Optimized Binary Constellation Design for Distributed Detection over Gaussian MAC Sensor Networks

by

Luca Sardellitti

A thesis submitted to the
Department of Mathematics and Statistics
in conformity with the requirements for
the degree of Master of Applied Science

Queen's University
Kingston, Ontario, Canada
September 2024

Abstract

We consider a distributed detection system transmitting a binary source over a Gaussian multiple access channel (MAC). We model the network via binary sensors whose outputs are generated by binary symmetric channels of different noise levels. We aim to find the optimal constellation design for each sensor under individual power constraints. We begin by analyzing a version of the problem where there are two sensors sending a uniform source over the channel. In an introductory investigation, we assume each sensor uses all available power, but can change the relative angle between the constellations. Although explicit analysis of error probability in this setup is infeasible, an upper bound on the error probability is optimized, and is numerically shown to coincide very well with the true optimal rotation angle. Interestingly, this upper bound ignores all information in the imaginary axis, and could be equivalently achieved by using less power if both sensors were sharing a one-dimensional MAC. This led to the problem formulation of optimal power allocation for one-dimensional constellation designs.

We next consider two binary sensors sending a non-uniform source over a one dimensional Gaussian MAC. We prove an optimal constellation design under individual sensor power constraints which minimizes the error probability of detecting the source. Three distinct cases arise for this optimization based on the parameters in the

problem setup. In the most notable case (Case III), the optimal signaling design is to not necessarily use all of the power allocated to the more noisy sensor (i.e., whose output has less correlation to the source). We compare the error performance of the optimal one dimensional constellation to orthogonal signaling. The results show that the optimal one dimensional constellation achieves lower error probability than using orthogonal channels.

Finally, we extend the problem to N sensors ($N \geq 2$) by making a simplifying assumption that the detection (fusion) center will not use maximum-a-posteriori (MAP) detection, but instead will use a simplified rule where the real line can only be split into left and right decision regions by a single decision boundary. Under this assumption, we characterize the optimization of any individual sensor's power allocation when the rest are fixed. This optimization is also divided into three distinct cases which directly mirror those from the two sensor problem. In the equivalent situation to Case III, the sensor should not necessarily use all of its power. However, unlike the two sensor case, the optimal power allocation does not generally have a closed form expression, and instead must be found numerically. We use the individual sensor optimization results to form an iterative algorithm for the joint optimization of all N sensors. We use numerical examples to compare the algorithm to other signaling designs, observing that the algorithm achieves consistently lower error probability than any design that also uses the simplified detection. Further, the algorithm is only slightly outperformed in some situations, particularly at high signal-to-noise ratio (SNR), by signaling techniques such as orthogonal signaling or using MAP detection, which typically require more power, bandwidth and overall implementation complexity.

Acknowledgments

I would first like to thank my supervisors, Professors Fady Alajaji and Glen Takahara for their support, guidance and advice throughout my degree. You have been invaluable mentors to me at every step of the way.

I would also like to thank my family and friends for all their support through my degree. Your encouragement means the world to me.

I am grateful to the Department of Mathematics and Statistics and Queen's University for fostering a warm and welcoming environment for my academic studies. Many thanks to Jennifer Read for her assistance throughout my degree.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Wireless sensor networks are widely used for monitoring the state of real world phenomena. This includes both the estimation of a real valued parameter (such as temperature or rain fall measurements) and the detection of an event occurring (such as the occurrence of forest fires or a security breach). In this thesis, we focus on the hypothesis testing problem described by distributed detection of an event occurring.

When working with generalized distributed detection problems, the error probability of the system cannot in general be expressed analytically. As a result, previous work on distributed detection typically uses related or proxy metrics for system error analysis. For example, [1] uses error exponents to evaluate the performance of various detection schemes, [2] uses the J-divergence (i.e., the Jeffreys-divergence [3]), while [4] and [5] use the deflection coefficient as the metric for optimization.

Previous work in this area employs a variety of signaling structures for the sensors. For example, [2] uses orthogonal channels for each sensor, while [5] uses a single MAC for the entire network. Works such as [5–8] fix a signaling design and analyze detection schemes at the fusion center. Other works optimize the sensors' signaling techniques under certain constraints. For example, [2] optimizes power allocation for

a network that uses orthogonal signaling and on-off keying for each sensor, under total and average power constraints. Alternatively, [4] assumes fixed power at each sensor, and analyzes the optimal rotation angle to send the signals. In [9], the problem of distributed mean-squared error estimation of a Gaussian source sent over a symmetric Gaussian sensor network is analyzed; it is shown that uncoded transmission (scalar coding) is optimal in the sense that it achieves Shannon's optimal performance theoretically achievable among all source-channel block codes of sufficiently large lengths. The distributed detection setup in this thesis can be seen as a discretization of the distributed estimation system in [9] if we use a uniformly distributed source and symmetric sensor channels. The source and sensor readings can be represented as one bit quantizations of their continuous counterparts in [9]. Further, [10] extends the distributed estimation problem of [9] to include fading and asymmetric sensors, providing a sufficient condition under which uncoded transmission is optimal.

Throughout the above mentioned works, there is not much emphasis put on generalized constellation design for distributed detection problems. We aim to solve the source-channel signaling problem of finding an optimized constellation design to minimize error probability under a given source and channel model. This is similar in principle to works such as [11–19], where general constellation design is optimized for a chosen criterion. In [11], the optimal joint binary constellation design for two correlated sources was derived. In [12] and [13], the authors used a minimum inter-constellation distance criterion for optimizing constellations for multiple sources. In contrast, [16–19] optimized M-ary constellations for a single source.

1.1 Contributions

We investigate constellation design and provide various optimality results for distributed detection problems involving a binary source sent over a Gaussian MAC. We model hypothesis testing for an event of interest occurring as a non-uniformly distributed binary source, and the sensor noises are modelled as passing the source through independent memoryless binary symmetric channels, introducing sensor errors.

In Chapter 2, we investigate the effect of rotation angle between binary constellations in a two sensor network transmitting a uniform binary source. We derive an explicit angle, θ^*_{ub}, which minimizes an upper bound on the error probability, and is numerically shown to perform identically to the true error probability. We also remark that the upper bound ignores all information in the imaginary axis; this motivates the study of optimal power allocation for one-dimensional constellation designs. More detailed derivations and analysis of this initial investigation can be found in [20].

Next, in Chapter 3, we analyze a two sensor communication network sharing a single MAC. With this setup, we analytically derive an optimal one dimensional constellation design to minimize the system error probability under individual power constraints for each sensor. In addition, we show using numerical and simulation results that the derived optimal one dimensional constellation achieves lower error probability than using orthogonal channels. Our most notable result is that in certain cases (which is dominant when the source is nearly uniformly distributed), the noisier sensor should use some but not all of its allocated power. The results of this chapter have been published in [21] and [22].

Finally, in Chapter 4, we extend the two sensor problem described in Chapter 3

to an arbitrary number of sensors. Under a simplifying assumption on the detection scheme, the optimization of any individual sensor's constellation is characterized. This individual optimization is iteratively applied to jointly optimize the constellations of all sensors. The performance of this algorithm is compared to other designs, where it is typically shown to be the best performing design, and is only slightly outperformed in some situations (typically high SNR) by schemes which use more power and bandwidth, such as orthogonal channels.

The MATLAB code used to generate all numerical results in this thesis can be found at https://github.com/lsardellitti/SensorNetworkMatlabTests.

Chapter 2

Preliminaries and Motivation

To begin formulation of the problem, we first define a generalized mathematical model of the situation, to give a unified notation for all specific versions of the problem which are addressed in this work.

2.1 General Problem Description and Notation

Let X be a binary event that is to be observed by a sensor network. Without loss of generality, we assume its distribution is such that $p_1 \triangleq \Pr(X = 1) \leq 0.5$. We also define $p_0 \triangleq \Pr(X = 0) = 1 - p_1$. For some integer $N \geq 2$, there are N sensors, $X_1, ..., X_N$ observing the source X. We will use the notation $X^N \triangleq (X_1, ..., X_N)$ to represent the random vector of the sensor readings. The sensors are modelled as passing X through memoryless binary symmetric channels. This is expressed as $X_i = X \oplus Z_i,\ i = 1, ..., N$, where $\oplus$ denotes addition modulo-2, with Z_i being independent Bernoulli noise processes with means (or channel crossover probabilities) ϵ_i for each $i = 1, ..., N$, respectively. It is also assumed that X is independent from $(Z_1, ..., Z_N)$. For additional notation, we also define the vectors or tuples $Z^N \triangleq (Z_1, ..., Z_N)$ and $\epsilon^N \triangleq (\epsilon_1, ..., \epsilon_N)$. Without loss of generality, we assume that the sensors are ordered in

increasing noise levels, which is equivalent to decreasing correlation to the source, X, expressed as: $0 < \epsilon_1 \leq ... \leq \epsilon_N < 0.5$. The sensors, unable to communicate with each other, encode their data independently using binary constellations. The constellations for the sensors are represented as follows: $\mathcal{C}_i = \{c_{0,i}, c_{1,i}\},\ i = 1, ..., N$, where for $s \in \{0, 1\}$, $c_{s,i} \in \mathcal{S}$ denotes the constellation point for Sensor i assigned to $X_i = s$ in the real or complex signal space $\mathcal{S} \in \{\mathbb{R}, \mathbb{C}\}$. For each $i = 1, ..., N$, let $S_i \in \mathcal{C}_i$ be the random variable associated to each sensor's chosen constellation point. Also let $P_i^{\max}$ be the power constraint for Sensor i, i.e., $E[\|S_i\|^2] \leq P_i^{\max},\ i = 1, ..., N$, where $\|\cdot\|$ is the ℓ_2-norm in the signal space $\mathcal{S}$. In this setup, each sensor has its own power allotment, as opposed to having a common power constraint on the entire network. To simplify notation, the vector of power allocations is defined as $P^{\max} \triangleq (P_1^{\max}, ..., P_N^{\max})$. The communication channel used to send the sensors' signals to the detection (fusion) center is modelled as a multiple access additive noise channel, where the received signal, R at the detection center is given by

$$R = \sum_{i=1}^{N} S_i + Z, \tag{2.1}$$

where Z is a random variable in the signal space $\mathcal{S}$ modelling the channel noise. The event X is reconstructed at the fusion center using some deterministic detection rule $\hat{x} : \mathcal{S} \to \{0, 1\}$, so the detected bit is represented by the random variable $\hat{X} \triangleq \hat{x}(R)$. Since the detection rule is deterministic, we partition the signal space into the subsets $\mathcal{D}_0$ and $\mathcal{D}_1 = \mathcal{D}_0^c$, called "decision regions", where $\mathcal{D}_s \triangleq \{r \in \mathcal{S} \mid \hat{x}(r) = s\}$, $s = 0, 1$.

We additionally introduce the following notation to simplify the algebraic expressions which will be used in the forthcoming analysis. First, we define the following notation for describing all possible combinations of sensor readings: $\mathcal{B} \triangleq \{0, 1\}^N$,

where any element in this set $\underline{b} = (b_1, ..., b_N) \in \mathcal{B}$ is such that $b_i \in \{0, 1\}$, $i \in 1, ..., N$ denotes the bit in position i, i.e., the reading from Sensor i. For ease of notation, we also use the string notation $\underline{b} = b_1...b_N$. We use this to define the following notation for the sensors' conditional probabilities $p_{\underline{b}|i} \triangleq \Pr(X^N = \underline{b} \mid X = i)$. We also define the combined constellation $\mathcal{C} = \{\sum_{i=1}^{N} c_i \,|\, c_i \in \mathcal{C}_i,\, i = 1, ..., N\}$ as all possible additive combinations of the binary constellations for each sensor. The combined constellation points are indexed by $\underline{b} \in \mathcal{B}$ using the notation $a_{\underline{b}} \in \mathcal{C}$, representing the constellation point associated to the combined sensor readings $X^N = \underline{b}$.

2.2 Preliminary Investigation - Two Sensor Uniform Source NOMA

To begin analysis into the problem of optimal constellation design, we start with a simplified problem to gain insight on the nature of the optimization. The first problem to investigate is a two sensor, uniform source, Gaussian channel distributed detection problem. Both sensors are assumed to share a 2-dimensional channel, and are able to use it in a non-orthogonal-multiple-access (NOMA) scheme by varying the angle between the constellations, similar to inter constellation rotation as described by [12]. These specifications are described by the following realizations of the parameters described in the Section 2.1: $p_1 = 0.5$, $N = 2$, $\mathcal{S} = \mathbb{C}$. The constellations for the sensors are parameterized as follows: $\mathcal{C}_1 = \{c_{0,1}, c_{1,1}\} = \{-\sqrt{P_1},\ \sqrt{P_1}\}$ and $\mathcal{C}_2 = \{c_{0,2}, c_{1,2}\} = \{-\sqrt{P_2}e^{j\theta},\ \sqrt{P_2}e^{j\theta}\}$, where j is the imaginary unit and $\theta \in [0, \frac{\pi}{2}]$ is the rotation between the two constellations. The signals are sent through a Gaussian MAC as described in (2.1), where the channel noise, Z, is a complex (bivariate) Gaussian noise variable with independent zero mean components of equal variance given by $\sigma^2 = \frac{N_0}{2}$. It is also assumed that Z is independent of the sensor signals S_1

and S_2. To recover the original data source, maximum-likelihood (ML) detection is used (which is optimal as the source is uniform), described by

$$\hat{x}(r) = \arg\max_{s\in\{0,1\}} f_R(r \mid X = s),$$

where $f_R(\cdot \mid X = s)$, $s = 0, 1$ are the conditional probability density functions (pdfs) of the received signal R given the source event $X = s$. The decision regions, $\mathcal{D}_0$ and $\mathcal{D}_1$, formed from ML detection are generally complex and have a curved boundary between them, as shown in Figure 2.1.

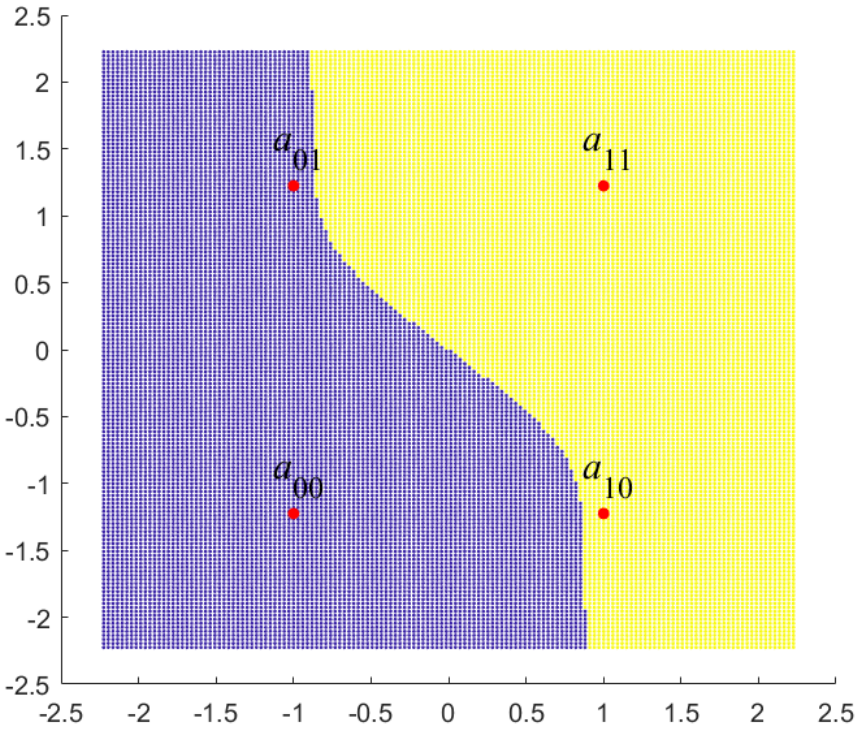


Figure 2.1: Decision regions for $\theta = \frac{\pi}{2}$ (yellow region is $\mathcal{D}_1$). The red points represent the superimposed constellation points $a_{lm} \in \mathcal{C}$ ($\epsilon_1 = 0.15$, $\epsilon_2 = 0.17$, $P_1 = 1$, $P_2 = 1.5$, $N_0 = 1$).

The complexity of the decision regions makes it infeasible to analyze the error probability of ML detection directly. An upper bound was formed based on simplifying the decision regions to being half planes, or so called "planar" decision regions. Instead of ML detection, the decision regions are $\mathcal{D}_0 = \mathcal{D}_{0,\text{planar}} \triangleq \{r \mid \text{Re}(r) \leq 0\}$. This yields the following expression for the upper bound on the error probability for any $\theta \in [0, \frac{\pi}{2}]$:

$$P_e^{\text{ub}}(\theta) = \epsilon_1 + (1-\epsilon_1-\epsilon_2)Q\left(\frac{\sqrt{P_1}+\sqrt{P_2}\cos(\theta)}{\sigma}\right) + (\epsilon_2-\epsilon_1)Q\left(\frac{\sqrt{P_1}-\sqrt{P_2}\cos(\theta)}{\sigma}\right),$$

where Q is the Gaussian tail distribution function, defined as

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{u^2}{2}}du.$$

This upper bound was optimized over $\theta \in [0, \frac{\pi}{2}]$, yielding the following minimizer:

$$\theta_{\text{ub}}^* = \cos^{-1}(\min\{pcf, 1\}),$$

where

$$pcf \triangleq \frac{N_0}{4\sqrt{P_1P_2}}\ln\left(\frac{1-\epsilon_2-\epsilon_1}{\epsilon_2-\epsilon_1}\right),$$

is called the *power-correlation-factor*. The error probability achieved by this optimized upper bound is indistinguishable from a numerical optimization over rotation angle using the ML detection scheme, seen in [20, Section IV]. Further, we note that the decision regions calculated for the optimal ML detection were indeed planar as well. The details of the derivation and optimization of this upper bound can be found in [20]. Figure 2.2 shows the combined constellation and ML decision regions at θ_{ub}^*.

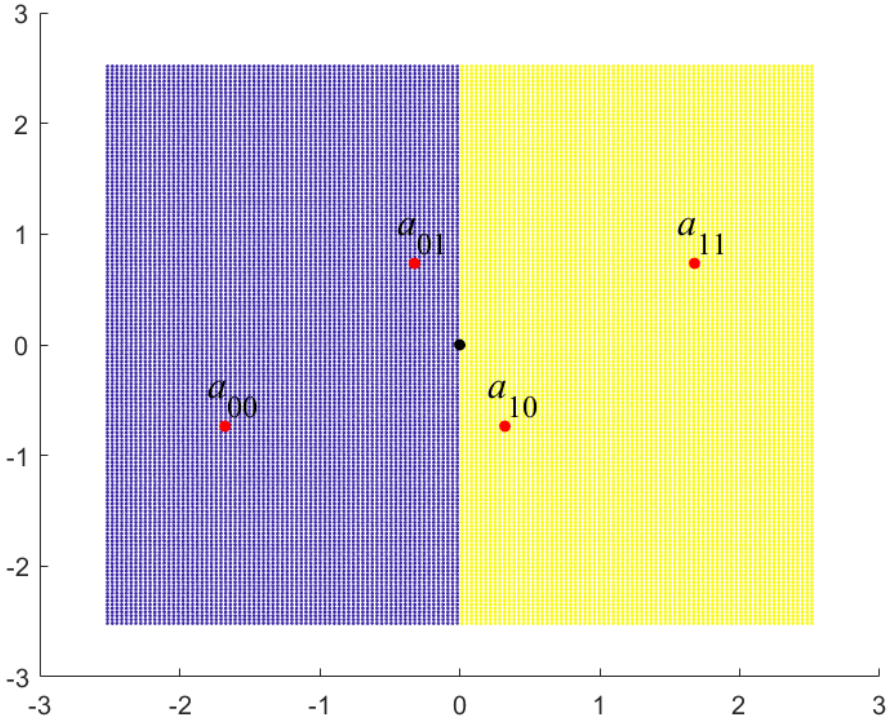


Figure 2.2: Combined constellation and decision regions at θ^*_{ub} (yellow region is $\mathcal{D}_1$). The red points represent the superimposed constellation points $a_{lm} \in \mathcal{C}$ ($\epsilon_1 = 0.1$, $\epsilon_2 = 0.15$, $P_1 = 1$, $P_2 = 1$, $N_0 = 1$).

We observe that the numerically optimized rotation angle and corresponding ML detection regions are identical to those derived from optimizing the upper bound. This implies that the optimal signaling design ignores all information in the imaginary axis. Hence, the same error performance could be recreated using less power by sending the points co-linearly. This initial investigation led to the main direction of this research to investigate optimal power allocation in a single multiple access channel instead of constellation rotation. The largest question that was raised from this investigation is whether this result extends to more complex situations such as non-uniform sources or more than two sensors. These questions are answered in the following chapters.

Chapter 3

Two User Gaussian MAC Sensor Networks

Informed by the initial investigation of Chapter 2, we approach a general version of the two sensor problem where the sensors share a single Gaussian MAC.

3.1 System Model

This chapter uses the problem set up as described in Section 2.1, with the specifications being $N = 2$ and $\mathcal{S} = \mathbb{R}$. A block diagram depicting the overall system model is shown in Figure 3.1.

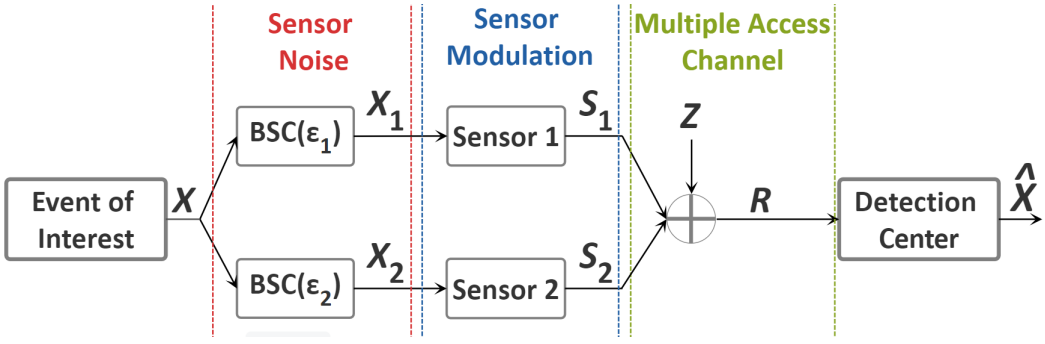


Figure 3.1: Block diagram showing the two sensor MAC system.

3.1.1 Channel Model

The sensors' signals are sent through a Gaussian MAC, where the received signal, R, is expressed as in (2.1), with the noise Z modelled as a real valued Gaussian noise variable with zero mean and variance $\sigma^2 = \frac{N_0}{2}$. It is assumed that Z is independent of the sensor signals S_1 and S_2.

3.1.2 Maximum-a-Posteriori Detection

The event X is reconstructed at the fusion center using (optimal) MAP detection. For a received signal $r \in \mathbb{R}$, the detected bit is determined as follows:

$$\begin{aligned}
\hat{x}(r) &= \arg\max_{i\in\{0,1\}} \Pr(X = i \mid R = r) \\
&= \arg\max_{i\in\{0,1\}} \Pr(X = i) f_R(r \mid X = i) \\
&= \arg\max_{i\in\{0,1\}} p_i \sum_{(l,m)\in\{0,1\}^2} p_{lm|i} f_R(r|S_1 + S_2 = a_{lm}) \\
&= \arg\max_{i\in\{0,1\}} p_i \sum_{(l,m)\in\{0,1\}^2} p_{lm|i} f_Z(r - a_{lm}),
\end{aligned} \tag{3.1}$$

where f_R and f_Z are the probability density functions (pdfs) of the received signal R and channel noise variable Z, respectively, $p_{lm|i} \triangleq \Pr(X_1 = l, X_2 = m|X = i)$, and $a_{lm} \in \mathcal{C}$ denotes the superimposed constellation symbol associated with $X_1 = l$ and $X_2 = m$. In the case of a tie, we arbitrarily choose to detect a 0, since the probability of a tie is zero because the noise is a continuous random variable. The conditional probabilities, $p_{lm|i}$, can be expressed as follows:

$$\begin{aligned}
&p_{11|0} = p_{00|1} = \epsilon_1\epsilon_2, \;\; p_{00|0} = p_{11|1} = (1-\epsilon_1)(1-\epsilon_2), \\
&p_{01|0} = p_{10|1} = (1-\epsilon_1)\epsilon_2, \;\; p_{10|0} = p_{01|1} = \epsilon_1(1-\epsilon_2).
\end{aligned} \tag{3.2}$$

3.2 Summary of Main Results

For fixed parameters p_1, ϵ_1, ϵ_2, N_0, $P_1^{\max}$ and $P_2^{\max}$, we show that the optimal constellation design for $\mathcal{C}_1$ and $\mathcal{C}_2$ which achieve the minimum error probability, P_e^*, are expressed as $\mathcal{C}_i = \{c_{0,i}, c_{1,i}\} = \left\{-\sqrt{\frac{p_1}{p_0}}P_i^*,\ \sqrt{\frac{p_0}{p_1}}P_i^*\right\}$ for $i \in \{1,2\}$, where the optimal power[1] allocations P_i^* are separated into three cases. The conditions for each case are given in Table 3.1 and the optimization results are summarized in Table 3.2.

Table 3.1: Case Characterization Conditions

Case	Condition	$\mathcal{D}_0$
I	$0 \le p_1 \le \frac{\epsilon_1\epsilon_2}{1-\epsilon_1-\epsilon_2+2\epsilon_1\epsilon_2}$	$\mathbb{R}$
II	$\frac{\epsilon_1\epsilon_2}{1-\epsilon_1-\epsilon_2+2\epsilon_1\epsilon_2} < p_1 \le \frac{\epsilon_1-\epsilon_1\epsilon_2}{\epsilon_1+\epsilon_2-2\epsilon_1\epsilon_2}$	$(-\infty, x^*]$
III	$\frac{\epsilon_1-\epsilon_1\epsilon_2}{\epsilon_1+\epsilon_2-2\epsilon_1\epsilon_2} < p_1 \le 0.5$	$(-\infty, x^*]$

For Cases II and III in Table 3.1, the decision boundary, x^*, is the unique root of (3.6) corresponding to P_1^* and P_2^*.

Table 3.2: Optimal Power Allocation Results

Case	P_1^*	P_2^*	P_e^*	$\lim_{N_0\to 0} P_e^*$
I	0	0	p_1	p_1
II	$\sqrt{P_1^{\max}}$	$\sqrt{P_2^{\max}}$	see (3.21)	see (3.35)
III	$\sqrt{P_1^{\max}}$	$\min(\sqrt{P_2^{\max}}, \tilde{P}_2)$	see (3.34)	ϵ_1

[1]Even though in this chapter, each P_i, $i \in \{1,2\}$, corresponds to the square root of the power $E[\|S_i\|^2]$ with upper bound constraint $P_i^{\max}$ as specified in Chapter 2, we will still refer to it as "power" for the sake of simplicity.

In Table 3.2, we used the quantities

$$\tilde{P}_2 \triangleq \frac{N_0 p_1 p_0}{2\sqrt{P_1^{\max}}} \ln \frac{(1-\epsilon_1-\epsilon_2)^2 - \Lambda}{(\epsilon_2-\epsilon_1)^2 - \Lambda}, \tag{3.3}$$

with

$$\Lambda \triangleq \frac{(p_0-p_1)^2}{p_0 p_1}(1-\epsilon_1)(1-\epsilon_2)\epsilon_1\epsilon_2.$$

Note that the expression for $\tilde{P}_2$ is always real valued when the conditions of Case III are met. Also note that because of the assumption $0 < \epsilon_1 \leq \epsilon_2 < 0.5$, the conditions for the three cases are numerically consistent, i.e., we have that

$$0 < \frac{\epsilon_1\epsilon_2}{1-\epsilon_1-\epsilon_2+2\epsilon_1\epsilon_2} < \frac{\epsilon_1-\epsilon_1\epsilon_2}{\epsilon_1+\epsilon_2-2\epsilon_1\epsilon_2} \leq 0.5,$$

where the last inequality holds with equality if and only if $\epsilon_1 = \epsilon_2$. The most interesting result is that in Case III, the optimal power allocation is not to necessarily use all of the available power for Sensor 2. The remainder of this chapter is dedicated to proving these results and illustrating them via numerical examples and simulations.

3.3 Proof of Main Results

First, we use Theorem 3.1 to show that the constellation design optimization problem can be restricted to a set of asymmetric constellations, parameterized by each sensor's power allocation. Then we analyze the boundary points between the decision regions $\mathcal{D}_0$ and $\mathcal{D}_1$ (decision boundaries) using this optimal asymmetric design. The characterization of these decision boundaries splits the problem into the three cases given in Table 3.1. For Case I, Proposition 3.2 shows the trivial nature of the decision

boundaries, and thus also the optimization in this case. For Case II, Propositions 3.6 and 3.7 derive bounds on the decision boundaries which are used in Theorem 3.8 to show that using all allocated power for both sensors is optimal. Finally, in Case III, Propositions 3.10 and 3.11 establish properties of the decision boundaries which are used in Theorem 3.14 to give the globally optimal power allocation for Sensor 2. Combining this result with Theorems 3.15 and 3.16, which show that the error probability decreases in both sensor powers until reaching the global minimum, yields the overall optimal power allocation under the given power constraints.

Theorem 3.1. *For any combination of binary constellations $\mathcal{C} = \mathcal{C}_1 + \mathcal{C}_2$, there exists a constellation pattern $\mathcal{C}^* = \mathcal{C}_1^* + \mathcal{C}_2^*$ which has equal error probability, equal or better power consumption, with the composition $\mathcal{C}_i^* = \left\{ -\sqrt{\frac{p_1}{p_0}} P_i,\ \sqrt{\frac{p_0}{p_1}} P_i \right\}$, for some $P_i \in \mathbb{R}$, $i \in \{1, 2\}$.*

Proof. In a Gaussian MAC using MAP detection, the error probability is the same for constellations that are translations of each other. Hence, constellations with the same distances between constellation points will have the same error performance. The distances between the points in the joint constellation $\mathcal{C}$ are determined by the distances between the points in the individual constellations $\mathcal{C}_i = \{c_{0,i}, c_{1,i}\}$, $i \in \{1, 2\}$. Let the constellation distance, d_i be defined as follows:

$$d_i \triangleq c_{1,i} - c_{0,i},\ i \in \{1, 2\}. \tag{3.4}$$

We will minimize the average power consumption for each sensor while maintaining constellation distance d_i. This is computed by using the constraint from (3.4):

$$P_i^2 = E[S_i^2] = p_0 c_{0,i}^2 + p_1 c_{1,i}^2 = c_{1,i}^2 - 2 p_0 d_i c_{1,i} + p_0 d_i^2.$$

This is a simple quadratic function of $c_{1,i}$, which is minimized at $c_{1,i} = p_0 d_i$, $c_{0,i} = -p_1 d_i$. Substituting this back into the expression of P_i^2, we see that the minimum power has the form $P_i^{*2} = p_1 p_0 d_i^2$. Finally, rearranging for $d_i = \frac{1}{\sqrt{p_1 p_0}} P_i^*$ shows that the minimum power constellation has the form:

$$c_{0,i} = -\sqrt{\frac{p_1}{p_0}} P_i, \qquad c_{1,i} = \sqrt{\frac{p_0}{p_1}} P_i, \qquad i \in \{1, 2\}.$$

□

Using Theorem 3.1, we can restrict the optimization search to constellations which take the following asymmetric form. $\mathcal{C}_i = \{c_{0,i}, c_{1,i}\} = \left\{ -\sqrt{\frac{p_1}{p_0}} P_i,\ \sqrt{\frac{p_0}{p_1}} P_i \right\}$, where $P_i \in [0, \sqrt{P_i^{\max}}]$ for $i \in \{1, 2\}$. To simplify notation, we define the following two symbols which represent these optimal asymmetric parameters:

$$\alpha \triangleq \sqrt{\frac{p_0}{p_1}}, \qquad \beta \triangleq \sqrt{\frac{p_1}{p_0}}. \tag{3.5}$$

The relationship $\alpha + \beta = \frac{1}{\sqrt{p_0 p_1}}$ will be used often in the remaining analysis. The problem has now been reduced to finding the optimal power allocations, $P_i \in [0, \sqrt{P_i^{\max}}]$ for $i \in \{1, 2\}$.

3.3.1 Decision Boundaries

To analyze the error probabilities, we first must understand the behaviour of the decision regions $\mathcal{D}_0$ and $\mathcal{D}_1$. To characterize these regions, we take the difference between the two terms in (3.1) and manipulate the expressions to give the following

explicit expression in terms of the problem parameters:

$$\begin{aligned}
&\Pr(X=1 \mid R=r) - \Pr(X=0 \mid R=r) \\
&= \sum_{(l,m)\in\{0,1\}^2} (p_1 p_{lm|1} - p_0 p_{lm|0}) f_Z(r - a_{lm}) \\
&= \frac{1}{\sigma\sqrt{2\pi}} \sum_{(l,m)\in\{0,1\}^2} (p_1 p_{lm|1} - p_0 p_{lm|0}) e^{\frac{-(r-a_{lm})^2}{N_0}}.
\end{aligned}$$

We are interested in the sign of this expression, so we simplify using the following forms of the restricted constellation points:

$$\begin{aligned}
a_{11} &= \alpha(P_1+P_2), \qquad & a_{01} &= -\beta P_1 + \alpha P_2, \\
a_{10} &= \alpha P_1 - \beta P_2, \qquad & a_{00} &= -\beta(P_1+P_2),
\end{aligned}$$

which gives the following function of x, which can be called a "detection discriminator" since it has the same sign as the original expression for any fixed P_1 and P_2:

$$w(x) = a e^{\frac{2(\alpha+\beta)(P_1+P_2)x}{N_0}} + b e^{\frac{2(\alpha+\beta)P_1 x}{N_0}} + c e^{\frac{2(\alpha+\beta)P_2 x}{N_0}} + d, \tag{3.6}$$

where

$$\begin{aligned}
a &\triangleq \bar{a} e^{-\frac{\alpha^2(P_1+P_2)^2}{N_0}}, \qquad & b &\triangleq \bar{b} e^{-\frac{(\alpha P_1 - \beta P_2)^2}{N_0}}, \\
c &\triangleq \bar{c} e^{-\frac{(\beta P_1 - \alpha P_2)^2}{N_0}}, \qquad & d &\triangleq \bar{d} e^{-\frac{\beta^2(P_1+P_2)^2}{N_0}},
\end{aligned}$$

and

$$\begin{aligned}
\bar{a} &\triangleq p_1 p_{11|1} - p_0 p_{11|0}, \qquad & \bar{b} &\triangleq p_1 p_{10|1} - p_0 p_{10|0}, \\
\bar{c} &\triangleq p_1 p_{01|1} - p_0 p_{01|0}, \qquad & \bar{d} &\triangleq p_1 p_{00|1} - p_0 p_{00|0}.
\end{aligned} \tag{3.7}$$

Using (3.6), we characterize the decision regions as $\mathcal{D}_0 = \{x \in \mathbb{R} \mid w(x) \leq 0\} = \mathcal{D}_1^c$. Note that by the assumptions on p_1, ϵ_1 and ϵ_2, we have that $\bar{d} < 0$, which implies that

$w(x)$ is negative as $x \to -\infty$ for any P_1 and P_2 (hence the MAP rule detects a 0). Thus, we can completely determine these regions by the boundary points between $\mathcal{D}_0$ and $\mathcal{D}_1$. These boundary points are the same as where $w(x)$ crosses from negative to positive, or vice-versa. Thus, it is relevant to analyze the set $\mathcal{X} = \{x \in \mathbb{R} \mid w(x) = 0\}$. Note that applying the results of [23, Corollary 3.2], we know that the size of this set is restricted to $|\mathcal{X}| \in \{0, 1, 2, 3\}$. Unfortunately, there is not a general way to solve for the values $x \in \mathcal{X}$ analytically. However, the problem can be split into three cases which can be analyzed without knowing these values explicitly. The decision regions can be expressed as unions of intervals using these boundary points. For example, if $\mathcal{X} = \{x\}$, then $\mathcal{D}_0 = (-\infty, x]$, whereas if $\mathcal{X} = \{x_1, x_2, x_3\}$, with $x_1 < x_2 < x_3$ then $\mathcal{D}_0 = (-\infty, x_1] \cup [x_2, x_3]$.

3.3.2 Case I: $0 \leq p_1 \leq \frac{\epsilon_1\epsilon_2}{1-\epsilon_1-\epsilon_2+2\epsilon_1\epsilon_2}$

The following proposition characterizes the decision boundaries in this case, which is an essential step of analyzing the error probability.

Proposition 3.2. *In Case I, there are no real solutions to the equation $w(x) = 0$, where $w(x)$ is given in* (3.6).

Proof. Each of the following inequalities hold due to the condition of Case I, and $p_1 \leq 0.5$, $0 < \epsilon_1 \leq \epsilon_2 < 0.5$:

$$\begin{aligned} \bar{a} &= p_1 p_{11|1} - p_0 p_{11|0} = p_1(1-\epsilon_1)(1-\epsilon_2) - (1-p_1)\epsilon_1\epsilon_2 \\ &= (1-\epsilon_1-\epsilon_2+2\epsilon_1\epsilon_2)\left(p_1 - \frac{\epsilon_1\epsilon_2}{1-\epsilon_1-\epsilon_2+2\epsilon_1\epsilon_2}\right) \leq 0 \\ \implies a &\leq 0, \end{aligned}$$

$$\begin{aligned}\bar{b} &= p_1 p_{10|1} - p_0 p_{10|0} = p_1(1-\epsilon_1)\epsilon_2 - (1-p_1)\epsilon_1(1-\epsilon_2) \\ &= (\epsilon_1+\epsilon_2-2\epsilon_1\epsilon_2)\left(p_1 - \frac{\epsilon_1-\epsilon_1\epsilon_2}{\epsilon_1+\epsilon_2-2\epsilon_1\epsilon_2}\right) < 0 \\ \implies b &< 0,\end{aligned}$$

$$\begin{aligned}\bar{c} &= p_1 p_{01|1} - p_0 p_{01|0} = p_1\epsilon_1(1-\epsilon_2) - (1-p_1)(1-\epsilon_1)\epsilon_2 \\ &= (\epsilon_1+\epsilon_2-2\epsilon_1\epsilon_2)\left(p_1 - \frac{\epsilon_2-\epsilon_1\epsilon_2}{\epsilon_1+\epsilon_2-2\epsilon_1\epsilon_2}\right) < 0 \\ \implies c &< 0,\end{aligned}$$

$$\begin{aligned}\bar{d} &= p_1 p_{00|1} - p_0 p_{00|0} = p_1\epsilon_1\epsilon_2 - (1-p_1)(1-\epsilon_1)(1-\epsilon_2) \\ &= (1-\epsilon_1-\epsilon_2+2\epsilon_1\epsilon_2)\left(p_1 - \frac{1-\epsilon_1-\epsilon_2+\epsilon_1\epsilon_2}{1-\epsilon_1-\epsilon_2+2\epsilon_1\epsilon_2}\right) < 0 \\ \implies d &< 0.\end{aligned}$$

Thus $w(x) < 0\ \forall x \in \mathbb{R}$ and $w(x)$ has no real roots. □

Since there are no roots of (3.6), there are also no decision boundaries in Case I. Thus, no matter what the received signal is, the optimal detection will always be $\hat{X} = 0$ (i.e., $\mathcal{D}_0 = \mathbb{R}$). Hence, the error probability is only dependent on the source probability p_1. In this case, the senors are not able to send any useful data, so they should not send anything at all. We conclude that the optimal power allocation and corresponding error performance are expressed as follows:

$$P_1^{*\text{Case I}} = 0, \quad P_2^{*\text{Case I}} = 0, \quad P_e^{*\text{Case I}} = p_1.$$

3.3.3 Case II: $\frac{\epsilon_1\epsilon_2}{1-\epsilon_1-\epsilon_2+2\epsilon_1\epsilon_2} < p_1 \leq \frac{\epsilon_1-\epsilon_1\epsilon_2}{\epsilon_1+\epsilon_2-2\epsilon_1\epsilon_2}$

Using a similar approach as in the proof of Proposition 3.2, we can show the following properties about the coefficients of (3.6) in this case:

$$a > 0, \quad b \leq 0, \quad c \leq 0, \quad d < 0.$$

We use these observations to characterize the decision regions in this case through the following proposition.

Proposition 3.3. *In Case II, there is exactly one real root to $w(x)$ in (3.6) for any $P_1, P_2 > 0$. Further, this root is also a decision boundary between $\mathcal{D}_0$ and $\mathcal{D}_1$.*

Proof. Let $P_1, P_2 > 0$. First we show that there exists at least one real solution to $w(x) = 0$. We use the fact that $w(x)$ is continuous and that its asymptotic behaviours are:

$$\lim_{x\to-\infty} w(x) = d, \quad \lim_{x\to\infty} w(x) = \infty.$$

Since $d < 0$ and $w(x)$ is continuous, it must have at least one root. Next we show that $w(x)$ can have at most one root by showing that once it becomes non-negative, the derivative is always strictly positive, so it can never have another zero. Assume $w(x) \geq 0$, we then have:

$$\begin{aligned}\frac{dw}{dx} &= \frac{2(\alpha+\beta)}{N_0}\left((P_1+P_2)ae^{\frac{2(\alpha+\beta)(P_1+P_2)x}{N_0}} + P_1be^{\frac{2(\alpha+\beta)P_1x}{N_0}} + P_2ce^{\frac{2(\alpha+\beta)P_2x}{N_0}}\right)\\ &= \frac{2(\alpha+\beta)}{N_0}\left((P_1+P_2)w(x) - P_2be^{\frac{2(\alpha+\beta)P_1x}{N_0}} - P_1ce^{\frac{2(\alpha+\beta)P_2x}{N_0}} - (P_1+P_2)d\right) > 0.\end{aligned}$$

This shows that $w(x)$ has exactly one real root, and this root is a boundary point between $\mathcal{D}_0$ and $\mathcal{D}_1$ as desired. □

Since there is a unique decision boundary, the decision regions will have the form $\mathcal{D}_0 = (-\infty, x]$, where x is the root of (3.6) corresponding to P_1 and P_2. We can express the error probability at any $P_1, P_2 > 0$ using the following expression:

$$P_e(P_1, P_2) = \sum_{(l,m)\in\{0,1\}^2} (p_1 p_{lm|1} - p_0 p_{lm|0}) Q\left(\frac{a_{lm} - x}{\sigma}\right) + p_0 p_{lm|0}, \tag{3.8}$$

where Q is the Gaussian tail distribution function, a_{lm} are the constellation points and x is the root of (3.6) corresponding to P_1 and P_2. We also give the following upper bound on the error probability.

Proposition 3.4. *If $P_1, P_2 > 0$ and x is the corresponding unique decision boundary, then for any $\hat{x} \in \mathbb{R}$, the following is an upper bound on the error probability*

$$P_{e,\hat{x}}^{UB}(P_1, P_2) \triangleq \sum_{(l,m)\in\{0,1\}^2} (p_1 p_{lm|1} - p_0 p_{lm|0}) Q\left(\frac{a_{lm} - \hat{x}}{\sigma}\right) + p_0 p_{lm|0}. \tag{3.9}$$

Proof. This expression corresponds to the error probability associated to using a decision boundary $\hat{x}$. Since the true decision boundary, x, from the MAP detection rule is optimal, this must be an upper bound. □

We now define the following functions, where $\bar{a}$, $\bar{b}$ and $\bar{c}$ are as defined in (3.7).

$$K(P_1) \triangleq \frac{N_0}{2(\alpha+\beta)P_1} \ln \frac{\bar{a}}{-\bar{c}} - \frac{\alpha-\beta}{2} P_1, \tag{3.10}$$

$$L(P_2) \triangleq \frac{N_0}{2(\alpha+\beta)P_2} \ln \frac{\bar{a}}{-\bar{b}} - \frac{\alpha-\beta}{2} P_2. \tag{3.11}$$

Note that these are well defined if $\bar{c} \neq 0$ and $\bar{b} \neq 0$, respectively. These functions are used often in the optimization proof, such as the in following proposition.

Proposition 3.5. *For any $P_1, P_2 > 0$ the following two statements are true if $\bar{b} \neq 0$ and $\bar{c} \neq 0$, respectively:*

$$x \lesseqgtr \alpha P_1 - L(P_2) \implies ae^{\frac{2(\alpha+\beta)P_2 x}{N_0}} + b \lesseqgtr 0, \tag{3.12}$$

$$x \lesseqgtr \alpha P_2 - K(P_1) \implies ae^{\frac{2(\alpha+\beta)P_1 x}{N_0}} + c \lesseqgtr 0, \tag{3.13}$$

where the symbol $\lesseqgtr$ means that the statements hold for any of the relations $<$, $>$ or $=$, consistently in the each line.

Proof. We prove (3.12) using the definition of L from (3.11). Assume $P_1, P_2 > 0$, $\bar{b} \neq 0$, then if x satisfies

$$\begin{aligned} & x \lesseqgtr \alpha P_1 - L(P_2) \\ \implies & ae^{\frac{2(\alpha+\beta)P_2 x}{N_0}} + b \lesseqgtr ae^{\frac{2(\alpha+\beta)P_2(\alpha P_1 - L(P_2))}{N_0}} + b, \end{aligned}$$

where

$$\begin{aligned} ae^{\frac{2(\alpha+\beta)P_2(\alpha P_1 - L(P_2))}{N_0}} + b &= ae^{\frac{2\alpha(\alpha+\beta)P_1P_2 - 2(\alpha+\beta)P_2 L(P_2)}{N_0}} + b \\ &= ae^{\frac{2\alpha(\alpha+\beta)P_1P_2 + (\alpha^2-\beta^2)P_2^2}{N_0} - \ln\frac{\bar{a}}{-\bar{b}}} + b \\ &= -\underbrace{\bar{b}e^{\frac{-(\alpha P_1 - \beta P_2)^2}{N_0}}}_{=b} \underbrace{\frac{a}{\bar{a}} e^{\frac{\alpha^2(P_1+P_2)^2}{N_0}}}_{=1} + b \\ &= -b + b = 0. \end{aligned}$$

The proof of (3.13) follows the same steps, using the definition of K from (3.10). □

This result is used to determine the properties of the decision boundaries by its

influence on the expression in (3.6), as seen in the following proposition.

Proposition 3.6. *In Case II, for $P_1, P_2 > 0$, if x is the corresponding root of* (3.6)*, $\bar{b} \neq 0$ and $\bar{c} \neq 0$, then the following two inequalities hold:*

$$x > \alpha P_1 - L(P_2), \tag{3.14}$$

$$x > \alpha P_2 - K(P_1). \tag{3.15}$$

Proof. Using Proposition 3.5, we have

$$\begin{aligned} x \leq \alpha P_1 - L(P_2) &\implies ae^{\frac{2(\alpha+\beta)P_2 x}{N_0}} + b \leq 0 \implies w(x) < 0, \\ x \leq \alpha P_2 - K(P_1) &\implies ae^{\frac{2(\alpha+\beta)P_1 x}{N_0}} + c \leq 0 \implies w(x) < 0. \end{aligned}$$

Hence, x could not be a zero of (3.6) in either of these cases, showing the desired result. □

As seen in the above proof, a direct consequence of applying Proposition 3.5 gives a lower bound on the values of the decision boundary. We also require the following additional result about these lower bounds.

Proposition 3.7. *In Case II, let $\bar{P}_1$, P_1', $\bar{P}_2$ and P_2' be arbitrary real numbers such that $0 < \bar{P}_1 < P_1'$, $0 < \bar{P}_2 < P_2'$. Then the root of* (3.6)*, x, satisfies the following two inequalities as a function of P_2 and P_1, respectively:*

$$\inf_{P_1 \in [\bar{P}_1, P_1']} x - \alpha P_1 + L(P_2) > 0, \quad \text{for } \bar{b} \neq 0, P_2 > 0, \tag{3.16}$$

$$\inf_{P_2 \in [\bar{P}_2, P_2']} x - \alpha P_2 + K(P_1) > 0, \quad \text{for } \bar{c} \neq 0, P_1 > 0. \tag{3.17}$$

Proof. To prove (3.16), let $\bar{P}_1$ and P_1' be arbitrary real numbers such that $0 < \bar{P}_1 < P_1'$. For any $P_1 \in [\bar{P}_1, P_1']$ and $P_2 > 0$, the root of (3.6), x, satisfies $x - \alpha P_1 + L(P_2) > 0$ from (3.14) of Proposition 3.6; therefore

$$\inf_{P_1 \in [\bar{P}_1, P_1']} x - \alpha P_1 + L(P_2) \geq 0. \tag{3.18}$$

Next, we define the function

$$w_L(x, P_1) \triangleq a e^{\frac{2(\alpha+\beta)(P_1+P_2)x}{N_0}} + b e^{\frac{2(\alpha+\beta)P_1 x}{N_0}},$$

which is uniformly continuous in both P_1 over $[\bar{P}_1, P_1']$ and x over $[\alpha\bar{P}_1 - L(P_2), \alpha P_1' - L(P_2)]$. By (3.12), $w_L(x, P_1) = 0$ at all points (x, P_1) on the line $x = \alpha P_1 - L(P_2)$. Thus, for any $d' > 0$, there exists $\delta > 0$ such that for any $P_1 \in [\bar{P}_1, P_1']$ and x that satisfies

$$\alpha P_1 - L(P_2) < x < \alpha P_1 - L(P_2) + \delta, \tag{3.19}$$

we have

$$w_L(x, P_1) < d'.$$

If (3.18) holds with equality, then there exists $P_1 \in [\bar{P}_1, P_1']$ with corresponding root of (3.6), x, that satisfies (3.19). In particular, for $d' = -\bar{d}e^{-\frac{\beta^2(P_1'+P_2)^2}{N_0}}$, which is a positive constant with respect to P_1 and x, we obtain

$$\begin{aligned}
& a e^{\frac{2(\alpha+\beta)(P_1+P_2)x}{N_0}} + b e^{\frac{2(\alpha+\beta)P_1 x}{N_0}} < -\bar{d}e^{-\frac{\beta^2(P_1'+P_2)^2}{N_0}} < -d \\
\implies & a e^{\frac{2(\alpha+\beta)(P_1+P_2)x}{N_0}} + b e^{\frac{2(\alpha+\beta)P_1 x}{N_0}} + d < 0 \\
\implies & w(x) < 0,
\end{aligned}$$

where the last inequality holds because $c < 0$ in the expression of $w(x)$ in (3.6) for Case II. This contradicts x being the root of (3.6), completing the proof. The proof of (3.17) is omitted as it follows the same steps as above. □

This result shows that the decision boundary is bounded away from the lower bound over any finite interval of power allocation. This is a small, but essential result for the main proof as it guarantees that the sequence of powers we use will grow large enough to reach the necessary value. The results of Propositions 3.5, 3.6 and 3.7 are used to prove the following main theorem characterizing the optimization in this case.

Theorem 3.8. *In Case II, $P_e(P_1, P_2)$ is decreasing in P_1 and P_2 for $P_1, P_2 > 0$.*

Proof. We will show the following two statements:

1. If $0 < P_1 < P_1'$, $0 < P_2$, then $P_e(P_1, P_2) > P_e(P_1', P_2)$.

2. If $0 < P_1$, $0 < P_2 < P_2'$, then $P_e(P_1, P_2) > P_e(P_1, P_2')$.

To show Statement 1, fix $0 < P_1 < P_1'$, $0 < P_2$. Let x and x' be the roots of (3.6) corresponding to the pairs (P_1, P_2) and (P_1', P_2), respectively. We define the following sequence $\{P_{1,i}\}_{i=0}^{\infty}$ recursively.

$$P_{1,0} = P_1, \qquad P_{1,i+1} = \begin{cases} P_1', & \bar{b} = 0 \\ \frac{1}{\alpha}\big(x_i + L(P_2)\big), & \bar{b} \neq 0 \end{cases}$$

where x_i is the root to (3.6) corresponding to $P_{1,i}$, $\bar{b}$ is from (3.7) and L is as defined in (3.11). Note that if $\bar{b} \neq 0$ and $P_{1,i} < P_1'$, applying (3.16) gives

$$P_{1,i+1} - P_{1,i} = \frac{1}{\alpha}\big(x_i + L(P_2)\big) - P_{1,i}$$

$$\geq \frac{1}{\alpha} \inf_{\bar{P}_1 \in [P_1, P_1']} \bar{x} - \alpha \bar{P}_1 + L(P_2) \overset{(3.16)}{>} 0,$$

where $\bar{x}$ denotes the root of (3.6) for $(\bar{P}_1, P_2)$. This shows that the sequence $\{P_{1,i}\}_{i=0}^{\infty}$ increases by at least this constant if $P_{1,i} < P_1'$. Therefore there exists i' large enough such that $P_{1,i'} \geq P_1'$. Hence, it is sufficient to show that for all i

$$P_e(P_{1,i}, P_2) - P_e(P_{1,i+1}, P_2) > 0.$$

Using the upper bound in Proposition 3.4, it is sufficient to show

$$\begin{aligned} & P_e(P_{1,i}, P_2) - P_{e,x_i}^{\mathrm{UB}}(P_{1,i+1}, P_2) > 0 \\ \iff & P_{e,x_i}^{\mathrm{UB}}(P_{1,i}, P_2) - P_{e,x_i}^{\mathrm{UB}}(P_{1,i+1}, P_2) > 0, \end{aligned}$$

since $P_e(P_{1,i}, P_2) = P_{e,x_i}^{\mathrm{UB}}(P_{1,i}, P_2)$ by its definition.

It is now sufficient to show P_{e,x_i}^{UB} is decreasing in P_1 over $(P_{1,i}, P_{1,i+1})$. The derivative of this expression is:

$$\frac{dP_{e,x_i}^{\mathrm{UB}}}{dP_1} = \frac{-e^{\frac{-x_i^2}{N_0}}}{\sigma\sqrt{2\pi}} \left(\alpha \left(a e^{\frac{2\alpha(P_1+P_2)x_i}{N_0}} + b e^{\frac{2(\alpha P_1 - \beta P_2)x_i}{N_0}} \right) - \beta \left(c e^{\frac{2(-\beta P_1 + \alpha P_2)x_i}{N_0}} + d e^{\frac{-2\beta(P_1+P_2)x_i}{N_0}} \right) \right). \tag{3.20}$$

If $\bar{b} = 0$, this derivative is negative for any P_1 (since this is equivalent to $b = 0$). If $\bar{b} \neq 0$, then we apply (3.12) from Proposition 3.5 and conclude that

$$\begin{aligned} & P_1 < P_{1,i+1} = \frac{1}{\alpha}\big(x_i + L(P_2)\big) \\ \iff & x_i > \alpha P_1 - L(P_2) \end{aligned}$$

$$\begin{aligned}
&\implies ae^{\frac{2(\alpha+\beta)P_2x_i}{N_0}} + b > 0 \\
&\implies ae^{\frac{2\alpha(P_1+P_2)x_i}{N_0}} + be^{\frac{2(\alpha P_1-\beta P_2)x_i}{N_0}} > 0 \\
&\implies \frac{dP_{e,x_i}^{\text{UB}}}{dP_1} < 0.
\end{aligned}$$

The proof of Statement 2 is omitted as it follows the exact same steps as above, replacing the roles of $L(P_2)$ with $K(P_1)$, $\bar{b}$ with $\bar{c}$, and applying (3.13) and (3.17) instead of (3.12) and (3.16). □

Thus the optimal power allocation and corresponding error performance can be expressed as follows, where $p_{lm|i}$ are as given in (3.2), x^* is the root to (3.6) and a_{lm}^* are the constellation points corresponding to P_1^* and P_2^*.

$$\begin{aligned}
P_1^{*\text{Case II}} &= \sqrt{P_1^{\max}}, \quad P_2^{*\text{Case II}} = \sqrt{P_2^{\max}}, \\
P_e^{*\text{Case II}} &= \sum_{(l,m)\in\{0,1\}^2} (p_1 p_{lm|1} - p_0 p_{lm|0}) Q\left(\frac{a_{lm}^* - x^*}{\sigma}\right) + p_0 p_{lm|0}.
\end{aligned} \tag{3.21}$$

The corresponding optimal decision region is $\mathcal{D}_0 = (-\infty, x^*]$.

3.3.4 Case III: $\frac{\epsilon_1-\epsilon_1\epsilon_2}{\epsilon_1+\epsilon_2-2\epsilon_1\epsilon_2} < p_1 \leq 0.5$

First note that the condition of this case implies $\epsilon_1 \neq \epsilon_2$. Also using the same reasoning as in the proof of Proposition 3.2, we make the following observations about the coefficients of $w(x)$ in (3.6):

$$a > 0, \quad b > 0, \quad c < 0, \quad d < 0.$$

We define the following functions of P_1, where $\bar{a}$, $\bar{b}$, $\bar{c}$ and $\bar{d}$ are as defined in (3.7):

$$\tilde{P}_2(P_1) \triangleq \frac{N_0}{2(\alpha+\beta)^2 P_1} \ln \frac{\bar{a}\bar{d}}{\bar{b}\bar{c}}, \tag{3.22}$$

$$K_\alpha(P_1) \triangleq \frac{N_0}{2(\alpha+\beta)P_1} \ln \frac{\bar{a}}{-\bar{c}} - \frac{\alpha-\beta}{2} P_1, \tag{3.23}$$

$$K_\beta(P_1) \triangleq \frac{N_0}{2(\alpha+\beta)P_1} \ln \frac{-\bar{d}}{\bar{b}} + \frac{\alpha-\beta}{2} P_1. \tag{3.24}$$

The condition for Case III combined with the other assumptions on the problem's parameters ensure that these functions are real valued for all $P_1 > 0$. Also note that expanding (3.22) at $P_1 = \sqrt{P_1^{\max}}$ gives the expression in (3.3). We begin with the following result which is used numerous times over the analysis of this case.

Proposition 3.9. *For any $P_1, P_2 > 0$ the following two statements are true:*

$$x \lessgtr \alpha P_2 - K_\alpha(P_1) \implies a e^{\frac{2(\alpha+\beta)P_1 x}{N_0}} + c \lessgtr 0, \tag{3.25}$$

$$x \lessgtr -\beta P_2 + K_\beta(P_1) \implies b e^{\frac{2(\alpha+\beta)P_1 x}{N_0}} + d \lessgtr 0. \tag{3.26}$$

Proof. These statements follow directly from rearranging these equations and using the definitions of K_α and K_β in (3.23) and (3.24), respectively. The steps are the same as in the proof of Proposition 3.5. □

This result is used for many of the remaining proofs in this section. For example, (3.25) and (3.26) directly apply to give the following bounds on the roots of (3.6).

Proposition 3.10. *In Case III, for any $P_1, P_2 > 0$, if x is a corresponding root*

of (3.6)*, then it must satisfy*

$$\begin{cases} x \in \big(\alpha P_2 - K_\alpha(P_1), -\beta P_2 + K_\beta(P_1)\big), & P_2 < \tilde{P}_2(P_1) \\ x = \alpha P_2 - K_\alpha(P_1) = -\beta P_2 + K_\beta(P_1), & P_2 = \tilde{P}_2(P_1) \\ x \in \big(-\beta P_2 + K_\beta(P_1), \alpha P_2 - K_\alpha(P_1)\big), & P_2 > \tilde{P}_2(P_1) \end{cases} \tag{3.27}$$

Proof. First we note that these intervals are valid and have non-zero length since

$$\begin{aligned} -\beta P_2 + K_\beta(P_1) - \big(\alpha P_2 - K_\alpha(P_1)\big) &= K_\alpha(P_1) + K_\beta(P_1) - (\alpha+\beta)P_2 \\ &= (\alpha+\beta)\big(\tilde{P}_2(P_1) - P_2\big), \end{aligned}$$

so

$$P_2 \lessgtr \tilde{P}_2(P_1) \implies \alpha P_2 - K_\alpha(P_1) \lessgtr -\beta P_2 + K_\beta(P_1).$$

If x is outside these intervals, then we apply (3.25) and (3.26) from Proposition 3.9 to see that $w(x) \neq 0$, so x could not be a root of (3.6). □

For a fixed $P_1 > 0$, the bounds described in Proposition 3.10 form two lines with respect to P_2, intersecting at $\tilde{P}_2(P_1)$. These bounds are illustrated in Figure 3.2.

Another important property about these bounds is given in the following proposition, which is similar to the results of Proposition 3.7.

Proposition 3.11. *In Case III, for any $P_1 > 0$, let $\bar{P}_2$ and P_2' be arbitrary real numbers such that $0 < \bar{P}_2 < P_2' < \tilde{P}_2(P_1)$, where $\tilde{P}_2(P_1)$ is as defined in* (3.22)*. Then the root of* (3.6)*, x, satisfies the following two inequalities:*

$$\inf_{P_2 \in [\bar{P}_2, P_2']} x - \alpha P_2 + K_\alpha(P_1) > 0, \tag{3.28}$$

$$\inf_{P_2 \in [\bar{P}_2, P_2']} K_\beta(P_1) - \beta P_2 - x > 0. \tag{3.29}$$

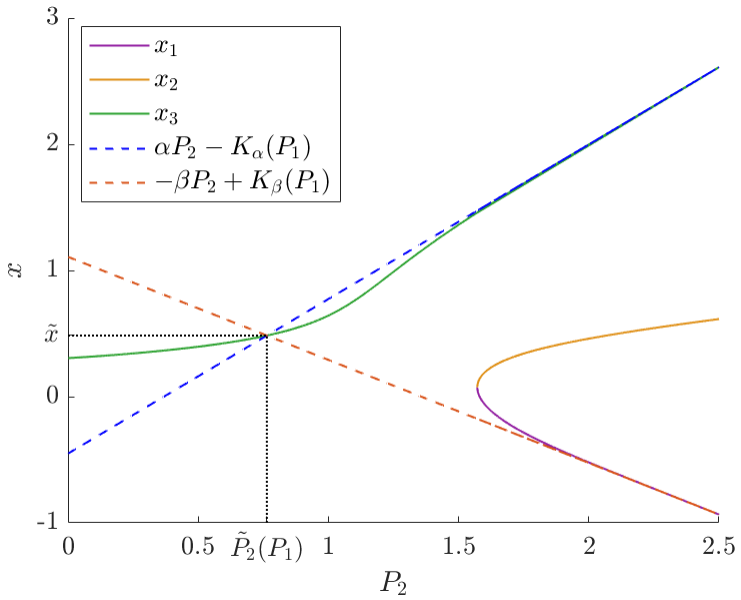


Figure 3.2: The roots of (3.6), x_1, x_2 and x_3, as a function of P_2 in Case III ($p_1 = 0.4$, $\epsilon_1 = 0.01$, $\epsilon_2 = 0.05$, $N_0 = 1$, $P_1 = 1$). The decision regions can be read as: for $P_2 < P_2^{\text{thr}} \overset{\triangle}{\approx} 1.6$, $\mathcal{D}_0 = (-\infty, x_3]$. Otherwise, $\mathcal{D}_0 = (-\infty, x_1] \cup [x_2, x_3]$.

Proof. The details of this proof follow the same steps as Proposition 3.7. □

The above results are referred to multiple times in the remaining proofs in this section. For example, Propositions 3.9 and 3.10 are used to prove the following result about the number of decision boundaries.

Proposition 3.12. *In Case III, for any $P_1, P_2 > 0$, there will be one or three boundary points between $\mathcal{D}_0$ and $\mathcal{D}_1$. Further, if $P_2 \in [0, \tilde{P}_2(P_1)]$ there is exactly one boundary point between $\mathcal{D}_0$ and $\mathcal{D}_1$.*

Proof. Let $P_1, P_2 > 0$. We have the following asymptotic behaviours of $w(x)$ in (3.6):

$$\lim_{x \to -\infty} w(x) = d, \quad \lim_{x \to \infty} w(x) = \infty.$$

Since $d < 0$, and w is continuous, we conclude that there must be an odd number of crossing points. Combing this fact with [23, Corollary 3.2] yields that there are one or three crossing points. Now let $P_2 \in [0, \tilde{P}_2(P_1)]$. We assume $w(x) = 0$, and perform the following derivative analysis:

$$\begin{aligned}\frac{dw}{dx} &= \frac{2(\alpha+\beta)}{N_0}\left((P_1+P_2)ae^{\frac{2(\alpha+\beta)(P_1+P_2)x}{N_0}} + P_1be^{\frac{2(\alpha+\beta)P_1x}{N_0}} + P_2ce^{\frac{2(\alpha+\beta)P_2x}{N_0}}\right)\\ &= \frac{2(\alpha+\beta)}{N_0}\left((P_1+P_2)w(x) - P_2be^{\frac{2(\alpha+\beta)P_1x}{N_0}} - P_1ce^{\frac{2(\alpha+\beta)P_2x}{N_0}} - (P_1+P_2)d\right)\\ &= \frac{-2(\alpha+\beta)}{N_0}\left(P_2(be^{\frac{2(\alpha+\beta)P_1x}{N_0}} + d) + P_1(ce^{\frac{2(\alpha+\beta)P_2x}{N_0}} + d)\right) > 0.\end{aligned}$$

This last inequality is true since Propositions 3.9 and 3.10 imply

$$be^{\frac{2(\alpha+\beta)P_1x}{N_0}} + d \leq 0.$$

Therefore $w(x)$ is strictly increasing at any zero, and hence must only have one root, and this root must be a boundary point between $\mathcal{D}_0$ and $\mathcal{D}_1$ as desired. □

Using Proposition 3.12, we conclude that the error probability in Case III can have two possible expressions based on the number of decision boundaries. If there is a single decision boundary, x, then the error expression will have the same form as in (3.8). If there are three decision boundaries, $x_1 < x_2 < x_3$, then the error expression has the following form:

$$\begin{aligned}P_e(P_1,P_2) = \sum_{(l,m)\in\{0,1\}^2} (p_1p_{lm|1} - p_0p_{lm|0})\Bigg(Q\left(\frac{a_{lm}-x_1}{\sigma}\right)\\ - Q\left(\frac{a_{lm}-x_2}{\sigma}\right) + Q\left(\frac{a_{lm}-x_3}{\sigma}\right)\Bigg) + p_0p_{lm|0}. \qquad (3.30)\end{aligned}$$

Now we define the following two functions of x, for any fixed P_1 and P_2:

$$g(x) \triangleq (p_1 p_{11|1} - p_0 p_{11|0}) Q\left(\frac{\alpha P_1 + \alpha P_2 - x}{\sigma}\right) + (p_1 p_{01|1} - p_0 p_{01|0}) Q\left(\frac{-\beta P_1 + \alpha P_2 - x}{\sigma}\right),$$
$$h(x) \triangleq (p_1 p_{10|1} - p_0 p_{10|0}) Q\left(\frac{\alpha P_1 - \beta P_2 - x}{\sigma}\right) + (p_1 p_{00|1} - p_0 p_{00|0}) Q\left(\frac{-\beta P_1 - \beta P_2 - x}{\sigma}\right).$$

These functions are significant because they decompose the error expression. For example, if P_1 and P_2 are fixed with a unique corresponding root of (3.6), x, the error probability can be written as

$$P_e(P_1, P_2) = g(x) + h(x) + \sum_{(l,m)\in\{0,1\}^2} p_0 p_{lm|0}.$$

Further, they have the following unique minimizers with analytic expressions corresponding to the bounds given in Proposition 3.10.

Proposition 3.13. *In Case III, for any $P_1, P_2 > 0$, $g(x)$ is minimized at $x = \alpha P_2 - K_\alpha(P_1)$ and $h(x)$ is minimized at $x = -\beta P_2 + K_\beta(P_1)$.*

Proof. Let $P_1, P_2 > 0$. The following are expressions for the derivatives:

$$\frac{dg}{dx} = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{N_0}} \left(a e^{\frac{2\alpha(P_1+P_2)x}{N_0}} + c e^{\frac{2(-\beta P_1 + \alpha P_2)x}{N_0}} \right)$$
$$\frac{dh}{dx} = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{N_0}} \left(b e^{\frac{2(\alpha P_1 - \beta P_2)x}{N_0}} + d e^{\frac{-2\beta(P_1+P_2)x}{N_0}} \right).$$

Next, we apply the results of Proposition 3.9 to show

$$x \lesseqgtr \alpha P_2 - K_\alpha(P_1) \implies 0 \gtreqless a e^{\frac{2(\alpha+\beta)P_1 x}{N_0}} + c$$
$$\implies 0 \gtreqless a e^{\frac{2\alpha(P_1+P_2)x}{N_0}} + c e^{\frac{2(-\beta P_1 + \alpha P_2)x}{N_0}}$$

$$\implies 0 \gtrless \frac{dg}{dx},$$

and

$$x \lessgtr -\beta P_2 + K_\beta(P_1) \implies 0 \gtrless be^{\frac{2(\alpha+\beta)P_1 x}{N_0}} + d$$
$$\implies 0 \gtrless be^{\frac{2(\alpha P_1 - \beta P_2)x}{N_0}} + de^{\frac{-2\beta(P_1+P_2)x}{N_0}}$$
$$\implies 0 \gtrless \frac{dh}{dx}.$$

□

The properties of the functions g and h make them useful in proving the following theorem about global optimality with respect to P_2.

Theorem 3.14. *In Case III,* $P_e\big(P_1, \tilde{P}_2(P_1)\big) < P_e(P_1, P_2)$, $\forall P_1 > 0, P_2 \neq \tilde{P}_2(P_1)$.

Proof. Fix $P_1 > 0$. Let $\tilde{a}_{lm}$ be the constellation points and $\tilde{x}$ be the root of (3.6) corresponding to P_1 and $\tilde{P}_2(P_1)$. Let $P_2 \neq \tilde{P}_2(P_1)$. Let a_{lm} be the constellation points and $\mathcal{X}$ denote the set of roots of (3.6) corresponding to P_1 and P_2. First we analyze the case where $|\mathcal{X}| = 1$. Let $\mathcal{X} = \{x\}$. In this case, the error expression, $P_e(P_1, P_2)$, will have the same form as given in (3.8). Also, since there is a unique root, $\tilde{x}$, at $\tilde{P}_2(P_1)$, $P_e\big(P_1, \tilde{P}_2(P_1)\big)$ takes the form of (3.8) as well. Therefore, we must show

$$\sum_{(l,m)\in\{0,1\}^2} (p_1 p_{lm|1} - p_0 p_{lm|0}) Q\left(\frac{\tilde{a}_{lm} - \tilde{x}}{\sigma}\right) < \sum_{(l,m)\in\{0,1\}^2} (p_1 p_{lm|1} - p_0 p_{lm|0}) Q\left(\frac{a_{lm} - x}{\sigma}\right)$$
$$\iff g\big(\alpha P_2 - K_\alpha(P_1)\big) + h\big(-\beta P_2 + K_\beta(P_1)\big) < g(x) + h(x).$$

This result follows immediately from Proposition 3.13, since $x \neq \alpha P_2 - K_\alpha(P_1)$ and $x \neq -\beta P_2 + K_\beta(P_1)$ for $P_2 \neq \tilde{P}_2(P_1)$ from Proposition 3.10.

For the case that $|\mathcal{X}| = 3$, we represent this as $\mathcal{X} = \{x_1, x_2, x_3\}$ such that $x_1 < x_2 < x_3$. Further we note that the decision regions are represented in terms of these boundaries as $\mathcal{D}_1 = (x_1, x_2) \cup (x_3, \infty)$. This gives rise to the following expression for the error probability, also noting that it can be expressed in terms of g and h:

$$\begin{aligned} P_e(P_1, P_2) = &\sum_{(l,m)\in\{0,1\}^2} (p_1 p_{lm|1} - p_0 p_{lm|0}) \Bigg(Q\bigg(\frac{a_{lm} - x_1}{\sigma}\bigg) \\ &- Q\bigg(\frac{a_{lm} - x_2}{\sigma}\bigg) + Q\bigg(\frac{a_{lm} - x_3}{\sigma}\bigg)\Bigg) + p_0 p_{lm|0} \\ &= g(x_1) - g(x_2) + g(x_3) + h(x_1) - h(x_2) + h(x_3) + \sum_{(l,m)\in\{0,1\}^2} p_0 p_{lm|0}. \end{aligned}$$

Hence we must show

$$\begin{aligned} &g\big(\alpha P_2 - K_\alpha(P_1)\big) + h\big(-\beta P_2 + K_\beta(P_1)\big) < \\ &\quad g(x_1) - g(x_2) + g(x_3) + h(x_1) - h(x_2) + h(x_3). \end{aligned} \tag{3.31}$$

First note that by applying Proposition 3.13, we have

$$g\big(\alpha P_2 - K_\alpha(P_1)\big) + h\big(-\beta P_2 + K_\beta(P_1)\big) < g(x_3) + h(x_1). \tag{3.32}$$

Now we will show

$$g(x_1) - g(x_2) + h(x_3) - h(x_2) \geq 0. \tag{3.33}$$

Since $|\mathcal{X}| = 3$ it is implied that $P_2 > \tilde{P}_2(P_1)$ by taking the contra-positive of Proposition 3.12. Then we can apply Proposition 3.10 which gives the bounds $x_2 \in \big(-\beta P_2 + K_\beta(P_1), \alpha P_2 - K_\alpha(P_1)\big)$. Finally, applying the same derivative analysis

as in Proposition 3.13 shows that $g(x)$ is decreasing on $[x_1, x_2]$ and $h(x)$ is increasing on $[x_2, x_3]$. This implies (3.33). Combining (3.32) with (3.33) implies (3.31). □

This gives an analytic expression for the global minimizer of the error probability with respect to P_2. The main intuition behind this result relies on splitting the error function into two simpler parts, which are shown to be minimized along the same linear bounds in Figure 3.2. Since the two lines intersect at $\tilde{P}_2(P_1)$, this value (along with its corresponding decision boundary) minimizes both parts of the error function, hence the entire error function is also minimized.

Although we have found a global minimum with respect to P_2, it is still necessary to determine the optimal constellation when there is not enough power to use the global minimum, i.e., if $\sqrt{P_2^{\max}} < \tilde{P}_2(P_1)$. The following theorem provides a sufficient result for this situation.

Theorem 3.15. *In Case III, if* $0 < P_1$, $0 < P_2 < P_2' < \tilde{P}_2(P_1)$, *then* $P_e(P_1, P_2) > P_e(P_1, P_2')$.

Proof. Fix $0 < P_1$, $0 < P_2 < P_2' < \tilde{P}_2$. Let x and x' be the roots of (3.6) corresponding to P_2 and P_2', respectively. We define the following sequence $\{P_{2,i}\}_{i=0}^{\infty}$ recursively.

$$\begin{aligned} P_{2,0} &= P_2, \\ P_{2,i+1} &= \min\left(\frac{1}{\alpha}\big(x_i + K_\alpha(P_1)\big), \frac{1}{\beta}\big(K_\beta(P_1) - x_i\big)\right), \end{aligned}$$

where x_i is the root of (3.6) corresponding to $P_{2,i}$. Note that (3.27) implies $P_{2,i} \leq \tilde{P}_2$ for any $i \geq 0$, so there is always a unique x_i. If $P_{2,i} < P_2'$, applying (3.28) and (3.29)

means one of the following two statements must be true:

$$\begin{aligned} P_{2,i+1} - P_{2,i} &= \frac{1}{\alpha}\big(x_i + K_\alpha(P_1)\big) - P_{2,i} \\ &\geq \underbrace{\frac{1}{\alpha} \inf_{\bar{P}_2 \in [P_2, P_2']} \bar{x} - \alpha \bar{P}_2 + K_\alpha(P_1)}_{\triangleq K'_\alpha} \overset{(3.28)}{>} 0, \end{aligned}$$

or

$$\begin{aligned} P_{2,i+1} - P_{2,i} &= \frac{1}{\beta}\big(K_\beta(P_1) - x_1\big) - P_{2,i} \\ &\geq \underbrace{\frac{1}{\beta} \inf_{\bar{P}_2 \in [P_2, P_2']} K_\beta(P_1) - \beta \bar{P}_2 - \bar{x}}_{\triangleq K'_\beta} \overset{(3.29)}{>} 0, \end{aligned}$$

so

$$P_{2,i+1} - P_{2,i} \geq \min(K'_\alpha, K'_\beta) > 0.$$

where $\bar{x}$ denotes the root of (3.6) for $(P_1, \bar{P}_2)$. This means that the sequence $\{P_{2,i}\}_{i=0}^{\infty}$ increases by at least this fixed constant if $P_{2,i} < P_2'$. Therefore, there exists i' large enough such that $P_{2,i'} \geq P_2'$. Hence, it is sufficient to show that for all i

$$P_e(P_1, P_{2,i}) - P_e(P_1, P_{2,i+1}) > 0.$$

Using the upper bound in Proposition 3.4 (which is valid for exactly the same reasons as before), it is sufficient to show

$$\begin{aligned} & P_e(P_1, P_{2,i}) - P_{e,x_i}^{\mathrm{UB}}(P_1, P_{2,i+1}) > 0 \\ \iff\ & P_{e,x_i}^{\mathrm{UB}}(P_1, P_{2,i}) - P_{e,x_i}^{\mathrm{UB}}(P_1, P_{2,i+1}) > 0, \end{aligned}$$

since $P_e(P_1, P_{2,i}) = P_{e,x_i}^{\mathrm{UB}}(P_1, P_{2,i})$ by definition in (3.9). It is now sufficient to show

P^{UB}_{e,x_i} is decreasing in P_2 over $(P_{2,i}, P_{2,i+1})$. We have

$$\frac{dP^{\text{UB}}_{e,x_i}}{dP_2} = \frac{-e^{\frac{-x_i^2}{N_0}}}{\sigma\sqrt{2\pi}}\left(\alpha\left(ae^{\frac{2\alpha(P_1+P_2)x_i}{N_0}} + ce^{\frac{2(-\beta P_1+\alpha P_2)x_i}{N_0}}\right) - \beta\left(be^{\frac{2(\alpha P_1-\beta P_2)x_i}{N_0}} + de^{\frac{-2\beta(P_1+P_2)x_i}{N_0}}\right)\right),$$

where applying Proposition 3.9 gives

$$\begin{aligned} P_2 < P_{2,i+1} &\leq \frac{1}{\alpha}\big(x_i + K_\alpha(P_1)\big) \\ &\implies x_i > \alpha P_2 - K_\alpha(P_1) \\ &\implies 0 < ae^{\frac{2(\alpha+\beta)P_1 x_i}{N_0}} + c \\ &\implies 0 < ae^{\frac{2\alpha(P_1+P_2)x_i}{N_0}} + ce^{\frac{2(-\beta P_1+\alpha P_2)x_i}{N_0}}, \end{aligned}$$

and

$$\begin{aligned} P_2 < P_{2,i+1} &\leq \frac{1}{\beta}\big(K_\beta(P_1) - x_i\big) \\ &\implies x_i < -\beta P_2 + K_\beta(P_1) \\ &\implies 0 > be^{\frac{2(\alpha+\beta)P_1 x_i}{N_0}} + d \\ &\implies 0 > be^{\frac{2(\alpha P_1-\beta P_2)x_i}{N_0}} + de^{\frac{-2\beta(P_1+P_2)x_i}{N_0}}, \end{aligned}$$

so

$$\implies \frac{dP^{\text{UB}}_{e,x_i}}{dP_2} < 0.$$

□

Combining Theorems 3.14 and 3.15 we can conclude that in Case III, the optimal

power allocation for P_2 given $P_1 > 0$ is

$$P_2^*(P_1) = \min\left(\sqrt{P_2^{\max}}, \tilde{P}_2(P_1)\right)$$

Now that we have the optimal P_2 allocation for a fixed P_1 we characterize the optimization with respect to P_1 using the following theorem.

Theorem 3.16. *In Case III, $P_e\big(P_1, P_2^*(P_1)\big)$ is decreasing in P_1, for all $P_1 > 0$.*

Proof. Let $P_1 > 0$. We can express the optimal power allocation for P_2 as the following function of P_1:

$$P_2^*(P_1) = \begin{cases} \sqrt{P_2^{\max}} & P_1 < P_1^{thresh} \\ \tilde{P}_2(P_1) & P_1 \geq P_1^{thresh} \end{cases}$$

where

$$P_1^{thresh} \triangleq \frac{N_0}{2(\alpha+\beta)^2\sqrt{P_2^{\max}}} \ln \frac{\bar{a}\bar{d}}{\bar{b}\bar{c}}.$$

We analyze this in two cases. First, assume $P_1 < P_1^{thresh}$. Since $P_2^*(P_1) = \sqrt{P_2^{\max}} < \tilde{P}_2(P_1)$, Proposition 3.12 implies that there is a unique root, x, to (3.6) corresponding to P_1 and P_2. Hence, by the same reasoning as in Theorem 3.8 it is sufficient to show $P_{e,x}^{\text{UB}}$ is decreasing in P_1. Unlike the previous proof, we can see immediately from the expression given in (3.20) that this derivative is negative for all P_1 because $\bar{b} > 0$ in Case III. Next, for $P_1 \geq P_1^{thresh}$ we have

$$P_2^*(P_1) = \frac{N_0}{2(\alpha+\beta)^2 P_1} \ln \frac{\bar{a}\bar{d}}{\bar{b}\bar{c}}.$$

Let x^* be the root to (3.6) for P_1 and $P_2^*(P_1)$. (3.27) implies $x^* = \alpha P_2^*(P_1) - K_\alpha(P_1) = -\beta P_2^*(P_1) + K_\beta(P_1)$. Substituting these relationships into the expression for the error

probability given in (3.8) yields

$$\begin{aligned} P_e\big(P_1, P_2^*(P_1)\big) = \bar{a}Q\bigg(\frac{1}{\sigma}\bigg(\alpha P_1 + P_a - \frac{\alpha-\beta}{2}P_1\bigg)\bigg) \\ + \bar{c}Q\bigg(\frac{1}{\sigma}\bigg(-\beta P_1 + P_a - \frac{\alpha-\beta}{2}P_1\bigg)\bigg) \\ + \bar{b}Q\bigg(\frac{1}{\sigma}\bigg(\alpha P_1 - P_b - \frac{\alpha-\beta}{2}P_1\bigg)\bigg) \\ + \bar{d}Q\bigg(\frac{1}{\sigma}\bigg(-\beta P_1 - P_b - \frac{\alpha-\beta}{2}P_1\bigg)\bigg) \\ + \sum_{(l,m)\in\{0,1\}^2} p_0 p_{lm|0}, \end{aligned}$$

where we define

$$P_a \triangleq \frac{N_0}{2(\alpha+\beta)P_1}\ln\frac{\bar{a}}{-\bar{c}}, \qquad P_b \triangleq \frac{N_0}{2(\alpha+\beta)P_1}\ln\frac{-\bar{d}}{\bar{b}},$$

noting that

$$\frac{dP_a}{dP_1} = -\frac{P_a}{P_1}, \qquad \frac{dP_b}{dP_1} = -\frac{P_b}{P_1}.$$

Then derivative analysis on the error probability yields

$$\begin{aligned} \frac{dP_e}{dP_1} &= -\frac{1}{\sigma\sqrt{2\pi}}\bigg(\bigg(\frac{\alpha+\beta}{2} - \frac{P_a}{P_1}\bigg)\bar{a}e^{\frac{-(\frac{\alpha+\beta}{2}P_1+P_a)^2}{N_0}} + \bigg(-\frac{\alpha+\beta}{2} - \frac{P_a}{P_1}\bigg)\bar{c}e^{\frac{-(-\frac{\alpha+\beta}{2}P_1+P_a)^2}{N_0}} \\ &\quad + \bigg(\frac{\alpha+\beta}{2} + \frac{P_b}{P_1}\bigg)\bar{b}e^{\frac{-(\frac{\alpha+\beta}{2}P_1-P_b)^2}{N_0}} + \bigg(-\frac{\alpha+\beta}{2} + \frac{P_b}{P_1}\bigg)\bar{d}e^{\frac{-(\frac{\alpha+\beta}{2}P_1+P_b)^2}{N_0}}\bigg) \\ &= -\frac{e^{\frac{-(\frac{\alpha+\beta}{2}P_1)^2}{N_0}}}{\sigma\sqrt{2\pi}}\bigg(e^{\frac{-P_a^2}{N_0}}\bigg(\big(\frac{\alpha+\beta}{2} - \frac{P_a}{P_1}\big)\bar{a}e^{\frac{-(\alpha+\beta)P_1P_a}{N_0}} + \big(-\frac{\alpha+\beta}{2} - \frac{P_a}{P_1}\big)\bar{c}e^{\frac{(\alpha+\beta)P_1P_a}{N_0}}\bigg) \\ &\quad + e^{\frac{-P_b^2}{N_0}}\bigg(\big(\frac{\alpha+\beta}{2} + \frac{P_b}{P_1}\big)\bar{b}e^{\frac{(\alpha+\beta)P_1P_b}{N_0}} + \big(-\frac{\alpha+\beta}{2} + \frac{P_b}{P_1}\big)\bar{d}e^{\frac{-(\alpha+\beta)P_1P_b}{N_0}}\bigg)\bigg) \end{aligned}$$

$$
\begin{aligned}
&= -\frac{e^{\frac{-(\frac{\alpha+\beta}{2}P_1)^2}{N_0}}}{\sigma\sqrt{2\pi}}\Bigg(e^{\frac{-P_a^2}{N_0}}\Big(\Big(\frac{\alpha+\beta}{2}-\frac{P_a}{P_1}\Big)\sqrt{-\bar{a}\bar{c}}-\Big(-\frac{\alpha+\beta}{2}-\frac{P_a}{P_1}\Big)\sqrt{-\bar{a}\bar{c}}\Big)\\
&\qquad + e^{\frac{-P_b^2}{N_0}}\Big(\Big(\frac{\alpha+\beta}{2}+\frac{P_b}{P_1}\Big)\sqrt{-\bar{b}\bar{d}}-\Big(-\frac{\alpha+\beta}{2}+\frac{P_b}{P_1}\Big)\sqrt{-\bar{b}\bar{d}}\Big)\Bigg)\\
&= -\frac{\alpha+\beta}{\sigma\sqrt{2\pi}}e^{\frac{-(\frac{\alpha+\beta}{2}P_1)^2}{N_0}}\Big(e^{\frac{-P_a^2}{N_0}}\sqrt{-\bar{a}\bar{c}}+e^{\frac{-P_b^2}{N_0}}\sqrt{-\bar{b}\bar{d}}\Big)\\
&< 0.
\end{aligned}
$$

□

The implication of this result is that the less noisy sensor should use all available power. Further, we note $\tilde{P}_2(P_1)$ is decreasing in P_1. This results in the generally intuitive behaviour to allocate more power to the more reliable sensor, and less to the worse sensor.

Combining the results of Theorems 3.14, 3.15 and 3.16 implies that the optimal power allocation and corresponding error performance can be expressed as follows, where x^* is the root of (3.6), and a_{lm}^* are the corresponding constellation points to P_1^* and P_2^*:

$$
\begin{aligned}
P_1^{*\text{Case III}} &= \sqrt{P_1^{\max}},\\
P_2^{*\text{Case III}} &= \min\big(\sqrt{P_2^{\max}}, \tilde{P}_2(\sqrt{P_1^{\max}})\big),\\
P_e^{*\text{Case III}} &= \sum_{(l,m)\in\{0,1\}^2}(p_1 p_{lm|1} - p_0 p_{lm|0})Q\left(\frac{a_{lm}^* - x^*}{\sigma}\right) + p_0 p_{lm|0}. \qquad (3.34)
\end{aligned}
$$

The corresponding decision region is $\mathcal{D}_0 = (-\infty, x^*]$. If $P_2^* = \tilde{P}_2(\sqrt{P_1^{\max}}) = \tilde{P}_2$ as in (3.3), then the decision boundary has the explicit form $x^* = \alpha\tilde{P}_2 - K_\alpha(\sqrt{P_1^{\max}})$.

3.3.5 High SNR Behaviour

For this analysis it is defined that high SNR means $N_0 \to 0$. This is a reasonable assumption since each sensor's SNR should be growing at similar rates, and one sensor should not have infinitely more power than the other. In Case I, there is nothing to consider, since the error probability is constant. In Cases II and III, the high SNR behaviour can be analyzed by considering a system that knows which point in the constellation $\mathcal{C}$ was sent. Note this is equivalent to knowing X_1 and X_2 perfectly, except if two constellation points are identical, i.e., $a_{lm} = a_{l'm'}$ for some $l \neq l'$ or $m \neq m'$. Also note that the high SNR behaviour of only sending Sensor i is always ϵ_i, $i \in \{1, 2\}$. In the case that both sensor values are used, the following is the MAP detection rule for knowing $X_1 = x_1$ and $X_2 = x_2$:

$$\begin{aligned}\hat{x}(x_1, x_2) &= \arg\max_{i \in \{0,1\}} \Pr(X = i \mid X_1 = x_1, X_2 = x_2) \\ &= \arg\max_{i \in \{0,1\}} p_i p_{x_1 x_2 | i}.\end{aligned}$$

The decision rules can be expressed in terms of the constants defined in (3.6) as

$$\bar{a} > 0 \iff \hat{x}(1,1) = 1, \qquad \bar{b} > 0 \iff \hat{x}(1,0) = 1,$$
$$\bar{c} > 0 \iff \hat{x}(0,1) = 1, \qquad \bar{d} > 0 \iff \hat{x}(0,0) = 1.$$

Case II

Based on the values of $\bar{a}, \bar{b}, \bar{c}, \bar{d}$ in this case, we have the following detection rule:

$$\hat{x}(1,1) = 1, \qquad \hat{x}(1,0) = 0,$$

$$\hat{x}(0,1) = 0, \qquad \hat{x}(0,0) = 0.$$

Since $\hat{x}(1,0) = \hat{x}(0,1)$, it does not matter if these two constellation points are identical (and these are the only two constellation points which can possibly be identical). Finally, the high SNR behaviour is calculated to be

$$\begin{aligned}\lim_{N_0 \to 0} P_e^*(P_1^*, P_2^*) &= p_1(p_{00|1} + p_{01|1} + p_{10|1}) + p_0 p_{11|0} \\ &= \epsilon_1\epsilon_2 + p_1(\epsilon_1 + \epsilon_2 - 2\epsilon_1\epsilon_2). \end{aligned} \tag{3.35}$$

Case III

In this case, there are two interesting cases to consider. First, for the optimal allocation $P_2 = P_2^*$, we have that $P_2^* \to 0$ for high SNR; so the error performance in this case approaches the performance of only sending Sensor 1, which is ϵ_1. In the alternative case that both sensors use all their power, the detection rule is as follows:

$$\begin{aligned}\hat{x}(1,1) &= 1, \qquad & \hat{x}(1,0) &= 1, \\ \hat{x}(0,1) &= 0, \qquad & \hat{x}(0,0) &= 0.\end{aligned}$$

If no constellation points are identical, $\hat{X} = X_1$, which implies

$$\lim_{N_0 \to 0} P_e(\sqrt{P_1^{\max}}, \sqrt{P_2^{\max}}) = \epsilon_1, \qquad P_1^{\max} \neq P_2^{\max}.$$

However, if $P_1^{\max} = P_2^{\max} = P^{\max}$ for some $P^{\max} > 0$, then $a_{01} = a_{10}$. Let P_e^s, $s \in \{0,1\}$ be the error probability if we decide to detect s for 10/01. Since $p_1 \leq p_0$

and

$$\begin{aligned} P_e^0 &= \epsilon_1\epsilon_2 + p_1(\epsilon_1 + \epsilon_2 - 2\epsilon_1\epsilon_2) \\ P_e^1 &= \epsilon_1\epsilon_2 + p_0(\epsilon_1 + \epsilon_2 - 2\epsilon_1\epsilon_2) \\ \implies P_e^0 &\leq P_e^1, \end{aligned}$$

we conclude to decide 0, and the final expression for the high SNR behaviour is

$$\lim_{N_0 \to 0} P_e\big(\sqrt{P^{\max}}, \sqrt{P^{\max}}\big) = \epsilon_1\epsilon_2 + p_1(\epsilon_1 + \epsilon_2 - 2\epsilon_1\epsilon_2).$$

These results are demonstrated in Figure 3.10, where the curve for using both sensors at their max power has a larger end behaviour than the derived optimal constellation design. Note that in Case II, $P_e^0 < \epsilon_1$, but in Case III, $\epsilon_1 < P_e^0 < \epsilon_2$.

3.4 Numerical and Simulation Results

In this section, we illustrate the results of this chapter numerically for specific parameter sets of the problem setup. We show that the theoretical results proven in the previous section are also supported by simulated experiments. In what follows, the SNR is defined as the geometric average of the available power allocations, reported in dB $\big($i.e., SNR (dB) $= 10\log_{10}(\text{SNR})\big)$:

$$\text{SNR}^{\max} \triangleq \frac{\sqrt{P_1^{\max} P_2^{\max}}}{N_0}. \tag{3.36}$$

Even though Sensor 2 does not necessarily use all of its allocated power, defining the SNR this way is sensible because the sensors have independent power constraints. If

the sensors had a joint power constraint, it would be more appropriate to use the true SNR.

3.4.1 Simulated Validation of Main Results

The experimental data is produced by sending 500,000 independent source bits via two simulated sensors and MAC, then using the MAP detection rule given in (3.1), the error probability is calculated. We will show in two ways that the simulations overlap with the theoretical results. First we show that the minimization problem is solved at the correct value of P_2 in Case III. Then we show that the error probability when using the derived optimal constellation design overlaps with the simulation results at any SNR in Case II. We always use the optimal asymmetric constellation designs for these simulations, $\mathcal{C}_i = \{c_{0,i}, c_{1,i}\} = \{-\beta P_i,\ \alpha P_i\}$ for $i \in \{1, 2\}$. To calculate the theoretical error probability, the decision boundaries are calculated by numerically solving for the roots of (3.6). Then, these values are used to calculate the appropriate error expression, (3.8) or (3.30), based on the number of roots.

The error probability as a function of P_2 and SNR are shown in Figures 3.3 and 3.4, respectively. These plots show that the simulated and theoretical error performance overlap very well, while also noting that the simulated minimum power allocation for P_2 coincides with the theoretical results.

3.4.2 Simulated Comparison to Orthogonal Signaling

In this section, we compare the error probability of the MAC signaling system derived in this paper to an alternative signaling method of using independent (orthogonal) channels for the sensors. To set up the orthogonal signaling, it is assumed that the

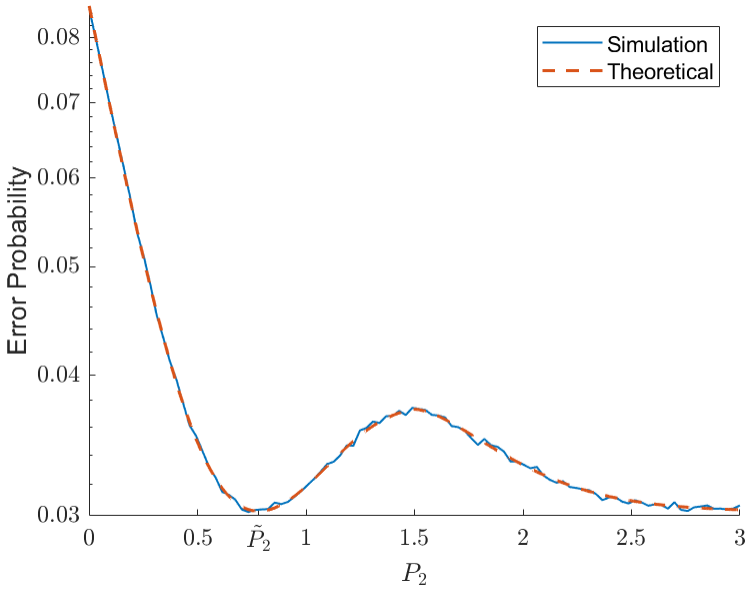


Figure 3.3: Theoretical and simulated error probability in Case III ($p_1 = 0.45, \epsilon_1 = 0.01, \epsilon_2 = 0.05, P_1 = 1, N_0 = 1$).

sensor network would have access to two independent zero-mean Gaussian communication channels with variance $\frac{N_0}{2}$. Note that we define the SNR in the orthogonal case to be the same as in (3.36). Even though there is more total noise when considering both orthogonal channels, this is a realistic comparison. If a system has access to two orthogonal channels with the same noise power, it can choose to only use one of the channels, which is exactly the equivalent MAC we are comparing to. We use two variations of orthogonal constellations as baseline comparisons. First, we use a simple symmetric binary phase-shift keying (BPSK) constellation design (i.e., $\mathcal{C}_i = \{c_{0,i}, c_{1,i}\} = \{-\sqrt{P_i^{\max}}, \sqrt{P_i^{\max}}\}$ for $i \in \{1, 2\}$). We also use the results of [11] which give an optimal orthogonal constellation design to be asymmetric BPSK with $\mathcal{C}_i = \{c_{0,i}, c_{1,i}\} = \{-\beta\sqrt{P_i^{\max}}, \alpha\sqrt{P_i^{\max}}\}$ for $i \in \{1, 2\}$, with α and β as defined in (3.5). To detect the source, the receiver uses the MAP detection rule which is the two dimensional extension of (3.1). Since we have not analyzed the orthogonal

channels case theoretically, we rely on simulation results to draw conclusions. In Figures 3.4 and 3.5, the error probabilities are compared under two parameter sets, to show Cases II and III, respectively. Each data point is generated from 500,000 independent simulated source bits being sent through the channel. Figure 3.5, also includes the error probabilities associated with using the maximum power symmetric constellation design over the MAC.

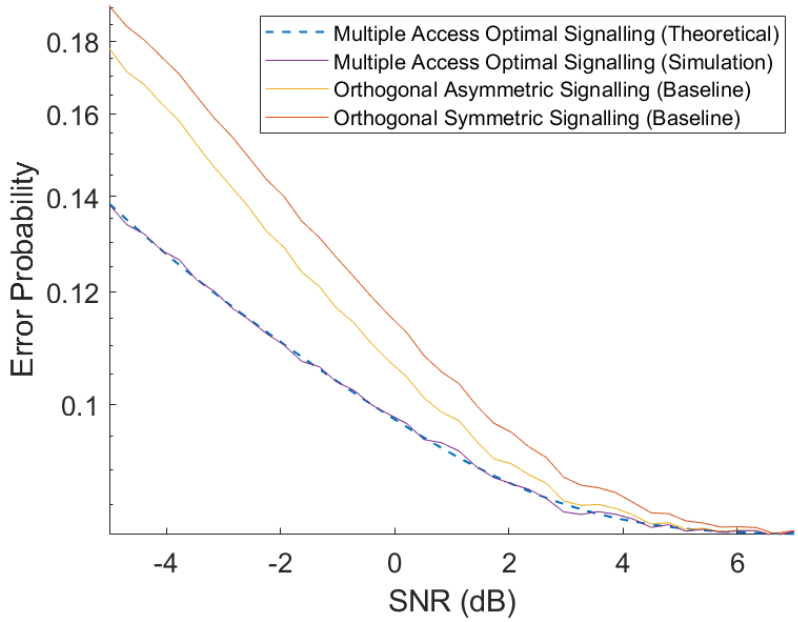


Figure 3.4: Error probability as a function of SNR in Case II ($p_1 = 0.3, \epsilon_1 = 0.1, \epsilon_2 = 0.15, P_1^{\max} = 1, P_2^{\max} = 1$).

From these graphs we can see that in both Cases II and III, the derived optimal MAC constellation has better error performance than orthogonal signaling. However, in Case III (Figure 3.5), orthogonal signaling can perform better than the sub-optimal multiple access symmetric constellation design. These results demonstrate that using a MAC optimally can have increased performance, while using less power and bandwidth. In Figure 3.4, the maximum SNR gain of the derived optimal constellation

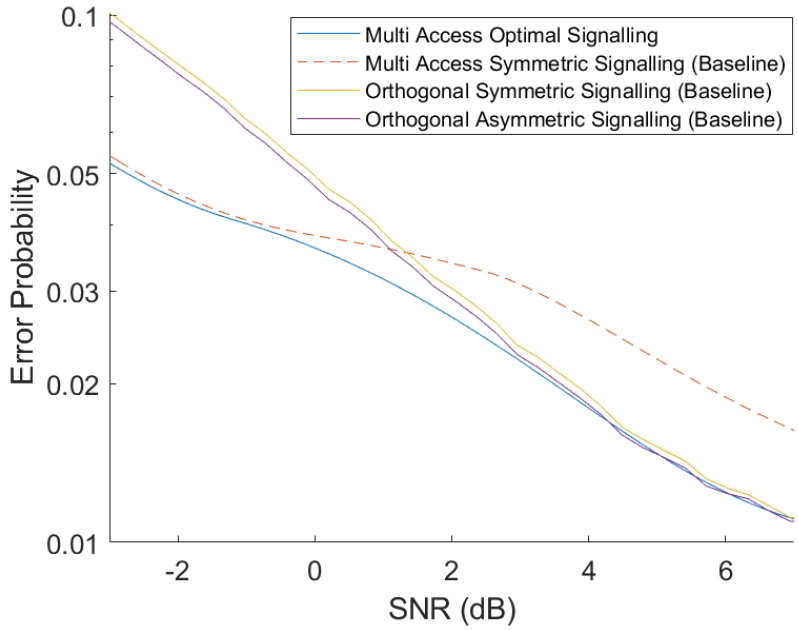


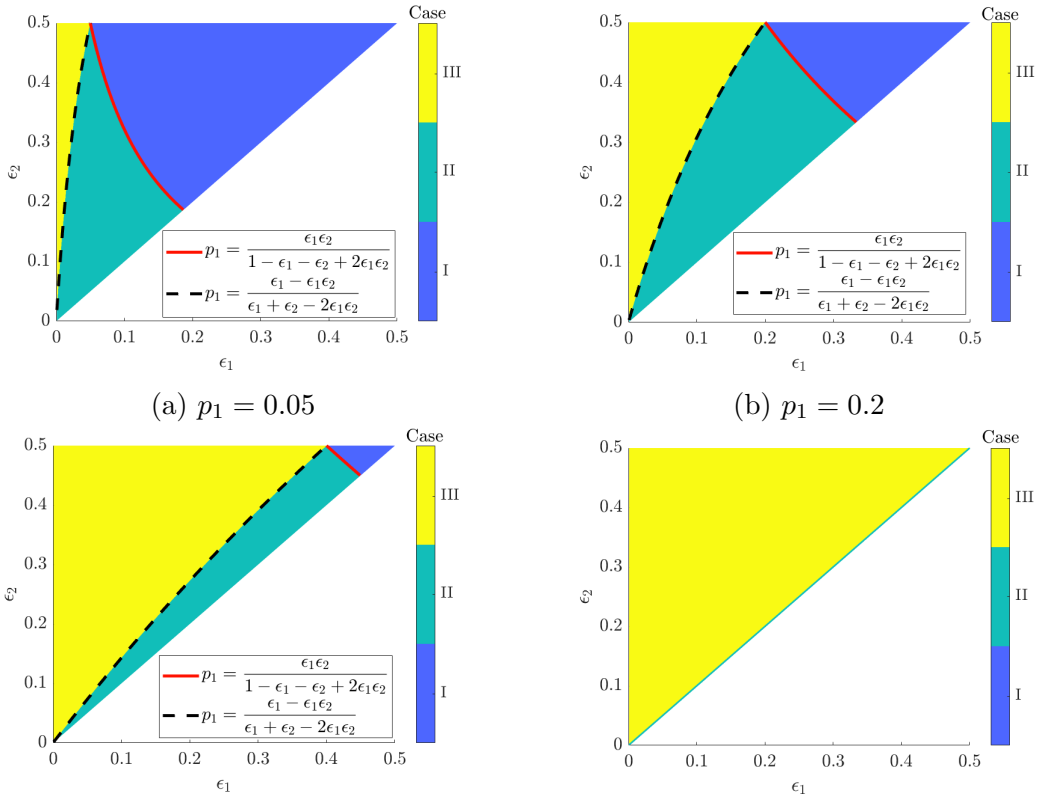
Figure 3.5: Error probability as a function of SNR in Case III ($p_1 = 0.4, \epsilon_1 = 0.01, \epsilon_2 = 0.05, P_1^{\max} = 1, P_2^{\max} = 2$).

compared to the next best option is about 2.4 dB. In Figure 3.5, the maximum SNR gain is approximately 0.97 dB, occurring around 0.036 error probability.

3.4.3 Analysis of Cases Based on Parameters p_1, ϵ_1 and ϵ_2

We analyze the behaviour of Cases I-III as a function of the parameters ϵ_1, ϵ_2 and p_1. By fixing p_1, we illustrate the case type regions as a colour map of ϵ_1 and ϵ_2. Examples of these graphs are shown in Figure 3.6.

We make the following observations from these diagrams. Case I occurs at large ϵ_1 and ϵ_2 values, while Case III is characterized by small ϵ_1 and large ϵ_2. The boundaries between these regions are given exactly by the threshold equations given in Table 3.1. As p_1 increases, the Case I region becomes smaller, while Case III becomes larger. Finally, at $p_1 = 0.5$, Case I disappears entirely, and Case II is equivalent to $\epsilon_1 = \epsilon_2$.



(a) $p_1 = 0.05$ (b) $p_1 = 0.2$

(c) $p_1 = 0.4$ (d) $p_1 = 0.5$

Figure 3.6: Case type regions for different values of p_1.

This can intuitively be explained by noting that for any $p_1 < 0.5$, as $\epsilon_1, \epsilon_2 \to 0.5$, X_1 and X_2 become uniformly distributed and independent from the source X. This effectively removes the source information, making it useless to send over the channel (Case I). However, if $p_1 = 0.5$, X_1 and X_2 are uniformly distributed for any ϵ_1 and ϵ_2, so there is no statistical redundancy in the source (as it is unbiased) that can be lost when observed by the sensors. Hence it is always beneficial to send the signals, which explains why Case I disappears at $p_1 = 0.5$.

3.4.4 Error Performance vs. (P_1, P_2) and (ϵ_1,ϵ_2)

For the following examples, the constellations are parameterized by the optimal asymmetric design, $\mathcal{C}_i = \{c_{0,i}, c_{1,i}\} = \{-\beta P_i,\ \alpha P_i\}$, $i \in \{1, 2\}$. Figures 3.7 and 3.8 show the error probability as a function of P_1 and P_2 in Cases II and III, respectively.

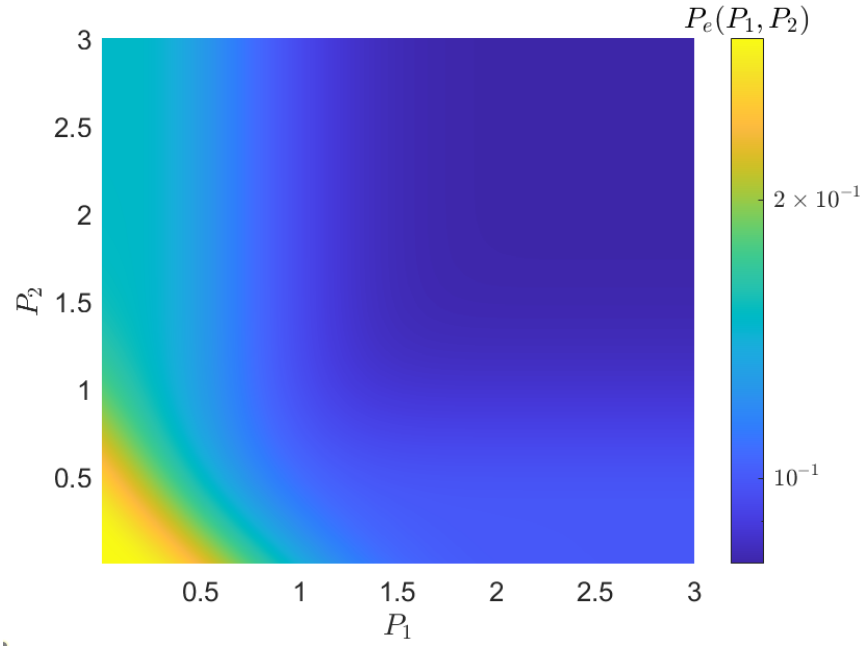


Figure 3.7: Error probability as a function of P_1 and P_2 in Case II ($p_1 = 0.3, \epsilon_1 = 0.1, \epsilon_2 = 0.15, N_0 = 1$).

In Figure 3.7, increasing P_1 and P_2 always decreases the error probability, which reinforces the dervied optimal power allocation to use all available power. Figure 3.8 illustrates the following properties of Case III. First, we can see that for any fixed P_1 (vertical slice of the graph), the minimum occurs at $P_2 = \tilde{P}_2(P_1)$, the red curve. Further, moving upward from $P_2 = 0$ to $P_2 = \tilde{P}_2(P_1)$, we also see that the error probability decreases in P_2. Finally, we can see that moving rightward along the optimal power allocation curve, $\tilde{P}_2(P_1)$, the optimal error probability decreases with P_1,

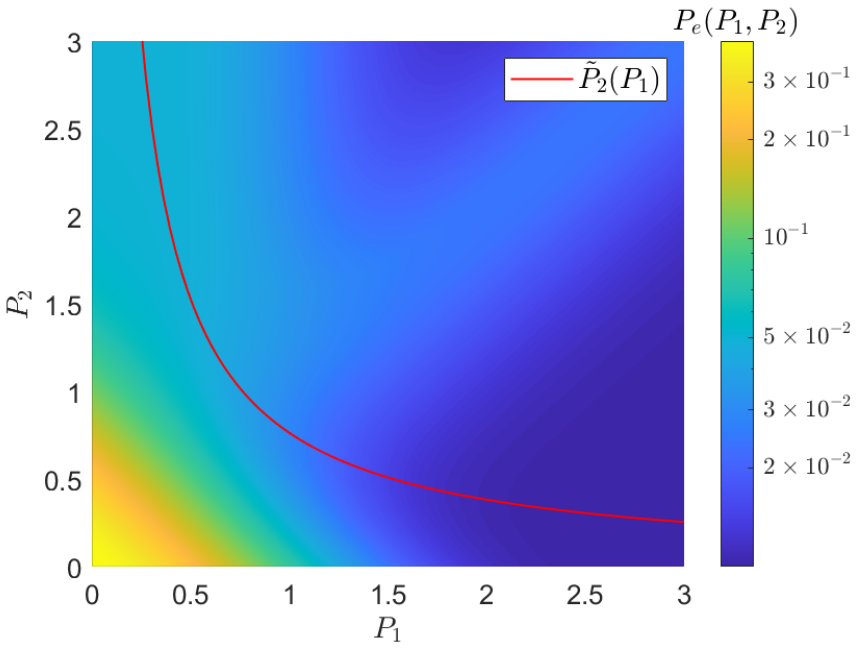


Figure 3.8: Error probability as a function of P_1 and P_2 in Case III ($p_1 = 0.4, \epsilon_1 = 0.01, \epsilon_2 = 0.05, N_0 = 1$).

which reinforces exactly the same optimal power allocation as proven. For example, if $P_1^{\max} = P_2^{\max} = 1$, then reading Figure 3.8 shows the optimal power allocations are $P_1^* = 1$ and $P_2^* = \tilde{P}_2(1) \approx 0.76$.

Next, we show how the optimal error probability changes with respect to the sensor noise parameters, ϵ_1 and ϵ_2. We fix the other parameters, p_1, N_0, $P_1^{\max}$ and $P_2^{\max}$, then for each pair (ϵ_1, ϵ_2), choose the constellation power allocations P_1^* and P_2^*. Figure 3.9 shows the error probability as function of ϵ_1 and ϵ_2.

When analyzing Figure 3.9, note that referencing Figure 3.6b, we can identify the regions of the three cases separated by the same line boundaries. We observe that in the Case I region (upper right corner) the error probability takes a constant value of 0.2, which is also the largest error probability across all regions. In the remaining regions (Cases II and III), it is intuitive that the error probability decreases for smaller

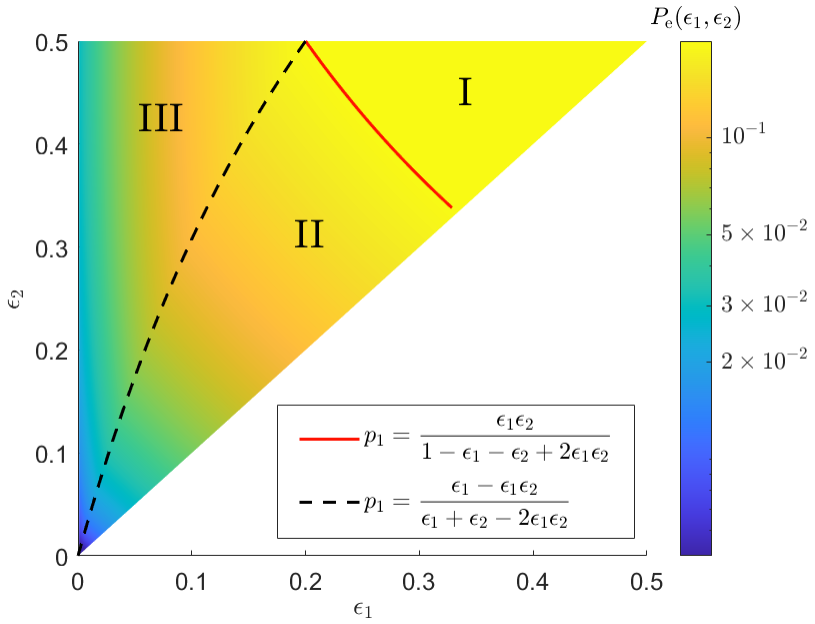


Figure 3.9: Error probability as a function of ϵ_1 and ϵ_2 ($p_1 = 0.2$, $P_1^{\max} = 1$, $P_2^{\max} = 1$, $N_0 = 1$). The region and boundary curves for each case are the same as in Figure 3.6b.

values of ϵ_1 and ϵ_2. A more insightful observation is that the error probability is more sensitive to ϵ_1 than ϵ_2. Especially in the Case III region, we see that varying ϵ_1 has a much larger impact on the error probability than ϵ_2. Thus, having one very reliable sensor and one very poor sensor can perform better than two moderately accurate sensors.

3.4.5 Error Probability vs. Signal to Noise Ratio

To demonstrate this system's SNR response, we vary N_0 to produce various SNR values as defined in (3.36). We also compare the derived optimal constellation design to other common power allocations. For the following example, the constellations are parameterized by the optimal asymmetric design, $\mathcal{C}_i = \{c_{0,i}, c_{1,i}\} = \{-\beta P_i,\ \alpha P_i\}$, $i \in \{1, 2\}$. Figure 3.10 shows the error probability of various constellation designs in

Case III as a function of N_0, expressed in terms of the SNR.

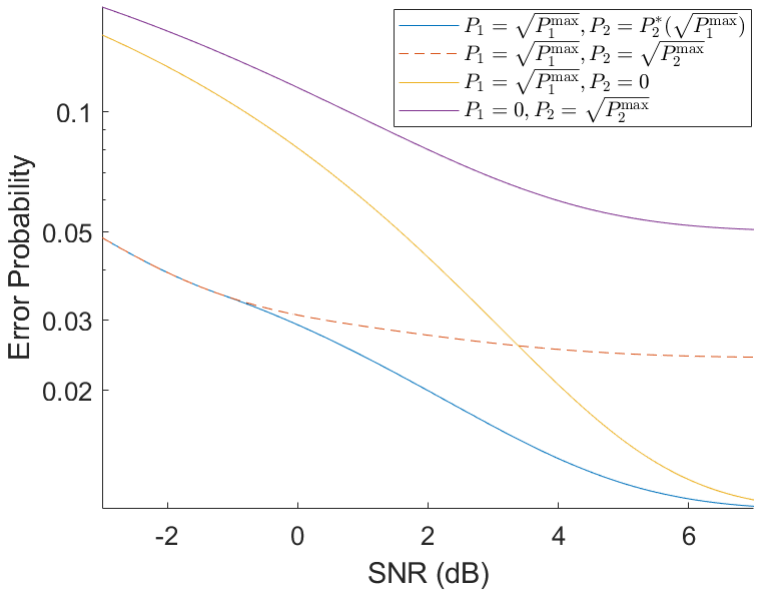


Figure 3.10: Error probability as a function of SNR in Case III ($p_1 = 0.4, \epsilon_1 = 0.01, \epsilon_2 = 0.05, P_1^{\max} = 1, P_2^{\max} = 1$).

We make the following observations from this plot. First, at low SNR, the optimal and both max curves are identical. This is because for large enough values of N_0, $\sqrt{P_2^{\max}} < \tilde{P}_2$ as defined in (3.3), so $P_2^*(\sqrt{P_1^{\max}}) = \sqrt{P_2^{\max}}$. Next, at high SNR, the optimal and P_1 max curves become asymptotically equal. This is because as $N_0 \to 0$, $P_2^*(\sqrt{P_1^{\max}}) \to 0$. At intermediate SNR (around 0-7 dB in this case), the optimal power allocation performs better than any of the alternatives. The largest SNR gain of using the derived optimal constellation is about 2.7 dB, occurring around 0.026 error probability.

Chapter 4

Multi User Gaussian MAC Sensor Networks

In this chapter, we investigate a generalized version of the problem setup in Chapter 3, where the number of sensors is an arbitrary $N \geq 2$. We first note that for $N > 2$, there is no guarantee that analytical solutions to the problem will exist. The focus of this chapter is on describing the existence of an optimization solution to an individual sensor optimization and how to numerically calculate it. Further, this individual minimization can be applied iteratively to form an algorithm for minimizing all N sensors.

4.1 Problem Setup

This chapter uses the most generalized setup described in Section 2.1 with signal space $\mathcal{S} = \mathbb{R}$. A block diagram showing the system model is given in Figure 4.1. We are interested in finding the optimal constellation design for each sensor, $\mathcal{C}_i$, $i \in \{1, ..., N\}$ which minimizes the error probability. Theorem 3.1 trivially generalizes to the N sensor case, hence we can again restrict the optimization search to constellations of the form $\mathcal{C}_i = \{-\beta P_i, \alpha P_i\}$, $i \in \{1, ..., N\}$, with α and β as defined in (3.5). Thus, the

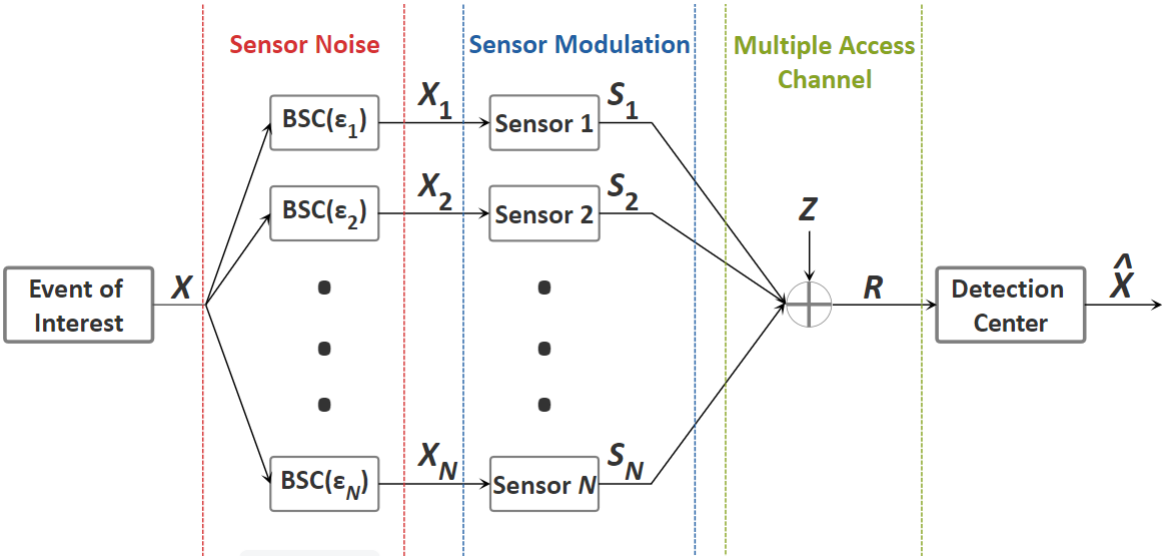


Figure 4.1: Block diagram showing the $N \geq 2$ sensor MAC system.

problem has also become a power allocation optimization for each $P_i \leq \sqrt{P_i^{\max}}$.[1] In conjunction with the sensor readings, $\mathcal{B}$, as defined in Section 2.1, for $i \in \{1, ..., N\}$, $s \in \{0, 1\}$ we define $\mathcal{B}_s^i \triangleq \{\underline{b} \in \mathcal{B} \mid b_i = s\}$, i.e., all possible combinations where Sensor i reads the value s.

We make the following simplifying assumption about the detection rule in the N sensor case. Instead of using MAP detection at the fusion center, we assume that there will be at most one decision boundary between $\mathcal{D}_0$ and $\mathcal{D}_1$, i.e., each these regions must be a single interval on the real line. This assumption ensures that the final solution is feasible to implement, as the number of possible decision boundaries grows with the number of sensors. Additionally, it will be shown numerically that the constellations optimized using this assumption will still perform comparably to using MAP detection, while generally using less power.

[1]Even though in this chapter, each P_i, $i \in \{1, ..., N\}$, corresponds to the square root of the power $E[\|S_i\|^2]$ with upper bound constraint $P_i^{\max}$ as specified in Chapter 2, we will still refer to it as "power" for the sake of simplicity.

With these assumptions, the error probability for any power allocation $P_1, ..., P_N$ will have the following expression:

$$P_e(P_1, ..., P_N) = \sum_{\underline{b} \in \mathcal{B}} k_{\underline{b}} Q\left(\frac{a_{\underline{b}} - x}{\sigma}\right) + p_0 p_{\underline{b}|0}, \tag{4.1}$$

where $x \in \mathbb{R}$ is a function of $P_1, ..., P_N$ such that this function is minimized (not necessarily uniquely), for each $\underline{b} \in \mathcal{B}$, $k_{\underline{b}} \triangleq p_1 p_{\underline{b}|1} - p_0 p_{\underline{b}|0}$, where $p_{\underline{b}|s}$, $s = 0, 1$ are the conditional probabilities and $a_{\underline{b}} \in \mathcal{C}$ represents the constellation points associated to sensor readings $\underline{b}$ and as defined in Section 2.1. Although there does not necessarily exist $x \in \mathbb{R}$ which minimizes (4.1), we give the following proposition to fully characterize this minimization.

Proposition 4.1. *For $P_1, ..., P_N > 0$ there exists $x \in \mathbb{R}$ which minimizes (4.1) if and only if $p_1 > \frac{\prod_{i=1}^N \epsilon_i}{\prod_{i=1}^N \epsilon_i + \prod_{i=1}^N (1-\epsilon_i)}$.*

Proof. Assume $p_1 > \frac{\prod_{i=1}^N \epsilon_i}{\prod_{i=1}^N \epsilon_i + \prod_{i=1}^N (1-\epsilon_i)}$. First, note that this is equivalent to $k_{1...1} > 0$. Also note that $k_{0...0} = p_1 \prod_{i=1}^N \epsilon_i - p_0 \prod_{i=1}^N (1 - \epsilon_i) < 0$. The derivative of (4.1) with respect to x is analyzed to be

$$\frac{d}{dx} P_e(P_1, ..., P_N) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{\underline{b} \in \mathcal{B}} k_{\underline{b}} e^{-\frac{(a_{\underline{b}} - x)^2}{N_0}}.$$

We arrive at the following expressions for the end behaviours of this derivative.

$$\lim_{x \to -\infty} \frac{d}{dx} P_e(P_1, ..., P_N) = \lim_{x \to -\infty} \frac{k_{0...0}}{\sigma\sqrt{2\pi}} e^{-\frac{(a_{0...0} - x)^2}{N_0}},$$
$$\lim_{x \to \infty} \frac{d}{dx} P_e(P_1, ..., P_N) = \lim_{x \to \infty} \frac{k_{1...1}}{\sigma\sqrt{2\pi}} e^{-\frac{(a_{0...0} - x)^2}{N_0}}.$$

By the above analysis of the signs of $k_{0...0}$ and $k_{1...1}$, we conclude $\frac{d}{dx} P_e(P_1, ..., P_N) \uparrow 0$

as $x \to -\infty$ and $\frac{d}{dx}P_e(P_1,...,P_N) \downarrow 0$ as $x \to \infty$. Since P_e is a smooth function of x, this immediately implies there must be some minimizer $x \in \mathbb{R}$. Conversely, if $p_1 \leq \frac{\prod_{i=1}^N \epsilon_i}{\prod_{i=1}^N \epsilon_i + \prod_{i=1}^N (1-\epsilon_i)}$, then $k_{1...1} \leq 0$. Further, for any $\underline{b} = b_1...b_N \in \mathcal{B}$,

$$\begin{aligned} k_{\underline{b}} &= p_1 \prod_{i|b_i=0} \epsilon_i \prod_{i|b_i=1} (1-\epsilon_i) - p_0 \prod_{i|b_i=1} \epsilon_i \prod_{i|b_i=0} (1-\epsilon_i) \\ &\leq p_1 \prod_{i=1}^N (1-\epsilon_i) - p_0 \prod_{i=1}^N \epsilon_i = k_{1...1} \leq 0. \end{aligned}$$

Combining this result with $k_{0...0} < 0$ gives $\frac{d}{dx}P_e(P_1,...,P_N) < 0$ for any $x \in \mathbb{R}$, and thus P_e could not admit a minimizer. Note that this corresponds to the optimal decision rule being always detecting a 0. □

Note that if there does not exist a minimizer of (4.1), this means that the optimal detection rule will be to trivially always detect a 0. This is analogous to Case I in Chapter 3, where the error probability is p_1 regardless of what any sensor sends, so they should not send anything at all.

4.2 Isolated Optimization of a Single Sensor

In this section, we assume every sensor except one has already fixed their power allocation, and analyze the optimization of the remaining sensor. First, we assume $p_1 > \frac{\prod_{i=1}^N \epsilon_i}{\prod_{i=1}^N \epsilon_i + \prod_{i=1}^N (1-\epsilon_i)}$, so that the optimization is not trivial and the error expression will always have the form as given in (4.1). For any fixed $P_1,...,P_N$, and $i \in \{1,...,N\}$ we define the following functions of x:

$$g_{0,i}^{P_1,...,P_N}(x) = \sum_{\underline{b} \in \mathcal{B}_0^i} k_{\underline{b}} Q\Big(\frac{a_{\underline{b}} - x}{\sigma}\Big), \tag{4.2}$$

$$g_{1,i}^{P_1,...,P_N}(x) = \sum_{\underline{b}\in\mathcal{B}_1^i} k_{\underline{b}} Q\Big(\frac{a_{\underline{b}} - x}{\sigma}\Big). \tag{4.3}$$

Note that these functions change based on the power allocations $P_1, ..., P_N$ since the constellation points depend on the powers. These functions decompose the error probability given in (4.1) and have many useful properties which make them feasible to analyze. For example, if x is a corresponding minimizer of (4.1) for $P_1, ..., P_N$, then for each i we have

$$P_e(P_1, ..., P_N) = g_{0,i}^{P_1,...,P_N}(x) + g_{1,i}^{P_1,...,P_N}(x) + \sum_{\underline{b}\in\mathcal{B}} p_0 p_{\underline{b}|0}, \tag{4.4}$$

which follows directly from their definitions. The derivatives of these functions with respect to x are

$$\frac{d}{dx} g_{0,i}^{P_1,...,P_N}(x) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{\underline{b}\in\mathcal{B}_0^i} k_{\underline{b}} e^{-\frac{(a_{\underline{b}}-x)^2}{N_0}}, \tag{4.5}$$

and

$$\frac{d}{dx} g_{1,i}^{P_1,...,P_N}(x) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{\underline{b}\in\mathcal{B}_1^i} k_{\underline{b}} e^{-\frac{(a_{\underline{b}}-x)^2}{N_0}}. \tag{4.6}$$

The most used property about these functions is given in the following proposition.

Proposition 4.2. *For $P_1, ..., P_N > 0$, for any $i = 1, ..., N$, $g_{0,i}$ and $g_{1,i}$ have the following scaled translation invariance properties:*

$$g_{0,i}^{P_1,...,P_N}\big(x - \beta(P_i - P_i')\big) = g_{0,i}^{P_1,...,P_i',...,P_N}(x), \quad \forall x \in \mathbb{R},\ P_i' > 0,$$
$$g_{1,i}^{P_1,...,P_N}\big(x + \alpha(P_i - P_i')\big) = g_{1,i}^{P_1,...,P_i',...,P_N}(x), \quad \forall x \in \mathbb{R},\ P_i' > 0.$$

Proof. Let $P_1, ..., P_N > 0$, $i \in \{1, ..., N\}$. These properties follow directly from the

fact that for every $\underline{b} \in \mathcal{B}_0^i$, the contribution of P_i to $a_{\underline{b}}$ is $-\beta P_i$, and for every $\underline{b} \in \mathcal{B}_1^i$, the contribution of P_i to $a_{\underline{b}}$ is αP_i. Combining this with the definitions given in (4.2) and (4.3) immediately implies the desired result. □

These properties are a fundamental part of the analysis when proving the proceeding results. For example, it is used to show the following optimization result, which is analogous to Case II from Chapter 3.

Theorem 4.3. *For $P_1, ..., P_{i-1}, P_{i+1}, ..., P_N > 0$ fixed, if*

$$\epsilon_i \leq \frac{p_0 \prod_{s \neq i} \epsilon_s}{p_0 \prod_{s \neq i} \epsilon_s + p_1 \prod_{s \neq i}(1-\epsilon_s)} < 1 - \epsilon_i, \tag{4.7}$$

then the error function given in (4.1) *is decreasing in P_i.*

Proof. First note that the second inequality in (4.7) is equivalent to the inequality condition in Proposition 4.1, to guarantee that the detection rule is not-trivial. Next, we will show (4.7) implies that $g_{0,i}^{P_1,...,P_N}$ is decreasing in x for any P_i. For any $b_1...0...b_N \in \mathcal{B}_0^i$

$$\begin{aligned} k_{b_1...0...b_N} &= p_1\epsilon_i\Big(\prod_{s|b_s=0}\epsilon_s \prod_{s|b_s=1}(1-\epsilon_s)\Big) - p_0(1-\epsilon_i)\Big(\prod_{s|b_s=1}\epsilon_s \prod_{s|b_s=0}(1-\epsilon_s)\Big) \\ &\leq p_1\epsilon_i \prod_{s\neq i}(1-\epsilon_s) - p_0(1-\epsilon_i)\prod_{s\neq i}\epsilon_s \\ &= \epsilon_i\Big(p_0\prod_{s\neq i}\epsilon_s + p_1\prod_{s\neq i}(1-\epsilon_s)\Big) - p_0\prod_{s\neq i}\epsilon_s \leq 0, \end{aligned}$$

where the last inequality is equivalent to a simple rearrangement of (4.7). Combining this result with $k_{0...0} < 0$ and the expression given in (4.5) shows $\frac{d}{dx} g_{0,i}^{P_1,...,P_N}(x) < 0$ for any P_i and x. Now let $0 \leq P_i < P_i'$, with x and x' being the respective minimizers

for the error expression as given in (4.1). Using Proposition 4.2, we have

$$g_{1,i}^{P_1,...,P_i',...P_N}\big(x+\alpha(P_i'-P_i)\big) = g_{1,i}^{P_1,...,P_i,...,P_N}(x).$$

Since $g_{0,i}^{P_1,...,P_N}$ is decreasing in x, we also have

$$\begin{aligned} g_{0,i}^{P_1,...,P_i',...P_N}\big(x+\alpha(P_i'-P_i)\big) &= g_{0,i}^{P_1,...,P_i,...,P_N}\big(x+(\alpha+\beta)(P_i'-P_i)\big) \\ &< g_{0,i}^{P_1,...,P_i,...,P_N}(x). \end{aligned}$$

Since x' must minimize the error function as expressed in (4.4), we conclude that

$$\begin{aligned} & g_{0,i}^{P_1,...,P_i',...,P_N}(x') + g_{1,i}^{P_1,...,P_i',...,P_N}(x') \\ \leq\ & g_{0,i}^{P_1,...,P_i',...,P_N}\big(x+\alpha(P_i'-P_i)\big) + g_{1,i}^{P_1,...,P_i',...,P_N}\big(x+\alpha(P_i'-P_i)\big) \\ <\ & g_{0,i}^{P_1,...,P_i,...,P_N}(x) + g_{1,i}^{P_1,...,P_i,...,P_N}(x) \\ \Longrightarrow\ & P_e(P_1,...,P_i',...,P_N) < P_e(P_1,...,P_i,...,P_N). \end{aligned}$$

□

Now we characterize the optimization in the remaining case using Proposition 4.4, followed by the main results of Theorems 4.5 and 4.6. Note that this case is analogous to Case III from Chapter 3.

Proposition 4.4. *For $P_1,...,P_{i-1},P_{i+1},...,P_N > 0$, if*

$$\epsilon_i > \frac{p_0\prod_{s\neq i}\epsilon_s}{p_0\prod_{s\neq i}\epsilon_s + p_1\prod_{s\neq i}(1-\epsilon_s)}, \tag{4.8}$$

then there exist values x_0^* and x_1^* with $x_0^* \geq x_1^*$ which minimize $g_{0,i}^{P_1,...,P_N}$ and $g_{1,i}^{P_1,...,P_N}$

for $P_i = 0$, respectively. Further, for $P_i > 0$, $g_{0,i}^{P_1,...,P_N}$ is minimized at $x = x_0^* - \beta P_i$ and $g_{1,i}^{P_1,...,P_N}$ is minimized at $x = x_1^* + \alpha P_i$.

Proof. Let $P_1, ..., P_{i-1}, P_{i+1}, ..., P_N > 0$, and $P_i = 0$. We analyze the end behaviours of the derivative expressions of (4.5) and (4.6). There are four limits of interest, corresponding to the left and right end behaviours of the two derivatives. We have

$$\begin{aligned}
\lim_{x\to-\infty} \frac{d}{dx} g_{0,i}^{P_1,...,P_N}(x) &= \lim_{x\to-\infty} \frac{1}{\sigma\sqrt{2\pi}} \sum_{\underline{b}\in\mathcal{B}_0^i} k_{\underline{b}} e^{-\frac{(a_{\underline{b}}-x)^2}{N_0}} \\
&= \lim_{x\to-\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(a_{0...0}-x)^2}{N_0}} \sum_{\underline{b}\in\mathcal{B}_0^i} k_{\underline{b}} e^{\frac{(a_{0...0}-x)^2-(a_{\underline{b}}-x)^2}{N_0}} \\
&= \lim_{x\to-\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(a_{0...0}-x)^2}{N_0}} \sum_{\underline{b}\in\mathcal{B}_0^i} k_{\underline{b}} e^{\frac{(a_{0...0}-a_{\underline{b}})(a_{0...0}+a_{\underline{b}}-2x)}{N_0}},
\end{aligned}$$

noting that $a_{0...0} < a_{\underline{b}}$ for any $\underline{b} \in \mathcal{B}_0^i$, $\underline{b} \neq 0...0$, so this limit simplifies to

$$\lim_{x\to-\infty} \frac{d}{dx} g_{0,i}^{P_1,...,P_N}(x) = \lim_{x\to-\infty} \frac{k_{0...0}}{\sigma\sqrt{2\pi}} e^{-\frac{(a_{0...0}-x)^2}{N_0}}.$$

This implies that if $k_{0...0} < 0$, then $\frac{d}{dx} g_{0,i}^{P_1,...,P_N} \uparrow 0$ as $x \to -\infty$. Using the exact same reasoning, combined with the facts that $a_{11...0...11} > a_{\underline{b}}$ for all $\underline{b} \in B_0^i$, $\underline{b} \neq 11...0...11$, $a_{00...1...00} < a_{\underline{b}}$ for all $\underline{b} \in B_1^i$, $\underline{b} \neq 00...1...00$ and $a_{1...1} > a_{\underline{b}}$ for all $\underline{b} \in B_1^i$, $\underline{b} \neq 1...1$, we arrive at the following conclusions about the end behaviours:

$$\begin{aligned}
\frac{d}{dx} g_{0,i}^{P_1,...,P_N} &\uparrow 0 \quad \text{as} \quad x \to -\infty \quad \text{if} \quad k_{0...0} < 0, \\
\frac{d}{dx} g_{0,i}^{P_1,...,P_N} &\downarrow 0 \quad \text{as} \quad x \to \infty \quad \text{if} \quad k_{11...0...11} > 0, \\
\frac{d}{dx} g_{1,i}^{P_1,...,P_N} &\uparrow 0 \quad \text{as} \quad x \to -\infty \quad \text{if} \quad k_{00...1...00} < 0, \\
\frac{d}{dx} g_{1,i}^{P_1,...,P_N} &\downarrow 0 \quad \text{as} \quad x \to \infty \quad \text{if} \quad k_{1...1} > 0.
\end{aligned}$$

Since these are both smooth functions, if the left end behaviour has a negative derivative, and the right behaviour has a positive derivative, then they each must admit a respective minimizer. These inequality conditions can be expressed as follows in terms of the parameters p_0, p_1 and ϵ_s, $s \in 1, ..., N$:

$$\begin{aligned} k_{0...0} < 0 &\iff p_1 \prod_{s=1}^{N} \epsilon_s - p_0 \prod_{s=1}^{N} (1-\epsilon_s) < 0 \\ &\iff (1-\epsilon_i)\Bigg(- p_1 \prod_{s\neq i} \epsilon_s - p_0 \prod_{s\neq i} (1-\epsilon_s)\Bigg) < -p_1 \prod_{s\neq i} \epsilon_s \\ &\iff 1-\epsilon_i > \frac{p_1 \prod_{s\neq i} \epsilon_s}{p_1 \prod_{s\neq i} \epsilon_s + p_0 \prod_{s\neq i} (1-\epsilon_s)} \triangleq K_1 \end{aligned}$$

By the same reasoning, we obtain

$$\begin{aligned} k_{00...1...00} < 0 &\iff \epsilon_i > \frac{p_1 \prod_{s\neq i} \epsilon_s}{p_1 \prod_{s\neq i} \epsilon_s + p_0 \prod_{s\neq i} (1-\epsilon_s)} \\ k_{11...0...11} > 0 &\iff \epsilon_i \overset{(4.8)}{>} \frac{p_0 \prod_{s\neq i} \epsilon_s}{p_0 \prod_{s\neq i} \epsilon_s + p_1 \prod_{s\neq i} (1-\epsilon_s)} \triangleq K_0 \\ k_{1...1} > 0 &\iff 1-\epsilon_i > \frac{p_0 \prod_{s\neq i} \epsilon_s}{p_0 \prod_{s\neq i} \epsilon_s + p_1 \prod_{s\neq i} (1-\epsilon_s)} \end{aligned}$$

By the assumption $0 < \epsilon_i < 0.5$, we have $1 - \epsilon_i > \epsilon_i$. Additionally, the assumption $p_0 \geq p_1$ gives $K_0 \geq K_1$. Thus, if (4.8) holds, then $k_{0...0} < 0$, $k_{11...0...11} > 0$, $k_{00...1...00} < 0$ and $k_{1...1} > 0$, implying that $g_{0,i}^{P_1,...,P_N}$ and $g_{1,i}^{P_1,...,P_N}$ admit minimizers x_0^* and x_1^*, respectively. To show $x_0^* \geq x_1^*$, we first show that for $P_i = 0$, $\frac{d}{dx} g_{1,i}^{P_1,...,P_N}(x) > \frac{d}{dx} g_{0,i}^{P_1,...,P_N}(x)$ for all $x \in \mathbb{R}$. First we note that any $\underline{b} \in \mathcal{B}_0^i$ has the form $\underline{b} = b_1...0...b_N$ and bijectively maps into $\mathcal{B}_1^i$ as $b_1...1...b_N$. Further for $P_i = 0$, it is clear that $a_{b_1...0...b_N} = a_{b_1...1...b_N}$.

Thus, it is sufficient to show $k_{b_1\ldots1\ldots b_N} > k_{b_1\ldots0\ldots b_N}$ for any $b_1, ..., b_N \in \{0, 1\}$. We have

$$\begin{aligned} k_{b_1\ldots1\ldots b_N} - k_{b_1\ldots0\ldots b_N} &= p_1 p_{b_1\ldots1\ldots b_N|1} - p_0 p_{b_1\ldots1\ldots b_N|0} - (p_1 p_{b_1\ldots0\ldots b_N|1} - p_0 p_{b_1\ldots0\ldots b_N|0}) \\ &= (1-\epsilon_i) p_1 p_{b_1\ldots b_{i-1} b_{i+1}\ldots b_N|1} - \epsilon_i p_0 p_{b_1\ldots b_{i-1} b_{i+1}\ldots b_N|0} \\ &\quad - \epsilon_i p_1 p_{b_1\ldots b_{i-1} b_{i+1}\ldots b_N|1} + (1-\epsilon_i) p_0 p_{b_1\ldots b_{i-1} b_{i+1}\ldots b_N|0} \\ &= (1-2\epsilon_i)(p_1 p_{b_1\ldots b_{i-1} b_{i+1}\ldots b_N|1} + p_0 p_{b_1\ldots b_{i-1} b_{i+1}\ldots b_N|0}) > 0. \end{aligned}$$

Now let x_1^* be a minimizer for $g_{1,i}^{P_1,\ldots,P_N}$. For any $x < x_1^*$ we must have $g_{0,i}^{P_1,\ldots,P_N}(x) > g_{0,i}^{P_1,\ldots,P_N}(x_1^*)$, otherwise

$$\begin{aligned} & g_{0,i}^{P_1,\ldots,P_N}(x) \le g_{0,i}^{P_1,\ldots,P_N}(x_1^*) \quad \text{and} \quad g_{1,i}^{P_1,\ldots,P_N}(x) \ge g_{1,i}^{P_1,\ldots,P_N}(x_1^*) \\ \implies & g_{1,i}^{P_1,\ldots,P_N}(x_1^*) - g_{0,i}^{P_1,\ldots,P_N}(x_1^*) \le g_{1,i}^{P_1,\ldots,P_N}(x) - g_{0,i}^{P_1,\ldots,P_N}(x), \end{aligned}$$

but this contradicts $\frac{d}{dx} g_{1,i}^{P_1,\ldots,P_N}(x) > \frac{d}{dx} g_{0,i}^{P_1,\ldots,P_N}(x)$. Thus any minimizer of $g_{0,i}^{P_1,\ldots,P_N}$ must satisfy $x_0^* \ge x_1^*$. Now, let $P_i > 0$. The desired result follows immediately from applying Proposition 4.2 (with $P_i' = 0$) since

$$g_{0,i}^{P_1,\ldots,P_N}(x - \beta P_i) = g_{0,i}^{P_1,\ldots,P_{i-1},0,P_{i+1},\ldots,P_N}(x) \quad \forall x \in \mathbb{R},$$

and

$$g_{1,i}^{P_1,\ldots,P_N}(x + \alpha P_i) = g_{1,i}^{P_1,\ldots,P_{i-1},0,P_{i+1},\ldots,P_N}(x), \quad \forall x \in \mathbb{R}.$$

□

Although the condition in (4.8) may be restrictive for small N, we observe the

following behaviour as N becomes large.

$$\lim_{N\to\infty} \frac{p_0 \prod_{s\neq i} \epsilon_s}{p_0 \prod_{s\neq i} \epsilon_s + p_1 \prod_{s\neq i}(1-\epsilon_s)} = \lim_{N\to\infty} \frac{1}{1+\frac{p_1}{p_0}\prod_{s\neq i}\frac{1-\epsilon_s}{\epsilon_s}} \leq \lim_{N\to\infty} \frac{1}{1+\frac{p_1}{p_0}(\frac{1-\epsilon_N}{\epsilon_N})^{N-1}} = 0,$$

where the last equality holds, so long as the value of ϵ_N does not approach 0.5. Thus, as the number of sensors grows, the condition of (4.8) becomes less restrictive. Another limitation when trying to practically apply the results of Proposition 4.4 is that the optimal values x_0^* and x_1^* generally do not have an analytical form. However, since the functions are smooth, we know that the minimizers must be at critical points of the respective functions. We denote the critical points of each function as

$$\mathcal{X}_s = \left\{ x \in \mathbb{R} \;\middle|\; \frac{d}{dx} g_{s,i}^{P_1,\ldots,P_N}(x) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{\underline{b}\in\mathcal{B}_s^i} k_{\underline{b}} e^{-\frac{(a_{\underline{b}}-x)^2}{N_0}} = 0 \right\}, \quad s = 0, 1. \tag{4.9}$$

Applying [23, Corollary 3.2], we get $|\mathcal{X}_s| < 2^{N-1}$, $s = 0, 1$. We also know that if (4.8) holds, then there must be at least one point in each of these sets. Thus, to find x_s^*, we can numerically calculate the values $x \in \mathcal{X}_s$, and exhaustively find the minimizer of the respective function, i.e. $x_s^* = \arg\min_{x\in\mathcal{X}_s} g_{s,i}^{P_1,\ldots,P_N}$. Although Proposition 4.4 only gives optimality with respect to x, it can easily be used to determine optimality with respect to P_i, as demonstrated in the following theorem.

Theorem 4.5. *For $P_1, \ldots, P_{i-1}, P_{i+1}, \ldots, P_N > 0$ fixed, if (4.8) holds, the error function given in (4.1) as a function of P_i is minimized at $P_i^* = \frac{x_0^* - x_1^*}{\alpha+\beta}$, with optimal decision boundary $x^* = \frac{\alpha x_0^* + \beta x_1^*}{\alpha+\beta}$, where x_0^* and x_1^* are minimizers of $g_{0,i}^{P_1,\ldots,0\ldots,P_N}$ and $g_{1,i}^{P_1,\ldots,0,\ldots,P_N}$, respectively. Further, if x_0^* and x_1^* are unique minimizers, then P_i^* is also*

a unique minimizer.

Proof. Assume $P_1, ..., P_{i-1}, P_{i+1}, ..., P_N > 0$, and (4.8) holds. The pair (P_i^*, x^*) as given in the statement of this theorem is the intersection of the two lines of optimality for $g_{0,i}^{P_1,...,P_N}$ and $g_{1,i}^{P_1,...P_N}$ as given in Proposition 4.4. I.e., x^* minimizes both functions:

$$g_{0,i}^{P_1,...,P_i^*,...,P_N}(x^*) \leq g_{0,i}^{P_1,...,P_i^*,...,P_N}(x), \quad g_{1,i}^{P_1,...,P_i^*,...,P_N}(x^*) \leq g_{1,i}^{P_1,...,P_i^*,...,P_N}(x), \quad \forall x \in \mathbb{R}.$$

Further, Proposition 4.2 implies that the minimum values of the functions $g_{0,i}^{P_1,...,P_N}$ and $g_{1,i}^{P_1,...P_N}$ remain the same for any P_i. Thus for any $P_i \geq 0$ and associated decision boundary, $x \in \mathbb{R}$, minimizing (4.1), we obtain

$$g_{0,i}^{P_1,...,P_i^*,...,P_N}(x^*) \leq g_{0,i}^{P_1,...,P_i,...,P_N}(x) \quad \text{and} \quad g_{1,i}^{P_1,...,P_i^*,...,P_N}(x^*) \leq g_{1,i}^{P_1,...,P_i,...,P_N}(x).$$

Noting that if x_0^* and x_1^* are unique minimizers and $P_i \neq P_i^*$, at least one of these must be strict inequalities. Since the expression for the error probability is decomposed as a sum of these functions as given in (4.4), we conclude

$$P_e(P_1, ..., P_i^*, ..., P_N) \leq P_e(P_1, ..., P_i, ..., P_N),$$

where the inequality is strict if x_0^* and x_1^* are unique minimizers. □

Although this result gives a minimizing power allocation, we must account for the problem's constraints, and give a characterization of optimality when there is not enough power to use the results of Theorem 4.5, i.e., if $P_i^* > \sqrt{P_i^{\max}}$. Unfortunately, there is not a general characterization that can be found for this situation, since the functions $g_{0,i}^{P_1,...,P_N}$ and $g_{1,i}^{P_1,...P_N}$ can be very complex and have numerous critical points.

However, using a simplifying assumption, we can deduce the following result which mirrors the fallback power allocation of Case III in Chapter 3.

Theorem 4.6. *Let $P_1, ..., P_{i-1}, P_{i+1}, ..., P_N > 0$ fixed and (4.8) holds. If $g_{0,i}^{P_1,...,0,...,P_N}$ and $g_{1,i}^{P_1,...,0,...,P_N}$ have unique critical points x_0^* and x_1^*, respectively, then the error function given in (4.1) is decreasing as function of P_i until reaching its (unique) minimum value attained at P_i^*, as given in Theorem 4.5.*

Proof. Let $P_1, ..., P_{i-1}, P_{i+1}, ..., P_N > 0$. Assume (4.8) holds, and $g_{0,i}^{P_1,...,0,...,P_N}$ and $g_{1,i}^{P_1,...,0,...,P_N}$ have unique critical points (thus also unique minimizers) x_0^* and x_1^*, respectively. The analysis of the end behaviours of the derivatives (4.5) and (4.6) in the proof of Proposition 4.4 combined with Proposition 4.2, gives that for any $P_i \geq 0$

$$\begin{aligned} x \lesseqgtr x_0^* - \beta P_i &\implies \frac{d}{dx} g_{0,i}^{P_1,...,P_N}(x) \lesseqgtr 0, \\ x \lesseqgtr x_1^* + \alpha P_i &\implies \frac{d}{dx} g_{1,i}^{P_1,...,P_N}(x) \lesseqgtr 0. \end{aligned}$$

Note that from the definition of P_i^*, for any $P_i < P_i^*$, $x_1^* + \alpha P_i < x_0^* - \beta P_i$. Further, since $P_e(P_1, ..., P_N) = g_{0,i}^{P_1,...,P_N}(x) + g_{1,i}^{P_1,...,P_N}(x) + \sum_{\underline{b} \in \mathcal{B}} p_0 p_{\underline{b}|0}$, for any $P_i < P_i^*$, the corresponding minimizer of (4.1), x, must satisfy $x \in (x_1^* + \alpha P_i, x_0^* - \beta P_i)$, otherwise the derivative would be non zero. Now, let $P_i < P_i' < P_i^*$, with x and x' the minimizers of (4.1) for P_i and P_i', respectively. We must break the following into two case. First we assume $x + (\alpha + \beta)(P_i' - P_i) \leq x_0^* - \beta P_i$. We apply Proposition 4.2 to obtain

$$g_{1,i}^{P_1,...,P_i',...P_N}\big(x + \alpha(P_i' - P_i)\big) = g_{1,i}^{P_1,...,P_i,...,P_N}(x),$$

and

$$g_{0,i}^{P_1,...,P_i',...P_N}\big(x+\alpha(P_i'-P_i)\big) = g_{0,i}^{P_1,...,P_i,...,P_N}\big(x+(\alpha+\beta)(P_i'-P_i)\big) < g_{0,i}^{P_1,...,P_i,...,P_N}(x),$$

where the last inequality holds since $g_{0,i}^{P_1,...,P_N}$ is decreasing over $(-\infty, x_0^* - \beta P_i)$ combined with the above assumption. Next, since x' must minimize (4.1) for P_i', which has the form as given in (4.4), we conclude

$$\begin{aligned} & g_{0,i}^{P_1,...,P_i',...,P_N}(x') + g_{1,i}^{P_1,...,P_i',...,P_N}(x') \\ \leq\ & g_{0,i}^{P_1,...,P_i',...,P_N}\big(x+\alpha(P_i'-P_i)\big) + g_{1,i}^{P_1,...,P_i',...,P_N}\big(x+\alpha(P_i'-P_i)\big) \\ <\ & g_{0,i}^{P_1,...,P_i,...,P_N}(x) + g_{1,i}^{P_1,...,P_i,...,P_N}(x) \\ \implies & P_e(P_1,...,P_i',...,P_N) < P_e(P_1,...,P_i,...,P_N). \end{aligned}$$

Alternatively, if we instead assume $x + (\alpha+\beta)(P_i' - P_i) > x_0^* - \beta P_i$, which notably, is equivalent to $x_0^* - \beta P_1' < x + \alpha(P_1' - P_1)$, we have

$$g_{1,i}^{P_1,...,P_i',...P_N}(x_0^* - \beta P_i') < g_{1,i}^{P_1,...,P_i',...,P_N}(x+\alpha(P_1'-P_1)) = g_{1,i}^{P_1,...,P_i,...,P_N}(x),$$

and

$$g_{0,i}^{P_1,...,P_i',...P_N}(x_0^* - \beta P_i') = g_{0,i}^{P_1,...,P_i,...,P_N}(x_0^* - \beta P_i) < g_{0,i}^{P_1,...,P_i,...,P_N}(x),$$

where the equalities are direct applications of Proposition 4.2 and the inequalities follow since $g_{1,i}^{P_1,\ldots,P_i',\ldots P_N}$ is increasing on $(x_1^* + \alpha P_1', \infty)$ combined with the relation $x_1^* + \alpha P_1' < x_0^* - \beta P_i'$, and $g_{0,i}^{P_1,\ldots,P_N}$ is decreasing over $(-\infty, x_0^* - \beta P_i)$. Using similar reasoning the the previous case, we conclude

$$\begin{aligned}
&g_{0,i}^{P_1,\ldots,P_i',\ldots,P_N}(x') + g_{1,i}^{P_1,\ldots,P_i',\ldots,P_N}(x') \\
&\leq g_{0,i}^{P_1,\ldots,P_i',\ldots,P_N}(x_0^* - \beta P_i') + g_{1,i}^{P_1,\ldots,P_i',\ldots,P_N}(x_0^* - \beta P_i') \\
&< g_{0,i}^{P_1,\ldots,P_i,\ldots,P_N}(x) + g_{1,i}^{P_1,\ldots,P_i,\ldots,P_N}(x) \\
\implies & P_e(P_1,\ldots,P_i',\ldots,P_N) < P_e(P_1,\ldots,P_i,\ldots,P_N).
\end{aligned}$$

□

Although this does not fully characterize all situations where there is not enough power to use the optimal allocation as given in Theorem 4.5, the assumption of both functions of interest having a single critical point is realized in many parameter combinations of the problem setup. As a result, a simple and practical way to handle not having enough power is to use all available power. There is no clear way to analyze the fallback allocation when there is not enough power and these functions have multiple critical points. The best approach is an exhaustive search of all critical point pairs, $(x_0, x_1) \in \mathcal{X}_0 \times \mathcal{X}_1$, where these sets are as defined in (4.9). For each pair, it must be verified whether the corresponding power allocation ($P_i = \frac{x_0 - x_1}{\alpha+\beta}$, with decision boundary $x = \frac{\alpha x_0 + \beta x_1}{\alpha+\beta}$) is valid (less than $\sqrt{P_i^{\max}}$). Then the error probability would be calculated for each valid pair, and the minimum over these values and $\sqrt{P_i^{\max}}$ would be chosen.

4.3 Extension of $N = 2$

Although the results in this section were derived under the simplifying assumption of a single decision boundary as opposed to MAP detection, we herein demonstrate how they can be applied to the $N = 2$ sensor case, and produce the same results as in Chapter 3 where MAP detection was used. First we note that the condition $p_1 \leq \frac{\prod_{i=1}^{N} \epsilon_i}{\prod_{i=1}^{N} \epsilon_i + \prod_{i=1}^{N} (1-\epsilon_i)}$ simplifies to $p_1 \leq \frac{\epsilon_1 \epsilon_2}{1-\epsilon_1-\epsilon_2+2\epsilon_1\epsilon_2}$, which is exactly the condition for Case I. Next, we note that the condition $\epsilon_1 \leq \frac{p_0 \prod_{s\neq 1} \epsilon_s}{p_0 \prod_{s\neq 1} \epsilon_s + p_1 \prod_{s\neq 1}(1-\epsilon_s)}$ simplifies to $\epsilon_1 \leq \frac{p_0\epsilon_2}{p_0\epsilon_2+p_1(1-\epsilon_2)}$, but $\frac{p_0\epsilon_2}{p_0\epsilon_2+p_1(1-\epsilon_2)} \geq \epsilon_2$, so this condition is always true (since $\epsilon_1 \leq \epsilon_2$), concluding that Sensor 1 should always use all of its power. Finally, the condition $\epsilon_2 \leq \frac{p_0 \prod_{s\neq 2} \epsilon_s}{p_0 \prod_{s\neq 2} \epsilon_s + p_1 \prod_{s\neq 2}(1-\epsilon_s)}$ simplifies to $\epsilon_2 \leq \frac{p_0\epsilon_1}{p_0\epsilon_1+p_1(1-\epsilon_1)}$ which can be rewritten as $p_1 \leq \frac{\epsilon_1-\epsilon_1\epsilon_2}{\epsilon_1+\epsilon_2-2\epsilon_1\epsilon_2}$, which is exactly the boundary between Cases II and III. In Case III, we note that $g_{0,2}^{P_1,P_2} = h$ and $g_{1,2}^{P_1,P_2} = g$. Thus the respective minimizers at $P_2 = 0$, x_0^* and x_1^* are $K_\beta(P_1)$ and $K_\alpha(P_1)$, respectively. Applying Theorem 4.5 gives the optimal power allocation for P_2 is $\tilde{P}_2(P_1)$. Finally, since the minimizers, x_0^* and x_1^* are unique, Theorem 4.6 shows that the error probability decreases until $\tilde{P}_2$, giving the same final optimization as in Chapter 3.

4.4 Algorithmic Optimization of All Sensors

The joint optimization problem of optimizing each sensor's power is infeasible to analyze and deduce results about global optimally. Instead, we apply the results of Section 4.2 to form an iterative algorithm to optimize individual sensor powers and converge to at least a locally optimal solution. First, we check if $p_1 \leq \frac{\prod_{i=1}^{N} \epsilon_i}{\prod_{i=1}^{N} \epsilon_i + \prod_{i=1}^{N} (1-\epsilon_i)}$. Proposition 4.1 implies that the optimization would be trivial in this case, as the detection rule is always 0, so each sensor should not send anything. Similarly to

Case I of the two sensor optimization, the optimal error probability would be p_1. Otherwise, if $p_1 > \frac{\prod_{i=1}^{N}\epsilon_i}{\prod_{i=1}^{N}\epsilon_i+\prod_{i=1}^{N}(1-\epsilon_i)}$, for each sensor, we must determine whether we can apply the optimization described by Theorems 4.3 or 4.5. This is equivalent to determining whether (4.7) or (4.8) holds for each $i = 1, ..., N$. To characterize which sensors satisfy which condition, we first show that

$$\frac{p_0\prod_{s\neq i'}\epsilon_s}{p_0\prod_{s\neq i'}\epsilon_s + p_1\prod_{s\neq i'}(1-\epsilon_s)} \leq \frac{p_0\prod_{s\neq i}\epsilon_s}{p_0\prod_{s\neq i}\epsilon_s + p_1\prod_{s\neq i}(1-\epsilon_s)}, \quad \text{for } i' > i.$$

Let $i' > i$. Since $0 < \epsilon_i \leq \epsilon_{i'} < 0.5$, we have

$$\begin{aligned}
&\frac{1-\epsilon_i}{\epsilon_i} \geq \frac{1-\epsilon_{i'}}{\epsilon_{i'}} \\
\implies &\frac{p_1}{p_0}\prod_{s\neq i'}\frac{1-\epsilon_s}{\epsilon_s} \geq \frac{p_1}{p_0}\prod_{s\neq i}\frac{1-\epsilon_s}{\epsilon_s} \\
\implies &\frac{1}{1+\frac{p_1}{p_0}\prod_{s\neq i'}\frac{1-\epsilon_s}{\epsilon_s}} \leq \frac{1}{1+\frac{p_1}{p_0}\prod_{s\neq i}\frac{1-\epsilon_s}{\epsilon_s}} \\
\implies &\frac{p_0\prod_{s\neq i'}\epsilon_s}{p_0\prod_{s\neq i'}\epsilon_s + p_1\prod_{s\neq i'}(1-\epsilon_s)} \leq \frac{p_0\prod_{s\neq i}\epsilon_s}{p_0\prod_{s\neq i}\epsilon_s + p_1\prod_{s\neq i}(1-\epsilon_s)}.
\end{aligned}$$

Combining this with $\epsilon_1 \leq \epsilon_2 \leq ... \leq \epsilon_N$ shows that once (4.8) holds for some Sensor n, each Sensor $i > n$ will also satisfy (4.8). The first sensors (if any) which satisfy (4.7) must use all available power. For the remaining sensors (Sensors n to N), we apply an initial condition, then beginning at Sensor n, each power allocation is individually optimized using Theorem 4.5. This process is repeated until either the error probability difference reaches a convergence threshold, or a maximum number of iterations have been completed. These steps are explained in detail in Algorithm 1. The parameters which can be changed when running this algorithm are the initial power allocations, P_0^N, maximum number of iterations, $m^{\max}$, objective function,

Algorithm 1 Numerical Optimization of All Sensors

if $p_1 \leq \frac{\prod_{i=1}^{N} \epsilon_i}{\prod_{i=1}^{N} \epsilon_i + \prod_{i=1}^{N} (1-\epsilon_i)}$ **then**
  $P_i \leftarrow 0, \forall i = 1, ..., N$
  -return
end if
$P^N \leftarrow P_0^N$
for $i \leftarrow 1$ to N **do**
  if $\epsilon_i > \frac{p_0 \prod_{s \neq i} \epsilon_s}{p_0 \prod_{s \neq i} \epsilon_s + p_1 \prod_{s \neq i} (1-\epsilon_s)}$ **then**
    break
  end if
  $P_i \leftarrow \sqrt{P_i^{\max}}$
end for
$i^{\min} \leftarrow i$
$P_e^{\text{prev}} \leftarrow 1$
for $m \leftarrow 1$ to $m^{\max}$ **do**
  for $i \leftarrow i^{\min}$ to N **do**
    $x_0^* \leftarrow \arg\min g_{0,i}^{P_1,...,P_N}$
    $x_1^* \leftarrow \arg\min g_{1,i}^{P_1,...,P_N}$
    $P_i \leftarrow \frac{x_0^* - x_1^*}{\alpha + \beta}$
    if $P_i > \sqrt{P_i^{\max}}$ **then**
      $P_i \leftarrow \sqrt{P_i^{\max}}$
    end if
  end for
  if $\text{obj}\big(P_e^{\text{prev}} - P_e(P_1, ..., P_N)\big) < c^{\text{thresh}}$ **then**
    break
  end if
  $P_e^{\text{prev}} \leftarrow P_e(P_1, ..., P_N)$
end for

obj$(\cdot)$ and the convergence threshold, c^{thresh}. For the numerical results demonstrated in the proceeding sections, we used the parameters $m^{\max} = 500$, $\text{obj}(x) = e^x - 1$ and $c^{\text{thresh}} = 10^{-9}$. Different initial power allocations, $P_0^N \triangleq (P_{0,1}, P_{0,2}, \ldots, P_{0,N})$, were investigated to determine an optimal starting point, the results of which can be seen in the following sections.

Since the corresponding error probability at each step in this algorithm cannot be

increasing (since each point is the solution of a respective minimization problem), and the error is bounded below by 0, this sequence of error probabilities must converge as $m \to \infty$. Thus, this algorithm will always approach at least a local minimum of the overall optimization problem. However, we note that the convergence of the error probability does not mean that the corresponding optimal power allocations necessarily converge to a finite value. For example, it is possible that the power allocation for some sensors (especially those with smaller ϵ_i values) will have constantly increasing values of P_i as the algorithm iterates until reaching its corresponding maximum power. Additionally, the convergence of this algorithm changes based on the starting power allocations for the sensors that have non-trivial optimizations. One way to deal with this variation is to try multiple starting points and use the one with the best convergence error probability.

4.4.1 Algorithm Visualization

We now give a visualization of how the algorithm performs the optimization, and how the power allocations evolve through the iterations. There are different ways that the algorithm can converge based the initial power allocations, as will be shown in the following section. We will be focusing on starting with low initial power, and show how the optimization proceeds in this case. Specifically, we set $N = 5$ and use an initial value of $P_0^N = (1, 0.1, 0.1, 0.1, 0.1)$. In the following figures, an "iteration" refers to a loop through individually optimizing each sensors power once, while a "step" refers to a single individual sensor's power optimization. Figure 4.2 shows how the power allocations for each sensor changes across each iteration, while Figure 4.3 shows how the error probability improves at each step.

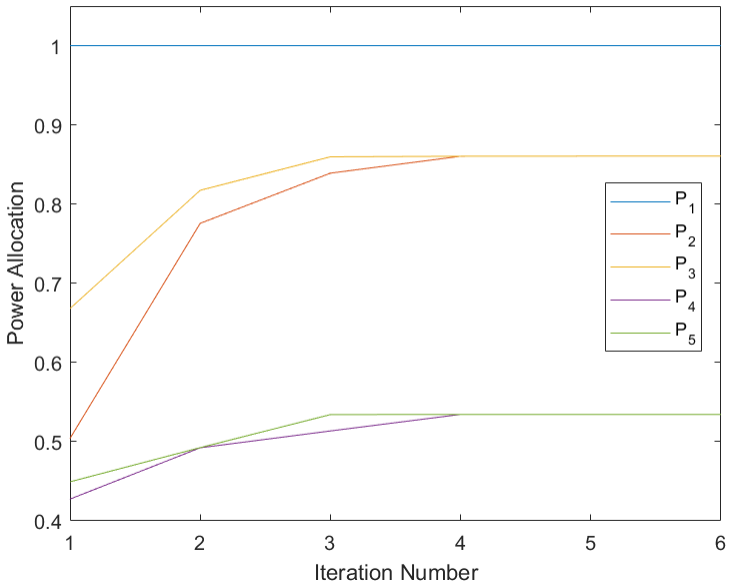


Figure 4.2: Power allocation as the algorithm is iterated ($N = 5$, $p_1 = 0.45$ $\epsilon^N = (0.1, 0.2, 0.2, 0.3, 0.3)$, $N_0 = 1$, $P_i^{\max} = 1$ $\forall i$, $P_0^N = (1, 0.1, 0.1, 0.1, 0.1)$).

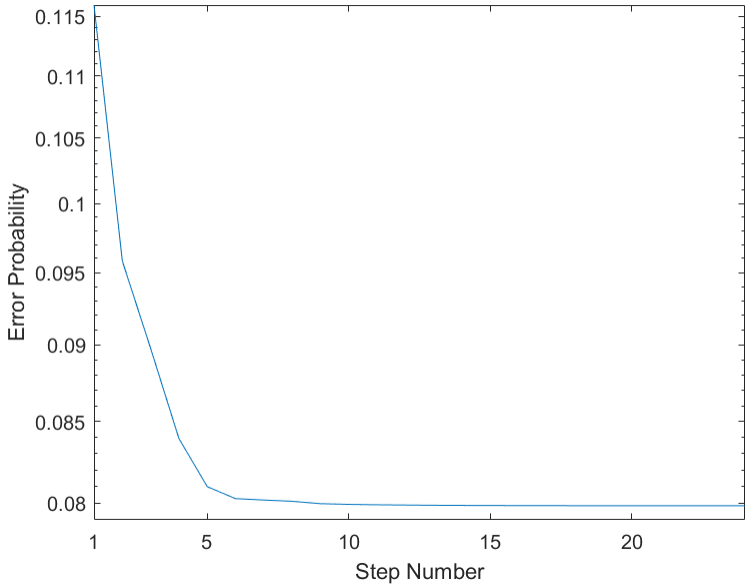


Figure 4.3: Error probability as the algorithm is iterated ($N = 5$, $p_1 = 0.45$ $\epsilon^N = (0.1, 0.2, 0.2, 0.3, 0.3)$, $N_0 = 1$, $P_i^{\max} = 1$ $\forall i$, $P_0^N = (1, 0.1, 0.1, 0.1, 0.1)$).

We observe the following general trends in these figures. First, we note that the algorithm converges within the first few iterations, and after the fourth iteration, the error fluctuations are negligible. Another interesting result about the power allocations is that the sensors with the same noise parameters converge to the same power allocation. This is intuitive since the symmetry of the situation should mean that one sensor is not preferred over the other. Finally, we see that the power allocations for each sensor are increasing as they converge to their optimal values. This is not the case for every starting position, but in general this is seen when using the small starting power allocations. This further explains why this is a good starting point for the optimization, since the interim results do not tend to overshoot where the algorithm will converge to.

4.5 High SNR Analysis

For this section we define high SNR as $N_0 \to 0$. First, we analyze the case where MAP detection is used, and we assume that the power allocations to each sensor are such that no constellation points are identical. Note that in this section, "constellation points" will always refer to the combined constellation, $\mathcal{C}$, from superimposing each sensor's binary constellation. With these assumptions, we can see that the situation can be viewed as the fusion center having full knowledge of each of the sensors' readings. In other words, we treat the problem as if there were no noise in the channel over which the sensors are sending their data. For any $\underline{b} \in \mathcal{B}$, the MAP detected bit will be

$$\hat{x}(\underline{b}) = \underset{i \in \{0,1\}}{\arg\max} \Pr(X = i \mid X^N = \underline{b})$$

$$
\begin{aligned}
&= \arg\max_{i\in\{0,1\}} \Pr(X = i, X^N = \underline{b}) \\
&= \arg\max_{i\in\{0,1\}} \Pr(X = i)\Pr(X^N = \underline{b} \mid X = i) \\
&= \arg\max_{i\in\{0,1\}} p_i p_{\underline{b}|i}, \qquad (4.10)
\end{aligned}
$$

where X is the binary source. Therefore, the high SNR error probability is calculated using the law of total probability to be

$$
\begin{aligned}
\lim_{N_0\to 0} P_e &= \sum_{\underline{b}\in\mathcal{B}} \Pr\big(X \neq \hat{x}(\underline{b}), X^N = \underline{b}\big) \\
&= \sum_{\underline{b}\in\mathcal{B}} \Pr\big(X \neq \hat{x}(\underline{b})\big)\Pr\big(X^N = \underline{b} \mid X \neq \hat{x}(\underline{b})\big) \\
&= \sum_{\underline{b}\in\mathcal{B}} \min_{i\in\{0,1\}} p_i p_{\underline{b}|i}, \qquad (4.11)
\end{aligned}
$$

where the last equality comes from the MAP detection rule given in (4.10) and the fact that the source is binary, so there is only one other option for detection. Note that this expression can have different explicit forms (in terms of the parameters ϵ_i) depending on how each sensor reading combination is detected as given in (4.10). Although some rules about this can be established, such as the all zero reading, $\underline{b} = 0...0$ with always be detected as a 0, there are many different cases that would need to be accounted for to write (4.11) as an explicit function of the parameters.

Next, we are interested in the high SNR behaviour of the simplified detection used in the algorithmic optimization. We first observe that since MAP detection is optimal, we already have a lower bound on this value, given in (4.11). The true expression for this error probability is much more complicated, and depends heavily on the relative positions of the constellation points for the sensor readings. Since there is only one

decision boundary, the high SNR behaviour is to split the constellation points into left and right groups to minimize a similar expression as in (4.11). To formalize this, let $\underline{b}_1, ..., \underline{b}_{2^N}$ be an ordering of $\mathcal{B}$ such that $a_{\underline{b}_1} < ... < a_{\underline{b}_{2^N}}$. Note that again for simplicity we are assuming that no constellation points are identical. We know that in all cases $b_1 = 0...0$, and as mentioned earlier, we always know the MAP detection rule will detect this as a 0. If the first M constellation points are detected as a 0 for some integer $M > 0$, then the high SNR error probability under this detection rule parameterized by M, $\lim_{N_0 \to 0} P_{e|M}$, is given by the following expression:

$$\begin{aligned}
\lim_{N_0 \to 0} P_{e|M} &= \sum_{i=1}^{M} \Pr(X = 1, X^N = b_i) + \sum_{i=M+1}^{2^N} \Pr(X = 0, X^N = b_i) \\
&= p_1 \sum_{i=1}^{M} \Pr(X^N = b_i \mid X = 1) + p_0 \sum_{i=M+1}^{2^N} \Pr(X^N = b_i \mid X = 0) \\
&= p_1 \sum_{i=1}^{M} p_{\underline{b}_i|1} + p_0 \sum_{i=M+1}^{2^N} p_{\underline{b}_i|0}.
\end{aligned}$$

Note that for a given M, this detection scheme could be practically implemented under simplified detection by placing the decision boundary anywhere in the interval $(a_{\underline{b}_M}, a_{\underline{b}_{M+1}})$. Thus, the best high SNR error probability achievable by the simplified detection scheme would be the solution to the following finite element minimization problem:

$$\lim_{N_0 \to 0} P_e = \min_{M \in \{1,...,2^N\}} \left(p_1 \sum_{i=1}^{M} p_{\underline{b}_i|1} + p_0 \sum_{i=M+1}^{2^N} p_{\underline{b}_i|0} \right). \tag{4.12}$$

Since this expression depends on the order of the constellation points, the high SNR error probability can vary based on how the algorithm converges to its optimal value. It is conceivable that the expressions given in (4.11) and (4.12) can be equal

for some power allocations, if the constellation points which are detected as 0's and 1's in the MAP detection rule can be arranged in a way that can be separated by a single boundary. However, it is also not obvious whether there always exists a power allocation which allows these expressions to be equal. To illustrate this, take a two sensor example, and suppose that the sensor readings "00" and "11" are MAP detected as a 0, and "01" and "10" are detected as 1. In this case, it is impossible for simplified detection to yield the same as optimal MAP detection, since a_{00} is always the left most constellation point, and a_{11} is always the right most constellation point. Note that this illustration is only to demonstrate the types of issues that can occur, and this exact setup could never happen with the parameter restrictions of the problem. This example illustrates the types of issues could arise with a larger number of sensors, but the exact specification and characterization of these situations becomes extremely complex.

To provide a better understanding of the behaviour of these two expressions, we compare how (4.11) and (4.12) perform as a function of the source distribution, p_1. To evaluate (4.12), we first perform the algorithmic optimization at a small noise power ($N_0 = 0.01$), then use the resulting optimized constellation to calculate this expression. Although this is not guaranteed to be the best constellation to use when evaluating (4.12), it is infeasible to try all admissible constellation patters, which grows like $2^N!$. Additionally, it is practical to see how the algorithm performs when used in a high SNR environment. Figure 4.4 shows the comparison of these two high SNR behaviours, where the right sub-figure is merely a restriction of the left sub-figure to less biased distribution to better display the discrepancies.

We see that for biased source distributions, these expressions become equal, while

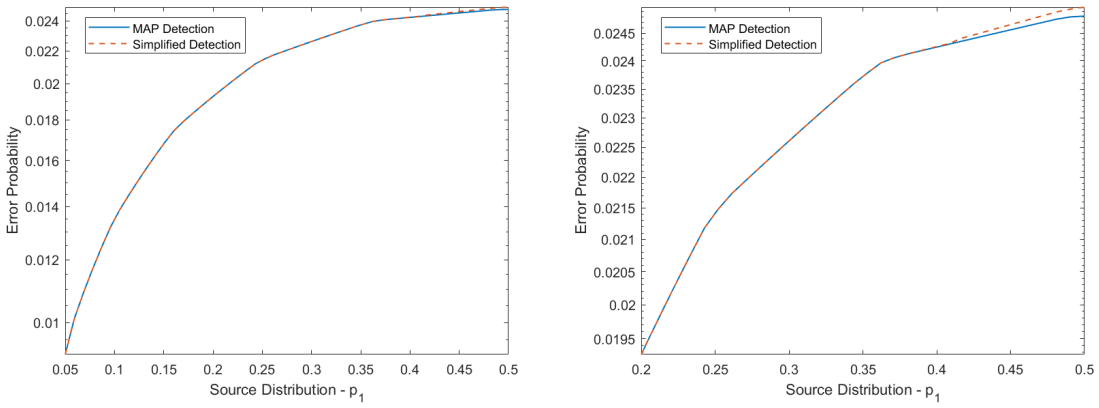


Figure 4.4: High SNR behaviour ($N_0 = 0.01$) of MAP and simplified detection schemes vs. p_1 for $N = 10$ and $\epsilon^N = (0.1, 0.2, 0.2, 0.2, 0.25, 0.25, 0.25, 0.3, 0.3, 0.3)$.

as the source becomes uniform, the simplified detection performs slightly worse than MAP detection. We note that these trends only arise at larger N values; when there are less sensors both expressions are identical at every source distribution. As N increases, the number of constellation points grows exponentially as 2^N. When there are more constellation points, it is less likely that they can be arranged in an ordering for simplified detection to achieve the same error probability as MAP detection.

4.6 Numerical and Simulated Comparisons

In this section, we compare the performance of the optimization algorithm (Algorithm 1) numerically to alternate signaling schemes. Additionally, since the algorithm does not specify its starting point, we compare how the algorithm performs from different starting points, and make general observations about what a good starting point should be. Finally, since we have been working under an assumption of simplifying the detection rules to a single decision boundary, we compare the performance of our algorithmic optimization to detection schemes which use the optimal MAP

detection rule.

In the following figures, we use the following signaling/detection schemes as a comparison to the algorithmic optimization. In the following, each sensor always uses the optimal asymmetric constellation designs, $\mathcal{C}_i = \{c_{0,i}, c_{1,i}\} = \{-\beta P_i,\ \alpha P_i\}$, for all $i = 1, ..., N$. First, we compare to using a Gaussian MAC, but each sensor using all of its available power. We consider both the single decision boundary detection described in this chapter, and the optimal MAP detection rule. Next, a naive pairwise implementation of the results from Chapter 3 are used. For a single Gaussian MAC, we look at each pair (ϵ_1, ϵ_i), $i = 2, ..., N$, and assign to Sensor i the optimal power result from Table 3.2. Finally, we compare to each sensor sending all of its power through orthogonal Gaussian channels and using MAP detection at the fusion center. These schemes are summarized in Table 4.1, with labels corresponding to how they are referred to in the figures of this chapter.

Table 4.1: Signaling Scheme Specifications

Scheme Name	Power Allocation, P^N	Other Specifications
Orthogonal Signaling	$(\sqrt{P_1^{\max}}, ..., \sqrt{P_N^{\max}})$	Orthogonal Channels
MAC Full Power	$(\sqrt{P_1^{\max}}, ..., \sqrt{P_N^{\max}})$	Simplified Detection
MAC Full Power MAP	$(\sqrt{P_1^{\max}}, ..., \sqrt{P_N^{\max}})$	MAP Detection
Pairwise MAC	$(P_1^*, P_2^*(P_1^*), ..., P_N^*(P_1^*))$	P_i^* from Table 3.2
MAC Algorithm	$(P_1^*, ..., P_N^*)$ (Alg. 1)	$P_0^N = (\sqrt{P_1^{\max}}, 0.1, ..., 0.1)$
MAC Algorithm Full Start	$(P_1^*, ..., P_N^*)$ (Alg. 1)	$P_0^N = (\sqrt{P_1^{\max}}, ..., \sqrt{P_N^{\max}})$

For orthogonal signaling, the detection rule is too complex to analyze numerically, so simulations are used. For each data point, the error probability is estimated by sending 1,000,000 source bits through simulated sensors and channels, then the detected bit is compared to the source, and the total number of errors is counted to

arrive at an error probability. To compare these designs, we fix the maximum powers $P_1^{\max}, ..., P_N^{\max}$, vary N_0 and report the SNR as the geometric average:

$$SNR = \frac{\sqrt[N]{\prod_{i=1,...,N} P_i^{\max}}}{N_0}. \tag{4.13}$$

Note that since each sensor has its own power allocation, using the maximum for SNR is a fair comparison, because even if a sensor does not use all of its power, it still has the same amount available to it. Additionally, for the orthogonal channels setup, this definition of SNR is still fair since even though there is more total noise power across all N channels, if the system chose to use only one of the channels, it would be exactly the same as the MAC's that are being compared to.

4.6.1 General Trends

We use the following example to illustrate the general trends in the comparisons between the various signaling/detection methods. Figure 4.5 compares the error probabilities of these schemes with respect to SNR.

We notice the following trends from Figure 4.5. First at low SNR, all of the MAC implementations perform the same, while the orthogonal channel approach performs significantly worse. The equal performance across all MAC implementations can be attributed to the fact that each MAC method at low SNR decides to use all power, i.e., the limit is every method using the same constellation. The orthogonal channels perform significantly worse at low SNR because the channels are too noisy to take advantage of the orthogonality, so it is more advantageous for the signals to constructively interfere with each other in the MAC channel, effectively increasing

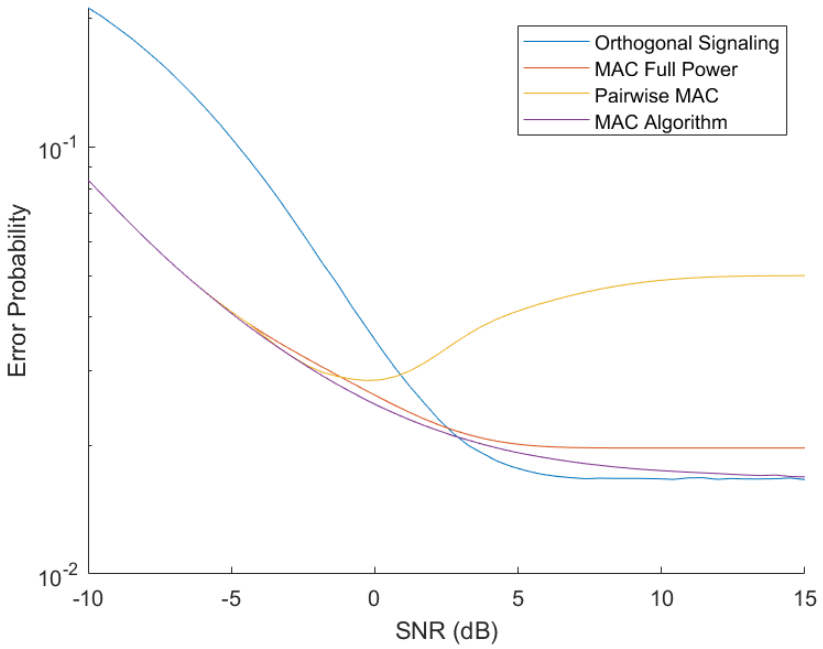


Figure 4.5: Error probability of different signaling schemes as a function of SNR for $N = 5$ $(p_1 = 0.35, (\epsilon_1, ..., \epsilon_5) = (0.05, 0.1, 0.15, 0.2, 0.2), P_i^{\max} = 1, i = 1, ..., 5)$.

power. As the SNR increases, we see that the pairwise approximation becomes worse and never recovers. This is because as SNR increases, the pairwise optimization always decreases the power to the noisier sensor, so all that remains is Sensor 1, which cannot perform as well as multiple sensors working together. Also, there comes a point where the orthogonal channels approach (around 2.5 dB) becomes the best performing solution, but not by very much. This can mainly be attributed to the fact that the orthogonal channel approach is using MAP detection, which allows it to take advantage of more information. Finally, at high SNR, many of the methods converge to the same error probability. Even though the algorithm is not necessarily always the best (only sometimes outperformed by orthogonal MAP detection), it always uses less power than the alternatives, and significantly less bandwidth than the orthogonal

approach.

4.6.2 Comparison of Starting Powers

In this example, we demonstrate how the performance of the optimization algorithm changes for different initial power allocations, and give an explanation of what starting powers are observed to give better overall performance of the converged results of the algorithm. In Figure 4.6, the performance various designs are shown in a case where different starting points for the algorithm give significantly different results.

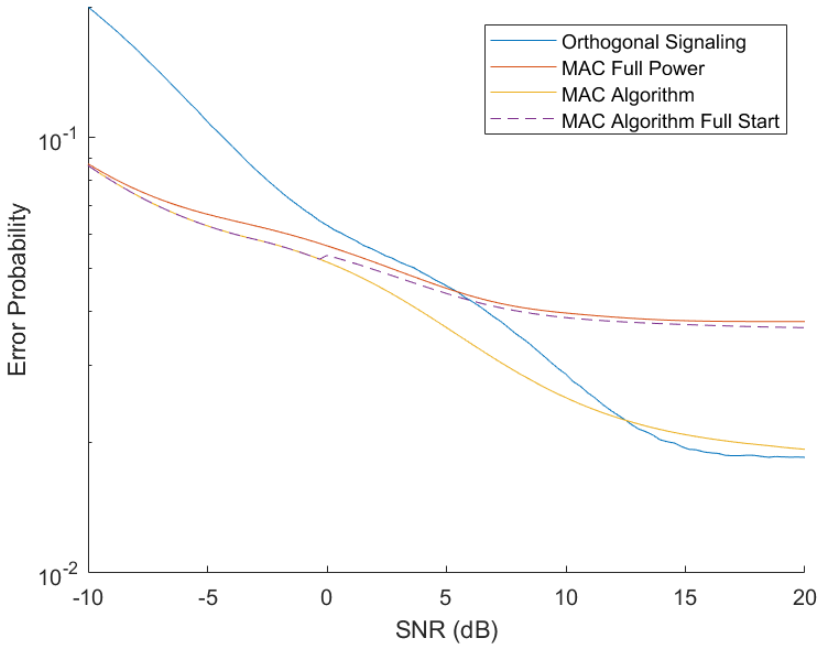


Figure 4.6: Error probability of different signaling schemes as a function of SNR for $N = 8$ ($p_1 = 0.4$, $\epsilon^N = (0.1, 0.1, 0.2, 0.2, 0.2, 0.2, 0.3, 0.3)$, $P^{\max} = (1, 1, 5, 5, 5, 5, 5, 5)$).

We see many of the same observations when comparing the different solutions as mentioned in the previous section. The most notable observation from this figure is that at moderate to high SNR (3 dB or greater), the algorithm when started at

maximum starting powers converges to significantly worse performance than when started at low initial powers. The main condition that causes this discrepancy is when the maximum power allocations for the noisy sensors is too large. This causes the algorithm when started too large to fall into a false minimum at very large powers, and remain there because it is shaped like a trough. Figure 4.7 gives a 2 dimensional interpretation of this phenomenon.

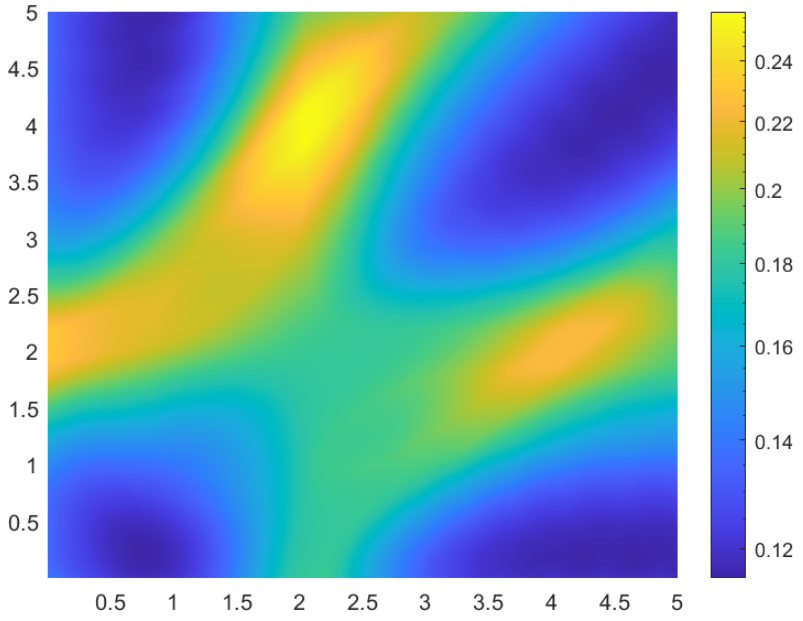


Figure 4.7: Error probability with respect to two powers.

We observe that even though there is a local minimum at small powers, at large powers there are low troughs that the algorithm can get trapped in. Although this does not necessarily mean that the false minimum will perform noticeably worse than the true local minimum. However, we still note that these false minimums will be using significantly more power than the true local minimum. This can be seen in Figure 4.8, where we see the comparison of the amount of available power used by

the two starting points of the algorithm, noting that in particular at high SNR, the maximum power starting point used much more that the smaller starting point. The power usage ratio is calculated as the total power used by the sensors divided by the sum of the maximum power allocations:

$$\text{Power Usage Ratio} \triangleq \frac{\sum_{i=1}^{N} P_i^2}{\sum_{i=1}^{N} P_i^{\max}}. \tag{4.14}$$

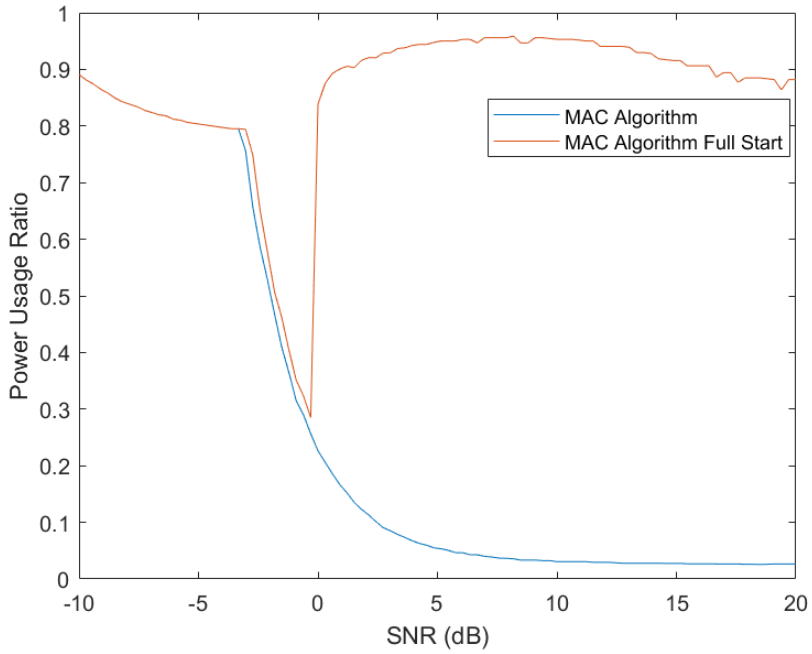


Figure 4.8: Power allocation of different starting points of the optimization algorithm for $N = 8$ ($p_1 = 0.4$, $\epsilon^N = (0.1, 0.1, 0.2, 0.2, 0.2, 0.2, 0.3, 0.3)$, $P^{\max} = (1, 1, 5, 5, 5, 5, 5, 5)$).

For these reasons, we conclude that using a small starting point is an ideal starting point for running the algorithm, as it performs better and uses less overall power than the alternatives. This can also be seen as artificially decreasing the maximum powers for the purpose of the algorithm (which is always a viable thing to do) can make the

algorithm perform better or equivalently.

4.6.3 Improvement Using MAP Detection

In this section, we show that using MAP detection instead of the simplified single boundary detection assumed in the theoretical analysis can lead to improved performance under some circumstances. Figure 4.9 shows the performance of the various signaling schemes with respect to SNR.

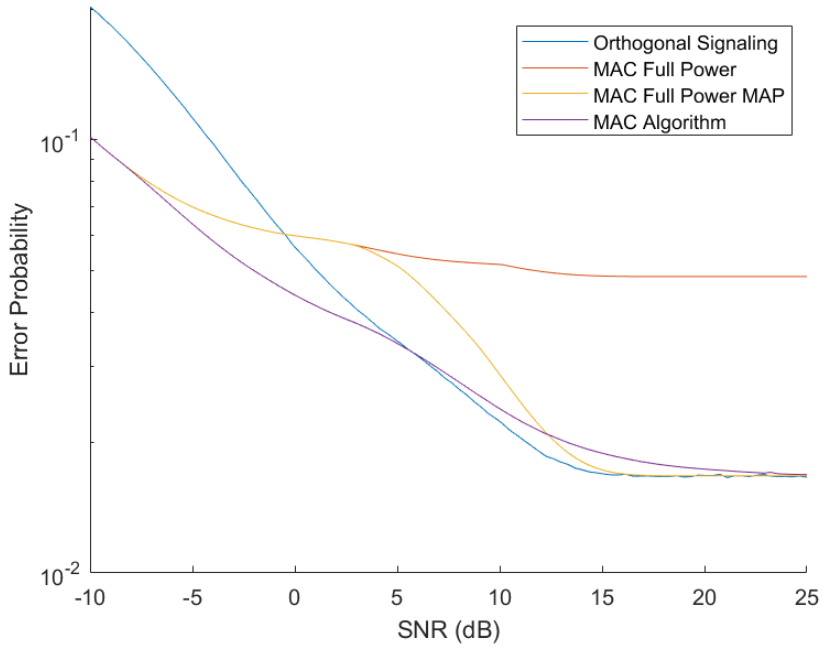


Figure 4.9: Error probability of different signaling schemes as a function of SNR for $N = 5$ ($p_1 = 0.35$, $\epsilon^N = (0.05, 0.1, 0.15, 0.2, 0.2)$, $P^{\max} = (1, 3, 3, 5, 5)$).

We see that the full power MAC approach can become asymptotically equal at high SNR to the better performing methods when using MAP detection, as opposed to staying much above it when using the simplifies detection rule. However, we note the complexity of MAP detection grows with the number of sensors, where in general,

to perform better than simplified detection, MAP detection must be using at least 3, and at most $2^N - 1$ decision boundaries, which can become infeasible to implement.

4.6.4 Large Number of Sensors ($N = 20$)

This illustration demonstrates that the general trends shown in the above examples continue to be prevalent even as the number of sensors grows larger. We take a basic version of the problem where $N = 20$, $p_1 = 0.5$ and half the sensors with crossover probability $\epsilon_i = 0.1$ and the remaining have crossover probability $\epsilon_i = 0.2$. Figure 4.10 shows the performance of the basic signaling/detection designs as SNR varies.

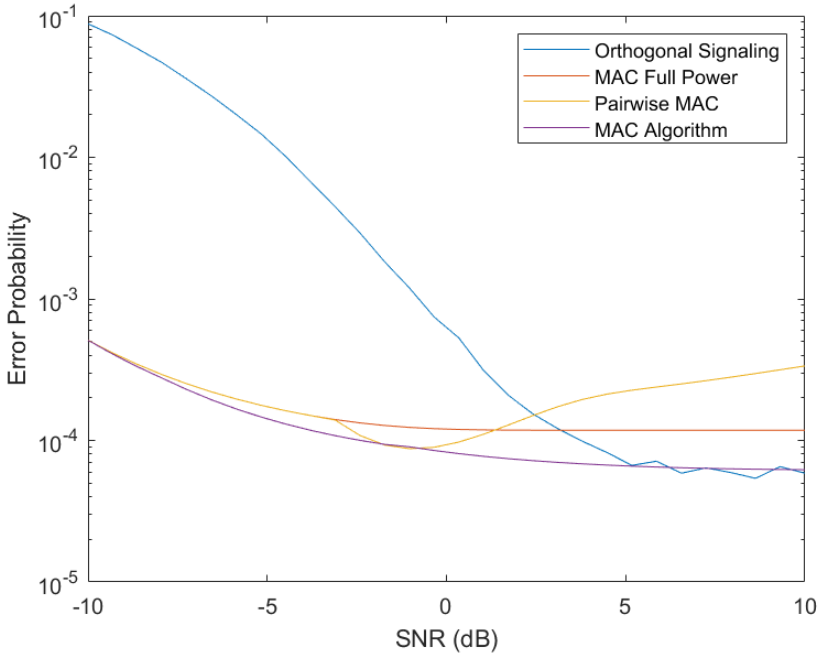


Figure 4.10: Error probability of different signaling schemes as a function of SNR for $N = 20$ ($p_1 = 0.5$, $\epsilon_i = 0.1$, $i = 1, ..., 10$, $\epsilon_i = 0.2$, $i = 11, ..., 20$, $P_i^{\max} = 1$, $\forall i$).

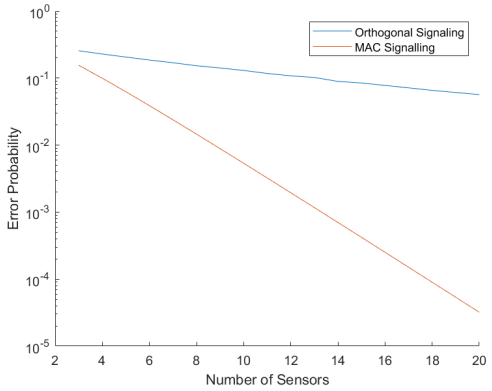
We observe all of the same trends as noted in the previous examples. Specifically, the orthogonal signaling performs worse than all MAC approaches at low SNR,

while at high SNR, orthogonal channels and algorithmic optimization are the better performers, while the simpler MAC schemes perform worse. We note that the performance difference between MAC and orthogonal channels at low SNR is amplified when there is a larger number of sensors. This can be explained by the low SNR region heavily favouring the MAC approaches which can work together and overcome the noise, whereas the orthogonal channels cannot adequately use the increased bandwidth it has access to because the noise power is too large. Further, another interesting feature which is amplified in this example with a large number of sensors is that the pairwise optimization strategy becomes very well performing, essentially identical to the algorithmic optimization at moderate SNR (around -2 dB).

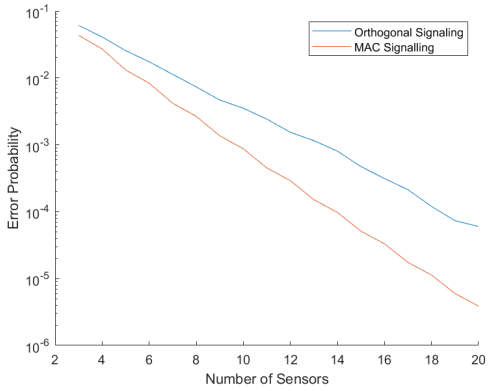
4.6.5 Reducing Number of Sensors

In this section, we demonstrate how the optimization algorithm can be used to reduce the number of sensors required to implement a system with a maximum allowable error probability. In the following examples, we assume each sensor has the same maximum power allocation, so the SNR does not vary when changing the number of sensors. We then choose an example SNR and vary the number of sensors. In the first example, we consider N identical sensors. The results for SNR values of -10, 0 and 10 dB are shown in Figure 4.11.

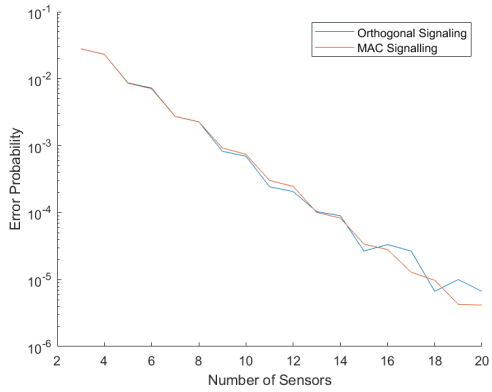
First, we note that in this case all MAC approaches are the same. This is because when all sensors are identical, the optimal allocation is max power to all sensors. Further, in accordance with the previous findings, orthogonal channels perform worse than MAC at low SNR and slightly better at high SNR. At SNR of -10 dB, we can see a 5 sensor gain at 0.15 error probability. At SNR 0 dB we see a two sensor gain



(a) SNR = -10 dB.



(b) SNR = 0 dB.



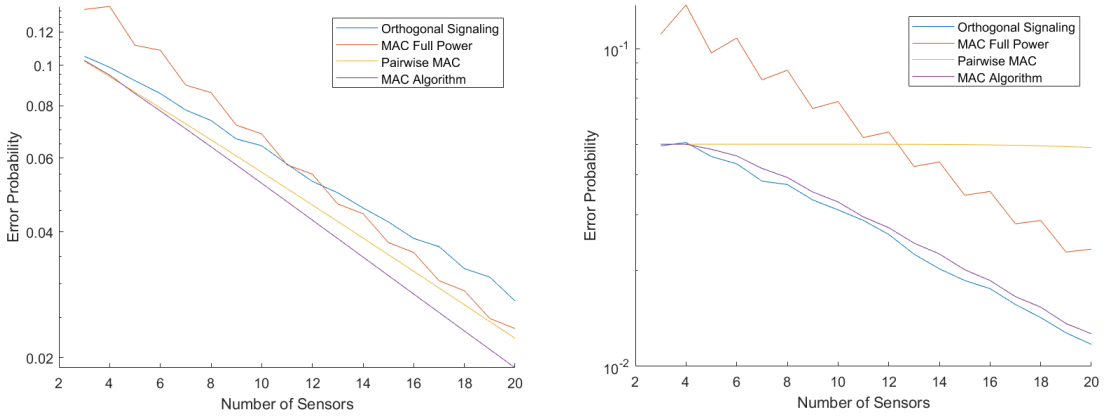
(c) SNR = 10 dB.

Figure 4.11: Error comparison at SNR values ($p_1 = 0.4$, $\epsilon_i = 0.1$, $P_i^{\max} = 1\ \forall i$).

at 0.11 error probability.

The next case that will be considered is where one sensor is better and the rest are worse. This case will allow the algorithmic optimization to benefit over the sub-optimal max power MAC constellations. Figure 4.12 shows how the error probability changes with number of sensors, if Sensor 1 has $\epsilon_1 = 0.05$ and all other sensors have $\epsilon_i = 0.3$, for $i > 1$

First we note that we are not showing a low SNR plot because the behaviour is identical to the previous example, where all MAC implementations perform the same, and consistently better than orthogonal channels. At SNR 0 dB, we remark that the

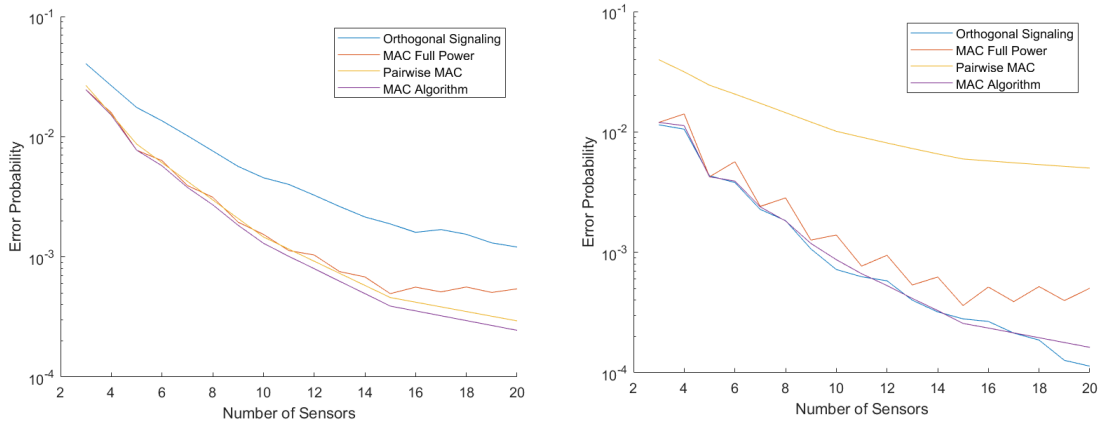


(a) SNR = 0 dB. (b) SNR = 10 dB.

Figure 4.12: Error comparison at SNR values ($p_1 = 0.5$, $\epsilon_1 = 0.05$ $\epsilon_i = 0.3$ $\forall i > 1$, $P_i^{\max} = 1$ $\forall i$).

algorithmic and pairwise optimization techniques perform the best, matching very well at small numbers of sensors (≤ 6). The sensor gain can be seen as 3 at error probability 0.08. Another interesting observation is that for the maximum power MAC approach, addition a sensor can possibly make error performance worse, as seen in the jump from 3 to 4 sensors at SNR 0 dB. This is because adding too much power from a noisy sensor can reduce the effectiveness of the information sent. This also explains the jagged appearance of the curves the max MAC approaches. Finally, at high SNR, we see that the orthogonal channels approach can again perform better than the algorithm, but the algorithm still outperforms the MAC full power approach. Also, we see that the pairwise approach does not change with number of sensors, further demonstrating the flaw in this simple approach as high SNR, as it does not use the additional sensors at high SNR since $P_i^* = 0$, for $i > 1$.

Finally, Figure 4.13 demonstrates a situation where the system has an array of sensors available to use, and for each N, the system makes use of the N least noisy sensors of this array.



(a) SNR = 0 dB. (b) SNR = 10 dB.

Figure 4.13: Error comparison at SNR values ($p_1 = 0.45$, $\epsilon_1 = \epsilon_2 = 0.05$, $\epsilon_3 = \epsilon_4 = \epsilon_5 = 0.1$, $\epsilon_i = 0.15$. $i = 6, ..., 10$ $\epsilon_i = 0.2$, $i = 11, ..., 15$ $\epsilon_i = 0.3$, $i = 16, ..., 20$, $P_i^{\max} = 1$ $\forall i$).

We observe the same trends as the previous two examples in these graphs. We observe that the sensor gain increases as the number of sensors increases, or the desired error probability decreases.

We conclude from this section that the optimization algorithm can achieve the same error performance as other common signaling designs while using less sensors. We also note that in general, the algorithm is only outperformed by other methods which are more complex, use more bandwidth, and especially use MAP detection.

4.6.6 Comparison to General Optimization Algorithms

We compare how our algorithm based on theoretical individual power optimization performs against generalized algorithms derived from gradient descent techniques. Specifically, we compare to a builtin MATLAB optimization function called "fminunc", with the "quasi-newton" option, which is based on the Nelder-Mead simplex algorithm as described in [24]. An important detail is that in general, gradient descent

type optimizations are for unconstrained problems, and do not account for boundary conditions. To handle this, we use the generally observed result from our derived algorithm that the first sensor always converges to using all of its power, so we fix the gradient descent method to use all of the power allocated to Sensor 1, then perform an unconstrained optimization on the remaining sensors, while ensuring that the unconstrained optimization converges within the constraints that the algorithm is bound within. Note that this is easily achieved by increasing the limits of the algorithms optimizations to anything larger than the converged gradient descent results. The gradient descent method is also sensitive to initial conditions, so we use the same values we found to work well for our algorithm, which is starting all sensors at small power allocations. Figures 4.14 and 4.15 show the comparison of error probability and power usage versus noise power, respectively. Note that since the gradient descent problem is unconstrained, it does not make sense to use SNR, so instead we use the inverse noise power: N_0^{-1}. The power usage ratio is calculated as defined in (4.14).

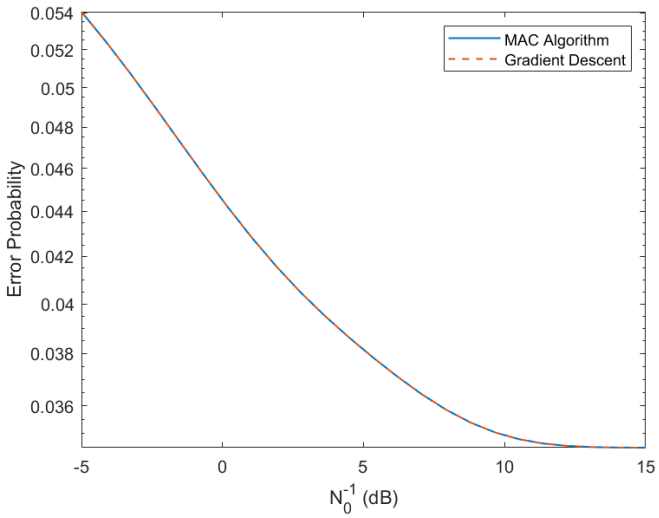


Figure 4.14: Error probability comparison to gradient descent ($N = 5$, $p_1 = 0.4$ $\epsilon^N = (0.1, 0.1, 0.2, 0.2, 0.3)$, $P^{\max} = (1, 4, 2, 2, 1)$).

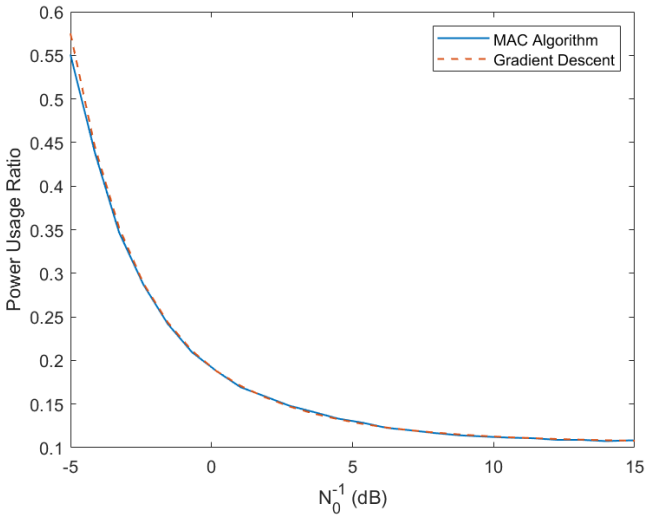


Figure 4.15: Power usage comparison to gradient descent ($N = 5$, $p_1 = 0.4$ $\epsilon^N = (0.1, 0.1, 0.2, 0.2, 0.3)$, $P^{\max} = (1, 4, 2, 2, 1)$).

We observe that both methods perform very similarly in error and power allocations. This strengthens the argument that our algorithm is indeed converging to an optimal value in the problem of MAC signaling. However, we note that the gradient descent method has the following issues with practical implementations. First, unconstrained optimization may not be feasible to implement if the power constraints are smaller than what the unconstrained optimization converges to. Furthermore, applying this unconstrained optimization required using knowledge from the algorithm to force Sensor 1 to use all of its power, otherwise the gradient descent method will not converge and keep increasing power usage. Finally, this optimization method is more complex than standard gradient descent, thus it may not be feasible to implement in a practical situation.

Chapter 5

Conclusion

In this work, various optimization results about constellation design for distributed detection of a binary source over a Gaussian MAC have been explored. A preliminary investigation into the effect of inter constellation rotation in a two sensor uniform source formulation of the problem yielded an optimized upper bound on the error probability which suggested that viewing the problem as a power allocation problem could give interesting and practical results. Following the lead of this result, the optimal one dimensional constellation design for a two sensor binary network was established. After reducing the problem to a power allocation optimization problem (with the appropriate asymmetric constellation designs from Theorem 3.1), it was proved that there are distinct cases that arise based on the fixed parameters of the problem, which are p_1, ϵ_1 and ϵ_2. In some cases (Cases I and II), the results are intuitive and not unexpected, as the optimal power allocations are to use none or all of the available power. However, in Case III, the most interesting and counter-intuitive result is that the optimal power allocation can be for Sensor 2 (with less correlation to the true data source) to use a portion, but not all of its available power. This is a significant result since Case III is prevalent for many parameter sets

of the problem setup. As shown in Figure 3.6, Case III becomes the dominant case as the binary source approaches a uniform distribution.

We used the insights gained from the two sensor problem to formulate results about distributed detection with $N > 2$ sensors. Under the assumption of simplifying the detection to a single decision boundary, the optimization of a single sensor's power allocation when the rest are fixed was characterized. For any sensor, the same three cases analogous to the two sensor problem formed. In the most significant case, a non-trivial optimal power allocation was shown to exist, and although an explicit expression cannot be given, it can be obtained by solving the equations given in (4.9). This individual sensor optimization was developed into an iterative algorithm for the joint optimization of all sensors. This algorithm, with appropriate starting power allocation, was shown to perform well in comparison to other common constellation designs, while consistently using less resources such as power and bandwidth. In practice, applying these optimization results can improve error performance while prolonging the battery life of a sensor network.

To expand further upon our results, the following future directions can be considered. First, if there were a sum power constraint instead of individual power constraints (e.g.., $P^{\max} = P_1 + P_2$), which applies to systems where sensors are co-located (such as connected motion sensors or powered exoskeleton suits), then we readily obtain that the optimal constellation will be of the form given in Chapter 3, with power allocations $P_1 \leq P^{\max}$ and $P_2 = P_2^*(P_1)$. There is no analytic solution to this optimization, so a numerical computation can be carried to determine the optimal values. Also, even without perfect information about the system parameters, such as p_1, ϵ_i, $i = 1, ..., N$ and N_0, our results can still be applied with some form of

estimates. The error probability can have different sensitivities to the ϵ-parameters. The robustness of our optimal signaling scheme can be analyzed when the parameters have estimation errors. Sensor network clustering problems such as those found in [25–28] could be considered. The performance of a cluster could be approximated using the expressions derived in this paper. The balance between error probability and energy efficiency could be investigated as a function of cluster organization.

Bibliography

[1] W. Li and H. Dai, "Distributed detection in wireless sensor networks using a multiple access channel," *IEEE Transactions on Signal Processing*, vol. 55, no. 3, pp. 822–833, 2007.

[2] X. Guo, Y. He, S. Atapattu, S. Dey, and J. S. Evans, "Power allocation for distributed detection systems in wireless sensor networks with limited fusion center feedback," *IEEE Transactions on Communications*, vol. 66, no. 10, pp. 4753–4766, 2018.

[3] H. Jeffreys, "An invariant form for the prior probability in estimation problems," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 186, no. 1007, pp. 453–461, 1946.

[4] C. Tepedelenlioglu and S. Dasarathan, "Distributed detection over Gaussian multiple access channels with constant modulus signaling," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2875–2886, 2011.

[5] S. Dasarathan and C. Tepedelenlioglu, "Distributed estimation and detection with bounded transmissions over Gaussian multiple access channels," *IEEE Transactions on Signal Processing*, vol. 62, no. 13, pp. 3454–3463, 2014.

[6] G. Ferrari, M. Martalò, and A. Abrardo, "Information fusion in wireless sensor networks with source correlation," *Information Fusion*, vol. 15, pp. 80–89, 2014.

[7] S. R. Panigrahi, N. Björsell, and M. Bengtsson, "Data fusion in the air with non-identical wireless sensors," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 5, no. 4, pp. 646–656, 2019.

[8] R. Jiang and B. Chen, "Fusion of censored decisions in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 2668–2673, 2005.

[9] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian "sensor" network," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5247–5251, 2008.

[10] H. Behroozi, F. Alajaji, and T. Linder, "On the optimal performance in asymmetric Gaussian wireless sensor networks with fading," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 2436–2441, 2010.

[11] J.-J. Weng, F. Alajaji, and T. Linder, "Optimized signaling of binary correlated sources over Gaussian multiple access channels," in *Proc. IEEE 88th Veh. Techn. Conf. (VTC-Fall)*, pp. 1–5, 2018.

[12] C.-H. Lin, S.-L. Shieh, T.-C. Chi, and P.-N. Chen, "Optimal inter-constellation rotation based on minimum distance criterion for uplink NOMA," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 525–539, 2019.

[13] H.-P. Liu, S.-L. Shieh, C.-H. Lin, and P.-N. Chen, "A minimum distance criterion based constellation design for uplink NOMA," in *Proc. IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, pp. 1–5, 2019.

[14] K.-H. Ngo, S. Yang, M. Guillaud, and A. Decurninge, "Joint constellation design for noncoherent mimo multiple-access channels," *IEEE Transactions on Information Theory*, vol. 68, no. 11, pp. 7281–7305, 2022.

[15] B. K. Ng and C.-T. Lam, "Joint power and modulation optimization in two-user non-orthogonal multiple access channels: A minimum error probability approach," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10693–10703, 2018.

[16] G. Takahara, F. Alajaji, N. Beaulieu, and H. Kuai, "Constellation mappings for two-dimensional signaling of nonuniform sources," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 400–408, 2003.

[17] B. Moore, G. Takahara, and F. Alajaji, "Pairwise optimization of modulation constellations for non-uniform sources," *Canadian Journal of Electrical and Computer Engineering*, vol. 34, no. 4, pp. 167–177, 2009.

[18] S. Vlajkov, A. Jovanović, and Z. Perić, "Approach in companding-quantisation-inspired PAM constellation design," *IET Communications*, vol. 12, no. 18, pp. 2305–2314, 2018.

[19] L. Wei, "Optimized M-ary orthogonal and bi-orthogonal signaling using coherent receiver with non-equal symbol probabilities," *IEEE Communications Letters*, vol. 16, no. 6, pp. 793–796, 2012.

[20] L. Sardellitti, G. Takahara, and F. Alajaji, "Optimized constellation design for two user binary sensor networks using NOMA," *arXiv:2305.19462*, 2023.

[21] L. Sardellitti, G. Takahara, and F. Alajaji, "Optimal binary signaling for a two sensor Gaussian MAC network," *IEEE Transactions on Communications*, 2024.

[22] L. Sardellitti, G. Takahara, and F. Alajaji, "Optimized distributed detection over a two sensor binary Gaussian MAC network," in *Proc. Canadian Conference of Electrical and Computer Engineering (CCECE)*, pp. 1–2, 2024.

[23] G. J. Jameson, "Counting zeros of generalised polynomials: Descartes' rule of signs and Laguerre's extensions," *The Mathematical Gazette*, vol. 90, no. 518, pp. 223–234, 2006.

[24] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the nelder-mead simplex method in low dimensions," *SIAM Journal of Optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[25] S. Yi, J. Heo, Y. Cho, and J. Hong, "Peach: Power-efficient and adaptive clustering hierarchy protocol for wireless sensor networks," *Computer Communications*, vol. 30, no. 14, pp. 2842–2852, 2007.

[26] E. Pei, J. Pei, S. Liu, W. Cheng, Y. Li, and Z. Zhang, "A heterogeneous nodes-based low energy adaptive clustering hierarchy in cognitive radio sensor network," *IEEE Access*, vol. 7, pp. 132010–132026, 2019.

[27] G. Chen, F. Nocetti, J. Gonzalez, and I. Stojmenovic, "Connectivity based k-hop clustering in wireless networks," in *Proc. 35th Annual Hawaii International Conference on System Sciences*, pp. 2450–2459, 2002.

[28] A. Bendjeddou, H. Laoufi, and S. Boudjit, "Leach-s: Low energy adaptive clustering hierarchy for sensor network," in *Proc. International Symposium on Networks, Computers and Communications (ISNCC)*, pp. 1–6, 2018.