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

Prepared by Wen-Chieh Chang

Advisory by Prof. Po-Ning Chen

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

Institute of Communications Engineering

National Chiao Tung University

Hsinchu, Taiwan 300, R.O.C.

E-mail: wenchiehcm@gmail.com

June, 2011

Abstract

The common criterion used in the power allocation problem for parallel memoryless Gaussian channels is to maximize overall capacity, resulting in the well-known water-filling policy. Such a capacity-optimizing 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 optimal 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 your invaluable instructions in the research. To Dr. Shih-Wei Wang, I am truly thankful for all the solid suggestions from you.

I would also like to thank all the members in the NTL lab for their company for 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). A typical optimization criterion for determining how best to distribute power among the channels is to maximize the system's capacity resulting in the well-known *water-filling* scheme, where optimal 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]. Recently, 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.

Besides, there is another more challenging power allocation problem over parallel Gaussian channels that has usually been discussed. Instead of knowing the noise variance in each channel, only the total noise variance is known. The worst case where for a given power allocation, noise variance for each channel is chosen such that the capacity is minimized (denoted by the worst-case capacity) is always considered. Signal power is then allotted to maximize the worst-case capacity. And the resulting power allocation policy is allotting equal signal power to each channel.

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. Second, 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 maximizes over the effective throughput (instead of capacity) and 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 giving a channel a power allotment that is smaller than a certain threshold 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 certain threshold. And we show that our proposed power allocation policy yields better gain than the traditional equal power allocation if there are more difference in the characteristics of the used codes between channels.

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-optimizing and worst-case-capacity-optimizing problem respectively. In Chapter 3, we introduce the system models and define the throughput-optimizing and worst-case-throughput-optimizing power allocation problem. We 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

2.1 Capacity-Optimizing Water-Filling with Known Noise Variance in Each Channel

Considering a system with K parallel AWGN channels, we assume that the noise sources in different channels are independent of each other. The noise variance for channel i is denoted by σ_i^2. The capacity of K parallel AWGN channels is

$$\sum_{i=1}^{K} \frac{1}{2} \log \left(1 + \frac{P_i}{\sigma_i^2}\right), \tag{2.1}$$

where P_i is the assigned power to channel i and should satisfies the power constraint $\sum_{i=1}^{K} P_i = P_{\text{total}}$. We denote $\underline{P} = \{P_i\}_{i=1}^{K}$ as its assemble format. Since (2.1) is always concave over P_i, the technique of Kuhn-Tucker conditions [7] is applied as the following. Let

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

where the constant λ is Lagrange's multiplier and is chosen such that the power constraint is satisfied. Taking the derivative of (2.2) with respective to P_i, we have

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

Hence,

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

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

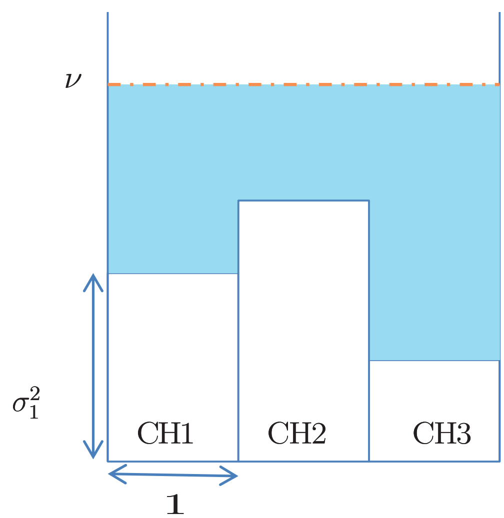

Figure 2.1: Example of water-filling power allocation for $K = 3$.

2.2 Worst-Case-Capacity-Optimizing Water-Filling with Only Total Noise Variance Available

Instead of knowing σ_i^2 of each channel, suppose that we obtain only the information of total noise variance, σ_{total}^2. Lacking of the knowledge of noise variance in each channel, we need to consider the worst case where for any given power allocation, σ_i^2 is always chosen such that the capacity is minimized. It is named the worst-case capacity and is defined as

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

The optimal power allocation is chosen such that (2.4) is maximized and the power constraint $\sum_{i=1}^{K} P_i = P_{\text{total}}$ is satisfied simultaneously. The first step toward this problem is to find the σ_i^2 that achieves the worst-case capacity. Since the capacity is convex (the second order derivative of (2.4) w.r.t. σ_i^2 is always positive) over σ_i^2, the technique of Kuhn-Tucker conditions can be applied to find the optimal σ_i^2. We let

$$C_1(\underline{P}) \quad = \quad \sum_{i=1}^{n} \frac{1}{2} \log \left(1 + \frac{P_i}{\sigma_i^2}\right) + \lambda_1 \left(\sum_{i=1}^{n} \sigma_i^2 - \sigma_{\text{total}}^2\right), \tag{2.5}$$

where λ_1 is chosen such that $\sum_{i=1}^{K} \sigma_i^2 = \sigma_{\text{total}}^2$. Taking derivative of (2.5) w.r.t. 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) could 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,

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

We then replace the σ_i^2 in (2.4) by (2.8) and thus take away the minimization. (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 over P_i, the technique of Kuhn-Tucker conditions can be applied again as the following. Let

$$C_2(\underline{P}) = \sum_{i=1}^{n} \frac{1}{2} \log \left(1 + \frac{P_i}{\frac{-P_i + \sqrt{(P_i)^2 - \frac{2P_i}{\lambda_1}}}{2}}\right) + \lambda_2 \left(\sum_{i=1}^{n} P_i - 1\right),$$

where λ_2 is chosen such that $\sum_{i=1}^{K} P_i = P_{\text{total}}$. The derivative of (2.10) w.r.t. P_i becomes

$$\frac{\partial C_2(\underline{P})}{\partial P_i} \left(\frac{1}{\left(-1+\sqrt{1-\frac{2}{\lambda_1 P_i}}\right)^2 + 2(-1+\sqrt{1-2\frac{2}{\lambda_1 P_i}})} \right) \left(\frac{1}{\sqrt{(1-\frac{2}{\lambda_1 P_i}}\lambda_1 P_i^2} \right) + \lambda_2. \quad (2.10)$$

And (2.10) should also satisfy

$$\begin{cases} \frac{\partial C_2(\underline{P})}{\partial P_i} = 0, & \text{if } P_i > 0; \\ \frac{\partial C_2(\underline{P})}{\partial P_i} < 0, & \text{if } P_i = 0. \end{cases}$$

For any $j \neq i$, we have

$$\frac{\partial C_2(\underline{P})}{\partial P_i} = \frac{\partial C_2(\underline{P})}{\partial P_j} \qquad (2.11)$$

if $P_i > 0$ and $P_j > 0$. One possible choice of P_j for (2.11) to hold and the power constraint to be satisfied is choosing $P_i = P_j$. Thus, the resulting power allocation policy is the equal power allocation.

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 AWGN. Let R_i be the rate of the code adopted by channel i, and denote by FER_i its corresponding frame error rate for frame size N_i. The system effective throughput is defined as

$$\mathrm{eff}\left(\underline{P}\right) \triangleq \sum_{i=1}^{K} R_i \left(1 - \mathrm{FER}_i\right), \tag{3.1}$$

which corresponds to the successfully transmitted information bits per channel use. Note that in the above formula, FER_i is a function of N_i, σ_i^2 and P_i. To simplify notation, we do not explicitly write FER_i as a function of N_i, P_i and σ_i^2.

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

$$\mathrm{eff}_{\mathrm{worst}}\left(\underline{P}\right) \triangleq \min_{\sum_{i=1}^{K} \sigma_i^2 = \sigma_{\mathrm{total}}^2} \sum_{i=1}^{K} R_i \left(1 - \mathrm{FER}_i\right) \tag{3.2}$$

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, FER_i does not exhibit a closed-form formula. Hence, the power allocation that maximizes $\text{eff}(\underline{P})$ and $\text{eff}_{\text{worst}}(\underline{P})$ can be obtained only via case-by-case simulation studies. It is thus hard to establish a general power allocation principle from such simulation-based power allocation results. One possible solution is to derive a good approximation for FER with a structure that can be facilitated in the analysis.

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

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

where d_{free} is the free distance of the convolutional code, $A_{d_{\text{free}}}$ is the number of codewords with Hamming weight equal to d_{free}, and P is the transmission power. However, the approximated FER in (3.3) is far from accurate for moderate SNRs and finite frame sizes(cf. the *approx. FER* curve in Figure. 3.1 for a convolutional code with rate $R = 1/4$ and memory order 6 where $RE_b = P$ 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 $A_{d_{\text{free}}}$ 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 (BER) to avoid considering the frame size, e.g.,

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

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,

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

is close to the true FER in the least squares sense over the range of operating SNRs(cf. the *adjusted approx. FER* 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 FER 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 FER 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 FER values in* log *scale. From* (3.4)*, the* $s[j]$ *can be modelled by*

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

or in matrix form

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

where

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

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 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_{total}.

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

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

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

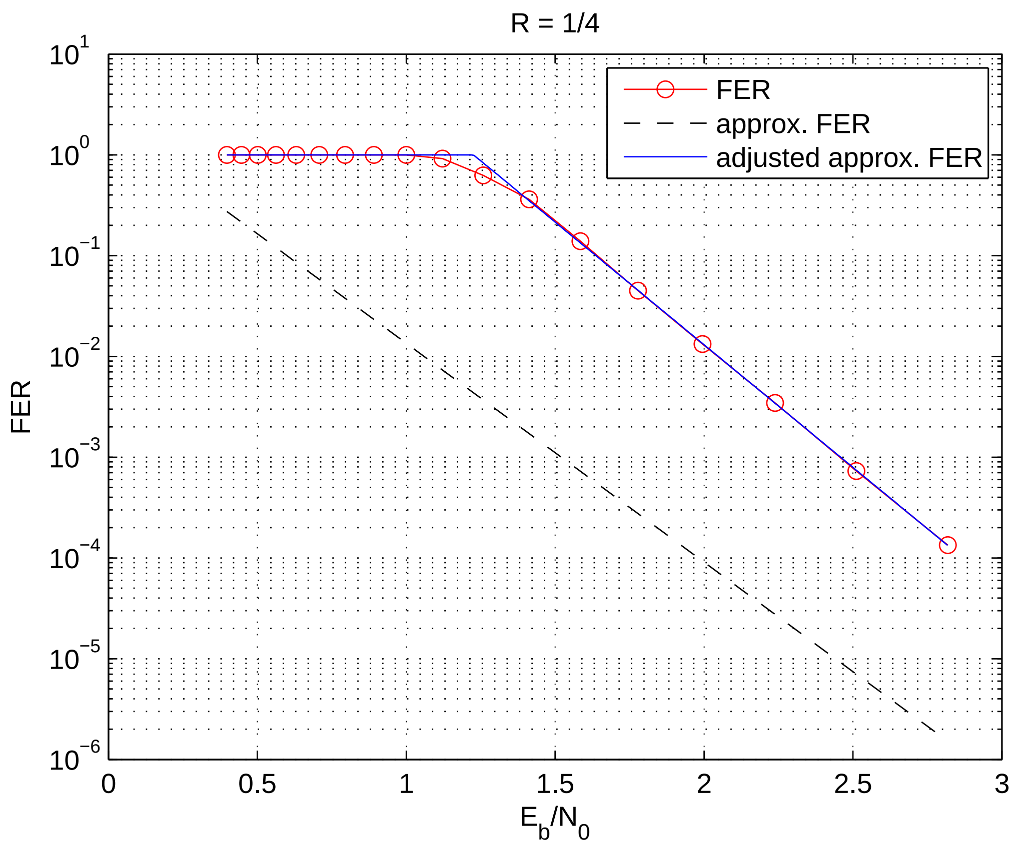

Figure 3.1: FER and its approximations for a $(4, 1, 6)$ convolutional code with generator polynomial(octal form) being [177 127 155 171], $d_{\text{free}} = 20$ and $A_{\text{dfree}} = 2$. The adjusted parameters are $d = 22.42$ and $A = 962.51$. The frame size is $N = 4(500 + 6)$. E_b/N_0 is plotted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

Chapter 4

Throughput-Oriented Water-Filling

In this chapter, the power allocation that maximizing (3.1) and (3.2) are discussed in two sub-sections respectively. The analyses are mainly based on the approximated FER defined in (3.4). Interestingly, both the proposed power allocation policies in two sections can be interpreted by the variation of water-filling. At the end of each section, we observe the power allocation when total power goes without bound. Several conclusions will be given.

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

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

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

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 technique of Kuhn-Tucker conditions as follows. Let

$$\mathrm{EFF}\left(\underline{P}\right) = \mathrm{eff}\left(\underline{P}\right) - \lambda \left(\sum_{i\in\mathcal{O}} P_i - P_{\mathrm{total}}\right),$$

where the constant λ is Lagrange's multiplier and is chosen such that $\sum_{i\in\mathcal{O}} P_i = P_{\text{total}}$. Taking differentiation with respect to P_i, we have

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

(4.2) should satisfy

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

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 $\sum_{i\in\mathcal{O}} P_i^* = P_{\text{total}}$. 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. $P_i^* \geq P_{\text{thres},i} \quad \forall\, i \in \mathcal{O}$).

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 defined above. We assume that* $\mathcal{O} = \{1, 2\}$ *has already been given. Thus, zero power will be allocated to channel* 3*. And at least* $P_{\textit{thres},1}$ *and* $P_{\textit{thres},2}$ *should be allocated to channel* 1 *and* 2 *respectively. The lowest water level* $\nu_{\min}$ *should then be chosen to be*

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

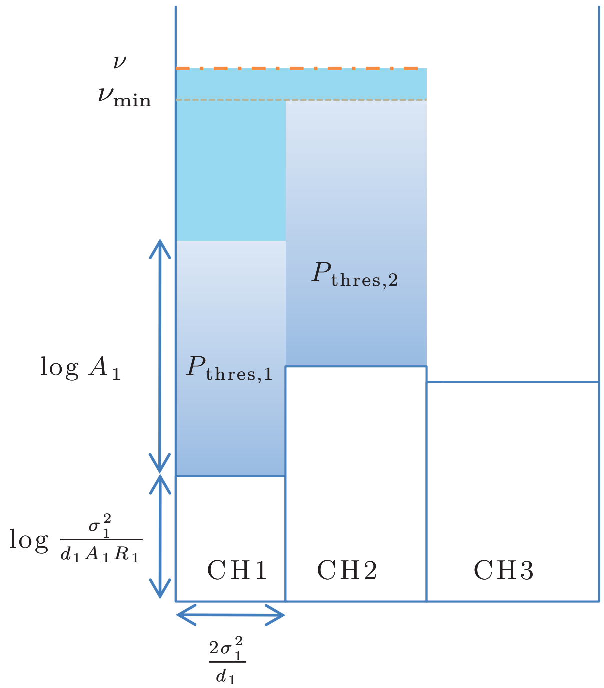

Figure 4.1: Example of the throughput-oriented water-filling I with $K = 3$ and $\mathcal{O} = \{1, 2\}$.

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-optimizing 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-optimizing 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_{total} 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

$$d_i\gamma_i^* = d_j\gamma_j^* + 2\log\frac{\sigma_j^2}{\sigma_i^2} + \log\frac{d_iA_iR_i}{d_jA_jR_j}$$

Thus

$$\lim_{P_{\text{total}}\to\infty}\frac{d_i\gamma_i^*}{d_j\gamma_j^*} = 1 + \lim_{P_{\text{total}}\to\infty}\frac{2\log\frac{\sigma_j^2}{\sigma_i^2} + \log\frac{d_iA_iR_i}{d_jA_jR_j}}{d_j\gamma_j^*} = 1.$$

Hence, when P_{total} 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_{total} 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 $P_{\text{thres},i}$ 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 II: Only Total Noise Variance is Available

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

$$\text{eff}_{\text{worst}}\left(\underline{P}\right) = \min_{\sum_{i=1}^{K}\sigma_i^2=\sigma_{\text{total}}^2} \sum_{i=1}^{K} R_i \left(1 - \min\left\{1, A_i e^{-\frac{d_i P_i}{2\sigma_i^2}}\right\}\right). \quad (4.6)$$

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

$$P_i \geq P_{\text{thres},i} \quad \forall\ 1 \leq i \leq K. \quad (4.7)$$

(4.6) becomes

$$\text{eff}_{\text{worst}}\left(\underline{P}\right) = \min_{\sum_{i=1}^{K}\sigma_i^2=\sigma_{\text{total}}^2} \sum_{i=1}^{K} R_i \left(1 - A_i e^{-\frac{d_i P_i}{2\sigma_i^2}}\right). \quad (4.8)$$

A straightforward approach to eliminate the minimization in (4.8) is by using the technique of Kuhn-Tucker conditions 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 over σ_i^2 from (4.9).

$$\frac{\partial^2 \text{eff}_{\text{worst}}\left(\underline{P}\right)}{\partial\left(\sigma_i^2\right)^2} = \frac{R_i A_i d_i P_i}{\left(\sigma_i^2\right)^3} e^{-\frac{d_i P_i}{2\sigma_i^2}} \left(1 - \frac{d_i P_i}{4\sigma_i^2}\right) \quad (4.9)$$

Thus, by letting

$$P_i \geq \frac{4\sigma_i^2}{d_i} \quad (4.10)$$

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 $\text{eff}_{\text{worst}}$ should be chosen as either 0 or σ_{total}^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 power allocation $P_i^\dagger$ which is defined as

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

where ν is chosen such that

$$\sum_{i=1}^{K} P_i^{\dagger} = P_{\text{total}}.$$

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 $\sigma_i^2 \leq \sigma_{\text{total}}^2$, we sufficiently choose $P_i^{\dagger}$ such that (4.12) is satisfied, which is

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

Note that what we did in (4.13) is reasonable since σ_i^2 is either 0 or σ_{total}^2 under our constraints. And we define the minimum power required in channel i as

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

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

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

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 choice of ν which is $\max\left\{2 - \log A_i R_i, -\log R_i\right\}\ \forall i$. From the definition of $P_i^{\dagger}$ and (4.14), we have

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

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

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

where $\gamma_{\text{thres}}^{\dagger}$ 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_{total}^2 when system SNR is greater than $\gamma_{\text{thres}}^{\dagger}$. Using this result, the worst-case effective throughput of $P_i^{\dagger}$ which is denoted by $\text{eff}_{\text{worst}}^{\dagger}$ can be computed as follows.

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

We define the noise variance that achieves $\text{eff}_{\text{worst}}^{\dagger}$ as

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

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

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

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 $\text{eff}_{\text{worst}}$ yields by any other power allocation is less than or equal to $\text{eff}^{\dagger}_{\text{worst}}$. 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

$$\hat{\sigma_i}^2 = \begin{cases} \sigma^2_{\text{total}} & \text{if} \quad i = k \\ 0 & \text{if} \quad i \neq k \end{cases},$$

where

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

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

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

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

$$\min \left\{ 1, A_k e^{-\frac{d_k \hat{P}_k}{2\sigma^2_{\text{total}}}} \right\} = 1,$$

(4.21) becomes $\sum_{i \neq k} R_i$ which is less than or equal to $\text{eff}^{\dagger}_{\text{worst}}$.

Else if

$$\min \left\{ 1, A_k e^{-\frac{d_k \hat{P}_k}{2\sigma^2_{\text{total}}}} \right\} = A_k e^{-\frac{d_k \hat{P}_k}{2\sigma^2_{\text{total}}}},$$

(4.21) becomes

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

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

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

Note that

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

There always exists at least a channel has its $\frac{d_i \triangle P_i}{2\sigma_{\text{total}}^2}$ 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

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

Thus

$$\begin{aligned} & \sum_{i \neq k} R_i + R_k - e^{-\left(\frac{d_k \hat{P}_k}{2\sigma_{\text{total}}^2} - \log(R_k A_k)\right)} \\ \leq\ & \sum_{i \neq k} R_i + \left(R_k - e^{-\nu}\right) \\ =\ & \text{eff}_{\text{worst}}^{\dagger}\left(\underline{P^{\dagger}}\right) \end{aligned}$$

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 justified. Thus, we could claim that when system SNR is greater than $\gamma_{\text{thres}}^{\dagger}$, the optimal power allocation that maximizes $\text{eff}_{\text{worst}}$ is by using $P_i^{\dagger}$ as power allocation. It is also worth knowing that the corresponding choice of σ_i^2 that achieves $\text{eff}_{\text{worst}}$ 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 $\frac{2\sigma_{\text{total}}^2}{d_i}$ 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 $P_{\text{thres},i}^{\dagger}$. 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 defined above. At least* $P_{thres,1}^{\dagger}$, $P_{thres,2}^{\dagger}$ *and* $P_{thres,3}^{\dagger}$ *should be allocated to three channels respectively. The lowest water level* $\nu_{\min}$ *should then be chosen to be*

$$\nu_{\min} = \max\left\{\max\{\log A_1, 2\} + \log\frac{1}{A_1 R_1}, \max\{\log A_2, 2\} + \log\frac{1}{A_2 R_2}, \max\{\log A_3, 2\} + \log\frac{1}{A_3 R_3}\right\}.$$

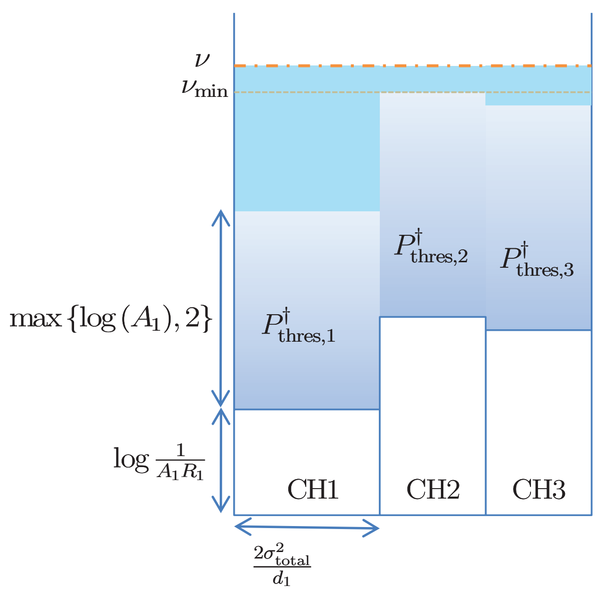

Figure 4.2: Example of the throughput-oriented water-filling II with $K = 3$.

We further look at $P_i^\dagger$ when it goes without bound. From (4.11), we have

$$\lim_{P_{\text{total}}\to\infty} \frac{P_i^\dagger}{P_j^\dagger} = \frac{\frac{1}{d_i}}{\frac{1}{d_j}}$$

The optimal allotted power in 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. Besides, our proposed power allocation scheme can be simplified to traditional equal power allocation scheme when all channels use the same code.

Chapter 5

Numerical and Simulation Results

In this chapter, we examine the proposed power allocation policies in two sections respectively. Several convolutional codes[2, 3] are adapted. The adjusted 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 I: Noise Variance in Each Channel is Known

In this section, three situations of parallel Gaussian channels with $K = 3$ are examined: 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, $R_1 = 1/2$, $R_2 = 1/3$ 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 $P_{\text{total}} < 5.14$, and both channels 1 and 2 should be active when $5.14 < P_{\text{total}} < 10.05$. Beyond the point $P_{\text{total}} = 10.05$, all three channels should be made active.

In Figure. 5.2, we compare the optimal effective throughput obtained from exhaustive

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 11.02	1478.07 4750.45	2(500+6) 2(1000+6)	[133 171]
$(3,1,6)$	14	1	15.79 16.12	593.83 1449.97	3(500+6) 3(1000+6)	[133 171 145]
$(4,1,6)$	20	2	22.42 22.13	962.51 1401.29	4(500+6) 4(1000+6)	[117 127 155 171]
$(2,1,2)$	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]

search with the effective throughputs obtained from our throughput-oriented water-filling I based on the FER approximation and from the capacity-optimizing water-filling policy. We clearly remark that our throughput-oriented water-filling I based on approximating the FER can achieve a near-optimal effective throughput, as anticipated. We also observe that the capacity-optimizing 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_{total}).

In Figure. 5.3, we plot the optimal power ratio P_2^*/P_{total} with respect to different power allocation policies. We note that a sudden increase for this ratio occurs in the exhaustive search curve at $P_{\text{total}} = 4.98$ 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 a little later than the jump determined by setting ν to equal $\nu_{\min} = \log \frac{\sigma_2^2}{d_2 R_2}$ in Figure. 4.1, i.e.,

$$P_{\text{total}} \quad > \quad \frac{2\sigma_1^2}{d_1}\left(\log\frac{\sigma_2^2}{d_2R_2} - \log\frac{\sigma_1^2}{d_1A_1R_1}\right) + \frac{2\sigma_2^2}{d_2}\left(\log\frac{\sigma_2^2}{d_2R_2} - \log\frac{\sigma_2^2}{d_2A_2R_2}\right) \quad = \quad 4.93.$$

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

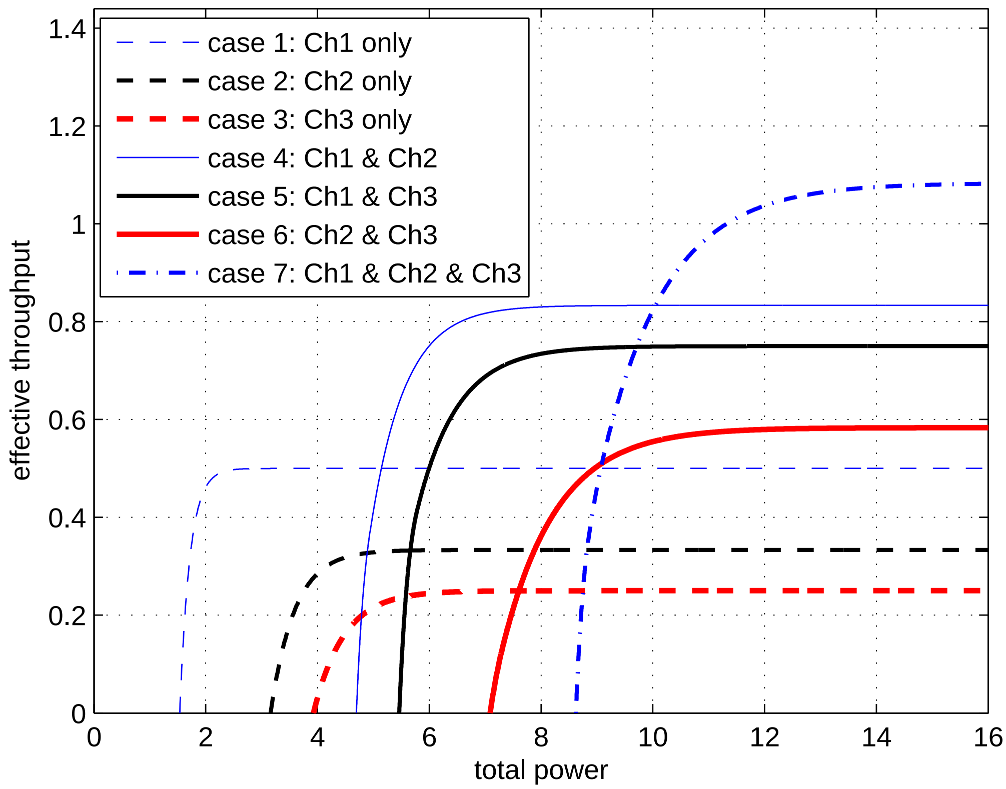

Figure 5.1: Case I: Effective throughputs for the seven choices of the active channel set $\mathcal{O}$.

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_{total}. In particular, channel 2 will have two cut-off regions given by $P_{\text{total}} < 3.76$ and $11.42 < P_{\text{total}} < 14.65$ as shown in

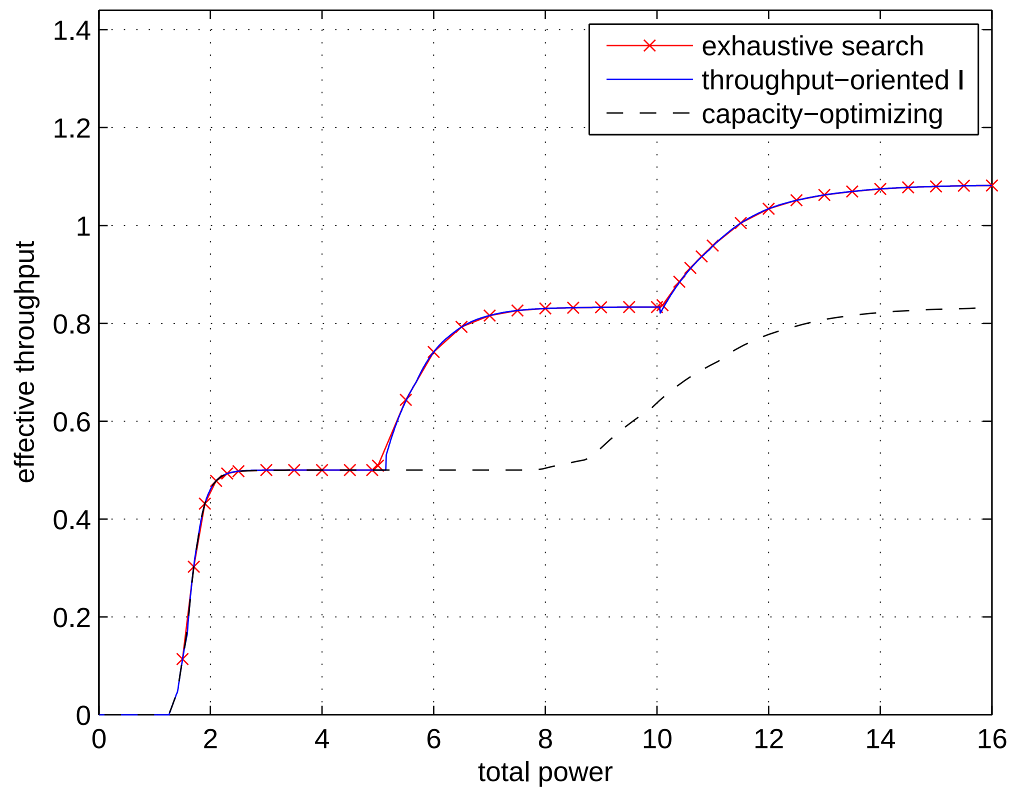

Figure 5.2: Case I: Optimal effective throughputs obtained from exhaustive search, the throughput-oriented water-filling I based on the FER approximation, and the capacity-optimizing water filling policy.

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

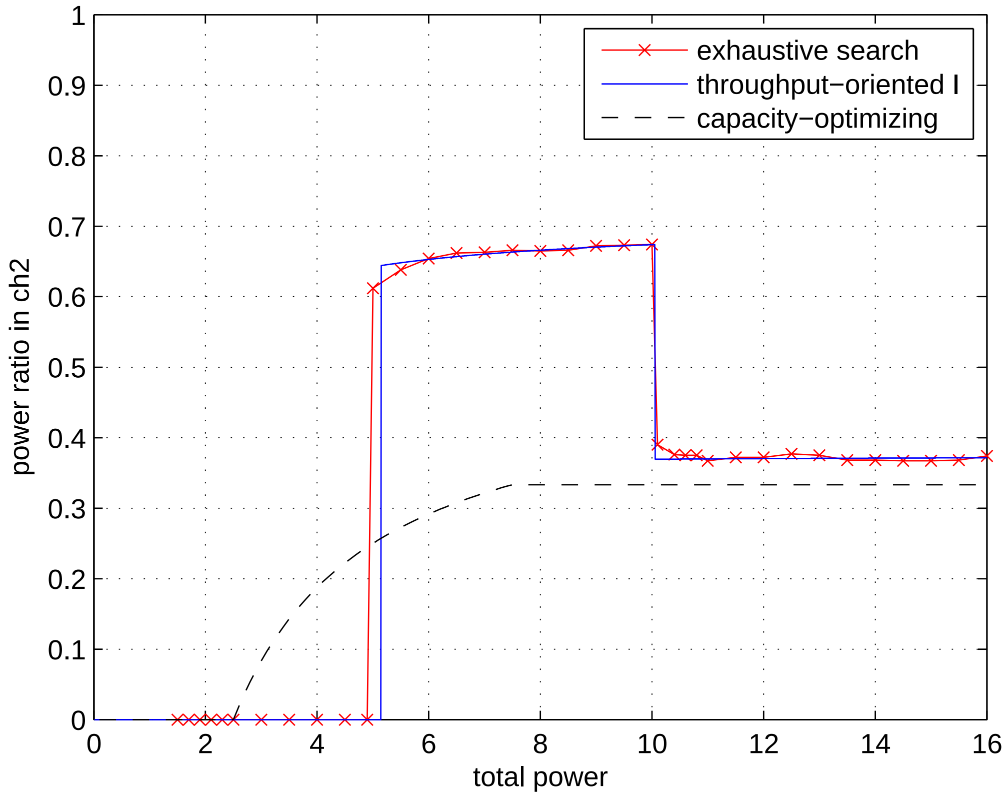

Figure 5.3: Case I: Optimal power ratio for channel 2.

gaps (between channel 1 and channels 2 or 3), the capacity-optimizing water-filling policy will yield a power allocation closer to the throughput-oriented water-filling I.

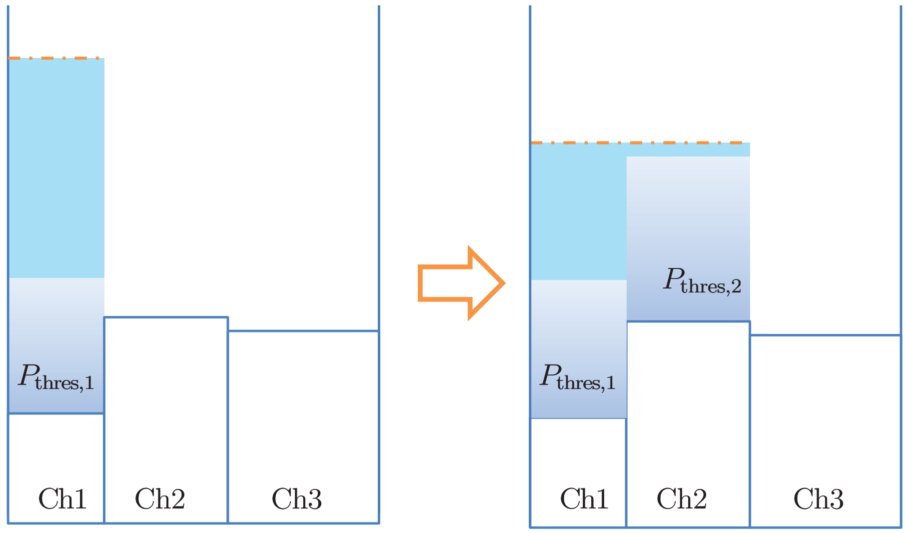

Figure 5.4: Case I: Illustration of the optimal active set $\mathcal{O}$ changing from $\{1\}$ to $\{1, 2\}$.

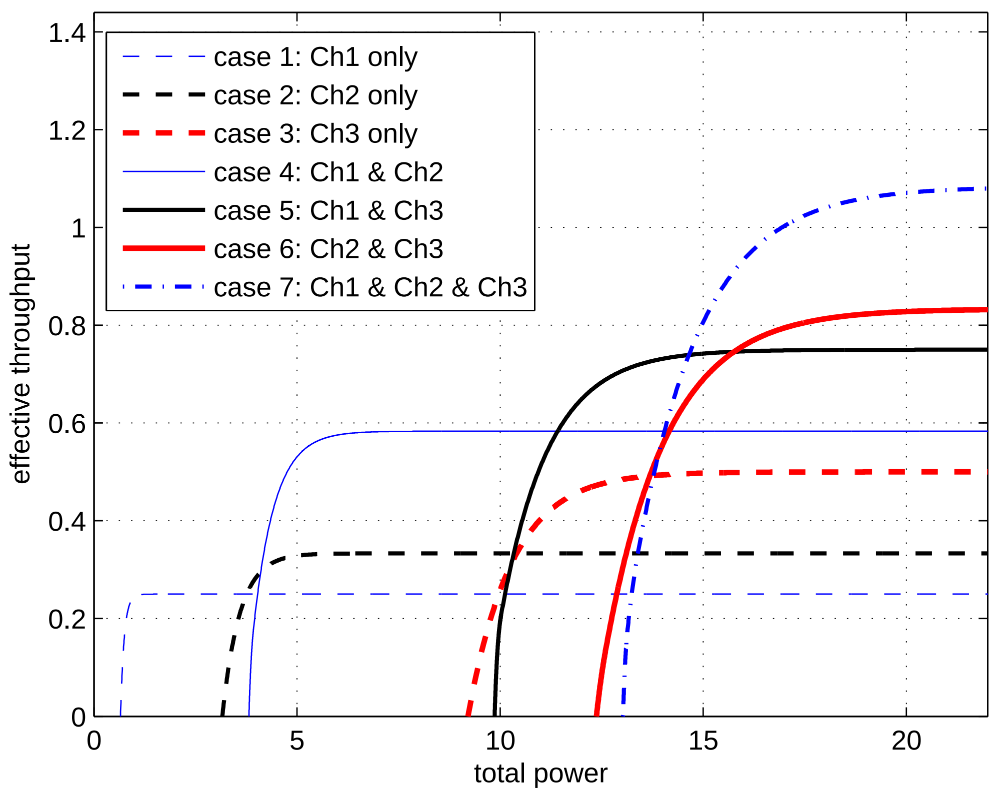

Figure 5.5: Case II: Effective throughputs for the seven choices of active channel set $\mathcal{O}$.

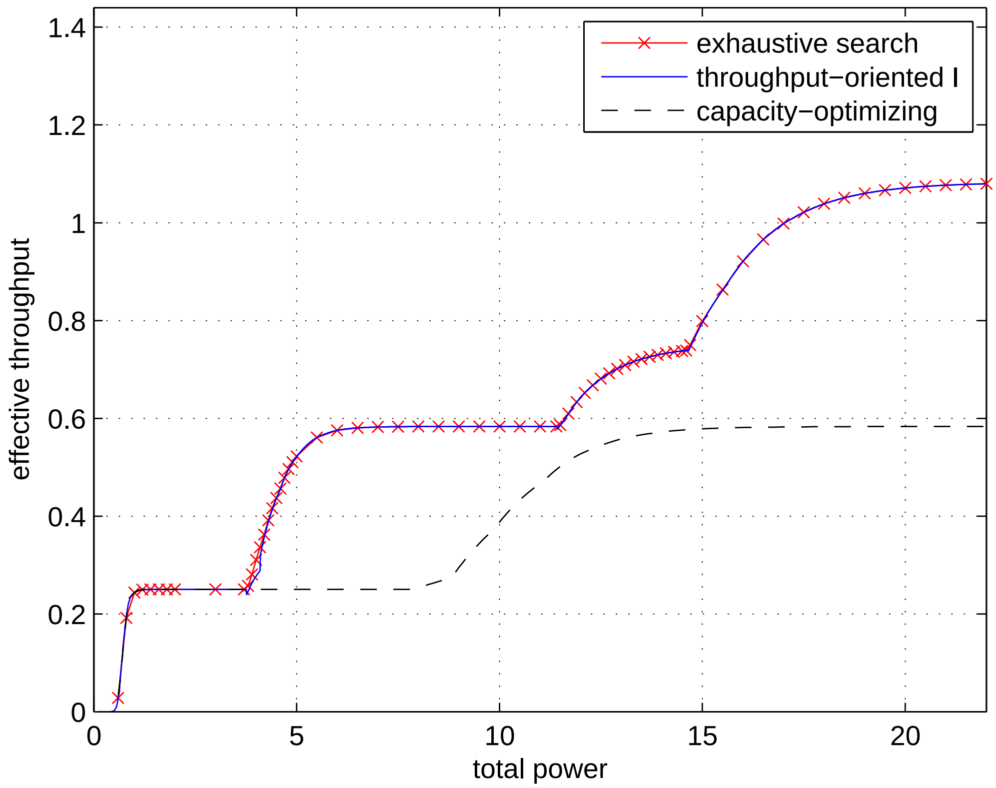

Figure 5.6: Case II: Optimal effective throughputs obtained from exhaustive search, the throughput-oriented water-filling I based on the FER approximation, and the capacity-optimizing water filling policy.

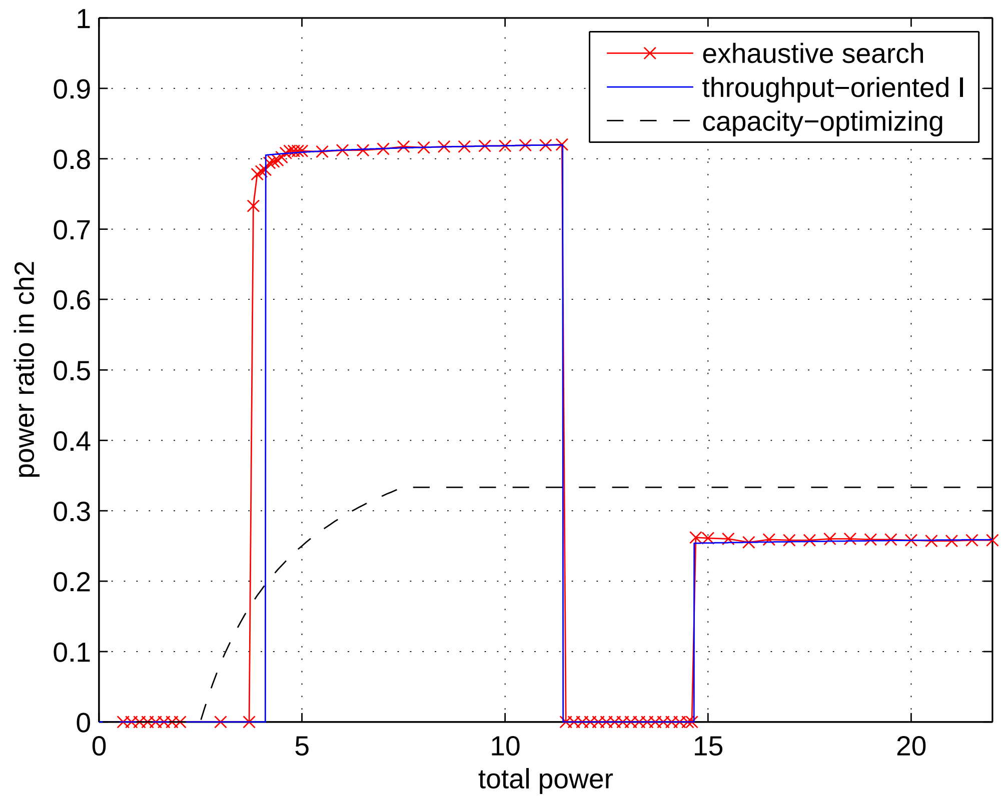

Figure 5.7: Case II: Optimal power ratio for channel 2.

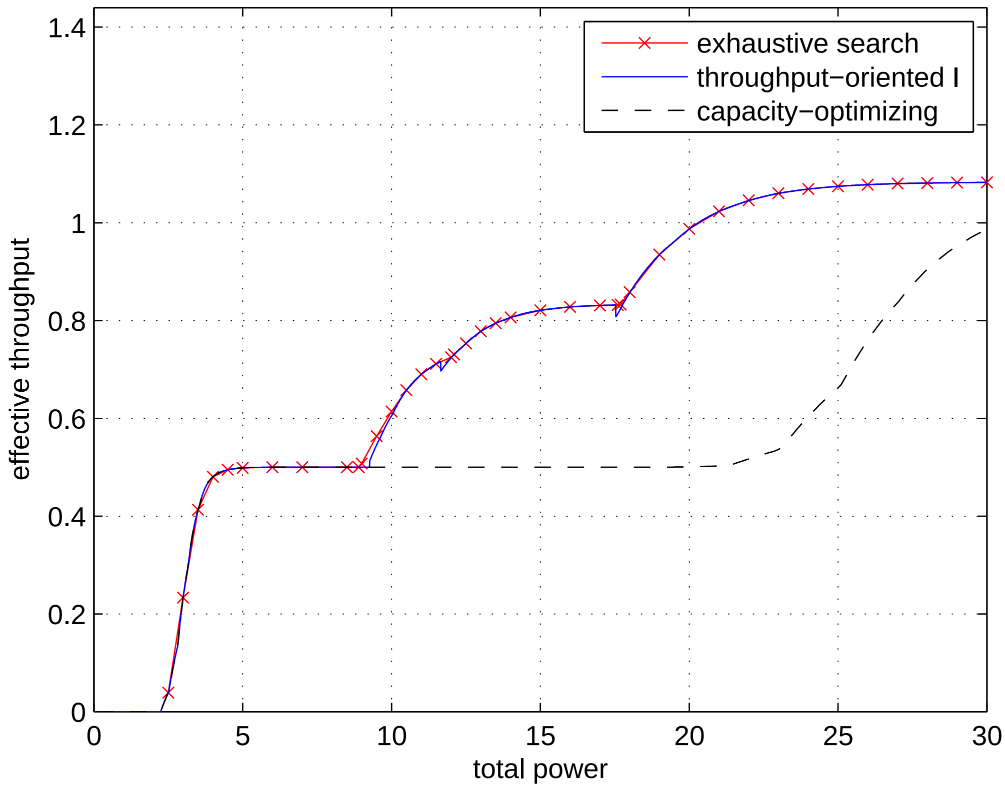

Figure 5.8: Case III: Optimal effective throughputs obtained from exhaustive search, the throughput-oriented water-filling I based on the FER approximation, and the capacity-optimizing water-filling policy.

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

In this section, three situations of parallel Gaussian channels are examined: 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^2_{total} is set to 10. Since the maximum effective throughput that could be achieved is the summation of the rate in each channel (denoted by maximum rate), we then compare the achieved percentage of maximum rate of different power allocation methods. In Figure. 5.9, we compare the achieved percentage of maximum rate obtained from exhaustive search with the throughput-oriented water-filling II in (4.11) and the traditional worst-case capacity-optimizing equal power allocation. The $\gamma^{\dagger}_{\text{thres}}$ where our proposed power allocation becomes optimal is 4.72 dB. We can see that almost all of the power allocation method achieve the maximum rate when system SNR is above $\gamma^{\dagger}_{\text{thres}}$. Although we cannot guarantee the optimality of using the throughput-oriented water-filling I for system SNR before $\gamma^{\dagger}_{\text{thres}}$, we can still see that it has around 1.4 dB gain over the equal power allocation when they achieve 85% of the maximum rate.

In Table 5.2 we provide the $\sigma_i^{2\dagger}$ in three channels for system SNR varying from 2 dB to 6 dB when using the throughput-oriented water-filling II. The result confirms our dissertation that the worst-case effective throughput is achieved by giving total noise power to only one channel for system SNR greater than $\gamma^{\dagger}_{\text{thres}}$. Also by our dissertation, allocating total noise power to either channel should yield the same effective throughput. However, when we allocate σ^2_{total} to channel 1 and 3 respectively, we yield slightly less effective throughput. We deduce that it is due to the inaccuracy from simulation and the approximation of the FER.

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

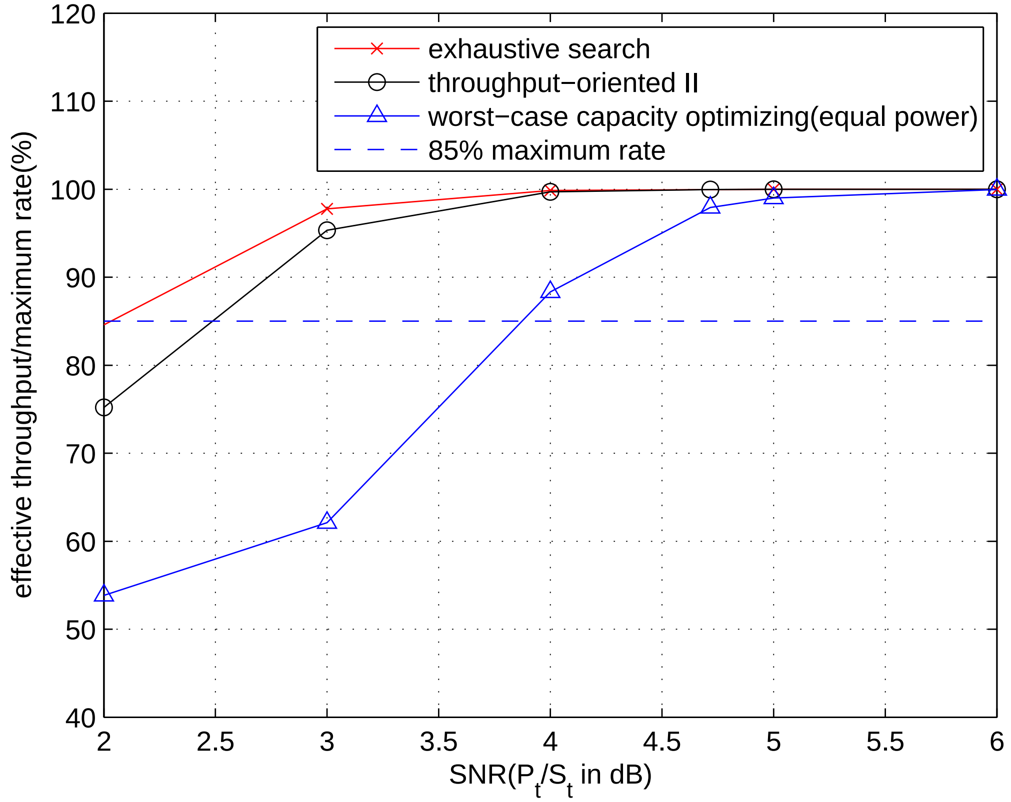

Figure 5.9: Case I: The worst-case effective throughputs obtained from exhaustive search, throughput-oriented water-filling II based on the FER approximation, and the worst-case-capacity-optimizing equal power allocation.

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 II should yield greater gain than equal power allocation 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 II yields around 2 dB gain when achieving 85% of the

Table 5.2: The $\sigma_i^{2\dagger}$ of each channel($K = 3$) for system SNR ranging from 2 dB to 6 dB.

	2 dB	3 dB	4 dB	5 dB	6 dB
$\sigma_1^{2\dagger}$	0	0	0	0	0
$\sigma_2^{2\dagger}$	10	10	10	10	10
$\sigma_3^{2\dagger}$	0	0	0	0	0

maximum rate. Besides, when we look at the situation when system SNR is equal to $\gamma_{\text{thres}}^{\dagger} = 5.19$ dB, throughput-oriented water-filling II almost achieves the maximum rate while the equal power allocation achieves only 83% of the maximum rate.

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 and the results are compared. 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 in gain when achieving 85% of the maximum rate(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 its puncture pattern as

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

The refined 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 becomes 1.89 dB when achieving 85% of the maximum rate, which is greater than the gain obtained in Case I. The result in this case confirms the fact that the throughput-oriented water-filling II yields larger gain from traditional equal power allocation when the characteristics of the used codes deviate largely from each other .

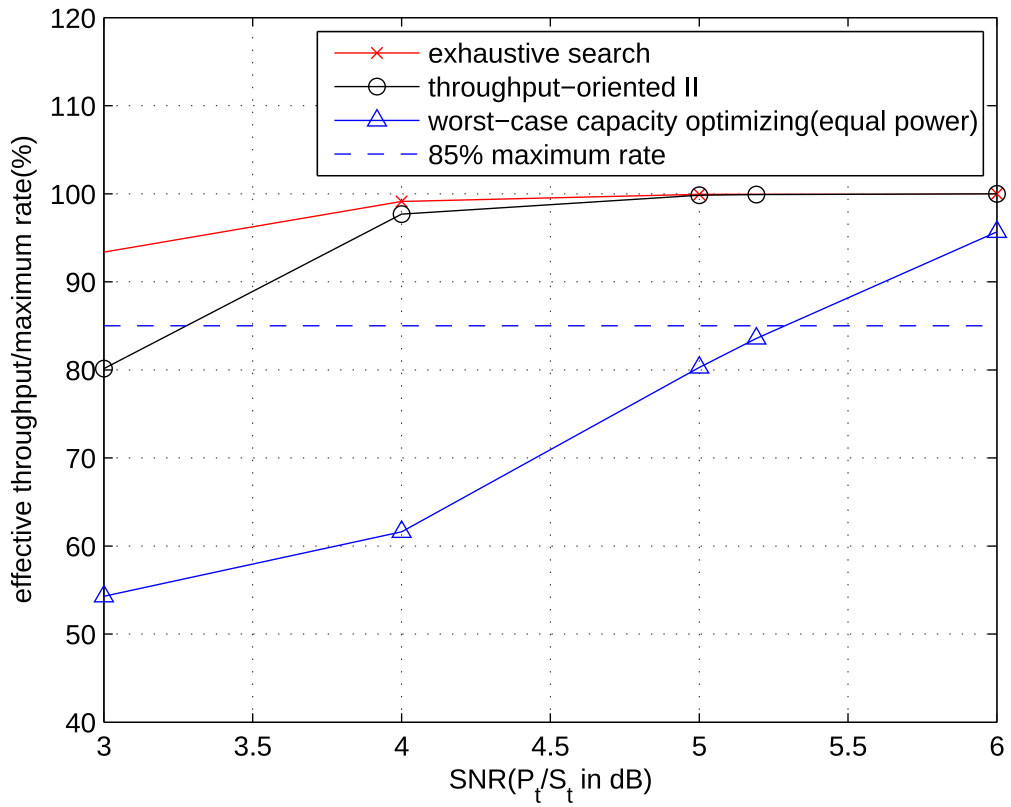

Figure 5.10: Case II: The worst-case effective throughputs obtained from exhaustive search, throughput-oriented water-filling II based on the FER approximation, and the the worst-case-capacity-optimizing equal power allocation.

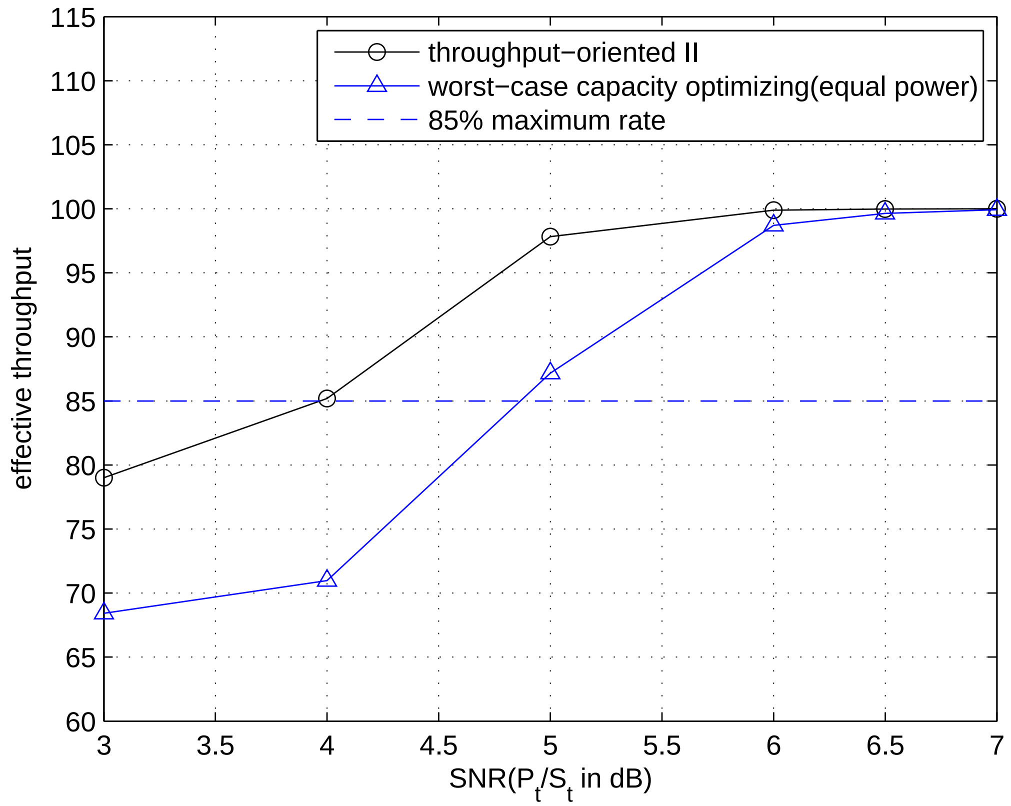

Figure 5.11: Case III: The worst-case effective throughput of using throughput-oriented water-filling II and equal power allocation. $K = 4$, $(2, 1, 6)$, $(3, 1, 6)$ and $(4, 1, 6)$ code are used in the first three channels and $(2, 1, 6)$ is used again in the fourth channel.

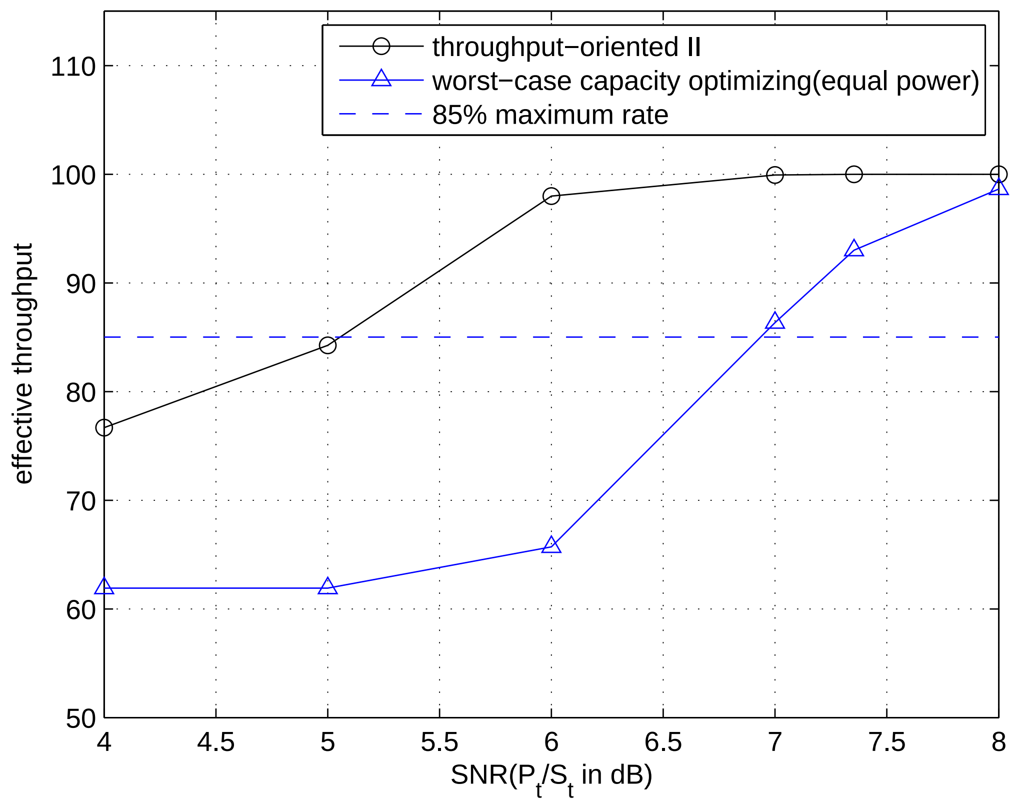

Figure 5.12: Case III: The worst-case effective throughput of using the throughput-oriented water-filling II and equal power allocation. $K = 4$, $(2, 1, 6)$, $(3, 1, 6)$ and $(4, 1, 6)$ code are used in the first three channels and $(3, 2, 6)$ is used in the fourth channel.

Chapter 6

Conclusion

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

In practice, standards usually provide a list of optional codes for each channel. For the case 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 $(2^K - 1)$ cases) such that our policy can readily determine the code to be used in each channel. For the case only total noise variance is available, the future work is to find the optimal power allocation policy for system SNR before 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.