumption that the fading distribution of the channels follows (2.2), where the average SNR for the source-relay and the relay-destination links are m times the average SNR for the source-destination link. We consider a scalable video source example consisting of two layers with a sum rate of 3 bps/Hz. The rates of the two source layers are respectively 1 and 2 bps/Hz. We consider two different utility functions; namely, $U(\bar{R}) = 1 - 2^{-2\bar{R}}$, which corresponds to minimizing the expected distortion of a Gaussian source and $U(\bar{R}) = \bar{R}$, which corresponds to maximizing the expected total rate at the destination.

Figures 2.1 and 2.2 show the relative power ratios of the layers at the source and the relay respectively, for different values of m, with the target of minimizing the expected distortion of a Gaussian source over Rayleigh fading channels. The optimal power ratios are plotted against the average SNR of the source-destination channel. It can be seen that for low average SNR values it is optimal to send only one layer. This is because the enhancement layer cannot be decoded reliably in this case. Therefore, it is better to discard it in order to get rid of its interference on the base layer, which enables the reception of the base layer at lower SNR values. On the other hand, when the average SNR is above a certain value, it becomes

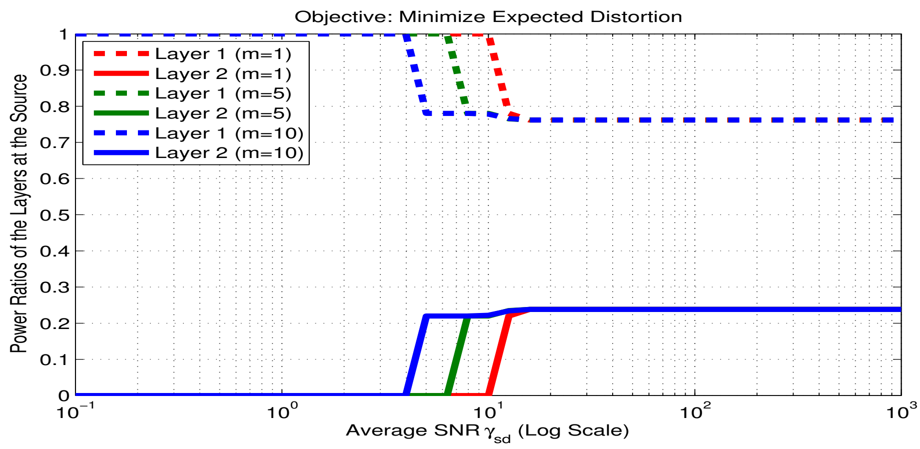


FIGURE 2.1: The relative power ratios of the layers at the source versus the average SNR value of a Rayleigh fading channel with $U(\bar{R}) = 1 - 2^{-2\bar{R}}$.

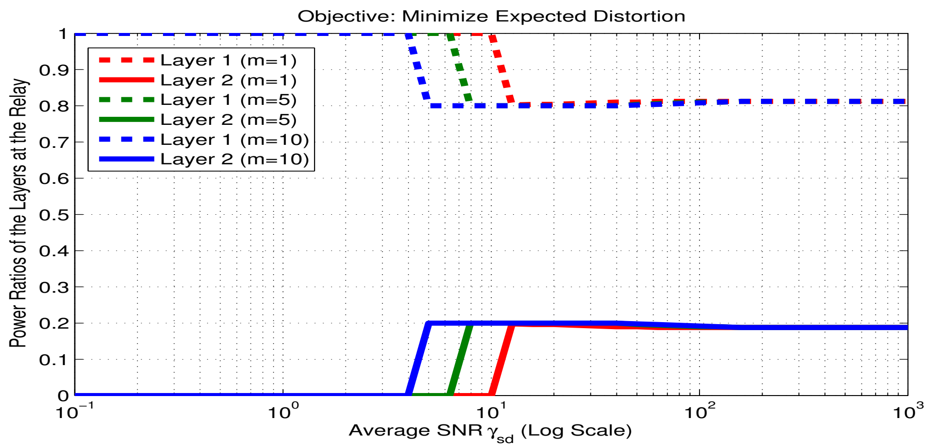


FIGURE 2.2: The relative power ratios of the layers at the relay versus the average SNR value of a Rayleigh fading channel with $U(\bar{R}) = 1 - 2^{-2\bar{R}}$.

optimal to send the two layers. This value for the average SNR is the same for the source and the relay. It is obvious that as the ratio m increases, the curves are shifted to the left, which means that for higher values of m it is optimal to send the two layers for lower values of the average SNR. This is intuitive because as m increases, the relay becomes more capable of enhancing the end-to-end performance, and hence the destination becomes more capable of decoding the enhancement layer even when its direct channel with the source has low SNR.

Figures 2.3 and 2.4 show the relative power ratios of the layers at the source and the relay respectively, for different values of m, with the target of maximizing the expected rate over Rayleigh fading channels. Since the utility function for maximizing the rate is linear, the

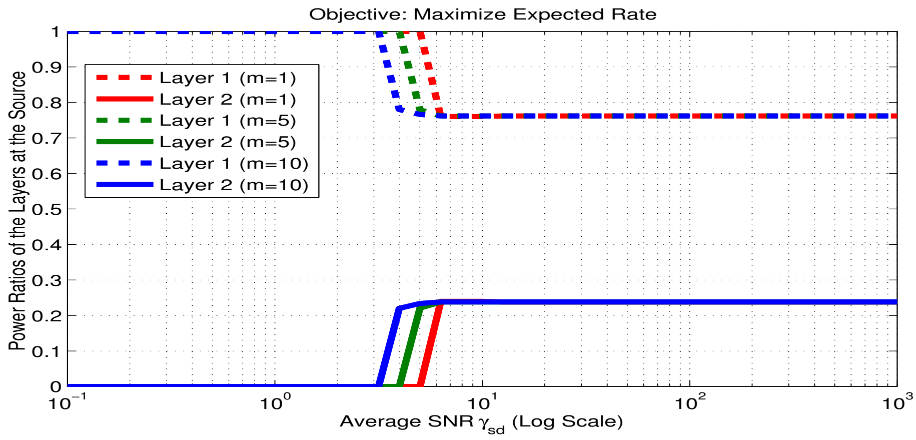


FIGURE 2.3: The relative power ratios of the layers at the source versus the average SNR value of a Rayleigh fading channel with $U(\bar{R}) = \bar{R}$.

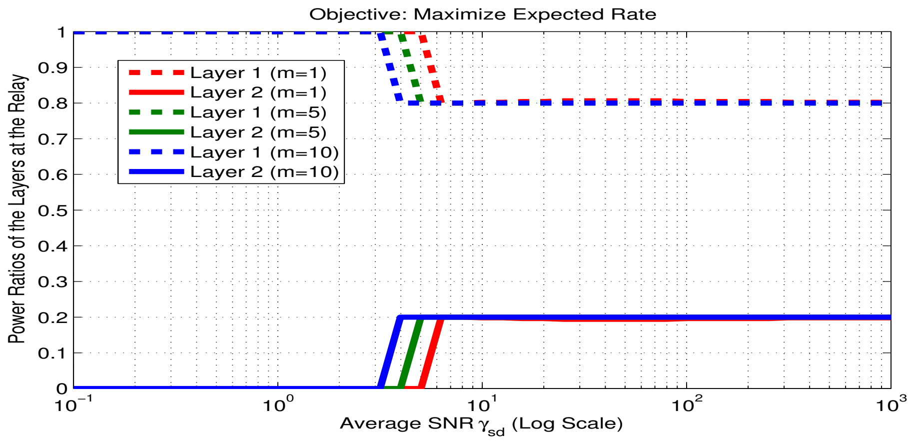


FIGURE 2.4: The relative power ratios of the layers at the relay versus the average SNR value of a Rayleigh fading channel with $U(\bar{R}) = \bar{R}$.

enhancement layer has more importance than in the case of distortion minimization. Consequently, a higher ratio of the power is allocated to the enhancement layer in this case. Furthermore, it becomes optimal to send both layers for lower values of the average SNR. In comparison, the solution for minimizing the average distortion gives more importance to the base layer, and hence it becomes optimal to send both layers for higher values of the average SNR. That is why the curves are shifted to the left in Figures 2.3 and 2.4 compared to Figures 2.1 and 2.2. Moreover, it is obvious that as the the ratio m increases, the curves are shifted to the left similar to the distortion minimization case.

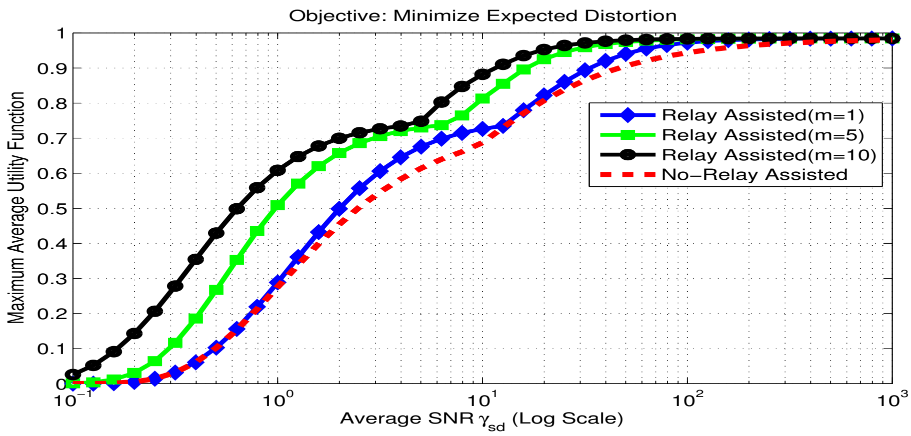


FIGURE 2.5: The maximized average utility function versus the average SNR value of a Rayleigh fading channel with $U(\bar{R}) = 1 - 2^{-2\bar{R}}$.

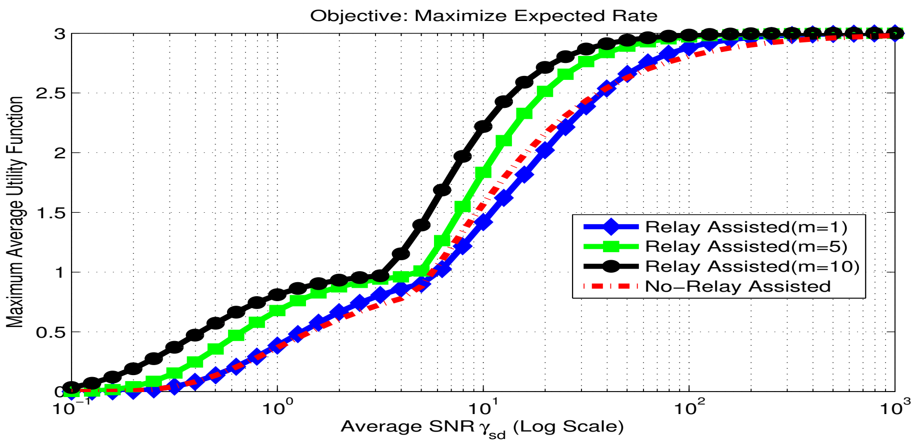


FIGURE 2.6: The maximized average utility function versus the average SNR value of a Rayleigh fading channel with $U(\bar{R}) = \bar{R}$.

In a comparison between relay-assisted transmission and direct channel with no relay assistance, over Rayleigh fading channels with different values of m, we can see from Figures 2.5 and 2.6 that there is an obvious gain in the maximum expected utility when the relay is involved. This is valid for both cases $U(\bar{R}) = 1 - 2^{-2\bar{R}}$ and $U(\bar{R}) = \bar{R}$. Furthermore, as m increases, the gain with respect to the no-relay case increases as well, as expected. For the case when $m = 1$, it can be seen that the maximum expected utility is close (and maybe less than) the no-relay case. This is because the channel gains of the relay channel are not high in this case. Therefore, the prospected gain due to channel diversity of the relay channel will be opposed by the multiplexing loss due to the transmission over two time slots.

Chapter 3

Optimal Power Allocation for Layered Broadcast Over AF Relay Channels

3.1 Introduction

In this chapter we are interested in the application of multilayer transmission using the broadcast approach, and how to optimally allocate power among the layers in the context of AF relay networks. This work is already published in [18], where we extended the solution algorithm presented in [19] for fixed and predetermined rates of the source layers. In [19] and originally in [28], [29], the problem of optimal power allocation for multilayer transmission was considered with fixed and predetermined rates for the source layers.

We also show in this chapter, unlike the DF relay case presented in Chapter 2, the application of the algorithm presented in [19] to solve the optimization problem assuming AF relaying is feasible. This means that we can solve both the optimal power allocation problem for the fixed rate case and the joint rate and power allocation problem for any number of source layers while maintaining a linear computation complexity with respect to the number of source layers. On the other hand, in the DF case the solution is obtained using numerical random search methods and for two source layers only [17]. The extension into more than two layers in the DF relay case becomes prohibitively more complex. Notice that the expected utility function in the AF case is a function of the channel statistics of the three links in the channel (i.e., source-destination, source-relay and relay-destination). So, we need to analytically characterize the end-to-end channel statistics in terms of the statistics of the three links of the channel model in order to be able to apply the algorithm presented in [19]. This is the main bottleneck in the problem. However, we proposed a simple and useful approximation of the end-to-end channel quality given that all three links in our channel model are Rayleigh faded. Furthermore, we provided several numerical examples to show both the optimal power allocation for the fixed rate case and the joint optimal rate and power allocation for two different utility functions and to demonstrate the gains of relaying over the case when the relay is not utilized.

3.2 System Model and Problem Formulation

3.2.1 System Model and Transmission Scheme

We consider a system that consists of three nodes: source, destination and relay. We assume that the source signal is Gaussian and it is encoded into independent M layers, $\mathbf{L} = [L_1, L_2, \ldots, L_M]$, with fixed rates $\mathbf{R} = [R_1, R_2, \ldots, R_M]$, with power ratios $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_M]$ of the total source power P_s, and with each layer successively refining the information from the lower layers. Therefore, the source transmits layer L_i with a power $P_i = \alpha_i P_s$. The relay is half-duplex and applies AF strategy [21]. Therefore, the transmission is carried over two consecutive time slots of equal duration and bandwidth. The source broadcasts the layers to the relay and the destination using superposition coding in the first time slot. In the second time slot, the relay forwards the signal that was received from the source after amplifying it. The power of the relay is denoted by P_r. Notice that the power ratios of the source layers at the relay preserve the same ratios like the source node since the relay just amplifies the layers without decoding and regenerating them.

Two copies of the layers are received at the destination in the two time slots. The destination utilizes both copies in order to decode the source information up to the number of layers that can be decoded reliably based on the end-to-end instantaneous channel quality. The layers are decoded with SIC. Thus, in order for the destination to decode layer L_i, it must be able to decode all "higher priority" layers first (i.e., all L_j where $j < i$).

We assume that the three nodes are equipped with a single antenna. We also denote the SNR over the three links of the relay channel using γ_{sr}, γ_{sd} and γ_{rd} for the source-relay, source-destination, and relay-destination links, respectively. We assume that the source and the relay transmit using constant power. Furthermore, we assume that the channel gains, and consequently the SNRs, stay constant for the duration of one transmission block, which consists of two consecutive time slots. However, γ_{sr}, γ_{sd} and γ_{rd} vary from one channel block to another randomly. Furthermore, we assume that the source and the relay do not know the instantaneous values of the SNRs while transmitting.

In this work, we assume that the variation (i.e. fading) of the channels' gain is Rayleigh distributed. Hence, the PDF of the channels follow an exponential distribution, and they are given as

$$f_{\mathsf{sd}}\left(\gamma_{\mathsf{sd}}\right) = \frac{1}{\bar{\gamma}}\exp\left(\frac{-\gamma_{\mathsf{sd}}}{\bar{\gamma}}\right), \tag{3.1a}$$

$$f_{\mathsf{sr}}\left(\gamma_{\mathsf{sr}}\right) = \frac{1}{m_1\bar{\gamma}}\exp\left(\frac{-\gamma_{\mathsf{sr}}}{m_1\bar{\gamma}}\right), \tag{3.1b}$$

$$f_{\mathsf{rd}}\left(\gamma_{\mathsf{rd}}\right) = \frac{1}{m_2\bar{\gamma}}\exp\left(\frac{-\gamma_{\mathsf{rd}}}{m_2\bar{\gamma}}\right), \tag{3.1c}$$

for the source-destination, source-relay and relay-destination channels, respectively. In (3.1), $\bar{\gamma}$ is the average SNR for the direct source-destination link and m_1 and m_2 denote the ratios between the average SNR of the source-relay and the relay-destination links to the source-destination link, respectively. We assume that $\bar{\gamma}$, m_1 and m_2 are known at the source node which utilizes its knowledge of the average channel qualities of the three links in the optimization of the power allocation α_i's over the layers in order to maximize the expected utility function, denoted U, of the total decoded rate, denoted $\bar{R}$, at the destination.

Since the quality of service of any link increases with the total decoded rate at the receiver, and since the lower layers are more important than the upper (enhancement) layers, we will assume that the utility function $U(\bar{R})$ that represents the user satisfaction is differentiable concave increasing.

3.2.2 End-to-End Channel Condition

The two copies of the layers y_{sd} and y_{rd} received from the source and the relay in the two time slots, respectively, are combined at the destination using *Maximum Ratio Combining MRC*. Therefore, the combined signal can be given as

$$y_c = ay_{sd} + by_{rd}, \tag{3.2}$$

where a and b are the combining ratios, and

$$y_{sd} = h_{sd}\Sigma_{i=1}^{M}L_i + n_{sd}, \tag{3.3a}$$

$$y_{sr} = h_{sr}\Sigma_{i=1}^{M}L_i + n_{sr}, \tag{3.3b}$$

$$y_{rd} = h_{rd}A_r y_{sr} + n_{rd}, \tag{3.3c}$$

where h_{sd}, h_{sr} and h_{rd} are the independent channel gains, n_{sd}, n_{sr} and n_{rd} are the independent noise levels with variance N_o for the three links, and A_r is the amplifying gain at the relay node that is a function of the power constraint at the relay P_r. Hence,

$$A_r = \sqrt{\frac{P_r}{|h_{sr}|^2 P_s + N_o}}. \tag{3.4}$$

It can be shown that the signal to noise ratio $\text{SNR}_c^{(L_i)}$ of the combined signal with SIC for layer L_i can be easily written as

$$\text{SNR}_c^{(L_i)} = \frac{|ah_{sd} + bh_{rd}h_{sr}A_r|^2 \alpha_i P_s}{N_o \left(|a|^2 + |b|^2 + |bh_{rd}A_r|^2\right) + |ah_{sd} + bh_{rd}h_{sr}A_r|^2 \sum_{m>i}^{M} \alpha_m P_s}. \tag{3.5}$$

In order to get the MRC, we need to find the combining ratios a and b that will maximize $\text{SNR}_c^{(L_i)}$. Therefore, we differentiate $\text{SNR}_c^{(L_i)}$ with respect to a^*, where a^* is the conjugate value of the combining ratio a, and find the values of a and b for the derivative $\frac{\delta \text{SNR}_c^{(L_i)}}{\delta a^*}$ to be equal zero, which can be found after some mathematical derivations, that are omitted here for brevity, as

$$a = Ch_{sd}^* \left(|h_{sr}|^2 P_s + |h_{rd}|^2 P_r + N_o\right), \tag{3.6a}$$
$$b = Ch_{sr}^* h_{rd}^* \sqrt{P_r \left(|h_{sr}|^2 P_s + N_o\right)}, \tag{3.6b}$$

where C is an arbitrary constant, and h^* is the conjugate value of the channel gain h. By substituting with (3.6) in (3.5), we can find the maximum SNR value for the layer L_i denoted by $\text{SNR}_{\text{MRC}}^{(L_i)}$, which yields

$$\text{SNR}_{\text{MRC}}^{(L_i)} = \frac{\alpha_i}{\frac{1}{\gamma} + \sum_{m>i}^{M} \alpha_m}, \tag{3.7a}$$
$$\gamma = \gamma_{sd} + \frac{\gamma_{sr}\gamma_{rd}}{\gamma_{sr} + \gamma_{rd} + 1}, \tag{3.7b}$$

where γ denotes the end-to-end SNR (i.e. the SNR at the destination after combining the direct and relayed signals optimally).

Since the receiver decodes the layers one by one using SIC, then for the destination to decode and make use of layer L_i, it must be able to decode this layer as well as all the previous layers.

Therefore, the value of γ must satisfy the relation

$$R_j \leq \frac{1}{2}\log\left(1+\frac{\alpha_j}{\frac{1}{\gamma}+\sum_{m>j}^{M}\alpha_m}\right) \quad \forall j \leq i. \tag{3.8}$$

This can be written as

$$\begin{aligned} \gamma \geq \bar{\gamma}_i &= \max\{\gamma_1, \gamma_2, \ldots, \gamma_i\} \\ &= \max\left\{\bar{\gamma}_{i-1}, \frac{1}{\frac{\alpha_i}{2^{2R_i}-1}-\sum_{m>i}^{M}\alpha_m}\right\}, \end{aligned} \tag{3.9}$$

where $\bar{\gamma}_i$, named as γ threshold, is the constraint on γ for the destination to be able to decode all the layers up to layer L_i, and γ_j is the minimum value for γ required to decode the layer L_j only after correctly cancelling all the previous layers, and can be written as

$$\gamma_j = \frac{1}{\frac{\alpha_j}{2^{2R_j}-1}-\sum_{m>j}^{M}\alpha_m}, \tag{3.10}$$

and hence the required γ decoding threshold would be the maximum of the γ threshold required to decode Layer i and that required to decode all of its previous layers.

It can be seen that the γ threshold values depends on the power allocated to each layer. The destination only decodes the layers whose thresholds are below the instantaneous end-to-end channel condition γ.

3.2.3 Problem Formulation

In this chapter, we need to find the optimal power allocation in order to maximize the expected user satisfaction that is defined by a utility function $U(\bar{R})$ of the total decoded rate $\bar{\mathbf{R}}$ at the destination, where $\bar{R} = \{\bar{R}_1, \bar{R}_2, \ldots, \bar{R}_M\}$, and $\bar{R}_i = \sum_{j=1}^{i} R_j$.

The algorithm used in this work can be utilized to employ many special definitions for the utility function such as minimizing the expected distortion of a Gaussian source or maximizing the expected rate. The optimization problem is to optimally allocate the power among the layers, i.e., $\alpha_i's$, such that we can obtain the maximum value of the expectation of the utility function $E[U(\bar{R})]$. Hence, we can formulate the optimization problem Similar to [19] as follows

$$\max_{\alpha} \quad \int_0^{\infty} f_\gamma(\gamma) \quad U\left(\bar{R}(\gamma,\alpha)\right) d\gamma \tag{3.11a}$$

$$\text{subject to} \qquad \sum_{i=1}^{M} \alpha_i = 1, \qquad \alpha_i \geq 0 \qquad \forall i, \tag{3.11b}$$

where $f_\gamma(\gamma)$ is the PDF of the end-to-end channel quality γ, $\bar{R}(\gamma,\alpha)$ is an indication that the total rate decoded successfully at a certain value of γ is a function of the power ratios $\alpha_i's$ of the layers.

Since the user only decodes the layers whose thresholds are below the received value of γ, then we have M possible levels for the total rate decoded successfully, and hence we can write the objective function in (3.11a) as:

$$\left(U_1 \int_{\bar{\gamma}_1}^{\bar{\gamma}_2} f_\gamma(\gamma) d\gamma + \ldots + U_M \int_{\bar{\gamma}_M}^{\infty} f_\gamma(\gamma) d\gamma\right), \tag{3.12}$$

where $U_i = U\left(\sum_{m=1}^{i} R_m\right)$, and $U_0 = 0$. Now we can write the problem in (3.11) using (3.12) as:

$$\min_{\alpha,\bar{\gamma}} \qquad \sum_{i=1}^{M} c_i \, F_\gamma(\bar{\gamma}_i(\bar{\gamma}_{i-1},\alpha)) \tag{3.13a}$$

$$\text{subject to} \qquad \sum_{i=1}^{M} \alpha_i = 1, \qquad \alpha_i \geq 0 \qquad \forall i, \tag{3.13b}$$

$$\bar{\gamma}_i = \max\left\{\bar{\gamma}_{i-1}, \frac{1}{\frac{\alpha_i}{2^{2R_i}-1} - \sum_{m>i}^{M} \alpha_m}\right\} \quad \forall i, \tag{3.13c}$$

where $c_i = U_i - U_{i-1}$, and $F_\gamma(\gamma)$ is the *Cumulative Distribution Function CDF* of the end-to-end channel quality γ. The constraint (3.13c) can be intuitively replaced with the following two constraints

$$\bar{\gamma}_i \quad \geq \quad \frac{1}{\frac{\alpha_i}{2^{2R_i}-1} - \sum_{m>i}^{M} \alpha_m} \quad \forall i, \tag{3.14a}$$

$$\bar{\gamma}_M \quad \geq \quad \bar{\gamma}_{M-1} \geq \ldots \geq \bar{\gamma}_1 > 0. \tag{3.14b}$$

It can be shown that the two constraints (3.14a) and (3.14b) are sufficient to guarantee that the constraints $\alpha_i \geq 0, \forall i$ in (3.13b) are satisfied. This becomes clear if we start by $i = M$ and apply (3.14a) and (3.14b), we find that α_M must be greater than or equal zero. Then, by applying (3.14a) and (3.14b) for $i = M - 1$, and using the fact that $\alpha_M \geq 0$ from the previous step, we find that α_{M-1} must also be greater than or equal zero. Repeating this process until $i = 1$, we can show that the constraint $\alpha_i \geq 0, \forall i$ is implicitly satisfied by the constraints (3.14a) and (3.14b).

Notice that the optimization problem in (3.13) has two variables α, and $\bar{\gamma}$. Since each variable is a function of the other one $\bar{\gamma}(\alpha)$, then for simplicity we can eliminate the variable $\bar{\gamma}$ from the optimization problem. Similar to what have been done in [19], it was shown that the constraint (3.14a) must be satisfied with equality, and by applying some simple manipulations on the constraint as in Appendix A, it can be shown that the optimization problem in (3.13) can be written as

$$\min_{\alpha,\bar{\gamma}} \qquad \sum_{i=1}^{M} c_i \, F_\gamma(\bar{\gamma}_i), \tag{3.15a}$$

$$\text{subject to} \qquad \sum_{i=1}^{M} \frac{\Delta b_i}{\bar{\gamma}_i} - 1 \leq 0, \tag{3.15b}$$

where $b_i = 2^{2\bar{R}_i} - 1$, $b_0 = 0$, and $\Delta b_i = b_i - b_{i-1}$. The constraint (3.14b) is removed, and will be implicitly considered when searching for the optimal values for the thresholds $\bar{\gamma}_i' s$.

Now there is only one optimization parameter for the problem $\bar{\gamma}_i$, and after solving the problem (3.15), we can calculate the optimal power ratios for each layer recursively starting from α_1 using the following relation derived in [19],

$$\alpha_i = \frac{b_i - b_{i-1}}{1 + b_i} \left(1 - \sum_{j=1}^{i-1} \alpha_j + \frac{1}{\bar{\gamma}_i} \right). \tag{3.16}$$

TABLE 3.1: k values for different values of $m_1\bar{\gamma}$, and $m_2\bar{\gamma}$

$m_1\bar{\gamma}, m_2\bar{\gamma}$	<0.5	0.5	0.8	1.2	2.2	4	7	10	25	50	70	150	250	500	1000	2000	6500	≥20000
<0.5	0.4	0.4	0.45	0.45	0.55	0.625	0.7	0.75	0.85	0.9	0.925	0.95	0.975	1	1	1	1	1
0.5	0.4	0.4	0.45	0.45	0.55	0.625	0.7	0.75	0.85	0.9	0.925	0.95	0.975	1	1	1	1	1
0.8	0.45	0.45	0.45	0.5	0.55	0.6	0.675	0.725	0.825	0.9	0.9	0.95	0.95	1	1	1	1	1
1.2	0.45	0.45	0.5	0.5	0.55	0.6	0.65	0.7	0.8	0.875	0.9	0.95	0.95	0.975	1	1	1	1
2.2	0.55	0.55	0.55	0.55	0.55	0.575	0.625	0.65	0.75	0.825	0.85	0.9	0.95	0.95	0.975	1	1	1
4	0.625	0.625	0.6	0.6	0.575	0.575	0.6	0.625	0.7	0.775	0.8	0.875	0.9	0.95	0.95	0.975	1	1
7	0.7	0.7	0.675	0.65	0.625	0.6	0.6	0.6	0.675	0.725	0.75	0.85	0.875	0.925	0.95	0.95	1	1
10	0.75	0.75	0.725	0.7	0.65	0.625	0.6	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	0.95	1	1
25	0.85	0.85	0.825	0.8	0.75	0.7	0.675	0.65	0.675	0.675	0.7	0.75	0.8	0.85	0.9	0.95	1	1
50	0.9	0.9	0.9	0.875	0.825	0.775	0.725	0.7	0.675	0.65	0.675	0.7	0.75	0.8	0.875	0.9	1	1
70	0.925	0.925	0.9	0.9	0.85	0.8	0.75	0.75	0.7	0.675	0.675	0.7	0.75	0.8	0.85	0.9	0.95	1
150	0.95	0.95	0.95	0.95	0.9	0.875	0.85	0.8	0.75	0.7	0.7	0.7	0.75	0.8	0.85	0.9	0.95	1
250	0.975	0.975	0.95	0.95	0.95	0.9	0.875	0.85	0.8	0.75	0.75	0.75	0.75	0.8	0.85	0.9	0.95	1
500	1	1	1	0.975	0.95	0.95	0.925	0.9	0.85	0.8	0.8	0.8	0.8	0.8	0.85	0.9	0.95	1
1000	1	1	1	1	0.975	0.95	0.95	0.95	0.9	0.875	0.85	0.85	0.85	0.85	0.85	0.9	0.95	1
2000	1	1	1	1	1	0.975	0.95	0.95	0.95	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.95	1
6500	1	1	1	1	1	1	1	1	1	1	0.95	0.95	0.95	0.95	0.95	0.95	0.95	1
≥20000	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

3.3 End-To-End Channel Approximation

Since obtaining the exact formula for the PDF of γ is not straightforward. So, alternatively, we propose to use an approximation for it. We aim in this Section to find the PDF (or equivalently CDF) for γ, given in (3.7b), in terms of the PDFs of γ_{sr}, γ_{sd} and γ_{rd}.

It can be easily shown that the value of γ can be bounded as

$$\gamma_{sd} < \gamma \leq \gamma_{sd} + \min\left(\gamma_{sr}, \gamma_{rd}\right). \tag{3.17}$$

The upper bound is obtained when the average SNR value for the source-relay link $m_1\bar{\gamma}$ or the relay destination link $m_2\bar{\gamma}$ (or both) is considerably high (around 10^4). While the lower bound is the case of low values (around 0.5) for both $m_1\bar{\gamma}$ and $m_2\bar{\gamma}$.

So, intuitively, we can in general rewrite the definition of γ approximately as

$$\gamma \approx \gamma_{sd} + k \ \min\left(\gamma_{sr}, \gamma_{rd}\right), \tag{3.18}$$

where the appropriate value for k should be used ($0 < k \leq 1$) such that the CDF of γ as defined in (3.18) becomes as close as possible to the exact CDF of γ as defined in (3.7b). We have done this task for different values of m_1 and m_2, and we have constructed Table 3.1 to show the most appropriate values for k for different values of $m_1\bar{\gamma}$, and $m_2\bar{\gamma}$ to get as close as possible approximation for the CDF of γ.

Notice that for $m_1\bar{\gamma} < 0.5$ or $m_2\bar{\gamma} < 0.5$, we find that we can choose any value for $0 < k \leq 1$ and get very close approximation for γ. Therefore, while constructing Table 3.1, we treated the case for $m_1\bar{\gamma} < 0.5$ as $m_1\bar{\gamma} = 0.5$ ($m_2\bar{\gamma} < 0.5$ as $m_2\bar{\gamma} = 0.5$). We assume that the PDF of the channels follows an exponential distribution as shown in (3.1).

Based on the proposed approximation formula, γ will be the sum of two independent exponential random variables, then we can easily write the CDF using the definition in (3.18) as

$$F_{\gamma}(\gamma) = 1 - \frac{\beta_1}{\beta_1 - \beta'} e^{-\gamma\beta'} + \frac{\beta'}{\beta_1 - \beta'} e^{-\gamma\beta_1}, \tag{3.19}$$

where $\beta' = \frac{\beta_2+\beta_3}{k}$, $\beta_1 = \frac{1}{\bar{\gamma}}$, $\beta_2 = \frac{1}{m_1\bar{\gamma}}$, and $\beta_3 = \frac{1}{m_2\bar{\gamma}}$. Figures 3.1 and 3.2 show the CDF for the approximated γ in (3.19) compared to the CDF of the exact γ as defined in (3.7b), which is obtained numerically, for different values of $\bar{\gamma}$, m_1, and m_2. The two figures demonstrate that the approximation given by (3.18) is appropriate. It can be also noticed that as the values of $m_1\bar{\gamma}$ and $m_2\bar{\gamma}$ decreases, the two curves of $k = 0$ and $k = 1$ get closer, and when the average SNR values decrease below a certain value (around 0.5), the two curves almost coincide. That is why we choose any value of $0 < k \leq 1$ when $m_1\bar{\gamma} < 0.5$ or $m_2\bar{\gamma} < 0.5$ and get very close approximation for γ.

3.4 Solution Structure

In this section we will apply the algorithm that was proposed in [20] to solve this problem. We can notice that the solution without the constraint (3.15b) satisfied with equality is not optimal. The reason is that in this case we can decrease the values of $\gamma_i's$ to have the constraint satisfied with equality while getting a lower value for the objective function. And since the constraint (3.15b) is equivalent to the constraint in (3.13b), then it is optimal to use the total power.

Now we can the write the dual problem as

$$\max_{\lambda} \quad g(\lambda), \tag{3.20}$$

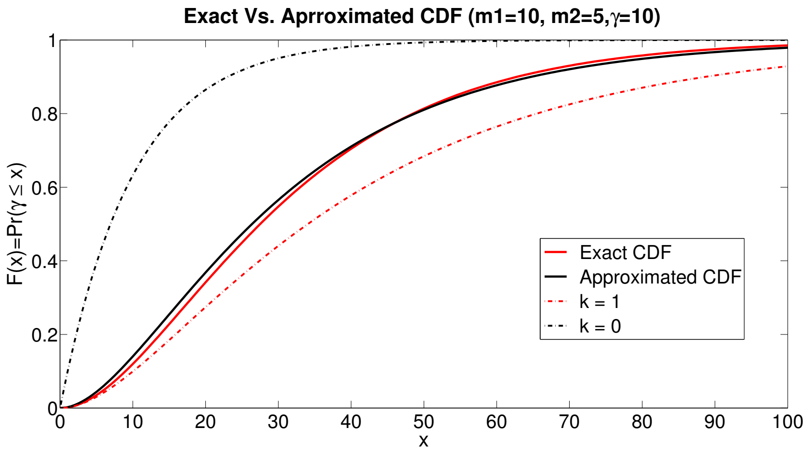


FIGURE 3.1: The approximated CDF Vs. the true CDF for γ with $\bar{\gamma} = 10$, $m_1 = 10$, $m_2 = 5$, and $k = 0.675$.

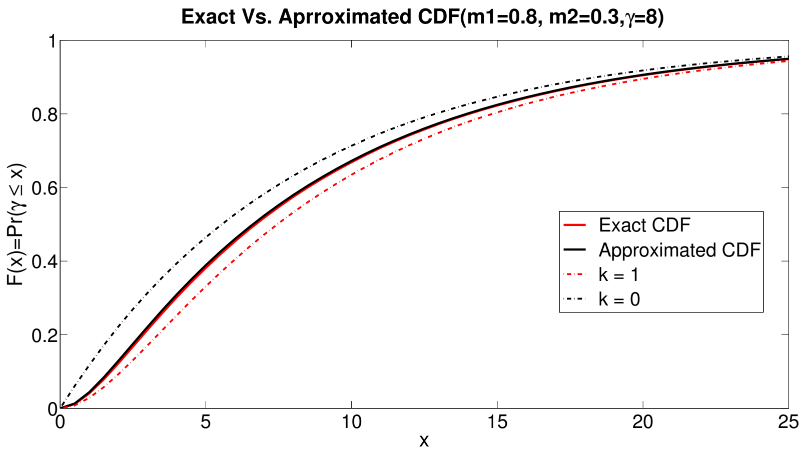


FIGURE 3.2: The approximated CDF Vs. the true CDF for γ with $\bar{\gamma} = 8$, $m_1 = 0.8$, $m_2 = 0.3$, and $k = 0.575$.

where λ is the Lagrangian dual variable, and

$$g(\lambda) = \min_{\bar{\gamma}} \quad L(\bar{\gamma}, \lambda), \tag{3.21}$$

where $L(\bar{\gamma}, \lambda)$ is the Lagrangian of (3.15) and can be written as $L(\bar{\gamma}, \lambda) = \sum_{i=1}^{M} c_i \, F_{\gamma}(\bar{\gamma}_i) + \lambda(\sum_{i=1}^{M} \frac{\Delta b_i}{\bar{\gamma}_i} - 1)$. It is important to notice that the value of $\bar{\gamma}_i's$ minimizing the Lagrangian are constrained to the region in (3.14b). The perturbation function can be defined as the solution

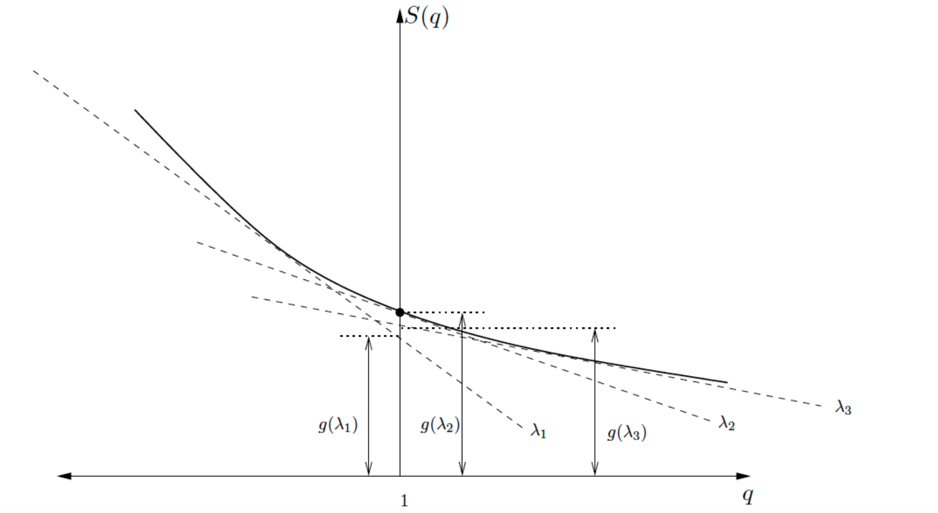


FIGURE 3.3: Changing λ to maximize $g(\lambda)$.

of the following problem as a function of q.

$$S(q) = \min_{\alpha, \bar{\gamma}} \quad \sum_{i=1}^{M} c_i \, F_\gamma(\bar{\gamma}_i), \tag{3.22a}$$

$$\text{subject to} \quad \sum_{i=1}^{M} \frac{\Delta b_i}{\bar{\gamma}_i} \leq q, \tag{3.22b}$$

In Figure 3.3, where the perturbation function is plotted against q, the solution of the primal problem in (3.15) is the intersection of the perturbation curve with the y-axis (i.e., $q = 1$). The value of $g(\lambda)$ can be found by drawing a line tangent to the perturbation curve with slope $-\lambda$, and the value of $g(\lambda)$ is the intersection of the line with the y-axis. By changing λ and noticing the corresponding values of $g(\lambda)$, we can find that the value of $g(\lambda)$ will increase until it reach a maximum value and then decrease again. Therefore, $g(\lambda)$ is a quasi-concave function, and hence we can find the optimal value of λ using a one-dimensional bisection search.

Since the primal problem in (3.15) is not convex, then the solution of the dual problem in (3.20) will be less than or equal the solution for the primal one (i.e., gives a lower bound). In order to find the solution of the primal problem we need to do an exhaustive study for the global perturbation curves.

Since the values of $\bar{\gamma}_i's$ minimizing the Lagrangian for a given value of λ must satisfy the region in (3.14b), then these values must be either on the gradient's null of the Lagrangian (i.e., gradient equals zero) or on one of the region boundaries defined by (3.14b); i.e., $\bar{\gamma}_i = \bar{\gamma}_{i+1}$. In Appendix B, it is shown that the Lagrangian will be increasing towards the region boundaries defined by (3.14b), when the following condition is satisfied

$$\frac{c_m}{\Delta b_m} > \frac{c_{m+1}}{\Delta b_{m+1}}, \forall m \in \{1, 2, \ldots, M-1\} \tag{3.23}$$

This condition is always satisfied if the utility function is concave increasing in the total decoded rate; i.e., $U(\bar{R}_j)$, which is assumed in our case. Therefore, the values of $\gamma_i's$ minimizing the Lagrangian in this case must be inside the region and have the gradient of the Lagrangian equal to zero as follows

$$\frac{\partial L}{\partial \bar{\gamma}_i} = c_i f_\gamma(\bar{\gamma}_i) - \lambda \frac{\Delta b_i}{\bar{\gamma}_i^2} = 0, \tag{3.24a}$$

$$\lambda = \frac{c_i}{\Delta b_i} \bar{\gamma}_i^2 f_\gamma(\bar{\gamma}_i) \quad \forall i = \{1, 2, \ldots, M\}, \tag{3.24b}$$

where $f_\gamma(\bar{\gamma}_i)$ is the PDF of γ defined in (3.18) and can be easily found as the first derivative of the CDF expression given in (3.19) as follows

$$f_\gamma(\gamma) = \frac{\beta_3 \beta'}{\beta_3 - \beta'} (e^{-\gamma\beta'} - e^{-\gamma\beta_3}). \tag{3.25}$$

It can be seen that $\bar{\gamma}_i = \infty$ is a valid solution for (3.24a), which means allocating zero power for the layer i, and that results in allocating zero power for all the layers above because of the constraint in (3.14b). Therefore, It may be optimal not to send all the layers.

Using the relation in (3.24b), the value of λ can be plotted against $\bar{\gamma}_i's$ as seen in Figure 3.4 for $M = 3$. It can be seen that the value of λ is increasing until it reaches a point of maximum λ_i^{max} for a corresponding value $\bar{\gamma}^{(p)}$ then it decreases again. The value of $\bar{\gamma}^{(p)}$ can be found from the relation $\frac{d\lambda}{d\bar{\gamma}} = 0$, which yields

$$\bar{\gamma}^{(p)} = -2 \frac{f_{\bar{\gamma}}(\bar{\gamma}^{(p)})}{f'_{\bar{\gamma}}(\bar{\gamma}^{(p)})}, \tag{3.26}$$

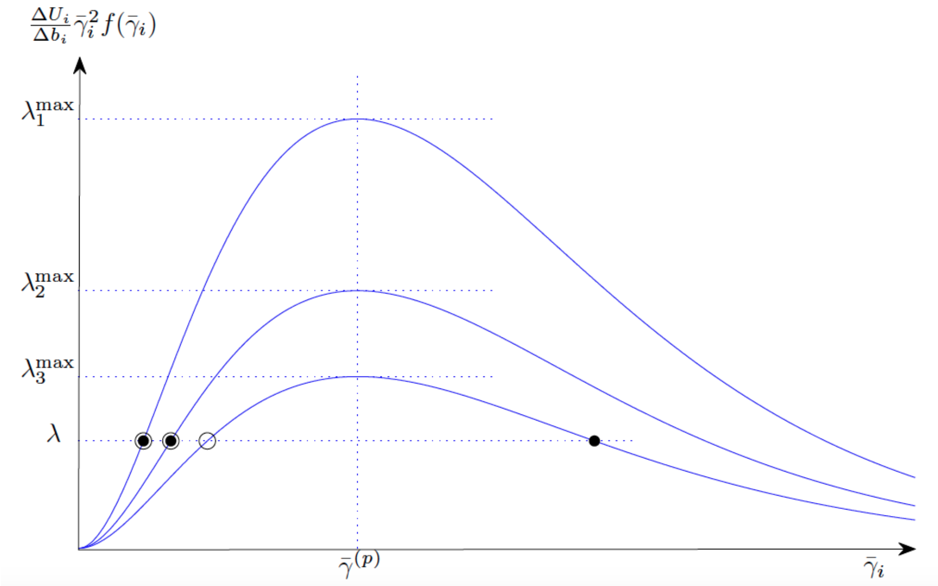


FIGURE 3.4: Sets of $\bar{\gamma}'s$ that minimize the Lagrangian for different $\lambda's$..

and the we can get the corresponding values of λ_i^{max} from the relation

$$\lambda_i^{max} = \frac{c_i}{\Delta b_i}(\bar{\gamma}^{(p)})^2 f_{\bar{\gamma}}(\bar{\gamma}^{(p)}) \quad \forall i = \{1, 2, \ldots, M\}, \tag{3.27}$$

Also, it can be noticed that for a certain value of λ, there is two corresponding values of $\bar{\gamma}_i$; one on the rising edge and the other on the falling edge. Therefore, if we denote the maximum possible number of layers for a given value of λ as $N^{max}(\lambda)$, then for a transmitted number of layers $N \leq N^{max}(\lambda)$, there are two possible values of $\bar{\gamma}_i' s$ satisfying (3.14b) and (3.24b); either we take all the values of $\bar{\gamma}_i' s$ on the rising edges, or we take $\bar{\gamma}_N$ on the falling edge and the rest of $\bar{\gamma}_i' s$ on the rising edges. Therefore, for a certain value of λ, we have possible $2N^{max}(\lambda)$ number of solutions for $\bar{\gamma}_i' s$, then we need an exhaustive search to find the solution minimizing the Lagrangian.

However, we can always find that the solution taking all γ thresholds on the rising edges results in lower value for the Lagrangian. Also, we can see in Appendix C that there is a unique value for γ threshold, denoted by $\bar{\gamma}^{(w)}$, to decide the number of layers N. Therefore, if we define $\lambda_i^{(w)}$ as the values of λ corresponding to the value of $\bar{\gamma}^{(w)}$

$$\lambda_i^{(w)} = \frac{c_i}{\Delta b_i}(\bar{\gamma}^{(w)})^2 f_{\bar{\gamma}}(\bar{\gamma}^{(w)}) \quad \forall i = \{1, 2, \ldots, M\}, \tag{3.28}$$

then if $\lambda_m^{(w)} \geq \lambda > \lambda_{m+1}^{(w)}$ it is optimal to send only m layers and to set $\bar{\gamma}_i = \infty, \forall i > m$. The value of the unique threshold $\bar{\gamma}^{(w)}$ can be obtained as the solution of the following relation in $(0, \bar{\gamma}^{(p)})$

$$1 - F_{\bar{\gamma}}(\bar{\gamma}^{(w)}) - \bar{\gamma}^{(w)} f_{\bar{\gamma}}(\bar{\gamma}^{(w)}) = 0. \tag{3.29}$$

Now we need do an exhaustive study for the perturbation curve to find the optimal solution for the primal problem (3.15). Although the perturbation function for a given number of layers is differentiable convex decreasing depicted in Figure 3.3, this might not be the case for the global perturbation function of the general problem which assumes that the optimal solution may involve any number of layers between 1 and M. In this case the perturbation function will be the minimum of all perturbation functions corresponding to different number of layers. This can be done by assuming first the case of sending $N \leq M$ layers. Considering the rising edges solution that minimizes the Lagrangian while changing λ from zero to λ_N^{max} will result in the solid part of the curve in Figure 3.5. Then if we change λ from λ_N^{max} to zero and consider the solution with $\bar{\gamma}_N$ is on the falling edge will result in the dashed part of the curve in Figure 3.5. Therefore, taking the minimum values of all the perturbation functions for $N = 1, 2, \ldots, M$ will result in the global perturbation function.

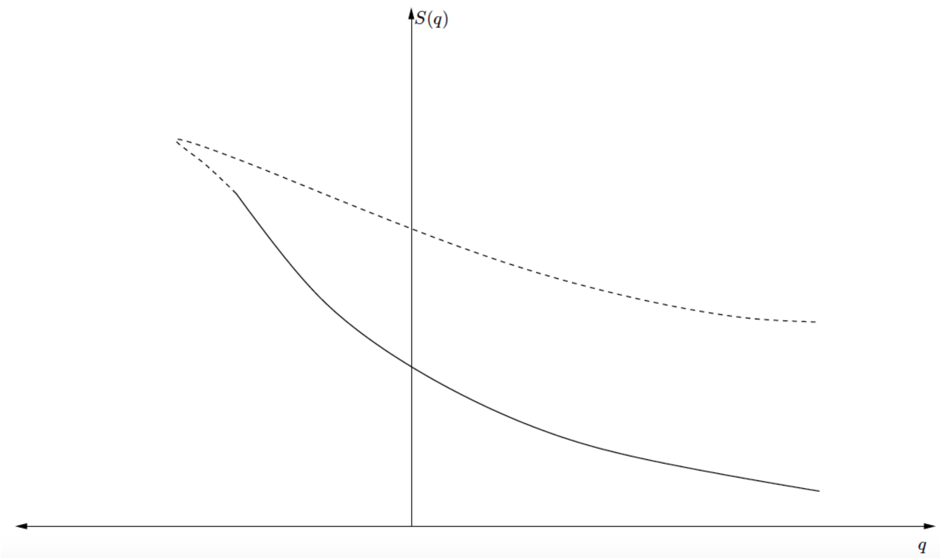
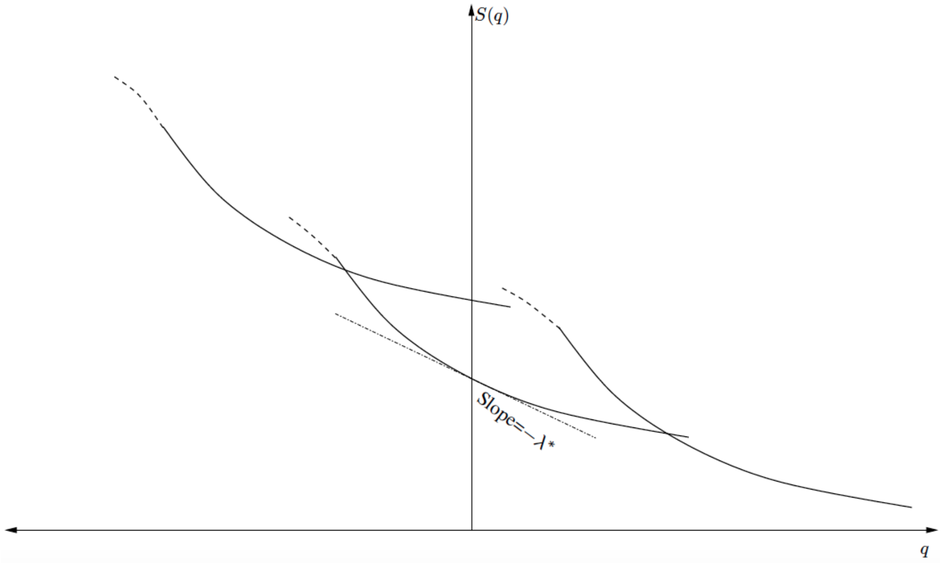
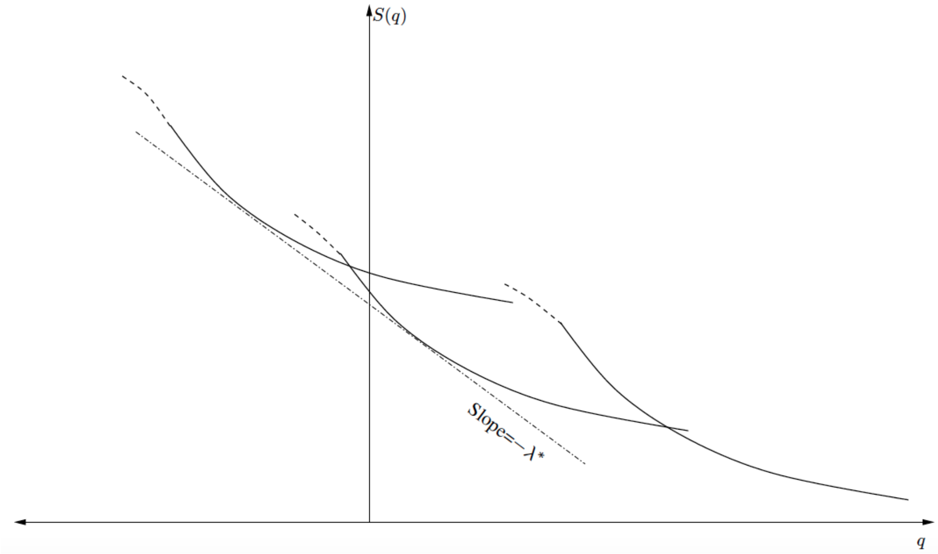


FIGURE 3.5: Changing λ to maximize $g(\lambda)$.

The global perturbation curve can take one of the two forms in Figures 3.6(a) and 3.6(b). The first form in Figure 3.6(a) corresponds the strong duality case where the solution for the primal problem in (3.15) is equal to the solution for the dual problem in (3.20). In this case, the optimal solution is the intersection of the global perturbation curve with the y-axis and can be found using bisection search over $\lambda \in [\lambda_{m+1}^{(w)}, \lambda_m^{(w)}]$ to find λ that satisfies the constraint (3.15b) with equality.



(a) Case: Strong Duality.



(b) Case: Weak Duality.

FIGURE 3.6: Global Perturbation Function for Strong and Weak Dualities

The second form in Figure 3.6(b) corresponds the weak duality case where the solution for the dual problem in (3.20) only gives a lower bound for the optimal solution of (3.15). In this case, the dual problem will have two solutions corresponding to transmit two successive

TABLE 3.2: Algorithm to Solve The Problem

---

1) INITIALIZATION:
   (a) Obtain $\mathbf{c}$ and $\mathbf{\Delta b}$ form $\mathbf{R}$ and $U(\bar{R})$.
   (b) Obtain $\bar{\gamma}^{(p)}$ by solving the relation in (3.26).
   (c) Obtain $\bar{\gamma}^{(w)}$ by solving the relation in (3.29) in $[0, \bar{\gamma}^{(p)}]$.
   (d) Declare the variable m as the index for the optimal number of transmitted layers and set m=M as an initial value.
2) CHECK DUALITY CASE AND FIND OPTIMAL NUMBER OF LAYERS:
   (a) If m=1, then $\bar{\gamma}_1 = \Delta b_1$ and $\bar{\gamma}_i = \infty, \forall i \neq 1$. GO TO step 5).
   (b) Use $\lambda = \lambda_m^{(w)}$ using (3.28), and get the corresponding $\bar{\gamma}_i' s$ by solving the relation (3.24b) in $[0, \bar{\gamma}^{(p)}]$.
   (c) Check if $\sum_{i=1}^{m} \frac{\Delta b_i}{\bar{\gamma}_i} < 1$, GO TO step 3). Otherwise continue.
   (d) Check if $\sum_{i=1}^{m-1} \frac{\Delta b_i}{\bar{\gamma}_i} < 1$, GO TO step 4). Otherwise, decrement m by 1, and repeat step 2).
3) STRONG DUALITY
   Using bisection search method, search for $\lambda \in [\lambda_{m+1}^{(w)}, \lambda_m^{(w)}]$ that satisfies $\sum_{i=1}^{m} \frac{\Delta b_i}{\bar{\gamma}_i}$ with equality. and we get the corresponding $\bar{\gamma}_i' s, \forall \lambda$ by solving the relation (3.24b) in $[0, \bar{\gamma}^{(p)}]$.
4) WEAK DUALITY
   (a) Obtain Solution A, which has $m-1$ transmitted layers, using bisection search for $\lambda_A \in [0, \lambda_m^{(w)}]$ that satisfies $\sum_{i=1}^{m-1} \frac{\Delta b_i}{\bar{\gamma}_i}$ with equality. and we get the corresponding $\bar{\gamma}_i' s$ by solving the relation (3.24b) in $[0, \bar{\gamma}^{(p)}]$.
   (b) Obtain Solution B (if exists), which has m transmitted layers, by searching for $\lambda_B \in [\lambda_m^{(w)}, \lambda_m^{(p)}]$ that satisfies $\sum_{i=1}^{m} \frac{\Delta b_i}{\bar{\gamma}_i}$ with equality.
      i. If $\lambda = \lambda_m^{(p)}$ will cause $\sum_{i=1}^{m} \frac{\Delta b_i}{\bar{\gamma}_i} < 1$, then use bisection search for λ over region λ_B that satisfies $\sum_{i=1}^{m} \frac{\Delta b_i}{\bar{\gamma}_i} = 1$, and we get the corresponding $\bar{\gamma}_i' s$ by solving the relation (3.24b) in $[0, \bar{\gamma}^{(p)}]$, then GO TO (c). Otherwise continue.
      ii. Search for λ over region λ_B that satisfies $\sum_{i=1}^{m} \frac{\Delta b_i}{\bar{\gamma}_i} = 1$, and we get the corresponding $\bar{\gamma}_m$ by solving the relation (3.24b) in $[\bar{\gamma}^{(p)}, \infty]$ and the rest of $\bar{\gamma}_i' s$ by solving the relation (3.24b) in $[0, \bar{\gamma}^{(p)}]$. In this case, the sum constraint will be monotonically increasing then at some point it will change to become monotonically decreasing. Therefore, either no solution at all exists, or there are two solutions in λ_B. In this case, choose the solution with the larger λ.
   (c) Compare solution A with solution B (if exists) and choose the minimum objective in (3.15a)
5) OPTIMAL POWER RATIOS
   Use (3.16) to obtain the values of the optimal power ratios for the layers.

---

layers m and $m-1$, with $\sum_{i=1}^{M} \frac{\Delta b_i}{\bar{\gamma}_i} > 1$ and $\sum_{i=1}^{M} \frac{\Delta b_i}{\bar{\gamma}_i} < 1$ respectively. Then both solutions will not satisfy the constraint (3.22b) with equality. Therefore, we need to search for the solution that satisfies the constraint (3.22b) with equality in the case of sending $m-1$ layers with rising edges using bisection over $\lambda \in [0, \lambda_m^{(w)}]$, and in the case of sending m layers using bisection over $\lambda \in [\lambda_m^{(w)}, \lambda_m^{(p)}]$, then take the solution with the minimum objective. It is important to notice that we may not find a solution that satisfies the constraint (3.22b) with equality in the case of sending m layers with rising edges. It happens if the value of $\lambda = \lambda_m^{(p)}$ will cause $\sum_{i=1}^{M} \frac{\Delta b_i}{\bar{\gamma}_i} > 1$. Then, we will need to find the solution that satisfies the constraint (3.22b) with equality in the case of sending m layers with $\bar{\gamma}_m$ is on the falling edge which also may not exist. If the minimum sum $\sum_{i=1}^{M} \frac{\Delta b_i}{\bar{\gamma}_i}$ of transmitting m layers is greater than one, then the optimal solution will be on the $(m-1)$-layers perturbation curve.

Now the previous algorithm steps to solve the primal optimization problem in (3.15) can be summarized in Table 3.2.

3.5 Numerical Results

In this section we present some numerical results for the case of Rayleigh fading channels described in (3.1). The proposed algorithm in Table 3.2 is applied for a scalable video source example consisting of four layers with a sum rate of 1 Mbps and transmitted over 1 MHz bandwidth. The rates of the source layers (obtained from Table IV in [30]) are 75.5, 80.4, 240 and 642 Kbps respectively. We consider two different utility functions; namely, $U(\bar{R}) = 1-2^{-2\bar{R}}$, which corresponds to minimizing the expected distortion of a Gaussian source and $U(\bar{R}) = \bar{R}$, which corresponds to maximizing the expected total rate at the destination.

Figures 3.7 and 3.8 show the optimal γ thresholds for the layers with the target of minimizing expected distortion and maximizing expected rate respectively, while Figures 3.9 and 3.10 show the optimal power ratios for these cases. The power ratios and γ thresholds are plotted against average SNR of the source-destination channel. The solid curves corresponds the case when the relay is used with $(m_1, m_2) = (16, 16)$ which is the best case for the relay position (i.e., relay in the mid point of the LOS between source and destination), with the assumption that the power of the signal $P \propto \frac{1}{d^4}$, where d is the distance. The dashed curves corresponds the case when no relay is used for comparison.

It can be seen that for some average SNR values it might be optimal to send only one layer, and as the average SNR value exceeds certain thresholds the number of layers increases. That is because as the average SNR increases, the channel condition becomes better, and the upper layers will be decoded reliably. It is obvious that the solid curves are shifted versions to the left for the dashed curves. Which means that it is optimal to send higher number of layers for lower values of SNR average when a relay is used even in the worst case. Therefore, the destination becomes more capable of decoding more layers refining the information even when its direct channel with the source has low SNR. However, we can notice that the γ threshold values for the layers when no relay is used are lower compared to the case for the relay-assisted. The reason is the multiplexing loss due to transmitting over two time slots. That becomes clear when comparing the expression for γ threshold in (3.10) with the case

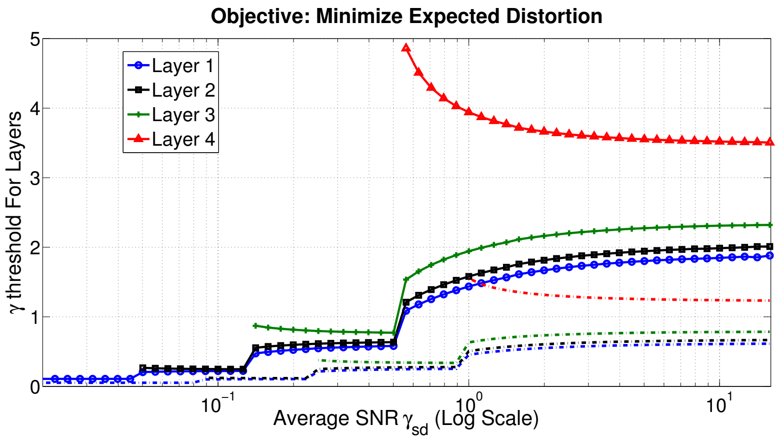


FIGURE 3.7: The threshold γ values of the layers versus the average SNR value of a Rayleigh fading channel with $U(\bar{R}) = 1 - 2^{-2\bar{R}}$.

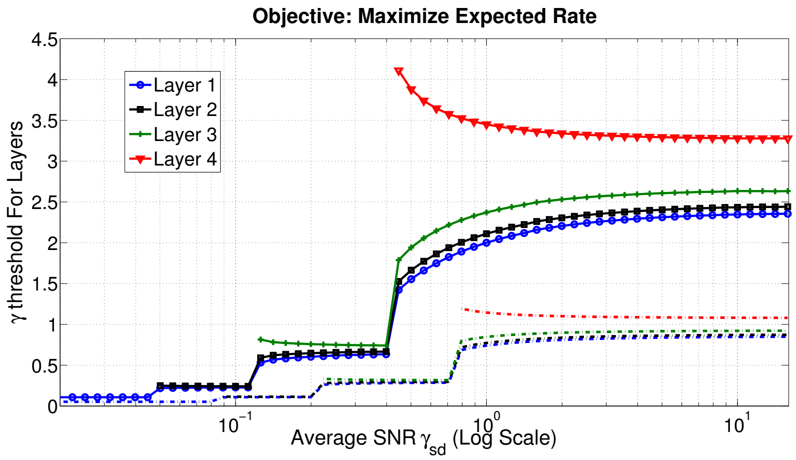


FIGURE 3.8: The threshold γ values of the layers versus the average SNR value of a Rayleigh fading channel with $U(\bar{R}) = \bar{R}$.

when no relay is used

$$\gamma_j = \frac{1}{\frac{\alpha_j}{2^{R_j}-1} - \sum_{m>j}^{M} \alpha_m}. \tag{3.30}$$

Since the utility function for maximizing the rate is linear, then the solution gives more importance for the higher layers as expected. This can be translated by allocating more power for the higher layers as seen in Figures 3.8 and 3.10. On the other hand, the solution for minimizing the average distortion gives more importance to the lower layers as in Figures 3.7 and 3.9, and hence the lower layers are allocated more power, and it becomes optimal to send

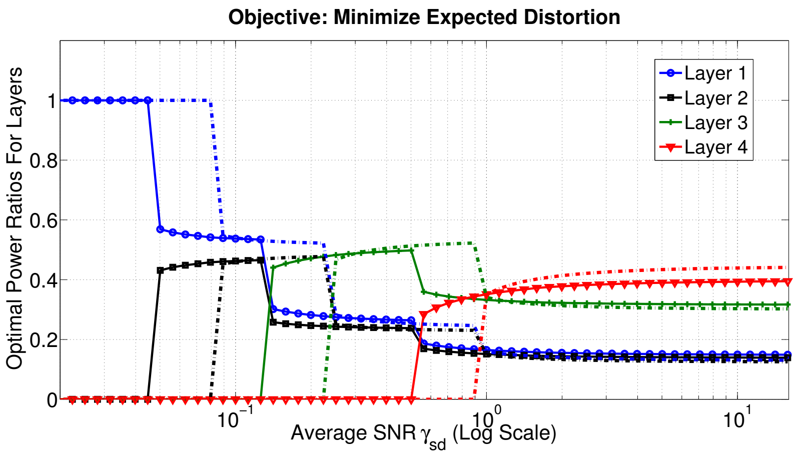


FIGURE 3.9: The relative power ratios of the layers versus the average SNR value of a Rayleigh fading channel with $U(\bar{R}) = 1 - 2^{-2\bar{R}}$.

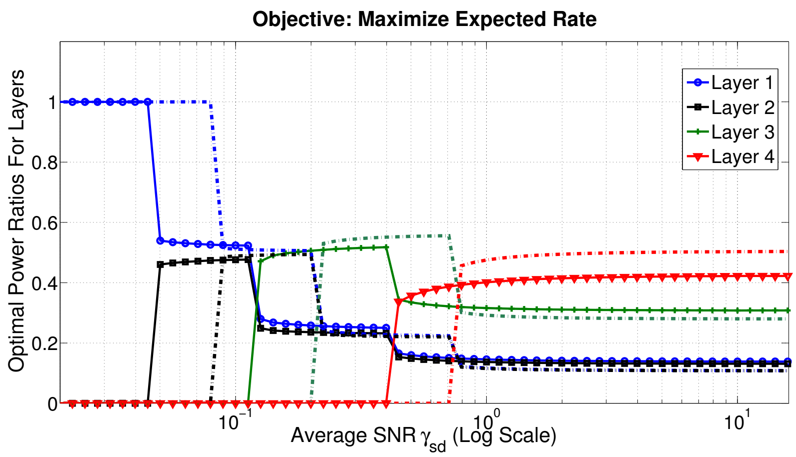


FIGURE 3.10: The relative power ratios of the layers versus the average SNR value of a Rayleigh fading channel with $U(\bar{R}) = \bar{R}$.

the higher layers for higher values of the average SNR compared to the case of maximizing the expected rate.

In Figures 3.11 and 3.12 we plot the maximized utility function with the target of minimizing expected distortion and maximizing expected rate respectively. It can be shown for the worst relay position case with $(m_1, m_2) = (100, 1)$ (i.e., relay near source or near destination), that the maximum expected utility is close(and maybe less than) the no-relay case. This is because the channel gains of the relay channel are not high in this case. Therefore, the

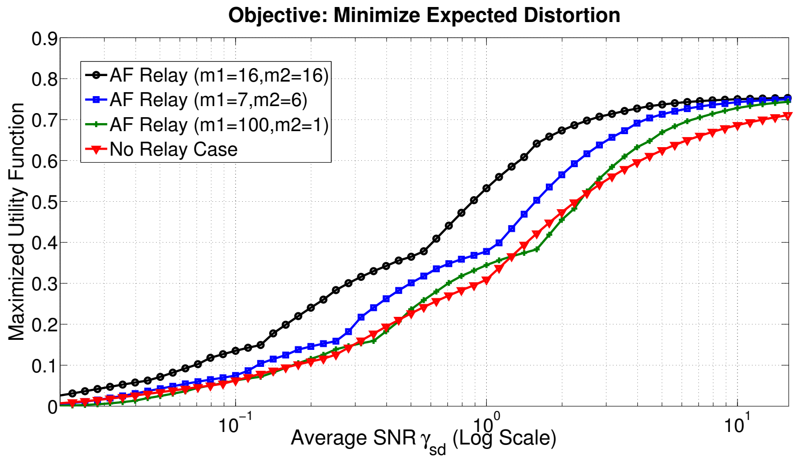


FIGURE 3.11: The maximized average utility function versus the average SNR value of a Rayleigh fading channel with $U(\bar{R}) = 1 - 2^{-2\bar{R}}$.

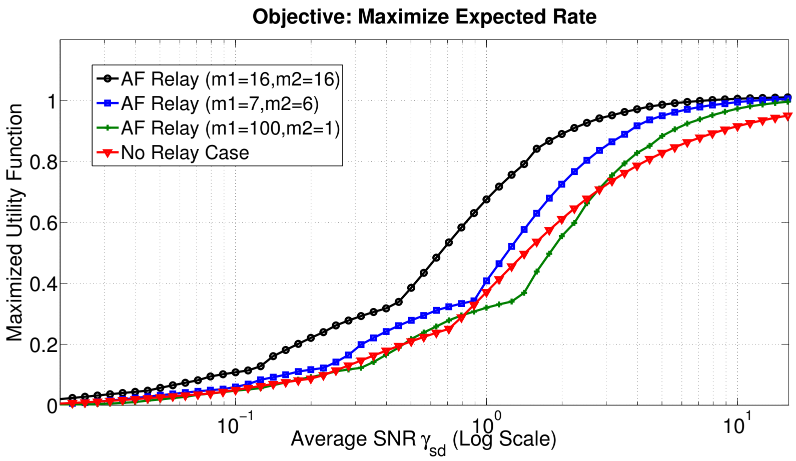


FIGURE 3.12: The maximized average utility function versus the average SNR value of a Rayleigh fading channel with $U(\bar{R}) = \bar{R}$.

prospected gain due to channel diversity of the relay channel will be opposed by the multiplexing loss due to the transmission over two time slots. Furthermore, the gain with respect to the no-relay case increases for the relay-assisted case with $(m_1, m_2) = (7, 6)$, and with $(m_1, m_2) = (16, 16)$ which is the best case (Relay in the mid point of the LOS between source and destination).

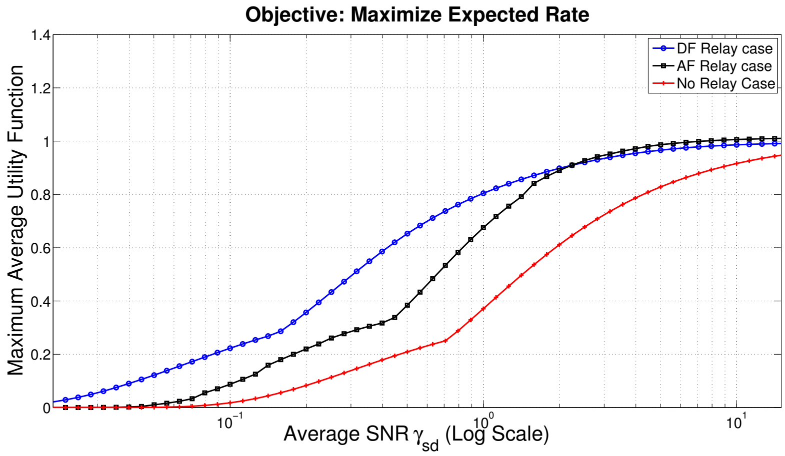


FIGURE 3.13: The maximized average utility function versus the average SNR value for different relaying strategies.

3.6 Comparison Between Different Relaying Strategies

Before we end this chapter we would like to study the effect of using different relaying strategies on the utility function. In Figure 3.13 we can see the average utility function is plotted versus the average source-destination channel quality for three different cases of relaying usage. We use the same example used in Chapter 2, where the source is transmitting using two layers only. The rates of the two source layers are respectively 1 and 2 bps/Hz. We consider $U(\bar{R}) = \bar{R}$, which corresponds to maximizing the expected total rate at the destination. Also, we use $m1 = m2 = 16$ which is the ratio between the average SNR of the source-relay and the relay-destination links to the source-destination link.

As discussed before, the DF relays is suffering from the error propagation because the relay node is decoding the layers before transmitting which is not the case for AF relays, because the AF relays do not perform any hard-decision operation on the received signal. However, in the AF protocol noise accumulates with the desired signal along the transmission path. That is why a better performance is obtained for DF relays. In both cases of AF and DF relays using no relays would be of lower values of expected utility function.

Another way to compare is by noting the small notches in the curves in Figure 3.13 , which corresponds the transition from sending one layer only to sending the two layers. For the DF relays, it can be noticed that it is optimal to send the enhancement layer for lower values of SNR compared with AF relays, which means better channel condition.

Chapter 4

Jointly Optimal Power and Rate Allocation for Layered Broadcast Over AF Relay Channels

4.1 Introduction

The joint power and rate allocation problem was considered in [31] and independently in [32], [33]. The authors solved the problem of minimizing the expected distortion of a Gaussian source with predetermined and fixed SNR thresholds of the source layers. The restriction of the SNR thresholds makes the source rates and power ratios dependent and hence finding one of them yields to the other. In [19] and originally in [28], [29], the problem was considered with fixed and predetermined rates for the source layers with no restrictions on the SNR thresholds. The solution algorithm proposed in [32], [33] has the worst case complexity $O(2^M)$, while the solution structure in [31], [19] has a linear complexity $O(M)$, where M is the number of layers. That is because the formulation in [31], [19] involves a change in optimization variables, where the power ratios of the layers were presented in terms of the rates and the SNR thresholds of the source layers. Other contributions include [34] in which the joint optimization problem was considered for the case of two layers with the objective of maximizing the expected rate. However, the global optimality is not guaranteed, and the solution proposed is not be extended for any number of layers greater than two. One possible solution for the optimal jointly power and rate allocation problem is by using the proposed solution algorithms considered in [31] and [19] iteratively; i.e., fix the rates and obtain the optimal SNR thresholds, then apply the obtained SNR thresholds from the previous step and solve for the rates. This can be repeated for many iterations until it converges to a solution. A similar idea was proposed in [35]. However, this algorithm is not optimal necessarily and can converge to a local maximum.

An efficient and generic algorithm for optimally joint rate and power allocation was considered in [20] for any finite number of layers with no restrictions on the source rates or SNR decoding thresholds, for any concave increasing utility function, and for any channel statistical model that fits some conditions. Moreover, the solution structure has a linear complexity with respect to the number of layers $O(M)$. The Lagrangian dual problem is applied where a two-dimensional bisection search is used to obtain the optimal solution satisfying the *Karush Kuhn Tucker KKT* conditions.[36–38]. The outer bisection search is done over the Lagrangian dual variable, while the inner bisection search is done over the SNR threshold of the last layer. Numerical results were provided for the objective of maximizing the long-term expected rate and minimizing the long-term average distortion. The case of infinite number of layers was considered in [20] providing an upper bound for the performance. However, it was shown that this upper bound can be approached for relatively small number of layers.

The analysis done in Chapter 3 is limited for pre-specified and un-controllable source layers. Although using fixed rates may be more practical specially when adapting the source coder is not feasible, it is a sub-optimal solution. In this chapter we extend the analysis done in Chapter 3 to the case of jointly optimal power and rate allocation among the layers in the context of AF relay networks. We use the same approximation done for the end-to-end channel statistics, so that we would be able to apply the algorithm presented in [20]. Also, the case of infinite number of layers is considered providing an upper bound for the performance.

4.2 System Model and Problem Formulation

We consider in this chapter the same system model as in Chapter 3. However, we assume the case of variable rates instead of fixed for the source layers. The relay is half-duplex and applies AF strategy, and the end-to-end channel quality is described using the same equations in (3.7). We need to find the jointly optimal power and rate allocation in order to maximize the expected user satisfaction. Therefore, the problem can be formulated similar to (3.13) as

$$\max_{\alpha,R,\bar{\gamma}} \sum_{i=1}^{M} U(\bar{R}_i)\ \left(F_\gamma(\bar{\gamma}_{i+1}(\alpha,R)) - F_\gamma(\bar{\gamma}_i(\alpha,R))\right) \tag{4.1a}$$

$$\text{subject to} \qquad \sum_{i=1}^{M} \alpha_i = 1, \qquad \alpha_i \geq 0 \qquad \forall i, \tag{4.1b}$$

$$\bar{\gamma}_i \geq \frac{1}{\frac{\alpha_i}{2^{2R_i}-1} - \sum_{m>i}^{M} \alpha_m} \qquad \forall i, \tag{4.1c}$$

$$\bar{\gamma}_M \geq \bar{\gamma}_{M-1} \geq \ldots \geq \bar{\gamma}_1 > 0, \tag{4.1d}$$

$$\bar{R}_M \geq \bar{R}_{M-1} \geq \ldots \geq \bar{R}_1 > 0, \tag{4.1e}$$

where $\bar{\gamma}_{M+1} = \infty$ for M number of layers, and $F_\gamma(\gamma)$ is the CDF of the end-to-end channel quality γ obtained in (3.19). The constraint (4.1e) was not considered before in Chapter 3, because the rates of the layers were predetermined.

Notice that the optimization problem in (3.13) has three variables α, R, and $\bar{\gamma}$. Since the the third variable is a function of the two other variables $\bar{\gamma}(\alpha, R)$, then for simplicity we can eliminate the variable $\bar{\gamma}$ from the optimization problem. Using the same procedures done in Chapter 3 to change the optimization variables, and using the same results obtained in Appendix A, we can reformulate the problem as follows

$$\max_{b,\bar{\gamma}} \quad \sum_{i=1}^{M} U(b_i)\ \left(F_\gamma(\bar{\gamma}_{i+1}) - F_\gamma(\bar{\gamma}_i)\right) \tag{4.2a}$$

$$\text{subject to} \quad \sum_{i=1}^{M} \frac{b_i - b_{i-1}}{\bar{\gamma}_i} = 1, \tag{4.2b}$$

$$\bar{\gamma}_M \geq \bar{\gamma}_{M-1} \geq \ldots \geq \bar{\gamma}_1 > 0, \tag{4.2c}$$

$$b_M \geq b_{M-1} \geq \ldots \geq b_1 > 0, \tag{4.2d}$$

where $b_i = 2^{2\bar{R}_i} - 1$, and $b_0 = 0$.

Now, the optimization problem is function only of the two variables b_i and γ_i, and after solving the problem (4.2), we can easily calculate the optimal rates for each layer using the relations $\bar{R}_i = \frac{1}{2}\log_2(1 + b_i)$, and $R_i = \bar{R}_i - \bar{R}_{i+1}$. Similarly, we can calculate the optimal power ratios for each layer recursively starting from α_1 using the following relation derived in [19],

$$\alpha_i = \frac{b_i - b_{i-1}}{1 + b_i}\left(1 - \sum_{j=1}^{i-1}\alpha_j + \frac{1}{\bar{\gamma}_i}\right) \tag{4.3}$$

Now, we can replace the optimization variables $\bar{\gamma}_i's$ with their reciprocal $\bar{n}_i's$ for simplicity. In this case, we have $F_\gamma(\bar{\gamma}_i) = 1 - F_n(\bar{n}_i)$, where F_n is the CDF of $n = \frac{1}{\gamma}$. Therefore, we can modify (4.2) to be

$$\max_{b,\bar{n}} \quad \sum_{i=1}^{M} U(b_i)\ \left(F_n(\bar{n}_i) - F_n(\bar{n}_{i+1})\right) \tag{4.4a}$$

$$\text{subject to} \quad \sum_{i=1}^{M} b_i(\bar{n}_i - \bar{n}_{i+1}) = 1, \tag{4.4b}$$

$$0 < \bar{n}_M \leq \bar{n}_{M-1} \leq \ldots \leq \bar{n}_1, \tag{4.4c}$$

$$b_M \geq b_{M-1} \geq \ldots \geq b_1 > 0, \tag{4.4d}$$

Where $n_{M+1} = 0$. The utility function $U(b_i)$ can have various definitions to employ many special cases for user satisfaction. For example, we use the definition $U(b_i) = \frac{1}{2}\log_2(1+b_i)$ in order to maximize the expected rate. Also, if the main target is to minimize the expected distortion of the Gaussian source, we the definition $U(b_i) = 1 - 2^{-2\left(\frac{1}{2}\log_2(1+b_i)\right)} = 1 + (1 + b_i)^{-1}$. The CDF (and equivalently the PDF) of n can be easily obtained using (3.19). as

$$F_n(n) = \frac{\beta_1}{\beta_1 - \beta'} e^{-\beta'/n} - \frac{\beta'}{\beta_1 - \beta'} e^{-\beta_1/n}, \tag{4.5a}$$

$$f_n(n) = \frac{1}{n^2}\frac{\beta_1\beta'}{\beta_1 - \beta'}\left(e^{-\beta'/n} - e^{-\beta_1/n}\right). \tag{4.5b}$$

where f_n is the PDF of n.

We can rearrange (4.4) by expanding the objective function and taking $F_n(\bar{n}_i)$ as a common factor. The problem can be written in this case as follows

$$\min_{b,\bar{n}} \quad \sum_{i=1}^{M} F_n(\bar{n}_i)\ \left(U(b_i) - U(b_{i-1})\right) \tag{4.6a}$$

$$\text{subject to} \quad \sum_{i=1}^{M} \bar{n}_i(b_i - b_{i-1}) = 1, \tag{4.6b}$$

$$0 < \bar{n}_M \leq \bar{n}_{M-1} \leq \ldots \leq \bar{n}_1, \tag{4.6c}$$

$$b_M \geq b_{M-1} \geq \ldots \geq b_1 > 0. \tag{4.6d}$$

By fixing the values of $\bar{\gamma}_i$, and hence $\bar{n}_i$ in (4.4), we get a similar problem formulation to [31], where a discrete channel model was considered. We can also get the same problem formulation as in Chapter 3, by fixing the rates for the layers in (4.6).

4.3 Problem Analysis

In this section we will apply the proposed solution algorithm in [20]. This solution is based on applying the KKT conditions to the dual problem. We start by removing the two constraints (4.4c) and (4.4d), and we can apply them implicitly when searching for the optimal values of $\bar{n}_i's$ and $b_i's$. Therefore, the optimization problem can be written as

$$\max_{b,\bar{n}} \quad \sum_{i=1}^{M} U(b_i) \ (F_n(\bar{n}_i) - F_n(\bar{n}_{i+1})) \tag{4.7a}$$

$$\text{subject to} \quad \sum_{i=1}^{M} b_i(\bar{n}_i - \bar{n}_{i+1}) = 1. \tag{4.7b}$$

Then, The Lagrangian dual problem of (4.7) can be written as

$$\min_{\lambda} \quad g(\lambda), \tag{4.8}$$

subject to $\lambda \geq 0$, where

$$g(\lambda) = \max_{\bar{n},b} \quad L(\bar{n}, b, \lambda), \tag{4.9}$$

where λ is the Lagrangian dual variable, and the Lagrangian $L(\bar{n}, b, \lambda)$ can be expressed in two equivalent forms according to the formulation of the optimization problem as

$$L(\bar{n}, b, \lambda) = \lambda + \sum_{i=1}^{M} \left(F_n(\bar{n}_i)\Delta U_i - \lambda \bar{n}_i \Delta b_i\right), \tag{4.10a}$$

$$L(\bar{n}, b, \lambda) = \lambda + \sum_{i=1}^{M} \left(U(b_i)\Delta F_i - \lambda b_i \Delta \bar{n}_i\right), \tag{4.10b}$$

where $\Delta U_i = U(b_i) - U(b_{i-1})$, $\Delta F_i = F_n(\bar{n}_i) - F_n(\bar{n}_{i+1})$, $\Delta b_i = b_i - b_{i-1}$, and $\Delta \bar{n}_i = \bar{n}_i - \bar{n}_{i+1}$, and where we apply the two constraints (4.4c) and (4.4d) when searching for

$b_i's$ and $\bar{n}_i's$ maximizing the Lagrangian. The formula in (4.10a) corresponds the problem formulation in (4.7), and (4.10b) corresponds the formulation in (4.6).

The solution of the Lagrangian problem in (4.8) must satisfy the KKT conditions, and the solution of the primal problem in (4.7) must satisfy the power constraint. Therefore we have the following $2M + 1$ equations

$$\frac{\Delta U_i}{\Delta b_i} f_n(\bar{n}_i) = \lambda \qquad \forall i, \tag{4.11a}$$

$$\frac{\Delta F_i}{\Delta \bar{n}_i} U'(b_i) = \lambda \qquad \forall i, \tag{4.11b}$$

$$\sum_{i=1}^{M} b_i(\bar{n}_i - \bar{n}_{i+1}) = 1, \tag{4.11c}$$

where $U'(b_i)$ is the first derivative of the utility function $U(b_i)$ with respect to b_i. We have in the dual problem (4.8) the Lagrangian variable λ in addition to the $2M$ variables $(b_i's, \bar{n}_i's)$ contained implicitly in $g(\lambda)$ as shown in (4.9). Therefore, we have $2M + 1$ variables must satisfy the $2M + 1$ equations in (4.11).

According to [20, Theorem 1] which states that:
Theorem (Existence and Uniqueness of the Solution). *For Rayleigh fading channels, and for any differentiable concave increasing utility function U(b), a strong duality exists between the primary and the dual problems.*

Therefore, we have only one solution for (4.11), and an efficient algorithm can be developed to find the solution of the joint power and rate allocation, as explained in Section 4.5.

4.4 Infinite Number Of Layers

In this section we will discuss the jointly optimization problem in the case of infinite number of layers, which guarantees achieving the upper bound for the utility function. Using infinite number of layers is not feasible, however, it gives us an indication for the maximum user satisfaction that can be achieved, and then we can use large number of layers sufficient to approach the upper bound as close as possible. In this section, we show that the framework developed in this thesis, for the case of finite number of layers, can be extended to the case of infinite number of layers with any differentiable concave increasing objective function $U(b)$.

In the case of infinite number of layers, we assume that each layer has infinitesimal small power ratio and rate. We will have a continuous range for the γ threshold $\bar{\gamma}$, and hence its reciprocal $\bar{n}$. Therefore, the decoding thresholds will be in the range $[\bar{n}_L, \bar{n}_F]$, where $\bar{n}_F$ is the maximum value of n sufficient to decode the first layer at least, and $\bar{n}_L$ is the maximum value of n in order to decode all the layers. The problem in this case turns to be finding the distribution of the sum rates $\bar{R}$ and its corresponding exponential $b = 2^{2\bar{R}} - 1$ as a function of the decoding thresholds; i.e., the solution will take the form $b(\bar{n})$. Therefore, the variable b will also have a continuous range $[b_F, b_L]$, where $b_F = 2^{2\bar{R}_1} - 1$ corresponding the case of decoding the first layer only, and $b_L = 2^{2\bar{R}_{sum}} - 1$ is the exponential in case of decoding all layers reliably.

Based on the continuous nature of the layers, we can replace $\Delta\bar{n}$, Δb, $\Delta F/\Delta\bar{n}$, and $\Delta U/\Delta b$ with $d\bar{n}$, db, $f_n(\bar{n})$, and $U'(b)$ respectively. Then the problem formulation in (4.4) can be rewritten for the case of continuum of layers as

$$\max_{\bar{n}_L, \bar{n}_F, b(\bar{n})} \quad U(b(\bar{n}_L))F_n(\bar{n}_L) + \int_{\bar{n}_L}^{\bar{n}_F} U(b(\bar{n}))F_n(\bar{n})d\bar{n} \tag{4.12a}$$

subject to

$$\bar{n}_L b(\bar{n}_L) + \int_{\bar{n}_L}^{\bar{n}_f} b(\bar{n})d\bar{n} = 1, \tag{4.12b}$$

$$\frac{d}{d\bar{n}}b(\bar{n}) \leq 0, \tag{4.12c}$$

where (4.12b) replaces the power constraint (4.4b), and (4.12c) replaces both the constraints (4.4c) and (4.4d). The constraint (4.12c) is satisfied when the objective utility function $U(b)$ is differentiable concave increasing, which is assumed in this work.

Similar to the case for finite number of layers, the KKT conditions must be satisfied in addition to the power constraint in (4.12b), and can be written as

$$U'(b(\bar{n}))f_n(\bar{n}) = \lambda, \quad \bar{n} \in (\bar{n}_L, \bar{n}_F) \tag{4.13a}$$

$$\frac{U(b(\bar{n}_F))}{b(\bar{n}_F)} f_n(\bar{n}_F) = \lambda, \tag{4.13b}$$

$$\frac{F_n(\bar{n}_L)}{\bar{n}_L} U'(b(\bar{n}_L)) = \lambda, \tag{4.13c}$$

$$\bar{n}_L b(\bar{n}_L) + \int_{\bar{n}_L}^{\bar{n}_F} b(\bar{n}) d\bar{n} = 1. \tag{4.13d}$$

From (4.13a) and (4.13c) with $\bar{n} = \bar{n}_L$, it can be shown that $F_n(\bar{n}_L) = \bar{n}_L f_n(\bar{n}_L)$, and using the proposed approximation formulas of the $F_n(n)$ and $f_n(n)$ found in (4.5) for Rayleigh fading channels, we can show that $\bar{n}_L$ is the solution for the relation

$$e^{-(\beta_1 - \beta')/\bar{n}_L} - \frac{\beta_1}{\beta'} \left(\frac{\bar{n}_L - \beta'}{\bar{n}_L - \beta_1} \right) = 0, \tag{4.14}$$

which can be solved numerically. Also, it can be shown from (4.13a) and (4.13b) with $\bar{n} = \bar{n}_F$ that $U(b(\bar{n}_F)) = b(\bar{n}_F) U'(b(\bar{n}_F))$, and since we assume for our work the utility function to be differential concave increasing function, then we must have

$$b(\bar{n}_F) = 0. \tag{4.15}$$

Therefore, substituting with $\bar{n} = \bar{n}_F$ in (4.13a), we have

$$U'(0) f_n(\bar{n}_F) = \lambda. \tag{4.16}$$

Finally, we can obtain the following relation

$$U(b(\bar{n})) = \frac{f_n(\bar{n}_F)}{f_n(\bar{n})} U'(0). \tag{4.17}$$

For a given definition of the utility function $U(b)$ and using the relation in (4.17), we can obtain an expression for $b(\bar{n})$. For example, If we consider the case of maximizing the expected rate, where $U(b) = \frac{1}{2} log_2(1 + b)$, then using the definition of $f_n(n)$ in (4.5b) we obtain

$$b(\bar{n}) = \frac{\bar{n}_F^2}{\bar{n}^2} \frac{e^{-\beta_3/\bar{n}} - e^{-\beta'/\bar{n}}}{e^{-\beta_3/\bar{n}_F} - e^{-\beta'/\bar{n}_F}} - 1, \quad n \in [\bar{n}_L, \bar{n}_F] \tag{4.18}$$

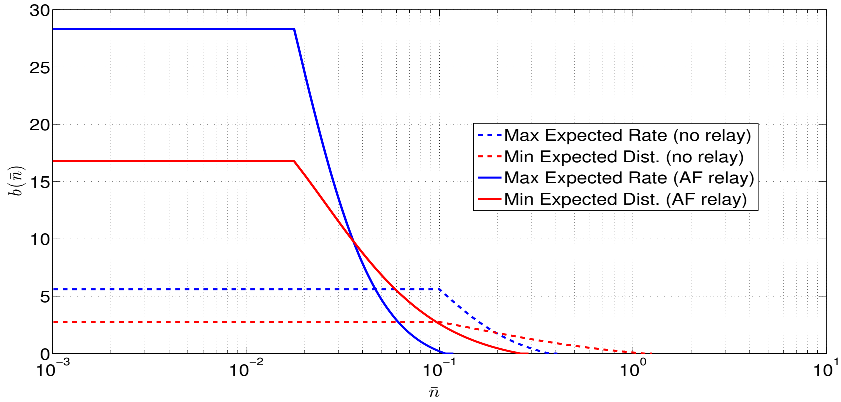


FIGURE 4.1: The optimal $b(\bar{n})$ versus the end-to-end quality for the case of infinite number of layers over a Rayleigh fading channel with $(m_1, m_2) = (16, 16)$.

Also, if we consider the case of minimizing the expected distortion of a Gaussian source, where, $U(b) = 1 + (1 + b_i)^{-1}$, then we can obtain $b(\bar{n})$ as follows

$$b(\bar{n}) = \frac{\bar{n}_F}{\bar{n}} \left(\frac{e^{-\beta_3/\bar{n}} - e^{-\beta'/\bar{n}}}{e^{-\beta_3/\bar{n}_F} - e^{-\beta'/\bar{n}_F}} \right)^{1/2} - 1, \quad n \in [\bar{n}_L, \bar{n}_F], \tag{4.19}$$

where $\bar{n}_L$ is obtained by solving the relation in (4.14), and $\bar{n}_F$ is obtained so that the power constraint in (4.13d) is achieved. This can be solved numerically using bisection search to obtain $\bar{n}_F$. Notice that $b(\bar{n}) = b(\bar{n}_L)$ for $n = [0, \bar{n}_L]$, and $b(\bar{n}) = 0$ for $n = [\bar{n}_F, \infty]$.

Figure 4.1 shows a numerical example for the optimal $b(\bar{n})$ with the objective of maximizing the expected rate and minimizing the expected distortion. The dotted curves represent the no-relay case, while the solid curves represent the AF relay case with $m_1 = m_2 = 16$, which is the best case for the relay position (i.e., relay in the mid point of the LOS between source and destination), with the assumption that the power of the signal $P \propto \frac{1}{d^4}$, where d is the distance. It can be observed that the curves with the objective of maximizing the expected rate have a narrower range of SNR, which means lower values for $\bar{n}_F$ (higher values for γ to decode the first layer), however, higher total rates are distributed among the layers. This is due to the nature of the utility function in that case which gives more importance for the higher layers.

Also, if we compare the solid curves with the dotted ones, we can notice lower values of $\bar{n}$ (higher values of γ thresholds) and higher total rates distributed among the layers for the case of the solid curves representing the AF relay case. The higher values for γ thresholds is due to the multiplexing loss due to transmitting over two time slots. We can easily prove that when comparing the expression for γ threshold for the case of AF relay in (3.10) with the case when no relay is used

$$\gamma_j = \frac{1}{\frac{\alpha_j}{2^{R_j}-1} - \sum_{m>j}^{M} \alpha_m}, \tag{4.20}$$

while the higher total rates for the AF relay case are due to the enhancement in the end-to-end performance.

4.5 Search Algorithm

In this section, we will apply a two-dimensional bisection search to find the solution of the equations in (4.11). We use an outer bisection search over λ to find the value satisfying the power constraint (4.11c) with equality. For each value of λ, we need to find the values of $b_i's$ and $\bar{n}_i's$ that satisfy the KKT conditions. Therefore, an inner bisection search is done to find the values satisfying the $2M$ equations in (4.11a) and (4.11b). First, we set an arbitrary value for $\bar{n}_M$, then we can get b_M from (4.11b) with $i = M$. The next step is to get b_{M-1} from (4.11a) by applying the values of $\bar{n}_M$ and b_M with $i = M$. Then, using $\bar{n}_M$ and b_{M-1} in (4.11b) with $i = M-1$ we can obtain $\bar{n}_{M-1}$. We can continue this procedure in the same manner until we get b_1 and $\bar{n}_1$ after solving $2M-1$ equations. There will be one remaining equation which is (4.11a) with $i = 1$. Therefore, the inner bisection search is done over $\bar{n}_M$ until we reach the values of b_1 and $\bar{n}_1$ satisfying this remaining equation. The algorithm for the inner and the outer bisection search is shown in Tables 4.1 and 4.2.

TABLE 4.1: Outer Bisection Search over λ

Initialize λ_L and λ_U (lower and upper bounds for bisection search).
Given a tolerance ϵ,

1. Set $\lambda_0 = (\lambda_L + \lambda_U)/2$.
2. Given λ_0, apply inner bisection in Table 4.2 to find $b_i, \bar{n}_i,\ i = 1, \ldots, M$.
3. Calculate $q = \sum_{i=1}^{M} b_i(\bar{n}_i - \bar{n}_{i+1})$.
4. If $q > 1 + \epsilon$, set $\lambda_L = \lambda_0$, return to Step 1.
5. If $q < 1 - \epsilon$, set $\lambda_U = \lambda_0$, return to Step 1.

Return $\lambda = \lambda_0$, and $b_i, \bar{n}_i,\ i = 1, \ldots, M$.

TABLE 4.2: Inner Bisection Search over $\bar{n}_M$

Initialize $\bar{n}_{M\_L}$ and $\bar{n}_{M\_U}$ (lower and upper bounds for bisection search).
Given a tolerance ϵ and λ_0,

1. Initialize i = M.
2. Set $\bar{n}_{M\_0} = (\bar{n}_{M\_L} + \bar{n}_{M\_U})/2$.
3. Calculate b_M using (4.11b) given $\bar{n}_{M\_0}$ and knowing that $\bar{n}_{M+1} = 0$.
4. Set $i = i - 1$.
5. Calculate b_i using (4.11a) given b_{i+1} and $\bar{n}_{i+1}$ obtained from previous steps.
6. Calculate $\bar{n}_i$ using (4.11b) given b_i and $\bar{n}_{i+1}$ obtained from previous steps.
7. If $i > 1$, return to Step 4.
8. Calculate $\omega = \frac{U(b_1)}{b_1} f_n(\bar{n}_1) - \lambda_0$.
9. If $\omega > \epsilon$, set $\bar{n}_{M\_L} = \bar{n}_{M\_0}$, return to Step 1.
10. If $\omega < -\epsilon$, set $\bar{n}_{M\_U} = \bar{n}_{M\_0}$, return to Step 1.

Return $b_i, \bar{n}_i,\ i = 1, \ldots, M$.

The choice of the upper and lower bounds for both inner and outer bisection search is completely arbitrary. However, it was shown by experiment that as the number of layers M increases, the optimal value of λ increases. Therefore, a suitable choice for the upper bound λ_U is the optimal λ for the case of infinite number of layers given by (4.16), and a suitable choice of the lower bound λ_L is the optimal λ for the one-layer case. Otherwise, we can choose $\lambda_L = 0$ and λ_U large enough such that the solution at λ_U results in a power $q < 0$. Also, a suitable choice for the lower bound $\bar{n}_{M\_L}$ for a Rayleigh channels can be given by (4.14), which is the minimum non-zero value of $\bar{n}_i$ independent from the number of layers M. While, the upper bound $\bar{n}_{M\_U}$ can be chosen arbitrarily high enough so that $\omega < 0$.

4.6 Numerical Results

In this section we present some numerical results for the case of Rayleigh fading channels described in (3.1). The search algorithm presented in Section 4.5 is applied for a source example consisting of three layers to find the optimal joint power and rate allocation. We use the proposed end-to-end channel approximation presented in Section 3.3. We also consider two different utility functions; namely, minimizing the expected distortion of a Gaussian source and maximizing the expected sum rate decoded at the destination.

Figures 4.2, 4.3, 4.4, and 4.5 show the optimal power ratios, rates, rate ratios, and $\bar{\gamma}$ decoding thresholds for the three layers respectively with $(m_1, m_2) = (16, 16)$, which is the best case for the relay position (Relay in the mid point of the LOS between source and destination), where the solid curves represent the case with the objective of maximizing the expected sum rate, and the dotted curves represent the case with the objective of minimizing the average distortion. Figure 4.2 shows the optimal power ratios for the three layers. It can be shown that the first layer for both cases of utility functions is given higher power allocation than the upper layers. This is clear since the multilayer systems give more protection for the first layer which is the base layer. Also, we can see that as $\bar{\gamma}$ increases the power ratio for the first layer increases.

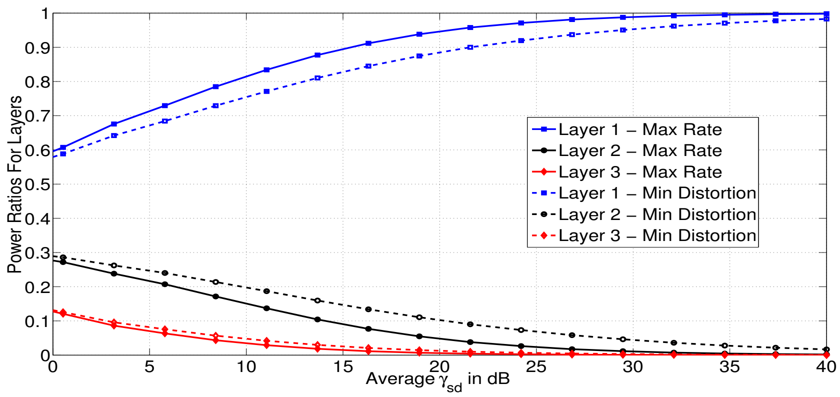


FIGURE 4.2: The optimal power ratios of the layers versus $\bar{\gamma}$ for three layers transmitted over a Rayleigh fading AF relay-assisted channel with $(m_1, m_2) = (16, 16)$.

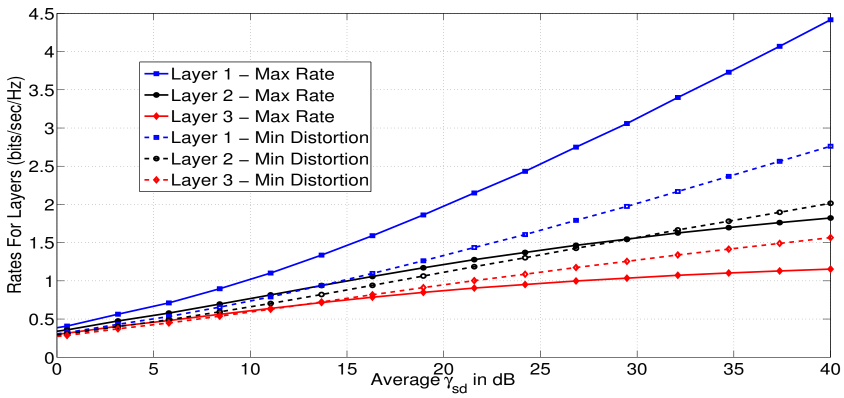


FIGURE 4.3: The optimal rates of the layers versus $\bar{\gamma}$ for three layers transmitted over a Rayleigh fading AF relay-assisted channel with $(m_1, m_2) = (16, 16)$.

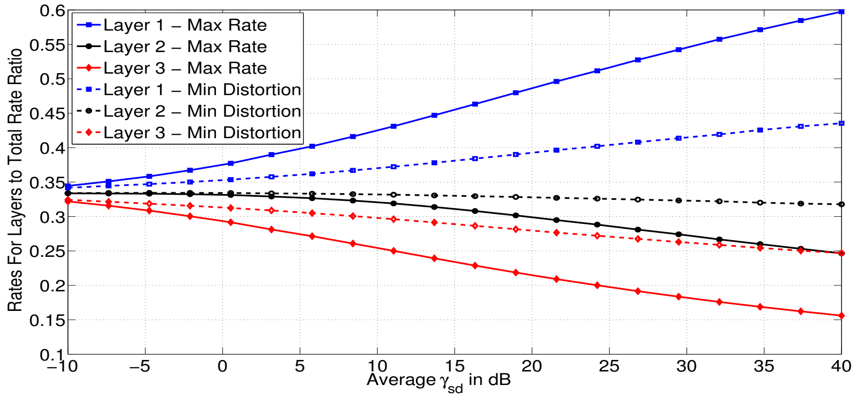


FIGURE 4.4: The optimal rate ratios of the layers versus $\bar{\gamma}$ for three layers transmitted over a Rayleigh fading AF relay-assisted channel with $(m_1, m_2) = (16, 16)$.

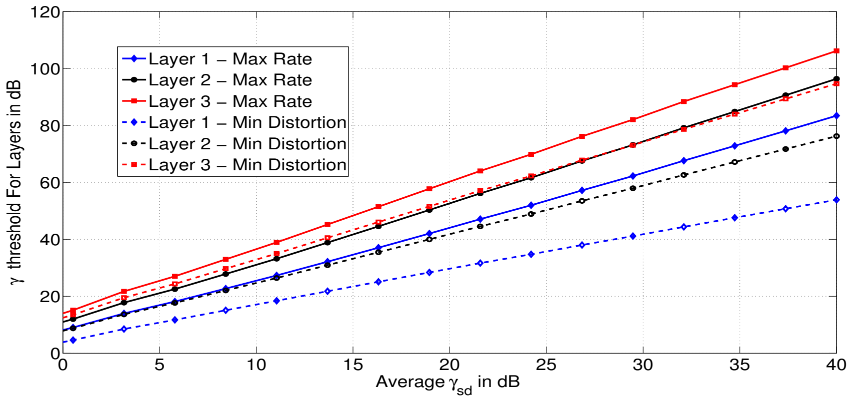


FIGURE 4.5: The optimal γ thresholds of the layers versus $\bar{\gamma}$ for three layers transmitted over a Rayleigh fading AF relay-assisted channel with $(m_1, m_2) = (16, 16)$.

We can notice from Figures 4.3 and 4.4 that Layer 1 is given lower rates, and Layer 3 is given higher rates for the case of minimizing the average distortion compared with the case of maximizing the expected sum rate. This is due to the nature of the utility function for both cases. The utility function for the former case is more convex in $\bar{\gamma}$ than the latter case. This gives more importance for the lower layers for the former case, which corresponds to lower rates for Layer 1 to increase the protection. On the other hand, the utility function for maximizing the rate is linear, then the solution gives more importance for the higher layers as expected. This can be represented by allocating lower rates for the higher layers.

Figure 4.5 shows the decoding thresholds $\bar{\gamma}$ for the three layers. It can be seen that the lower layers in the case of minimum distortion have lower SNR decoding thresholds than

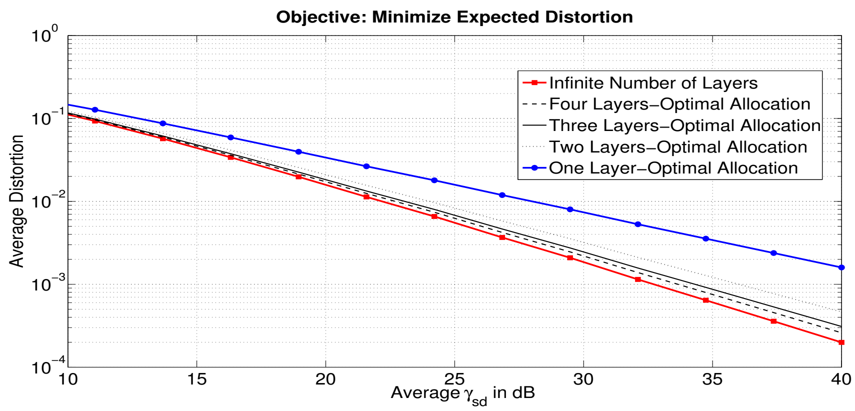


FIGURE 4.6: The average distortion versus $\bar{\gamma}$ for various number of layers transmitted over a Rayleigh fading AF relay-assisted channel with $(m_1, m_2) = (16, 16)$ with the objective of minimizing the average distortion.

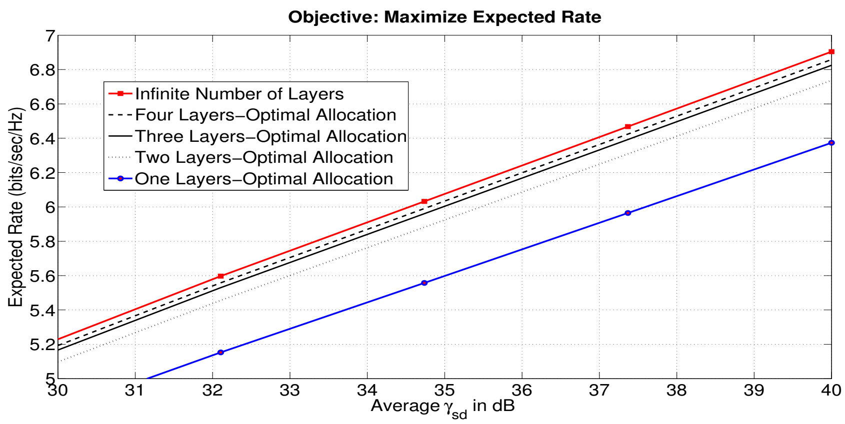


FIGURE 4.7: The expected sum rate versus $\bar{\gamma}$ for various number of layers transmitted over a Rayleigh fading AF relay-assisted channel with $(m_1, m_2) = (16, 16)$ with the objective of maximizing the expected sum rate.

the corresponding layers in the case of maximizing the expected rates. This is consistent with the previous results where the case of minimum distortion gives more importance for the lower layers. Therefore, the power ratios for Layer 1 in this case is lower compared with the case of maximum sum rate as shown in Figure 4.2 because it has lower rate. Moreover, It can be seen that at optimality the γ threshold value of Layer 1 is lower than those of the upper layers. This is due to the fact that the base layer must be given the highest protection.

In Figures 4.6 and 4.7, It can be shown the effect of increasing the number of layers for the case of minimum average distortion and maximum expected sum rate respectively. For the

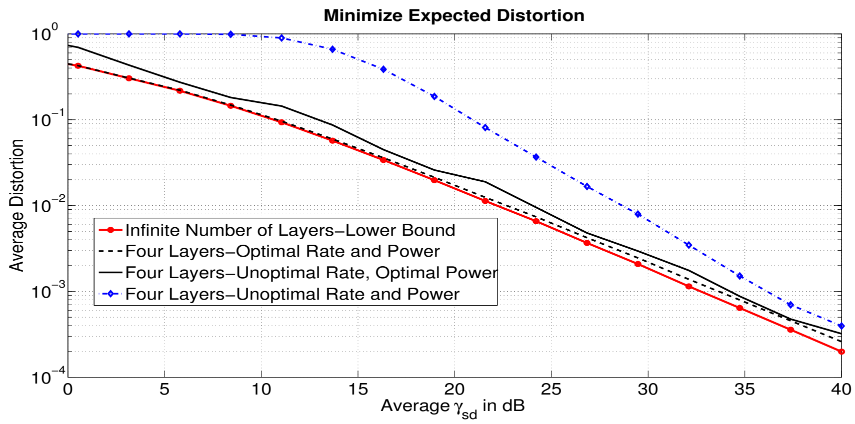


FIGURE 4.8: The average distortion versus $\bar{\gamma}$ over a Rayleigh fading AF relay-assisted channel with $(m_1, m_2) = (16, 16)$ with the objective of minimizing the average distortion.

former case, by increasing the number of layers, the average distortion decreases until we reach the lower bound achieved theoretically by transmitting infinite number of layers, while for the latter case, by increasing the number of layers, the expected sum rate increases until we reach the upper bound achieved theoretically by transmitting infinite number of layers. However, we can notice that we can achieve a close values to the lower and upper bounds by transmitting a small number of layers.

Figure 4.8 shows a comparison for the minimum average distortion in a four layer system between the optimal power and rate allocation, fixed power and rate allocation, and optimal power allocation with fixed sub-optimal equal rates. We can see that the jointly optimal power and rate allocation decreases the average distortion compared with the other sub-optimal allocations. We can also notice that the infinite number of layers provide a theoretical lower bound for the average distortion which was considered in [31].

Figure 4.9 shows a comparison for the maximum expected sum rate in a four layer system between the optimal power and rate allocation, fixed power and rate allocation, and optimal power allocation with fixed sub-optimal equal rates. We can see that the jointly optimal power and rate allocation increases the expected sum rate compared with the other sub-optimal allocations. We can also notice that the infinite number of layers provide a theoretical upper bound for the expected sum rate.

In Figures 4.8 and 4.9, we can notice that for low values of $\bar{\gamma}$ that the equal allocation of power and/or rate gives a close utility function to the optimal power and rate allocation, and as $\bar{\gamma}$ increases, the optimal solution moves away from the equal allocation. This is because

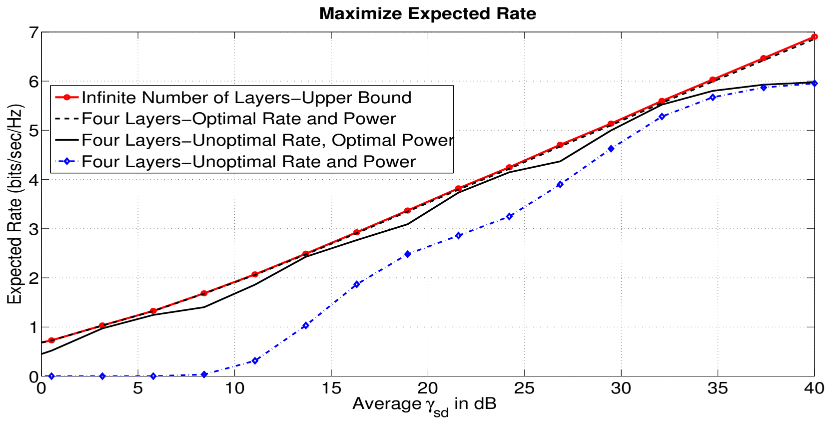


FIGURE 4.9: The expected sum rate versus $\bar{\gamma}$ over a Rayleigh fading AF relay-assisted channel with $(m_1, m_2) = (16, 16)$ with the objective of maximizing the expected sum rate.

for low values of $\bar{\gamma}$ the optimal solution for rates and power ratios are close for all the layers and nearly equal; Therefore, the optimal solution will be close to the equal allocation, and as $\bar{\gamma}$ increases, the gap between the optimal rates and power ratios for the layers increases, and hence the performance of the equal allocation becomes gradually worse than the optimal allocation at high SNR.

In Figure 4.10, we plot the minimum average distortion for different cases of relay positions. It can be shown that for the worst relay position case with $(m_1, m_2) = (100, 1)$ (i.e., relay near source or near destination), that the minimum average distortion is close the no-relay case. This is because the channel gains of the relay channel are not high in this case. Therefore, the prospected gain due to channel diversity of the relay channel will be opposed by the multiplexing loss due to the transmission over two time slots. Furthermore, the gain with respect to the no-relay case increases and the average distortion decreases for the relay-assisted case with $(m_1, m_2) = (8, 4)$, and with $(m_1, m_2) = (16, 16)$ which is the best case (Relay in the mid point of the LOS between source and destination).

In Figure 4.11, we plot the maximum expected sum rate for different cases of relay positions. It can be shown that even for the best relay position case with $(m_1, m_2) = (16, 16)$ that the maximum expected sum rate is greater than the no-relay case only for low values of $\bar{\gamma}$, and as $\bar{\gamma}$ increases above a certain level, the expected sum rate for the no-relay case will be greater than the relay-assisted channel case. This can be explained by comparing the relation in (3.8) which gives the layer rates for the case of the relay-assisted channel with the relation (4.21)

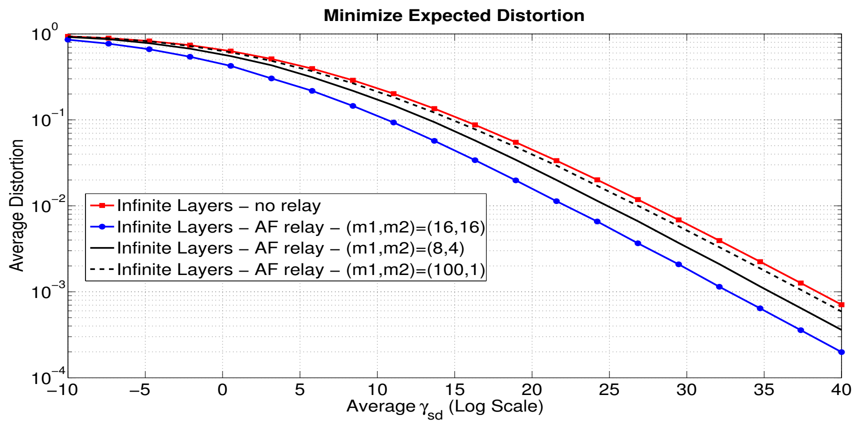


FIGURE 4.10: The average distortion versus $\bar{\gamma}$ with infinite number of layers transmitted over a Rayleigh fading channel with and without using a relay with the objective of minimizing the average distortion.

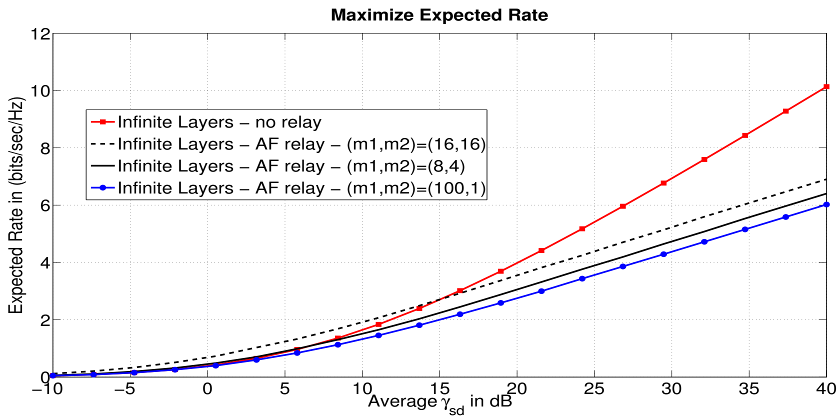


FIGURE 4.11: The expected sum rate versus $\bar{\gamma}$ with infinite number of layers transmitted over a Rayleigh fading channel with and without using a relay with the objective of maximizing the expected sum rate.

for the no-relay case.

$$R_j \leq \log\left(1 + \frac{\alpha_j}{\frac{1}{\gamma} + \sum_{m>j}^{M} \alpha_m}\right) \quad \forall j \leq i. \tag{4.21}$$

Due to the concave nature of the log function, and for high values of $\bar{\gamma}$, the gain due to the enhancement in the end-to-end channel quality caused by using the relay will be less than the multiplexing loss presented by the multiplication by $\frac{1}{2}$ in (3.8). Therefore, the expected sum rate decreases for high $\bar{\gamma}$ values.

Chapter 5

Conclusion and Future Work

In this work, we considered a fading relay channel where the source uses layered source coding with successive refinement. The layers are transmitted using superposition coding at the transmitter with successive interference cancellation at the receiver. We considered three cases of optimal resource allocation for different relaying strategies. We have investigated the Amplify-and-Forward, and Decode-and-Forward strategies in particular.

The first case, a relay has been considered to assist the channel using selection relaying DF strategy. The random search method has been applied to find the optimal power allocation at the source and at the relay in order to maximize the expected user satisfaction that is defined by a utility function of the total decoded rate at the destination. Several numerical examples were obtained for two different utility functions, which are maximizing the expected rate and minimizing the expected distortion of a Gaussian source. It has been shown that it may be optimal not to transmit both layers for low average SNR values of the channels. In this case, all the power is allocated to the base layer and the enhancement layer is discarded. An obvious gain was observed for the relay channel in comparison with the direct transmission case with no relay assistance. These gain increases as the ratio between the average SNR of the source-relay and relay-destination links to the average SNR of the source-destination link increases.

For the second case, the transmission is relay-aided, and the relay applies AF strategy. The optimal power allocation over the source layers can be obtained for any number of layers using the algorithm that is known in the literature for this class of problems. The objective is to maximize the expected user satisfaction that is defined by a utility function of the total decoded rate at the destination. However, we needed to obtain the end-to-end channel statistics analytically. So, we have proposed a simple and appropriate approximation for the AF relay scenario. We also obtained several numerical examples for two different utility functions, which are maximizing the expected rate and minimizing the expected distortion of a Gaussian source. Furthermore, we obtained that for some cases, it is optimal not to send all the layers depending on the channel condition. The numerical results demonstrate the gains of relaying for different relay positions.

In the final case, the relay applies AF strategy, and we solve the optimal rate and power allocation problem over the source layers using an algorithm that is known in the literature. We used the same approximation done for the end-to-end channel statistics from the previous case. The objective was to maximize a utility function that represents the receiver satisfaction. Moreover, we have shown that with a relatively small number of layers, we can approach the upper bound corresponding the infinite number of layers case. Several numerical examples were obtained for two different utility functions, which are maximizing the expected rate and minimizing the expected distortion of a Gaussian source. The numerical results demonstrate that relaying causes gain for the case of minimizing the expected distortion. However, it was shown that for high values of SNR and for the objective of maximizing the expected rate, the no-relay case shows better performance.

Future work for the problem considered in this thesis may include finding an approximation and a closed form for the end-to-end channel quality for the DF relay case, so that the optimal power allocation problem can be solved for any number of layers and with linear complexity. Some other directions for future research may be using different relaying strategies such as the dynamic DF strategy, and trying to solve the optimal resource allocation for these cases.

Appendix A

Change of Optimization Variables

We can reformulate the optimization problem in (3.13) such that the γ thresholds are the only optimization parameter, which would renders the problem easier to solve. Therefore, we reformulate the constraints (3.14a) and (3.13b) by assuming that all layers above layer i are allocated zero power, which means that layer i is decoded with no interference, as follows:

$$\begin{aligned} \bar{\gamma}_i &\geq \frac{2^{2R_i}-1}{\alpha_i} && \text{(A.1a)} \\ &= \frac{2^{2R_i}-1}{1-\sum_{m<i}\alpha_m} && \text{(A.1b)} \\ &= \frac{b_i - b_{i-1}}{1-\sum_{m<i}\frac{b_m-b_{m-1}}{\bar{\gamma}_m}} \quad \forall i, && \text{(A.1c)} \end{aligned}$$

where $b_i = 2^{\sum_{m<i} 2R_m} - 1$ and $b_0 = 0$. Then we can easily rewrite this constraint as:

$$\sum_{m<i} \frac{b_m - b_{m-1}}{\bar{\gamma}_m} \leq 1 \quad \forall i \leq M. \tag{A.2}$$

It can be shown that if the constraint in (A.2) is satisfied for $i = M$, then it will be satisfied for all $i < M$. Therefore, the constraint can now be written as

$$\sum_{i=1}^{M} \frac{\Delta b_i}{\bar{\gamma}_i} - 1 \leq 0, \tag{A.3}$$

where $\Delta b_i = b_i - b_{i-1}$.

Appendix B

No Solution On The Boundary

At any point on the boundary $\bar{\gamma}_i = \bar{\gamma}_{i+1}$, we move towards the infeasible region if either $\bar{\gamma}_i$ is increasing or $\bar{\gamma}_{i+1}$ is decreasing. Thus, in order for the Lagrangian $L(\bar{\gamma}, \lambda)$ to be increasing towards the boundaries (i.e., no solution on the boundaries), it must be increasing either as $\bar{\gamma}_i$ increases or as $\bar{\gamma}_{i+1}$ decreases $\forall i$. This can be expressed as

$$\frac{\partial L}{\partial \bar{\gamma}_i}|_{\bar{\gamma}_i = \bar{\gamma}_{i+1}} > 0 \quad \textbf{OR} \quad \frac{\partial L}{\partial \bar{\gamma}_{i+1}}|_{\bar{\gamma}_i = \bar{\gamma}_{i+1}} < 0. \tag{B.1}$$

Substituting with the Lagrangian in (B.1), we get the following condition

$$\frac{c_i}{\Delta b_i}\bar{\gamma}_i{}^2 f_\gamma(\bar{\gamma}_i) - \lambda > 0 \quad \textbf{OR} \quad \frac{c_{i+1}}{\Delta b_{i+1}}\bar{\gamma}_{i+1}^2 f_\gamma(\bar{\gamma}_{i+1}) - \lambda < 0. \tag{B.2}$$

We can see if the condition (3.23) is satisfied, then if one of the conditions in (B.2) is satisfied, then the other one must be not satisfied. This can be shown by considering the first condition in (B.2) is not satisfied as follows

$$\frac{c_i}{\Delta b_i}\bar{\gamma}_i{}^2 f_\gamma(\bar{\gamma}_i) - \lambda < 0, \tag{B.3a}$$

$$\frac{c_i}{\Delta b_i} < \frac{\lambda}{\bar{\gamma}_i{}^2 f_\gamma(\bar{\gamma}_i)}, \tag{B.3b}$$

then for $\frac{c_i}{\Delta b_i} > \frac{c_{i+1}}{\Delta b_{i+1}}$,

$$\frac{c_{i+1}}{\Delta b_{i+1}} < \frac{c_i}{\Delta b_i} < \frac{\lambda}{\bar{\gamma}_i{}^2 f_\gamma(\bar{\gamma}_i)}, \tag{B.4a}$$

$$\frac{c_{i+1}}{\Delta b_{i+1}}\bar{\gamma}_{i+1}{}^2 f_\gamma(\bar{\gamma}_{i+1}) - \lambda < 0. \tag{B.4b}$$

Therefore, the other condition is satisfied. In a similar way, we can prove if the second condition is not satisfied, the first one is satisfied. We can conclude that if the condition (3.23) is satisfied, then the Lagrangian will be increasing towards the boundaries (i.e., infeasible regions). In other words, no solution will be on the boundaries $\bar{\gamma}_i = \bar{\gamma}_{i+1}, \forall i$.

Appendix C

Unique SNR Threshold to Decide Number of Layers

First we can reformulate the Lagrangian in (3.21) as $L(\bar{\gamma}, \lambda) = \sum_{i=1}^{M} L_i(\bar{\gamma}_i, \lambda) - \lambda$, where $L_i(\bar{\gamma}_i, \lambda)$ can be expressed as

$$L_i(\bar{\gamma}_i, \lambda) = c_i \, F_\gamma(\bar{\gamma}_i) + \lambda \frac{\Delta b_i}{\bar{\gamma}_i} \tag{C.1}$$

Substituting with the critical points in (3.24b) of λ minimizing the Lagrangian, we can write (C.1) as

$$L_i(\bar{\gamma}_i, \lambda) = c_i \, \left(F_\gamma(\bar{\gamma}_i) + \bar{\gamma}_i f_\gamma(\bar{\gamma}_i)\right) \tag{C.2}$$

It was shown that $\bar{\gamma}_i = \infty$ is a valid solution for (3.24a), which corresponds $L_i(\bar{\gamma}_i, \lambda) = c_i$, and $F_\gamma(\bar{\gamma}_i) + \bar{\gamma}_i f_\gamma(\bar{\gamma}_i) = 0$. However, making $\bar{\gamma}_i = \infty$ implies as well that $\bar{\gamma}_j = \infty, \, \forall j > i$ Another possible solution for $\bar{\gamma}_i$ is on the rising edges satisfying (3.24b) as seen in Figure 3.4, which corresponds $L_i(\bar{\gamma}_i, \lambda)$ satisfying (C.2). The solution on the falling edges will always results in higher values for the Lagrangian. So we need $F_\gamma(\bar{\gamma}_i) + \bar{\gamma}_i f_\gamma(\bar{\gamma}_i) < 0$, for $\bar{\gamma}_i$ to be on the rising edges so that the Lagrangian is minimized. Also, for the same reason, it must be that $F_\gamma(\bar{\gamma}_j) + \bar{\gamma}_j f_\gamma(\bar{\gamma}_j) > 0$, for $\bar{\gamma}_j = \infty, \, \forall j > i$.

We can define the following function $Y(\bar{\gamma}) = 1 - F_\gamma(\bar{\gamma}_j) - \bar{\gamma}_j f_\gamma(\bar{\gamma}_j)$. Let's define $\bar{\gamma}_w$ where $Y(\bar{\gamma}_w) = 0$. It can be easily shown that $Y(\bar{\gamma})$ is a quasi convex function with a single zero value for bounded $\bar{\gamma}_w$. We can also observe that $Y(0) = 1$ and $Y(\infty) = 0$. Thus, we can conclude that $Y(\bar{\gamma})$ starts with a positive value at $\bar{\gamma} = 0$ and decreases monotonically until it reaches zero at $\bar{\gamma}_w$. Then, it reaches a global minimum value at $\bar{\gamma}_p$ defined by (3.26), and then starts to increase approaching zero at $\bar{\gamma} = \infty$. Also, $Y(\bar{\gamma})$ is positive-valued for the region $[0, \bar{\gamma}_w[$, and negative in $]\bar{\gamma}_w, \infty[$.

Therefore, we can have the following conclusion as described in [19]. If for a given λ, the critical point obtained by (3.24b) results in $\bar{\gamma}_i \in [0, \bar{\gamma}_w[$, then using this finite $\bar{\gamma}_i$ is better than using $\bar{\gamma} = \infty$. On the other hand, if at the critical point (3.24b), $\bar{\gamma}_i > \bar{\gamma}_w$, then for all layers $j \geq i$, using the infinite threshold results in a lower value for $L_i(\bar{\gamma}_i, \lambda)$ and thus a lower value for the Lagrangian (3.21).

By looking at Figure 3.4, it can be seen that increasing λ causes the values of all $\bar{\gamma}_i$ on the rising edges to increase while keeping their order, and hence the optimal number of transmitted layers will decrease gradually by one layer at a time as we increase λ.

Bibliography

[1] J. G. Proakis. *Digital Communications*. McGraw-Hill International Editions, third edition, 1995.

[2] T. S. Rappaport. *Wireless Communications: Principles & Practice*. Prentice-Hall, 2002.

[3] W. Equitz and T. Cover. Successive refinement of information. *IEEE Transactions on Information Theory*, 37(3):269–275, March 1991.

[4] B. Rimoldi. Successive refinement of information: Characterization of the achievable rates. *IEEE Transactions on Information Theory*, 40(1):253–259, January 1994.

[5] P. Bergmans. Random coding theorem for broadcast channels with degraded components. *IEEE Transactions on Information Theory*, 19(2):197–207, March 1973.

[6] S. Shamai. A broadcast approach for the Gaussian slowly fading channel. In *Proceedings IEEE International Symposium on Information Theory (ISIT)*, page 150, Ulm, Germany, July 1997.

[7] T. Cover and J. Thomas. *Elements of Information Theory*. Wiley-Interscience, second edition, July 2006.

[8] M. Shaqfeh, W. Mesbah, and H. Alnuweiri. Utility maximization for layered multimedia transmission via orthogonal multiplexing. In *Proceedings IEEE 17th International Conference on Telecommunications (ICT)*, pages 211–218, Doha, Qatar, April 2010.

[9] S. Sesia, G. Caire, and G. Vivier. Lossy transmission over slow-fading AWGN channels: A comparison of progressive, superposition and hybrid approaches. In *Proceedings IEEE International Symposium on Information Theory (ISIT)*, pages 224–228, September 2005.

[10] U. Mittal and N. Phamdo. Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communications. *IEEE Transactions on Information Theory*, 48(5):1082–1102, May 2002.

[11] W. Mesbah, M. Shaqfeh, and H. Alnuweiri. Rate maximization of multilayer transmission over Rayleigh fading channels. In *Proceedings IEEE International Symposium on Information Theory (ISIT)*, pages 2074–2078, Istanbul, Turkey, July 2013.

[12] S. Shamai and A. Steiner. A broadcast approach for a single-user slowly fading MIMO channel. *IEEE Transactions on Information Theory*, 49(10):2617–2635, October 2003.

[13] C. Ng, D. Gunduz, A. Goldsmith, and E. Erkip. Distortion minimization in Gaussian layered broadcast coding with successive refinement. *IEEE Transactions on Information Theory*, 55(11):5074–5086, November 2009.

[14] W. Mesbah, M. Shaqfeh, and H. Alnuweiri. Distortion minimization in layered broadcast transmission of a Gaussian source over Rayleigh channels. In *Proceedings IEEE Information Theory Workshop (ITW)*, Seville, Spain, September 2013.

[15] E. C. Van Der Meulen. Three-terminal communication channels. *Advances in Applied Probability*, 3:120–154, 1971.

[16] T. Cover and A. El Gamal. Capacity theorems for the relay channel. *IEEE Transactions on Information Theory*, 25(5):572–584, September 1979.

[17] M. A. Attia, M. Shaqfeh, K. Seddik, and H. Alnuweiri. Power optimization for layered transmission over decode-and-forward relay channels. In *10th International Wireless Communications and Mobile Computing Conference (IWCMC)*, Nicosia, Cyprus, August 2014.

[18] M. A. Attia, M. Shaqfeh, K. Seddik, and H. Alnuweiri. Optimal power allocation for layered broadcast over amplify-and-forward relay channels. In *2nd Global Conference on Signal and Information Processing (GlobalSIP)*, Atlanta, Georgia, USA, December 2014.

[19] M. Shaqfeh, W. Mesbah, and H. Alnuweiri. Utility maximization for layered transmission using the broadcast approach. *IEEE Transactions on Wireless Communications*, 11(3):1228–1238, March 2012.

[20] W. Mesbah, M. Shaqfeh, and H. Alnuweiri. Jointly optimal rate and power allocation for multilayer transmission. *IEEE Transactions on Wireless Communications*, 13(2): 834–845, February 2014.

[21] J. Nicholas Laneman, D. Tse, and G. Wornell. Cooperative diversity in wireless networks: Efficient protocols and outage behavior. *IEEE Transactions on Information Theory*, 50(12):3062–3080, December 2004.

[22] L. Lai, K. Liu, and H. El Gamal. The three node wireless network: Achievable rates and cooperation strategies. *IEEE Transactions on Information Theory*, 52(3):805–828, March 2006.

[23] M. Shaqfeh and H. Alnuweiri. Joint power and resource allocation for block-fading relay-assisted broadcast channels. *IEEE Transactions on Wireless Communications*, 10 (6):1904–1913, June 2011.

[24] G. Kramer, M. Gastpar, and P. Gupta. Cooperative strategies and capacity theorems for relay networks. *IEEE Transactions on Information Theory*, 51(9):3037–3063, September 2005.

[25] A. Zafar, M. Shaqfeh, M.-S. Alouini, and H. Alnuweiri. Exploiting multi-user diversity and multi-hop diversity in dual-hop broadcast channels. *IEEE Transactions on Wireless Communications*, 12(7):3314–3325, July 2013.

[26] D. Gunduz and E. Erkip. Source and channel coding for cooperative relaying. *IEEE Transactions on Information Theory*, 53(10):3453–3475, October 2007.

[27] S. Chapra and R. Canale. *Numerical Methods for Engineers*. McGraw-Hill Science/Engineering/Math, sixth edition, 2009.

[28] M. Shaqfeh, W. Mesbah, and H. Alnuweiri. Utility maximization for layered broadcast over Rayleigh fading channels. In *Proceedings IEEE International Conference on Communications (ICC)*, pages 1–6, Cape Town, South Africa, May 2010.

[29] M. Shaqfeh, W. Mesbah, and H. Alnuweiri. Optimal power allocation for layered multimedia transmission via broadcast over Rayleigh fading channels. In *Proceedings IEEE International Conference on Multimedia Computing and Information Technology (MCIT 2010)*, pages 113–116, Sharjah, UAE, March 2010.

[30] E. Dahlman, S. Parkvall, J. Sköld, and P. Beming. *3G Evolution: HSPA and LTE for Mobile Broadband*. Academic Press, Elsevier, second edition, 2008.

[31] C. Tian, A. Steiner, S. Shamai, and S. Diggavi. Successive refinement via broadcast: Optimizing expected distortion of a Gaussian source over a Gaussian fading channel. *IEEE Transactions on Information Theory*, 54(7):2903–2918, July 2008.

[32] C. Ng, D. Gunduz, A. Goldsmith, and E. Erkip. Minimum expected distortion in Gaussian layered broadcast coding with successive refinement. In *Proceedings IEEE International Symposium on Information Theory (ISIT)*, pages 2226–2230, Nice, France, June 2007.

[33] C. Ng, D. Gunduz, A. Goldsmith, and E. Erkip. Recursive power allocation in Gaussian layered broadcast coding with successive refinement. In *Proceedings IEEE International Conference on Communications (ICC)*, pages 889–896, Glasgow, Scotland, June 2007.

[34] Y. Liu, K. Lau, O. Takeshita, and M. Fitz. Optimal rate allocation for superposition coding in quasi-static fading channels. In *Proceedings IEEE International Symposium on Information Theory (ISIT)*, page 111, Lausanne, Switzerland, July 2002.

[35] F. Etemadi and H. Jafarkhani. Rate and power allocation for layered transmission with superposition coding. *IEEE Signal Processing Letters*, 14(11):773–776, November 2007.

[36] P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*. Academic Press, 1981.

[37] W. Yu and R. Lui. Dual methods for nonconvex spectrum optimization of multicarrier systems. *IEEE Transactions on Communications*, 54(7):1310–1322, July 2006.

[38] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, March 2004.