國立交通大學

電信工程研究所

碩士論文

在平行高斯通道下考慮固定碼長與碼率的編碼條件之吞吐量基準資源分配方式

Throughput-Oriented Power Allocation Policies for Parallel Gaussian Channels Under Finite-Length and Fixed-Rate Coding Constraints

研 究 生：張紊傑

指導教授：陳伯寧

中 華 民 國 一百 年 六 月

在平行高斯通道下考慮固定碼長與碼率的編碼條件之吞吐量基準資源分配方式

Throughput-Oriented Power Allocation Policies for Parallel Gaussian Channels Under Finite-Length and Fixed-Rate Coding Constraints

研究生：張紊傑　　Student：Wen-Chieh Chang

指導教授：陳伯寧　　Advisor：Po-Ning Chen

國 立 交 通 大 學
電信工程研究所

碩 士 論 文

A Thesis

Submitted to Institute of Computer and Information Science

College of Electrical Engineering and Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer and Information Science

June 2011

Hsinchu, Taiwan, Republic of China

中華民國一百年六月

在平行高斯通道下考慮固定碼長與碼率的編碼條件之吞吐量基準資源分配方式

學生：張紊傑　　　　　　　　　　　　　　　　　指導教授：陳伯寧

國立交通大學電信工程研究所碩士班

摘　　　要

傳統探討平行無記憶性高斯通道資源分配的問題都是以最大化整體通道容量為目的，而所推得的最佳能量分配法為著名的充水(Water-Filling)演算法。但是此一能量分配法則的問題為，各管道達到最佳能量分配的碼率是總能量的函數，所得的最佳碼率常是一個實作上不易實現的實數，同時隨著訊號雜訊比增大還需不斷更換最佳編碼方式，如此才能達到原本設定的最大通道容量，另外最大通道容量是成立於碼長趨近於無限大的情況，在實際有限的碼長下，此種能量分配方式是否可以達到最佳系統效能(即系統有效吞吐量)值得探討，故而本論文直接探討在有限碼長、固定編碼方式下以達到最大有效系統吞吐量為目的的傳送能量分配策略。結果顯示，我們所提出的資源分配方式可以以類似充水演算法的概念以圖形詮釋，並且幾乎達到最大系統有效吞吐量。

Throughput-Oriented Power Allocation Policies for Parallel Gaussian Channels Under Finite-Length and Fixed-Rate Coding Constraints

Student: Wen-Chieh Chang Advisor: Po-Ning Chen

Institute of Communications Engineering

National Chiao Tung University

Abstract

The common criterion used in the power allocation problem for parallel memoryless Gaussian channels is to maximize overall mutual information (namely, to achieve the capacity), resulting in the well-known water-filling policy. Such a capacity-achieving power allocation, although theoretically interesting and beneficial in conceptually elucidating the behavior of coding systems, does not match well with practical situations as capacity is an asymptotic rate requiring the codeword length to grow to infinity. In addition, the overall system capacity can only be achieved when the coding scheme of each channel is optimally and continuously adapted to the allotted power. However in a practical system, the adopted codes are by no means optimal in terms of achieving capacity and have only a finite number of rate choices. Furthermore, a common quantity of interest is the effective system throughput. In light of these observations, we study in this paper the problem of determining the power allocation strategy for a system of coded parallel Gaussian channels with the objective of maximizing effective throughput under finite-length and fixed-rate coding constraints. An approximating formula of the system's effective throughput is proposed for the case of convolutional codes and used to identify the optimal power allocation for each parallel channel. Our results show that the proposed power allocation policies can be graphically represented as a variation of the water-filling principle and achieves a near-optimal throughput.

Acknowledgements

I would like to express my heartily gratitude to my advisor, Professor Po-Ning Chen. This thesis would not be possible without his wisdom in guidance and his patience to me. I am deeply encouraged by his enthusiasm in research and the way he looks at life. To Professor Chung-Hsuan Wang and Professor Fady Alajaji, I deeply appreciate their invaluable instructions in the research. To Dr. Shih-Wei Wang, I am truly thankful for all the solid suggestions from him.

I would also like to thank all the members in the NTL lab as well as all the people that have ever helped me in the past two years. Finally, I would like to dedicate this thesis to my family, who always provide me with love and spiritual support.

Contents

List of Figures

Chapter 1

Introduction

1.1 Overview

Finding the best strategy for allocating power over parallel independent additive white Gaussian noise (AWGN) channels is a classical problem in information theory (e.g., cf. [1, 6] and the references therein). For this problem, a typical optimization criterion for the distribution of power is to maximize the system's mutual information (namely, to achieve the system's capacity), which results in the well-known *water-filling* scheme. In this scheme, the capacity is achieved when the input of each parallel channel is Gaussian distributed and has a power allotment given by the "water level" of its respective "vessel" with base height equal to the channel's noise variance [1]. In 2006, Lozano, Tulio and Verdú re-visited this problem by judiciously constraining the input to be drawn from discrete modulation constellations used in practice such as phase-shift keying (PSK) [9] and quadrature amplitude modulation. They concluded the study with a refined optimal power allocation policy, referred to as *mercury water-filling* [6]. The result was obtained based on two key observations regarding parallel Gaussian channels: (*i*) both the mutual information and the minimum mean-square error are functions of the signal-to-noise ratios (SNRs), and (*ii*) the derivative of the former measure with respect to the SNR is equal to the latter one.

In literature, there is another challenging power allocation problem over parallel Gaussian channels. Instead of knowing the noise variance in each channel, only the total noise variance is known. In such case, the criterion becomes to maximize the so-called *worst-case mutual information*, defined as the smallest mutual information among all possible noise variance distributions with variance sum equal to a given constant. Signal power is then allotted to achieve the *worst-case capacity*, which is the maximum of the *worst-case mutual information* among all power allocations. The resulting power allocation policy is to allot equal signal power to each channel, regardless of the value of the total noise variance.

The capacity-oriented power allocation, although theoretically interesting and useful for the analysis of channel coded systems, is not realistic in several aspects. First, channel capacity is a function of the total system power, and the optimal coding scheme that achieves capacity may be different for different capacity values. Hence, optimality can be achieved only when the coding scheme of each parallel channel can be optimally adapted to the power allotment, which is difficult to fulfill in practice. Secondly, the optimal rate obtained from a capacity-based power allocation is often a concretely unrealizable real number; this is in contrast with practical systems whose code rates are usually restricted to only a few rational numbers such as 1/2, 1/3, 2/3, 1/4, etc. Finally, capacity is an asymptotic quantity that requires the coding blocklength or frame size to grow without bound; yet, in practical systems, the blocklength is finite (typically preset as a function of the system's delay requirements).

In view of the above points, we herein investigate power allocation policies that respectively achieve the maximum effective throughput (instead of capacity) and maximum worst-case effective throughput (instead of worst-case capacity) for convolutionally coded parallel memoryless Gaussian channels with finite-length and fixed-rate coding constraints, where effective throughput is defined as the number of successfully transmitted information bits per channel use. Since in general, there is no closed-form formula for the error rate (and hence effective throughput) of a coded system, the optimal solution can only be obtained via case-by-case simulation. Our study however shows that it is possible to obtain a good approximating expression for the error rate of each coded channel and then use these approximations to derive the near-optimal power allocation policies as a function of the system's total power and noise variances. The resulting near-optimal-throughput power allocation policies are reminiscent of a variation of the traditional water-filling principle, where the base width and height of each individual vessel (corresponding to each parallel channel) now become functions of the code characteristics. However, unlike the case of water-filling, we obtain that when a channel is in use (or active), a minimal power should be allocated to it. In the effective-throughput-optimizing problem, we show that the optimal power assigned to each channel may experience a sudden jump (or discontinuity) when the total system power increases. This is due to the practical constraint requiring the code rates to be fixed and positive, and hence for a given channel, a power allotment that is smaller than a certain value can only result in an inferior overall throughput. In the worst-case-effective-throughput-optimizing problem, we provided a near-optimal power allocaton policy for system SNR greater than a certain threshold. We show that our proposed power allocation policy yields better gain than the traditional equal power allocation if the differences in the characteristics of the used codes between channels are larger.

The rest of the thesis is organized as follows. In Chapter 2, we prove the optimality of water-filling policy and equal power allocation policy for capacity-achieving and worst-case-capacity-achieving problems, respectively. In Chapter 3, we introduce the system model and define the throughput-optimizing and worst-case-throughput-optimizing power allocation problems. We then propose the near-optimal power allocation policies based on convolutional codes in Chapter 4 and present numerical and simulation results in Chapter 5. Finally, we conclude the thesis in Chapter 6.

Chapter 2

Preliminaries

As aforementioned, the traditional power allocation policy is to maximize the overall system mutual information. Under a common assumption that the noise variance of each individual channel is known, the resulting optimal power allocation that achieves the system capacity is the water-filling power allocation scheme. For an alternative case where only the total noise variance is known, the optimal power allocation that achieves the system worst-case capacity is to allocate equal power to each channel. For the two cases mentioned above, proofs of the optimality of the claimed power allocation policies will be respectively given in the two sections in this chapter.

2.1 Capacity-Achieving Water-Filling Power Allocation Policy with Known Noise Variance in Each Channel

Consider a system with K parallel AWGN channels. Assume that the noises are independent of each other. Denote the noise variance for channel i by σ_i^2, $1 \leq i \leq K$. The capacity of K parallel AWGN channels is then give by

$$\max_{\sum_{i=1}^{K} P_i \leq P_{\mathrm{t}}} \sum_{i=1}^{K} \frac{1}{2} \log\left(1 + \frac{P_i}{\sigma_i^2}\right). \tag{2.1}$$

We denote for convenience $\underline{P} = \{P_i\}_{i=1}^{K}$ as its assemble format. Since (2.1) is concave over $\underline{P}$, the technique of Lagrange multiplier [7] can be applied as the following. Let

$$f\left(\underline{P}\right) = \sum_{i=1}^{K} \frac{1}{2} \log\left(1 + \frac{P_i}{\sigma_i^2}\right) - \lambda \left(\sum_{i=1}^{K} P_i - P_{\mathrm{t}}\right), \tag{2.2}$$

where the constant λ is the so-called Lagrange multiplier and is always chosen such that the power-sum constraint $\sum_{i=1}^{K} P_i = P_{\mathrm{t}}$ is satisfied. Taking the derivative of (2.2) with respective to P_i, we have from the Kuhn-Tucker condition that

$$\begin{cases} \frac{1}{2\left(P_i + \sigma_i^2\right)} = 0, & \text{if} \quad P_i > 0; \\ \frac{1}{2\left(P_i + \sigma_i^2\right)} \leq 0, & \text{if} \quad P_i = 0. \end{cases}$$

Hence,

$$P_i = \left(\nu - \sigma_i^2\right)^+, \tag{2.3}$$

where $(x)^+ \triangleq \max\{0, x\}$ and ν is chosen such that the power-sum constraint is satisfied (and is equal to $-\frac{1}{2\lambda}$). (2.3) can be graphically interpreted via a water-filling scheme as shown in Figure. 2.1.

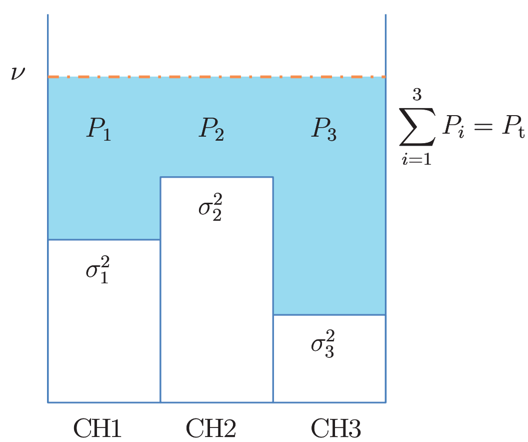

Figure 2.1: An example of the water-filling power allocation for $K = 3$.

2.2 Worst-Case-Capacity-Achieving Water-Filling Power Allocation Policy Subject to a Known Sum of Noise Variances

Instead of knowing the noise variance, σ_i^2, of each channel, we suppose that we obtain only the information of total noise variance, $\sigma_{\mathrm{t}}^2 = \sum_{i=1}^K \sigma_i^2$. Lacking the knowledge of noise variance in each channel, we need to consider the worst case scenario, where for any given power allocation, σ_i^2 is always chosen such that the system mutual information is minimized. It is named the *worst-case mutual information* and is defined as

$$\min_{\sum_{i=1}^K \sigma_i^2 = \sigma_{\mathrm{t}}^2} \sum_{i=1}^K \frac{1}{2} \log\left(1 + \frac{P_i}{\sigma_i^2}\right). \tag{2.4}$$

Based on (2.4), the optimal power allocation is chosen such that (2.4) is maximized subject to the power constraint $\sum_{i=1}^K P_i = P_{\mathrm{t}}$. The first step toward this problem is to find the $\{\sigma_i^2\}$ that achieves the worst-case mutual information. It can be derived that the second-order derivative of (2.4) with respect to each σ_i^2 is positive; hence, (2.4) is a convex function of σ_i^2. We can then apply the Lagrange multiplier technique to find the optimal $\{\sigma_i^2\}$ that achieves the worst-case mutual information. Let

$$f_1(\underline{P}) = \sum_{i=1}^K \frac{1}{2} \log\left(1 + \frac{P_i}{\sigma_i^2}\right) + \lambda_1 \left(\sum_{i=1}^K \sigma_i^2 - \sigma_{\mathrm{t}}^2\right), \tag{2.5}$$

where λ_1 is chosen such that $\sum_{i=1}^K \sigma_i^2 = \sigma_{\mathrm{t}}^2$. Taking derivative of (2.5) with respect to σ_i^2, we have

$$\begin{cases} \frac{P_i}{2\left(\left(\sigma_i^2\right)^2 + P_i \sigma_i^2\right)} + \lambda_1 = 0, & \text{if } \sigma_i^2 > 0; \\ \frac{P_i}{2\left(\left(\sigma_i^2\right)^2 + P_i \sigma_i^2\right)} + \lambda_1 < 0, & \text{if } \sigma_i^2 = 0. \end{cases} \tag{2.6}$$

(2.6) can be reorganized into a second-order polynomial function of σ_i^2 as the following:

$$\begin{cases} \left(\sigma_i^2\right)^2 + P_i \sigma_i^2 + \frac{P_i}{2\lambda_1} = 0, & \text{if } \sigma_i^2 > 0; \\ \left(\sigma_i^2\right)^2 + P_i \sigma_i^2 + \frac{P_i}{2\lambda_1} < 0, & \text{if } \sigma_i^2 = 0. \end{cases} \tag{2.7}$$

Hence, we obtain

$$\sigma_i^2 = \left(\frac{-P_i + \sqrt{(P_i)^2 - \frac{2P_i}{\lambda_1}}}{2} \right)^+ . \tag{2.8}$$

We then replace the σ_i^2 in (2.4) by (2.8) and take away the minimization, and (2.4) becomes

$$\sum_{i=1}^{n} \frac{1}{2} \log \left(1 + \frac{P_i}{\frac{-P_i + \sqrt{(\gamma P_i)^2 - \frac{2P_i}{\lambda_1}}}{2}} \right) . \tag{2.9}$$

Since (2.9) is concave with respect to P_i, the Lagrange multiplier technique can be applied again as the following. Let

$$f_2(\underline{P}) = \sum_{i=1}^{K} \frac{1}{2} \log \left(1 + \frac{P_i}{\frac{-P_i + \sqrt{(P_i)^2 - \frac{2P_i}{\lambda_1}}}{2}} \right) + \lambda_2 \left(\sum_{i=1}^{K} P_i - 1 \right), \tag{2.10}$$

where λ_2 is chosen such that $\sum_{i=1}^{K} P_i = P_{\mathrm{t}}$. The derivative of (2.10) with respect to P_i becomes

$$\frac{\partial f_2(\underline{P})}{\partial P_i} = \left(\frac{1}{\left(-1 + \sqrt{1 - \frac{2}{\lambda_1 P_i}}\right)^2 + 2(-1 + \sqrt{1 - 2\frac{2}{\lambda_1 P_i}})} \right) \left(\frac{1}{\sqrt{(1 - \frac{2}{\lambda_1 P_i}} \lambda_1 P_i^2} \right) + \lambda_2. \tag{2.11}$$

(2.11) should satisfy

$$\begin{cases} \frac{\partial f_2(\underline{P})}{\partial P_i} = 0, & \text{if } P_i > 0; \\ \frac{\partial f_2(\underline{P})}{\partial P_i} < 0, & \text{if } P_i = 0. \end{cases}$$

For any $j \neq i$, we accordingly have that if $P_i > 0$ and $P_j > 0$,

$$\frac{\partial f_2(\underline{P})}{\partial P_i} = \frac{\partial f_2(\underline{P})}{\partial P_j}. \tag{2.12}$$

This concludes that choosing $P_i = P_j$ for every i and j will be one of the optimal power allocation that maximizes (2.4).

Chapter 3

System Model and Problem Formulation

Consider a system with K parallel channels or links, each of which has a binary-antipodal-input (realized via binary PSK modulation) and suffers independent AWGN noise. Let R_i be the rate of the code adopted by channel i, and denote by $P_{\mathrm{e},i}$ its corresponding frame error rate for frame size N_i, $1 \le i \le K$. The system effective throughput is then defined as

$$T\left(\underline{P}\right) \triangleq \sum_{i=1}^{K} R_i \left(1 - P_{\mathrm{e},i}\right), \tag{3.1}$$

which corresponds to the successfully transmitted information bits per channel use. Note that in the above formula, $P_{\mathrm{e},i}$ is a function of N_i, σ_i^2 and P_i. To simplify the notations, we do not explicitly write $P_{\mathrm{e},i}$ as a function of N_i, P_i and σ_i^2.

Corresponding to the capacity-achieving problem that σ_i^2 in each channel is known to the system, we will find P_i such that (3.1) is maximized under the power constraint $\sum_{i=1}^{K} P_i = P_{\mathrm{t}}$. Similarly, corresponding to the wort-case-capacity-achieving problem that only the total noise variance is available, we will find P_i such that the worst-case effective throughput defined in (3.2) is maximized.

$$T_{\mathrm{w}}\left(\underline{P}\right) \triangleq \min_{\sum_{i=1}^{K} \sigma_i^2 = \sigma_{\mathrm{t}}^2} \sum_{i=1}^{K} R_i \left(1 - P_{\mathrm{e},i}\right) \tag{3.2}$$

In this thesis, we implicitly assume that an error-detection scheme is applied to each frame such that information is successfully transmitted only when no decoding error occurs within a frame.[1] We also assume that the time needed to transmit one code bit is identical for all channels.

In general, $P_{\mathrm{e},i}$ does not exhibit a closed-form formula. Hence, the power allocation that maximizes $T\left(\underline{P}\right)$ (respectively, $T_{\mathrm{w}}\left(\underline{P}\right)$) can be obtained only via case-by-case simulation studies. It is thus hard to establish a general power allocation principle from such a simulation-based power allocation result. One possible solution is to derive a good approximation for $P_{\mathrm{e},i}$ with a structure that can facilitate its analysis.

When transmitting a convolutional code over an AWGN channel with noise variance σ^2, the frame error rate at high SNRs can be well approximated by the event error rate [5] as

$$P_{\mathrm{e}} \approx A_{d_{\mathrm{free}}} e^{-\frac{1}{2} d_{\mathrm{free}} \frac{P}{\sigma^2}}, \tag{3.3}$$

where d_{free} is the free distance of the convolutional code, $A_{d_{\mathrm{free}}}$ is the number of codewords with Hamming weight equal to d_{free}, and P is the transmission power. However, the approximated P_{e} in (3.3) is far from accurate for moderate SNRs and finite frame sizes (cf. the *approx.* P_e curve in Figure. 3.1 for a convolutional code with rate $R = 1/4$ and memory order 6, in which $RE_b = P$ and $\sigma^2 = N_0/2$). Instead of adding more rectifying terms to (3.3) that may later introduce analytical obstacles, we choose to fix this inaccuracy by replacing $A_{d_{\mathrm{free}}}$ and d_{free} with the refined parameters A and d respectively such that the adjusted curve

[1] Alternatively, one may define the effective throughput based on the (information) bit error rate (P_b) to avoid considering the frame size, e.g.,

$$\sum_{i=1}^{K} \frac{\text{\# of info. bits successfully recovered at receiver } i}{\text{\# of total info. bits transmitted via channel } i} = \sum_{i=1}^{K} (1 - P_{b,i}).$$

This however may introduce an impractical situation where a high bit error rate (e.g., nearly one half) at the receiver can still provide a non-trivial throughput to the system. Such an impractical situation can be avoided under the definition in (3.1) since almost all frames fail the error detection check under a high bit error rate.

defined below,

$$\log(P_{\mathrm{e}}) \approx \min\left\{0, \log(A) - \frac{P}{2\sigma^2}d\right\}, \tag{3.4}$$

is close to the true P_{e} in the least squares sense over the range of operating SNRs (cf. the *adjusted approx.* P_e curve in Figure. 3.1). For details of the procedure for retrieving A and d, please see Example 3.1.

Example 3.1. *From* (3.4), *we know that the approximated* P_e *in* log *scale is a linear combination of* d_i *and* $\log A$ *in the operating SNR region. Thus linear least square estimator [8] can be applied to retrieve* d *and* $\log A$*. We let* $\boldsymbol{x} = [x[0]\, x[1] \ldots x[M-1]]^T$ *be a vector composed by* M *true* P_e *values which are in* log *scale and* $\boldsymbol{g} = [g[0]\, g[1] \ldots g[M-1]]^T$ *denote its corresponding* $\frac{E_b}{N_0}$*. We also let* $\boldsymbol{s} = [s[0]\, s[1] \ldots s[M-1]]^T$ *denote the approximated* P_e *values in* log *scale. From* (3.4), *the* $s[j]$ *can be modelled by*

$$s[j] = -R\ g[j]\ d + \log A \quad \forall\ 1 \leq j \leq M$$

or in matrix form

$$\boldsymbol{s} = \boldsymbol{H\theta},$$

where

$$\boldsymbol{H} = \begin{bmatrix} -Rg[0] & 1 \\ -Rg[1] & 1 \\ \cdot & \\ \cdot & \\ \cdot & \\ -Rg[M-1] & 1 \end{bmatrix}, \boldsymbol{\theta} = \begin{bmatrix} d \\ \log A \end{bmatrix}.$$

The least square estimator is found by minimizing

$$J(\boldsymbol{\theta}) = (\boldsymbol{x} - \boldsymbol{H\theta})^T (\boldsymbol{x} - \boldsymbol{H\theta}). \tag{3.5}$$

The gradient of (3.5) *is*

$$\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = -2\boldsymbol{H}^T\boldsymbol{x} + 2\boldsymbol{H}^T\boldsymbol{H\theta}$$

Setting the gradient to be zero yields the least square estimator

$$\hat{\boldsymbol{\theta}} = \left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{x}$$

The refined parameters A and d for the code is then found.

We denote the refined parameters in channel i by A_i and d_i, respectively. Note that the effect of N_i to the frame error rate is included in the choice of d_i and A_i. Even if we use the same code, d_i and A_i of a code will be different if we use different frame size. Once d_i and A_i for a given code is determined, it can be later used universally to find the throughput-optimizing power allocation policy for every value of $P_{\rm t}$.

An immediate consequence from the adjusted approximation formula in (3.4) is that the contribution to the system effective throughput from channel i will be zero if

$$P_i < P_{{\rm th},i} \triangleq \frac{2\sigma_i^2}{d_i}\log(A_i).$$

In Chapter 5, our simulations will confirm that for a given code assigned to channel i, allocating a power value smaller than $P_{{\rm th},i}$ indeed provides very limited contribution to the system throughput since almost all frames will fail the implicitly assumed error-detection check.

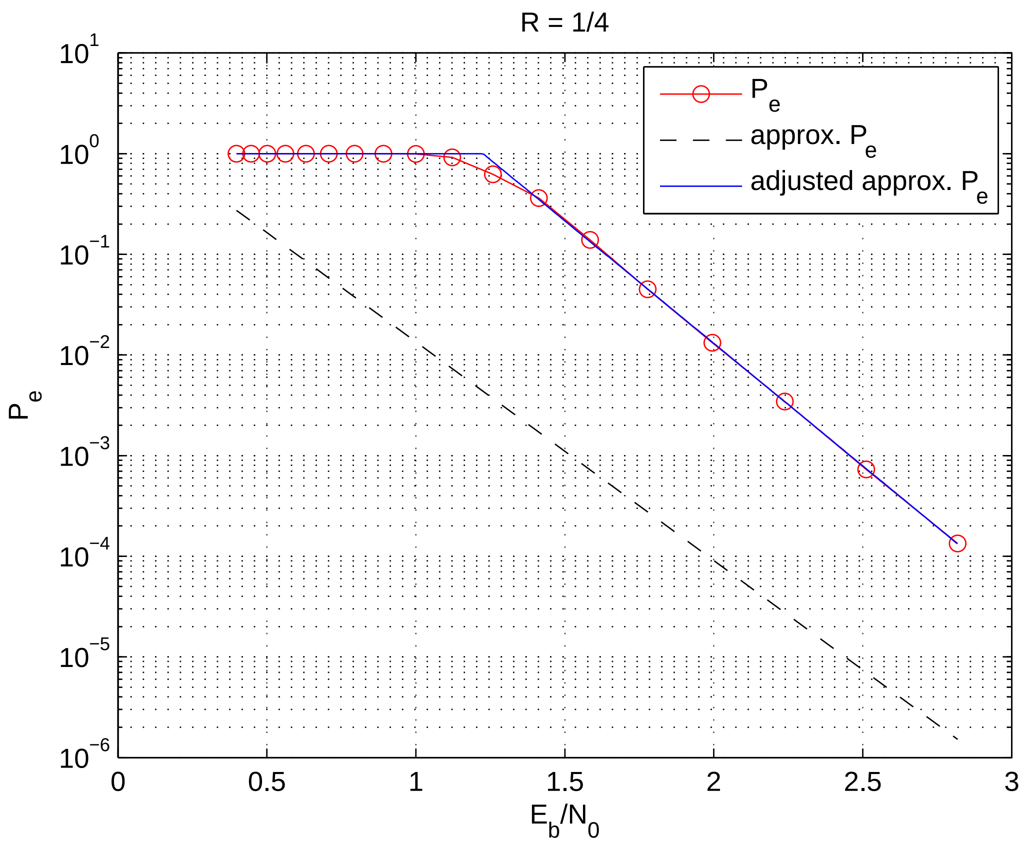

Figure 3.1: P_e and its approximations for a $(4, 1, 6)$ convolutional code with generator polynomial (in octal) being [177 127 155 171], $d_\mathrm{free} = 20$ and $A_\mathrm{dfree} = 2$. The adjusted parameters are $d = 22.42$ and $A = 962.51$. The frame size is $N = 4(500 + 6)$. E_b/N_0 is plotted in linear scale.

Chapter 4

Throughput-Oriented Water-Filling

In this chapter, instead of maximizing the overall mutual information, we suggest a power allocation that aims at maximizing the effective throughput defined in (3.1) and (3.2). The analyses in Sections 4.1 and 4.2 are mainly based on the approximated P_{e} defined in (3.4). Interestingly, both the proposed power allocation policies in the two sections can be interpreted by some variations of water-filling. At the end of each section, we will remark on the near-optimal power allocation policies we proposed when total power goes without bound, and several conclusions will be given.

4.1 Throughput-Oriented Water-Filling: Noise Variance in Each Channel is Known

Based on the adjusted approximation formula in (3.4), (3.1) becomes

$$\begin{aligned} T\left(\underline{P}\right) &= \sum_{i=1}^{K} R_i \left(1 - \min\left\{1, A_i e^{-\frac{d_i P_i}{2\sigma_i^2}}\right\}\right) \\ &= \sum_{i\in\mathcal{O}} R_i \left(1 - A_i e^{-\frac{d_i P_i}{2\sigma_i^2}}\right), \end{aligned} \tag{4.1}$$

provided that the optimal set of active channels in use, denoted by $\mathcal{O}$, can be *priori* determined. Since (4.1) is a concave function over P_i, the power allocation can be obtained by using the Lagrange multiplier technique and the Kuhn-Tucker condition as follows. Let

$$\mathrm{E}\left(\underline{P}\right) = T\left(\underline{P}\right) - \lambda\left(\sum_{i\in\mathcal{O}} P_i - P_{\mathrm{t}}\right),$$

where the constant λ is the Lagrange's multiplier and is chosen such that $\sum_{i\in\mathcal{O}} P_i = P_{\mathrm{t}}$. Taking derivative with respect to P_i, we have

$$\frac{\partial \mathrm{E}\left(\underline{P}\right)}{\partial P_i} = \frac{d_i}{2\sigma_i^2} A_i e^{-\frac{d_i P_i}{2\sigma_i^2}} - \lambda. \tag{4.2}$$

By the Kuhn-Tucker condition, (4.2) should satisfy

$$\begin{cases} \frac{\partial \mathrm{E}(\underline{P})}{P_i} = 0, & \text{if} \quad P_i > 0, \\ \frac{\partial \mathrm{E}(\underline{P})}{P_i} < 0, & \text{if} \quad P_i = 0. \end{cases} \tag{4.3}$$

By (4.2) and (4.3), the optimal P_i can be shown to have the following form:

$$P_i^* = \frac{2\sigma_i^2}{d_i}\left(\nu - \log\frac{\sigma_i^2}{d_i A_i R_i}\right)^+ \tag{4.4}$$

where ν is chosen such that $\sum_{i\in\mathcal{O}} P_i^* = P_{\mathrm{t}}$. Note that ν should also satisfy

$$\nu \geq \nu_{\min} \triangleq \max_{i\in\mathcal{O}} \log\frac{\sigma_i^2}{d_i R_i}$$

for the reason that all the channels in $\mathcal{O}$ should be activated (i.e. $P_i^* \geq P_{\mathrm{th},i} \quad \forall\, i \in \mathcal{O}$).

Interestingly, the above result can be interpreted graphically as a variation of the water-filling principle. For channels outside $\mathcal{O}$, zero power will be allocated. For each channel in $\mathcal{O}$, a vessel with base width $\frac{2\sigma_i^2}{d_i}$ and base height $\log\frac{\sigma_i^2}{d_i A_i R_i}$ will be used for water filling. The resulting water level ν must be no less than the base height $\log\frac{\sigma_i^2}{d_i A_i R_i}$ plus $\log(A_i)$ for every $i \in \mathcal{O}$. The water inside each vessel is then the optimal power to be allotted. An example is illustrated in Figure. 4.1.

Example 4.1. *A three channels ($K = 3$) system is considered. Each channel has its base height and base width as stated in the paragraph immediately above this example. We assume that $\mathcal{O} = \{1, 2\}$ has already been given. Thus, zero power will be allocated to channel 3. And* *at least* $P_{th,1}$ *and* $P_{th,2}$ *should be allocated to channels* 1 *and* 2*, respectively. The lowest water level* $\nu_{\min}$ *should then be chosen as*

$$\nu_{\min} = \max\left\{\log\frac{\sigma_1^2}{d_1 R_1}, \log\frac{\sigma_2^2}{d_2 R_2}\right\}.$$

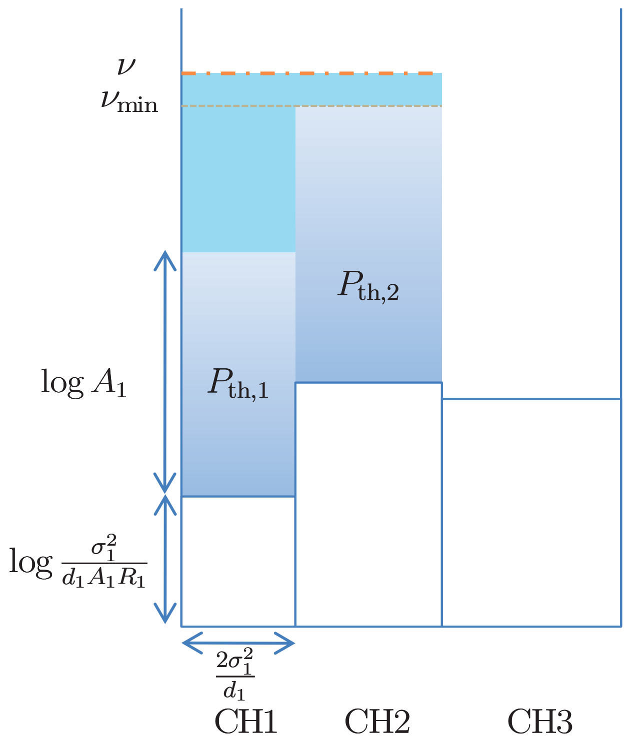

Figure 4.1: An example of the throughput-oriented water-filling with $K = 3$ and $\mathcal{O} = \{1, 2\}$.

By (4.4), we obtain that for $i \in \mathcal{O}$ (hence $\nu \geq \log \frac{\sigma_i^2}{d_i R_i}$),

$$d_i \gamma_i^* = 2\nu + 2\log(d_i A_i R_i) - 2\log(\sigma_i^2) \tag{4.5}$$

where $\gamma_i^* \triangleq P_i^*/\sigma_i^2$ denotes the SNR of channel i. Equation (4.5) then indicates that the optimal power allocation should make the SNR γ_i^* inversely proportional to the logarithm of noise power σ_i^2. This is in stark contrast with the capacity-achieving water-filling policy (with Gaussian inputs), which results in an SNR that is inversely proportional to the noise power itself. For example, when two active channels i and j adopt the same code with

$d_i = d_j = 5$, and $\sigma_j^2/\sigma_i^2 = 2$, (4.5) implies that

$$\gamma_i^* = \gamma_j^* + \frac{2}{d_i} \log \frac{\sigma_j^2}{\sigma_i^2} = \gamma_j^* + 0.12,$$

while the capacity-achieving power allocation formula $P_i^* = (\nu - \sigma_i^2)^+$ requires that

$$\gamma_i^* = 2\gamma_j^* + 1.$$

From our simulations, we indeed observe that the latter power assignment actually yields a poor system throughput.

When the total power P_t is adequately large, all channels become active. We then obtain from (4.5) that the SNRs of any two channels, say channels i and j, are characterized by

$$d_i \gamma_i^* = d_j \gamma_j^* + 2 \log \frac{\sigma_j^2}{\sigma_i^2} + \log \frac{d_i A_i R_i}{d_j A_j R_j}$$

Thus

$$\lim_{P_\mathrm{t} \to \infty} \frac{d_i \gamma_i^*}{d_j \gamma_j^*} = 1 + \lim_{P_\mathrm{t} \to \infty} \frac{2 \log \frac{\sigma_j^2}{\sigma_i^2} + \log \frac{d_i A_i R_i}{d_j A_j R_j}}{d_j \gamma_j^*} = 1.$$

Hence, when P_t is large, our result indicates that the allotted powers should make the $d_i \gamma_i^*$ products equal across all channels. As in most cases, the approximate d_i is close to the free distance of the code used by channel i; this suggests that, when P_t grows without bound, the optimal SNR γ_i^* should in general be chosen as the reciprocal of the code's free distance.

As already mentioned, our result also indicates that there is a minimum power required for each channel to be activated. In other words, if the allocated power P_i is less than $P_{\mathrm{th},i}$ then re-assigning this power to other channels will generally result in a better throughput.

A remaining question is how to determine the optimal $\mathcal{O}$. A straightforward approach is to examine each of the choices of $\mathcal{O}$, which is by no means complex. To examine one possible choice of $\mathcal{O}$, for a given total power, P_i is then determined by (4.4) since the relationship of P_i and total power is a one-to-one mapping. The corresponding effective throughput is obtained by simple calculation according to (4.1).

4.2 Throughput-Oriented Water-Filling: Only Total Noise Variance is Available

Based on the adjusted approximation formula in (3.4), (3.2) becomes

$$T_{\mathrm{w}}\left(\underline{P}\right) = \min_{\sum_{i=1}^{K}\sigma_i^2=\sigma_{\mathrm{t}}^2} \sum_{i=1}^{K} R_i \left(1 - \min\left\{1, A_i e^{-\frac{d_i P_i}{2\sigma_i^2}}\right\}\right). \tag{4.6}$$

We focus on the situation that all channels are active, which means that

$$P_i \geq P_{\mathrm{th},i} \quad \forall\, 1 \leq i \leq K. \tag{4.7}$$

(4.6) becomes

$$T_{\mathrm{w}}\left(\underline{P}\right) = \min_{\sum_{i=1}^{K}\sigma_i^2=\sigma_{\mathrm{t}}^2} \sum_{i=1}^{K} R_i \left(1 - A_i e^{-\frac{d_i P_i}{2\sigma_i^2}}\right). \tag{4.8}$$

A straightforward approach to eliminate the minimization in (4.8) is to use the Lagrange multiplier technique and Kuhn-Tucker condition to find σ_i^2 such that the worst-case effective throughput is achieved. However, in general, the worst-case effective throughput is not a concave function of σ_i^2 since

$$\frac{\partial^2 T_{\mathrm{w}}\left(\underline{P}\right)}{\partial\left(\sigma_i^2\right)^2} = \frac{R_i A_i d_i P_i}{\left(\sigma_i^2\right)^3} e^{-\frac{d_i P_i}{2\sigma_i^2}} \left(1 - \frac{d_i P_i}{4\sigma_i^2}\right). \tag{4.9}$$

Thus, by letting

$$P_i \geq \frac{4\sigma_i^2}{d_i}, \tag{4.10}$$

we further constrain our problem to be concave over σ_i^2 such that (4.9) becomes always negative. Under this constraint, we know that the σ_i^2 to achieve T_{w} should be chosen as either 0 or σ_{t}^2. Since the total noise variance is a given value, only one channel will be allocated the whole noise power and the rest of the channels is allocated zero noise power.

Next we consider one possible power allocation $P_i^{\dagger}$, which is defined as

$$P_i^{\dagger} = \frac{2\sigma_{\mathrm{t}}^2}{d_i}\left(\nu - \log\frac{1}{A_i R_i}\right), \tag{4.11}$$

where ν is chosen such that

$$\sum_{i=1}^{K} P_i^{\dagger} = P_{\mathrm{t}}.$$

Our aim is to prove that this power allocation performs better than any other power allocation policies and hence is optimal.

From the power constraints in (4.7) and (4.10), $P_i^{\dagger}$ should satisfy

$$P_i^{\dagger} \geq \max\left\{\frac{2\sigma_i^2}{d_i}\log\left(A_i\right), \frac{4\sigma_i^2}{d_i}\right\} \quad \forall\ 1 \leq i \leq K. \tag{4.12}$$

Since $\sigma_i^2 \leq \sigma_{\mathrm{t}}^2$, we can further increase P_{t} (equivalently, ν) such that

$$P_i^{\dagger} \geq \max\left\{\frac{2\sigma_{\mathrm{t}}^2}{d_i}\log\left(A_i\right), \frac{4\sigma_{\mathrm{t}}^2}{d_i}\right\} \quad \forall\ 1 \leq i \leq K. \tag{4.13}$$

Define the minimum power required in channel i as

$$P_{\mathrm{th},i}^{\dagger} \triangleq \max\left\{\frac{2\sigma_{\mathrm{t}}^2}{d_i}\log\left(A_i\right), \frac{4\sigma_{\mathrm{t}}^2}{d_i}\right\} \quad \forall\ 1 \leq i \leq K.$$

Replacing $P_i^{\dagger}$ by (4.11), we further deduce (4.13) as a condition on ν,

$$\frac{2\sigma_{\mathrm{t}}^2}{d_i}\left(\nu - \log\frac{1}{A_i R_i}\right) \geq \max\left\{\frac{2\sigma_{\mathrm{t}}^2}{d_i}\log A_i, \frac{4\sigma_{\mathrm{t}}^2}{d_i}\right\},$$

thereby implying

$$\nu \geq \max\left\{-\log R_i, 2 - \log\left(A_i R_i\right)\right\} \quad \forall\ 1 \leq i \leq K. \tag{4.14}$$

We let $\nu_{\min}$ denote the minimum value of the choice of ν, which satisfies (4.14) for every i. From the definition of $P_i^{\dagger}$ and (4.14), we have

$$P_i^{\dagger} \geq \frac{2\sigma_{\mathrm{t}}^2}{d_i}\left(\nu_{\min} - \log\left(\frac{1}{A_i R_i}\right)\right) \quad \forall\ 1 \leq i \leq K. \tag{4.15}$$

(4.15) equivalently implies a constraint in system SNR by taking summation over $P_i^{\dagger}$ and dividing it by σ_{t}^2, which is

$$\frac{1}{\sigma_{\mathrm{t}}^2}\sum_{i=1}^{K} P_i^{\dagger} \geq \sum_{i=1}^{K}\frac{2}{d_i}\left(\nu_{\min} - \log\left(\frac{1}{A_i R_i}\right)\right) = \gamma_{\mathrm{th}}^{\dagger},$$

where $\gamma_{\text{th}}^{\dagger}$ is the threshold system SNR. From above, we have claimed that by using $P_i^{\dagger}$ as power allocation, the optimal choice of σ_i^2 is either 0 or σ_{t}^2, when system SNR is greater than $\gamma_{\text{th}}^{\dagger}$. Using this result, the worst-case effective throughput due to $\{P_i^{\dagger}\}$, which is denoted by $T_{\text{w}}^{\dagger}$ can be computed as follows.

$$\begin{aligned} T_{\text{w}}^{\dagger}\left(\underline{P}^{\dagger}\right) &\triangleq \min_{\sum_{\sigma_i^2=\sigma_{\text{t}}^2}} \sum_{i=1}^{K} R_i \left(1 - A_i e^{-\frac{d_i P_i^{\dagger}}{2\sigma_i^2}}\right) \\ &= \min_{\sum_{\sigma_i^2=\sigma_{\text{t}}^2}} \sum_{i=1}^{K} R_i \left(1 - A_i e^{-\frac{d_i \frac{2\sigma_{\text{t}}^2}{d_i}\left(\nu - \log \frac{1}{R_i A_i}\right)}{2\sigma_i^2}}\right) \end{aligned} \tag{4.16}$$

We define the noise variance that achieves $T_{\text{w}}^{\dagger}$ as

$$\sigma_i^{2\dagger} \triangleq \begin{cases} \sigma_{\text{t}}^2, & \text{if} \quad i = m \\ 0, & \text{if} \quad i \neq m \end{cases}, \tag{4.17}$$

where m can be chosen to be $1 \leq m \leq K$. We will later show that any value of m yields the same $T_{\text{w}}^{\dagger}$. We then take away the minimization in (4.16) by using $\sigma_i^{2\dagger}$ as the noise power. (4.16) becomes

$$\begin{aligned} &\sum_{i=1}^{K} R_i \left(1 - A_i e^{-\frac{d_i \frac{2\sigma_{\text{t}}^2}{d_i}\left(\nu - \log \frac{1}{R_i A_i}\right)}{2(\sigma_i^{\dagger})^2}}\right) \\ &= \sum_{i \neq m} R_i + R_m - e^{-\left(\nu - \log \frac{1}{R_m A_m}\right) + \log R_m A_m} \\ &= \sum_{i=1}^{K} R_i - e^{-\nu}. \end{aligned} \tag{4.18}$$

By (4.18), it is noted that the worst-case effective throughput is independent of m. Thus definition of $\sigma_i^{2\dagger}$ is justified. Besides, we know that $R_m - e^{-\nu}$ is always non-negative for all possible value of m from (4.14).

The optimality of using $P_i^{\dagger}$ as power allocation is proved by the method of contradiction. We will show that the T_{w} obtained from any other power allocation is less than or equal to

$T_{\mathrm{w}}^{\dagger}$. The proof is as follows. Consider any power allocation $\hat{P}_i = P_i^{\dagger} + \triangle P_i^{\dagger}$, where $\triangle P_i^{\dagger} \neq 0$ for at least one channel and

$$\sum_{i=1}^{K} \triangle P_i^{\dagger} = 0. \tag{4.19}$$

We also consider a specific noise power allocation

$$\hat{\sigma_i}^2 = \begin{cases} \sigma_{\mathrm{t}}^2 & \text{if} \quad i = k \\ 0 & \text{if} \quad i \neq k \end{cases},$$

where

$$k \triangleq \arg \min_{1 \leq i \leq K} \left(\frac{d_i P_i}{2\sigma_{\mathrm{t}}^2} - \log A_i R_i \right). \tag{4.20}$$

An upper bound for the worst-case effective throughput of $\underline{\hat{P}}$ can be found as the following:

$$\begin{aligned} & \min_{\sum \sigma_i^2 = \sigma_{\mathrm{t}}^2} \sum_{i=1}^{K} R_i \left(1 - \min \left\{ 1, A_i e^{-\frac{d_i \hat{P}_i}{2\sigma_i^2}} \right\} \right) \\ \leq & \sum_{i=1}^{K} R_i \left(1 - \min \left\{ 1, A_i e^{-\frac{d_i \hat{P}_i}{2\hat{\sigma_i}^2}} \right\} \right) \\ = & \sum_{i \neq k} R_i + R_k \left(1 - \min \left\{ 1, A_k e^{-\frac{d_k \hat{P}_k}{2\sigma_{\mathrm{t}}^2}} \right\} \right). \end{aligned} \tag{4.21}$$

The first inequality holds for the reason that the minimization over the effective throughput is always less than or equal to the effective throughput using the noise power $\hat{\sigma_i}^2$ in our case. For the second equality, it is obvious that the minimization over 1 and $A_i e^{-\frac{d_i \hat{P}_i}{2\hat{\sigma_i}^2}}$ is always greater than zero. Moreover, if we have

$$\min \left\{ 1, A_k e^{-\frac{d_k \hat{P}_k}{2\sigma_{\mathrm{t}}^2}} \right\} = 1,$$

(4.21) becomes $\sum_{i \neq k} R_i$ which is less than or equal to $T_{\mathrm{w}}^{\dagger}$.

Else if

$$\min \left\{ 1, A_k e^{-\frac{d_k \hat{P}_k}{2\sigma_{\mathrm{t}}^2}} \right\} = A_k e^{-\frac{d_k \hat{P}_k}{2\sigma_{\mathrm{t}}^2}},$$

(4.21) becomes

$$\sum_{i\neq k} R_i + R_k - e^{-\left(\frac{d_k\hat{P}_k}{2\sigma_{\rm t}^2} - \log\left(R_k A_k\right)\right)}. \tag{4.22}$$

By comparing (4.22) with $T_{\rm w}^{\dagger}$ in (4.18), the only difference is in the exponential term. From the definition of k in (4.20), we know that

$$\frac{d_k\hat{P}_k}{2\hat{\sigma_k^2}} - \log\left(R_k A_k\right) = \min_{1\leq i\leq K}\left(\frac{d_i\hat{P}_i}{2\sigma_{\rm t}^2} - \log A_i R_i\right). \tag{4.23}$$

Note that

$$\begin{aligned} \frac{d_i\hat{P}_i}{2\sigma_{\rm t}^2} - \log R_i A_i &= \frac{d_i\left(P_i^{\dagger} + \triangle P_i\right)}{2\hat{\sigma}_{\rm t}{}^2} - \log R_i A_i \\ &= \nu + \frac{d_i\triangle P_i}{2\sigma_{\rm t}^2} \quad \forall\ 1\leq i\leq K. \end{aligned} \tag{4.24}$$

There always exists at least a channel that has its $\frac{d_i\triangle P_i}{2\sigma_{\rm t}^2}$ being negative since $\hat{P}_i \neq P_i^{\dagger}$ and $\triangle P_i$ should satisfy the constraint in (4.19). Taking minimization over (4.25), we have

$$\min_{1\leq i\leq K}\left(\frac{d_i\hat{P}_i}{2\sigma_{\rm t}^2} - \log A_i R_i\right) < \nu \tag{4.25}$$

Thus

$$\sum_{i\neq k} R_i + R_k - e^{-\left(\frac{d_k\hat{P}_k}{2\sigma_{\rm t}^2} - \log\left(R_k A_k\right)\right)} \leq \sum_{i\neq k} R_i + \left(R_k - e^{-\nu}\right) = T_{\rm w}^{\dagger}\left(\underline{P^{\dagger}}\right)$$

From the discussion above, we have proved that the worst-case effective throughput of any power allocation $\hat{P}_i$ is less than that of $P_i^{\dagger}$. The optimality of $P_i^{\dagger}$ is then justified. Thus, we can claim that when system SNR is greater than $\gamma_{\rm th}^{\dagger}$, the optimal power allocation that maximizes $T_{\rm w}$ is by using $P_i^{\dagger}$ as the power allocation. It is also worth knowing that the corresponding choice of σ_i^2 that achieves $T_{\rm w}$ is to put total noise power to any one of the channel.

The power allocation scheme can also be interpreted as a variation of water filling principle. For each channel, a vessel with base width $\frac{2\sigma_{\rm t}^2}{d_i}$ and base height $\log\frac{1}{A_i d_i}$ will be used for water filling. From our constraints in power in (4.13), each channel should be allocated at least $P^{\dagger}_{\text{th},i}$. The resulting ν must be no less than $\nu_{\min}$. The water filling inside each vessel is then the optimal power to be allotted. An example with three channels($K = 3$) is illustrated in Fig. 4.2.

Example 4.2. *A three channels* ($K = 3$) *system is considered. Each channel has its base height and base width as defined in the paragraph immediately above this example. At least* $P^{\dagger}_{th,1}$, $P^{\dagger}_{th,2}$ *and* $P^{\dagger}_{th,3}$ *should be allocated to three channels, respectively. The lowest water level* $\nu_{\min}$ *should then be chosen as*

$$\nu_{\min} = \max_{1 \le i \le 3} \left\{ \max \{\log A_i, 2\} + \log \frac{1}{A_i R_i} \right\}.$$

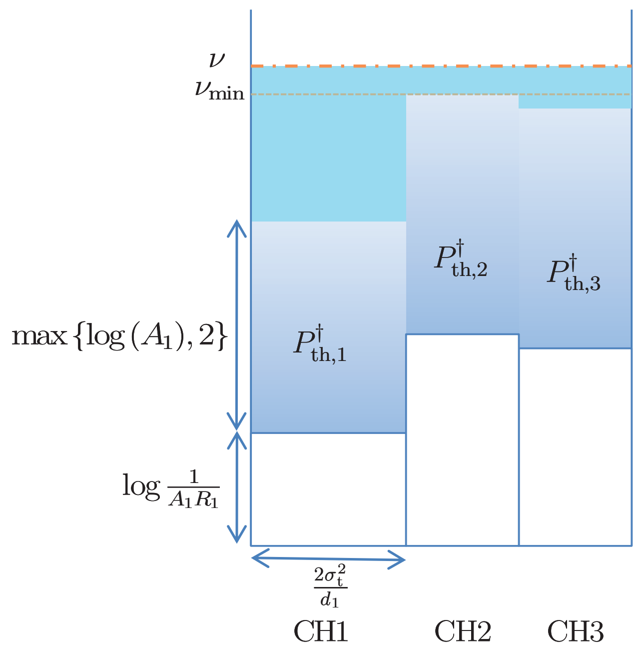

Figure 4.2: An example of the throughput-oriented water-filling with $K = 3$.

From the definition of $P^{\dagger}_i$, we can see that the proposed power allocation depends on σ_t^2. Alternatively, we can say that the proposed power allocation scheme depends on system SNR if we look at the power allocation ratio for each channel, which is derived as the following:

$$p_i^\dagger = \frac{P_i^\dagger}{P_\mathrm{t}} = \frac{2}{d_i \gamma_\mathrm{t}} \left(\nu - \log \frac{1}{A_i R_i} \right), \tag{4.26}$$

where $\gamma_\mathrm{t} = P_\mathrm{t}/\sigma_\mathrm{t}^2$ is the system SNR. In practice, the information of system SNR can be obtained by applying feedback technique to the system. Besides, the effect of system SNR in the power allocation can be eliminated if the code used in each channel can be chosen such that the product of A_i and R_i is the same for different channels. In this way, the ratio of power between different channels becomes a constant, which is as the following:

$$p_i^\dagger : p_j^\dagger = \frac{1}{d_i} : \frac{1}{d_j},$$

for $i \neq j$. Thus, the allocated power ratio for channel i becomes

$$p_i^\dagger = \frac{\frac{1}{d_i}}{\sum_{m=1}^{K} \frac{1}{d_m}},$$

which is only related to d_i. Furthermore, we look at $p_i^\dagger$ when the system SNR goes without bound. From (4.26), we have

$$\lim_{\gamma_\mathrm{t} \to \infty} \frac{p_i^\dagger}{p_j^\dagger} = \frac{\frac{1}{d_i}}{\frac{1}{d_j}}.$$

The optimal allotted power to channel i should be inversely proportional to its d_i, which is closed to its free distance. It coincides with the fact that the free distance dominates frame error rate and thus dominates worst-case effective throughput when SNR tends to be infinity. Finally, our proposed power allocation scheme can be simplified to the traditional equal power allocation scheme when all channels use the same code.

Chapter 5

Numerical and Simulation Results

In this chapter, we compare the effective throughput retrieved by using different power allocation policies. Several convolutional codes [2, 3] are adopted. The parameters in the sense of the approximation given by (3.4) for these codes are listed in Table 5.1.

5.1 Throughput-Oriented Water-Filling: Noise Variance in Each Channel is Known

In this section, three situations of parallel Gaussian channels with $K = 3$ are examined. They are respectively referred to as Cases I, II and III.

In Case I, the noise variances for the three channels are $\sigma_1^2 = 1$, $\sigma_2^2 = 3.5$ and $\sigma_3^2 = 6$, respectively. Here, codes with higher code rates are naturally assigned to less noisy channels; hence we have $R_1 = 1/2$, $R_2 = 1/3$ and $R_3 = 1/4$. The frame sizes for the three channels are $N_1 = 2(1000 + 6)$, $N_2 = 3(1000 + 6)$ and $N_3 = 4(1000 + 6)$, respectively. In Figure. 5.1, we depict the effective throughputs for the seven possible choices of the active channel set $\mathcal{O}$. The figure indicates that all the power should be allocated to channel 1 if $P_\mathrm{t} < 5.14$, and both channels 1 and 2 should be active when $5.14 < P_\mathrm{t} < 10.05$. Beyond the point $P_\mathrm{t} = 10.05$, all three channels should be made active.

Table 5.1: The information of the used codes in the simulation.

code	d_{free}	$A_{d_{\text{free}}}$	adjusted d	adjusted A	codeword length N	generator polynomial (octal)
$(2,1,6)$	10	11	10.63	1478.07	2(500+6)	[133 171]
			11.02	4750.45	2(1000+6)	
$(3,1,6)$	14	1	15.79	593.83	3(500+6)	[133 171 145]
			16.12	1449.97	3(1000+6)	
$(4,1,6)$	20	2	22.42	962.51	4(500+6)	[117 127 155 171]
			22.13	1401.29	4(1000+6)	
$(2,1,2)$	5	1	5.31	111.56	2(500+2)	[5 7]
$(3,1,11)$	24	13	29.04	41373.67	3(500+11)	[5475 6471 7553]
$(4,1,10)$	29	3	35.54	11266.62	4(500+10)	[2565 2747 3311 3723]

In Figure. 5.2, we compare the optimal effective throughput obtained from exhaustive search with that obtained from our throughput-oriented water-filling based on the FER approximation and from the capacity-achieving water-filling policy. We remark that our throughput-oriented water-filling can achieve a near-optimal effective throughput as anticipated. We also observe that the capacity-achieving water-filling policy yields a good throughput only when all the power is allocated to a single channel (which is the optimal choice only for small values of P_{t}).

In Figure. 5.3, we plot the optimal power ratio P_2^*/P_{t} with respect to different power allocation policies. We note that a sudden increase for this ratio occurs in the exhaustive search curve at $P_{\text{t}} = 4.98$ which is exactly the instance the active channel set $\mathcal{O}$ changes from $\{1\}$ to $\{1,2\}$ as shown in Figure. 5.4. This jump occurs when the total power is a little bit larger than the total power corresponding to $\nu = \nu_{\min} = \log \frac{\sigma_2^2}{d_2 R_2}$ in Figure. 4.1, i.e.,

$$P_{\text{t}} \quad > \quad \frac{2\sigma_1^2}{d_1}\left(\log\frac{\sigma_2^2}{d_2 R_2} - \log\frac{\sigma_1^2}{d_1 A_1 R_1}\right) \;+\; \frac{2\sigma_2^2}{d_2}\left(\log\frac{\sigma_2^2}{d_2 R_2} - \log\frac{\sigma_2^2}{d_2 A_2 R_2}\right) \quad = \quad 4.93.$$

This is because Channel 2 can provide a solid contribution to the system effective throughput only when P_2 is adequately larger than $P_{\text{th},2}$. Figure 5.3 also indicates that the predicted jump point from the throughput-oriented water-filling based on the FER approximation, i.e., $P_t = 5.14$, is very close to the true jump point, $P_t = 4.98$, while the capacity-achieving water-filling policy always suggests a continuous increase in the power ratio.

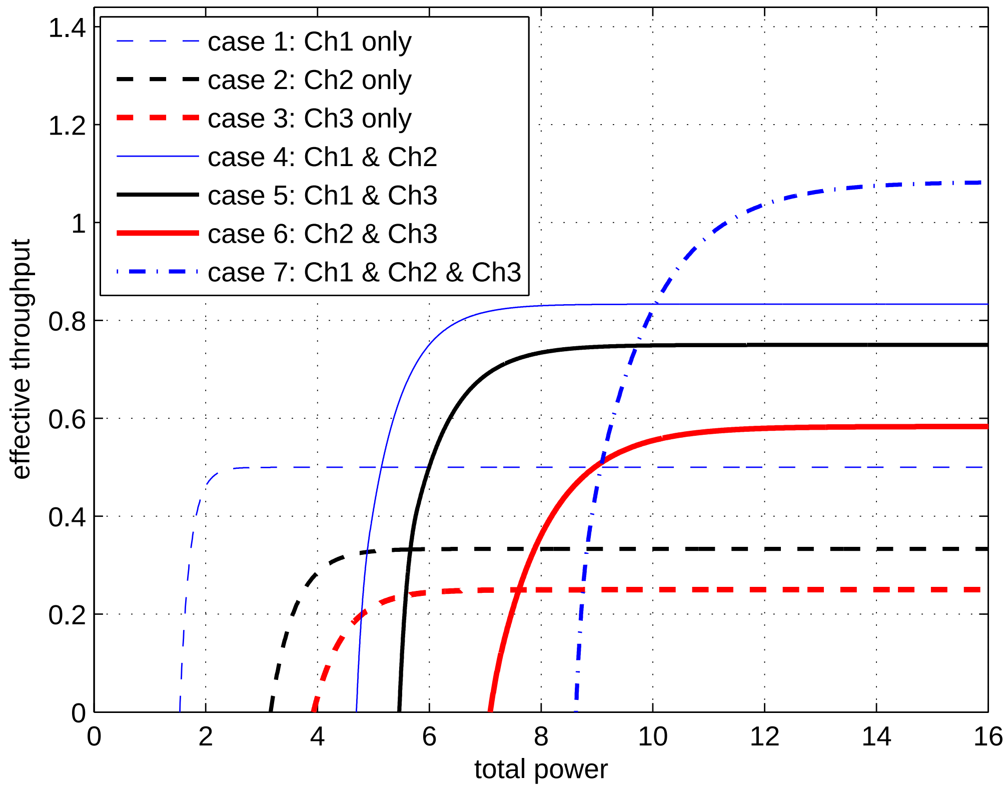

Figure 5.1: Case I: Effective throughputs for the seven choices of the active channel set $\mathcal{O}$.

For Case II, we exchange the codes used in Channels 1 and 3 in Case I. Hence, $R_1 = 1/4$ and $R_3 = 1/2$. The results are summarized in Figures. 5.5, 5.6 and 5.7. These figures point out that using a lower code rate for a less noisy channel will yield a better throughput only when the total power is very small. For moderate to high total power, exchanging the codes between channels 1 and 3 never results in a better effective throughput. This confirms the common intuition that when a channel is less noisy, a code with a higher rate should be used. A side observation is that when assigning a code with lower rate to a less noisy channel, the set of active channels changes more often with respect to P_t. In particular, Channel 2 will

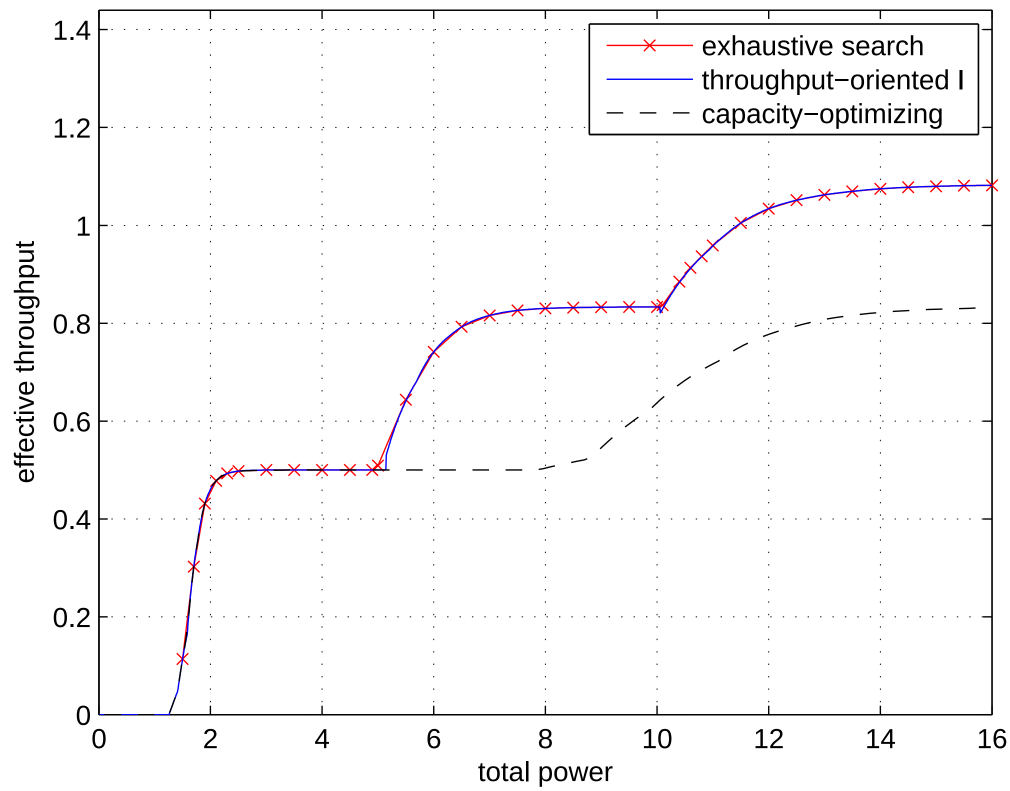

Figure 5.2: Case I: Optimal effective throughputs obtained from exhaustive search, the throughput-oriented water-filling based on the FER approximation, and the capacity-achieving water filling policy.

have two cut-off regions given by $P_t < 3.76$ and $11.42 < P_t < 14.65$ as shown in Figure. 5.5. In addition, Figure. 5.5 shows that adopting a wrong $\mathcal{O}$ will noticeably degrade the effective throughput. Hence, exchanging the codes between Channels 1 and 3 will make complicated the optimization of the throughput.

Finally for Case III, the codes used for three channels are the same as those used in Case I, but the frame sizes are changed to $N_1 = 2(500 + 6)$, $N_2 = 3(500 + 6)$ and $N_3 = 4(500 + 6)$. Thus d_i and A_i are changed simultaneously. Besides, the noise variances are changed to $\sigma_1^2 = 2$, $\sigma_2^2 = 8$, $\sigma_3^2 = 9$. Similar behaviors can be observed from Figure 5.8 except that the capacity-achieving water-filling policy gives an effective throughput closer to the optimal

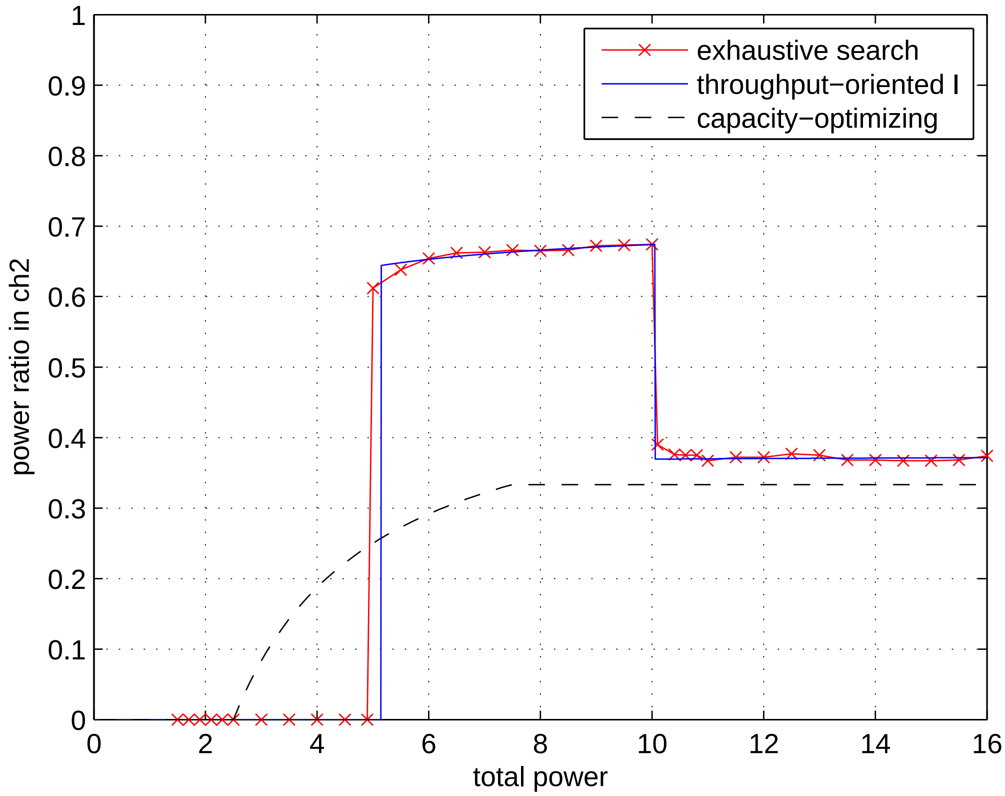

Figure 5.3: Case I: Optimal power ratio for channel 2.

one for high values of the total power. This can be somehow anticipated from the discussion following (4.5) as when the noise variances of the active channels have larger gaps (between Channel 1 and Channels 2 or 3), the capacity-achieving water-filling policy will yield a power allocation closer to the throughput-oriented water-filling.

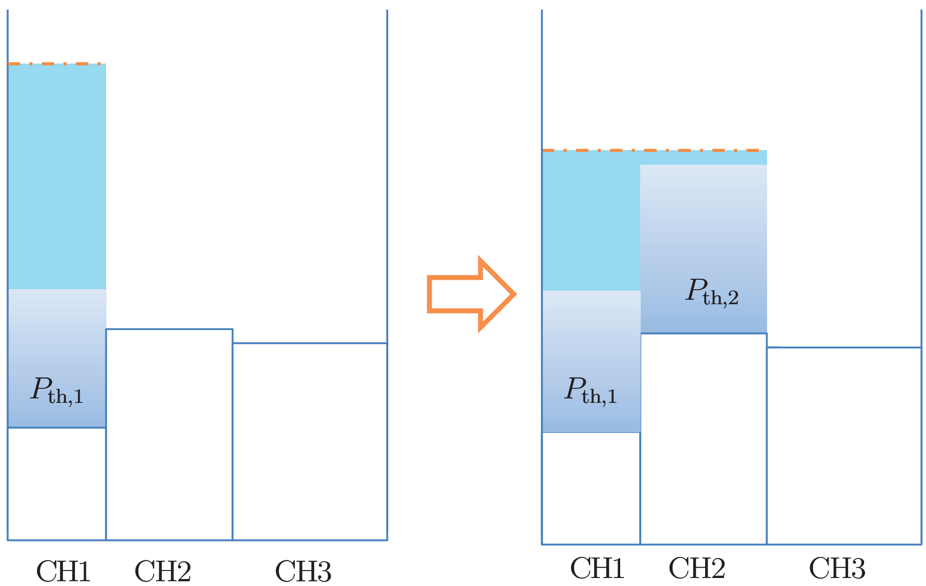

Figure 5.4: Case I: Illustration of the optimal active set $\mathcal{O}$ changing from $\{1\}$ to $\{1, 2\}$.

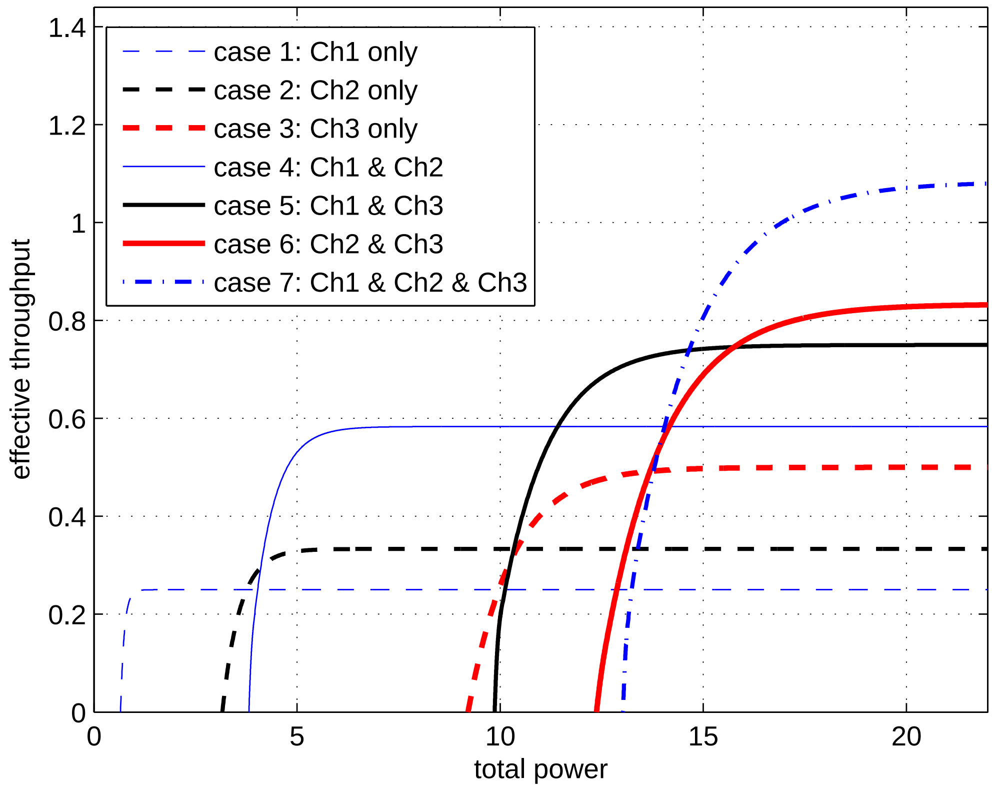

Figure 5.5: Case II: Effective throughputs for the seven choices of active channel set $\mathcal{O}$.

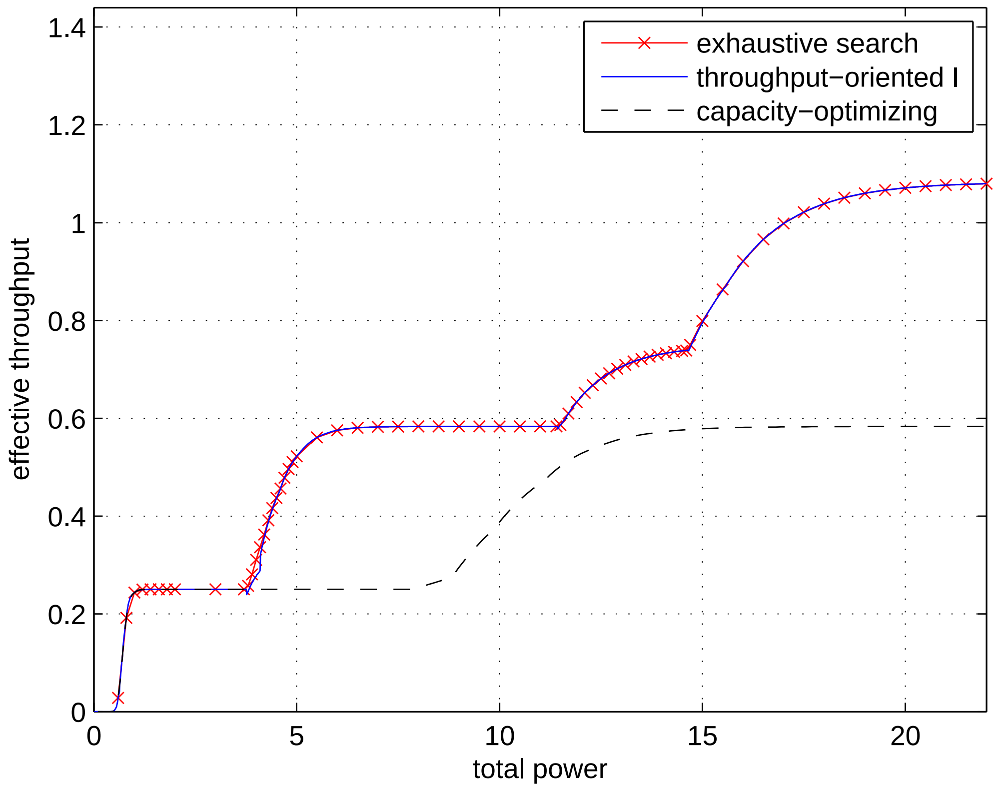

Figure 5.6: Case II: Optimal effective throughputs obtained from exhaustive search, the throughput-oriented water-filling based on the FER approximation, and the capacity-achieving water filling policy.

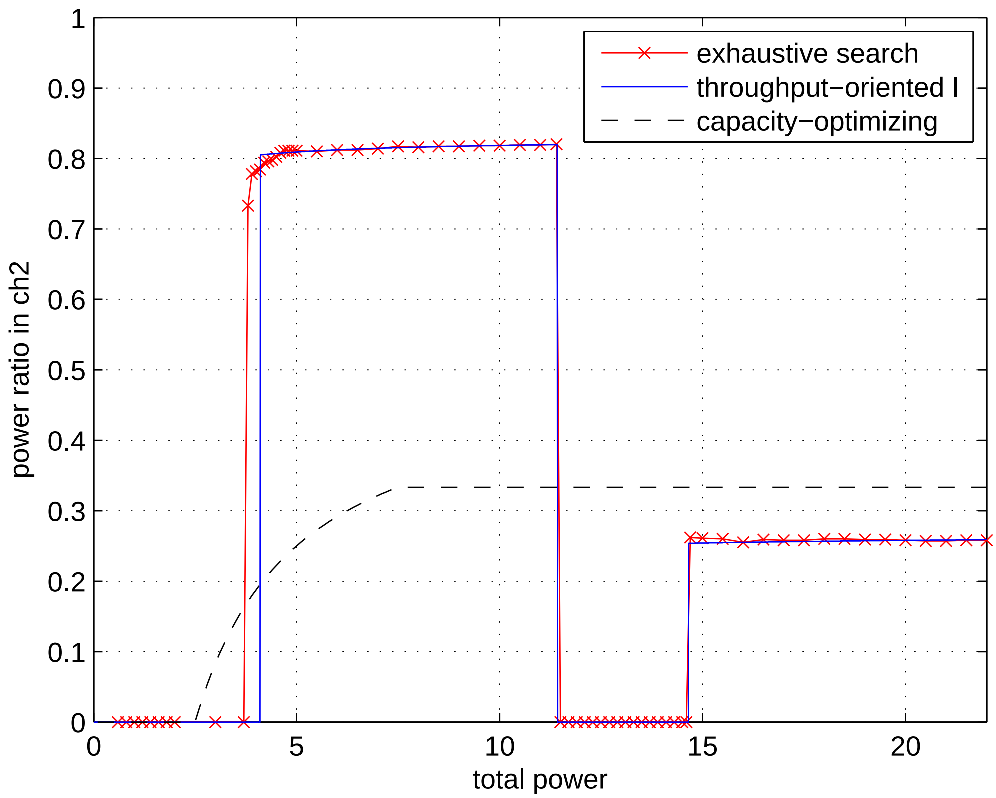

Figure 5.7: Case II: Optimal power ratio for Channel 2.

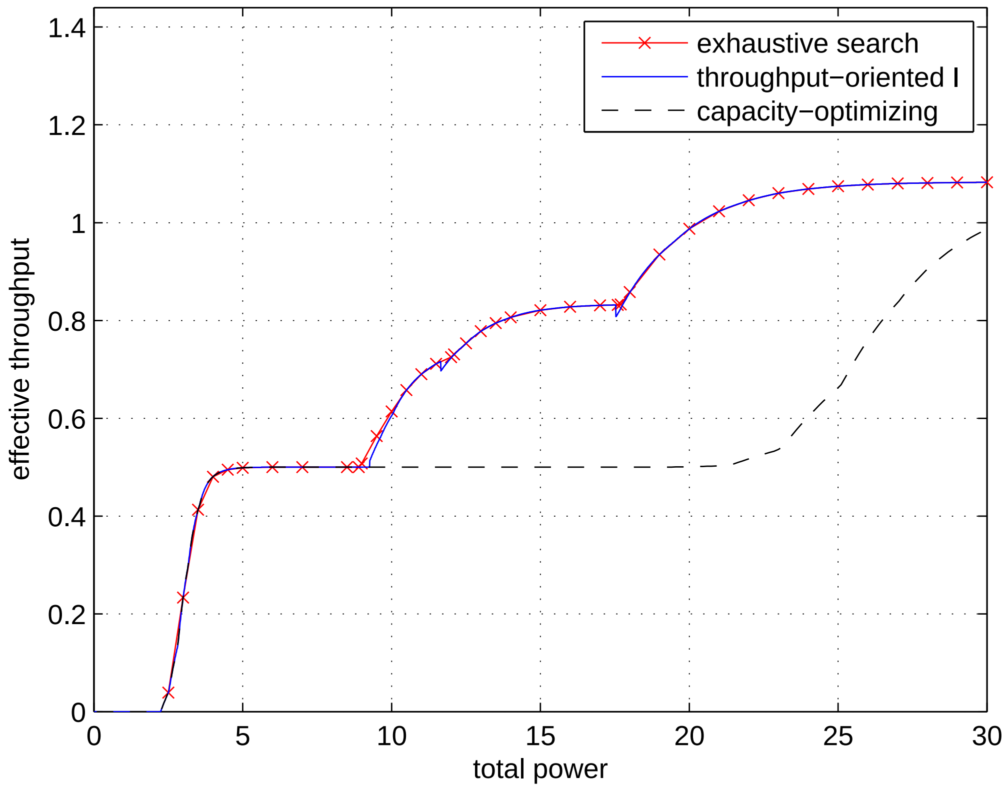

Figure 5.8: Case III: Optimal effective throughputs obtained from exhaustive search, the throughput-oriented water-filling based on the FER approximation, and the capacity-achiving water-filling policy.

5.2 Throughput-Oriented Water-Filling: Only Total Noise Variance is Available

In this section, three situations of parallel Gaussian channels are examined. They are respectively referred to as Cases I, II and III.

In Case I, we consider $K = 3$ and use the $(2, 1, 6)$, $(3, 1, 6)$ and $(4, 1, 6)$ convolutional code in three channels, respectively. The frame size of the codes are $N_1 = 2(500 + 6)$, $N_2 = 3(500 + 6)$ and $N_3 = 4(500 + 6)$. The total noise variance σ_{t}^2 is set to 10. For convenience, we will plot the ratios of the effective throughput against the maximum rate, which is the sum of the rates of the three channels in the following figures.

In Figure. 5.9, we compare the ratios of the effective throughputs against the maximum rate, obtained from the throughput-oriented water-filling in (4.11) and the traditional worst-case capacity-achieving equal power allocation. The $\gamma_{\mathrm{th}}^{\dagger}$, at which value our proposed power allocation becomes optimal, is 4.72 dB. We can see that almost all of the power allocation methods achieve the maximum rate when system SNR is above $\gamma_{\mathrm{th}}^{\dagger}$. Although we cannot guarantee the optimality of using the throughput-oriented water-filling for system SNR smaller than $\gamma_{\mathrm{th}}^{\dagger}$, we can still see that it has around 1.4 dB gain over the equal power allocation when the effective throughput of the system is required to achieve 85% of the maximum rate.

We also observe the distribution of worst-case noise variances for system SNR varying from 2 dB to 6 dB when using throughput-oriented water-filling as the power allocation method. The result shows that we should always give total noise power to Channel 2. This confirms our claim that the worst-case effective throughput is achieved by giving total noise power to only one channel for system SNR greater than $\gamma_{\mathrm{th}}^{\dagger}$.

In Case II, the $(2, 1, 2)$, $(3, 1, 11)$ and $(4, 1, 10)$ convolutional codes are used in three

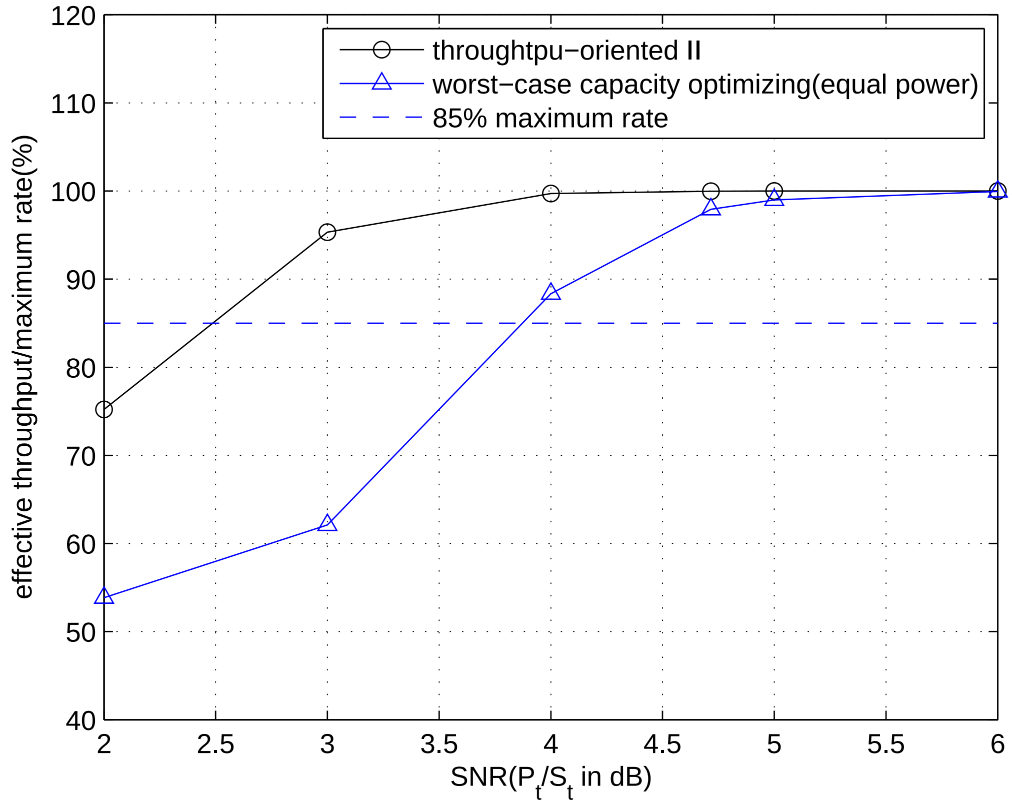

Figure 5.9: Case I: The worst-case effective throughputs obtained from throughput-oriented water-filling based on the FER approximation, and the worst-case-capacity-achieving equal power allocation.

channels, respectively. Compared with the codes used in Case I, the codes used in Case II has larger gaps in d_i, where $d_1 = 5.31$, $d_2 = 29.40$ and $d_3 = 35.54$. It is anticipated that by using this set of codes, throughput-oriented water-filling should yield a greater gain than equal power allocation, when being compared with Case I. A simple way to prove this anticipation is by looking at the situation when system SNR is large. The proposed power allocation policy suggests that P_i should be allocated inversely proportional to d_i. For larger difference in the amount of d_i's, the proposed power allocation deviates greatly from the equal power allocation, and thus yields better gain. Figure 5.10 confirms our deduction. We see that throughput-oriented water-filling yields around 2 dB gain when the effective throughput achieves 85% of the maximum rate. Besides, when we look at the situation when system SNR is equal to $\gamma_{\text{th}}^{\dagger} = 5.19$ dB, throughput-oriented water-filling almost achieves the maximum rate while the equal power allocation achieves only 83% of the maximum rate.

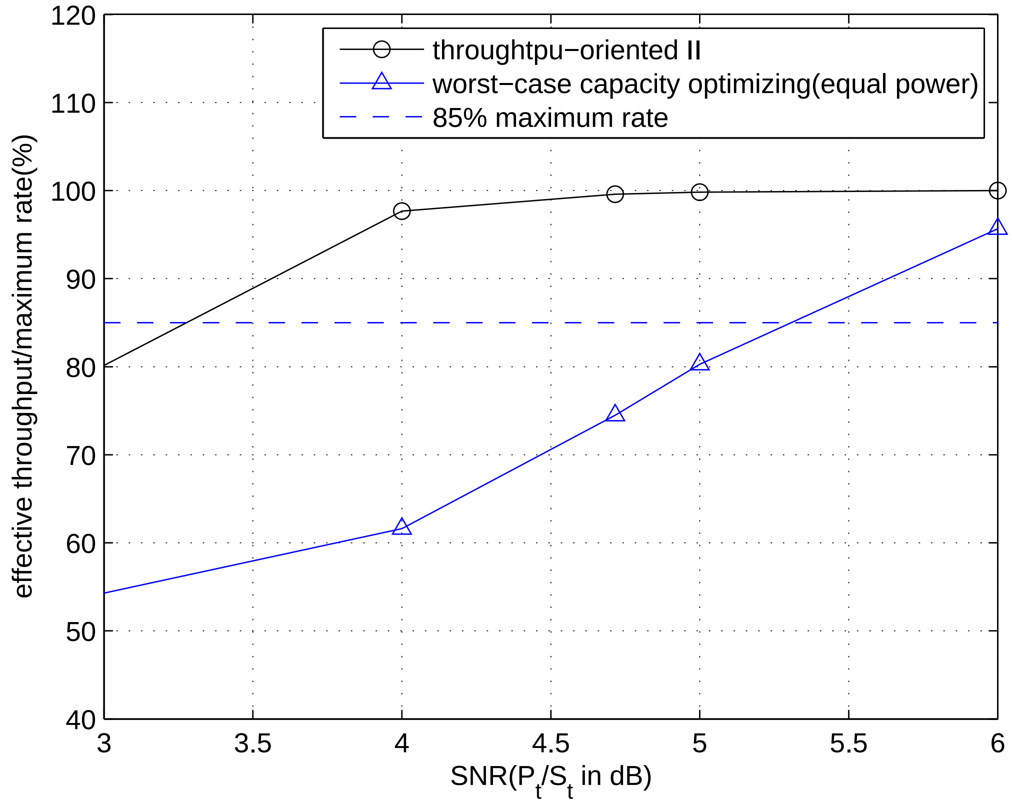

Figure 5.10: Case II: The worst-case effective throughputs obtained from throughput-oriented water-filling based on the FER approximation, and the the worst-case-capacity-achieving equal power allocation.

For Case III, we increase the number of channels to be $K = 4$. We use the $(2, 1, 6)$, $(3, 1, 6)$ and $(4, 1, 6)$ convolutional codes in the first three channels as in Case I. Two different codes are chosen to be used in Channel 4 for comparison.

Firstly, we use the $(2, 1, 6)$ in Channel 4, which is the same code as that used in channel 1. We yield only 0.89 dB gain when the effective throughput is required to achieve 85% of the maximum rate (See Figure 5.11), which is less than the gain in Case I. Secondly, we use the $(3, 2, 6)$ punctured convolutional code in Channel 4. It is punctured from $(2, 1, 6)$ code with puncture pattern

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

The adjusted parameters for the punctured $(3, 2, 6)$ code is $d_4 = 7.89$ and $A_4 = 6469.15$, where d_4 is much less than the d_i of other used codes. From Figure 5.12, we could see that the gain enlarges to 1.89 dB when the effective throughput achieves 85% of the maximum rate, which is greater than the gain obtained in Case I. The result in this case confirms the anticipation that the throughput-oriented water-filling yields a larger gain from traditional equal power allocation when the characteristics of the used codes deviate largely from each other .

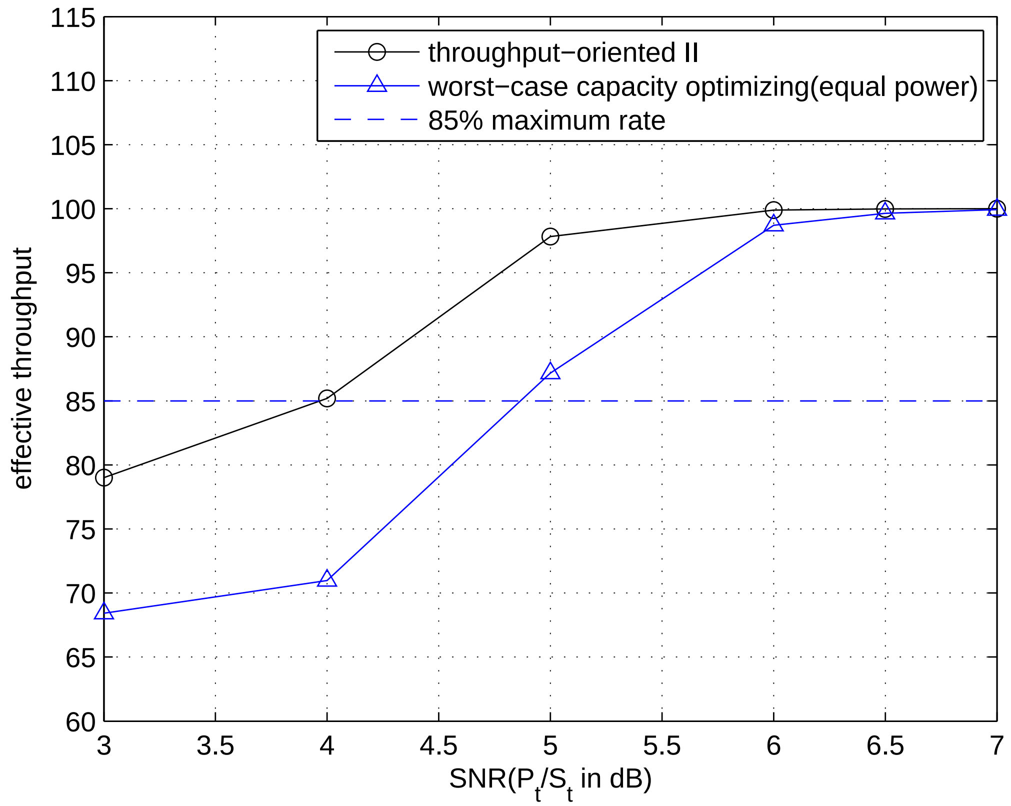

Figure 5.11: Case III: The worst-case effective throughput of using throughput-oriented water-filling and equal power allocation. $K = 4$. $(2, 1, 6)$, $(3, 1, 6)$ and $(4, 1, 6)$ codes are used in the first three channels, and $(2, 1, 6)$ code is used again in the fourth channel.

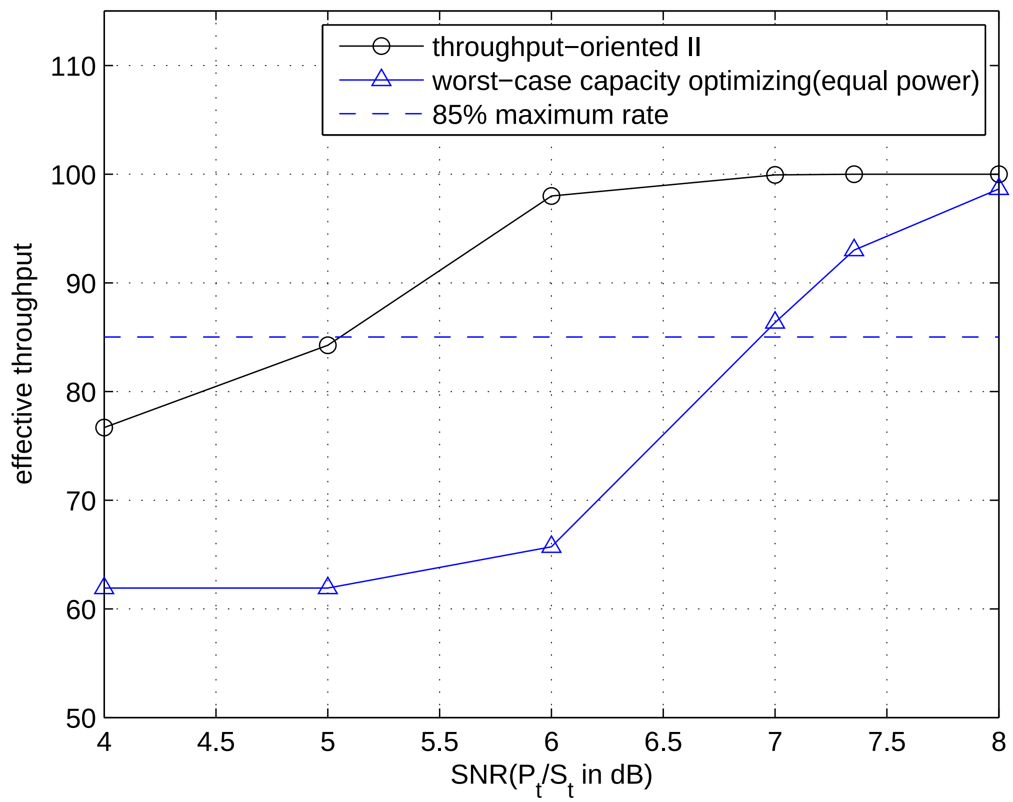

Figure 5.12: Case III: The worst-case effective throughput of using the throughput-oriented water-filling and equal power allocation. $K = 4$. $(2, 1, 6)$, $(3, 1, 6)$ and $(4, 1, 6)$ code sare used in the first three channels, and $(3, 2, 6)$ code is used in the fourth channel.

Chapter 6

Conclusion

In this paper, two power allocation policies are respectively proposed for the two situations: (i) noise variance is known to each channel and (ii) only total noise variance is known. We aim to maximize the effective throughput and the so-defined worst-case effective throughput of the K coded parallel AWGN channels, subject to practical finite-length and fixed-rate coding constraints. These policies preserve the notion of the water-filling principle by additionally taking into consideration the code characteristics. Simulation and numerical results show that the proposed policy for the situation that noise variance is known to each channel can achieve a near-optimal effective throughput for all values of the total power. When only the total noise variance is known, the proposed policy can also achieve a near-optimal effective throughput for system SNR greater than a certain threshold.

In practice, standards usually provide a list of optional codes for each channel. For the case where noise variance in each channel is known, a natural future work is thus to provide a quick determination of the optimal active channel set $\mathcal{O}$ (instead of examining all $(2^K - 1)$ possibilities) such that our policy can readily determine the suitable code to be used in each channel. For the case where only total noise variance is available, the future work is to find the optimal power allocation policy for system SNR below the threshold system SNR.

Bibliography

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

[2] P. Frenger, P. Orten and T. Ottosson, "Convolutional codes with optimum distance spectrum," *IEEE Commun. Lett.*, vol. 3, no. 11, pp. 317-319, Nov. 1999.

[3] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC Codes) and their applications," *IEEE Trans. Commun.*, vol. 36, no. 4, pp. 389-400, Apr. 1988.

[4] R. Johannesson and K. Sh. Zigangirov, *Fundamentals of Convolutional Coding*. IEEE Press, 1999.

[5] S. Lin and D. J. Costello, *Error Control Coding, 2nd ed.*, Prentice Hall, 2004.

[6] A. Lozano, A. M. Tulino and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033-3051, July 2006.

[7] S. Boyd and L.Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[8] S M. Kay, *Fundamentals of Statistical Signal of Processing, vol.2 : Estimation Theory*, Prentice Hall PTR, 1993.

[9] J. Proakis, *Digital Communications*, McGraw-Hill, fifth edition, 2007.