SIGNAL AND DETECTOR RANDOMIZATION FOR MULTIUSER AND MULTICHANNEL COMMUNICATION SYSTEMS

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND

ELECTRONICS ENGINEERING

AND THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

By

Mehmet Emin Tutay

November 2013

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Assoc. Prof. Dr. Sinan Gezici (Supervisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Prof. Dr. Orhan Arıkan

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Assoc. Prof. Dr. İbrahim Körpeoğlu

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Asst. Prof. Dr. Behçet Uğur Töreyin

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Assoc. Prof. Dr. Ali Cafer Gürbüz

Approved for the Graduate School of Engineering and Science:

Prof. Dr. Levent Onural
Director of the Graduate School of Engineering and Science

ABSTRACT

SIGNAL AND DETECTOR RANDOMIZATION FOR MULTIUSER AND MULTICHANNEL COMMUNICATION SYSTEMS

Mehmet Emin Tutay

Ph.D. in Electrical and Electronics Engineering

Supervisor: Assoc. Prof. Dr. Sinan Gezici

November 2013

Randomization can be considered as a possible approach to enhance error performance of communication systems subject to average power constraints. In the first part of this dissertation, we consider downlink of a multiuser communications system subject to an average power constraint, where randomization is employed at the transmitter and receiver sides by modeling signal levels as random variables (stochastic signals) and employing different sets of detectors via time-sharing (detector randomization), respectively. In the second part, we consider single-user systems, where we assume that there exist multiple channels between the transmitter and receiver with arbitrary noise distributions over each of them and only one of the channels can be employed for transmission at any given time. In this case, randomization is performed by choosing the channel in use according to some probability mass function and employing stochastic signaling at the transmitter.

First, the jointly optimal power control with signal constellation randomization is proposed for the downlink of a multiuser communications system. Unlike a conventional system in which a fixed signal constellation is employed for all the

bits of a user (for given channel conditions and noise power), power control with signal constellation randomization involves randomization/time-sharing among different signal constellations for each user. A formulation is obtained for the problem of optimal power control with signal constellation randomization, and it is shown that the optimal solution can be represented by a randomization of $(K+1)$ or fewer distinct signal constellations for each user, where K denotes the number of users. In addition to the original nonconvex formulation, an approximate solution based on convex relaxation is derived. Then, detailed performance analysis is presented when the receivers employ symmetric signaling and sign detectors. Specifically, the maximum asymptotical improvement ratio is shown to be equal to the number of users, and the conditions under which the maximum and minimum asymptotical improvement ratios are achieved are derived. In the literature, it is known that employing different detectors with corresponding deterministic signals via time-sharing may enhance error performance of communications systems subject to average power constraints. Motivated by this result, as a second approach, we study optimal detector randomization for the downlink of a multiuser communications system. A formulation is provided to obtain optimal signal amplitudes, detectors, and detector randomization factors. It is shown that the solution of this joint optimization problem can be calculated in two steps, resulting in significant reduction in computational complexity. It is proved that the optimal solution is achieved via randomization among at most $\min\{K, N_\mathrm{d}\}$ detector sets, where K is the number of users and N_d is the number of detectors at each receiver. Lower and upper bounds are derived on the performance of optimal detector randomization, and it is proved that the optimal detector randomization approach can reduce the worst-case average probability of error of the optimal approach that employs a single detector for each user by up to K times. Various sufficient conditions are obtained for the improvability and nonimprovability via detector randomization. In the special case of equal crosscorrelations and noise powers, a simple solution is developed for the optimal detector randomization problem, and necessary and sufficient conditions are presented for the uniqueness of that solution.

Next, a single-user $M-$ary communication system is considered in which the transmitter and the receiver are connected via multiple additive (possibly non-Gaussian) noise channels, any one of which can be utilized for a given symbol transmission. Contrary to deterministic signaling (i.e., employing a fixed constellation), a stochastic signaling approach is adopted by treating the signal values transmitted for each information symbol over each channel as random variables. In particular, the joint optimization of the channel switching (i.e., time-sharing among different channels) strategy, stochastic signals, and decision rules at the receiver is performed in order to minimize the average probability of error under an average transmit power constraint. It is proved that the solution to this problem involves either one of the following: (i) deterministic signaling over a single channel, (ii) randomizing (time-sharing) between two different signal constellations over a single channel, or (iii) switching (time-sharing) between two channels with deterministic signaling over each channel. For all cases, the optimal strategies are shown to employ corresponding maximum *a posteriori* probability (MAP) decision rules at the receiver.

Keywords: Multiuser, Downlink, Probability of Error, Minimax, Detection, Stochastic Signaling, Detector Randomization, Channel Switching, M-ary Communications, Gaussian Noise, Multimodal Noise, Power Constraint.

ÖZET

ÇOK KULLANICILI VE ÇOK KANALLI HABERLEŞME SİSTEMLERİ İÇİN SİNYAL VE SEZİCİ RASTGELELEŞTİRME

Mehmet Emin Tutay
Elektrik ve Elektronik Mühendisliği, Doktora
Tez Yöneticisi: Doç. Dr. Sinan Gezici
Kasım 2013

Rastgeleleştirme, ortalama güç kısıtlı iletişim sistemlerinde hata performansını artırmak için muhtemel bir yaklaşım olarak düşünülebilir. Bu tezin ilk kısmında, rastgeleleştirmenin verici ve alıcıda sırasıyla, işaret seviyelerinin rastgele değişkenler (stokastik işaretler) olarak modellenerek ve farklı sezici kümelerinin zaman paylaşımı (sezici rastgeleleştirme) ile kullanılarak uygulandığı çok kullanıcılı sistemlerin aşağı bağlantısı ortalama güç kısıtı altında ele alınmaktadır. Tezin ikinci kısmında, tek kullanıcılı sistemler ele alınmakta, verici ve alıcı arasında çeşitli gürültü dağılımlara sahip çoklu kanalların olduğu ve anlık olarak bu kanallardan sadece birinin kullanılabildiği varsayılmaktadır. Bu durumda rastgeleleştirme, iletişim yapılacak kanalın belirli olasılık yığın fonksiyonuna göre seçilmesiyle ve vericide stokastik işaretleme kullanılmasıyla gerçekleştirilmektedir.

İlk olarak, çok kullanıcılı sistemlerin aşağı bağlantısı için işaret yıldız kümelerinin rastgeleleştirilmesiyle optimal güç kontrolü önerilmektedir. Kullanıcının bütün bitleri için sabit bir işaret yıldız kümesinin (verilen kanal şartları ve gürültü gücü için) kullanıldığı geleneksel sistemlerin aksine, işaret yıldız

kümelerinin rastgeleleştirilmesiyle yapılan güç kontrolü, her bir kullanıcı için farklı işaret yıldız kümeleri arasında rastgeleleştirme/zaman paylaşımı gerektirebilmektedir. İşaret yıldız kümelerinin rastgeleleştirilmesiyle optimal güç kontrolü problemi için bir formülasyon elde edilmekte ve her bir kullanıcı için optimal çözümün $(K+1)$- burada K kullanıcı sayısını göstermekte - veya daha az işaret yıldız kümeleri arasında rastgeleleştirme ile ifade edilebileceği gösterilmektedir. Özgün dış bükey olmayan formülasyona ek olarak, dışbükey gevşetme metoduna dayanan yaklaşık bir çözüm elde edilmektedir. Daha sonra, alıcılarda simetrik işaretleme ve işaret sezicileri kullanıldığı durumda detaylı başarım analizi sunulmaktadır. Daha açık bir ifadeyle, en büyük asimptotik gelişim oranının kullanıcı sayısına eşit olduğu gösterilmekte ve en büyük ve en küçük asimptotik gelişim oranlarına erişilmesi için koşullar elde edilmektedir.

Literatürde, ortalama güç kısıtlı iletişim sistemlerinde deterministik işaretleme ile çalışan sezicilerin zaman paylaşımı ile kullanımının, hata performansını iyileştirebileceği bilinmektedir. Bu sonuçtan hareketle, ikinci yaklaşım olarak, çok kullanıcılı sistemlerin aşağı bağlantısı için optimal sezici rastgeleleştirme problemi çalışılmaktadır. Optimal işaret genliklerinin, sezicilerin ve sezici rastgeleleştirme oranlarının elde edilebilmesi için bir formülasyon sunulmaktadır. Bu ortak eniyileme probleminin çözümünün, hesaplama karmaşıklığının önemli ölçüde daha az olduğu iki aşamada hesaplanabildiği gösterilmektedir. Optimal çözüme en fazla $\min\{K, N_{\mathrm{d}}\}$ sezici seti arasında rastgeleleştirme ile ulaşıldığı -burada K kullanıcı sayısını ve N_{d} her bir alıcıdaki sezici sayısını göstermekte- ispatlanmaktadır. Optimal sezici rastgeleleştirme başarımı için alt ve üst sınırlar elde edilmekte ve optimal sezici rastgeleleştirme yaklaşımının en kötü durumdaki hata olasılığını her bir kullanıcı için bir sezici kullanan optimal yaklaşıma göre K oranında azaltabildiği ispatlanmaktadır. Sezici rastgeleleştirme ile iyileşmenin olabileceği ve olamayacağı yeter koşullar elde edilmektedir. Çapraz ilintilerin ve gürültü güçlerinin eşit olduğu özel durumlarda, optimal sezici rastgeleleştirme problemi için basit bir çözüm geliştirilmekte ve çözümün tek olması için gerek ve yeter koşullar sunulmaktadır.

Daha sonra, verici ve alıcı arasında, verilen bir sembol iletimi için herhangi birinin kullanılabildiği çoklu toplanabilir gürültü kanalları (Gaussian olmayabilir) bulunan, tek kullanıcılı M-li iletişim sistemleri ele alınmaktadır. Deterministik işaretlemenin (sabit yıldız kümesi kullanmanın) aksine, her bilgi sembolü için her bir kanal üstünden gönderilen işaret değerlerini rastgele değişkenler olarak ele alan stokastik işaretleme benimsenmektedir. Özellikle, ortalama güç kısıtı altında ortalama hata olasılığını enküçültmek için kanal anahtarlama yöntemi, stokastik işaretler ve alıcıdaki karar kurallarının ortak eniyilemesi gerekleştirilmektedir. Bu problemin çözümünün şunlardan herhangi biri olduğu ispat edilmektedir: (i) tek kanal üstünden deterministik işaretleme, (ii) tek kanal üstünden iki farklı yıldız kümesi arasında rastgeleleştirme (zaman paylaşımı), (iii) her biri deterministik işaretleme kullanan iki kanal arasında anahtarlama (zaman paylaşımı). Bütün durumlarda, optimal yöntemlerin alıcıda maksimum sonsal olasılık karar kurallarını kullandığı gösterilmektedir.

Anahtar Kelimeler: Çok Kullanıcı, Aşağı Bağlantı, Hata Olasılığı, En Küçük En Büyük, Sezim, Stokastik İşaretleme, Sezici Rastgeleleştirme, Kanal Anahtarlama, M-li iletişim, Gauss Gürültüsü, Çok Doruklu Gürültü, Güç Kısıtı.

ACKNOWLEDGMENTS

I would like to thank my supervisor Assoc. Prof. Sinan Gezici and Prof. Orhan Arıkan for their help and advices.

I am very thankful to Assoc. Prof. İbrahim Körpeoğlu for accepting to be in my dissertation monitoring committee. I would also like to thank Asst. Prof. Behçet Uğur Töreyin and Assoc. Prof. Ali Cafer Gürbüz for accepting to be in my exam committee.

This thesis study was supported in part by the National Young Researchers Career Development Programme (project no. 110E245) of the Scientific and Technological Research Council of Turkey (TUBITAK). I appreciate the financial support of TUBITAK.

Contents

List of Figures

List of Tables

Dedicated to the memory of my teacher,

Erdal Can

Chapter 1

Introduction

The main motivation behind this study is the recent results in which randomization is shown to be an effective method for performance improvement in terms of average error probability. Specifically, communications systems subject to average power constraints are studied for single-user scenarios in [2–6], where randomization is performed by modeling transmitted signal levels as random variables (also referred to as *stochastic signaling*), employing different detectors with corresponding deterministic signals via time-sharing (called *detector randomization*), and employing different channels via time-sharing (i.e., channel-switching). In the first part of this dissertation, downlink of a multiuser communications system subject to some average power constraint is considered under stochastic signaling and detector randomization approaches in the presence of Gaussian noise. In the second part, a single-user scenario is considered in the presence of multiple channels with any generic noise probability density functions (PDFs), when stochastic signaling can be employed at the transmitter for each channel. In both parts, it is shown that the optimal randomization strategy can be represented by discrete probability distributions with certain numbers of point masses. In the following, the previous related work in the literature and the main contributions of this dissertation are presented.

1.1 Multiuser Case

Recently, the effects of randomization or time-sharing have been investigated in various studies such as [2–13]. In [2], the convexity properties of error probability in terms of signal and noise power are investigated for binary-valued scalar signals over additive unimodal noise channels under an average power constraint. Based on the convexity results, the scenarios in which power randomization can or cannot be useful for improving error performance are determined, and optimal strategies for jammer power randomization are developed. The study in [3] generalizes the results of [2] by exploring the convexity properties of the error probability for constellations with arbitrary shape, order, and dimensionality for a maximum likelihood (ML) detector in the presence of additive Gaussian noise with no fading and with frequency-flat slowly fading channels. For communications systems that operate over time-invariant non-Gaussian channels [14], randomization (time-sharing) among multiple signal constellations can improve performance of a given receiver in terms of error probability. Specifically, it is shown in [4] that randomization among up to three distinct signal constellations can reduce the average probability of error of a communications system that operates under second and fourth moment constraints. In addition, [5] investigates the joint optimization of the signal constellation randomization and detector design under an average power constraint and shows that the use of at most two distinct signal constellations and the corresponding maximum *a posteriori* probability (MAP) detector minimizes the average probability of error.

In a different context, time-varying or random signal constellations are utilized in [15–20] for the purpose of enhancing error performance or achieving diversity. In [15], the author proposes (pseudo)randomly rotating the signal constellation for each transmitted vector in order to improve the coded frame-error-rate of spatial multiplexing in block fading. The advantages of this approach in reducing the outage probability are investigated in [16]. Although a form of constellation randomization is performed in [15, 16], they are different from the work in Chapter 2 of this thesis since a (pseudo)random rotation of the signal constellation is proposed for a single-user system in those studies, whereas we obtain *optimal randomization* of signal constellations for a *multiuser* system in this thesis. In addition, the studies in [17–20] consider random signal mapping, random rotations, or time-varying phase shifts to transmitted signals in order to achieve diversity.

1.1.1 Optimal Randomization of Signal Constellations on the Downlink of a Multiuser DS-CDMA System

In the first part of this thesis, we consider a generic problem on the signal constellation design for the downlink of a binary multiuser communications system in which users can randomize or time-share among multiple signal constellations. Unlike conventional systems in which a fixed signal constellation is employed for all the bits of a user (for given channel conditions and noise power) [21], we formulate a generic problem that can involve randomization/time-sharing among different signal constellations for each user. Due to such randomization/time-sharing, the signal amplitude corresponding to each bit of a user can be modeled as a generic random variable in this approach. Therefore, the problem can be formulated as obtaining the optimal probability distribution for the signal amplitude corresponding to each bit of each user in a multiuser system.

Since the signal amplitudes for all bits of all users are modeled as generic random variables in the power control with signal constellation randomization problem in this study, the proposed approach can also be considered as a generalization of randomized power control algorithms in the literature from various perspectives [22–27].[1] First, as the power control with signal constellation randomization approach can result in strategies in which different power allocation strategies are employed for different bits of a given user, it is a more generic approach than randomized power control in general. Second, the proposed approach is employed for each state of the channel whereas power control algorithms are used with respect to varying channel conditions. In other words, the power control strategies in the literature adapt the power as the channel state changes, whereas the proposed approach performs constellation randomization for a given (fixed) channel state. Third, even for the symmetric signaling case (in which signal amplitudes for bit 0 and bit 1 are negatives of each other, and the same power allocation strategy is employed for bit 0 and bit 1 for each user), the proposed approach is different from those in the literature [22–27] since it models the signal amplitudes (powers) of the users as generic random variables and obtains the optimal probability distributions of those random variables that minimize a probability of error metric. (The main intuition behind the benefits of this approach is that when the signal amplitudes (powers) are modeled as random variables, various time-sharing (randomization) strategies can be implemented in order to optimize the error performance of the system, as investigated in Sections 2.3-2.5.) For example, in [22], transmit powers are selected from a discrete set of power levels, namely, zero and peak power, and optimal power randomization strategies are obtained under that specification for a two-hop interference channel.[2] [23] considers the same strategy for power control in ad-hoc sensor networks, and works on the optimization of transmission (on-state) probability to meet certain quality of service requirements. In another study [24], a random power control algorithm is proposed, in which the transmitter selects its power level *randomly* from a uniform distribution. It is shown that this approach can improve network connectivity over the fixed power control approach in the case of static channels. However, the performance of this uniform power selection approach deteriorates

[1]Please refer to [28–30] for surveys on power control in wireless networks.

[2]Please refer to [31] and [32] for other game theoretic approaches for power control.

in fading channels, as investigated in [25]. In [26], random power allocation according to a certain probability distribution is proposed. Namely, the transmit power is modeled by a truncated inverted exponential distribution, and the parameter of this distribution is updated at certain intervals based on feedback. The connectivity analysis of this approach is presented in [27] for wireless sensor networks, and improvements in energy efficiency are observed.[3]

Motivated by the recent results that illustrate the improvements obtained via randomization [2–10, 15, 35, 36], the aim of this study is to formulate a generic power control problem with signal constellation randomization for the downlink of a multiuser communications system in which the signal amplitude for each bit of a user is modeled as a random variable. In other words, by adopting the approach in [4], the aim is to jointly design the optimal randomization of signal constellations for all users in the downlink of a direct sequence code division multiple access (DS-CDMA) system in order to optimize error performance for given receiver structures. The main challenge in the joint design of signal constellation randomization is that signal amplitudes of each user affect not only its own error performance but also the performance of all other users via interference.

The main contributions of Chapter 2 can be summarized as follows:

- The joint design of optimal randomization of signal constellations is performed in a multiuser system for the first time.
- It is shown that the optimal power control with signal constellation randomization results in a randomization among up to $(K+1)$ different signal constellations for each user, where K is the number of users.

[3] In [33] and [34], the term "stochastic power control" is used in a different meaning from "randomized power control" in [22–27]. Specifically, [33] and [34] do not employ any power or signal randomization but apply an approach that is based on measurements (which are inherently random) instead of known deterministic parameters.

- In addition to the generic problem formulation, which needs to be solved via global optimization algorithms due to its nonconvex nature, an approximate convex solution is obtained based on convex relaxation.

1.1.2 Optimal Detector Randomization for Multiuser Communications Systems

In the previous scenario, the downlink of a multiuser system is considered in which randomization is employed at the transmitter by modeling transmitted signal levels as random quantities, while at the receiver of each user a fixed decision rule (e.g., sign detector) is employed. Another technique for enhancing error performance of some communications systems that operate over time-invariant channels is to perform detector randomization, which involves the use of multiple detectors at the receiver with certain probabilities (certain fractions of time) [6–8], [37, 38]. In other words, a receiver can randomize (time-share) among multiple detectors in order to reduce the average probability of error. In [7], randomization between two antipodal signal pairs and the corresponding MAP detectors is performed for an average power constrained binary communications system, and significant performance improvements are observed as a result of detector randomization in some cases in the presence of symmetric Gaussian mixture noise. In [6], the results in [7] and [5] are extended by considering both detector randomization and signal constellation randomization for an average power constrained M-ary communications system. It is proved that the joint optimization of detector and signal constellation randomization results in a randomization between at most two MAP detectors corresponding to two deterministic signal constellations. The study in [6] is extended to the Neyman-Pearson (NP) framework in [37] by considering a power constrained on-off keying communications systems. As discussed in [39], detector randomization can be regarded as a generalization of noise enhanced detection with a fixed detector [9, 13]. In addition, when variable detectors are considered, noise enhanced detection and detector randomization can be considered as alternative approaches.[4] In [8], probability distributions of optimal additive noise components are investigated for variable detectors, and the optimal randomization between detector and additive noise pairs is investigated for optimal noise enhancement.

Although detector randomization has recently been investigated, e.g., in [6–8, 37], no previous studies have considered detector randomization for multiuser communications systems. In Chapter 3 of this dissertation, we study optimal detector randomization for multiuser communications systems. In particular, we consider the downlink of a direct sequence spread spectrum (DSSS) communications system under an average power constraint, and propose an optimization problem to obtain optimal signal amplitudes (corresponding to information symbols for different users), detectors, and detector randomization factors (probabilities) that minimize the worst-case (maximum) average probability of error of the users. Since this joint optimization problem is quite complex in its original formulation, a low-complexity approach is developed in order to obtain the optimal solution in two steps, where the optimal signal amplitudes and detector randomization factors are calculated in the first step, and the corresponding ML detectors are obtained in the second step. Also, it is shown that the optimal solution requires randomization among at most $\min\{K, N_{\mathrm{d}}\}$ detectors for each user, where K is the number of users and N_{d} is the number of detectors at each receiver. In addition, the performance of the optimal detector randomization approach is investigated, and a lower bound is presented for the minimum worst-case average probability of error. It is proved that the optimal detector randomization approach can improve the performance of the optimal approach that employs a single detector for each user (i.e., no detector randomization)

[4]The main difference is that an additive noise component is employed at the detector in the noise enhanced detection approach whereas the transmitted signal values are adapted according to the detector randomization strategy in the detector randomization approach.

by up to K times. Sufficient conditions are derived for the improvability and nonimprovability via detector randomization. Furthermore, in the special case of equal crosscorrelations and noise powers, a simple solution is proposed for the optimal detector randomization problem, and necessary and sufficient conditions are obtained for the uniqueness of that solution. Finally, numerical examples are presented in order to illustrate the improvements achieved via detector randomization. Although the results in this study are obtained for the downlink of a binary DSSS system, possible extensions to uplink scenarios and M-ary systems are discussed in Section 3.5.

It should be emphasized that detector randomization in this study is designed for time-invariant channels; equivalently, detector randomization is performed for each channel realization assuming that channel statistics do not change for a certain number of symbols [6, 7, 37]. Therefore, the proposed approach is different from power control (and detector adaptation) algorithms that are developed for varying channel conditions [28–30]. In addition, randomized power control algorithms in the literature, such as [22–27], employ significantly different approaches than that in this study.

1.2 Single-User Case

1.2.1 Optimal Signaling and Detector Design for $M-$ary Communications Systems in the Presence of Multiple Additive Noise Channels

When multiple channels are present between a transmitter and a receiver, it may be advantageous to perform channel switching; that is, to transmit over one channel for a certain fraction of time, and then switch to another channel during the next transmission period even if the channel statistics are not varying with

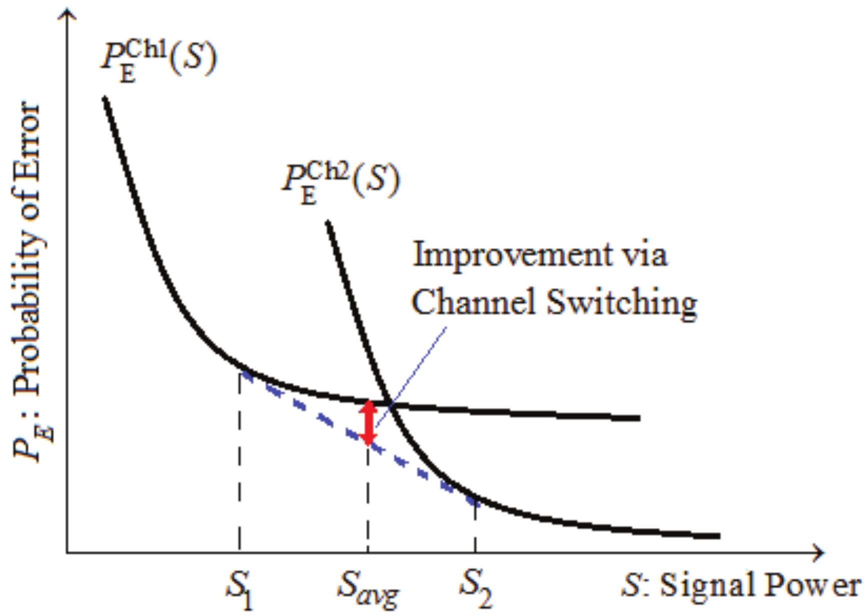

Figure 1.1: Illustrative example demonstrating the benefits of switching between two channels under an average power constraint [1].

time [2, 40, 41]. Figure 1.1 illustrates this fact for an average power constrained binary communications system which employs antipodal signaling with $-\sqrt{\mathrm{S}}, \sqrt{\mathrm{S}}$ for a given signal power S. It is seen that the average probability of error can be reduced by switching (time-sharing) between channel 1 and channel 2 with respective power levels S_1 and S_2 in comparison to the constant power transmission scheme that employs power $\mathrm{S}_{\mathrm{avg}}$ exclusively over channel 1. More precisely, time-sharing exploits the nonconvexity of the plot for the *minimum* of the error probabilities over both channels as a function of the signal power. The resulting strategy yields the convex hull of the individual error probability functions. This observation is first noted in [2] while studying the convexity properties of error probability with respect to the transmit signal power for the optimal detection of *antipodal* signals corrupted by additive *unimodal* noise. It is shown that the optimum performance under an average power constraint can be achieved by time-sharing between at most two channels and power levels.

In Chapter 4 of this dissertation, we study the optimal signaling and detection strategy that minimizes the average probability of error for an average power constrained M-ary communications system in which the transmitter and the receiver are connected via multiple additive noise channels. Similar to [2], it is assumed that only a single channel is used for symbol transmission at any given time instant. Although the analysis in [2] is restricted to unimodal noise distributions and deterministic binary antipodal signals, we consider generic noise PDFs (i.e., including non-Gaussian or multimodal cases), and a stochastic signaling approach by assuming that the transmitter can perform signal randomization for each information symbol sent over any one of the channels. More specifically, we investigate the *joint* optimization of the channel switching strategy, stochastic signals (employed for the transmission of each symbol over each channel), and decision rules (used for each channel at the receiver) in order to minimize the average probability of error under an average transmit power constraint.

The main contributions of Chapter 4 can be summarized as follows:

- A generic problem formulation is proposed for the optimal signaling and detection problem in the presence of multiple additive noise channels by considering the joint optimization of the channel switching strategy, stochastic signals, and detectors without imposing any restrictions on probability distributions of channel noise.

- It is proved that the solution to this generic problem corresponds to either (i) deterministic signaling (i.e., employing a fixed constellation) over a single channel with the corresponding MAP detector, (ii) randomizing (time-sharing) between two different signal constellations over a single channel with the corresponding MAP detector, or (iii) switching (time-sharing) between the MAP detectors of two channels with deterministic signaling over each channel.

In addition, numerical examples are provided to illustrate the improvements that can be achieved via the optimal signaling and detection strategy. The results in this study generalize some of the previous studies in the literature and cover them as special cases. For example, in the absence of channel switching (i.e., in the presence of a single channel between the transmitter and the receiver) and for binary communications, the results reduce to those in [5]. In addition, in the absence of stochastic signaling and when the channel noise is assumed to have a unimodal differential PDF for a binary communications system, the problem considered in this study covers the one in [2] as a special case.

1.3 Organization of the Dissertation

This dissertation is organized as follows. In Chapter 2, downlink of a multiuser communications system is considered in the presence of Gaussian noise when fixed decision rules (specifically, sign detectors) are employed at the receiver of each user [42]. The system is subject to an average power constraint and the objective is to find the optimal signal constellation randomization to minimize the worst-case average error probability. Chapter 3 considers the scenario in Chapter 2 based on a different approach [43]. Namely, It is assumed that each user has N_{d} detectors at the receiver and the objective is to jointly optimize randomization factors, detectors and corresponding deterministic signals to minimize the worst-case error probability. Another important difference is that power is assumed to be limited for a bit duration, while in Chapter 2 the time average power constraint is considered. In Chapter 4, single-user systems are considered in the presence of multiple channels with any generic noise PDFs when stochastic signaling is adopted at the transmitter for each channel. The objective is to optimize stochastic signals, channel switching factors, and detectors to minimize the average error probability. Finally, Chapter 5 concludes this dissertation by providing an overall summary.

Chapter 2

Optimal Randomization of Signal Constellations on the Downlink of a Multiuser DS-CDMA System

This chapter is organized as follows. In Section 2.1, the system model is introduced and receiver structures are described. In Section 2.2, the optimal power control with signal constellation randomization problem is formulated and theoretical results are obtained for generic detector structures at the receivers. Specific results are obtained for sign detectors in Section 2.3. In Section 2.4, numerical examples are provided to illustrate the improvements obtained via the proposed power control with signal constellation randomization approach. Concluding remarks are made and possible extensions to uplink scenarios and M-ary systems are discussed in Section 2.5.

2.1 System Model

Consider the downlink of a multiuser DS-CDMA binary communications system, in which the baseband model for the transmitted signal is given by

$$p(t) = \sum_{k=1}^{K} S_k^{(i_k)} c_k(t) \ , \tag{2.1}$$

where K is the number of users, $S_k^{(i_k)}$ denotes the amplitude of the kth user's signal corresponding to bit i_k, with $i_k \in \{0, 1\}$, and $c_k(t)$ is the real pseudo-noise signal for user k. The pseudo-noise signals spread the spectra of users' signals and provide multiple-access capability [21]. Information intended for user k is carried by $S_k^{(i_k)}$, which corresponds to bit 0 for $i_k = 0$ and bit 1 for $i_k = 1$. $S_k^{(i_k)}$'s are modeled as real numbers, and they scale the amplitudes of the pseudo-noise signals, $c_k(t)$'s. It is assumed that bit 0 and bit 1 are equally likely (i.e., the prior probabilities of the bits are equal to 0.5) for all users, and the information bits for different users are independent.

The signal in (2.1) is transmitted to K users, and the received signal at user k is represented by

$$r_k(t) = \sum_{l=1}^{K} S_l^{(i_l)} c_l(t) + n_k(t) \ , \tag{2.2}$$

for $k = 1, \ldots, K$, where $n_k(t)$ denotes the noise at the receiver of user k, which is modeled as a zero-mean white Gaussian process with spectral density σ_k^2. It is assumed that the noise processes at different receivers are independent. Although a simple additive noise model is employed in (2.2), multipath channels with slow *frequency-flat fading* can also be covered by the considered model if perfect channel estimation is assumed at the receivers [4]. In that case, the signal in (2.2) can be considered as the scaled version of the received signal by the inverse of the channel coefficient; hence, the average power of the noise component in (2.2), σ_k^2, would correspond to the average noise power in the received signal divided by the channel power gain. (In other words, the effects of frequency-flat

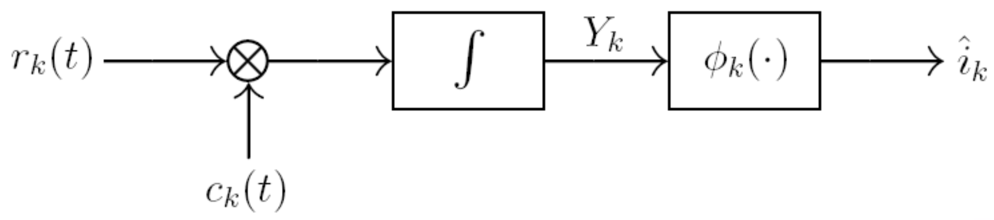

Figure 2.1: Receiver structure for user k.

fading can be taken into account by incorporating channel power gains into the σ_k^2 terms in (2.2).)

The receiver for user k processes the signal in (2.2) as shown in Figure 2.1. Specifically, the received signal $r_k(t)$ is correlated with the pseudo-noise signal for user k, $c_k(t)$, which effectively corresponds to a despreading operation, and then the correlator output is used by a generic detector in order to estimate the transmitted bit for user k. Based on (2.2), the correlator output for user k can be expressed as

$$Y_k = S_k^{(i_k)} + \sum_{\substack{l=1 \\ l \neq k}}^{K} \rho_{k,l} S_l^{(i_l)} + N_k \ , \tag{2.3}$$

for $k = 1, \ldots, K$, where $\rho_{k,l} \triangleq \int c_k(t) c_l(t) dt$ denotes the crosscorrelation between the pseudo-noise signals for user k and l (it is assumed without loss of generality that $\rho_{k,k} = 1$ for $k = 1, \ldots, K$), and $N_k \triangleq \int n_k(t) c_k(t) dt$ is the noise component. It can be shown that $N_1, \ldots, N_K$ form a sequence of independent zero-mean Gaussian random variables with variances, $\sigma_1^2 \ldots, \sigma_K^2$, respectively. In (2.3), the first term corresponds to the desired signal component, the second term represents the multiple-access interference (MAI), and the last term is the noise component.

The correlator output in (2.3) is used by a generic detector (decision rule) ϕ_k to generate an estimate of the transmitted information bit, as shown in Figure 2.1.

Specifically, for a given correlator output $Y_k = y_k$, the bit estimate is denoted as

$$\hat{i}_k = \phi_k(y_k) = \begin{cases} 0 \ , & y_k \in \Gamma_{k,0} \\ 1 \ , & y_k \in \Gamma_{k,1} \end{cases} \tag{2.4}$$

for $k = 1, \ldots, K$, where $\Gamma_{k,0}$ and $\Gamma_{k,1}$ denote the decision regions for bit 0 and bit 1, respectively, and they form a partition of the observation space [44]. In the next section, theoretical results are obtained for generic detectors at the receivers; that is, ϕ_k's can be arbitrary decision rules.

2.2 Power Control with Signal Constellation Randomization for Multiuser Systems

2.2.1 Optimal Power Control with Signal Constellation Randomization

In conventional systems, $S_k^{(i_k)}$ in (2.1) corresponds to a fixed value for each bit of a given user; in other words, a signal constellation is selected for each user, and it is employed for all the bits in the multiuser system (for given channel conditions and noise power). For example, consider a two-user system, in which bit 0 and bit 1 are represented by -1 and 1, respectively, for user 1, and by -0.5 and 0.5, respectively, for user 2. Then, the *joint signal constellation* for the two users is represented by $\left(S_1^{(0)}, S_1^{(1)}, S_2^{(0)}, S_2^{(1)}\right) = (-1, 1, -0.5, 0.5)$. In this case, there is no randomization or time-sharing among multiple signal constellations, and a fixed signal constellation is employed for all the bits of each user in the system for given channel conditions and noise power. A specific example is illustrated in Table 2.1-(A) when 12 bits are transmitted for each user.

Unlike conventional systems, we consider *power control with signal constellation randomization* in this study and model $S_k^{(i_k)}$ in (2.1) as generic random

Table 2.1: (A) Example of a conventional system in which no signal constellation randomization is employed. Joint signal constellation $\left(S_1^{(0)}, S_1^{(1)}, S_2^{(0)}, S_2^{(1)}\right) = (-1, 1, -0.5, 0.5)$ is used for all the bits. (B) Example of power control with signal constellation randomization in which half of the bits are transmitted according to joint signal constellation $(-0.7, 0.7, -0.4, 0.4)$ and the remaining half are transmitted according to $(-1.1, 1.1, -0.8, 0.8)$.

(A)

Bit of User 1 (i_1)	0	1	0	0	1	0	1	1	0	0	1	1
Amplitude of User 1's Signal $\left(S_1^{(i_1)}\right)$	-1	1	-1	-1	1	-1	1	1	-1	-1	1	1
Bit of User 2 (i_2)	1	0	1	0	0	1	1	0	1	0	0	1
Amplitude of User 2's Signal $\left(S_2^{(i_2)}\right)$	0.5	-0.5	0.5	-0.5	-0.5	0.5	0.5	-0.5	0.5	-0.5	-0.5	0.5

(B)

Bit of User 1 (i_1)	0	1	0	0	1	0	1	1	0	0	1	1
Amplitude of User 1's Signal $\left(S_1^{(i_1)}\right)$	-0.7	0.7	-1.1	-0.7	1.1	-1.1	0.7	1.1	-0.7	-1.1	0.7	1.1
Bit of User 2 (i_2)	1	0	1	0	0	1	1	0	1	0	0	1
Amplitude of User 2's Signal $\left(S_2^{(i_2)}\right)$	0.4	-0.4	0.8	-0.4	-0.8	0.8	0.4	-0.8	0.4	-0.8	-0.4	0.8

variables [4]. In this case, it is possible to employ different signal constellations for different bits in the system (for given channel conditions and noise power). In other words, randomization/time-sharing among multiple signal constellations is possible. For example, in a two-user system, one can time-share between joint signal constellations $\left(S_1^{(0)}, S_1^{(1)}, S_2^{(0)}, S_2^{(1)}\right) = (-0.7, 0.7, -0.4, 0.4)$ and $\left(S_1^{(0)}, S_1^{(1)}, S_2^{(0)}, S_2^{(1)}\right) = (-1.1, 1.1, -0.8, 0.8)$. Specifically, if half of the bits are sent according to the first set of signal constellations and the remaining half are sent according to the second one, the overall joint signal constellation, $\left(S_1^{(0)}, S_1^{(1)}, S_2^{(0)}, S_2^{(1)}\right)$, can be represented by a discrete random variable which is equal to $(-0.7, 0.7, -0.4, 0.4)$ or $(-1.1, 1.1, -0.8, 0.8)$ with equal probabilities. In Table 2.1-(B), this example of power control with signal constellation randomization is illustrated when 12 bit are transmitted for each user. As observed from the table, for user 1 (user 2), half of bits 0 are represented by -0.7 (-0.4) and the remaining half are represented by -1.1 (-0.8); similarly, half of bits 1 are represented by 0.7 (0.4) and the remaining half are represented by 1.1 (0.8) in order to implement the desired signal constellation randomization.

In order to provide a generic formulation of the proposed power control with signal constellation randomization approach in multiuser systems, let $\boldsymbol{S}$ denote the vector of random variables corresponding to the amplitudes of all users' signals for bit 0 and bit 1; that is,

$$\boldsymbol{S} = \left(S_1^{(0)}, S_1^{(1)}, S_2^{(0)}, S_2^{(1)}, \cdots, S_K^{(0)}, S_K^{(1)}\right) , \tag{2.5}$$

where $S_k^{(i_k)}$ is as in (2.1). In other words, $\boldsymbol{S}$ is the *joint signal constellation*, which is a $2K$ dimensional vector consisting of signal constellations for all users (as exemplified in the previous paragraphs), and it is modeled as a generic random vector in order to facilitate any type of signal constellation randomization. In addition, let $p_{\boldsymbol{S}}$ represent the probability density function (PDF) of $\boldsymbol{S}$. According to this definition, the conventional approach of *no constellation randomization* (or, *fixed signal constellations*) corresponds to a PDF in the form of $p_{\boldsymbol{S}}(\mathbf{s}) = \delta(\mathbf{s} - \mathbf{s}_0)$, where $\delta(\cdot)$ represents the Dirac delta function. (For instance, $p_{\boldsymbol{S}}(\mathbf{s}) = \delta\left(\mathbf{s} - (-1, 1, -0.5, 0.5)\right)$ for the example in Table 2.1-(A).) On the other hand, *any generic PDF* can be employed in the power control with signal constellation randomization approach considered in this study. (For instance, $p_{\boldsymbol{S}}(\mathbf{s}) = 0.5\,\delta(\mathbf{s} - (-0.7, 0.7, -0.4, 0.4)) + 0.5\,\delta(\mathbf{s} - (-1.1, 1.1, -0.8, 0.8))$ for the example in Table 2.1-(B).)

Based on the definition in (2.5), the aim is to find the optimal PDF of $\boldsymbol{S}$, i.e., the *optimal randomization of signal constellations*, in a given multiuser system. Considering a generic approach in the sense that the PDF of $\boldsymbol{S}$, $p_{\boldsymbol{S}}$, can be in any form (corresponding to discrete, continuous, or mixed random variables), we formulate the following power control with signal constellation randomization problem:

$$\min_{p_{\boldsymbol{S}}} \max_{k\in\{1,\dots,K\}} \mathrm{P}_k \tag{2.6}$$

$$\text{subject to } \mathrm{E}\left\{\int |p(t)|^2 dt\right\} \leq A \tag{2.7}$$

where P_k denotes the average probability of error for user k, $p(t)$ is as in (2.1), and A is a constraint on the average power of the transmitted signal. In other words, the aim is to find the optimal PDF for the joint signal constellation that minimizes the maximum of the average probabilities of error under a constraint on the average transmitted power. The minimax approach is adopted for fairness [45–48]; that is, for preventing scenarios in which the average probabilities of error are very low for some users whereas they are (unacceptably) high for others. Extensions to cases in which different users have different levels of importance are also possible as discussed in Section 2.5. It is noted that the formulation in (2.6)-(2.7) is similar to a max-min SINR problem [46]. However, the main differences are that the optimization in (2.6)-(2.7) is performed over the set of possible PDFs for the joint signal constellation, and that the considered probability of error metric leads to different solutions than the max-min SINR problem in general.

In order to express the optimization problem in (2.6)-(2.7) more explicitly, we first manipulate the average power expression in (2.7) based on (2.1) as follows:

$$\mathrm{E}\left\{\int |p(t)|^2 dt\right\} = \sum_{k=1}^{K}\sum_{l=1}^{K} \rho_{k,l}\, \mathrm{E}\left\{S_k^{(i_k)} S_l^{(i_l)}\right\} = \mathrm{E}\{H(\boldsymbol{S})\} \tag{2.8}$$

where $H(\boldsymbol{S})$ is defined as

$$H(\boldsymbol{S}) \triangleq \sum_{k=1}^{K}\sum_{l=1}^{K} \rho_{k,l} S_k^{(i_k)} S_l^{(i_l)} \ . \tag{2.9}$$

In some scenarios, *symmetric signaling* is used, that is, the amplitudes of users' signals corresponding to bit 0 and bit 1 are selected as $S_k^{(0)} = -S_k^{(1)}$ for $k = 1, \ldots, K$.[1] In that case, $\mathrm{E}\big\{S_k^{(i_k)} S_l^{(i_l)}\big\} = \mathrm{E}\big\{\big|S_k^{(1)}\big|^2\big\}$ if $k = l$ and $\mathrm{E}\big\{S_k^{(i_k)} S_l^{(i_l)}\big\} = 0$ if $k \neq l$ since information bits are equally likely. Then, $H(\boldsymbol{S})$ in (2.9) becomes $H(\boldsymbol{S}) = \sum_{k=1}^{K} \big|S_k^{(1)}\big|^2$.

Next, the average probability of error for user k, P_k, is obtained as follows (please see Appendix 2.6.1 for details):

$$\mathrm{P}_k = \mathrm{E}\{G_k(\boldsymbol{S})\} \ , \tag{2.10}$$

[1] For the example in Table 2.1, symmetric signaling is employed.

where the expectation is over the random vector $\boldsymbol{S}$ in (2.5), and $G_k(\boldsymbol{S})$ is defined as

$$G_k(\boldsymbol{S}) \triangleq \frac{1}{2^K} \sum_{m\in\{0,1\}} \sum_{\mathbf{i}_k\in\{0,1\}^{K-1}} \mathrm{P}\left\{ \left(N_k + S_k^{(m)} + \sum_{\substack{l=1\\ l\neq k}}^{K} \rho_{k,l} S_l^{(i_l)} \right) \in \Gamma_{k,1-m} \,\middle|\, \boldsymbol{S} \right\} . \tag{2.11}$$

The probabilities in (2.11) are calculated with respect to the PDF of N_k for given values of $S_k^{(i_k)}$'s, and $\mathbf{i}_k$ is defined as $\mathbf{i}_k \triangleq [i_1 \cdots i_{k-1}\ i_{k+1} \cdots i_K]$; i.e., the vector of all the bit indices except for the kth one. In (2.11), we consider fixed (given) decision rules at the receivers; that is, the decision regions, $\Gamma_{k,1-m}$'s, are independent of $p_{\boldsymbol{S}}$.

Based on (2.8) and (2.10), the optimization problem in (2.6)-(2.7) can be stated as

$$\min_{p_{\boldsymbol{S}}} \max_{k\in\{1,\dots,K\}} \mathrm{E}\{G_k(\boldsymbol{S})\} \tag{2.12}$$

$$\text{subject to } \ \mathrm{E}\left\{H(\boldsymbol{S})\right\} \leq A \ . \tag{2.13}$$

The optimization problem in (2.12)-(2.13) can be quite complex in its current form since it requires optimization over all possible PDFs for a random vector of size $2K$ (see (2.5)).[2] However, various approaches can be taken in order to provide a simpler formulation of the optimization problem. To that end, the following proposition is presented first.

Proposition 2.2.1. *Suppose G_k's are continuous functions and the elements of $\boldsymbol{S}$ take values from finite closed intervals. Then, an optimal solution to* (2.12)-(2.13) *can be expressed as*

$$p_{\boldsymbol{S}}(\mathbf{s}) = \sum_{j=1}^{K+1} \lambda_j \, \delta(\mathbf{s} - \mathbf{s}_j) \ , \tag{2.14}$$

where $\sum_{j=1}^{K+1} \lambda_j = 1$ *and* $\lambda_j \geq 0$ *for* $j = 1, \dots, K+1$.

[2]The dimension of vector $\boldsymbol{S}$ can be reduced to K if symmetric signaling is employed.

Proof: Please see Appendix 2.6.2. □

Proposition 2.2.1 states that an optimal joint signal constellation $\boldsymbol{S}$ can be represented as a discrete random variable which corresponds to a randomization of $(K+1)$ or fewer distinct signal constellations for each user. In other words, for each information bit of each user, an optimal solution can be obtained by performing randomization among up to $(K+1)$ different signal amplitudes. This is unlike the conventional case in which a fixed amplitude value is transmitted for each information bit of a user.

Another implication of Proposition 2.2.1 can be provided as follows. Since a generic formulation is considered, the set of G_k's and H corresponding to all possible joint signal constellations is not a convex set in general. Hence, the optimal solution of (2.12)-(2.13) can require randomization (time-sharing), as expressed in (2.14), in order to achieve the points on the convex hull of this set. (Please see the proof of the proposition in Appendix 2.6.2 for a mathematical statement of this observation.)

In practice, randomization of signal constellations can be performed, for example, via time-sharing by employing each signal constellation for a certain number of information bits in proportion to the probability of that constellation. A simple example was provided in the second paragraph of this section and in Table 2.1-(B). More generally, if N_{I} information bits are to be transmitted to each user, $\lambda_1 N_{\mathrm{I}}$ bits are generated according to $\mathbf{s}_1$, $\lambda_2 N_{\mathrm{I}}$ bits are generated according to $\mathbf{s}_2$, $\ldots$, and $\lambda_{K+1} N_{\mathrm{I}}$ bits are generated according to $\mathbf{s}_{K+1}$ in order to realize the PDF of the joint signal constellation in (2.14). It should be emphasized that *the receivers do not need to know this randomization structure* since the signal constellation randomization is optimized by the transmitter for *fixed (given)* detectors at the receivers of different users (see (2.4)) based on the optimization problem in (2.6)-(2.7). In particular, the average probability of error for user k, P_k, in (2.6) is given by (2.10) and (2.11), which indicate that the decision regions $\Gamma_{k,0}$ and $\Gamma_{k,1}$ (equivalently, the detector) for each user are independent of the probability distribution of the joint signal constellation, $\boldsymbol{S}$; hence, the receiver implements its detector without knowing the randomization structure.

Proposition 2.2.1 implies that it is not necessary to search over all PDFs in (2.12)-(2.13). Instead, only the PDFs in the form of (2.14) can be considered, and the problem in (2.12)-(2.13) can be reduced to

$$\min_{\{\lambda_j,\mathbf{s}_j\}_{j=1}^{K+1}} \max_{k\in\{1,\ldots,K\}} \sum_{j=1}^{K+1} \lambda_j\, G_k(\mathbf{s}_j) \tag{2.15}$$

$$\text{subject to} \quad \sum_{j=1}^{K+1} \lambda_j\, H(\mathbf{s}_j) \leq A\ , \quad \sum_{j=1}^{K+1} \lambda_j = 1\ ,\ \lambda_j \geq 0\ ,\ j = 1, \ldots, K+1\ . \tag{2.16}$$

Since this optimization problem is over a number of variables instead of functions, it provides a significant simplification over the problem in (2.12)-(2.13). However, it can still be a nonconvex optimization problem in general. The structure of the optimization problem in (2.15)-(2.16) can be utilized in order to obtain close-to-optimal solutions with low complexity. Namely, as discussed in the next subsection, a convex relaxation approach can be employed to provide an approximate solution of (2.15)-(2.16).

Remark: In order to realize the proposed approach of power control with signal constellation randomization in practice, the transmitter needs to know the noise powers at the receivers (or, the signal-to-noise ratios (SNRs) at the receivers, considering a flat-fading scenario, as discussed after (2.2)), which can be sent via feedback to the transmitter. Such a feedback is commonly available in multiuser systems for power control purposes [28]. In addition, if the randomization is implemented via time-sharing, the channel conditions should be (almost) constant for a number of bit durations; hence, slowly fading channels are well-suited for the power control with signal constellation randomization approach. □

Power Control with Constellation Randomization versus Conventional Power Control

The main difference of the proposed power control with constellation randomization approach from conventional power control algorithms is that the former is employed for each state of the channel whereas the latter is used with respect to varying channel conditions. In other words, the power control strategies in the literature adapt the power as the channel state changes, whereas the proposed approach performs constellation (power) randomization for a given (fixed) channel state. Therefore, these two approaches are different in the sense that they are employed in different scenarios. In addition, it is possible to employ these two approaches jointly: conventional power control as the channel conditions change, and power control with constellation randomization for each channel state. In such a scenario, the conventional power control strategy will determine the power that is allocated for each channel state, which in effect sets the value of A in (2.7), and the proposed approach will employ the optimal constellation randomization under the power limit based on the optimization problem in (2.6)-(2.7). Therefore, the proposed power control with constellation randomization approach is well-suited for slow fading channels, where the channel state is (almost) constant for a certain number of bit durations and then changes to a different value after a certain amount of time (i.e., block fading scenarios).

2.2.2 Approximate Solution Based on Convex Relaxation

Although the optimization problem in (2.15)-(2.16) can be solved via global optimization techniques in general, it becomes challenging for an optimization technique to achieve the global optimum as the number K of users increases.[3]

[3]Specifically, there are a total of $(2K+1)(K+1)$ unknown variables in (2.15)-(2.16) (which reduces to $(K+1)^2$ for symmetric signaling).

Therefore, it is desirable to obtain a convex version of the problem, which always converges to its global optimum. In the following, an approximate formulation of the problem is provided based on convex relaxation [49].

First, consider a set of possible joint signal constellations for $\boldsymbol{S}$ in (2.5) and denote them as $\tilde{\mathbf{s}}_1, \ldots, \tilde{\mathbf{s}}_{N_m}$. Then, the PDF of the joint signal constellation is approximately modeled as

$$p_{\boldsymbol{S}}(\mathbf{s}) \approx \sum_{j=1}^{N_m} \tilde{\lambda}_j \, \delta(\mathbf{s} - \tilde{\mathbf{s}}_j) \ , \tag{2.17}$$

where $\sum_{j=1}^{N_m} \tilde{\lambda}_j = 1$, $\tilde{\lambda}_j \geq 0$ for $j = 1, \ldots, N_m$, and $\tilde{\mathbf{s}}_1, \ldots, \tilde{\mathbf{s}}_{N_m}$ are *known* joint signal constellations. Then, the approximate version of (2.12)-(2.13) can be formulated as follows:

$$\min_{\tilde{\boldsymbol{\lambda}}} \ \max_{k \in \{1,\ldots,K\}} \ \tilde{\boldsymbol{\lambda}}^T \mathbf{g}_k \tag{2.18}$$

$$\text{subject to} \ \ \tilde{\boldsymbol{\lambda}}^T \mathbf{h} \leq A \ , \ \ \tilde{\boldsymbol{\lambda}}^T \mathbf{1} = 1 \ , \ \ \tilde{\boldsymbol{\lambda}} \geq \mathbf{0} \ , \tag{2.19}$$

where $\tilde{\boldsymbol{\lambda}} \triangleq \left[\tilde{\lambda}_1 \cdots \tilde{\lambda}_{N_m}\right]$, $\mathbf{g}_k \triangleq [G_k(\tilde{\mathbf{s}}_1) \cdots G_k(\tilde{\mathbf{s}}_{N_m})]$, $\mathbf{h} \triangleq [H(\tilde{\mathbf{s}}_1) \cdots H(\tilde{\mathbf{s}}_{N_m})]$, and $\mathbf{0}$ and $\mathbf{1}$ denote vectors of zeros and ones, respectively. In other words, instead of considering all possible PDFs as in (2.15)-(2.16), a number of known joint signal constellations are considered, and the optimal weights, $\tilde{\boldsymbol{\lambda}}$, corresponding to those joint signal constellations are searched for. In general, the solution of (2.18)-(2.19) provides an approximation to the optimal solution that is obtained from (2.15)-(2.16). The approximation accuracy can be improved by increasing N_m, i.e., by considering a larger number of elements in the set of possible signal values, $\tilde{\mathbf{s}}_1, \ldots, \tilde{\mathbf{s}}_{N_m}$, in (2.17). (In effect, for a larger N_m, the optimization in (2.18)-(2.19) is performed based on a discrete random variable with a larger number of point masses. If these point masses are selected appropriately, a larger N_m results in an error rate that is never higher than that for a smaller N_m.) In addition, if $\tilde{\mathbf{s}}_1, \ldots, \tilde{\mathbf{s}}_{N_m}$ contain all the possible joint signal constellations (e.g., for a digital system), then the solution of (2.18)-(2.19) becomes exact.

By defining an auxiliary variable t, an equivalent form of (2.18)-(2.19) can be obtained as follows:

$$\min_{t,\tilde{\boldsymbol{\lambda}}} \; t \tag{2.20}$$

$$\text{subject to} \;\; \tilde{\boldsymbol{\lambda}}^T \mathbf{g}_k \leq t \;, \;\; k = 1, \ldots, K \tag{2.21}$$

$$\tilde{\boldsymbol{\lambda}}^T \mathbf{h} \leq A \;, \;\; \tilde{\boldsymbol{\lambda}}^T \mathbf{1} = 1 \;, \;\; \tilde{\boldsymbol{\lambda}} \geq \mathbf{0} \;. \tag{2.22}$$

It is noted that (2.20)-(2.22) corresponds to *linearly constrained linear programming (LCLP)*. Therefore, it can be solved efficiently in polynomial time [49].

2.2.3 Optimal Selection of Fixed Signal Constellations as a Special Case of Optimal Power Control with Signal Constellation Randomization

Conventionally, a fixed signal constellation is employed for each user in a multiuser system [21, 28]. This conventional scenario can be considered as a special case of power control with signal constellation randomization in which the PDF of $\boldsymbol{S}$ in (2.5), $p_{\boldsymbol{S}}$, is modeled as $p_{\boldsymbol{S}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{s})$. Then, the optimization problem in (2.12)-(2.13) reduces to the *optimal selection of fixed signal constellations* problem, which is expressed as

$$\min_{\mathbf{s}} \max_{k\in\{1,\ldots,K\}} G_k(\mathbf{s}) \;\; \text{subject to} \;\; H(\mathbf{s}) \leq A \;. \tag{2.23}$$

In other words, the optimal fixed signal constellations that minimize the maximum probability of error are obtained under the average power constraint. As investigated in Section 2.4, the optimal fixed signal constellations approach can result in degraded performance in certain scenarios compared to the optimal power control with signal constellation randomization However, it has lower computational complexity, which can be desirable in certain applications.

2.3 Special Case: Sign Detectors

In this section, optimal power control with signal constellation randomization is studied in detail for symmetric signaling when sign detectors are employed at the receivers. In addition to the statistical characterization of the optimal solution, performance improvements that can be achieved via constellation randomization are quantified for interference limited scenarios.

Although sign detectors may not be optimal in the presence of interference [50], they facilitate simple implementation as they have low complexity and do not need any prior information about the interference. The use of sign detectors is justified also by the zero mean nature of the noise and interference (see (2.3)). It should be noted that the interference has zero mean since symmetric signaling and equally likely information bits are assumed. For these reasons, sign detectors are employed in many binary communications systems, such as in various wireless sensor network applications due to their low complexity and practicality [51].

For sign detectors, the decision rules at the receivers (see (2.4)) become

$$\hat{i}_k = \phi_k(y_k) = \begin{cases} 0\ , & y_k < 0 \\ 1\ , & y_k > 0 \end{cases} \tag{2.24}$$

for $k = 1, \ldots, K$. In the case of $y_k = 0$, the detector decides for bit 0 or bit 1 with equal probabilities. Then, for symmetric signaling (i.e., $S_k^{(1)} = -S_k^{(0)}$ for $k = 1, \ldots, K$), $G_k(\boldsymbol{S})$ in (2.11) can be expressed, after some manipulation, as

$$G_k(\boldsymbol{S}) = \frac{1}{2^{K-1}} \sum_{\mathbf{i}_k \in \{0,1\}^{K-1}} Q\left(\frac{S_k^{(1)} + \sum_{l=1, l\neq k}^{K} \rho_{k,l} S_l^{(i_l)}}{\sigma_k}\right). \tag{2.25}$$

In order to provide intuitions about the performance of constellation randomization in MAI limited scenarios, an asymptotical analysis is performed as $\sigma_k \to 0$ for $i = 1, \ldots, K$. In this case, $G_k(\boldsymbol{S})$ in (2.25) can be expressed as

$$G_k(\boldsymbol{S}) = \frac{1}{2^{K-1}} \sum_{\mathbf{i}_k \in \{0,1\}^{K-1}} u\left(-S_k^{(1)} - \sum_{l=1,\, l\neq k}^{K} \rho_{k,l} S_l^{(i_l)}\right)\ , \tag{2.26}$$

where $u(\cdot)$ represents the unit step function defined as $u(x) = 1$ for $x > 0$, $u(x) = 0.5$ for $x = 0$ and $u(x) = 0$ for $x < 0$.

First, the following corollary to Proposition 2.2.1 is presented related to the probability distribution of the optimal joint signal constellation when sign detectors are employed.

Corollary 2.3.1. *Assume that signal amplitudes take values from finite closed intervals, and $\sigma_k \to 0$ for $k = 1, \ldots, K$. Then, an optimal solution to* (2.12)-(2.13) *can be expressed, for sign detectors and symmetric signaling, as*

$$p_{\boldsymbol{S}}(\mathbf{s}) = \sum_{j=1}^{K} \lambda_j \, \delta(\mathbf{s} - \mathbf{s}_j) \ , \tag{2.27}$$

where $\sum_{j=1}^{K} \lambda_j = 1$ and $\lambda_j \geq 0$ for $j = 1, \ldots, K$.

Proof: Please see Appendix 2.6.3. □

In other words, instead of the generic solution in (2.14), which specifies a randomization among up to $(K+1)$ different signal constellations for each user, a randomization among up to K different signal constellations is sufficient in this scenario. This is mainly due to the fact that, as $\sigma_k \to 0$ for $k = 1, \ldots, K$, $G_k(\boldsymbol{S})$ in (2.26) depends only on the relative signal amplitudes, which makes the average power constraint in (2.13) ineffective (i.e., signal amplitudes can be scaled by the same positive number without affecting $G_k(\boldsymbol{S})$'s and $H(\boldsymbol{S})$ in (2.9) can be adjusted appropriately).

Next, the aim is to compare the performance of the power control with signal constellation randomization and fixed signal constellations approaches for sign detectors in the absence of noise. Assume without loss of generality that $S_k^{(1)}$'s are positive. Then, it is observed that both approaches can achieve zero probability of error if there exists a joint signal constellation $\boldsymbol{S}$ such that[4]

$$S_k^{(1)} > \sum_{l=1,\, l\neq k}^{K} |\rho_{k,l}| S_l^{(1)} \ , \quad \forall k \in \{1, \ldots, K\} \ . \tag{2.28}$$

This simple condition follows from (2.26) since it guarantees that the argument of the unit step function is negative for all bit combinations (recalling that $S_l^{(0)} = -S_l^{(1)}$ as symmetric signaling is considered). This is similar to the *no error floor* condition in classical multiuser systems [21]. (However, we still state it explicitly in order to employ it in Proposition 2.3.1 and Proposition 2.3.2 below.)

The condition in (2.28) corresponds to scenarios in which MAI is not significant and no error floor occurs due to interference. However, this condition may not be satisfied in certain cases and the MAI can be significant. For those cases, it is important to quantify the maximum amount of improvement that can be achieved via the power control with signal constellation randomization approach over the fixed signal constellations approach. Let $\mathrm{P}_{\mathrm{rnd}}$ denote the minimum value of the maximum probability of error corresponding to the optimal power control with signal constellation randomization, which is obtained as the solution of (2.12)-(2.13). In addition, let $\mathrm{P}_{\mathrm{fix}}$ denote the minimum value of the maximum probability of error for the optimal fixed signal constellations approach, which is obtained from (2.23). Then, the following proposition specifies the maximum asymptotical improvement due to signal constellation randomization.

Proposition 2.3.1. *Suppose there exist no signal amplitudes that satisfy* (2.28). *Then, for sign detectors and symmetric signaling, the maximum asymptotical improvement ratio is equal to the number of users. In other words,*

$$1 \leq \lim_{\sigma_1,\ldots,\sigma_K \to 0} \frac{\mathrm{P}_{\mathrm{fix}}}{\mathrm{P}_{\mathrm{rnd}}} \leq K \ . \tag{2.29}$$

[4] It can be assumed without loss of generality that $\boldsymbol{S}$ satisfies the power constraint in (2.13) since scaling the joint signal constellation $\boldsymbol{S}$ by any positive number does not affect the inequalities in (2.28).

Also, the maximum asymptotical improvement ratio, K, is achieved if there exist signal amplitudes such that

$$S_k^{(1)} > \sum_{l=1,\, l\neq k}^{K} |\rho_{k,l}|\, S_l^{(1)} \ , \ \forall k \in \{1,\ldots,K\} \setminus \{k^*\} \quad \textit{and} \tag{2.30}$$

$$-\,2 \min_{l\in\{1,\ldots,K\}\setminus\{k^*\}} \left\{ |\rho_{k^*,l}|\, S_l^{(1)} \right\} < S_{k^*}^{(1)} - \sum_{l=1,\, l\neq k^*}^{K} |\rho_{k^*,l}|\, S_l^{(1)} < 0 \tag{2.31}$$

for any $k^* \in \{1,\ldots,K\}$.

Proof: In order to prove the inequality in (2.29), it is first observed that $\mathrm{P_{fix}}/\mathrm{P_{rnd}} \geq 1$ is satisfied in all cases (even for finite σ_k's) since the fixed signal constellations approach is a special case of the power control with signal constellation randomization approach, as discussed in Section 2.2.3. To prove the upper bound in (2.29), consider the case in which there exist signal amplitudes that satisfy the conditions in (2.30)-(2.31).

For fixed signal constellations, the average probability of error for user k is given by $\mathrm{P}_k = G_k(\mathbf{s})$ for $k = 1,\ldots,K$ (see (2.10)). Let $\mathbf{s}_{k^*}^{(1)}$ denote a joint signal constellation that satisfies the conditions in (2.30)-(2.31) for $k^* \in \{1,\ldots,K\}$. Based on the expression for G_k in (2.26), it is obtained that $G_k\big(\mathbf{s}_{k^*}^{(1)}\big) = 0$, $\forall k \in \{1,\ldots,K\}\backslash\{k^*\}$ since the argument of the unit step function, $-S_k^{(1)} - \sum_{l=1,\, l\neq k}^{K} \rho_{k,l} S_l^{(i_l)}$, is always negative due to the conditions in (2.30).[5] On the other hand, the value of $G_{k^*}\big(\mathbf{s}_{k^*}^{(1)}\big)$ is obtained as follows. The condition in (2.31) can be expressed as

$$\sum_{l=1,\, l\neq k^*}^{K} |\rho_{k^*,l}|\, S_l^{(i_l)} - 2 \min_{l\in\{1,\ldots,K\}\setminus\{k^*\}} \left\{ |\rho_{k^*,l}|\, S_l^{(1)} \right\} < S_{k^*}^{(1)} < \sum_{l=1,\, l\neq k^*}^{K} |\rho_{k^*,l}|\, S_l^{(i_l)} \ . \tag{2.32}$$

Due to symmetric signaling, $\sum_{l=1,\, l\neq k^*}^{K} |\rho_{k^*,l}|\, S_l^{(i_l)}$ corresponds to the maximum value of $-\sum_{l=1,\, l\neq k^*}^{K} \rho_{k^*,l} S_l^{(i_l)}$ for $\mathbf{i}_{k^*} \in \{0,1\}^{K-1}$ (see (2.26)). Similarly,

[5] It is recalled that $S_l^{(1)}$'s are assumed to be positive without loss of generality and $S_l^{(0)} = -S_l^{(1)}$ due to symmetric signaling.

$\sum_{l=1,\, l\neq k^*}^{K} |\rho_{k^*,l}|\, S_l^{(i_l)} - 2 \min_{l\in\{1,\ldots,K\}\setminus\{k^*\}} \left\{ |\rho_{k^*,l}|\, S_l^{(1)} \right\}$ is equal to the second largest value of $-\sum_{l=1,\, l\neq k^*}^{K} \rho_{k^*,l} S_l^{(i_l)}$ since that value is achieved when all the $-\rho_{k^*,l}\, S_l^{(i_l)}$ terms are taken to be positive except for the one with the smallest absolute value. Therefore, under the condition in (2.32), $S_{k^*}^{(1)}$ is between the maximum and the second largest value of $-\sum_{l=1,\, l\neq k^*}^{K} \rho_{k^*,l} S_l^{(i_l)}$, which implies that the argument of the unit step function in (2.26), $-S_{k^*}^{(1)} - \sum_{l=1,\, l\neq k^*}^{K} \rho_{k^*,l} S_l^{(i_l)}$, is negative for all possible signal combinations except for one of them. Hence, the unit step function in (2.26) becomes zero for $(2^{K-1}-1)$ combinations and becomes one only for one combination, which results in $G_{k^*}\big(\mathbf{s}_{k^*}^{(1)}\big) = 1/2^{K-1}$. Overall, the maximum value of the average probability of error is given by $\max_k \mathrm{P}_k = \max_k G_k\big(\mathbf{s}_{k^*}^{(1)}\big) = 1/2^{K-1}$ for the fixed signal constellations approach when a joint signal constellation that satisfies the conditions in (2.30)-(2.31) is employed. Since it is impossible to set all G_k's to zero simultaneously due to the assumption in the proposition, $1/2^{K-1}$ presents the minimum value for the maximum average probability of error. Therefore, the solution of (2.23) is given by $\mathrm{P}_{\text{fix}} = 1/2^{K-1}$ under the conditions in (2.30)-(2.31).

For the power control with signal constellation randomization approach, the average probability of error for user k is given by $\mathrm{P}_k = \mathrm{E}\{G_k(\boldsymbol{S})\}$ for $k = 1, \ldots, K$ (see (2.10)). Due to the assumption in the proposition, there does not exist any signal amplitudes that set all G_k's to zero simultaneously. Therefore, it is impossible to set all the P_k values to zero even in the signal constellation randomization approach. However, signal constellation randomization can be used to reduce the maximum average probability of error by means of randomization/time-sharing. To explain this point, consider joint signal constellations $\mathbf{s}_{k^*}^{(1)}$ that satisfy the conditions in (2.30)-(2.31). As discussed in the previous paragraph, these vectors result in $G_k\big(\mathbf{s}_{k^*}^{(1)}\big) = 0$, $\forall k \in \{1, \ldots, K\} \setminus \{k^*\}$ and $G_{k^*}\big(\mathbf{s}_{k^*}^{(1)}\big) = 1/2^{K-1}$ for $k^* \in \{1, \ldots, K\}$. Since the aim is to minimize $\max_k \mathrm{E}\{G_k(\boldsymbol{S})\}$ over all possible PDFs for the joint signal constellation, the optimal solution is obtained by an equalizer rule [44], which sets $\mathrm{E}\{G_1(\boldsymbol{S})\} = \mathrm{E}\{G_2(\boldsymbol{S})\} = \cdots = \mathrm{E}\{G_K(\boldsymbol{S})\}$. For this equalizer rule, the optimal PDF for the joint signal constellation can be expressed as

$$p_{\boldsymbol{S}}(\mathbf{s}) = \frac{1}{K}\sum_{k^*=1}^{K}\delta\big(\mathbf{s}-\mathbf{s}_{k^*}^{(1)}\big)\ . \tag{2.33}$$

Therefore,

$$\mathrm{E}\{G_k(\boldsymbol{S})\} = \frac{1}{K}\sum_{k^*=1}^{K}G_k\big(\mathbf{s}_{k^*}^{(1)}\big) = \frac{1}{K2^{K-1}} \tag{2.34}$$

is obtained for all $k\in\{1,\dots,K\}$. Hence, $\max_k \mathrm{P}_k = \max_k \mathrm{E}\left\{G_k\left(\boldsymbol{S}\right)\right\} = 1/(K2^{K-1})$. Since it is impossible to set all $G_k\left(\mathbf{s}\right)$'s to zero for a given $\mathbf{s}$ due to the assumption in the proposition and setting $(K-1)$ of them to zero and one of them to $1/2^{K-1}$ corresponds to the optimal scenario for a given $\mathbf{s}$, the solution in (2.33) presents the optimal solution of $\min_{p_{\boldsymbol{S}}}\max_k \mathrm{P}_k$, which is equal to $1/(K2^{K-1})$. Hence, $\mathrm{P}_{\mathrm{rnd}} = 1/(K2^{K-1})$ is obtained.

Overall, an improvement ratio of $\mathrm{P}_{\mathrm{fix}}/\mathrm{P}_{\mathrm{rnd}} = K2^{K-1}/2^{K-1} = K$ is achieved under the conditions in the proposition. Finally, it is shown that K presents an upper limit on the asymptotical improvement ratio for the scenario in the proposition. To that aim, let the probability distribution of the joint signal constellation corresponding to the optimal power control with signal constellation randomization approach be expressed as in (2.27). Then, the minimum value of the maximum probability of error in the power control with signal constellation randomization approach is given by $\mathrm{P}_{\mathrm{rnd}} = \max_k\ \sum_{j=1}^{K}\lambda_j\,G_k(\mathbf{s}_j)$, where $\sum_{j=1}^{K}\lambda_j = 1$. Next, the following inequalities are obtained:

$$\mathrm{P}_{\mathrm{rnd}} = \max_k\sum_{j=1}^{K}\lambda_j\,G_k(\mathbf{s}_j) \geq \frac{1}{K}\sum_{k=1}^{K}\sum_{j=1}^{K}\lambda_j\,G_k(\mathbf{s}_j) \tag{2.35}$$

$$\geq \frac{1}{K}\sum_{j=1}^{K}\lambda_j\left(\min_{\mathbf{s}}\sum_{k=1}^{K}G_k(\mathbf{s})\right) = \frac{1}{K}\ \min_{\mathbf{s}}\sum_{k=1}^{K}G_k(\mathbf{s}) \tag{2.36}$$

$$\geq \frac{1}{K}\ \min_{\mathbf{s}}\ \max_k\ G_k(\mathbf{s}) = \frac{1}{K}\ \mathrm{P}_{\mathrm{fix}} \tag{2.37}$$

The inequalities in (2.35) and (2.37) follow from the fact that $K \max_k y_k \geq \sum_{k=1}^{K} y_k \geq \max_k y_k$ for $y_k \geq 0\ \forall k$, and the inequality in (2.36) is obtained by performing an additional minimum operation. Based on (2.35)-(2.37), $\mathrm{P}_{\text{fix}}/\mathrm{P}_{\text{rnd}} \leq K$ is obtained. $\square$

Proposition 2.3.1 states that in interference-limited scenarios, the maximum average probability of error can be reduced by a factor of up to K via signal constellation randomization. This improvement ratio is related to the result in Corollary 2.3.1, which states that a randomization among up to K joint signal constellations can be employed to reduce the maximum average probability of error compared to the fixed signal constellations case. By employing randomization among multiple different joint signal constellations, the average probabilities of error for different users can be equalized to a certain extent, which can reduce the maximum value of the average probabilities of error. In practice, the randomization operation can be implemented in the time domain via time-sharing (or in the frequency domain for multichannel systems) by employing each joint signal constellation for a certain fraction of time.

In Proposition 2.3.1, the upper and lower bounds on the asymptotical improvements that can be achieved via signal constellation randomization are presented, and the conditions under which the upper bound is achieved are specified. In the following proposition, conditions are obtained to specify when the lower bound in (2.29) is achieved; that is, when the use of signal constellation randomization does not provide any performance improvements over the use of fixed signal constellations.

Proposition 2.3.2. *Consider sign detectors and symmetric signaling, and assume that there exist no signal amplitudes that satisfy* (2.28). *In addition, define* $\mathbf{s}^*$ *as a joint signal constellation that minimizes the sum of the average error probabilities of the users. Then, if* $G_1(\mathbf{s}^*) = G_2(\mathbf{s}^*) = \cdots = G_K(\mathbf{s}^*)$, $\mathbf{s}^*$ *is a solution of the optimal power control with signal constellation randomization* *problem, and the asymptotical improvement ratio is equal to one; that is,*

$$\lim_{\sigma_1,\ldots,\sigma_K\to 0} \frac{\mathrm{P}_{\text{fix}}}{\mathrm{P}_{\text{rnd}}} = 1 \ . \tag{2.38}$$

Proof: The joint signal constellation $\mathbf{s}^*$ defined in the proposition can be expressed as

$$\mathbf{s}^* = \arg\min_{\mathbf{s}} \sum_{k=1}^{K} G_k(\mathbf{s}) \ .$$

Also, by definition, $\mathrm{P}_{\text{fix}} = \min_{\mathbf{s}} \max_{k} \ G_k(\mathbf{s})$, which can be bounded from below as follows:

$$\mathrm{P}_{\text{fix}} = \min_{\mathbf{s}} \max_{k} \ G_k(\mathbf{s}) \geq \frac{1}{K} \min_{\mathbf{s}} \sum_{k=1}^{K} G_k(\mathbf{s}) = G_1(\mathbf{s}^*) \tag{2.39}$$

where the condition in the proposition, $G_1(\mathbf{s}^*) = G_2(\mathbf{s}^*) = \cdots = G_K(\mathbf{s}^*)$, is used to obtain the last equality in (2.39). Since $\min_{\mathbf{s}} \max_{k} \ G_k(\mathbf{s})$ is lower bounded by $G_1(\mathbf{s}^*)$ as stated in (2.39) and this lower bound can be achieved for $\mathbf{s} = \mathbf{s}^*$, $\mathrm{P}_{\text{fix}} = G_1(\mathbf{s}^*)$ is obtained. Therefore, $\mathbf{s}^*$ is a solution for the optimal selection of fixed signal constellations problem, as claimed in the proposition. In addition, from (2.35) and (2.36), $\mathrm{P}_{\text{rnd}} \geq \frac{1}{K} \min_{\mathbf{s}} \sum_{k=1}^{K} G_k(\mathbf{s})$, which becomes $\mathrm{P}_{\text{rnd}} \geq G_1(\mathbf{s}^*) = \mathrm{P}_{\text{fix}}$ under the conditions in the proposition. Since $\mathrm{P}_{\text{rnd}} \leq \mathrm{P}_{\text{fix}}$ is also satisfied by definition (as the fixed signal constellations approach is a special case of power control with signal constellation randomization), $\mathrm{P}_{\text{rnd}} = \mathrm{P}_{\text{fix}}$ is obtained. $\square$

Proposition 2.3.2 implies that if a joint signal constellation that minimizes the sum of the average error probabilities of the users also equalizes those average error probabilities, then it is a solution of both the optimal selection of fixed signal constellations and the optimal power control with signal constellation randomization problems for the scenario in the proposition. In other words, the signal constellation randomization approach cannot provide any performance improvements over the fixed signal constellations approach, and the two approaches yield the same solution, namely, $p_{\mathbf{s}}(\mathbf{s}) = \delta(\mathbf{s} - \mathbf{s}^*)$.

2.4 Performance Evaluation

In this section, simulations are performed in order to compare the performance of the power control with signal constellation randomization approach against various approaches that employ fixed signal constellations. Namely, the following techniques are investigated in the simulations.

Power Control with Signal Constellation Randomization: Randomization of signal constellations is performed in an *optimal* or *suboptimal* manner based on the formulations in (2.15)-(2.16) or (2.20)-(2.22), respectively. In the following, the former approach is called *optimal randomization of signal constellations*, whereas the latter is named *constellation randomization with relaxation*. Optimal randomization of signal constellations can have prohibitive computational complexity when the number of users is high. Therefore, constellation randomization with relaxation is employed for large numbers of users in order to reduce the computational complexity.

Optimal Fixed Signal Constellations: In this case, fixed signal constellations are considered for all users, and the optimal solution is obtained from (2.23), as discussed in Section 2.2.3.

Fixed Signal Constellations at Power Limit: Instead of obtaining the optimal fixed signal constellations from (2.23), one can also consider a fixed signal constellations scheme that equalizes signal-to-interference-plus-noise ratios (SINRs) at different receivers, and utilizes all the available power at the transmitter [29]. The SINR at the receiver of user k is calculated from (2.3) as $\text{SINR}_k = \mathrm{E}\{|S_k^{(i_k)}|^2\}/(\mathrm{E}\{|\sum_{l\neq k}\rho_{k,l}S_l^{(i_l)}|^2\}+\sigma_k^2)$, which becomes $\text{SINR}_k = |S_k^{(1)}|^2/(\sum_{l\neq k}\rho_{k,l}^2|S_l^{(1)}|^2+\sigma_k^2)$ for symmetric signaling and fixed signal constellations. In the *fixed signal constellations at the power limit* approach, $S_1^{(1)},\ldots,S_k^{(K)}$ are chosen such that $\text{SINR}_1 = \cdots = \text{SINR}_K$ and $\sum_{k=1}^{K}|S_k^{(1)}|^2 = A$. Although this approach can provide very low complexity solutions, its performance is inferior to both the optimal fixed signal constellations and optimal randomization of signal constellations approaches in general, as investigated below.

In the following, the approaches proposed in this study, optimal randomization of signal constellations and constellation randomization with relaxation, are compared to the existing approaches in the literature, optimal fixed signal constellations and fixed signal constellations at power limit.

In the simulations, equally likely information bits are assumed, and symmetric signaling is considered. Also, the users employ sign detectors at the receivers, and the standard deviations of the noise at the receivers are taken to be equal, that is, $\sigma_k = \sigma$, $k = 1, \ldots, K$. In addition, as stated after (2.3), $\rho_{k,l}$'s are set to one for $k = l$; that is, $\rho_{k,k} = 1$ for $k = 1, \ldots, K$.

First, a 3-user scenario is considered, that is, $K = 3$, and the crosscorrelations between the pseudo-noise signals for different users are set to $\rho_{1,2} = 0.1$, $\rho_{1,3} = 0.2$, and $\rho_{2,3} = 0.3$. Also, the average power constraint A in (2.7) is taken as 3. In Figure 2.2, the maximum probabilities of error are plotted versus $1/\sigma^2$ for the optimal randomization of signal constellations, constellation randomization with relaxation, optimal fixed signal constellations, and fixed signal constellations at the power limit approaches. For the optimal randomization of signal constellations approach, the PSO algorithm is employed with 2000 iterations and 50 particles in order to obtain the solution of (2.15)-(2.16) (please refer to [52] for details of the PSO algorithms). For the constellation randomization with relaxation approach, the possible signal values for bit 1 are selected as 32 different amplitudes equally spaced between 0 and 1.4, and the negatives of these possible values are employed for bit 0. From the figure, it is observed that the optimal randomization of signal constellations, the constellation randomization with relaxation, and the optimal fixed signal constellations approaches have almost the same performance, and the fixed signal constellations at the power

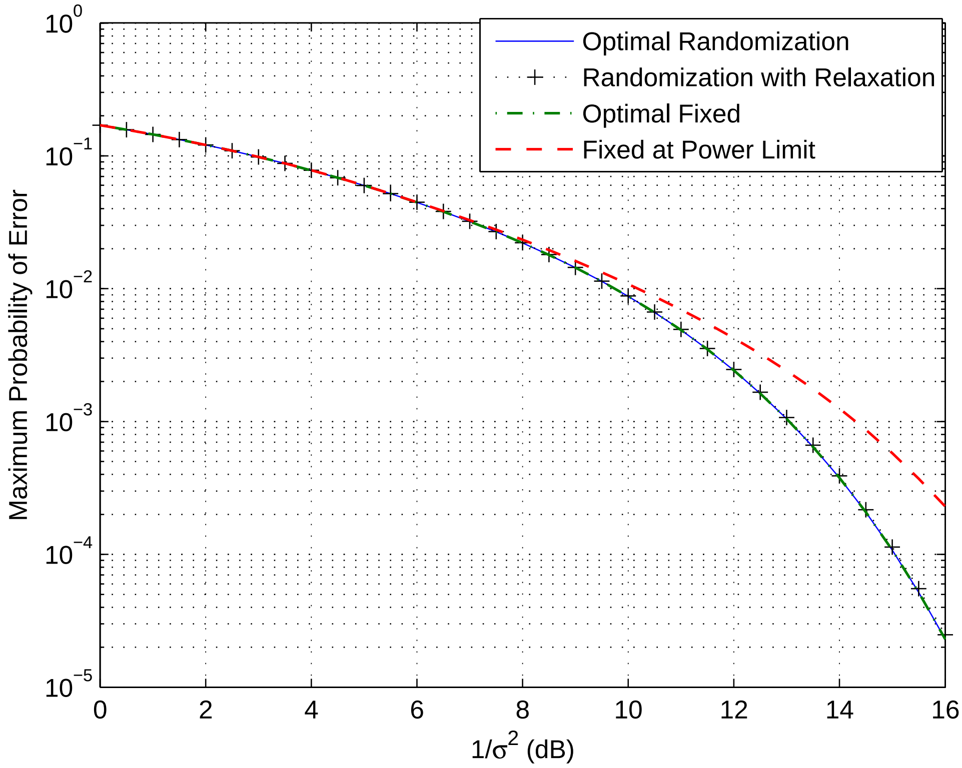

Figure 2.2: Maximum probabilities of error versus $1/\sigma^2$ for the optimal randomization of signal constellations ("Optimal Randomization"), constellation randomization with relaxation ("Randomization with Relaxation"), optimal fixed signal constellations ("Optimal Fixed"), and fixed signal constellations at the power limit ("Fixed at Power Limit") approaches, where $K = 3$, $\rho_{1,2} = 0.1$, $\rho_{1,3} = 0.2$, $\rho_{2,3} = 0.3$, and $A = 3$.

limit approach has higher maximum error probabilities for small values of σ^2, i.e., for low noise powers. On the other hand, all the approaches have similar performance in the noise limited scenarios. It is concluded that it is not optimal in general to employ fixed signal constellations that equate the SINRs of different users.

Next, a 6-user scenario is considered, that is, $K = 6$, and the crosscorrelations between the pseudo-noise signals for different users are set to 0.21; i.e., $\rho_{k,l} = 0.21$ for $k \neq l$. Also, the average power constraint A in (2.7) is taken as 6. In Figure 2.3, the maximum probabilities of error are illustrated for the constellation randomization with relaxation, optimal fixed signal constellations,

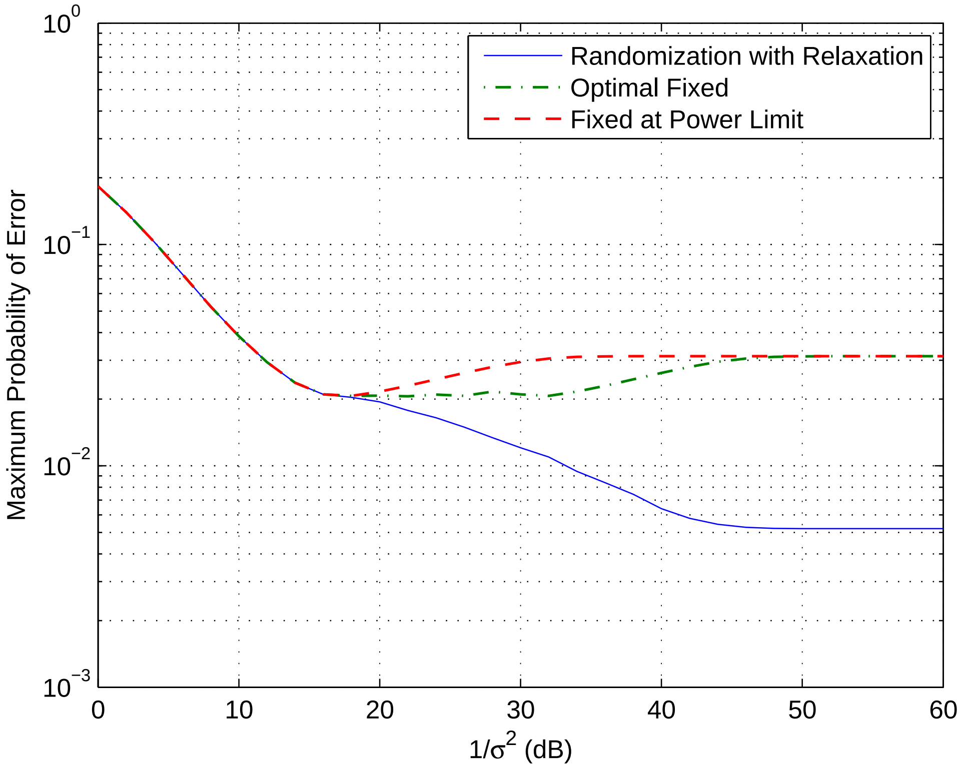

Figure 2.3: Maximum probabilities of error versus $1/\sigma^2$ for the constellation randomization with relaxation, optimal fixed signal constellations, and fixed signal constellations at the power limit approaches, where $K = 6$, $\rho_{k,l} = 0.21$ for all $k \neq l$, and $A = 6$.

and fixed signal constellations at the power limit approaches. Since the solution of (2.15)-(2.16) requires a search over a $(K+1)^2 = 49$ dimensional space, global optimization techniques may not be employed to obtain the optimal randomization of signal constellations solution in this scenario. Therefore, randomization of signal constellations is performed only via the constellation randomization with relaxation approach, which is based on (2.20)-(2.22). In obtaining the solution for this approach, the signal amplitude for information bit 1 of each user is modeled to take values from 0 to 1.4 with an increment of 0.2.[6] Then, the optimal weights for these possible signal amplitudes are obtained from (2.20)-(2.22) via CVX: Matlab Software for Disciplined Convex Programming [53]. The use of a finite set of signal amplitudes can be justified by considering a digital system

[6]Since symmetric signaling is considered, the possible signal amplitudes for bit 0 are from -1.4 to 0 with an increment of 0.2.

in which a number of bits are used to represent each signal amplitude. In this scenario, a 4-bit representation is considered as there are 8 possible signal values, $\{0, 0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4\}$, for information bit 1 and the negative of these values for information bit 0. From Figure 2.3, it is observed that the constellation randomization with relaxation approach outperforms the approaches that employ fixed signal constellations for small noise variances; that is, for MAI limited scenarios. In addition, the optimal fixed signal constellations approach achieves lower maximum probabilities of error than the fixed signal constellations at the power limit approach for medium range of σ values.[7] Another important observation from the figure is that, for small values of σ, the constellation randomization approach achieves a 6 times improvement in the maximum probability of error compared to the optimal fixed signal constellations approach, as claimed in Proposition 2.3.1. In fact, it can be shown that the assumptions in the proposition are satisfied in this scenario. Namely, there exist no signal amplitudes that satisfy (2.28), and the conditions in (2.30)-(2.31) are satisfied, for example, when all $S_k^{(1)}$'s are 1.2 except for one of them, which is equal to 0.8.

In addition, consider the same scenario as for Figure 2.3, but assume that $\rho_{k,l} = 0.15$ for $k \neq l$. In this case, the conditions in (2.28) are satisfied. Therefore, no error floors are expected and the MAI does not become a limiting factor. The error performances are illustrated in Figure 2.4 for this scenario. It is observed that the maximum probabilities of error decrease towards zero as the noise variance is reduced, and all the algorithms have almost the same error performance. As another example, the results in Figure 2.5 are presented when $\rho_{k,l} = 0.25$ for $k \neq l$. In this case, since the crosscorrelation is high, the MAI is very effective and very high error probabilities are encountered. Also, it can be shown that the

[7]It is also observed that the error probabilities of the approaches that employ fixed constellations can increase in some cases even when the noise variance decreases. This is mainly because of the multi-modal nature of the overall noise, which is the sum of zero-mean Gaussian background noise and MAI. Please see [5] for a detailed discussion.

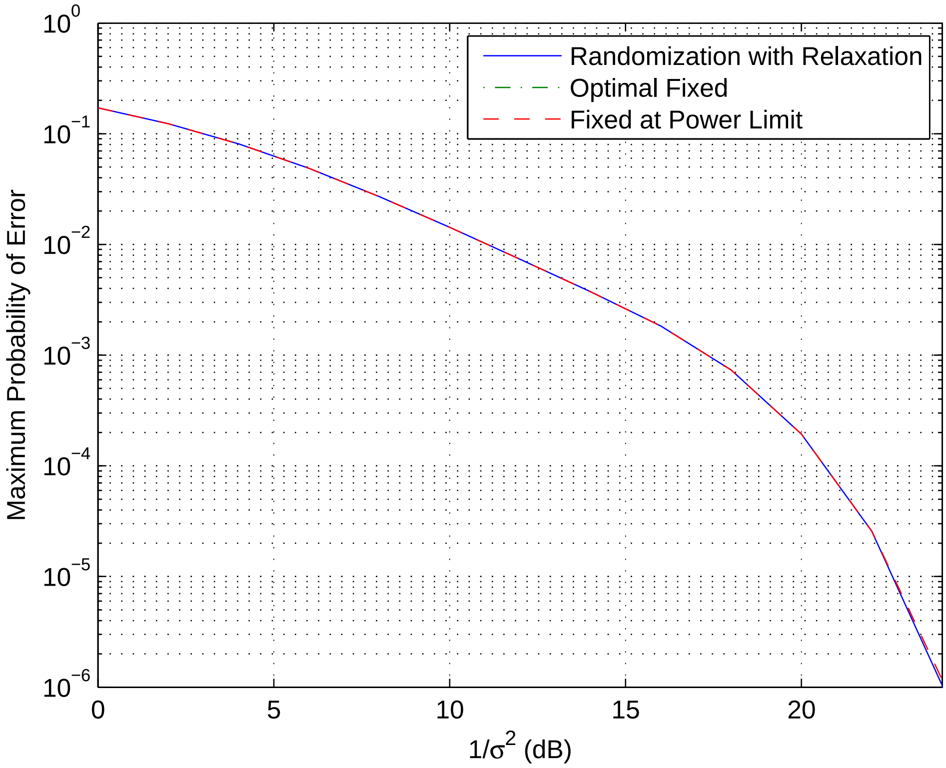

Figure 2.4: Maximum probabilities of error versus $1/\sigma^2$ for the constellation randomization with relaxation, optimal fixed signal constellations, and fixed signal constellations at the power limit approaches, where $K = 6$, $\rho_{k,l} = 0.15$ for all $k \neq l$, and $A = 6$.

conditions in (2.28) and those in (2.30)-(2.31) are not satisfied for this scenario. In Figure 2.5, the constellation randomization approach provides improvements over the approaches with fixed signal constellations, which have the same performance. However, the improvement ratio is smaller than 6 in this scenario, which is about 1.4 at low σ values.

In Figure 2.6, the error probabilities of the different approaches are plotted versus ρ, where $\rho_{k,l} = \rho$ for $k \neq l$. In addition, the other parameters are set to $A = 6$, $K = 6$, and $\sigma = 10^{-3}$. It is observed that the constellation randomization approach has lower error probabilities than the other approaches for $\rho \in [0.2, 0.29]$ and $\rho \in [0.33, 0.57]$. The improvement region and the amount of

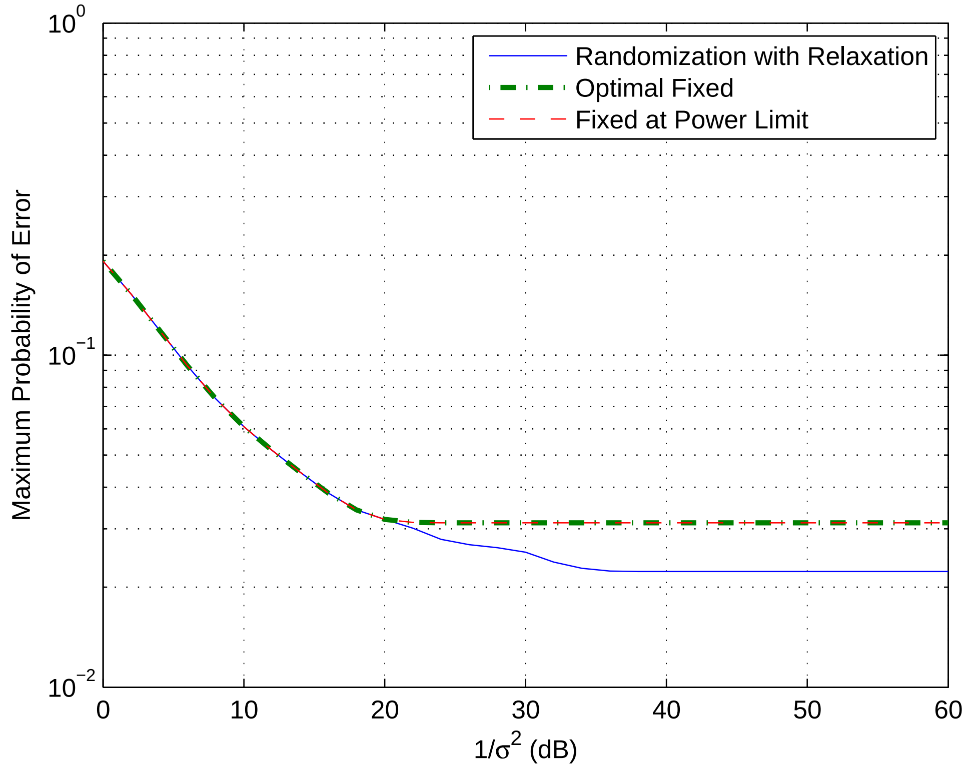

Figure 2.5: Maximum probabilities of error versus $1/\sigma^2$ for the constellation randomization with relaxation, optimal fixed signal constellations, and fixed signal constellations at the power limit approaches, where $K = 6$, $\rho_{k,l} = 0.25$ for all $k \neq l$, and $A = 6$.

improvement depend on the relation among the system parameters. For example, as investigated in Section 2.3, an improvement ratio of K is achieved for $\rho \in [0.2, 0.215]$ (which can be obtained from the conditions in (2.30)-(2.31)), and lower improvement ratios are observed in other regions. Also, the optimal fixed signal constellations approach outperforms the fixed signal constellations at the power limit approach for certain range of ρ values. However, it does not provide significant improvements in general.

In order to compare the error performance of the three approaches for different numbers of users, Figure 2.7 is presented, where $A = 6$, $\sigma = 10^{-3}$, and $\rho_{k,l} = 0.35$ for $k \neq l$. It is observed that the constellation randomization with relaxation approach provides improvements over the approaches that employ fixed signal

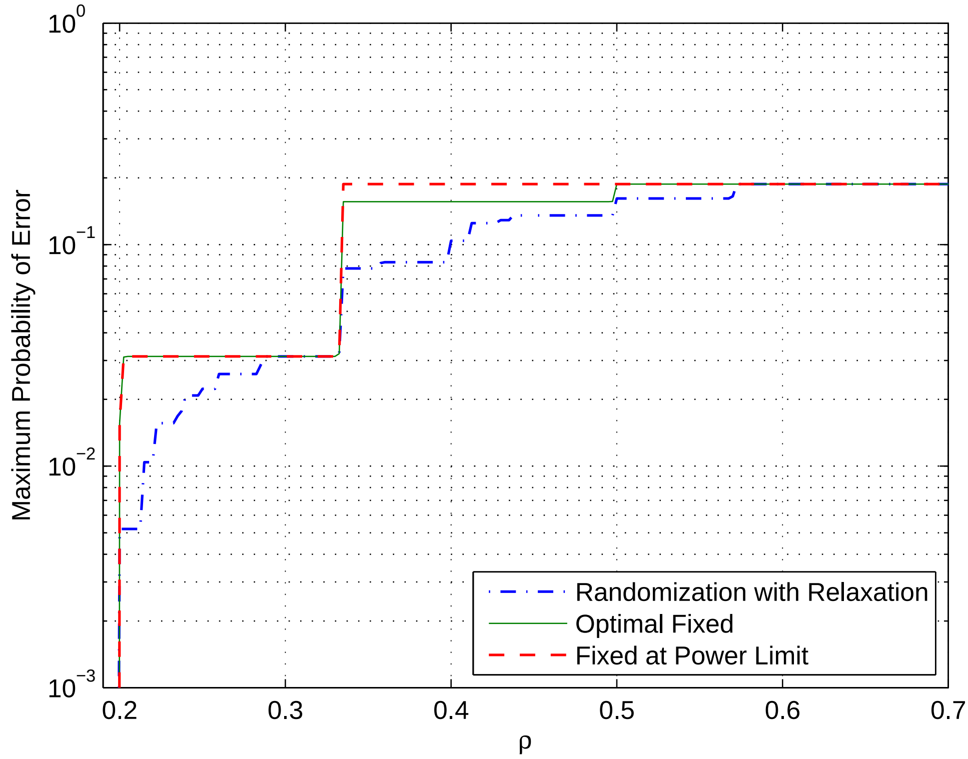

Figure 2.6: Maximum probabilities of error versus ρ for the constellation randomization with relaxation, optimal fixed signal constellations, and fixed signal constellations at the power limit approaches, where $K = 6$, $A = 6$, and $\sigma = 10^{-3}$.

constellations when the number of users is larger than three, in which case the MAI becomes a dominating factor. Also, the approaches that employ fixed signal constellations achieve similar maximum probabilities of error in most cases. In addition, their error performance is observed to be a non-monotonic function of the number of users. For example, the errors are lower for $K = 5$ than those for $K = 4$. The reason for this seemingly counterintuitive behavior can be explained from the expression in (2.25), or more simply from (2.26) since σ is sufficiently small. Considering the fixed signal constellations at the power limit approach, the signal amplitudes are set to $S_k^{(1)} = -S_k^{(0)} = \sqrt{A/K}$ for $k = 1, \ldots, K$. Since $\rho_{k,l} = 0.35$ for $k \neq l$, it can be shown for $K = 4$ and $K = 5$ that there is only one combination of the information bits of interfering users for which the argument of the unit step function in (2.26) becomes positive. Namely, when

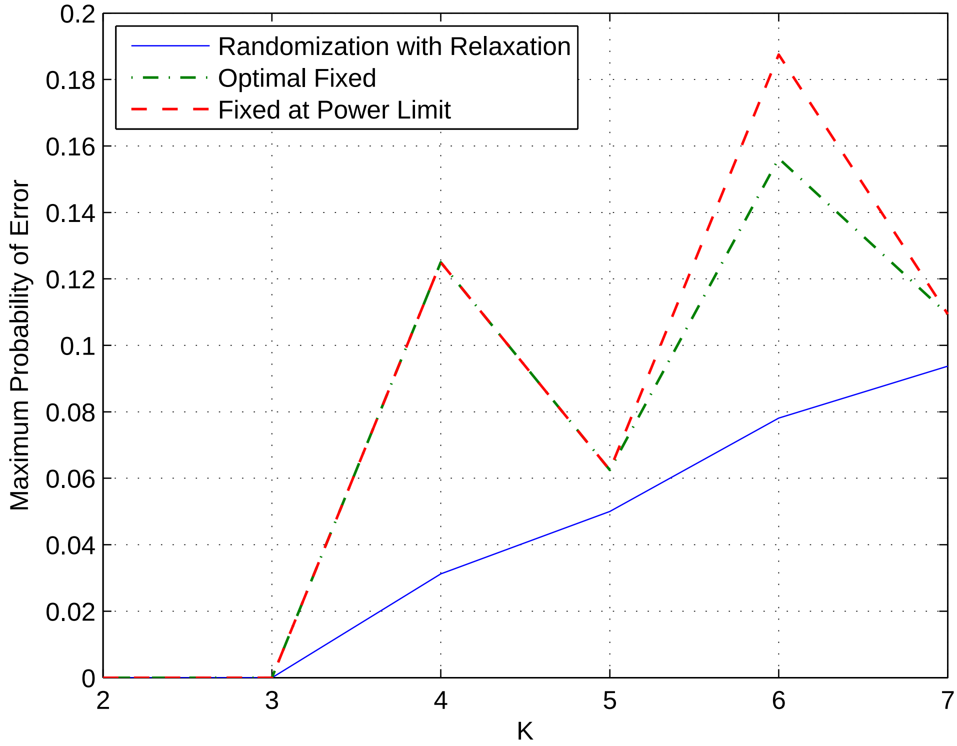

Figure 2.7: Maximum probabilities of error versus the number of users, K, for the constellation randomization with relaxation, optimal fixed signal constellations, and fixed signal constellations at the power limit approaches, where $\sigma = 10^{-3}$, $\rho_{k,l} = 0.35$ for all $k \neq l$, and $A = 6$.

all the interfering signals are $-\sqrt{A/K}$, the argument of the unit step function becomes $-\sqrt{A/K} + 0.35(K-1)\sqrt{A/K}$, which is positive for $K \geq 4$. On the other hand, when one of the interfering signals is set to $\sqrt{A/K}$, the argument becomes $-\sqrt{A/K} + 0.35(K-3)\sqrt{A/K}$, which is negative for $K \leq 5$. (Of course, the result is still negative when more than one interfering signals are set to $\sqrt{A/K}$.) Therefore, for $K = 4$ and $K = 5$, $G_k(\boldsymbol{S})$ in (2.26) is equal to $1/2^{K-1}$ for $k = 1, \ldots, K$ since the unit step function is 1 only for one combination and 0 otherwise. Hence, the maximum probability of error for $K = 5$, is lower than that for $K = 4$, as observed in Figure 2.7. However, for $K = 6$, there are multiple combinations of interfering signals for which the unit step function in (2.26) equal to one. Therefore, larger errors are observed in that case.

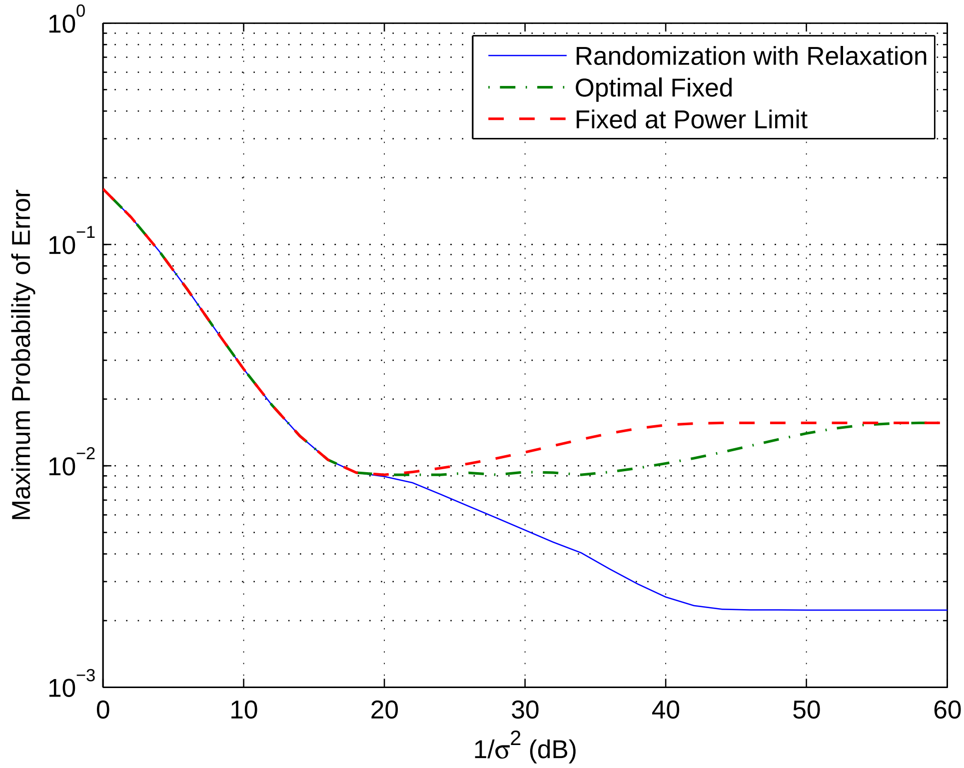

Figure 2.8: Maximum probabilities of error versus $1/\sigma^2$ for the constellation randomization with relaxation, optimal fixed signal constellations, and fixed signal constellations at the power limit approaches, where $K = 7$, $\rho_{k,l} = 0.17$ for all $k \neq l$, and $A = 7$.

Finally, a scenario with $K = 7$ users is considered, where $\rho_{k,l} = 0.17$ for $k \neq l$, and $A = 7$. In Figure 2.8, the maximum probabilities of error are illustrated for the constellation randomization with relaxation, optimal fixed signal constellations, and fixed signal constellations at the power limit approaches. Similar observations to those for Figure 2.3 can be made. In particular, it is observed that an improvement ratio of 7 is achieved at low noise variances; that is, the maximum probability of error is reduced by 7 times via the randomization of signal constellations, as claimed in Proposition 2.3.1.

The main observations from the simulation results can be summarized as follows: (i) Signal constellation randomization can provide performance improvements over the approaches that employ fixed signal constellations and the amount of improvement depends mainly on the noise level, the number of users, and the crosscorrelations between the pseudo-noise signals. (ii) The worst-case error rate of the optimal fixed signal constellations approach can be reduced by up to K times via the optimal randomization of signal constellation approach. (iii) The fixed signal constellation approach that equalizes the SINRs of the users and utilizes all the available power has the worst performance among all the considered approaches.

2.5 Concluding Remarks and Extensions

The optimal power control with signal constellation randomization has been proposed for the downlink of a multiuser DS-CDMA system. After presenting a formulation for the optimal power control with signal constellation randomization problem, it has been shown that an optimal joint signal constellation can be obtained by a randomization of $(K+1)$ or fewer distinct joint signal constellations, where K denotes the number of users. In addition to the original nonconvex formulation, an approximate solution based on convex relaxation has been obtained. Then, detailed performance analysis has been performed when the receivers employ symmetric signaling and sign detectors. Specifically, the maximum asymptotical improvement ratio has been shown to be equal to the number of users, and the conditions under which the maximum and minimum asymptotical improvement ratios are achieved have been derived. Numerical examples have been presented to investigate the theoretical results.

Although the problem formulation is based on the minimax approach in (2.6), the results in this study can directly be extended to cover cases in which the users have different levels of importance. In that case, the expression in (2.6) can be replaced with $\min_{\boldsymbol{p_S}} \max_{k\in\{1,\dots,K\}} w_k \mathrm{P}_k$, where w_k's are non-negative weighting factors that are set according to the importance levels. Then, the definition of G_k in

(2.11) can be updated by multiplying the expression by w_k, and all the theoretical results in the remaining parts can be extended accordingly.

Finally, the theoretical approach employed for the binary multiuser systems in this work can also be utilized for $M-$ary systems with $M > 2$. In that case, the definitions of the joint signal constellation in (2.5), and the auxiliary functions in (2.9) and (2.11) should be updated. Then, the results in Section 2.2 can be extended to M-ary systems as well.

2.6 Appendices

2.6.1 Derivation of (2.10)

For the generic decision rule in (2.4), the average probability of error for user k can be expressed as $\mathrm{P}_k = 0.5\,\mathrm{P}\{Y_k \in \Gamma_{k,0}\,|\,i_k = 1\} + 0.5\,\mathrm{P}\{Y_k \in \Gamma_{k,1}\,|\,i_k = 0\}$, which, based on (2.3), becomes

$$\mathrm{P}_k = 0.5\,\mathrm{P}\left\{S_k^{(1)} + \sum_{\substack{l=1\\l\neq k}}^{K}\rho_{k,l}S_l^{(i_l)} + N_k \in \Gamma_{k,0}\right\} + 0.5\,\mathrm{P}\left\{S_k^{(0)} + \sum_{\substack{l=1\\l\neq k}}^{K}\rho_{k,l}S_l^{(i_l)} + N_k \in \Gamma_{k,1}\right\}. \tag{2.40}$$

Since bits are equally likely, (2.40) can be expressed, by defining $\mathbf{i}_k \triangleq [i_1 \cdots i_{k-1}\ i_{k+1} \cdots i_K]$, as

$$\mathrm{P}_k = \frac{1}{2^K}\sum_{m\in\{0,1\}}\ \sum_{\mathbf{i}_k\in\{0,1\}^{K-1}} \mathrm{P}\left\{S_k^{(m)} + \sum_{\substack{l=1\\l\neq k}}^{K}\rho_{k,l}S_l^{(i_l)} + N_k \in \Gamma_{k,1-m}\right\} \tag{2.41}$$

In the signal constellation randomization approach, $S_k^{(i_k)}$'s are random variables. Hence, the probability expression in (2.41) can be calculated by first conditioning on given values of $S_k^{(i_k)}$'s and then taking the expectation with respect to the PDF of $\boldsymbol{S}$; that is,

$$\mathrm{P}\left\{S_k^{(m)} + \sum_{\substack{l=1\\l\neq k}}^{K} \rho_{k,l} S_l^{(i_l)} + N_k \in \Gamma_{k,1-m}\right\} = \mathrm{E}\left\{\mathrm{P}\left\{N_k + S_k^{(m)} + \sum_{\substack{l=1\\l\neq k}}^{K} \rho_{k,l} S_l^{(i_l)} \in \Gamma_{k,1-m} \,\middle|\, \boldsymbol{S}\right\}\right\}. \tag{2.42}$$

It is noted that the probability in (2.42) is calculated according to the PDF of N_k. By defining the expression inside the expectation in (2.42) as $G_k(\boldsymbol{S})$, (2.10) and (2.11) are obtained from (2.41) and (2.42). □

2.6.2 Proof of Proposition 2.2.1

The proof can be obtained based on Carathéodory's theorem [54, 55] similarly to the proofs in [4], [9], [12]. First, define the following set: $U = \{(u_0, u_1, \ldots, u_K) : u_0 = H(\mathbf{s}),\, u_1 = G_1(\mathbf{s}), \ldots, u_K = G_K(\mathbf{s}) \text{ for } \mathbf{s} \in \mathcal{S}\}$, where $\mathcal{S} \triangleq [\mathrm{s}_{\min}, \mathrm{s}_{\max}]^{2K}$, with $\mathrm{s}_{\min}$ and $\mathrm{s}_{\max}$ denoting the minimum and maximum signal amplitude values, respectively. Since the functions are continuous and $\mathcal{S}$ is a closed set, U is closed and bounded; hence, it is a compact set. Therefore, the convex hull of U, denoted by V, is a closed subset of $\mathbb{R}^{K+1}$ [56]. Next, define set W as follows: $W = \big\{(w_0, w_1, \ldots, w_K) : w_0 = \mathrm{E}\{H(\boldsymbol{S})\},\, w_1 = \mathrm{E}\{G_1(\boldsymbol{S})\}, \ldots, w_K = \mathrm{E}\{G_K(\boldsymbol{S})\},\ \forall p_{\boldsymbol{S}}(\mathbf{s}),\, \mathbf{s} \in \mathcal{S}\big\}$. Similar arguments as in [4], [9, 12, 57] can be used to conclude that set W is equal to the convex hull of U; that is, $W = V$. Therefore, due to Carathéodory's theorem [54, 55], any point in V (equivalently, in W) can be expressed as the convex combination of $(K+2)$ or fewer points in U since the dimension of U is smaller than or equal to $(K+1)$. Since the optimization problem in (2.12)-(2.13) aims to minimize the maximum of $\mathrm{E}\{G_k(\boldsymbol{S})\}$'s, the optimal solution must correspond to the boundary of W. (Note that W contains its boundary as it is a closed set.) Since any point at the boundary of W can be expressed as the convex combination of at most $(K+1)$ elements in U [54], an optimal PDF can be represented as in (2.14). $\square$

2.6.3 Proof of Corollary 2.3.1

As $\sigma_k \to 0$ for $k = 1, \ldots, K$, $G_k(\boldsymbol{S})$'s are expressed as in (2.26). Due to the unit step function in (2.26), scaling a joint signal constellation by a positive value does not affect the probabilities of error; that is, $G_k(\mathbf{s}) = G_k(c\,\mathbf{s})$ for all $c > 0$. Therefore, for each $\mathbf{s}$, there exists a positive constant c for which G_k's are unchanged but $H(c\,\mathbf{s}) = c^2 H(\mathbf{s}) \leq A$ (see (2.9)). Hence, the average power constraint in (2.13) becomes ineffective in this scenario. Therefore, the proof of Proposition 2.2.1 in Appendix 2.6.2 can be applied in this case by redefining sets U and W as $U = \{(u_0, u_1, \ldots, u_{K-1}) \,:\, u_0 = G_1(\mathbf{s}), \ldots, u_{K-1} = G_K(\mathbf{s}) \text{ for } \mathbf{s} \in \mathcal{S}\}$ and $W = \{(w_0, w_1, \ldots, w_{K-1}) \,:\, w_0 = \mathrm{E}\{G_1(\boldsymbol{S})\}, \ldots, w_{K-1} = \mathrm{E}\{G_K(\boldsymbol{S})\},\ \forall p_{\boldsymbol{S}}(\mathbf{s}),\, \mathbf{s} \in \mathcal{S}\}$, respectively. Since the dimension of W reduces to K in this case, the optimal PDF can be obtained as in (2.27) in this scenario based on similar arguments to those in Appendix 2.6.2. $\square$

Chapter 3

Optimal Detector Randomization for Multiuser Communications Systems

This chapter is organized as follows. In Section 3.1, the system model is introduced and receiver structures are described. In Section 3.2, the optimal detector randomization problem is formulated, and a low-complexity approach is presented. Analysis of optimal detector randomization is performed in Section 3.3, and lower bounds and upper bounds are obtained on the performance of optimal detector randomization. In addition, various conditions for improvability or nonimprovability via detector randomization are derived, and simple solution is provided for equal crosscorrelations and noise powers. Numerical examples are presented in Section 3.4. In Section 3.5, concluding remarks are made and possible extensions to uplink scenarios and M-ary systems are discussed.

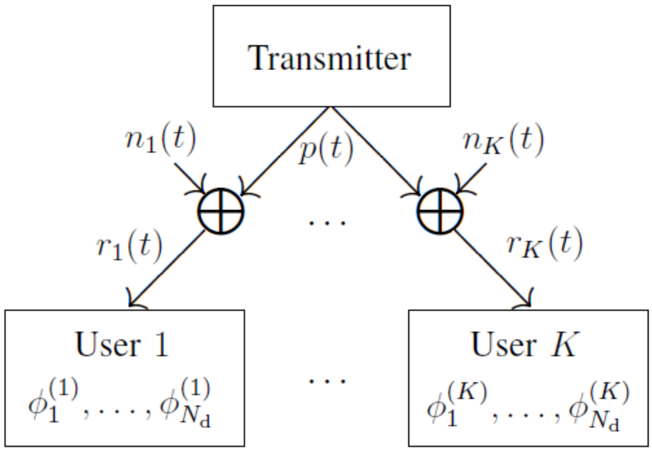

Figure 3.1: System model. The transmitter sends information bearing signals to K users over additive noise channels, and each user estimates the transmitted symbol by performing detector randomization among N_d detectors.

3.1 System Model

Consider the downlink of a multiuser communications system in which the transmitter (e.g., base station or access point) sends information bearing signals to K users simultaneously via code division multiple access (CDMA). In addition, assume that the users can perform detector randomization [6, 7] in coordination with the transmitter by employing different detectors for certain fractions of time. In particular, suppose that each user can time-share (randomize) among N_d detectors; namely, user k employs detector $\phi_1^{(k)}$ for the first $N_{\mathrm{s},1}$ symbols, detector $\phi_2^{(k)}$ for the next $N_{\mathrm{s},2}$ symbols, … , and detector $\phi_{N_\mathrm{d}}^{(k)}$ for the last $N_{\mathrm{s},N_\mathrm{d}}$ symbols[1], where $k \in \{1, 2, \ldots, K\}$. The described scenario is also depicted in Figure 3.1, which illustrates a K-user system with N_d detectors for each user.

[1]Such a coordination can be achieved in practice by employing a communications protocol that informs the users about this randomization (time-sharing) structure by including the related information in the header of the communications packet [6].

For the downlink of a DSSS binary[2] communications system as in Figure 3.1, the baseband model of the transmitted signal can be expressed as

$$p(t) = \sum_{k=1}^{K} S_{k,l}^{(i_k)} c_k(t) \ , \tag{3.1}$$

for $l \in \{1, \ldots, N_{\mathrm{d}}\}$ and $i_k \in \{0, 1\}$, where K is the number of users, $S_{k,l}^{(i_k)}$ denotes the transmitted signal amplitude for information bit i_k that is intended for detector l of user k, and $c_k(t)$ is the real pseudo-noise signal for user k. Pseudo-noise signals are employed to spread the spectra of users' signals and provide multiple-access capability [21]. It is assumed that the prior probabilities of bit 0 and bit 1 are equal to 0.5 for all users, and that the information bits of different users are independent.

The signal in (3.1) is transmitted to K users over the additive noise channels as in Figure 3.1, and the received signal at user k is modeled as

$$r_k(t) = \sum_{j=1}^{K} S_{j,l}^{(i_j)} c_j(t) + n_k(t) \ , \tag{3.2}$$

for $k = 1, \ldots, K$, where $n_k(t)$ is the noise at the receiver of user k, which is a zero-mean white Gaussian process with spectral density σ_k^2. The noise processes at different receivers are supposed to be independent. Although a simple additive noise model is employed in (3.2), multipath channels with slow frequency-flat fading can also be incorporated into the model under the assumption of perfect channel estimation by adjusting the average powers of the noise components in (3.2), equivalently, the σ_k^2 terms, accordingly [4].

The receiver structure for user k is illustrated in Figure 3.2. The received signal $r_k(t)$ in (3.2) is first correlated with the pseudo-noise signal for user k, $c_k(t)$. Then, the correlator output is processed by one of the detectors according to the detector randomization strategy and the transmitted bit of user k is estimated. (Although N_{d} detectors are shown in Figure 3.2, the receiver can also be

[2] As mentioned in Section 3.5, the results can be extended to M-ary communications systems as well.

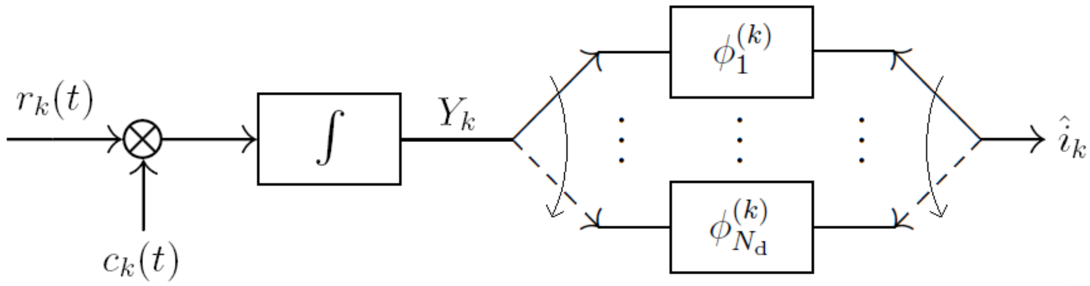

Figure 3.2: Receiver structure for user k. The received signal is first despread by the pseudo-noise signal, and the resulting signal, Y_k, is processed by one of the detectors according to a detector randomization strategy.

implemented by adapting the parameters of one detector over time.) From (3.2) and Figure 3.2, the correlator output for user k, Y_k, can be expressed as

$$Y_k = S_{k,l}^{(i_k)} + \sum_{\substack{j=1 \\ j\neq k}}^{K} \rho_{k,j} S_{j,l}^{(i_j)} + N_k \ , \tag{3.3}$$

for $k = 1, \ldots, K$, where $\rho_{k,j} \triangleq \int c_k(t)c_j(t)dt$ denotes the crosscorrelation between the pseudo-noise signals for user k and j (it is assumed that $\rho_{k,k} = 1$ for $k = 1, \ldots, K$), and $N_k \triangleq \int n_k(t)c_k(t)dt$ is the noise component. The noise components $N_1, \ldots, N_K$ form a sequence of independent zero-mean Gaussian random variables with variances, $\sigma_1^2, \ldots, \sigma_K^2$, respectively. It is noted from the expression for Y_k in (3.3) that the first term corresponds to the desired signal component, the second term denotes the multiple-access interference (MAI), and the last term is the noise component.

As shown in Figure 3.2, the correlator output in (3.3) is processed by detectors $\phi_1^{(k)}, \ldots, \phi_{N_\mathrm{d}}^{(k)}$ according to a detector randomization strategy, and an estimate of the transmitted information bit, $\hat{i}_k$, is generated. Mathematically, for a given correlator output $Y_k = y_k$, the bit estimate is obtained as

$$\hat{i}_k = \phi_l^{(k)}(y_k) = \begin{cases} 1 \ , & \text{if } y_k \in \Gamma_l^{(k)} \\ 0 \ , & \text{otherwise} \end{cases} \tag{3.4}$$

if the l^{th} detector is employed for user k, where $l \in \{1, \ldots, N_\mathrm{d}\}$ and $k \in \{1, \ldots, K\}$. In (3.4), $\Gamma_l^{(k)}$ denotes the decision region in which bit 1 is selected by the l^{th} detector of user k. The receiver of user k can perform randomization among these N_{d} detectors in order to optimize the error performance. Let v_l denote the randomization (or time-sharing) factor for detector $\phi_l^{(k)}$, where $\sum_{l=1}^{N_{\mathrm{d}}} v_l = 1$ and $v_l \geq 0$ for $l = 1, \ldots, N_{\mathrm{d}}$. In other words, user k employs detector $\phi_l^{(k)}$ for $100v_l$ percent of the time, where $l \in \{1, \ldots, N_{\mathrm{d}}\}$ and $k \in \{1, \ldots, K\}$.[3] It should be noted that employing the same randomization factors for all users does not cause any loss of generality since the cases in which different randomization factors are used for different users can be covered by the preceding formulation by considering an updated value of N_{d} with corresponding detectors and randomization factors.

3.2 Optimal Detector Randomization

The aim in this study is to jointly optimize the randomization factors, the detectors (decision regions), and the transmitted signal amplitudes for all the users under an average power constraint. In order to formulate this generic problem, we first define the following signal vector $\boldsymbol{S}_l$ that consists of the signal amplitudes intended for detector l for bit 0 and bit 1 of all users: $\boldsymbol{S}_l = \left[S_{1,l}^{(0)}\; S_{1,l}^{(1)} \cdots S_{K,l}^{(0)}\; S_{K,l}^{(1)}\right]$. In addition, let $\boldsymbol{\phi}_l$ denote the set of the l^{th} detectors of the users, which is defined as $\boldsymbol{\phi}_l = \left[\phi_l^{(1)} \cdots \phi_l^{(K)}\right]$ for $l \in \{1, \ldots, N_{\mathrm{d}}\}$. For a randomization strategy specified by randomization factors $\{v_1, \ldots, v_{N_{\mathrm{d}}}\}$ (as described in the previous paragraph), the system in Figure 3.1 operates as follows: For v_l fraction of the time, the transmitter sends the signal vector $\boldsymbol{S}_l$ and the users employ the corresponding detectors in $\boldsymbol{\phi}_l$ for $l = 1, \ldots, N_{\mathrm{d}}$. Therefore, the aim is to obtain the optimal set $\{v_l, \boldsymbol{\phi}_l, \boldsymbol{S}_l\}_{l=1}^{N_{\mathrm{d}}}$ that optimizes the error performance of the system under an

[3] It is assumed that statistics of channel noise do not change during this randomization (time-sharing) operation. Therefore, the detector randomization approach is well-suited for block fading channels, where detector randomization can be performed for each channel realization [58].

average power constraint. Specifically, the following optimization problem is proposed:

$$\min_{\{v_l,\boldsymbol{\phi}_l,\boldsymbol{S}_l\}_{l=1}^{N_\mathrm{d}}} \max_{k\in\{1,\ldots,K\}} \mathrm{P}_k \tag{3.5}$$

$$\text{subject to } \mathrm{E}\left\{\int |p(t)|^2 dt\right\} \leq A \tag{3.6}$$

where P_k is the average probability of error for user k, A specifies an average power constraint, and $p(t)$ is as in (3.1). The minimax approach is adopted for fairness [45–48] by preventing scenarios in which the average probabilities of error are very low for some users whereas they are (unacceptably) high for others.[4]

The constraint in (3.6) is defined in such a way that the average power is limited in each bit duration. In other words, the expectation operation in (3.6) is over the equiprobable information bits of the users. Hence, from (3.1), (3.6) can be expressed as

$$\sum_{k=1}^{K}\sum_{j=1}^{K} \rho_{k,j}\, \mathrm{E}\left\{S_{k,l}^{(i_k)} S_{j,l}^{(i_j)}\right\} \leq A\ , \tag{3.7}$$

where $\mathrm{E}\left\{S_{k,l}^{(i_k)} S_{j,l}^{(i_j)}\right\}$ is given by

$$\mathrm{E}\left\{S_{k,l}^{(i_k)} S_{j,l}^{(i_j)}\right\} = \begin{cases} 0.25 S_{k,l}^{(0)} S_{j,l}^{(0)} + 0.25 S_{k,l}^{(0)} S_{j,l}^{(1)} & \\ \quad + 0.25 S_{k,l}^{(1)} S_{j,l}^{(0)} + 0.25 S_{k,l}^{(1)} S_{j,l}^{(1)}, & k \neq j \\ 0.5\big|S_{k,l}^{(0)}\big|^2 + 0.5\big|S_{k,l}^{(1)}\big|^2, & k = j \end{cases} \tag{3.8}$$

for $l \in \{1,\ldots,N_\mathrm{d}\}$. If *symmetric signaling* is employed, (i.e., if signal amplitudes are selected as $S_{k,l}^{(0)} = -S_{k,l}^{(1)}$ for $k = 1,\ldots,K$ and $l = 1,\ldots,N_\mathrm{d}$), then $\mathrm{E}\{S_{k,l}^{(i_k)} S_{j,l}^{(i_j)}\} = \big|S_{k,l}^{(1)}\big|^2$ for $k = j$ and $\mathrm{E}\{S_{k,l}^{(i_k)} S_{j,l}^{(i_j)}\} = 0$ for $k \neq j$. Then, the expression in (3.7) becomes $\sum_{k=1}^{K} \big|S_{k,l}^{(1)}\big|^2 \leq A$. (We consider the generic case in this study and the results for symmetric signaling can be obtained as a special case.)

[4] It is possible to extend the results to cases in which different users have different levels of importance by multiplying each P_k with a weighting factor.

For notational simplicity in the following analysis, we define

$$h(\boldsymbol{S}_l) \triangleq \sum_{k=1}^{K}\sum_{j=1}^{K} \rho_{k,j}\,\mathrm{E}\left\{S_{k,l}^{(i_k)} S_{j,l}^{(i_j)}\right\} \tag{3.9}$$

where $\boldsymbol{S}_l$ is as defined in the first paragraph of this section. Then, the average power constraint in (3.7) (hence, in (3.6)) is given by

$$h(\boldsymbol{S}_l) \leq A \quad \text{for} \quad l \in \{1, \ldots, N_{\mathrm{d}}\}. \tag{3.10}$$

In order to calculate the average probability of error for user k, P_k, we first express, from (3.3) and (3.4), the error probability of the l^{th} detector of user k when the signal vector $\boldsymbol{S}_l$ is employed as follows:

$$\begin{aligned} &g_{k,l}(\boldsymbol{S}_l) = \\ &\frac{1}{2^K} \sum_{\mathbf{i}_k \in \{0,1\}^{K-1}} \Bigg(\mathrm{P}\Bigg\{ \Bigg(N_k + S_{k,l}^{(1)} + \sum_{\substack{j=1 \\ j \neq k}}^{K} \rho_{k,j} S_{j,l}^{(i_j)} \Bigg) \notin \Gamma_l^{(k)} \Bigg\} \\ &\qquad + \mathrm{P}\Bigg\{ \Bigg(N_k + S_{k,l}^{(0)} + \sum_{\substack{j=1 \\ j \neq k}}^{K} \rho_{k,j} S_{j,l}^{(i_j)} \Bigg) \in \Gamma_l^{(k)} \Bigg\} \Bigg), \end{aligned} \tag{3.11}$$

with $\mathbf{i}_k \triangleq [i_1 \cdots i_{k-1}\; i_{k+1} \cdots i_K]$ (the vector of all the bit indices except for the k^{th} one), and $\Gamma_l^{(k)}$ denoting the decision region of the l^{th} detector of user k for information symbol 1; that is, $\phi_l^{(k)}$, as specified in (3.4). In (3.11), the probabilities are with respect to the distribution of the noise component N_k for a given value of $\boldsymbol{S}_l$. Also, it should be noted that the decision region $\Gamma_l^{(k)}$ can be a function of $\boldsymbol{S}_l$ in general due to the joint optimization in (3.5) and (3.6).

Since $g_{k,l}(\boldsymbol{S}_l)$ in (3.11) denotes the error probability of the l^{th} detector of user k when signal vector $\boldsymbol{S}_l$ is employed, the average probability of user k for a randomization strategy that employs signal vector $\boldsymbol{S}_l$ and detectors $\boldsymbol{\phi}_l$ with probability v_l for $l = 1, \ldots, N_{\mathrm{d}}$ can be expressed as

$$\mathrm{P}_k = \sum_{l=1}^{N_{\mathrm{d}}} v_l\, g_{k,l}(\boldsymbol{S}_l)\ . \tag{3.12}$$

From (3.10) and (3.12), the optimization problem in (3.5) and (3.6) can be stated as

$$\min_{\{v_l,\boldsymbol{\phi}_l,\boldsymbol{S}_l\}_{l=1}^{N_\mathrm{d}}} \max_{k\in\{1,\ldots,K\}} \sum_{l=1}^{N_\mathrm{d}} v_l\, g_{k,l}(\boldsymbol{S}_l) \tag{3.13}$$

$$\text{subject to } \; h(\boldsymbol{S}_l) \leq A\,,\ \forall\, l\in\{1,\ldots,N_\mathrm{d}\} \tag{3.14}$$

$$\sum_{l=1}^{N_\mathrm{d}} v_l = 1\,,\ v_l \geq 0\,,\ \forall\, l\in\{1,\ldots,N_\mathrm{d}\}\,. \tag{3.15}$$

This problem is very challenging in general since it requires joint optimization of the signal amplitudes, the detectors, and the detector randomization factors. However, a significant simplification can be achieved based on the following proposition:

Proposition 3.2.1. *The optimization problem in* (3.13)-(3.15) *can be expressed as*

$$\min_{\{v_l,\boldsymbol{S}_l\}_{l=1}^{N_\mathrm{d}}} \max_{k\in\{1,\ldots,K\}} \sum_{l=1}^{N_\mathrm{d}} \frac{v_l}{2} \int_{-\infty}^{\infty} \min\left\{p_0^{(k)}(y|\boldsymbol{S}_l), p_1^{(k)}(y|\boldsymbol{S}_l)\right\} dy \tag{3.16}$$

$$\textit{subject to } h(\boldsymbol{S}_l) \leq A\,,\ \forall\, l\in\{1,\ldots,N_\mathrm{d}\} \tag{3.17}$$

$$\sum_{l=1}^{N_\mathrm{d}} v_l = 1\,,\ v_l \geq 0\,,\ \forall\, l\in\{1,\ldots,N_\mathrm{d}\} \tag{3.18}$$

where $p_{i_k}^{(k)}(y|\boldsymbol{S}_l)$ *is given by*

$$p_{i_k}^{(k)}(y|\boldsymbol{S}_l) = \frac{1}{\sigma_k\sqrt{2\pi}\,2^{K-1}} \sum_{\mathbf{i}_k\in\{0,1\}^{K-1}} \exp\left\{-\frac{1}{2\sigma_k^2}\left(y - S_{k,l}^{(i_k)} - \sum_{\substack{j=1\\ j\neq k}}^{K}\rho_{k,j}S_{j,l}^{(i_j)}\right)^2\right\} \tag{3.19}$$

for $i_k = 0,1$ *with* $\mathbf{i}_k \triangleq [i_1 \cdots i_{k-1}\; i_{k+1} \cdots i_K]$.

Proof: Consider the optimization problem in (3.13)-(3.15), where $g_{k,l}(\boldsymbol{S}_l)$ is defined as in (3.11) and represents the error probability of the l^{th} detector of user k when signal vector $\boldsymbol{S}_l$ is employed. Since the aim is to minimize $\max_{k\in\{1,\ldots,K\}} \sum_{l=1}^{N_\mathrm{d}} v_l\, g_{k,l}(\boldsymbol{S}_l)$ over all possible $\{v_l,\boldsymbol{\phi}_l,\boldsymbol{S}_l\}_{l=1}^{N_\mathrm{d}}$ under the specified constraints, optimal decision rules, $\boldsymbol{\phi}_l$, that minimize $g_{k,l}(\boldsymbol{S}_l)$ must be employed for each signal vector $\boldsymbol{S}_l$. For any signal vector, it is known that the ML detector minimizes the error probability when the information symbols are equally likely [44]. Therefore, it is concluded that the optimal solution to (3.13)-(3.15) results in the use of ML detectors at the receivers. Considering the l^{th} detector of user k, the ML decision rule can be specified as $i_k = 1$ if $p_1^{(k)}(y|\boldsymbol{S}_l) \geq p_0^{(k)}(y|\boldsymbol{S}_l)$ and $i_k = 0$ otherwise, where $p_{i_k}^{(k)}(y|\boldsymbol{S}_l)$ is the conditional probability density function (PDF) of observation Y_k when the information bit i_k is transmitted for the l^{th} detector of user k (see (3.3)). Therefore, the error probability of the ML detector can be calculated from $\frac{1}{2}\int \min\left\{p_0^{(k)}(y|\boldsymbol{S}_l), p_1^{(k)}(y|\boldsymbol{S}_l)\right\}dy$ [2], which corresponds to $g_{k,l}(\boldsymbol{S}_l)$ when the l^{th} detector of user k employs the ML decision rule. Hence, the expression in (3.16) is obtained from (3.13). (It is noted that the optimization space is reduced from $\{v_l, \boldsymbol{\phi}_l, \boldsymbol{S}_l\}_{l=1}^{N_{\text{d}}}$ to $\{v_l, \boldsymbol{S}_l\}_{l=1}^{N_{\text{d}}}$ since the error probabilities of the optimal detectors are expressed in terms of the signal vectors.) In addition, based on (3.3), $p_{i_k}^{(k)}(y|\boldsymbol{S}_l)$ can be expressed as in (3.19) considering equally likely information bits. $\square$

Based on Proposition 3.2.1, it is concluded that for the joint optimization problem in (3.13)-(3.15), where the detectors are modeled as generic ones, the joint optimal solution always results in the use of ML detectors at all the users. It is also noted that the results of Proposition 3.2.1 will be valid for any non-Gaussian PDF as well when the conditional PDF expression in (3.19) is updated accordingly.

Comparison of the optimization problems in (3.13)-(3.15) and in (3.16)-(3.18) reveals that Proposition 3.2.1 provides a significant simplification in obtaining the optimal solution as it reduces the optimization space from $\{v_l, \boldsymbol{\phi}_l, \boldsymbol{S}_l\}_{l=1}^{N_{\text{d}}}$ to $\{v_l, \boldsymbol{S}_l\}_{l=1}^{N_{\text{d}}}$. Namely, instead of searching over all possible signal amplitudes, detectors, and detector randomization factors, (3.16)-(3.18) requires a search over possible signal amplitudes and detector randomization factors. Once the optimal signal amplitudes and detector randomization factors are obtained from (3.16)-(3.18), the optimal detectors are specified by the corresponding ML decision rules. In particular, if $\{\hat{\boldsymbol{S}}_l\}_{l=1}^{N_\mathrm{d}}$ denote the optimal signal amplitudes obtained from (3.16)-(3.18), the l^{th} detector of user k outputs bit 1 if $p_1^{(k)}(y|\hat{\boldsymbol{S}}_l) \geq p_0^{(k)}(y|\hat{\boldsymbol{S}}_l)$ and bit 0 otherwise for $k \in \{1, \dots, K\}$ and $l \in \{1, \dots, N_\mathrm{d}\}$, where $p_0^{(k)}(y|\hat{\boldsymbol{S}}_l)$ and $p_1^{(k)}(y|\hat{\boldsymbol{S}}_l)$ are obtained from (3.19).

Although the formulation in (3.16)-(3.18) provides a significant simplification over that in (3.13)-(3.15), it can still have high computational complexity when the number of detectors and/or the number of users are high. In particular, it is noted from (3.16)-(3.18) that the optimal solution of the signal amplitudes and the randomization factors requires a search over a $(2K+1)N_\mathrm{d}$ dimensional space ($(K+1)N_\mathrm{d}$ dimensional space if symmetric signaling is employed). In the following proposition, it is stated that employing more than K detectors at a receiver is not needed for the optimal solution.

Proposition 3.2.2. *The optimization problem in* (3.16)-(3.18) *achieves the same minimum value as the following problem:*

$$\min_{\{v_l, \boldsymbol{S}_l\}_{l=1}^{\min\{K,N_\mathrm{d}\}}} \max_{k \in \{1,\dots,K\}} \sum_{l=1}^{\min\{K,N_\mathrm{d}\}} \frac{v_l}{2} \int_{-\infty}^{\infty} \min\left\{p_0^{(k)}(y|\boldsymbol{S}_l), p_1^{(k)}(y|\boldsymbol{S}_l)\right\} dy \tag{3.20}$$

$$\text{subject to } h(\boldsymbol{S}_l) \leq A\,, \ \forall\, l \in \{1, \dots, \min\{K, N_\mathrm{d}\}\} \tag{3.21}$$

$$\sum_{l=1}^{\min\{K,N_\mathrm{d}\}} v_l = 1\,, \ v_l \geq 0\,, \ \forall\, l \in \{1, \dots, \min\{K, N_\mathrm{d}\}\} \tag{3.22}$$

where $p_{i_k}^{(k)}(y|\boldsymbol{S}_l)$ *is as in* (3.19).

Proof: Define

$$\tilde{g}_k(\boldsymbol{S}_l) \triangleq 0.5 \int_{-\infty}^{\infty} \min\left\{p_0^{(k)}(y|\boldsymbol{S}_l), p_1^{(k)}(y|\boldsymbol{S}_l)\right\} dy \tag{3.23}$$

and express the objective function in (3.16) as $\sum_{l=1}^{N_\mathrm{d}} v_l\, \tilde{g}_k(\boldsymbol{S}_l) = \mathrm{E}\{\tilde{g}_k(\boldsymbol{S})\}$, where $\boldsymbol{S}$ is a discrete random vector that takes the value of $\boldsymbol{S}_l$ with probability v_l for $l = 1, \dots, N_\mathrm{d}$ (cf. (3.18)). Let $p_{\boldsymbol{S}}$ denote the probability mass function (PMF) of

$\boldsymbol{S}$. In addition, define $\mathcal{P}_A$ as the set of all PMFs with N_d point masses for which $p_{\boldsymbol{S}}(\boldsymbol{S}) = 0$ whenever $h(\boldsymbol{S}) > A$. Then, (3.16)-(3.18) can be expressed as

$$\min_{p_{\boldsymbol{S}} \in \mathcal{P}_A} \max_{k \in \{1,\dots,K\}} \mathrm{E}\{\tilde{g}_k(\boldsymbol{S})\} \,. \tag{3.24}$$

Optimization problems that are in similar forms to (3.24) have been studied in the literature, such as in [12] and [11]. First, the following set is defined: $U = \{(\tilde{g}_1(\boldsymbol{S}), \dots, \tilde{g}_K(\boldsymbol{S}))\,, \forall \boldsymbol{S} \in \mathcal{S}_A\}$, where $\mathcal{S}_A$ is the set of $\boldsymbol{S}$ for which $h(\boldsymbol{S}) \leq A$. Then, it can be observed that set W, defined as $W = \{(\mathrm{E}\{\tilde{g}_1(\boldsymbol{S})\}, \dots, \mathrm{E}\{\tilde{g}_K(\boldsymbol{S})\})\,, \forall p_{\boldsymbol{S}} \in \mathcal{P}_A\}$, corresponds to the convex hull of set U. Therefore, based on Carathéodory's theorem [54], any K-tuple at the boundary of set W can be obtained as the convex combination of at most K elements in U. (The boundary is considered since a minimization operation is to performed.) Hence, the optimal solution to (3.24) can be expressed in the form of a discrete random vector with at most K non-zero point masses. For this reason, if N_d is larger than K, it is sufficient to perform the search over probability distributions with K point masses. $\square$

Based on Proposition 3.2.2, it is concluded that there is no need for employing more than K detectors at a receiver in a K-user system for achieving the optimal error performance. In other words, randomization among more than K detectors cannot provide any additional performance improvements. In addition, as observed from (3.20)-(3.22), the dimension of the search space in obtaining the optimal solution is specified by $(2K+1)\min\{K, N_\mathrm{d}\}$ (by $(K+1)\min\{K, N_\mathrm{d}\}$ for symmetric signaling). It is also noted that the results of Proposition 3.2.2 will be valid for non-Gaussian PDFs as well when the conditional PDF expression in (3.19) is updated accordingly.

3.3 Analysis of Optimal Detector Randomization

In this section, we investigate the performance of the optimal detector randomization approach specified by (3.20)-(3.22), and determine scenarios in which performance improvements can be obtained over the optimal approach that does not employ any detector randomization, which is called as the *optimal single detectors* approach in the following.

The optimal single detectors approach can be considered as a special case of the detector randomization approach when there is only one detector at each receiver; that is, $N_{\rm d} = 1$. Therefore, based on (3.13)-(3.15), the optimal single detectors approach can be specified by the following optimization problem:

$$\begin{aligned} &\min_{\boldsymbol{\phi},\,\boldsymbol{S}} \max_{k\in\{1,\dots,K\}} g_k(\boldsymbol{S}) \\ &\quad\text{subject to}\;\; h(\boldsymbol{S}) \leq A \end{aligned} \tag{3.25}$$

where $g_k(\boldsymbol{S})$ can be expressed as in (3.11) by removing the dependence on l in the expressions (since there is only one detector for each user), $\boldsymbol{\phi} = \left[\phi^{(1)} \cdots \phi^{(K)}\right]$ represents the detectors of the users, and $\boldsymbol{S}$ is the vector of signal amplitudes for bit 0 and bit 1 of all users; i.e., $\boldsymbol{S} = \left[S_1^{(0)}\; S_1^{(1)} \cdots S_K^{(0)}\; S_K^{(1)}\right]$.

Since (3.25) is a special case of (3.13)-(3.15), its solution can be obtained from Proposition 3.2.1 by setting $N_{\rm d} = 1$ in (3.16)-(3.18). Hence, the optimal single detectors approach can also be formulated as

$$\begin{aligned} &\min_{\boldsymbol{S}} \max_{k\in\{1,\dots,K\}} \tilde{g}_k(\boldsymbol{S}) \\ &\quad\text{subject to}\;\; h(\boldsymbol{S}) \leq A \end{aligned} \tag{3.26}$$

where $\tilde{g}_k(\boldsymbol{S})$ is as defined in (3.23). In other words, the optimal single detectors approach requires the calculation of the optimal signal amplitudes from (3.26). Then, each user employs the corresponding ML detector, which selects bit 1 if

$p_1^{(k)}(y|\boldsymbol{S}^\diamond) \geq p_0^{(k)}(y|\boldsymbol{S}^\diamond)$ and bit 0 otherwise, where $\boldsymbol{S}^\diamond$ denotes the solution of (3.26).

Let $\mathrm{P_{SD}}$ denote the optimal value achieved by the optimization problem in (3.26) (equivalently, (3.25)); that is, the minimum worst-case (maximum) average probability of error corresponding to the optimal single detectors approach. Similarly, let $\mathrm{P_{DR}}$ represent the solution of the optimization problem in (3.20)-(3.22) (equivalently, (3.13)-(3.15)), which is the minimum worst-case average probability of error achieved by the optimal detector randomization approach. The main purpose of this section is to provide bounds on $\mathrm{P_{DR}}$, and to specify various relations between $\mathrm{P_{SD}}$ and $\mathrm{P_{DR}}$. First, the following proposition is obtained to provide a lower bound on $\mathrm{P_{DR}}$.

Proposition 3.3.1. *The minimum worst-case average probability of error achieved by the optimal detector randomization approach in* (3.20)-(3.22), $\mathrm{P_{DR}}$, *is lower bounded as follows:*

$$\mathrm{P_{DR}} \geq \frac{1}{K}\sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^*) \triangleq \mathrm{P_{LB}} \tag{3.27}$$

with

$$\boldsymbol{S}^* = \arg\min_{\boldsymbol{S}\in\mathcal{S}_A} \sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}) \tag{3.28}$$

where $\mathcal{S}_A$ *is defined as* $\mathcal{S}_A \triangleq \{\boldsymbol{S} : h(\boldsymbol{S}) \leq A\}$ *and* $\tilde{g}_k(\boldsymbol{S})$ *is as in* (3.23). *In addition, the lower bound in* (3.27) *is achieved; that is,* $\mathrm{P_{DR}} = \mathrm{P_{LB}}$, *if and only if there exists feasible* $\{v_l, \boldsymbol{S}_l\}_{l=1}^{\min\{K,N_\mathrm{d}\}}$ *(i.e., satisfying* (3.21) *and* (3.22)*) such that* $\sum_{l=1}^{\min\{K,N_\mathrm{d}\}} v_l\, \tilde{g}_k(\boldsymbol{S}_l) = \mathrm{P_{LB}}$, $\forall k \in \{1,\ldots,K\}$.

Proof: Consider a modified version of the optimization problem in (3.20)-(3.22), which is described as

$$\min_{\{v_l,\boldsymbol{S}_l\}_{l=1}^{\min\{K,N_\mathrm{d}\}}} \frac{1}{K}\sum_{k=1}^{K}\sum_{l=1}^{\min\{K,N_\mathrm{d}\}} v_l\,\tilde{g}_k(\boldsymbol{S}_l) \tag{3.29}$$

$$\text{subject to } \; h(\boldsymbol{S}_l)\leq A\,,\ \forall\, l\in\{1,\ldots,\min\{K,N_\mathrm{d}\}\} \tag{3.30}$$

$$\sum_{l=1}^{\min\{K,N_\mathrm{d}\}} v_l = 1\,,\ v_l\geq 0\,,\ \forall\, l\in\{1,\ldots,\min\{K,N_\mathrm{d}\}\} \tag{3.31}$$

where $\tilde{g}_k(\boldsymbol{S}_l)$ is given by (3.23). Define $g_{\mathrm{avg}}(\boldsymbol{S}) \triangleq \frac{1}{K}\sum_{k=1}^{K}\tilde{g}_k(\boldsymbol{S})$ and express the problem in (3.29)-(3.31) as

$$\min_{\{v_l,\,\boldsymbol{S}_l\in\mathcal{S}_A\}_{l=1}^{\min\{K,N_\mathrm{d}\}}} \sum_{l=1}^{\min\{K,N_\mathrm{d}\}} v_l\,g_{\mathrm{avg}}(\boldsymbol{S}_l) \tag{3.32}$$

$$\text{subject to } \sum_{l=1}^{\min\{K,N_\mathrm{d}\}} v_l = 1,\ v_l\geq 0,\ \forall\, l\in\{1,\ldots,\min\{K,N_\mathrm{d}\}\} \tag{3.33}$$

where $\mathcal{S}_A$ is as described in the proposition. The optimal solution of (3.32)-(3.33) is obtained by assigning all the weight to the minimizer of $g_{\mathrm{avg}}(\boldsymbol{S})$ over $\mathcal{S}_A$, which corresponds to $\boldsymbol{S}^*$ defined in (3.28). For example, $v_1 = 1$, $v_l = 0$ for $l = 2,\ldots,N_\mathrm{d}$, and $\boldsymbol{S}_1 = \boldsymbol{S}^*$ achieves the minimum value of the objective function in (3.32)-(3.33). Therefore, the minimum value achieved by the optimization problem in (3.29)-(3.31) is equal to $g_{\mathrm{avg}}(\boldsymbol{S}^*) = \frac{1}{K}\sum_{k=1}^{K}\tilde{g}_k(\boldsymbol{S}^*)$. When the optimization problems in (3.20)-(3.22) and in (3.29)-(3.31) are compared, it is observed that the latter provides a lower bound on the former since the average of the error probabilities of the users is considered in (3.29) whereas the maximum of the error probabilities is employed in (3.20). (Please note the $\frac{1}{K}\sum_{k=1}^{K}$ and $\max\limits_{k\in\{1,\ldots,K\}}$ operators, respectively.) Therefore, the solution of (3.29)-(3.31), which is specified by $\frac{1}{K}\sum_{k=1}^{K}\tilde{g}_k(\boldsymbol{S}^*)$, provides a lower bound on the solution of (3.20)-(3.22), P_{DR}. Hence, (3.27) is obtained.

In order to prove the sufficiency of the achievability condition in Proposition 3.3.1, assume that there exists feasible $\{v_l,\boldsymbol{S}_l\}_{l=1}^{\min\{K,N_\mathrm{d}\}}$ (i.e., satisfying (3.21) and (3.22)) such that $\sum_{l=1}^{\min\{K,N_\mathrm{d}\}} v_l\,\tilde{g}_k(\boldsymbol{S}_l) = \mathrm{P}_{\mathrm{LB}}$, $\forall k\in\{1,\ldots,K\}$. Then, it is easy to verify from (3.20) and (3.23) that the summation term in (3.20) becomes equal to P_{LB}, $\forall k \in \{1,\ldots,K\}$, for the specified solution. Hence, (3.20)-(3.22) achieves the lower bound in this case, and $\mathrm{P}_{\mathrm{DR}} = \mathrm{P}_{\mathrm{LB}}$ is obtained. For proving the necessity of the achievability condition in the proposition via contradiction, assume that $\mathrm{P}_{\mathrm{DR}} = \mathrm{P}_{\mathrm{LB}}$ and the optimal solution of (3.20)-(3.22), denoted by $\{\hat{v}_l, \hat{\boldsymbol{S}}_l\}_{l=1}^{\min\{K,N_{\mathrm{d}}\}}$, results in a scenario in which the $\sum_{l=1}^{\min\{K,N_{\mathrm{d}}\}} \hat{v}_l\, \tilde{g}_k(\hat{\boldsymbol{S}}_l)$ terms are not all the same. In particular, assume that $\exists k' \in \{1,\ldots,K\}$ such that $\sum_{l=1}^{\min\{K,N_{\mathrm{d}}\}} \hat{v}_l\, \tilde{g}_{k'}(\hat{\boldsymbol{S}}_l) < \mathrm{P}_{\mathrm{LB}}$ and that $\sum_{l=1}^{\min\{K,N_{\mathrm{d}}\}} \hat{v}_l\, \tilde{g}_k(\hat{\boldsymbol{S}}_l) = \mathrm{P}_{\mathrm{LB}}$, $\forall k \in \{1,\ldots,K\}\backslash\{k'\}$.[5] Then, the following inequality is obtained:

$$\frac{1}{K}\sum_{k=1}^{K}\sum_{l=1}^{\min\{K,N_{\mathrm{d}}\}} \hat{v}_l\, \tilde{g}_k(\hat{\boldsymbol{S}}_l) < \mathrm{P}_{\mathrm{LB}}\ . \tag{3.34}$$

However, this implies a contradiction since

$$\frac{1}{K}\sum_{k=1}^{K}\sum_{l=1}^{\min\{K,N_{\mathrm{d}}\}} \hat{v}_l\, \tilde{g}_k(\hat{\boldsymbol{S}}_l) = \sum_{l=1}^{\min\{K,N_{\mathrm{d}}\}} \hat{v}_l \left(\frac{1}{K}\sum_{k=1}^{K}\tilde{g}_k(\hat{\boldsymbol{S}}_l)\right) \geq \mathrm{P}_{\mathrm{LB}} \tag{3.35}$$

where the inequality follows from (3.27). Therefore, when the lower bound is achieved, i.e., $\mathrm{P}_{\mathrm{DR}} = \mathrm{P}_{\mathrm{LB}}$, all the $\sum_{l=1}^{\min\{K,N_{\mathrm{d}}\}} \hat{v}_l\, \tilde{g}_k(\hat{\boldsymbol{S}}_l)$ terms must be equal to P_{LB}. Hence, in order to achieve the lower bound in (3.27), there must exist feasible $\{v_l, \boldsymbol{S}_l\}_{l=1}^{\min\{K,N_{\mathrm{d}}\}}$ such that $\sum_{l=1}^{\min\{K,N_{\mathrm{d}}\}} v_l\, \tilde{g}_k(\boldsymbol{S}_l) = \mathrm{P}_{\mathrm{LB}}$, $\forall k \in \{1,\ldots,K\}$, as stated in the proposition. □

Proposition 3.3.1 presents a bound on the performance of the optimal detector randomization approach in (3.20)-(3.22). The advantage of this lower bound is that it is calculated based on the solution of the minimization problem in (3.28), which is much simpler than the optimization problem in (3.20)-(3.22). In addition, the achievability condition in Proposition 3.3.1 implies that the worst-case average probability of error achieved by the optimal detector randomization

[5]Note that none of the $\sum_{l=1}^{\min\{K,N_d\}} \hat{v}_l\, \tilde{g}_k(\hat{\boldsymbol{S}}_l)$ terms can be larger than P_{LB} since it is assumed that $\mathrm{P}_{\mathrm{DR}} = \mathrm{P}_{\mathrm{LB}}$; i.e., the maximum of these terms is equal to P_{LB} (see (3.20) and (3.23)). Therefore, either all these terms can be equal to P_{LB} or some of them can be smaller than P_{LB}. The latter is shown to be impossible in the remaining part of the proof.

approach attains the lower bound if and only if there exists an *equalizer* solution for the optimal detector randomization problem in (3.20)-(3.22), which equates the average error probabilities of all users to the lower bound in (3.27). As a simple example, if $\boldsymbol{S}^*$ in (3.28) satisfies that $\tilde{g}_1(\boldsymbol{S}^*) = \cdots = \tilde{g}_K(\boldsymbol{S}^*)$, then $v_1 = 1$, $v_l = 0$ for $l = 2, \ldots, \min\{K, N_{\rm d}\}$, and $\boldsymbol{S}_1 = \boldsymbol{S}^*$ results in $\sum_{l=1}^{\min\{K,N_{\rm d}\}} v_l\, \tilde{g}_k(\boldsymbol{S}_l) = \tilde{g}_k(\boldsymbol{S}^*) = \mathrm{P}_{\rm LB}$, $\forall k \in \{1, \ldots, K\}$; hence, the lower bound is achieved in this scenario; i.e., $\mathrm{P}_{\rm DR} = \mathrm{P}_{\rm LB}$, as a result of Proposition 3.3.1. As investigated in the following, there also exist other scenarios in which $\mathrm{P}_{\rm DR} = \mathrm{P}_{\rm LB}$ is satisfied when all $\tilde{g}_k(\boldsymbol{S}^*)$'s are not the same.

Next, improvements that can be achieved via the optimal detector randomization approach over the optimal single detectors approach are quantified in the following proposition.

Proposition 3.3.2. *Let* $\mathrm{P}_{\rm SD}$ *and* $\mathrm{P}_{\rm DR}$ *denote the minimum worst-case error probabilities obtained from the solutions of* (3.26) *and* (3.20)-(3.22), *respectively. Then, the following relations hold between* $\mathrm{P}_{\rm SD}$ *and* $\mathrm{P}_{\rm DR}$.

(i) The improvement ratio, defined as $\mathrm{P}_{\rm SD}/\mathrm{P}_{\rm DR}$, *is bounded as follows:*

$$1 \leq \frac{\mathrm{P}_{\rm SD}}{\mathrm{P}_{\rm DR}} \leq K\,. \tag{3.36}$$

(ii) The maximum improvement ratio, K, *is achieved if and only if* $\mathrm{P}_{\rm DR} = \mathrm{P}_{\rm LB}$ *(where* $\mathrm{P}_{\rm LB}$ *is as defined in* (3.27)*), and* $\boldsymbol{S}^*$ *in* (3.28) *is the optimal solution to the optimization problem in* (3.26) *with* $\tilde{g}_k(\boldsymbol{S}^*) = 0$, $\forall k \in \{1, \ldots, K\} \setminus \{k^*\}$ *and* $\tilde{g}_{k^*}(\boldsymbol{S}^*) > 0$, *where* $\tilde{g}_k$ *is given by* (3.23) *and* k^* *is any value in* $\{1, \ldots, K\}$.

(iii) No improvement is achieved; that is, $\mathrm{P}_{\rm DR} = \mathrm{P}_{\rm SD}$, *if* $\tilde{g}_1(\boldsymbol{S}^*) = \cdots = \tilde{g}_K(\boldsymbol{S}^*)$.

(iv) Improvement is guaranteed; that is, $\mathrm{P}_{\rm DR} < \mathrm{P}_{\rm SD}$, *if* $\mathrm{P}_{\rm DR} = \mathrm{P}_{\rm LB}$ *and* $\tilde{g}_k(\boldsymbol{S}^\diamond) \neq \tilde{g}_l(\boldsymbol{S}^\diamond)$ *for some* $k, l \in \{1, \ldots, K\}$, *where* $\boldsymbol{S}^\diamond$ *denotes the solution of* (3.26).

Proof: *(i)* Since the optimal single detectors approach is a special case of the detector randomization approach, $\mathrm{P}_{\mathrm{DR}} \leq \mathrm{P}_{\mathrm{SD}}$ is always satisfied; hence, the lower bound in (3.36) is directly obtained. In order to derive the upper bound in (3.36), the following inequalities are considered first:

$$\mathrm{P}_{\mathrm{SD}} = \max_k \tilde{g}_k(\boldsymbol{S}^{\diamond}) \leq \max_k \tilde{g}_k(\boldsymbol{S}^{*}) \leq \sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^{*}) \tag{3.37}$$

where $\boldsymbol{S}^{\diamond}$ is the solution of (3.26), and $\boldsymbol{S}^{*}$ is given by (3.28). Note that the first inequality follows by definition since $\boldsymbol{S}^{\diamond}$ and $\boldsymbol{S}^{*}$ are the solutions of (3.26) and (3.28), respectively, and the second inequality follows from the identity $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1$, $\forall \mathbf{x}$, where $\|\mathbf{x}\|_\infty$ and $\|\mathbf{x}\|_1$ are the maximum and Manhattan norms, respectively. Then, the upper bound in (3.36) is obtained as follows:

$$\frac{\mathrm{P}_{\mathrm{SD}}}{\mathrm{P}_{\mathrm{DR}}} \leq \frac{\sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^{*})}{\mathrm{P}_{\mathrm{DR}}} \leq \frac{\sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^{*})}{\mathrm{P}_{\mathrm{LB}}} = K \tag{3.38}$$

where the first inequality is obtained from (3.37), and the second inequality and the equality follow from (3.27).

(ii) In order to achieve the maximum improvement ratio of K in (3.36), the inequalities in (3.37) and (3.38) should hold with equality. Then, from (3.37), it is concluded that $\boldsymbol{S}^{*}$ in (3.28) should also be a solution of (3.26) (so that $\max_k \tilde{g}_k(\boldsymbol{S}^{\diamond}) = \max_k \tilde{g}_k(\boldsymbol{S}^{*})$), and $\tilde{g}_k(\boldsymbol{S}^{*})$ should be zero for all k except for one of them (so that $\max_k \tilde{g}_k(\boldsymbol{S}^{*}) = \sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^{*})$). In addition, for the second inequality in (3.38) to hold with equality, $\mathrm{P}_{\mathrm{DR}} = \mathrm{P}_{\mathrm{LB}}$ should be satisfied. Hence, the conditions in Part *(ii)* of Proposition 3.3.2 are obtained.

(iii) Consider a scenario in which $\tilde{g}_1(\boldsymbol{S}^{*}) = \cdots = \tilde{g}_K(\boldsymbol{S}^{*})$. In order to prove that $\mathrm{P}_{\mathrm{DR}} = \mathrm{P}_{\mathrm{SD}}$ via contradiction, first suppose that $\max_k \tilde{g}_k(\boldsymbol{S}^{\diamond}) < \max_k \tilde{g}_k(\boldsymbol{S}^{*})$. Then, the following relation is obtained:

$$\sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^{\diamond}) \leq K \max_k \tilde{g}_k(\boldsymbol{S}^{\diamond}) < K \max_k \tilde{g}_k(\boldsymbol{S}^{*}) = \sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^{*}). \tag{3.39}$$

Note that the second inequality and the equality in (3.39) are due to the assumptions of $\max_k \tilde{g}_k(\boldsymbol{S}^{\diamond}) < \max_k \tilde{g}_k(\boldsymbol{S}^{*})$ and $\tilde{g}_1(\boldsymbol{S}^{*}) = \cdots = \tilde{g}_K(\boldsymbol{S}^{*})$, respectively. Since (3.39) implies that $\sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^{\diamond}) < \sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^{*})$, it results in a contradiction due to the definition of $\boldsymbol{S}^*$ in (3.28). Therefore, when $\tilde{g}_1(\boldsymbol{S}^*) = \cdots = \tilde{g}_K(\boldsymbol{S}^*)$, the relation $\max_k \tilde{g}_k(\boldsymbol{S}^\diamond) < \max_k \tilde{g}_k(\boldsymbol{S}^*)$ cannot be true. This implies that $\max_k \tilde{g}_k(\boldsymbol{S}^\diamond) = \max_k \tilde{g}_k(\boldsymbol{S}^*)$ must be satisfied in this scenario since $\max_k \tilde{g}_k(\boldsymbol{S}^\diamond) \leq \max_k \tilde{g}_k(\boldsymbol{S}^*)$ is always satisfied by definition (as $\boldsymbol{S}^\diamond$ is the solution of (3.26)). Then, $\mathrm{P}_{\mathrm{DR}} = \mathrm{P}_{\mathrm{SD}}$ is obtained as follows:

$$\mathrm{P}_{\mathrm{SD}} = \max_k \tilde{g}_k(\boldsymbol{S}^\diamond) = \max_k \tilde{g}_k(\boldsymbol{S}^*) = \frac{1}{K}\sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^*) = \mathrm{P}_{\mathrm{LB}} \tag{3.40}$$

where the third equality is due to $\tilde{g}_1(\boldsymbol{S}^*) = \cdots = \tilde{g}_K(\boldsymbol{S}^*)$ and the last equality is from (3.27). Since in general $\mathrm{P}_{\mathrm{LB}} \leq \mathrm{P}_{\mathrm{DR}} \leq \mathrm{P}_{\mathrm{SD}}$ holds (see (3.27) and Part *(i)* of Proposition 3.3.2), (3.40) implies that $\mathrm{P}_{\mathrm{DR}} = \mathrm{P}_{\mathrm{SD}} = \mathrm{P}_{\mathrm{LB}}$ when $\tilde{g}_1(\boldsymbol{S}^*) = \cdots = \tilde{g}_K(\boldsymbol{S}^*)$.

(iv) Assume that $\mathrm{P}_{\mathrm{DR}} = \mathrm{P}_{\mathrm{LB}}$ and $\tilde{g}_k(\boldsymbol{S}^\diamond) \neq \tilde{g}_l(\boldsymbol{S}^\diamond)$ for some $k, l \in \{1, \ldots, K\}$. Then, the result is derived as follows:

$$\mathrm{P}_{\mathrm{SD}} = \max_k \tilde{g}_k(\boldsymbol{S}^\diamond) > \frac{1}{K}\sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^\diamond) \geq \frac{1}{K}\sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^*) = \mathrm{P}_{\mathrm{LB}} = \mathrm{P}_{\mathrm{DR}}\ , \tag{3.41}$$

where the first inequality is obtained from the assumption that $\tilde{g}_k(\boldsymbol{S}^\diamond) \neq \tilde{g}_l(\boldsymbol{S}^\diamond)$ for some $k, l \in \{1, \ldots, K\}$, the second inequality and the second equality follow from Proposition 3.3.1, and the final equality is due to the assumption of $\mathrm{P}_{\mathrm{DR}} = \mathrm{P}_{\mathrm{LB}}$. $\square$

Proposition 3.3.2 quantifies the improvements that can be achieved via the optimal detector randomization approach and states that the worst-case average probability of error can be reduced by a factor of K compared to the optimal single detectors approach that does not perform any detector randomization. Therefore, significant gains can be possible in the presence of a large number of users. In addition, the scenarios in which this maximum improvement ratio can be achieved are specified based on the conditions in Part *(ii)* of the proposition. It should be noted that the condition of $\tilde{g}_k(\boldsymbol{S}^*) = 0$, $\forall k \in \{1 \ldots K\} \setminus \{k^*\}$ and $\tilde{g}_{k^*}(\boldsymbol{S}^*) > 0$ cannot hold exactly for ML detectors that operate in the presence of Gaussian noise, which has an infinite support. Therefore, the maximum improvement ratio of K may not be achieved exactly in practice; however, it can be quite close to K in certain scenarios (see, e.g., Figure 3.3 at 28 dB). Proposition 3.3.2 also provides some simple conditions to determine if the optimal detector randomization approach can or cannot provide any improvements over the optimal single detectors approach.

Remark 1: Although the results in Proposition 3.3.1 and Proposition 3.3.2 are obtained when all the users employ ML detectors, which are specified by the error probability expression $\tilde{g}_k$ in (3.23), the results are also valid for other types of detectors; e.g., the sign detector or the optimal single-threshold detector. In other words, Proposition 3.3.1 and Proposition 3.3.2 hold for arbitrary $\tilde{g}_k$ corresponding to any type of detectors. □

In the following proposition, the structure of the optimal detector randomization solution obtained from (3.20)-(3.22) is specified in the case of equal crosscorrelations and noise powers.

Proposition 3.3.3. *Assume that there are at least K detectors at each receiver; that is, $N_{\mathrm{d}} \geq K$. If the crosscorrelations between the pseudo-noise signals for different users are equal; i.e., $\rho_{k,j} = \rho$, $\forall k \neq j$, and the standard deviations of the noise at the receivers are the same; i.e., $\sigma_k = \sigma$, $\forall k$, then an optimal solution to* (3.20)-(3.22)*, which achieves the lower bound in* (3.27)*, can be expressed as*

$$v_l = \frac{1}{K} \;, \; \boldsymbol{S}_l = \mathrm{CS}_{2l-2}(\boldsymbol{S}^*) \;\; \textit{for } l = 1, \ldots, K \tag{3.42}$$

where $\boldsymbol{S}^*$ is as in (3.28) *and $\mathrm{CS}_{2l-2}(\boldsymbol{S}^*)$ denotes the circular shift of the elements of $\boldsymbol{S}^*$ by $2l - 2$ positions.*[6]

[6] Since $\boldsymbol{S}^*$ is feasible; i.e, satisfies $h(\boldsymbol{S}^*) \leq A$ by definition (see (3.28)), $\mathrm{CS}_{2l}(\boldsymbol{S}^*)$'s are feasible as well due to the definition of h in (3.9).

Proof: When the solution in (3.42) is employed, the objective function in (3.20) becomes

$$\max_{k\in\{1,\dots,K\}} \frac{1}{2K}\int_{-\infty}^{\infty}\sum_{l=1}^{K}\min\Big\{p_0^{(k)}(y|\,\mathrm{CS}_{2l-2}(\boldsymbol{S}^*))\,,p_1^{(k)}(y|\,\mathrm{CS}_{2l-2}(\boldsymbol{S}^*))\Big\}dy\,. \tag{3.43}$$

In addition, for equal crosscorrelations and noise variances, $p_{i_k}^{(k)}(y|\boldsymbol{S}_l)$ in (3.19) is given by

$$p_{i_k}^{(k)}(y|\boldsymbol{S}_l)=\frac{1}{\sigma\sqrt{2\pi}\,2^{K-1}}\sum_{\mathbf{i}_k\in\{0,1\}^{K-1}}\exp\left\{-\frac{1}{2\sigma^2}\left(y-S_{k,l}^{(i_k)}-\rho\sum_{\substack{j=1\\j\neq k}}^{K}S_{j,l}^{(i_j)}\right)^2\right\} \tag{3.44}$$

for $i_k=0,1$, where $\boldsymbol{S}_l=\big[S_{1,l}^{(0)}\ S_{1,l}^{(1)}\cdots S_{K,l}^{(0)}\ S_{K,l}^{(1)}\big]$ and $\mathbf{i}_k=[i_1\cdots i_{k-1}\ i_{k+1}\cdots i_K]$. Then, if $\boldsymbol{S}_l=\mathrm{CS}_{2l-2}(\boldsymbol{S}^*)$ is employed for $l=1,\dots,K$, where $\boldsymbol{S}^*\triangleq[S_{1,*}^{(0)}\ S_{1,*}^{(1)}\cdots S_{K,*}^{(0)}\ S_{K,*}^{(1)}]$, it can be shown based on (3.44) that the $\sum_{l=1}^{K}\min\Big\{p_0^{(k)}(y|\,\mathrm{CS}_{2l-2}(\boldsymbol{S}^*)),p_1^{(k)}(y|\,\mathrm{CS}_{2l-2}(\boldsymbol{S}^*))\Big\}$ terms in (3.43) become equal for $k=1,\dots,K$.[7] Therefore, the overall expression in (3.43) can be stated as

$$\frac{1}{2K}\sum_{l=1}^{K}\int_{-\infty}^{\infty}\min\Big\{p_0^{(k)}(y|\,\mathrm{CS}_{2l-2}(\boldsymbol{S}^*))\,,p_1^{(k)}(y|\,\mathrm{CS}_{2l-2}(\boldsymbol{S}^*))\Big\}dy \tag{3.45}$$

for *any* $k\in\{1,\dots,K\}$. From (3.44), it is easy to verify that (3.45) is also equal to

$$\frac{1}{2K}\sum_{k=1}^{K}\int_{-\infty}^{\infty}\min\Big\{p_0^{(k)}(y|\boldsymbol{S}^*),p_1^{(k)}(y|\boldsymbol{S}^*)\Big\}\,dy\ , \tag{3.46}$$

which can be expressed as $\frac{1}{K}\sum_{k=1}^{K}\tilde{g}_k(\boldsymbol{S}^*)\triangleq \mathrm{P}_{\mathrm{LB}}$ based on the definitions in (3.23) and (3.27). Hence, it is observed that for the solution in (3.42), the optimization problem in (3.20)-(3.22) achieves the lower bound in Proposition 3.3.1; i.e., (3.42) provides an optimal solution to (3.20)-(3.22) that achieves the lower bound in (3.27), as claimed in the proposition. □

[7]For example, if $K=2$, then $\mathrm{CS}_0(\boldsymbol{S}^*)=[S_{1,*}^{(0)}\ S_{1,*}^{(1)}\ S_{2,*}^{(0)}\ S_{2,*}^{(1)}]$ and $\mathrm{CS}_2(\boldsymbol{S}^*)=[S_{2,*}^{(0)}\ S_{2,*}^{(1)}\ S_{1,*}^{(0)}\ S_{1,*}^{(1)}]$, for which $\min\{p_0^{(k)}(y|\,\mathrm{CS}_0(\boldsymbol{S}^*)),p_1^{(k)}(y|\,\mathrm{CS}_0(\boldsymbol{S}^*))\}+\min\{p_0^{(k)}(y|\,\mathrm{CS}_2(\boldsymbol{S}^*)),p_1^{(k)}(y|\,\mathrm{CS}_2(\boldsymbol{S}^*))\}$ is the same for $k=1$ and $k=2$, as can be observed from (3.44).

Although the optimal solution to the generic problem in (3.20)-(3.22) requires a search over a $(2K+1)K$ dimensional space (assuming $N_{\mathrm{d}} \geq K$), a significantly simpler solution can be obtained under the conditions in Proposition 3.3.3; namely, the following algorithm can be employed: (i) Calculate $\boldsymbol{S}^*$ from (3.28). (ii) Obtain the optimal solution as in (3.42).[8] It is noted that this algorithm requires a search over a $2K$ dimensional space in order to calculate $\boldsymbol{S}^*$. In addition, if symmetric signaling is employed, the search space dimensions reduce to $(K+1)K$ and K for the problems in (3.20)-(3.22) and in (3.28), respectively.

Remark 2: Under the conditions in Proposition 3.3.3, if $\boldsymbol{S}^*$ is a solution of (3.28), any permutation of the signal amplitude pairs for different users is a solution as well.[9] For example, if $\boldsymbol{S}^* = \left[S_{1,*}^{(0)}\ S_{1,*}^{(1)}\ S_{2,*}^{(0)}\ S_{2,*}^{(1)}\ S_{3,*}^{(0)}\ S_{3,*}^{(1)}\right]$ $= [-1\ 1\ -2\ 2\ -3\ 3]$, then $[-1\ 1\ -3\ 3\ -2\ 2]$, $[-2\ 2\ -1\ 1\ -3\ 3]$, $[-2\ 2\ -3\ 3\ -1\ 1]$, $[-3\ 3\ -1\ 1\ -2\ 2]$, and $[-3\ 3\ -2\ 2\ -1\ 1]$ are solutions of (3.28), too. $\square$

The following proposition presents necessary and sufficient conditions for the uniqueness of the solution in (3.42).

Proposition 3.3.4. *Consider scenarios in which performance improvements are achieved via optimal detector randomization over the optimal single detectors approach. Under the conditions in Proposition 3.3.3, the optimal solution in* (3.42) *is unique if and only if*

- *the solution of* (3.28), $\boldsymbol{S}^*$, *is unique up to permutations of signal amplitude pairs (see Remark 2), and*
- *the signal amplitude pairs in* $\boldsymbol{S}^*$ *are the same except for one of them.*[10]

[8]The definition of the circular shift in Proposition 3.3.3 can be a right circular shift or a left circular shift without affecting the optimality of the solution in (3.42).

[9]This is implied by the proof of Proposition 3.3.3 based on the equivalence of (3.45) and (3.46) (see (3.23) and (3.28) as well).

[10]The case in which $\boldsymbol{S}^*$ is unique and the signal amplitude pairs in $\boldsymbol{S}^*$ are all the same is not considered since no improvement is achieved via detector randomization in that scenario

Proof: Please see Appendix 3.6.1.

Proposition 3.3.4 guarantees the uniqueness of the optimal solution in (3.42) based on the uniqueness of the solution $\boldsymbol{S}^*$ of (3.28) and the structure of this optimal solution. As an example, for $K = 4$, if $\boldsymbol{S}^* = [-1\ 1\ -1\ 1\ -1\ 1\ -2\ 2]$ is the unique solution of (3.28) up to permutations of signal amplitude pairs (i.e., the only solutions of (3.28) are $[-1\ 1\ -1\ 1\ -1\ 1\ -2\ 2]$, $[-1\ 1\ -1\ 1\ -2\ 2\ -1\ 1]$, $[-1\ 1\ -2\ 2\ -1\ 1\ -1\ 1]$, and $[-2\ 2\ -1\ 1\ -1\ 1\ -1\ 1]$), then the optimal solution is unique as a result of Proposition 3.3.4 since the signal amplitude pairs in $\boldsymbol{S}^*$ are the same except for one of them. Also, from Proposition 3.3.3, the optimal solution in (3.42) is given by $v_1 = v_2 = v_3 = v_4 = 0.25$, $\boldsymbol{S}_1 = [-1\ 1\ -1\ 1\ -1\ 1\ -2\ 2]$ $\boldsymbol{S}_2 = [-2\ 2\ -1\ 1\ -1\ 1\ -1\ 1]$, $\boldsymbol{S}_3 = [-1\ 1\ -2\ 2\ -1\ 1\ -1\ 1]$, and $\boldsymbol{S}_4 = [-1\ 1\ -1\ 1\ -2\ 2\ -1\ 1]$ in this example.

3.4 Performance Evaluation

In this section, numerical results are presented to investigate the theoretical results obtained in the previous sections and to compare the proposed optimal detector randomization approach against other approaches that do not perform any detector randomization. Specifically, the following approaches are considered in the simulations.

Optimal Detector Randomization: This scheme refers to the proposed optimization problem in (3.13)-(3.15), which can be solved via (3.20)-(3.22), as stated in Proposition 3.2.2. It is noted that when the conditions in Proposition 3.3.3 are satisfied, the optimal solution can also be obtained via (3.42), which has significantly lower computational complexity.

(i.e., the condition in Part *(iii)* of Proposition 3.3.2 is satisfied). Specifically, $\boldsymbol{S}^*$ is employed all the time and each user runs a single ML detector corresponding to $\boldsymbol{S}^*$.

Optimal Single Detectors: In this approach, a single detector is employed by each user; hence, no detector randomization is performed. The solution is obtained from (3.25) (equivalently, (3.26)). Namely, the optimal signals and the corresponding single detectors (ML rules) are calculated in this approach.

Single Detectors at Power Limit: This approach employs a single detector for each user, and equalizes the signal-to-interference-plus-noise ratios (SINRs) at all the detectors. In addition, all the available power is utilized. Specifically, in this scheme, the signal amplitudes are chosen in such a way that $\mathrm{SINR}_1 = \cdots = \mathrm{SINR}_K$ and $h(\boldsymbol{S}) = A$, where SINR_k is the SINR for user k and $h(\boldsymbol{S})$ is as in (3.9). The SINR for user k can be calculated from (3.3) as $\mathrm{SINR}_k = \mathrm{E}\{|S_k^{(i_k)}|^2\}/(\mathrm{E}\{|\sum_{j\neq k}\rho_{k,j}S_j^{(i_j)}|^2\} + \sigma_k^2)$ for $k = 1, \ldots, K$, which becomes $\mathrm{SINR}_k = |S_k^{(1)}|^2/(\sum_{j\neq k}\rho_{k,j}^2|S_j^{(1)}|^2 + \sigma_k^2)$ for symmetric signaling. In general, the *single detectors at power limit* approach has low computational complexity compared to the other approaches; however, it can result in degraded performance as investigated in the following.

In the simulations, symmetric signaling with equiprobable information symbols is considered for all users, and the standard deviations of the noise at the receivers are set to the same value; i.e., $\sigma_k = \sigma$, $k = 1, \ldots, K$. In addition, as stated after (3.3), $\rho_{k,j}$'s are taken as one for $k = j$; that is, $\rho_{k,k} = 1$ for $k = 1, \ldots, K$.

First, a 5-user scenario is considered (that is, $K = 5$), and the crosscorrelations between the pseudo-noise signals for different users are set to 0.27; i.e., $\rho_{k,j} = 0.27$ for $k \neq j$. Also, the average power constraint A in (3.6) is taken as 5. In Figure 3.3, the maximum average probability of error is plotted versus $1/\sigma^2$ for the optimal detector randomization, optimal single detectors, and single detectors at power limit approaches. From the figure, it is observed that the optimal detector randomization approach achieves the best performance among all the approaches, and the optimal single detectors approach outperforms the

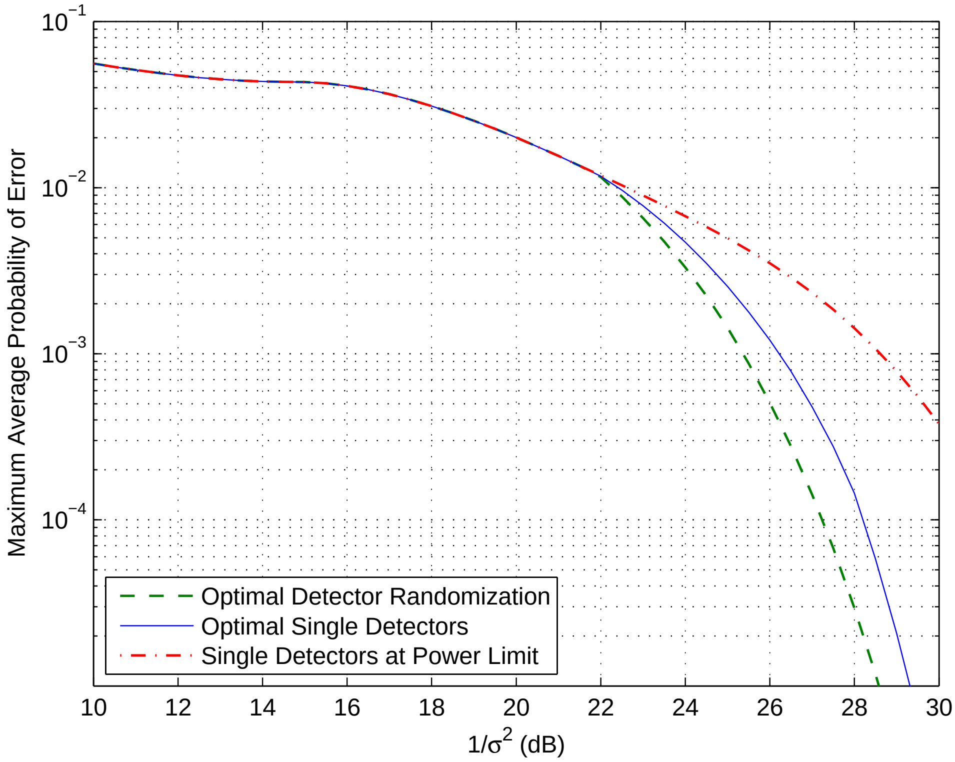

Figure 3.3: Maximum average probability of error versus $1/\sigma^2$ for the optimal detector randomization, optimal single detectors, and single detectors at power limit approaches, where $K = 5$, $\rho_{k,j} = 0.27$ for all $k \neq j$, and $A = 5$.

single detectors at power limit approach for small noise variances. In addition, the calculations show that for high noise variances the *nonimprovability* condition in Part *(iii)* of Proposition 3.3.2 is satisfied, while for small noise variances the *improvability* condition stated in Part *(iv)* of the same proposition is valid. It is also noted that the improvement ratio, which is the ratio between the maximum error probabilities of the optimal single detectors and optimal detector randomization approaches, satisfies the inequality (3.36) in Proposition 3.3.2. In particular, the maximum improvement ratio of 5 is approximately achieved at $1/\sigma^2 = 28\,\text{dB}$.

In order to investigate the results in Figure 3.3 in more detail, Table 3.1 presents the solution $\boldsymbol{S}^\diamond$ of the optimal single detectors approach in (3.26) for various noise variances, where $\boldsymbol{S}^\diamond = \left[S_{1,\diamond}^{(0)}\ S_{1,\diamond}^{(1)} \cdots S_{K,\diamond}^{(0)}\ S_{K,\diamond}^{(1)}\right]$. Since symmetric signaling is employed, only the signal amplitudes corresponding to bit 1 of the users

Table 3.1: Solution of the optimal single detectors approach in (3.26) for the scenario in Figure 3.3. (Only the signal amplitudes for bit 1 of the users are shown due to symmetry.)

$1/\sigma^2$ (dB)	$S_{1,\diamond}^{(1)}$	$S_{2,\diamond}^{(1)}$	$S_{3,\diamond}^{(1)}$	$S_{4,\diamond}^{(1)}$	$S_{5,\diamond}^{(1)}$
18	1	1	1	1	1
20	1	1	1	1	1
22	1.1167	0.9686	0.9686	0.9686	0.9686
24	1.1321	0.9642	0.9642	0.9642	0.9642
26	1.1421	0.9612	0.9612	0.9612	0.9612
28	0.1514	1.1154	1.1154	1.1154	1.1154

Table 3.2: Solution of (3.28), $\boldsymbol{S}^*$, for the scenario in Figure 3.3. (Only the signal amplitudes for bit 1 of the users are shown due to symmetry.) Note that $\boldsymbol{S}^*$ specifies the solution of the optimal detector randomization approach as in (3.42).

$1/\sigma^2$ (dB)	$S_{1,*}^{(1)}$	$S_{2,*}^{(1)}$	$S_{3,*}^{(1)}$	$S_{4,*}^{(1)}$	$S_{5,*}^{(1)}$
18	1	1	1	1	1
20	1	1	1	1	1
22	0.1531	1.1154	1.1154	1.1154	1.1154
24	0.1522	1.1154	1.1154	1.1154	1.1154
26	0.1516	1.1155	1.1155	1.1155	1.1155
28	0.1513	1.1155	1.1155	1.1155	1.1155

are shown in the table. (The signal amplitudes for bit 0 are given by $S_{k,\diamond}^{(0)} = -S_{k,\diamond}^{(1)}$ for $k = 1, 2, 3, 4, 5.$). In addition, Table 3.2 illustrates the solution of (3.28), $\boldsymbol{S}^*$, which specifies the solution of the optimal detector randomization approach as described in (3.42) in Proposition 3.3.3. Again only the signal amplitudes corresponding to bit 1 of the users are shown due to symmetry. From Tables 3.1 and 3.2, it is observed that both the optimal single detectors and the optimal detector randomization approaches converge to the single detectors at power limit approach for high noise variances. This is due to the fact that the Gaussian noise becomes dominant as the noise variance increases and the multiuser interference plus noise term becomes approximately a Gaussian random variable, in which case the optimal solution is to assign equal powers for all users at the maximum power limit. Also, it is noted that the *nonimprovability* condition in Part *(iii)* of Proposition 3.3.2 is satisfied for that scenario. On the other hand, for small noise variances, the solutions become different from that of the single detectors at power limit approach, and improvements are achieved as observed in Figure 3.3. In addition, Table 3.2 implies that the conditions in Proposition 3.3.4 are satisfied for small noise variances; hence, the solution of the optimal detector randomization approach specified in (3.28) is unique in those scenarios. For example, at $1/\sigma^2 = 24\,\text{dB}$, the unique solution of the optimal detector randomization approach is specified by $v_l = 0.2$ and $\boldsymbol{S}_l = \text{CS}_{2l-2}([-0.1522\ 0.1522\ -1.1154\ 1.1154\ -1.1154\ 1.1154\ -1.1154\ 1.1154\ -1.1154\ 1.1154])$ for $l = 1, 2, 3, 4, 5$. Another important observation can be made from Table 3.2 regarding the signal values for the optimal detector randomization approach. When the noise variance is smaller than a certain value, the optimal solution does not vary significantly with the noise level. Hence, perfect knowledge of the noise level may not be required for achieving a near optimal performance. Finally, it is observed from Tables 3.1 and 3.2 that the optimal signal values are the same for many (or, all) of the users at a given noise variance, which is mainly due to the the structures of the optimization problems in (3.26) and (3.28), and the facts that the

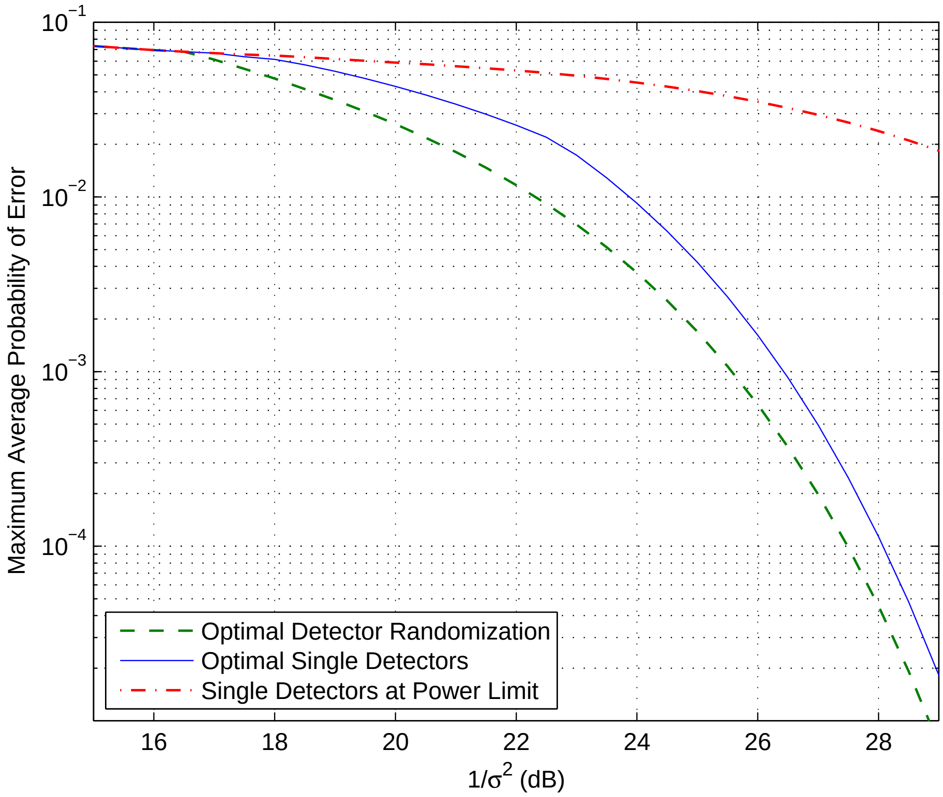

Figure 3.4: Maximum average probability of error versus $1/\sigma^2$ for the optimal detector randomization, optimal single detectors, and single detectors at power limit approaches, where $K = 5$, $\rho_{k,j} = 0.35$ for all $k \neq j$, and $A = 5$.

crosscorrelations between the pseudo-noise signals for different users are equal, and the standard deviations of the noise at the receivers are the same. In other words, the optimization problems in (3.26) and (3.28) tend to yield equalizer rules (for all or some of the users) in the considered scenario.

Next, another scenario with $K = 5$ users is considered, where $\rho_{k,j} = 0.35$ for $k \neq j$, and $A = 5$. In Figure 3.4, the maximum average probability of error is illustrated for the optimal detector randomization, optimal single detectors, and single detectors at power limit approaches. Similar observations to those for Figure 3.3 can be made. The main difference is that improvements are achieved for a larger range of noise variances in this scenario. In addition, the solutions of the optimal single detectors and the optimal detector randomization approaches are specified in Tables 3.3 and 3.4 for the scenario in Figure 3.4 for some values

Table 3.3: Solution of the optimal single detectors approach in (3.26) for the scenario in Figure 3.4.

$1/\sigma^2$ (dB)	$S^{(1)}_{1,\diamond}$	$S^{(1)}_{2,\diamond}$	$S^{(1)}_{3,\diamond}$	$S^{(1)}_{4,\diamond}$	$S^{(1)}_{5,\diamond}$
15	1	1	1	1	1
20	1.1099	1.1099	0.9195	0.9195	0.9195
25	0.2180	0.2180	1.2787	1.2787	1.2787
30	0.2218	0.2218	1.2782	1.2782	1.2782

Table 3.4: Solution of (3.28), $\boldsymbol{S}^*$, for the scenario in Figure 3.4. Note that $\boldsymbol{S}^*$ specifies the solution of the optimal detector randomization approach as in (3.42).

$1/\sigma^2$ (dB)	$S^{(1)}_{1,*}$	$S^{(1)}_{2,*}$	$S^{(1)}_{3,*}$	$S^{(1)}_{4,*}$	$S^{(1)}_{5,*}$
15	1	1	1	1	1
20	0.2084	0.2084	1.2797	1.2797	1.2797
25	0.2180	0.2180	1.2787	1.2787	1.2787
30	0.2218	0.2218	1.2782	1.2782	1.2782

of $1/\sigma^2$. Again similar observations to those in the previous scenario can be made. However, in this case, the solution in (3.28) is not unique since the second uniqueness condition in Proposition 3.3.4 is not satisfied, as observed from Table 3.4.

Then, a scenario with $K = 6$ users is considered, where $\rho_{k,j} = 0.21$ for $k \neq j$, and $A = 6$. In Figure 3.5, the maximum average probability of error is illustrated for the optimal detector randomization, optimal single detectors, and single detectors at power limit approaches. Similar observations as in the previous scenarios are made. The main difference in this scenario is that the improvement ratio is smaller than those in Figure 3.3 and Figure 3.4. Also, the solutions of the optimal single detectors and the optimal detector randomization approaches are specified in Tables 3.5 and 3.6 for the scenario in Figure 3.5 for some values of $1/\sigma^2$.

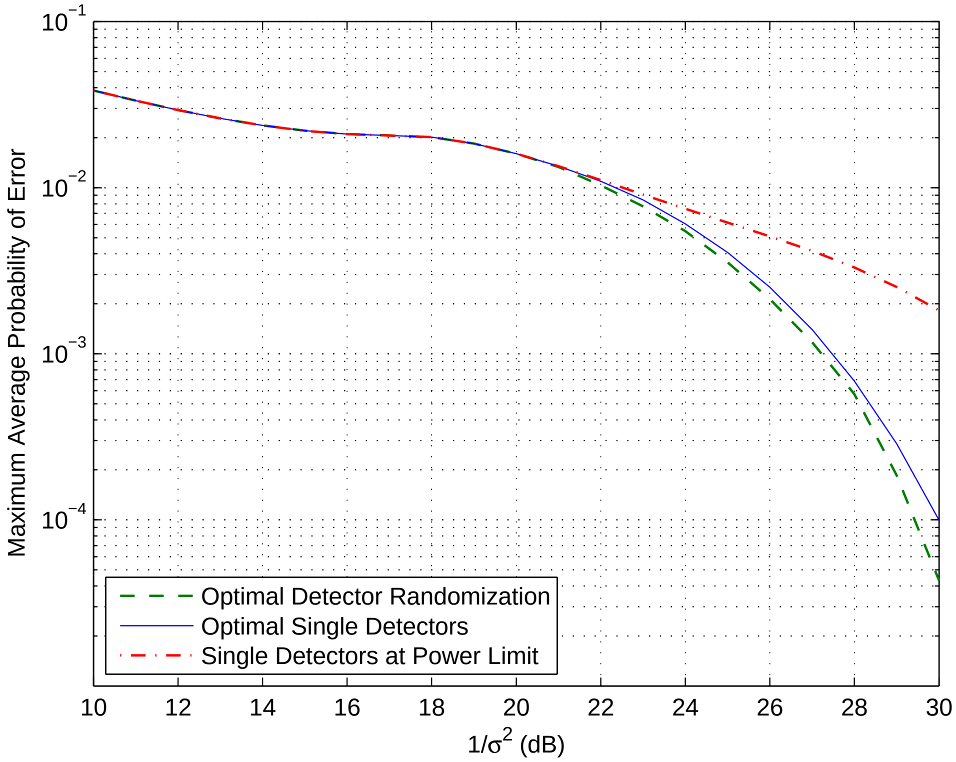

Figure 3.5: Maximum average probability of error versus $1/\sigma^2$ for the optimal detector randomization, optimal single detectors, and single detectors at power limit approaches, where $K = 6$, $\rho_{k,j} = 0.21$ for all $k \neq j$, and $A = 6$.

Table 3.5: Solution of the optimal single detectors approach in (3.26) for the scenario in Figure 3.5.

$1/\sigma^2$ (dB)	$S_{1,\diamond}^{(1)}$	$S_{2,\diamond}^{(1)}$	$S_{3,\diamond}^{(1)}$	$S_{4,\diamond}^{(1)}$	$S_{5,\diamond}^{(1)}$	$S_{6,\diamond}^{(1)}$
18	1	1	1	1	1	1
20	1	1	1	1	1	1
22	1.0662	0.9862	0.9862	0.9862	0.9862	0.9862
24	1.0978	0.9793	0.9793	0.9793	0.9793	0.9793
26	1.1353	0.9707	0.9707	0.9707	0.9707	0.9707
28	1.1602	0.9648	0.9648	0.9648	0.9648	0.9648

Table 3.6: Solution of (3.28), $\boldsymbol{S}^*$, for the scenario in Figure 3.5. Note that $\boldsymbol{S}^*$ specifies the solution of the optimal detector randomization approach as in (3.42).

$1/\sigma^2$ (dB)	$S_{1,*}^{(1)}$	$S_{2,*}^{(1)}$	$S_{3,*}^{(1)}$	$S_{4,*}^{(1)}$	$S_{5,*}^{(1)}$	$S_{6,*}^{(1)}$
18	1	1	1	1	1	1
20	1	1	1	1	1	1
22	1.1117	0.9761	0.9761	0.9761	0.9761	0.9761
24	1.1283	0.9723	0.9723	0.9723	0.9723	0.9723
26	1.1430	0.9689	0.9689	0.9689	0.9689	0.9689
28	1.1606	0.9647	0.9647	0.9647	0.9647	0.9647

3.5 Conclusions and Extensions

Optimal detector randomization has been studied for the downlink of a DSSS system. An optimization problem has been formulated in order to obtain the optimal signal amplitudes, detectors, and detector randomization factors. Since this joint optimization problem is quite challenging in general, a simplified problem has been proposed, in which the search is performed over signal amplitudes and detector randomization factors only, and then the ML detectors corresponding to the optimal signal amplitudes are employed at the receivers. It has been shown that this simplified approach provides the optimal solution to the generic problem when detector randomization is performed over at most $\min\{K, N_{\mathrm{d}}\}$ detector sets, where K is the number of users and N_{d} is the number of detectors at each receiver. Then, the performance of the optimal detector randomization approach has been investigated, and a lower bound has been obtained for the minimum worst-case average probability of error. Also, it has been shown that the optimal detector randomization approach can improve the performance of the optimal single detectors approach by up to K times. In addition, various sufficient conditions have been obtained for the improvability and nonimprovability via detector randomization. Furthermore, in the special case of equal crosscorrelations and noise powers, a simple solution has been provided for the optimal detector randomization problem, and necessary and sufficient conditions have been presented for the uniqueness of that solution. Finally, numerical examples have been presented in order to illustrate the improvements achieved via detector randomization.

Although the downlink of a DSSS system is considered in this study, the results can also be applied to the uplink of a synchronous DSSS under certain assumptions. Specifically, suppose that the receiver (the base station or the access point) employs a bank of K correlators corresponding to the pseudo-noise signals of the users and then performs the bit decision for user k based on the k^{th} correlator output via detector randomization among N_{d} detectors, where $k \in \{1, \ldots, K\}$. In this scenario, the theoretical results in Section 3.2 and Section 3.3 can be extended to the uplink as well. However, when an asynchronous system is considered or when the receiver employs multiuser detection approaches [21], the results in this study cannot be directly applied. Therefore, optimal detector randomization in such scenarios is considered as a future work.

The results in this study can also be extended to cover scenarios in which each user performs M-ary modulation for $M > 2$. In that case, the definitions of $\phi_l^{(k)}$ in (3.4), $\boldsymbol{S}_l$ at the beginning of Section 3.2, h in (3.9), $g_{k,l}$ in (3.11), and $\tilde{g}_k$ in (3.23) can be updated accordingly, and the theoretical results in the previous sections can still be employed based on these new definitions.

3.6 Appendices

3.6.1 Proof of Proposition 3.3.4

First, it is shown that the optimal solution in (3.42) is unique if the conditions in the proposition are satisfied. To that aim, define the following sets

$$\mathcal{S}_{\text{equ}} = \left\{ \boldsymbol{S} \in \mathcal{S}_A : \sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}) = \sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^*) \right\} \tag{3.47}$$

$$\mathcal{S}_{\text{lar}} = \left\{ \boldsymbol{S} \in \mathcal{S}_A : \sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}) > \sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^*) \right\} \tag{3.48}$$

where $\tilde{g}_k$ is given by (3.23) and $\mathcal{S}_A$ is as defined in Proposition 3.3.1. Note that each $\boldsymbol{S} \in \mathcal{S}_A$ must belong to either $\mathcal{S}_{\text{equ}}$ or $\mathcal{S}_{\text{lar}}$ due to the definition of $\boldsymbol{S}^*$ in (3.28). Let $\{v_l, \boldsymbol{S}_l\}_{l=1}^{K}$ denote the optimal solution of (3.20)-(3.22). Then, it is proved that $\boldsymbol{S}_l \in \mathcal{S}_{\text{equ}}$ must hold for all $l \in \{1, \ldots, K\}$ since it would otherwise lead to a scenario in which the optimal solution of (3.20)-(3.22), P_{DR}, could not achieve the lower bound in (3.27) as shown below:

$$\begin{aligned} \mathrm{P}_{\mathrm{DR}} = \max_{k \in \{1,\ldots,K\}} \sum_{l=1}^{K} v_l \tilde{g}_k(\boldsymbol{S}_l) &\geq \frac{1}{K} \sum_{l=1}^{K} v_l \sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}_l) \\ &> \frac{1}{K} \sum_{l=1}^{K} v_l \sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^*) = \mathrm{P}_{\mathrm{LB}} \, . \end{aligned} \tag{3.49}$$

Here, the strict inequality is obtained under the assumption that there exists $l \in \{1, \ldots, K\}$ such that $\boldsymbol{S}_l \notin \mathcal{S}_{\text{equ}}$ (i.e., $\boldsymbol{S}_l \in \mathcal{S}_{\text{lar}}$). However, as stated in Proposition 3.3.3, the lower bound must be achieved in the considered scenario. Therefore, (3.49) presents a contradiction, implying that $\boldsymbol{S}_l \in \mathcal{S}_{\text{equ}}$ must hold for all $l \in \{1, \ldots, K\}$.

Next, define set $\mathcal{S}_{\text{per}}$ as follows: $\mathcal{S}_{\text{per}} = \{\boldsymbol{S} \in \mathcal{S}_A : \boldsymbol{S}$ is a permutation of signal amplitude pairs in $\boldsymbol{S}^*\}$. From Remark 2, it is noted that the elements of $\mathcal{S}_{\text{per}}$ correspond to all possible $\boldsymbol{S} \in \mathcal{S}_A$ that satisfy $\sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}) = \sum_{k=1}^{K} \tilde{g}_k(\boldsymbol{S}^*)$; hence, $\mathcal{S}_{\text{per}} = \mathcal{S}_{\text{equ}}$. Then, based on the argument in the previous paragraph, it is concluded that the optimal solution of (3.20)-(3.22), $\{v_l, \boldsymbol{S}_l\}_{l=1}^{K}$, must satisfy $\boldsymbol{S}_l \in \mathcal{S}_{\text{per}}$ for all $l \in \{1,\ldots,K\}$. If the conditions in the proposition are satisfied (i.e., $\boldsymbol{S}^*$ is unique up to permutations of the signal amplitude pairs, which are the same except for one of them), there exist exactly K elements in $\mathcal{S}_{\text{per}}$, which correspond to the circular shifts of $\boldsymbol{S}^*$ by $2l-2$ elements for $l = 1,\ldots,K$; that is, $\mathcal{S}_{\text{per}} = \{\text{CS}_{2l-2}(\boldsymbol{S}^*),\ l = 1,\ldots,K\}$. In order to specify the randomization factors, $v_1,\ldots,v_K$, of the optimal solution in this scenario, define $\mathbf{v}$ as $\mathbf{v} = [v_1\ \cdots v_K]^T$ and $\mathbf{G}$ as a $K \times K$ matrix with its element in row k and column l being equal to $\tilde{g}_k(\text{CS}_{2l-2}(\boldsymbol{S}^*))$.[11] Then, based on Proposition 3.3.1, the optimal weights must satisfy

$$\mathbf{G}\mathbf{v} = \mathbf{p}_{\text{LB}} \quad \text{and} \quad \mathbf{1}^T\mathbf{v} = 1\ , \tag{3.50}$$

where $\mathbf{1} \triangleq [1\cdots 1]^T$ and $\mathbf{p}_{\text{LB}} \triangleq [\text{P}_{\text{LB}}\cdots \text{P}_{\text{LB}}]^T$ with $\text{P}_{\text{LB}} = \frac{1}{K}\sum_{k=1}^{K}\tilde{g}_k(\boldsymbol{S}^*)$ as in (3.27). Note that each element of $\mathbf{G}\mathbf{v}$ corresponds to the average error probability of a user, which should be equal to P_{LB}, since the lower bound in (3.27) is achieved, i.e., $\text{P}_{\text{DR}} = \text{P}_{\text{LB}}$, in this scenario (see the achievability condition in Proposition 3.3.1). It can be shown that $\mathbf{G}$ is a circulant matrix [59] based on the following lemma:

Lemma 3.6.1. *Under the conditions in Proposition 3.3.3,* $\tilde{g}_k(\text{CS}_{2l-2}(\boldsymbol{S}^*)) = \tilde{g}_j(\text{CS}_{2m-2}(\boldsymbol{S}^*))$ *if* $(l-k)_{\text{mod}\,K} = (m-j)_{\text{mod}\,K}$ *for* $k,l,j,m \in \{1,\ldots,K\}$.

Proof: From (3.23), $\tilde{g}_k(\text{CS}_{2l-2}(\boldsymbol{S}^*))$ can be expressed as

$$\tilde{g}_k(\text{CS}_{2l-2}(\boldsymbol{S}^*)) = 0.5\int_{-\infty}^{\infty}\min\left\{p_0^{(k)}(y|\,\text{CS}_{2l-2}(\boldsymbol{S}^*))\,,p_1^{(k)}(y|\,\text{CS}_{2l-2}(\boldsymbol{S}^*))\right\}dy \tag{3.51}$$

where $p_{i_k}^{(k)}$ is given by (3.44) under the conditions in Proposition 3.3.3. From (3.44) and (3.51), it is observed that $\tilde{g}_k(\text{CS}_{2l-2}(\boldsymbol{S}^*))$ and $\tilde{g}_j(\text{CS}_{2m-2}(\boldsymbol{S}^*))$ are equal if the k^{th} signal amplitude pair in $\text{CS}_{2l-2}(\boldsymbol{S}^*)$ is the same as the j^{th} signal

[11] Note that the elements of $\mathbf{G}$ are strictly positive based on (3.23) and (3.44).

amplitude pair in $\mathrm{CS}_{2m-2}(\boldsymbol{S}^*)$. Since the k^{th} and the j^{th} signal amplitude pairs in $\mathrm{CS}_{2l-2}(\boldsymbol{S}^*)$ and $\mathrm{CS}_{2m-2}(\boldsymbol{S}^*)$, respectively, become the same for $(l-k)_{\bmod K} = (m-j)_{\bmod K}$ due to the nature of the circular shift operator, $\tilde{g}_k(\mathrm{CS}_{2l-2}(\boldsymbol{S}^*)) = \tilde{g}_j(\mathrm{CS}_{2m-2}(\boldsymbol{S}^*))$ is obtained for $(l-k)_{\bmod K} = (m-j)_{\bmod K}$, where $k, l, j, m \in \{1, \dots, K\}$. $\square$

In addition to being a circulant matrix, $\mathbf{G}$ also has the property that its elements in each row are either all the same or the same except for one of them under the second condition in the proposition (i.e., when the signal amplitude pairs in $\boldsymbol{S}^*$ are the same except for one of them). This observation follows directly from (3.23) and (3.44). Therefore, one of the rows of $\mathbf{G}$, say the first one, is in the form of $[a \ b \cdots b]$, and the other rows are the circular shifts of this row in such a way that $\mathbf{G}$ is a circulant matrix. First consider the case in which $a \neq b$. Then, it is concluded that $\mathbf{G}$ is nonsingular since its eigenvalues are all nonzero. (In particular, one eigenvalue is $a + (K-1)b$ and the remaining ones are $a - b$.) Hence, there exists a unique solution of (3.50). Based on the fact that $\frac{1}{K}\sum_{l=1}^{K} \tilde{g}_k(\mathrm{CS}_{2l-2}(\boldsymbol{S}^*)) = \frac{1}{K}\sum_{j=1}^{K} \tilde{g}_j(\boldsymbol{S}^*)$ for each $k \in \{1, \dots, K\}$ (which can be verified from (3.23) and (3.44)), the unique solution of (3.50) is obtained as $\mathbf{v} = \left[\frac{1}{K} \cdots \frac{1}{K}\right]^T$. For $a = b$, all the elements of $\mathbf{G}$; hence, all the $\tilde{g}_k(\mathrm{CS}_{2l-2}(\boldsymbol{S}^*))$ terms, are the same. Therefore, no improvement is achieved via detector randomization in that scenario, and the optimal solution can be achieved by employing $\boldsymbol{S}^*$ all the time. Hence, this trivial scenario is excluded as stated at the beginning of Proposition 3.3.4; that is, $a = b$ does not hold for scenarios considered in the proposition.

In order to prove the necessity of the conditions in the proposition, first assume that $\boldsymbol{S}^*$ is not unique up to permutations of signal amplitude pairs. Then, a different solution can be obtained for each distinct $\boldsymbol{S}^*$ as described above. Namely, a distinct solution is calculated as in (3.42) for each $\boldsymbol{S}^*$. Therefore, the solution is not unique if the first condition in Proposition 3.3.4 is not satisfied.

Next, assume that $\boldsymbol{S}^*$ is unique up to permutations of signal amplitude pairs but it does not satisfy the second condition in the proposition; that is, there are at least three distinct signal amplitude pairs in $\boldsymbol{S}^*$ or two distinct signal amplitude pairs each with multiple repetitions. Then, the permutations of the signal amplitude pairs in $\boldsymbol{S}^*$ result in more than K different signal vectors; i.e., there exist more than K elements in set $\mathcal{S}_{\mathrm{per}}$, which is as defined above. (In particular, if there exist N_{p} distinct signal amplitude pairs in $\boldsymbol{S}^*$, each of which has $R_1, \ldots, R_{N_{\mathrm{p}}}$ repetitions, respectively, then there are $K!/(R_1! \cdots R_{N_{\mathrm{p}}}!)$ different permutations of signal amplitude pairs; i.e., $|\mathcal{S}_{\mathrm{per}}| = K!/(R_1! \cdots R_{N_{\mathrm{p}}}!)$.) In this case, there exist at least two distinct signal vectors $\boldsymbol{S}^*_{x_1}$ and $\boldsymbol{S}^*_{x_2}$, which are not circular shifts of each other. Then, the circular shifts of $\boldsymbol{S}^*_{x_1}$ and $\boldsymbol{S}^*_{x_2}$ can be employed in order to obtain two distinct solutions based on (3.42). Hence, it is concluded that the solution is not unique if the second condition in the proposition is not satisfied. $\square$

Chapter 4

Optimal Signaling and Detector Design for $M-$ary Communications Systems in the Presence of Multiple Additive Noise Channels

This chapter is organized as follows. In Section 4.1, the optimal signaling and detection problem is formulated in the presence of multiple additive noise channels under an average transmit power constraint, and the form of the solution to this optimization problem is obtained. Numerical examples are presented in Section 4.2, which is followed by some concluding remarks in Section 4.3.

4.1 Stochastic Signaling and Channel Switching

Consider an M-ary communications system, in which the information can be conveyed between the transmitter and receiver over K additive non-varying and independent noise channels as illustrated in Figure 4.1. The transmitter is allowed to switch or time share among these K channels to improve the correct decision performance at the receiver. A relay at the transmitter controls access to the channels so that only one of the channels can be used for symbol transmission at any given time. Furthermore, a *stochastic signaling* approach is adopted by treating the signal transmitted from each channel for each information symbol as a random vector instead of a constant value [4, 6]. In other words, the transmitter can perform randomization of signal values for each information symbol, which also corresponds to a form of constellation randomization [5, 15, 16]. The transmitter and the receiver are assumed to be synchronized so that the receiver knows which channel is currently in use, and employs the optimal decision rule for the corresponding channel and the stochastic signaling scheme. In practice, this assumption can be realized by employing a communications protocol that allocates the first $N_{s,1}$ symbols in the payload for channel 1, the next $N_{s,2}$ symbols in the payload for channel 2, and so on. The information on the numbers of symbols for different channels can be included in the header of a communications packet [6].

Multiple channels can be available between a transmitter and a receiver, for example, in cognitive radio systems, where secondary users sense the spectrum in order to determine available frequency bands for communications [60, 61]. In the presence of multiple available frequency bands between a transmitter-receiver pair in a cognitive radio system (see, e.g., [62]), channel switching can be performed in order to improve the error performance of the secondary system. Therefore, one application of the scenario in Figure 4.1 can be the communications of secondary users in a cognitive radio system.

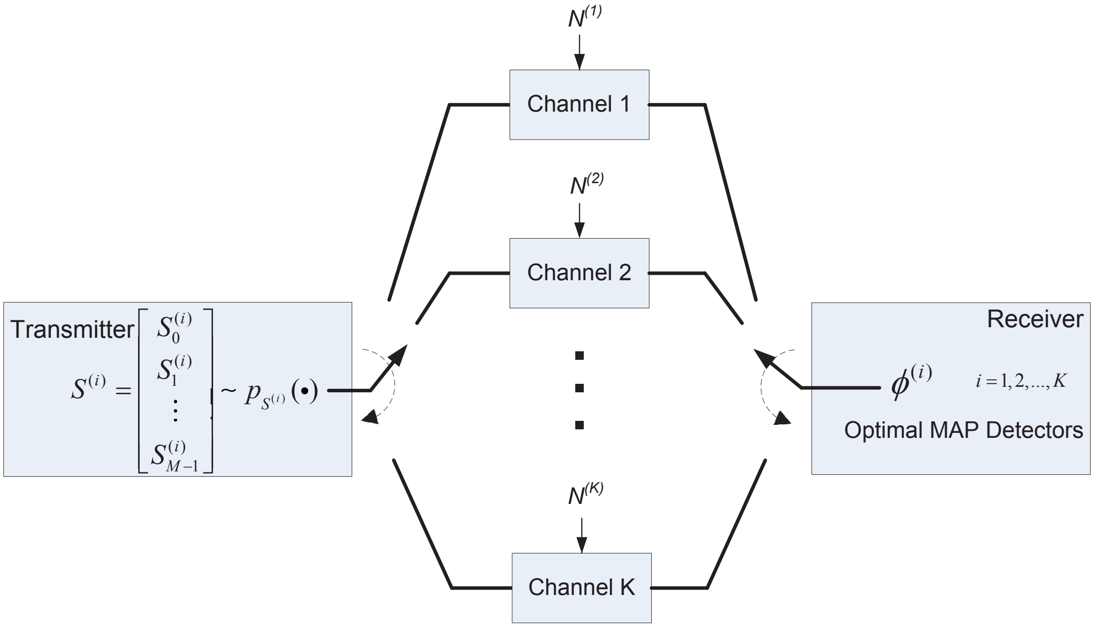

Figure 4.1: M-ary communications system that employs stochastic signaling and channel switching.

As pointed out in [2], for a binary-valued scalar communications system that employs antipodal signaling and the corresponding optimal MAP detector at the receiver, error probability is a nonincreasing convex function of the signal-to-noise ratio (SNR) when the channel has a continuously differentiable unimodal noise PDF with a finite variance. The more general case of arbitrary signal constellations is investigated in [3] by concentrating on the maximum likelihood (ML) detection over additive white Gaussian noise (AWGN) channels. The symbol error rate (SER) is shown to be always convex in SNR for 1-D and 2-D constellations, and also for higher dimensional constellations at high SNR regime. As a result, it is impossible to improve the error performance of an optimal detector via stochastic signaling under an average transmit power constraint in the above mentioned cases due to the convexity of the error probability. On the other hand, nonconvexity can be observed at low to intermediate SNRs in the presence of multimodal noise and even unimodal (including Gaussian) noise for high dimensional constellations.[1] As an example, it is reported in [4] and [5] that employing stochastic signaling; that is, modeling signals for different symbols as random variables instead of deterministic quantities, can provide significant performance improvements under Gaussian mixture noise. Motivated by this observation, we consider additive noise channels with generic PDFs and aim to obtain the optimal signaling and detection strategy when multiple channels are available for symbol transmission and stochastic signaling can be performed over each channel. In this scenario, the noisy observation vector $\mathbf{Y}$ received by the detector corresponding to the ith channel can be modeled as follows.

$$\mathbf{Y} = \mathbf{S}_j^{(i)} + \mathbf{N}^{(i)}\,, \quad j \in \{0, 1, \ldots, M-1\} \text{ and } i \in \{1, \ldots, K\}\,, \tag{4.1}$$

where $\mathbf{S}_j^{(i)}$ represents the N-dimensional signal vector transmitted for symbol j over channel i, and $\mathbf{N}^{(i)}$ is the noise component of the ith channel that is independent of $\mathbf{S}_j^{(i)}$ and all the noise components of the remaining channels. It should be emphasized that $\mathbf{S}_j^{(i)}$ is modeled as a random vector to employ stochastic signaling. Also, the prior probabilities of the symbols, denoted by $\pi_0, \pi_1, \ldots, \pi_{M-1}$, are assumed to be known. The vector channel model given above provides the discrete-time equivalent representation of a continuous-time system that processes the received signal by an orthonormal set of linear filters, samples the output of each filter once per symbol interval and concatenates the sampled values into a vector, thereby capturing the effects of modulator, additive noise channel and receiver front-end processing on the noisy observation signal. The resulting digital signal vector is fed to the designated detector to finalize the demodulation task. In addition, although the signal model in (4.1) is in the form of a simple additive noise channel, it is sufficient to incorporate various effects such as thermal noise, multiple-access interference, and jamming [2]. It is also valid in the case of flat-fading channels assuming perfect channel estimation

[1] Non-Gaussian and multimodal noise distributions are observed in some practical systems due to effects such as interference and jamming [21, 63, 64].

[4]. Note that the probability distribution of the noise component in (4.1) is not necessarily Gaussian since it is modeled to include the effects of interference and jamming as well. Hence, the noise component can have a significantly different probability distribution from the Gaussian distribution [21, 63, 64].

The receiver uses the observation in (4.1) in order to determine the transmitted information symbol. For that purpose, a generic decision rule (detector) is considered for each channel making a total of K detectors getting utilized at the receiver. That is, for a given observation vector $\mathbf{Y} = \mathbf{y}$, the detector of the ith channel $\phi^{(i)}(\mathbf{y})$ can be characterized as

$$\phi^{(i)}(\mathbf{y}) = j\,, \quad \text{if } \mathbf{y} \in \Gamma_j^{(i)}\,, \tag{4.2}$$

for $j \in \{0, 1, \dots, M-1\}$, where $\Gamma_0^{(i)}, \Gamma_1^{(i)}, \dots, \Gamma_{M-1}^{(i)}$ are the decision regions (i.e., a partition of the observation space $\mathbb{R}^N$) for the detector of the ith channel [44]. The transmitter and the receiver can switch between these K channels in any manner in order to optimize the probability of error performance. Let v_i denote the probability that channel i is selected for a given symbol transmission by the communications system. In the remainder of this study, v_i is called the *channel switching factor* for channel i, where $\sum_{i=1}^{K} v_i = 1$ and $v_i \geq 0$ for $i = 1, \dots, K$. In the context of time sharing, the transmitter and the receiver communicate over channel i for $100v_i$ percent of the time.

The aim of this study is to *jointly* optimize the channel switching strategy $(v_1, \dots, v_K)$, stochastic signals, and detectors in order to achieve the minimum average probability of error, or equivalently, the maximum average probability of correct decision. The average probability of correct decision can be expressed as $\mathrm{P_c} = \sum_{i=1}^{K} v_i\, \mathrm{P_c^{(i)}}$, where $\mathrm{P_c^{(i)}}$ represents the corresponding probability of correct decision for channel i under M-ary signaling; that is

$$\mathrm{P_c^{(i)}} = \sum_{j=0}^{M-1} \pi_j \int_{\Gamma_j^{(i)}} p_j^{(i)}(\mathbf{y})\, \mathrm{d}\mathbf{y} \tag{4.3}$$

for $i = 1, 2, \ldots, K$, with $p_j^{(i)}(\mathbf{y})$ denoting the conditional PDF of the observation when the jth symbol is transmitted over the ith channel. Since stochastic signaling is considered, $\mathbf{S}_j^{(i)}$ in (4.1) is modeled as a random vector. Recalling that the signals and the noise are independent, the conditional PDF of the observation can be calculated as $p_j^{(i)}(\mathbf{y}) = \int_{\mathbb{R}^N} p_{\mathbf{S}_j^{(i)}}(\mathbf{x})\, p_{\mathbf{N}^{(i)}}(\mathbf{y}-\mathbf{x})\, \mathrm{d}\mathbf{x} = \mathbb{E}\left\{p_{\mathbf{N}^{(i)}}\left(\mathbf{y} - \mathbf{S}_j^{(i)}\right)\right\}$, where the expectation is over the PDF of $\mathbf{S}_j^{(i)}$. Then, the average probability of correct decision can be expressed as

$$\mathrm{P_c} = \sum_{i=1}^{K} v_i \left(\sum_{j=0}^{M-1} \int_{\Gamma_j^{(i)}} \pi_j\, \mathbb{E}\left\{p_{\mathbf{N}^{(i)}}\left(\mathbf{y} - \mathbf{S}_j^{(i)}\right)\right\} \mathrm{d}\mathbf{y} \right) . \tag{4.4}$$

In practical systems, there is a constraint on the average power emitted from the transmitter. Under the framework of stochastic signaling and channel switching, this constraint on the average power can be expressed in the following form [44].

$$\sum_{i=1}^{K} v_i \left(\sum_{j=0}^{M-1} \pi_j\, \mathbb{E}\left\{\left\|\mathbf{S}_j^{(i)}\right\|_2^2\right\} \right) \leq \mathrm{A} , \tag{4.5}$$

where A denotes the average power limit.

In this study, we primarily concentrate on obtaining the optimal signaling and detection strategy in terms of the correct decision probability for an M-ary communications system in the presence of multiple channels. The novelty of the problem introduced here arises from the following two aspects: (i) signals transmitted over each channel corresponding to different symbols are modeled as random vectors subject to an overall average power constraint, (ii) no restrictions are imposed on the noise PDFs of the channels available for switching, and (iii) optimal detectors are designed jointly with the optimal signaling and switching strategies. This formulation, in turn translates into a design problem over the channel switching factors $\{v_i\}_{i=1}^K$, channel specific signal PDFs employed at the transmitter $\left\{p_{\mathbf{S}_0^{(i)}}, p_{\mathbf{S}_1^{(i)}}, \ldots, p_{\mathbf{S}_{M-1}^{(i)}}\right\}_{i=1}^K$, and the corresponding optimal detectors used at the receiver $\left\{\phi^{(i)}\right\}_{i=1}^K$. Stated more formally, the aim is to solve the following optimization problem.

$$\begin{aligned}
\max_{\left\{\phi^{(i)}, v_i, p_{\mathbf{S}_0^{(i)}}, p_{\mathbf{S}_1^{(i)}}, \dots, p_{\mathbf{S}_{M-1}^{(i)}}\right\}_{i=1}^{K}} \quad & \sum_{i=1}^{K} v_i \left(\sum_{j=0}^{M-1} \int_{\Gamma_j^{(i)}} \pi_j \, \mathbb{E}\left\{ p_{\mathbf{N}^{(i)}}\left(\mathbf{y} - \mathbf{S}_j^{(i)}\right) \right\} \mathrm{d}\mathbf{y} \right) \\
\text{subject to} \quad & \sum_{i=1}^{K} v_i \left(\sum_{j=0}^{M-1} \pi_j \, \mathbb{E}\left\{ \left\| \mathbf{S}_j^{(i)} \right\|_2^2 \right\} \right) \leq \mathrm{A}\,, \\
& \sum_{i=1}^{K} v_i = 1\,, \quad v_i \geq 0\,, \quad \forall\, i \in \{1, 2, \dots, K\}\,. \quad (4.6)
\end{aligned}$$

Included in the above statement are the implicit assumptions stating that each $p_{\mathbf{S}_j^{(i)}}(\cdot)$ should represent a PDF. Therefore, $p_{\mathbf{S}_j^{(i)}}(\mathbf{x}) \geq 0$, $\forall \mathbf{x} \in \mathbb{R}^N$, and $\int_{\mathbb{R}^N} p_{\mathbf{S}_j^{(i)}}(\mathbf{x})\, \mathrm{d}\mathbf{x} = 1$ are required $\forall\, j \in \{0, 1, \dots, M-1\}$ and $\forall\, i \in \{1, \dots, K\}$.

The signals for all the M symbols that are transmitted over channel i can be expressed as the elements of a random vector as follows: $\mathbf{S}^{(i)} \triangleq \left[\mathbf{S}_0^{(i)}\, \mathbf{S}_1^{(i)}\, \dots\, \mathbf{S}_{M-1}^{(i)}\right] \in \mathbb{R}^{MN}$, where $\mathbf{S}_j^{(i)}$'s are N-dimensional row vectors $\forall j \in \{0, 1, \dots, M-1\}$. More explicitly, each realization of $\mathbf{S}^{(i)}$ represents a signal constellation for M-ary symbol transmission in an N-dimensional space. Then, the optimization problem in (4.6) can be expressed in a more compact form as follows:

$$\begin{aligned}
\max_{\left\{\phi^{(i)}, v_i, p_{\mathbf{S}^{(i)}}\right\}_{i=1}^{K}} \quad & \sum_{i=1}^{K} v_i \, \mathbb{E}\left\{G_i(\mathbf{S}^{(i)})\right\} \\
\text{subject to} \quad & \sum_{i=1}^{K} v_i \, \mathbb{E}\left\{H(\mathbf{S}^{(i)})\right\} \leq \mathrm{A}\,, \quad \sum_{i=1}^{K} v_i = 1\,, \quad v_i \geq 0\,, \quad \forall\, i \in \{1, 2, \dots, K\}\,,
\end{aligned} \tag{4.7}$$

where $G_i(\mathbf{S}^{(i)}) = \sum_{j=0}^{M-1} \int_{\Gamma_j^{(i)}} \pi_j \, p_{\mathbf{N}^{(i)}} \left(\mathbf{y} - \mathbf{S}_j^{(i)}\right) \mathrm{d}\mathbf{y}$, $H(\mathbf{S}^{(i)}) = \sum_{j=0}^{M-1} \pi_j \left\|\mathbf{S}_j^{(i)}\right\|_2^2$, and each expectation is taken with respect to $p_{\mathbf{S}^{(i)}}(\cdot)$, which denotes the PDF of the signal constellation employed for symbol transmission over channel i. Specifically, $G_i(\mathbf{s}^{(i)})$ represents the probability of correct decision when the signal constellation represented by the deterministic vector $\mathbf{s}^{(i)}$ is used for the transmission of M symbols over the additive noise channel i and the corresponding detector $\phi^{(i)}$ is employed at the receiver. Then, $\mathbb{E}\{G_i(\mathbf{S}^{(i)})\}$ can be interpreted as the probability of correct decision for a generic stochastic signaling scheme over channel i. The exact number of signal constellations employed by this scheme is determined by the number of distinct values that the random vector $\mathbf{S}^{(i)}$ can take. The expression for $H(\cdot)$ is the same irrespective of which channel is used, and an explicit reference to the channel number as in the subscript of $G_i(\cdot)$ is not necessary.

Let $\mathrm{P}_{\mathrm{c}}^{\dagger}$ denote the maximum average probability of correct decision obtained as the solution of the optimization problem in (4.7). To provide a simpler formulation of this problem, an upper bound on $\mathrm{P}_{\mathrm{c}}^{\dagger}$ will be derived first, and then the achievability of that bound will be investigated.

Suppose that $G(\mathbf{x})$ denotes the maximum of the probabilities of correct decision when the deterministic signal constellation $\mathbf{x}$ is used for the transmission of M symbols over the additive noise channels $i = 1, 2, \ldots, K$ and the corresponding detectors for all K channels are employed at the receiver. That is, $G(\mathbf{x}) = \max\left\{G_i(\mathbf{x}) : i = 1, 2, \ldots, K \text{ and } \mathbf{x} \in \mathbb{R}^{MN}\right\}$, from which $G(\mathbf{x}) \geq G_i(\mathbf{x})$ follows $\forall\, i \in \{1, 2, \ldots, K\}$ and $\forall\, \mathbf{x} \in \mathbb{R}^{MN}$. This inequality can be applied to the objective function in (4.7) to obtain a new optimization problem that provides an upper bound on the solution of the optimization problem in (4.7) as follows.

$$\begin{aligned} &\max_{\left\{\phi^{(i)}, v_i, p_{\mathbf{S}^{(i)}}\right\}_{i=1}^{K}} \quad \sum_{i=1}^{K} v_i\, \mathbb{E}\left\{G(\mathbf{S}^{(i)})\right\} \\ &\text{subject to} \quad \sum_{i=1}^{K} v_i\, \mathbb{E}\left\{H(\mathbf{S}^{(i)})\right\} \leq \mathrm{A}\,, \quad \sum_{i=1}^{K} v_i = 1\,, \quad v_i \geq 0\,, \quad \forall\, i \in \{1, 2, \ldots, K\}\,, \end{aligned} \tag{4.8}$$

where the expectations are taken with respect to $p_{\mathbf{S}^{(i)}}(\cdot)$'s. Note that by replacing $G_i(\mathbf{S}^{(i)})$ with $G(\mathbf{S}^{(i)})$, the reference to individual channels inside the expectation operator is dropped which will prove useful in the foregoing analysis.

Let $\mathrm{P}_\mathrm{c}^\star$ denote the maximum average probability of correct decision obtained as the solution to the optimization problem in (4.8). From the definition of function $G(\cdot)$, $\mathrm{P}_\mathrm{c}^\star \geq \mathrm{P}_\mathrm{c}^\dagger$ is always satisfied. In order to achieve further simplification of the problem in (4.8), define $p_{\breve{\mathbf{S}}}(\breve{\mathbf{s}}) \triangleq \sum_{i=1}^K v_i \, p_{\mathbf{S}^{(i)}}(\breve{\mathbf{s}})$, where $\breve{\mathbf{s}} \triangleq [\breve{\mathbf{s}}_0 \; \breve{\mathbf{s}}_1 \; \cdots \; \breve{\mathbf{s}}_{M-1}] \in \mathbb{R}^{MN}$, and $\breve{\mathbf{s}}_j$'s are N-dimensional row vectors $\forall j \in \{0, 1, \ldots, M-1\}$. Since $\sum_{i=1}^K v_i = 1$, $v_i \geq 0 \; \forall i$, and $p_{\mathbf{S}^{(i)}}(\cdot)$'s are valid PDFs on $\mathbb{R}^{MN}$, $p_{\breve{\mathbf{S}}}(\breve{\mathbf{s}})$ satisfies the conditions to be a PDF. Then, the optimization problem in (4.8) can be written in the following equivalent form.

$$\max_{p_{\breve{\mathbf{S}}},\, \{\phi^{(i)}\}_{i=1}^K} \quad \mathbb{E}\{G(\breve{\mathbf{S}})\} \quad \text{subject to} \quad \mathbb{E}\{H(\breve{\mathbf{S}})\} \leq \mathrm{A}\,, \tag{4.9}$$

where $G(\breve{\mathbf{s}}) = \max\big\{G_i(\breve{\mathbf{s}}) \, : \, i = 1, 2, \ldots, K \text{ and } \breve{\mathbf{s}} \in \mathbb{R}^{MN}\big\}$, and the expectations are taken with respect to $p_{\breve{\mathbf{S}}}(\cdot)$, which denotes the PDF of the signal constellation employed for transmission of symbols $\{0, 1, \ldots, M-1\}$.

In (4.9), $G(\breve{\mathbf{s}})$ represents the maximum of the probabilities of correct decision when the deterministic signal constellation $\breve{\mathbf{s}}$ is used for the transmission of M symbols over the additive noise channels $i = 1, 2, \ldots, K$ and the corresponding detectors are employed at the receiver. Then, $\mathbb{E}\{G(\breve{\mathbf{S}})\}$ can be interpreted as a randomization among channels with respect to the PDF $p_{\breve{\mathbf{S}}}(\cdot)$, where the probability of correct decision corresponding to each component of $p_{\breve{\mathbf{S}}}$ (i.e., for each signal constellation $\breve{\mathbf{s}}$ in the support of $p_{\breve{\mathbf{S}}}$) is maximized by transmitting it over the *most favorable channel* (i.e., the channel with the highest probability of correct decision for the given signal constellation $\breve{\mathbf{s}}$), and altogether they maximize the average probability of correct decision.

Optimization problems in the form of (4.9) have been investigated in various studies in the literature [4–9, 11, 12, 37, 65]. Assume that $G_i(\mathbf{s})$ in (4.7) is a continuous function $\forall\, i \in \{1, 2, \ldots, K\}$ and $\boldsymbol{a} \preceq \mathbf{s} \preceq \boldsymbol{b}$ where $\boldsymbol{a}$ and $\boldsymbol{b}$ are finite real vectors in $\mathbb{R}^{MN}$. Then, $G(\mathbf{s}) = \max\{G_i(\mathbf{s}) \, : \, i = 1, 2, \ldots, K\}$ is also continuous on $[\boldsymbol{a}, \boldsymbol{b}]$, and the optimal solution of (4.9) can be represented by a randomization of at most two signal constellations as a result of Carathéodory's theorem [54]; that is, $p_{\breve{\mathbf{S}}}^{\mathrm{opt}}(\breve{\mathbf{s}}) = \lambda\,\delta(\breve{\mathbf{s}}-\mathbf{s}_1) + (1-\lambda)\,\delta(\breve{\mathbf{s}}-\mathbf{s}_2)$. Therefore, the problem in (4.9) can be solved over such signal PDFs resulting in the following optimization problem.

$$\begin{aligned} \max_{\left\{\lambda,\,\mathbf{s}_1,\,\mathbf{s}_2,\,\left\{\phi^{(i)}\right\}_{i=1}^{K}\right\}} \quad & \lambda\, G(\mathbf{s}_1) + (1-\lambda)\, G(\mathbf{s}_2) \\ \text{subject to} \quad & \lambda\, H(\mathbf{s}_1) + (1-\lambda)\, H(\mathbf{s}_2) \ \leq \mathrm{A}\ , \quad \lambda \in [0,1] \end{aligned} \tag{4.10}$$

where $G(\mathbf{s}_k) = \max\left\{G_i(\mathbf{s}_k) \ :\ G_i(\mathbf{s}_k) = \sum_{j=0}^{M-1}\int_{\Gamma_j^{(i)}}\pi_j\ p_{\mathbf{N}^{(i)}}(\mathbf{y}-\mathbf{s}_{k,j})\,\mathrm{d}\mathbf{y} \text{ and } i = 1,\,2,\,\ldots\,,\,K\right\}$, $H(\mathbf{s}_k) = \sum_{j=0}^{M-1}\pi_j\left\|\mathbf{s}_{k,j}\right\|_2^2$, and $\mathbf{s}_k = [\,\mathbf{s}_{k,0}\ \ \mathbf{s}_{k,1}\ \cdots\ \mathbf{s}_{k,M-1}] \in \mathbb{R}^{MN}$ with $\mathbf{s}_{k,j}$ denoting the N-dimensional vector representing the jth symbol in constellation k. Therefore, optimal performance can be achieved by randomizing between at most two signal constellations, $\mathbf{s}_1$ and $\mathbf{s}_2$. From (4.10), it is deduced that the objective function is maximized under the specified constraints by either one of the following strategies:

1. transmitting exclusively over a single channel via deterministic signaling, i.e., $\lambda \in \{0,1\}$,

2. randomizing (time sharing) between two signal constellations over a single channel, i.e., $\lambda \in (0,1)$ and $\underset{i\in\{1,2,\ldots,K\}}{\arg\max}\ G_i(\mathbf{s}_1) = \underset{i\in\{1,2,\ldots,K\}}{\arg\max}\ G_i(\mathbf{s}_2)$,

3. switching (time sharing) between two channels and deterministic signaling over each channel, i.e., $\lambda \in (0,1)$ and $\underset{i\in\{1,2,\ldots,K\}}{\arg\max}\ G_i(\mathbf{s}_1) \neq \underset{i\in\{1,2,\ldots,K\}}{\arg\max}\ G_i(\mathbf{s}_2)$.

Three distinct cases mentioned above can also be grouped under two overlapping cases as follows:

1. randomizing between *at most* two signal constellations over a single channel,

2. switching between *at most* two channels and deterministic signaling over each channel.

It is noted that randomizing between *at most* two signal constellations over a single channel covers deterministic signaling since the former reduces to the latter for $\lambda \in \{0, 1\}$. Similarly, switching between *at most* two channels and deterministic signaling over each channel also reduces to deterministic signaling over a single channel when $\lambda \in \{0, 1\}$. This form is introduced because it provides an ease of notation in the following analysis.

The last step in the simplification of the optimization problem in (4.10) comes from an observation about the structure of optimal detectors. For a given channel i and the corresponding signaling scheme over the channel (deterministic or randomization between two signal constellations), the conditional probability of the observation $\mathbf{y}$ given that symbol j is transmitted can be expressed as

$$\begin{aligned} p_j^{(i)}(\mathbf{y}) &= \mathbb{E}\{p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{S}_j^{(i)})\} \\ &= \begin{cases} p_{\mathbf{N}^{(i)}}\big(\mathbf{y} - \mathbf{s}_j^{(i)}\big), & \text{if deterministic} \\ \lambda\, p_{\mathbf{N}^{(i)}}\big(\mathbf{y} - \mathbf{s}_{1,j}^{(i)}\big) + (1-\lambda)\, p_{\mathbf{N}^{(i)}}\big(\mathbf{y} - \mathbf{s}_{2,j}^{(i)}\big), & \text{if randomized} \end{cases} . \end{aligned} \tag{4.11}$$

When deciding among M symbols based on observation $\mathbf{y}$ at detector i, the MAP decision rule selects symbol j if $j = \underset{l \in \{0,1,\ldots,M-1\}}{\arg\max}\ \pi_l\, p_l^{(i)}(\mathbf{y})$, and it maximizes the probability of correct decision [44]. Therefore, it is not necessary to search over all decision rules in (4.10); only the MAP decision rule should be determined for the detector of each channel and its corresponding probability of correct decision should be considered. The probability of correct decision for a generic decision rule is given in (4.3). Using the decision regions corresponding to the MAP detector, i.e., $\Gamma_j^{(i)} = \{\mathbf{y} \in \mathbb{R}^N \mid \pi_j\, p_j^{(i)}(\mathbf{y}) \geq \pi_l\, p_l^{(i)}(\mathbf{y}),\ \forall l \neq j\}$, the average probability of correct decision for ith channel becomes

$$
\begin{aligned}
\mathrm{P}_{\mathrm{c,\,MAP}}^{(i)} &= \int_{\mathbb{R}^N} \max_{j\,\in\{0,\,1,\,\ldots\,,\,M-1\}} \left\{\pi_j \; p_j^{(i)}(\mathbf{y})\right\} \,\mathrm{d}\mathbf{y} \\
&= \int_{\mathbb{R}^N} \max_{j\,\in\{0,\,1,\,\ldots\,,\,M-1\}} \left\{\pi_j \; \mathbb{E}\left\{p_{\mathbf{N}^{(i)}}\big(\mathbf{y}-\mathbf{S}_j^{(i)}\big)\right\}\right\} \,\mathrm{d}\mathbf{y} \\
&= \begin{cases} \displaystyle\int_{\mathbb{R}^N} \max_{j\,\in\{0,\,1,\,\ldots\,,\,M-1\}} \left\{\pi_j \; p_{\mathbf{N}^{(i)}}\big(\mathbf{y}-\mathbf{s}_j^{(i)}\big)\right\} \mathrm{d}\mathbf{y}, & \text{if deterministic} \\
\displaystyle\int_{\mathbb{R}^N} \max_{j\,\in\{0,\,1,\,\ldots\,,\,M-1\}} \Big\{\pi_j \; \Big(\lambda\, p_{\mathbf{N}^{(i)}}\big(\mathbf{y}-\mathbf{s}_{1,j}^{(i)}\big) & \\
\qquad\qquad\qquad + (1-\lambda)\, p_{\mathbf{N}^{(i)}}\big(\mathbf{y}-\mathbf{s}_{2,j}^{(i)}\big)\Big)\Big\}\mathrm{d}\mathbf{y}, & \text{if randomized} \end{cases}
\end{aligned}
\tag{4.12}
$$

Below, more explicit forms of the optimization problem stated in (4.10) are given for all possible scenarios mentioned previously.

1. Transmitting exclusively over a single channel via deterministic signaling:

In this case, a single channel is utilized exclusively, and the transmitted signal for each symbol is deterministic, i.e., a fixed signal constellation is employed for symbol transmission over the channel. Without loss of generality, channel i is considered. The optimization problem in (4.10) becomes

$$
\max_{\left\{\mathbf{s}^{(i)},\,\phi^{(i)}\right\}} \sum_{j=0}^{M-1} \int_{\Gamma_j^{(i)}} \pi_j \; p_{\mathbf{N}^{(i)}}\left(\mathbf{y}-\mathbf{s}_j^{(i)}\right) \,\mathrm{d}\mathbf{y} \;\; \text{subject to} \; \sum_{j=0}^{M-1} \pi_j \left\|\mathbf{s}_j^{(i)}\right\|_2^2 \;\leq \mathrm{A}\,. \tag{4.13}
$$

Using the result given in (4.12) for the deterministic case, the equivalent optimization problem can be written as follows.

$$
\max_{\mathbf{s}^{(i)}} \int_{\mathbb{R}^N} \max_{j\,\in\{0,\,1,\,\ldots\,,\,M-1\}} \left\{\pi_j \; p_{\mathbf{N}^{(i)}}\left(\mathbf{y}-\mathbf{s}_j^{(i)}\right)\right\} \mathrm{d}\mathbf{y} \;\;\text{subject to}\; \sum_{j=0}^{M-1} \pi_j \left\|\mathbf{s}_j^{(i)}\right\|_2^2 \;\leq \mathrm{A} \tag{4.14}
$$

2. Randomizing (time sharing) between at most two signal constellations over a single channel:

Similarly to the previous case, the transmission occurs over a single channel exclusively, but in this case the transmitted signal for each symbol is a randomization between at most two different signal vectors. Without loss of generality, channel i is considered. The optimization problem in (4.10) is expressed as follows.

$$\begin{aligned} \max_{\left\{\lambda,\, \mathbf{s}_1^{(i)},\, \mathbf{s}_2^{(i)},\, \phi^{(i)}\right\}} \quad & \lambda\, G_i(\mathbf{s}_1^{(i)}) + (1-\lambda)\, G_i(\mathbf{s}_2^{(i)}) \\ \text{subject to} \quad & \lambda\, H(\mathbf{s}_1^{(i)}) + (1-\lambda)\, H(\mathbf{s}_2^{(i)}) \ \leq \mathrm{A}\,, \quad \lambda \in [0,1] \end{aligned} \tag{4.15}$$

where $G_i(\mathbf{s}_k^{(i)}) = \sum_{j=0}^{M-1} \int_{\Gamma_j^{(i)}} \pi_j\, p_{\mathbf{N}^{(i)}}(\mathbf{y}-\mathbf{s}_{k,j}^{(i)})\, \mathrm{d}\mathbf{y}\,,\ H(\mathbf{s}_k) = \sum_{j=0}^{M-1} \pi_j \left\|\mathbf{s}_{k,j}^{(i)}\right\|_2^2$, and $k \in \{1,\, 2\}$. As stated earlier, it is assumed that a single detector is employed for each channel at the receiver. Using the result for randomized signaling case given in (4.12), the equivalent optimization problem can be written as

$$\begin{aligned} & \max_{\left\{\lambda,\, \mathbf{s}_1^{(i)},\, \mathbf{s}_2^{(i)}\right\}} \quad \int_{\mathbb{R}^N} \max_{j\,\in\{0,\,1,\,\ldots\,,\,M-1\}} \left\{\pi_j\ p_j^{(i)}(\mathbf{y})\right\}\, \mathrm{d}\mathbf{y} \\ & \text{subject to} \quad \lambda \left(\sum_{j=0}^{M-1} \pi_j \left\|\mathbf{s}_{1,j}^{(i)}\right\|_2^2\right) \ + \ (1-\lambda) \left(\sum_{j=0}^{M-1} \pi_j \left\|\mathbf{s}_{2,j}^{(i)}\right\|_2^2\right) \ \leq \mathrm{A}\,, \quad \lambda \in [0,1] \end{aligned} \tag{4.16}$$

where $p_j^{(i)}(\mathbf{y}) = \lambda\, p_{\mathbf{N}^{(i)}}\left(\mathbf{y}-\mathbf{s}_{1,j}^{(i)}\right) + (1-\lambda)\, p_{\mathbf{N}^{(i)}}\left(\mathbf{y}-\mathbf{s}_{2,j}^{(i)}\right)$. It is recalled that the optimization problem in (4.16) reduces to that of (4.14) when $\lambda \in \{0,1\}$.

3. Switching (time sharing) between at most two channels and deterministic signaling over each channel:

In this case, optimum performance is investigated while transmitting over at most two channels and the transmission over each channel is deterministic, i.e., a fixed signal constellation is employed for symbol transmission over each channel but the channels are switched in time. Without loss of generality, channels i and l are considered ($i \neq l$ and $i\,, l \in \{1, 2, \ldots, K\}$). The optimization problem in (4.10) takes the following form.

$$\begin{aligned} \max_{\left\{\lambda,\, \mathbf{s}^{(i)},\, \mathbf{s}^{(l)},\, \phi^{(i)},\, \phi^{(l)}\right\}} \quad & \lambda\, G_i(\mathbf{s}^{(i)}) + (1-\lambda)\, G_l(\mathbf{s}^{(l)}) \\ \text{subject to} \quad & \lambda\, H(\mathbf{s}^{(i)}) + (1-\lambda)\, H(\mathbf{s}^{(l)}) \ \leq \mathrm{A}\,, \quad \lambda \in [0,1] \end{aligned} \tag{4.17}$$

where $G_i(\mathbf{s}^{(i)}) = \sum_{j=0}^{M-1} \int_{\Gamma_j^{(i)}} \pi_j \; p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{s}_j^{(i)})\, \mathrm{d}\mathbf{y}$, $H(\mathbf{s}^{(i)}) = \sum_{j=0}^{M-1} \pi_j \left\|\mathbf{s}_j^{(i)}\right\|_2^2$, $G_l(\mathbf{s}^{(l)})$ and $H(\mathbf{s}^{(l)})$ are defined similarly by replacing i with l in the preceding equations. Since deterministic signaling is employed in each channel, the result given in (4.12) for the deterministic case should be applied for each channel. Then, an equivalent optimization problem can be written as

$$\begin{aligned} &\max_{\left\{\lambda,\, \mathbf{s}^{(i)},\, \mathbf{s}^{(l)}\right\}} && \lambda\, G_{i,\mathrm{MAP}}(\mathbf{s}^{(i)}) + (1-\lambda)\, G_{l,\mathrm{MAP}}(\mathbf{s}^{(l)}) \\ &\text{subject to} && \lambda\, H(\mathbf{s}^{(i)}) + (1-\lambda)\, H(\mathbf{s}^{(l)}) \;\leq \mathrm{A}\ , \quad \lambda \in [0,1] \end{aligned} \tag{4.18}$$

where $G_{i,\mathrm{MAP}}(\mathbf{s}^{(i)}) = \int_{\mathbb{R}^N} \max_{j\,\in\{0,\,1,\,\ldots\,,\,M-1\}} \left\{\pi_j \; p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{s}_j^{(i)})\right\} \mathrm{d}\mathbf{y}$, $H(\mathbf{s}^{(i)}) = \sum_{j=0}^{M-1} \pi_j \left\|\mathbf{s}_j^{(i)}\right\|_2^2$, $G_l(\mathbf{s}^{(l)})$ and $H(\mathbf{s}^{(l)})$ are defined similarly by replacing i with l in the respective equations.

It is noted that the optimization space is considerably reduced in (4.14), (4.16) and (4.18) compared to those in (4.13), (4.15) and (4.17), respectively since there is no need to search over the detectors in (4.14), (4.16) and (4.18).

In the rest of the analysis, only the second and third cases will be investigated since they cover deterministic signaling over a single channel as a special case. In the view of the above analysis, the solution of the optimization problem in (4.10) can be decomposed into two parts. First, randomizing between at most two signal constellations over a single channel is considered. Let $\mathrm{P}_{\mathrm{c,\,Opt}}^{(i)}$ be the solution of the optimization problem in (4.16) for ith channel; that is, $\mathrm{P}_{\mathrm{c,\,Opt}}^{(i)}$ denotes the maximum average probability of correct decision that can be achieved by stochastic signaling over channel i under the average power constraint. Secondly, switching between at most two channels with deterministic signaling over each channel is considered. Let $\mathrm{P}_{\mathrm{c,\,Opt}}^{(i,\,l)}$ be the solution of the optimization problem in (4.18) for channels i and l; that is, $\mathrm{P}_{\mathrm{c,\,Opt}}^{(i,\,l)}$ denotes the maximum average probability of correct decision that can be achieved by switching between channels i and l under the average power constraint. Then, the solution of the optimization problem in (4.10) can be obtained by solving the following set of optimization problems and computing their maximum.

$$\mathrm{P}_{\mathrm{c}}^{\mathrm{Stoc}} = \max\left\{\mathrm{P}_{\mathrm{c,\,Opt}}^{(i)} \;:\; i \in \{1,\,2,\,\ldots\,,\,K\}\right\} \tag{4.19}$$

$$\mathrm{P}_{\mathrm{c}}^{\mathrm{CS}} = \max\left\{\mathrm{P}_{\mathrm{c,\,Opt}}^{(i,\,l)} \;:\; i \in \{1,\,2,\,\ldots\,,\,K\},\, l \in \{1,\,2,\,\ldots\,,\,K\},\ \text{and } i < l\right\} \tag{4.20}$$

$$\mathrm{P}_{\mathrm{c}}^{\star} = \max\left\{\mathrm{P}_{\mathrm{c}}^{\mathrm{Stoc}},\, \mathrm{P}_{\mathrm{c}}^{\mathrm{CS}}\right\} \tag{4.21}$$

where the superscript Stoc denotes stochastic signaling over a single channel and CS abbreviates channel switching. Note that the preceding results are obtained without assuming any specific form on the noise PDFs $p_{\mathbf{N}^{(i)}}$'s of the channels present in the communications system. For example, when the noise PDFs on all the channels are different, the solution of the optimization problem is given by (4.21) without any further simplifications. Namely, $K(K+1)/2$ optimization problems must be solved in total to obtain the optimal set of parameters and the resulting performance score. In the cases where some channels share the same noise PDF, the results are still valid but the optimization sets given in (4.19) and (4.20) over which the maximum values are computed can be refined to avoid repeated computations of the same expressions.[2]

The following proposition states that the expressions in (4.19)-(4.21) provides the solution of the generic problem in (4.7).

Proposition 4.1.1. *The maximum average probabilities of correct decision achieved by the solutions of the optimization problems in* (4.7) *and* (4.21) *are equal, i.e.,* $\mathrm{P}_{\mathrm{c}}^{\dagger} = \mathrm{P}_{\mathrm{c}}^{\star}$.

Proof: First, consider the optimization problem in (4.7) when $K = 2$ channels are used, and deterministic signaling is employed for each channel, i.e., $p_{\mathbf{S}^{(1)}}(\mathbf{s}^{(1)}) = \delta(\mathbf{s}^{(1)} - \mathbf{s}_1)$ and $p_{\mathbf{S}^{(2)}}(\mathbf{s}^{(2)}) = \delta(\mathbf{s}^{(2)} - \mathbf{s}_2)$. Suppose also that

[2]Detector randomization as discussed in [6, 7] can also be analyzed using our framework. Specifically, it can be modeled by assuming that some channels have identical noise distributions. That is, each channel appears in the system model with a certain multiplicity.

the symbols transmitted over each channel are decoded using the MAP detector corresponding to that channel. In that case, (4.7) reduces to the optimization problem in (4.18); hence, (4.7) covers (4.18) as a special case. Secondly, consider the optimization problem in (4.7) when $K = 1$ channel is used, and a randomization between at most two signal constellations is employed, i.e., $p_{\mathbf{S}}(\mathbf{s}) = \lambda\,\delta(\mathbf{s}-\mathbf{s}_1) + (1-\lambda)\,\delta(\mathbf{s}-\mathbf{s}_2)$. Suppose also that a single MAP detector is employed at the receiver. Then, (4.7) reduces to the optimization problem in (4.16); hence, (4.7) covers (4.16) as a special case. Since both (4.16) and (4.18) are special cases of (4.7) for any choice of the channels $i \in \{1, 2, \ldots, K\}$, $l \in \{1, 2, \ldots, K\}$ and $i \neq l$, the maximum value of the objective function in (4.7) should be larger than or equal to the maximum given by (4.21). This, in turn, implies that $\mathrm{P}_\mathrm{c}^\dagger \geq \mathrm{P}_\mathrm{c}^\star$. On the other hand, the optimization problem in (4.7) has been replaced with the upper bound given in (4.8), the solution of which is shown to reduce to that given in (4.21); that is, $\mathrm{P}_\mathrm{c}^\dagger \leq \mathrm{P}_\mathrm{c}^\star$. Therefore, it is concluded that $\mathrm{P}_\mathrm{c}^\dagger = \mathrm{P}_\mathrm{c}^\star$. $\square$

Proposition 4.1.1 implies that the solution of the original optimization problem stated in (4.7), which considers the joint optimization of switching factors among channels, channel specific signal PDFs employed at the transmitter and the corresponding detectors used at the receiver, can be obtained as the solution of the much simpler optimization problem specified in (4.21). Formally, when multiple channels are available for signal transmission (i.e., $K \geq 2$), it is sufficient to either employ switching between two channels with deterministic signaling over each channel (i.e., there is no need to employ stochastic signaling over a channel to achieve the optimal solution while switching channels); or randomize between at most two signal constellations over a single channel, whichever results in the highest average probability of correct decision.

The solution of the optimization problem in (4.21) can be obtained via global optimization techniques (since it is a nonlinear nonconvex optimization problem in general due to arbitrary noise PDFs), or a convex relaxation approach as in [12] can be employed to obtain approximate solutions in polynomial time.

4.2 Numerical Results

In this section, numerical examples are presented to evaluate the performance of the proposed signaling strategies in the presence of multiple channels. A scalar binary communications system with equiprobable information symbols is considered and the average power limit is set to $\mathrm{A} = 1$. It is assumed that $K \geq 2$ channels are available between the transmitter and the receiver, and only one of them can be used for transmission at any given time. The following four strategies are considered for performance comparison.

Gaussian solution over the best channel: In this approach, antipodal signals $\{-\sqrt{\mathrm{A}}, \sqrt{\mathrm{A}}\}$ are transmitted for binary information symbols over the most favorable channel, i.e., the one that yields the highest probability of correct decision, and the corresponding MAP detector is employed at the receiver. Since deterministic antipodal signaling is optimal in the presence of Gaussian noise (not necessarily optimal for other types of noise), this approach is called Gaussian solution over the best channel.

Optimal deterministic solution over the best channel: In this scheme, the optimal deterministic signal constellation and the corresponding MAP decision rule are obtained to maximize the probability of correct decision in the absence of stochastic signaling and channel switching. K optimization problems in the form of (4.14) are solved and the most favorable channel is employed for symbol transmission.

Optimal stochastic solution over the best channel: This scheme employs a single MAP detector at the receiver and randomizes between at most two signal constellations. The optimization problem in (4.16) is solved for all K channels and the most favorable channel is selected for symbol transmission as shown in (4.19).

Optimal channel switching with deterministic signaling: In this scheme, switching is performed between at most two channels with deterministic signaling over each channel. $K(K-1)/2$ optimization problems in the form of (4.18) are solved and the most favorable channel pair is selected as shown in (4.20).

It should be noted that the maximum of the last two strategies constitute the solution to the optimal signaling and detector design problem in the presence of multiple channels, as stated in (4.21).

In the following numerical examples, it is assumed that the channel noise is modeled by a Gaussian mixture distribution [9, 12, 63, 64], which is represented by

$$p_{\mathrm{N}^{(i)}}(n) = \frac{1}{\sqrt{2\pi}\sigma_i L_i} \sum_{l=1}^{L_i} \exp\left\{ -\frac{\left(n - \mu_l^{(i)}\right)^2}{2\sigma_i^2} \right\} \tag{4.22}$$

for $i \in \{1 \dots K\}$, where L_i is the number of components in the mixture for channel i. As noted from (4.22), the components of the Gaussian mixture noise have the same weight $1/L_i$ and the same variance σ_i^2. For notational simplicity, the component means of the Gaussian mixture for channel i are collected in the vector $\boldsymbol{\mu}^{(i)} = \left[\mu_1^{(i)} \dots \mu_{L_i}^{(i)}\right]$. Based on (4.22), the average noise power of the ith channel can be calculated as $\mathbb{E}\left\{|\mathrm{N}^{(i)}|^2\right\} = \sigma_i^2 + \frac{1}{L_i}\|\boldsymbol{\mu}^{(i)}\|_2^2$, where $\|\boldsymbol{\mu}^{(i)}\|_2$ denotes the L_2 norm of vector $\boldsymbol{\mu}^{(i)}$.

First, we consider a scenario in which $K \geq 2$ identical channels (i.e., channels with the same noise PDF) are available; i.e., $\sigma_i = \sigma$, $L_i = L$, and $\boldsymbol{\mu}^{(i)} = \boldsymbol{\mu}$, $\forall i \in \{1 \dots K\}$, where $\boldsymbol{\mu} = [\mu_1 \dots \mu_L]$. Since identical channels are considered and at most two channels are required for the optimum solution as discussed in

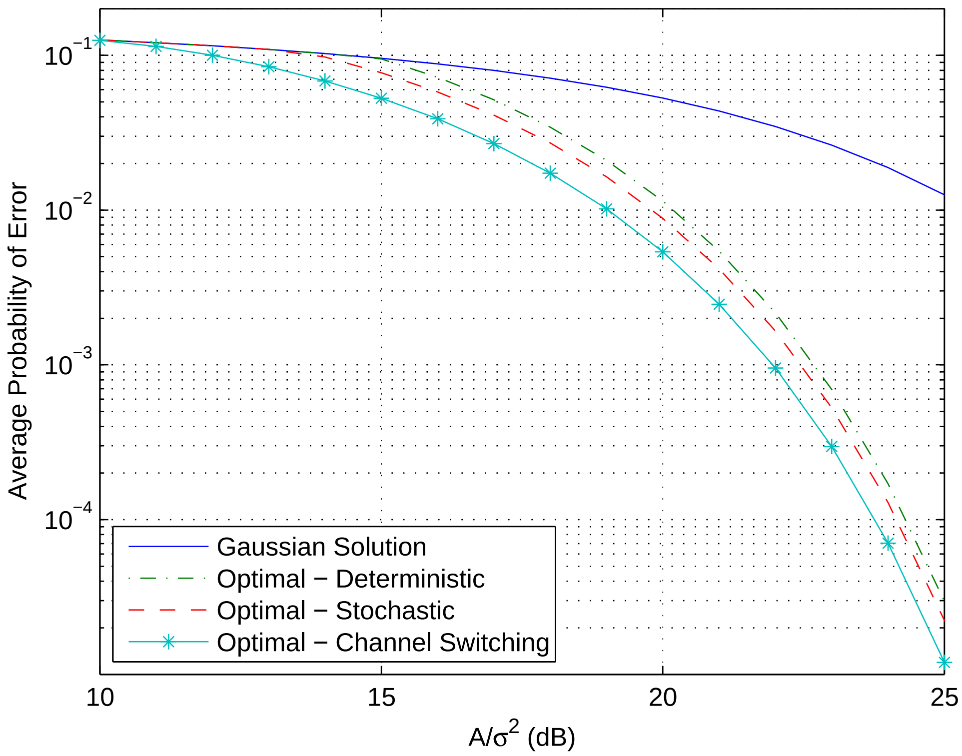

Figure 4.2: Average probability of error versus A/σ^2 for various strategies, where $L = 3$ and $\boldsymbol{\mu} = [-0.9\ 0\ 0.9]$ for the Gaussian mixture noise.

Section 4.1, K can be any number that is larger than or equal to 2. Hence, the results in this part are valid for all $K \geq 2$. In Figure 4.2, the average probabilities of error corresponding to the four strategies discussed above are plotted versus A/σ^2 for $L = 3$ and $\boldsymbol{\mu} = [-0.9\ 0\ 0.9]$. From the figure, it is observed that the Gaussian solution has the worst performance among all the approaches as expected since it is optimized for Gaussian noise and is not expected to achieve good performance in the presence of multimodal channel noise. When optimal deterministic signaling is employed, significant gains can be achieved over the Gaussian solution in this example. In addition, further improvements are possible when stochastic signaling is used instead of deterministic signaling. However, the best performance is achieved when switching is performed between two MAP detectors corresponding to two signal constellations. Since identical channels are considered in this example, channel switching can also be regarded as

Table 4.1: Optimal signal parameters for the scenario in Figure 4.2.

	Deterministic Sig.	Stochastic Signaling			Channel Switching		
A/σ^2 (dB)	$\mathbf{s}_1$	λ	$\mathbf{s}_{1,1}$	$\mathbf{s}_{2,1}$	λ	$\mathbf{s}_1^{(1)}$	$\mathbf{s}_1^{(2)}$
10	1	N/A	1	1	0.1533	0.7271	1.0418
15	0.7239	0.7885	0.7160	1.6783	0.4492	0.7060	1.1870
20	0.6904	0.7650	0.6894	1.6456	0.4254	0.6880	1.1790
25	0.6799	0.7482	0.6798	1.6120	0.3843	0.6796	1.1558

detector randomization via time-sharing for this scenario [6]. Furthermore, the performance of detector randomization is guaranteed to exceed that of stochastic signaling in the case of identical channels, which is also evident from Figure 4.2.[3]

In order to further investigate the results in Figure 4.2, the parameters for the proposed strategies are presented in Table 4.1 for some values of A/σ^2. Due to the symmetry of the Gaussian mixture noise, antipodal signaling is employed for binary communications. More explicitly, for optimal deterministic signaling, $\mathbf{s}_0$ and $\mathbf{s}_1$ denote the signals transmitted for information symbols 0 and 1, respectively, and we have $\mathbf{s}_0 = -\mathbf{s}_1$. For optimal stochastic signaling, the optimal signal for information symbol $i \in \{0, 1\}$ is expressed in the form of $p_{\mathbf{S}_i}(\mathbf{s}) = \lambda\,\delta(\mathbf{s}-\mathbf{s}_{1,i})+(1-\lambda)\,\delta(\mathbf{s}-\mathbf{s}_{2,i})$ with $\mathbf{s}_{1,0} = -\mathbf{s}_{1,1}$ and $\mathbf{s}_{2,0} = -\mathbf{s}_{2,1}$. Finally, the optimal channel switching solution employs the signal pair $\left\{-\mathbf{s}_1^{(1)},\ \mathbf{s}_1^{(1)}\right\}$ and the corresponding MAP detector with probability λ, and the signal pair $\left\{-\mathbf{s}_1^{(2)},\ \mathbf{s}_1^{(2)}\right\}$ and the corresponding MAP detector with probability $1-\lambda$. From Table 4.1, it is observed that all the solutions converge to the Gaussian solution as the noise variance increases. This is due to the fact that the Gaussian mixture noise approximates a unimodal PDF at high values of the variance for which the Gaussian solution is optimal. However, as the noise variance decreases (i.e., A/σ^2 increases), the multimodal nature of the noise PDF prevails and the best performance is achieved by the optimal channel switching solution.

[3] Additional examples are obtained for $\boldsymbol{\mu} = [-0.9\ -0.2\ 0.2\ 0.9]$, $\boldsymbol{\mu} = [-0.9\ -0.2\ 0\ 0.2\ 0.9]$, and $\boldsymbol{\mu} = [-1.2\ -0.6\ -0.1\ 0.1\ 0.6\ 1.2]$ as well, and similar observations to those for Figure 4.2 are made. The resulting figures are not presented since they are quite similar to Figure 4.2.

The results depicted in Figure 4.2 and Table 4.1 can also be verified by plotting the error probability of the optimal MAP detector as a function of the signal power in the presence of deterministic antipodal signaling, i.e., $\mathbf{s}_1 = -\mathbf{s}_0 = \mathbf{s}$. This is shown in Figure 4.3 for the channel characterized by the parameters $L = 3$, $\boldsymbol{\mu} = [-0.9\ 0\ 0.9]$ and $\mathrm{A}/\sigma^2 = 15$ dB, where $\mathrm{A} = 1$ as specified before. Due to multimodal noise, the error probability is a nonmonotonic and nonconvex function of the signal power [5, 7]. From Figure 4.3, it is seen that the optimal deterministic solution is obtained as $\mathbf{s}_1 = -\mathbf{s}_0 = \sqrt{0.524} = 0.7239$, which corresponds to the minimum value (0.0948) of the error probability curve for $\mathbf{s}^2 \leq 1$. The best performance is achieved by switching between two power levels 0.4984 and 1.409 using the corresponding antipodal signal pairs $\{-0.7060, 0.7060\}$ and $\{-1.1870, 1.1870\}$, which are in compliance with Table 4.1. Also, the switching factor λ can be calculated based on the average power limit, $\mathrm{A} = 1$, as follows: $0.4984\lambda + 1.409(1 - \lambda) = 1$, which yields $\lambda = 0.4492$ as in Table 4.1. It is observed from Figure 4.3 that switching between two MAP detectors can reduce the average probability of error down to nearly 0.05, which is indicated by the red circle in the figure.

Next, we consider a scenario in which all the channels have distinct noise PDFs. In this case, the best performance can be achieved by either the optimal channel switching with deterministic signaling approach or the optimal stochastic solution over the best channel approach. For the Gaussian mixture noise model in (4.22), it is assumed that $\sigma_i = \sigma$ and $L_i = L$, $\forall i \in \{1, \ldots, K\}$, and that the component means of the Gaussian mixture are chosen as

$$\boldsymbol{\mu}^{(i)} = \sqrt{E}\frac{\mathbf{v}_i}{\|\mathbf{v}_i\|_2} \tag{4.23}$$

for $i = 1, \ldots, K$, where E is a constant and $\mathbf{v}_i$'s are L-dimensional distinct vectors. It is noted that $\|\boldsymbol{\mu}^{(i)}\|_2^2 = E$. Hence, the average noise power is the same for all the channels. Namely, $\mathbb{E}\left\{|\mathrm{N}^{(i)}|^2\right\} = \sigma^2 + \frac{E}{L}$, $\forall i \in \{1, \ldots, K\}$. In Figure 4.4, the average probabilities of error for the four strategies are plotted versus A/σ^2 for

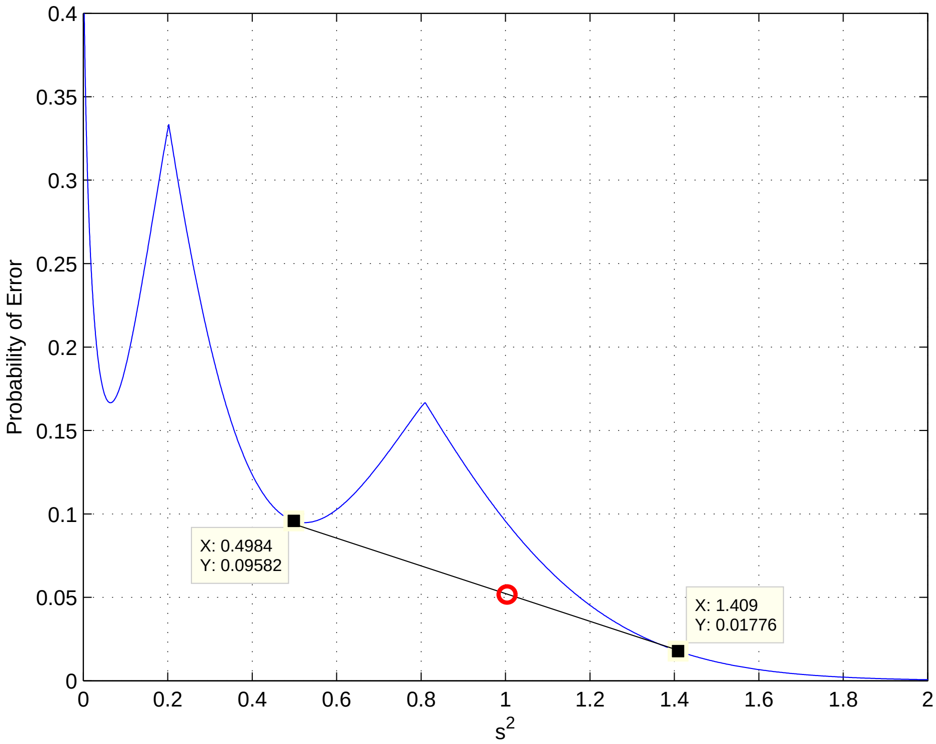

Figure 4.3: Error probability versus signal power $\mathbf{s}^2$ for the channel characterized by the parameters $L = 3$ and $\boldsymbol{\mu} = [-0.9\ 0\ 0.9]$ and $\mathrm{A}/\sigma^2 = 15$ dB (cf. Figure 4.2 and Table 4.1).

$K = 3$, $\mathbf{v}_1 = [-3\ -2\ 0\ 2\ 3]$, $\mathbf{v}_2 = [-4\ -3\ 0\ 3\ 4]$, $\mathbf{v}_3 = [-5\ -3\ 0\ 3\ 5]$, and $E = 3$. From Figure 4.4, it is concluded that the optimal channel switching strategy achieves the lowest average probability of error and the Gaussian solution has the worst performance over the whole range of A/σ^2 values.

The optimal parameters of the strategies in Figure 4.4 are shown for some values of A/σ^2 in Table 4.2. For the Gaussian solution and the optimal deterministic solution, the channel that results in the lowest probability of error is indicated in the first column of the respective area in the table and the second column specifies the scalar signal value employed for the transmission of information symbol 1. Again, antipodal signals are considered for symbol 0 and symbol 1. It is observed that either channel 2 or channel 3 is employed for these solutions depending on the noise level. For the optimal stochastic solution, the

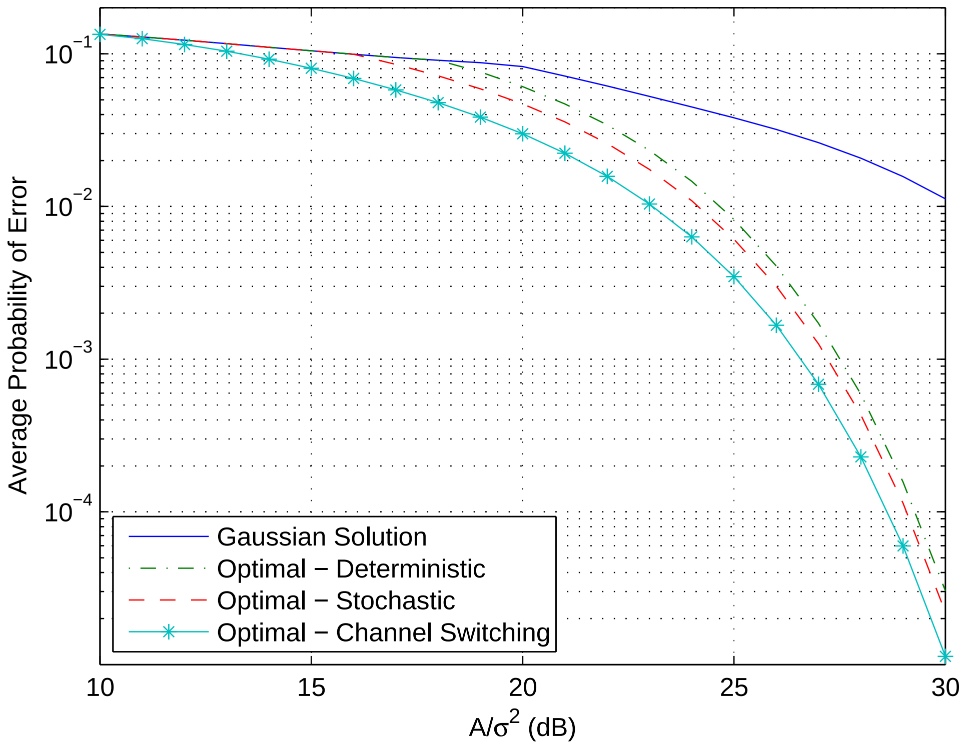

Figure 4.4: Average probability of error versus A/σ^2 for various approaches, where $K = 3$, $\mathbf{v}_1 = [-3\ \ -2\ 0\ 2\ 3]$, $\mathbf{v}_2 = [-4\ \ -3\ 0\ 3\ 4]$, $\mathbf{v}_3 = [-5\ \ -3\ 0\ 3\ 5]$, and $E = 3$ (see (4.23)).

same notation is employed as in Table 4.1 together with the channel index employed for communications. In the case of optimal channel switching, Table 4.2 shows the two channels between which switching is performed (the "X" mark indicates that the corresponding channel is not utilized). As an example, for $\mathrm{A}/\sigma^2 = 20\,\mathrm{dB}$ in Figure 4.4, the optimal channel switching strategy transmits over channel 1 using the constellation $\{-1.2108,\ 1.2108\}$ with probability 0.5614 (i.e., 56.14% of the time), and transmits over channel 2 using the constellation $\{-0.6353,\ 0.6353\}$ with probability 0.4386. Since the average noise power is the same for all channels, the optimal parameters for each strategy are determined by the variance and the means of the Gaussian mixture components. As A/σ^2 increases, the overlap between the class conditional PDFs corresponding

Table 4.2: Optimal signal parameters for the scenario in Figure 4.4.

	Gaussian solution		Deterministic Sig.		Stochastic Signaling				Channel Switching			
A/σ^2 (dB)	Channel	$\mathbf{s}_1$	Channel	$\mathbf{s}_1$	Channel	λ	$\mathbf{s}_{1,1}$	$\mathbf{s}_{2,1}$	λ	$\mathbf{s}_1^{(1)}$	$\mathbf{s}_1^{(2)}$	$\mathbf{s}_1^{(3)}$
10	2	1	2	1	2	N/A	1	1	0.8450	1.0601	0.5697	X
15	2	1	2	1	2	0.0502	1.0078	0.9996	0.5642	1.202	0.6509	X
20	3	1	2	0.6405	2	0.7547	0.6381	1.6805	0.5614	1.2108	0.6353	X
25	3	1	2	0.6213	2	0.7348	0.6210	1.6439	0.6023	1.1848	0.6206	X
30	3	1	2	0.6152	2	0.7222	0.6152	1.6174	0.6369	1.1638	0.6151	X

to symbols $i \in \{0, 1\}$ decreases and there is more room in the signal space for performance improvement via randomized approaches.

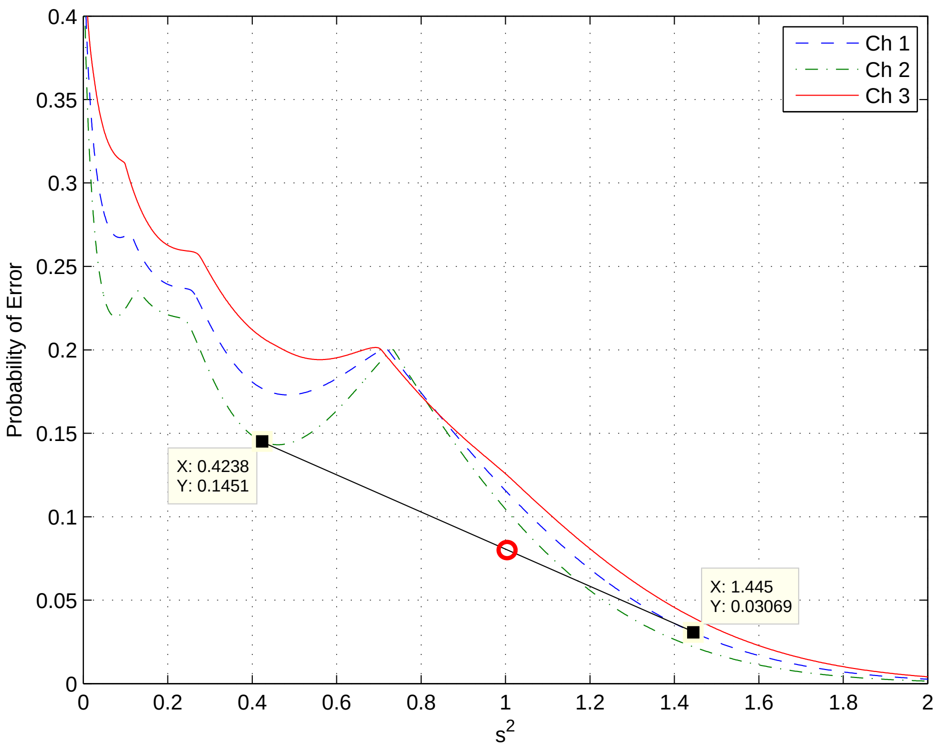

Figure 4.5: Error probability versus signal power $\mathbf{s}^2$ for the three channels when $A/\sigma^2 = 15$ dB (cf. Figure 4.4 and Table 4.2).

In order to illustrate the improvements via channel switching, Figure 4.5 presents the error probabilities of the three channels considered in Figure 4.4 and Table 4.2 as a function of the signal power in the presence of antipodal signaling when $A/\sigma^2 = 15$ dB. As shown in the figure, the optimal channel

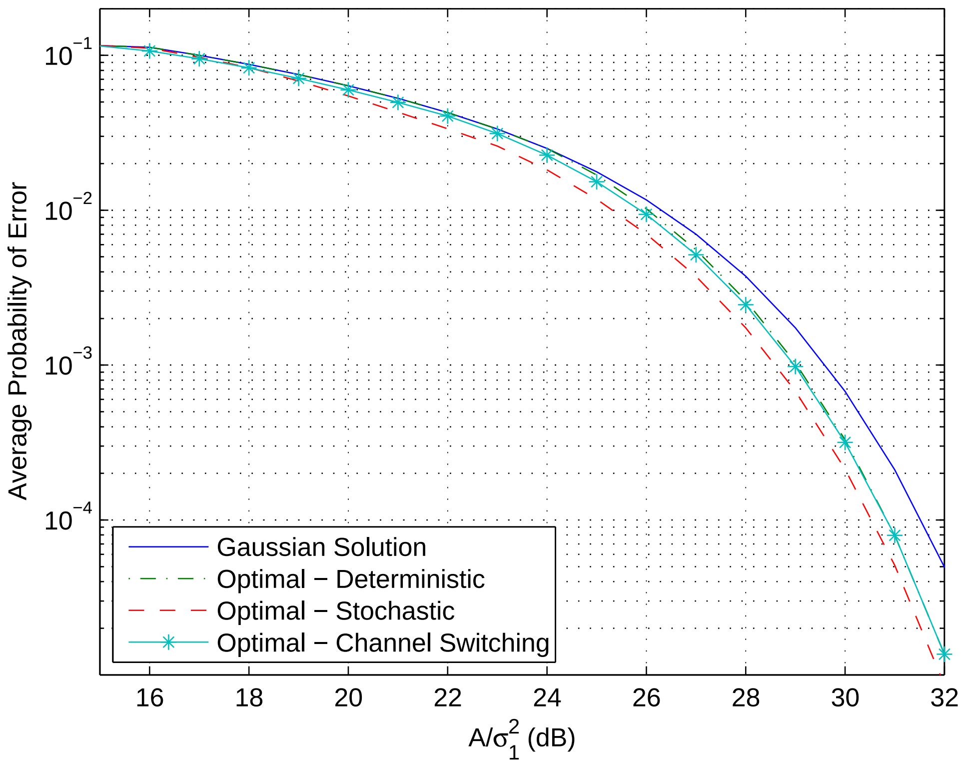

Figure 4.6: Average probability of error versus A/σ_1^2 for various approaches, where the first channel is characterized by the parameters $K = 2$, $\mathbf{v}_1 = [-6\ \ -3\ -2\ 2\ 3\ 6\ ]$, $E = 4$ (see (4.23)), and the second channel has zero-mean Gaussian noise with the same average power as the first channel.

switching strategy performs time sharing between Channel 1 and Channel 2 with power levels 1.445 and 0.4238 (i.e., signal constellations $\{-1.202, 1.202\}$ and $\{-0.6509, 0.6509\}$), respectively. The results are in compliance with Table 4.2, as expected. It should also be noted that a lower average probability of error can be achieved for the scenario in Figure 4.5 if detector randomization is allowed for each channel; that is, if multiple detectors can be implemented and time shared for the detection of symbols acquired over each channel. In that case, a randomization between two constellations and the corresponding MAP detectors over Channel 2 can result in a lower average probability of error. Fortunately, as previously stated in Footnote 2, such scenarios can be covered using the proposed framework in this study by considering multiple channels with identical distributions.

Table 4.3: Optimal signal parameters for the scenario in Figure 4.6.

	Gaussian solution		Deterministic sig.		Stochastic Signaling				Channel Switching		
A/σ_1^2 (dB)	Channel	$\mathbf{s}_1$	Channel	$\mathbf{s}_1$	Channel	λ	$\mathbf{s}_{1,1}$	$\mathbf{s}_{2,1}$	λ	$\mathbf{s}_1^{(1)}$	$\mathbf{s}_1^{(2)}$
15	2	1	2	1	2	N/A	1	1	0.1823	0.6683	1.0599
20	1	1	1	1	1	0.0857	0.2068	1.0439	0.9134	1.0266	0.6576
25	1	1	1	0.6963	1	0.6725	0.6964	1.4344	0.8810	0.6961	2.1951
30	1	1	1	0.7037	1	0.6378	0.7037	1.3743	0.9495	0.7037	3.2388

Finally, a scenario with just two channels is considered. The parameters of the first channel are given by $\mathbf{v}_1 = [-6\ -3\ -2\ 2\ 3\ 6]$, $L_1 = 6$, and $E = 4$ (see (4.23)). The second channel is modeled to have zero-mean Gaussian noise with the same average power as the first one; i.e., $L_2 = 1$, $\mu^{(2)} = 0$, and $\sigma_2^2 = \sigma_1^2 + \frac{E}{L_1}$ in (4.22). The average probabilities of error for the proposed strategies are plotted versus A/σ_1^2 in Figure 4.6. Unlike the cases in Figure 4.2 and Figure 4.4, the best performance is achieved by stochastic signaling over the best channel in this scenario. It should be emphasized that the possibility of an optimal solution in the form of stochastic signaling is stated in Section 4.1 (see (4.19)-(4.21)). It is also observed that the optimal channel switching strategy performs very closely to the optimal deterministic signaling strategy. In other words, channel switching does not provide significant performance improvements due to the poor error performance of Channel 2 with respect to that of Channel 1 over the given range of A/σ_1^2 values. The optimal parameters of the strategies depicted in Figure 4.6 are presented for some values of A/σ_1^2 in Table 4.3.

4.3 Concluding Remarks

Optimal signaling and detector design has been studied under an average transmit power constraint for generic noise distributions in the presence of multiple channels and stochastic signaling. It has been shown that the optimal solution to the joint channel switching, stochastic signaling, and detector design problem corresponds to one of the following strategies: (i) deterministic signaling over a single channel, (ii) randomizing (time sharing) between at most two signal constellations over a single channel, or (iii) switching (time sharing) between at most two channels with deterministic signaling over each channel. For all cases, the optimal strategies employ the corresponding MAP detectors at the receiver. Optimization problems have been formulated to obtain the parameters of the proposed strategies. In addition, sufficient conditions have been provided to specify whether the proposed strategy can or cannot improve the error performance over the conventional approach, in which a single channel is employed with deterministic signaling at the average power limit. Various numerical examples have been presented to illustrate the theoretical results. It has been observed that significant performance improvements can be achieved in some cases via the proposed optimal approach in the presence of multimodal noise.

Chapter 5

Conclusions and Future Work

In this dissertation, single-user and multiuser communications systems subject to average power constraints have been studied. In Chapter 2, the downlink of a multiuser communications system has been considered under the assumptions that the transmitter can randomize among different signal constellations and a fixed decision rule is employed at the receiver of each user. It has been shown that the optimal strategy is to randomize among at most $(K+1)$ different signal constellations, where K is the number of users. Since the original optimization problem is nonconvex, an approximate solution based on convex relaxation has been obtained. In the case of binary symmetric signaling and employment of sign detectors at the receiver of each user, the maximum improvement ratio achieved via the proposed approach compared to the conventional approach has been calculated in the interference limited scenario as K. Sufficient conditions have been provided for the maximum and minimum improvement ratios. In Chapter 3, the scenario in Chapter 2 has been reconsidered under the assumption that each user has N_{d} detectors and the receiver can switch among them according to some probability distribution. In that scenario, the objective has been the joint optimization of signal constellations, detector randomization factors and detectors under an average power constraint. It has been noted that the power is limited in any bit duration as opposed to the previous scenario, in which the time average power constraint is considered. The conditions under which the maximum and minimum improvement ratios are achieved have been provided. It has been shown that the optimal detector randomization approach has a lower bound, and a simple solution to achieve that bound has been presented in the case of equal crosscorrelations and noise powers. The extensions to M-ary communications systems and uplink scenarios have been discussed for both scenarios in Chapters 2 and 3.

In Chapter 4, single user systems have been considered in the presence of multiple channels where each channel can have any arbitrary noise PDF. It has been assumed that at any given time only one channel can be used for transmission and the receiver knows which channel is in use. Stochastic signaling has been considered at the transmitter for each channel. In other words, the two approaches, stochastic signaling and channel switching, have been considered jointly for single user M-ary communications systems subject to an average power constraint. The objective has been to jointly optimize stochastic signals, channel switching factors, and detectors to minimize the average probability of error. It has been shown that the optimal solution is to randomize among two distinct signal levels over the same channel (stochastic signaling) or to switch among two channels with deterministic signaling over each channel (channel switching). Therefore, it has been concluded that considering the two approaches jointly does not provide any further improvement.

For the first part of this dissertation, a future work can be to consider the downlink of a multiuser communications system in block fading channels. For the second part of the dissertation, a possible future work is to study a multichannel scenario, where each channel has a transmission cost and the objective is to minimize the average transmission cost under average power and average error probability constraints.

Bibliography

[1] B. Dulek, M. E. Tutay, S. Gezici, and P. K. Varshney, "Optimal signaling and detector design for M-ary communication systems in the presence of multiple additive noise channels," *Digital Signal Processing*, DOI: 10.1016/j.dsp.2013.10.012, 2014.

[2] M. Azizoglu, "Convexity properties in binary detection problems," *IEEE Trans. Inform. Theory*, vol. 42, no. 4, pp. 1316–1321, July 1996.

[3] S. Loyka, V. Kostina, and F. Gagnon, "Error rates of the maximum-likelihood detector for arbitrary constellations: Convex/concave behavior and applications," *IEEE Trans. Information Theory*, vol. 56, no. 4, pp. 1948–1960, April 2010.

[4] C. Goken, S. Gezici, and O. Arikan, "Optimal stochastic signaling for power-constrained binary communications systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3650–3661, Dec. 2010.

[5] ——, "Optimal signaling and detector design for power-constrained binary communications systems over non-Gaussian channels," *IEEE Commun. Letters*, vol. 14, no. 2, pp. 100–102, Feb. 2010.

[6] B. Dulek and S. Gezici, "Detector randomization and stochastic signaling for minimum probability of error receivers," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 923–928, Apr. 2012.

[7] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson and inequality-constrained signal detection," *IEEE Trans. Sig. Processing*, vol. 57, no. 5, pp. 1655–1669, May 2009.

[8] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, "Theory of the stochastic resonance effect in signal detection: Part II–Variable detectors," *IEEE Trans. Sig. Processing*, vol. 56, no. 10, pp. 5031–5041, Oct. 2007.

[9] ——, "Theory of the stochastic resonance effect in signal detection: Part I–Fixed detectors," *IEEE Trans. Sig. Processing*, vol. 55, no. 7, pp. 3172–3184, July 2007.

[10] H. Chen, P. K. Varshney, and J. H. Michels, "Noise enhanced parameter estimation," *IEEE Trans. Sig. Processing*, vol. 56, no. 10, pp. 5074–5081, Oct. 2008.

[11] S. Bayram and S. Gezici, "Noise-enhanced M-ary hypothesis-testing in the minimax framework," in *Proc. International Conference on Signal Processing and Commun. Systems*, Omaha, Nebraska, Sep. 2009, pp. 31–36.

[12] S. Bayram, S. Gezici, and H. V. Poor, "Noise enhanced hypothesis-testing in the restricted Bayesian framework," *IEEE Trans. Sig. Processing*, vol. 58, no. 8, pp. 3972–3989, Aug. 2010.

[13] M. D. McDonnell, "Is electrical noise useful?" *Proceedings of the IEEE*, vol. 99, no. 2, pp. 242–246, Feb. 2011.

[14] G. Guo, M. Mandal, and Y. Jing, "A robust detector of known signal in non-Gaussian noise using threshold systems," *Signal Processing*, vol. 92, pp. 2676–2688, 2012.

[15] E. G. Larsson, "Improving the frame-error-rate of spatial multiplexing in block fading by randomly rotating the signal constellation," *IEEE Commun. Letters*, vol. 8, no. 8, pp. 514–516, Aug. 2004.

[16] ——, "Constellation randomization (CoRa) for outage performance improvement on MIMO channels," in *IEEE Global Telecommunications Conference*, vol. 1, Nov.-Dec. 2004, pp. 386–390.

[17] Y. Li, C. N. Georghiades, and G. Huang, "Transmit diversity over quasi-static fading channels using multiple antennas and random signal mapping," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1918–1926, Nov. 2003.

[18] C. Lamy and J. Boutros, "On random rotations diversity and minimum MSE decoding of lattices," *IEEE Trans. Information Theory*, vol. 46, pp. 1584–1589, July 2000.

[19] A. Hiroike, F. Adachi, and N. Nakajima, "Combined effects of phase sweeping transmitter diversity and channel coding," *IEEE Trans. Vehicular Tech.*, vol. 41, pp. 170–176, May 1992.

[20] X. Ma and G. B. Giannakis, "Space-time-multipath coding using digital phase sweeping," in *IEEE Global Communications Conference*, vol. 1, Nov. 2002, pp. 384–388.

[21] S. Verdu, *Multiuser Detection.* 1st ed. Cambridge, UK: Cambridge University Press, 1998.

[22] Y. Shi, R. K. Mallik, and K. B. Letaief, "Randomized power control for two-hop interference channels," *IEEE Commun. Letters*, vol. 14, no. 12, pp. 1128–1130, Dec. 2010.

[23] N. Tarhuni, M. Elmusrati, and A. Abouda, "Two-state discretized transmit power control for ad-hoc sensor networks," in *3rd International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Oct. 2011, pp. 1–6.

[24] T.-S. Kim and S.-L. Kim, "Random power control in wireless ad hoc networks," *IEEE Commun. Letters*, vol. 9, no. 12, pp. 1046–1048, Dec. 2005.

[25] M. Elmusrati, N. Tarhuni, and R. Jantti, “Performance analysis of random uniform power allocation for wireless networks in Rayleigh fading channels,” *Eur. Trans. Telecomms.*, vol. 20, pp. 457–462, 2009.

[26] ——, “Random power control for uncorrelated Rayleigh fading channels,” in *IEEE International Conference on Signal Processing and Communications (ICSPC)*, Nov. 2007, pp. 360–363.

[27] Y. Zhu, Y. Xu, S. Guan, and J. He, “Connectivity analysis of random power control in wireless sensor networks,” in *12th IEEE International Conference on Communication Technology (ICCT)*, Nov. 2010, pp. 1248–1251.

[28] D. M. Novakovic and M. L. Dukic, “Evolution of the power control techniques for DS-CDMA toward 3G wireless communication systems,” *IEEE Communications Surveys*, vol. 3, no. 4, pp. 2–15, 2010.

[29] S. Koskie and Z. Gajic, “SIR-based power control algorithms wireless CDMA networks: An overview,” *Dynamics of Continuous, Discrete and Impulsive Systems B: Applications and Algorithms*, no. 10-S, p. 286, 2003.

[30] M. Chiang, P. Hande, T. Lan, and C. W. Tan, “Power control in wireless cellular networks,” *Foundations and Trends in Networking*, now Publishers Inc., 2008.

[31] Z. Han, D. Niyato, W. Saad, T. Basar, and A. Hjorungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications.* Cambridge University Press, 2012.

[32] K. Akkarajitsakul, E. Hossain, D. Niyato, and D. I. Kim, “Game theoretic approaches for multiple access in wireless networks: A Survey,” *IEEE Communications Surveys Tutorials*, vol. 13, no. 3, pp. 372–395, Quarter 2011.

[33] S. Ulukus and R. D. Yates, “Stochastic power control for cellular radio systems,” *IEEE Transactions on Communications*, vol. 46, no. 6, pp. 784–798, 1998.

[34] J. Luo, S. Ulukus, and A. Ephremides, “Standard and quasi-standard stochastic power control algorithms,” *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2612–2624, July 2005.

[35] C. Goken, S. Gezici, and O. Arikan, “On the optimality of stochastic signaling under an average power constraint,” in *Proc. 48th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sep. 29-Oct. 1 2010, pp. 1158–1164.

[36] R. K. Mallik, R. A. Scholtz, and G. P. Papavassilopoulos, “Analysis of an on-off jamming situation as a dynamic game,” *IEEE Trans. Commun*, vol. 48, no. 8, pp. 1360–1373, Aug. 2000.

[37] B. Dulek and S. Gezici, “Optimal stochastic signal design and detector randomization in the Neyman-Pearson framework,” in *37th IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP’12)*, Kyoto, Japan, March 2012.

[38] E. L. Lehmann, *Testing Statistical Hypotheses*, 2nd ed. New York: Chapman & Hall, 1986.

[39] S. M. Kay, “Noise enhanced detection as a special case of randomization,” *IEEE Sig. Processing Lett.*, vol. 15, pp. 709–712, 2008.

[40] Y. Ma and C. C. Chai, “Unified error probability analysis for generalized selection combining in Nakagami fading channels,” *IEEE Jour. Select. Areas Commun.*, vol. 18, no. 11, pp. 2198–2210, Nov. 2000.

[41] J. A. Ritcey and M. Azizoglu, “Performance analysis of generalized selection combining with switching constraints,” *IEEE Commun. Letters*, vol. 4, no. 5, pp. 152–154, May 2000.

[42] M. E. Tutay, S. Gezici, and O. Arikan, “Optimal randomization of signal constellations on the downlink of a multiuser DS-CDMA system,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4878–4891, 2013.

[43] ——, “Optimal detector randomization for multiuser communications systems,” *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2876–2889, 2013.

[44] H. V. Poor, *An Introduction to Signal Detection and Estimation.* New York: Springer-Verlag, 1994.

[45] S. A. Grandhi, J. Zander, and R. Yates, “Constrained power control,” *Wireless Personal Communications*, vol. 1, no. 4, pp. 257–270, 1995.

[46] C. W. Tan, M. Chiang, and R. Srikant, “Maximizing sum rate and minimizing MSE on multiuser downlink: Optimality, fast algorithms and equivalence via max-min SINR,” *IEEE Trans. on Signal Proc.*, vol. 59, no. 12, pp. 6127–6143, Dec. 2011.

[47] C. Yang and J. Li, “Mixed-strategy based discrete power control approach for cognitive radios: A matrix game-theoretic framework,” in *International Conference on Future Computer and Communication*, vol. 3, May 2010, pp. V3–806–V3–810.

[48] T. Heikkinen, “A minimax game of power control in a wireless network under incomplete information,” DIMACS Technical Report 99-43, Tech. Rep., 1999.

[49] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, UK: Cambridge University Press, 2004.

[50] S. Bayram and S. Gezici, “On the performance of single-threshold detectors for binary communications in the presence of Gaussian mixture noise,” *IEEE Trans. Commun.*, vol. 58, no. 11, pp. 3047–3053, Nov. 2010.

[51] J.-J. Xiao and Z.-Q. Luo, “Universal decentralized detection in a bandwidth-constrained sensor network,” *IEEE Trans. on Signal Processing*, vol. 53, no. 8, pp. 2617–2624, aug. 2005.

[52] K. E. Parsopoulos and M. N. Vrahatis, *Particle swarm optimization method for constrained optimization problems.* IOS Press, 2002, pp. 214–220, in Intelligent Technologies–Theory and Applications: New Trends in Intelligent Technologies.

[53] "CVX: Matlab software for disciplined convex programming, Version 1.21 (April 2011, Build 808)." [Online]. Available: http://cvxr.com/cvx

[54] R. T. Rockafellar, *Convex Analysis.* Princeton, NJ: Princeton University Press, 1968.

[55] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization.* Boston, MA: Athena Specific, 2003.

[56] C. C. Pugh, *Real Mathematical Analysis.* New York: Springer-Verlag, 2002.

[57] L. Huang and M. J. Neely, "The optimality of two prices: Maximizing revenue in a stochastic network," in *Proc. 45th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sep. 2007.

[58] A. Goldsmith, *Wireless Communications.* Cambridge University Press, 2005.

[59] I. Kra and S. Simanca, "On circulant matrices," *Notices of the American Mathematical Society*, vol. 59, no. 3, March 2012.

[60] J. Mitola and J. Maguire, G.Q., "Cognitive radio: making software radios more personal," *Personal Communications, IEEE*, vol. 6, no. 4, pp. 13–18, 1999.

[61] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 2, pp. 201–220, 2005.

[62] S. Gezici, H. Celebi, H. Poor, and H. Arslan, “Fundamental limits on time delay estimation in dispersed spectrum cognitive radio systems,” *Wireless Communications, IEEE Transactions on*, vol. 8, no. 1, pp. 78–83, 2009.

[63] V. Bhatia and B. Mulgrew, “Non-parametric likelihood based channel estimator for Gaussian mixture noise,” *Signal Processing*, vol. 87, pp. 2569–2586, Nov. 2007.

[64] T. Erseghe, V. Cellini, and G. Dona, “On UWB impulse radio receivers derived by modeling MAI as a Gaussian mixture process,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2388–2396, June 2008.

[65] B. Dulek and S. Gezici, “Optimal signaling and detector design for power constrained on-off keying systems in Neyman-Pearson framework,” in *Proc. 16th IEEE Wksp. Stat. Signal Process. (SSP'11)*, Nice, France, June 2011, pp. 93–96.