Distributed Detection and Coding in Information Networks

by

Shan-Yuan Ho

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2006

Author ....
Department of Electrical Engineering and Computer Science
January 31, 2006

Certified by ....
Robert G. Gallager
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by ....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Distributed Detection and Coding in Information Networks

by

Shan-Yuan Ho

Submitted to the Department of Electrical Engineering and Computer Science on January 31, 2006, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

This thesis investigates the distributed information and detection of a binary source through a parallel system of relays. Each relay observes the source output through a noisy channel, and the channel outputs are independent conditional on the source input. The relays forward limited information through a noiseless link to the final destination which makes an optimal decision. The essence of the problem lies in the loss of information at the relays. To solve this problem, the characteristics of the error curve are established and developed as a tool to build a fundamental framework for analysis. For understanding, the simplest non-trivial case of two relays, each forwarding only a single binary digit to the final destination is first studied. If the binary output constraint is removed and the output alphabet size for one relay is M, then no more than $M + 1$ alphabet symbols are required from the other relay for optimal operation. For arbitrary channels, a number of insights are developed about the structure of the optimal strategies for the relay and final destination. These lead to a characterization of the optimal solution. Furthermore, the complete solution to the Additive Gaussian Noise channel is also provided.

Thesis Supervisor: Robert G. Gallager
Title: Professor of Electrical Engineering and Computer Science

Acknowledgments

"At times our own light goes out and is rekindled by a spark from another person. Each of us has cause to think with deep gratitude of those who have lighted the flame within us."

- Albert Schweitzer

This is the close of a continuously disrupted and extremely long chapter in my life. Therefore, these acknowledgements are long and yet cannot fully express my gratitude. I am most deeply indebted to my advisor and mentor, **Professor Robert G. Gallager**. His extreme patience, generosity, and genuine concern for the welfare of his students is unparalleled. Through his mentorship, I received an education in many facets far exceeding the most hopeful expectations of any graduate student. This research experience has been an extraordinary journey, which I have thoroughly enjoyed. I was shown the beauty and style of simplicity in the way of fundamental research and was given the freedom to independently explore my ideas carefully guided by his unprecedented research instinct - where to explore, where to explore further, and when to stop. Few possess a passion for research and even fewer possess the type of deep insight and understanding of both the grand perspective and the minute detail that Prof. Gallager demonstrates. He taught me the clarity of thinking and the profoundity of thought through the simplicity of solution during the research process. He then instructed me in the lucidity and elegance of exposition during the writing process. I was also the beneficiary of Prof. Gallager's wisdom and counsel in various aspects of life, which deeply impacted me. One of the most humble, generous, and honorable people I have ever known, Prof. Gallager's honesty is always stated with graciousness. His example is an inspiration, and his influence has changed my life. I am immensely fortunate to have had the privilege of being a student of Prof. Gallager.

I gratefully acknowledge my committee members from whom I have learned so much: **Professor John Tsitsiklis** provided invaluable insight, depth of knowledge, and precise statements. I admire his clarity of understanding and brilliance. **Professor Emre Telatar** gave a meticulous reading of my thesis and enormous help in many ways. His insight, clear explanations, depth of understanding in so many fields, and love of research are exceptional. **Professor Daniel Spielman** provided astute suggestions and many helpful discussions over the years. A number of other professors have had a signficant influence on my education. I am particularly privileged to have had many enlivening interactions with **Professor Daniel Kleitman**, and I thoroughly enjoyed both our technical and non-technical discussions. His ability to create valuable learning environments anywhere and to find no problem uninteresting is uncanny. I have learned so many tools of the trade and richly benefitted from his remarkable advice and stories. **Professor Al Drake** was always willing to provide varied input. He included humor in every interaction and understood the soothing of the soul after a long peaceful motorcycle ride. **Professor Pierre Humblet** first introduced me to the blending of mathematics with engineering in research. His sharp technical insight and kindness is more than appreciated. I thank **Professor Yih-Jian Tai** for reviving in me my love of literature and history. Through our many scholarly discussions, I felt privileged to to be his student in the modern and classical sense. To **Professor Eric Feron** who generously gave me his kind help and constant encouragement. To **Professor Dimitri Bertsekas** for many useful and pleasant interactions. To **Professor Martha Gray** for her invaluable advice and honest perspective. **Professor Dave Forney** and **Professor Sanjoy Mitter's** enthusiasm for research and fruitful interactions have inspired me. To **Professor and Mrs. Vernon Ingram** for the generosity of their love and kind support and wise counsel at all times.

I am especially thankful to **Marilyn Pierce** for her caring support, kindness, and everything she has done for me throughout my graduate career. She took care of every administrative problem that came my way and allowed me to forget the definition of "red tape." The completion of this degree would have been impossible without her. I thank **Lisa Bella**, **Bob Bruen**, **Peggy Carney**, **Jennifer Craig**, **Lisa Gaumond**, **Michael Lewy**, **Fifa Monserrate**, **Kathleen O'Sullivan**, **Dennis Porche**, **Peter Swedock** for their expert help at needed times.

Many splendid friends, colleagues, and officemates in the Department of Mathematics, Department of EECS, the Laboratory for Information and Decision Systems, and MIT have contributed richly to my graduate experience, during and beyond.

To **Emre Telatar** who is the best friend that anyone could possibly ask for, who puts up with my antics at all times, who always provides a listening ear, and who has always been there for me. His sensitivity, graciousness, kindness, humbleness, and integrity have never wavered. To **Meng-Yong Goh** whose demonstration of utmost loyalty and true friendship since our college days is unsurpassed, for not tiring of hearing my nonsense at all odd hours of the night, and who always has my best interest at heart. I enjoyed my friendship and the many thoughtful discussions with **Tom Richardson** and **Roy Yates** and appreciate their honest, down-to-earth perspective, including Roy's sometimes useful advice. **Erdal Arikan** whose positive encouragement, advice, and insight is fresh water to a weary soul every time we meet, and whose love of research and learning is infectious. **John Keen** for his constant encouragement, support and friendship during my roughest times, without a hint of negativity. I had some unforgettable officemates who have certainly slowed down the research and dissertation process but created many fond memories in the meantime. They include **Murat Azizoglu, Rick Barry, Angela Chiu, Ken Chou, Ying Li, Zhi-Quan (Tom) Luo, Muriel Médard, Rajesh Pankaj, Abhay Parekh, John Spinelli, I.E. Telatar, Paul Tseng, Cynara Wu**. I thank **Alex Postnikov** for the discourses in mental agility; **Diana Dabby** for reminding me that I am still alive; **Sanjeev Kulkarni** for always keeping an even keel; **John Spinelli** and **Mark Wang** whose humor is unmatched; **Rocky Romano** for honing me in the art of procrastination; **Chi-Fang Chen** whose ability to relax in just about every situation had a calming effect; **Kathy Misovec** who frequently questioned my sanity but never doubted my courage or integrity.

I thank **Shelene Chang** for her creative and brilliant solutions to some of my out-of-the-ordinary problems. **Tom Klemas** - fellow coach, instructor, researcher, and student (both undergrad and grad) for sharing all sorts of thoughts and experiences. We share a rare common language and ground to talk in breadth and depth. **Dan and Lauren Crews** and **Rita and John Bonugli** for the stress relief and the many delicious home cooked meals. My good friend **Michelle Bonugli** for her loyal friendship, hidden talents, unforgettable laughter, and unrepeatable jokes (which we incessantly repeat anyway!). **J.P. Mattia, Paparao, Dana Pascovici, Richard Stone** created many pleasant distractions through this sometimes tedious ordeal. I have enjoyed the camaraderie of **Constantinos Boussios, Randy Berry, Stefano Casadei, Constantine Caramanis, Emre Koksal, Olivier Leveque, Marina Meila, Angelia Nédic, Asu Ozdaglar, Anant Sahai, Tengo Saengudomlert, Huizhen Yu**.

I am very appreciative of **Georgios Kotsalis** for his unfailing support, friendship and strength of character. His insight into people is amazing and his courage and integrity are admirable. I am also very grateful for the kind understanding, patience, and constant support of **Fang Zhao's** friendship and appreciate her kind help with many LaTeX issues in this document, particularly during the final stages. I will fondly miss the midnight crew of **Selcuk Bayraktar, Georgios Kotsalis, Nuno Martins** where the language and topics of conversations were somehow much more creative than during the day. Fellow instructors, TA's, and former students (academic and athletic): **Andrew Brzezinski, Jenn Cheng, Karolina Corin, Chuying Fang, Brian Heng, O. Pat Kreidl, Charlotte Lau, Josh Lifton, Chunmei Liu, Karen Sachs, David Sachs, Keith Santarelli, Ayanna Samuels, Jolene Singh, Heelo Sudo, Xiaoyan Xie, Fangyun Yang** made my experiences at MIT rich and rewarding.

I sincerely thank MIT athletic director **Dr. Richard Hill** for his friendship and support; Aikido

Sensei Dick Stroud for teaching me the finer points of harmony, balance and sensitivity, not only on the mat but in life; the MIT fencing team coaches **Eric Sollee, Jarek Koniuz**, and **Rob Hupp**, MIT Ice Hockey Team coaches **Katia Paskevitch** and **Tony Gray**, for the many hours of instruction, fun, and collegiate competitions. I especially enjoyed many enlightening conversations with my friend and men's gymnastic coach, **Noah Riskin**. I shall remember members of the **MIT Men's Gymnastics Team** who provided entertainment and also helped me with my gymnastics.

A number of wonderful people outside the world of MIT have been especially kind to me. These very special individuals have inspired and supported me at various points during my graduate studies and in my life.

To **Master Joe Pina** who is the most dedicated and devoted coach that any athlete could ever ask for, who taught me the art of martial art through the sport of Tae-Kwon-Do, and whose intense competition training was no less gruesome than any part of my academic education. I had to dig deep within my soul to deliver when I felt there was nothing left to offer. **Master Martin Mackowski** who is without doubt the finest instructor I have ever had and who built my solid foundation in the basics. To **Master Deb Pederson** for whom I would have no other coach in my corner during my competitions. To **Master Chris Jensen** for refining my techniques and basics from the beginning, whose precision and martial analysis is unsurpassed. To my teammates and training partners in California and Massachusetts, thanks for the sweat, pain, and agonizing workouts where we learned that *whether sport develops character is questionable, but one thing we all know for certain, SPORT REVEALS CHARACTER.* Many thanks to: **Eric Lee** who is without question the finest sparring partner I have ever had and who has prepared me for many competitions since our teenage years; World Champions **Valery Long** who taught me almost everything I know about distancing, timing, and footwork and prepared me firsthand in our many sparring sessions throughout my competitive training; **Marcia Hall** for the high standards and inspiration; **Dana Hee, Deb Pederson, Val Long, Henry Parker, Pedro Xavier, Alberto Montrond, Noel Bennett** for the day-in-day-out training and the numerous competitions. **Sean Clark** whom our coach calls "the World Champion yet to be", whose support, friendship, and loyalty is doubly appreciated. **AJ Dunbar** the best physical trainer I have come across. **Heather Tallman** who is perhaps the most diverse and talented individual I have ever met and who has demonstrated that true friendship is timeless.

To my dear friend **Dana Hee**, Olympic Gold Medalist and Hollywood Stuntwoman, I shall always treasure those sweet memories of intense training, competition, and travel. It was a glorious adventure. Since our youth, we deeply share and dearly embrace the invaluable lessons life and experience have taught us. We painfully learned to maximize working with our pre-birth lottery (genetics) in the circumstance of our post-birth lottery (environment) and came to the revelation that no one can take away what is within us.

I am indebted to **Kim McGlothin** and **John Blumenstiel** for their expert advice on a variety of issues, pragmatic help at times of crisis, and for teaching me a perspective on human nature that I will carry into the future chapters of my life. I would also like to acknowledge the kind support and insightful understanding of **Grace Chiang** and **Ming-Wen Lu**.

I thank my entire medical team at Spaulding Rehabilitation Hospital. Their comprehensive guidance through the rehabilitation process over 3 years helped me rebuild my life after my near fatal accident. This process was extremely painful and trying; however, the support of the entire team gave me hope, comfort, and motivation. I am especially grateful to my Speech Language Pathologist, **R. Richard Sanders** whose expertise, insightfulness, and patience during our speech therapy sessions were critical to my rehabilitation.

My deep gratitude to **Grandpa and Grandma Liu** for their warm love and keen interest in all aspects of my life since my youth. I especially thank **Grandpa Liu**, my maternal grandfather's best friend, for his insightful advice over the years, his deep enthusiasm in all my undertakings, and for always standing behind me in all my pursuits. I am indebted to my **Maternal Grandfather** for loving me and for believing in me even before I was cognizant of it. His high standards of integrity, fairness, generosity, and strength of character will forever be imprinted on my heart. I thank my sister and brother-in-law, **Della and Steve Lau** for their help and support at various times.

To my loving **Mother** for demonstrating a zeal for life, a passion for learning, the courage to endure, for teaching me never to take anything in life for granted, and for instilling in me a love for and appreciation of mathematics. But most of all, for loving me deeply at all times, for always believing in me especially during my lowest times, and for enthusiastically sharing the experiences in all my endeavors. She always patiently listened to me, generously shared her experiences, and provided a keen insight and type of wisdom and understanding that few possess. I will never forget her sacrifices. Her strength of spirit has sustained me, her wisdom has guided me, and her love has nurtured me.

This research is partly supported by a MIT Alumni Fellowship, General Electric Fellowship, MIT EECS Graduate Fellowship, and Grass Foundation Fellowship, and is kindly appreciated.

Wisdom is supreme; therefore get wisdom.
Though it costs all you have, get understanding.

Turn your ear to wisdom and apply your heart to understanding,
Call out for insight, and cry aloud for understanding,
Look for it as for silver and search for it as for hidden treasure...

Then wisdom will enter your heart and
Knowledge will be pleasant to your soul.
Discretion will protect you and understanding will guard you.
Then you will understand what is right and just and fair – every good path.
Preserve sound judgement and discernment,
Do not let them out of your sight;
They will be life for you.

Blessed is the man who finds wisdom, the man who gains understanding,
For she is more profitable than silver and yields better returns than gold;
She is more precious than rubies;
And nothing you desire can compare with her.

Proverbs 2:2-4,9-11; 3:13-15,21-22; 4:5-7

謹獻給

我可敬的　摯愛的母親

我深受崇敬的導師

我榮耀的天父

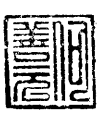
因為他們

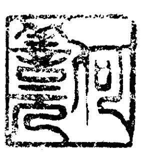
在我人生最艱難、低谷的時刻支持我。

將他們敏銳的智慧、珍貴的閱歷和深刻的洞察力賦予我。

並且對於我的領悟、成長、成功和備受啓迪而深感歡欣。

何善元





Dedicated to:

My Honorable and Loving Mother
My Admirable and Respected Professor
The God of the Cosmos

Because they

Supported me during my darkest hours and lowest times,

Provided streams of profound wisdom, invaluable knowledge and experience, and deep insight,

And are truly happy for my understanding and growth, my successes, and my enlightenment.

Shan-Yuan Ho

Contents

List of Figures

Chapter 1

Introduction

1.1 Thesis Problem

There are many communication network situations where relays must each transmit a limited amount of information to a single destination which must fuse all the information together. For example, a large number of sensors could all be observing some source data which must be conveyed to a fusion center through limited communication channels. The measurements of the data are typically noisy, and the sensors usually are not able to communicate with each other since they are at physically different locations. Since each sensor is allowed to transmit very limited information to the destination, without knowing what the other sensors observe, it is not clear what information each sensor should send to the single destination. Another common example lies in a network of nodes where a number of nodes must act as relays but have no knowledge, except that from the source, about what the other relays receive. These relays all receive the same data, each corrupted by different noise, and must relay partial information about the data to a single destination. The destination fuses the information from these nodes to determine the original data. All relays and the destination have full knowledge of the code used at each relay and the decoding rule at the destination. It is assumed that no coding is permitted at the source. We assume, however, that all sensors know the statistics of the source, the noise statistics at all sensors, and the strategies of all sensors. Figure 1-1 depicts some typical scenarios.

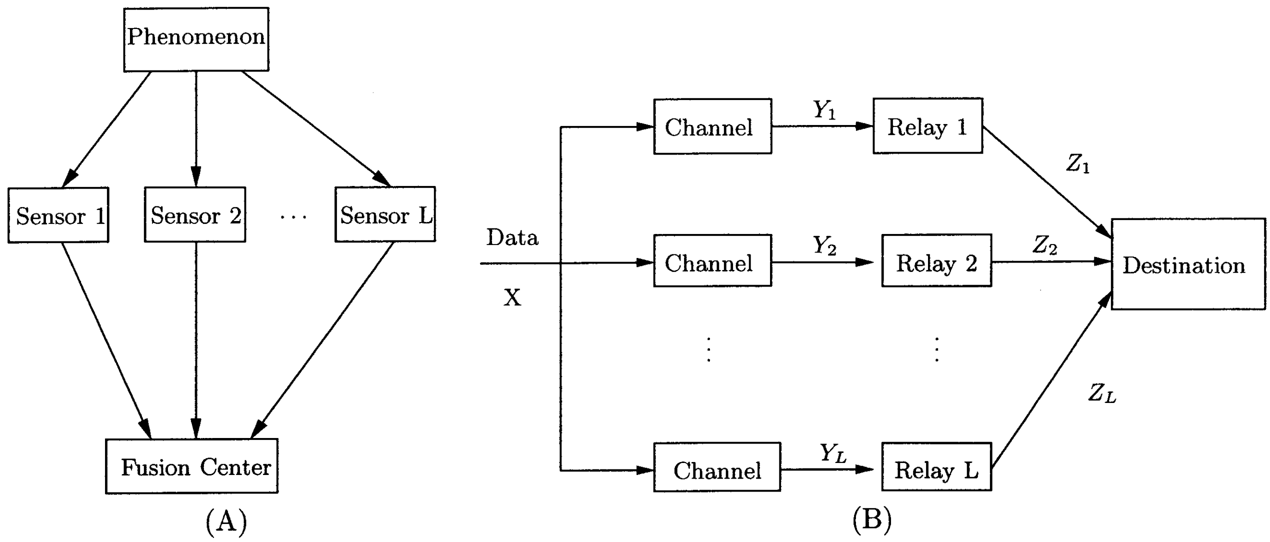


Figure 1-1: Sensors (A) and Relays (B) obtain noisy information and transmit/relay limited information to the Fusion Center/Destination

If the source X originates from a discrete distribution Q_X, the problem involves distributed detection. If the data originates from a continuous distribution F_X, then the problem involves a distributed estimation problem. We restrict ourselves to distributed detection where the source is an independent identically distributed (i.i.d.) sequence generated from a fixed distribution Q_X. In fact, we will usually restrict the source to be binary. The relay/sensor's observation space Y, which can be either discrete or continuous or both, will be quantized or coded and then sent to the destination or fusion point. The communication constraint between relay and fusion point determines the fineness of the relay quantization levels which we assume will be transmitted noiselessly. Otherwise, coding could be implemented on the quantized bits, and from the Coding Theorem, these bits could eventually be received by the fusion center with arbitrarily small probability of error.

Each channel can be modeled as a general discrete-time memoryless channel characterized by a conditional probability distribution for each input. One example of interest is the additive Gaussian noise channel where the noise has zero mean and unit variance. Another example of interest is the class of discrete memoryless channels (DMC) where the noise is described by a probability transition matrix $P(Y \mid X)$.

Before studying the bigger problem of the distributed detection of an iid source sequence $X^n = \{X_i\}_{i=1}^n$, we first study the one shot distributed detection problem of a single source symbol X. The one shot problem of detecting a single source symbol does have some advantages over encoding an entire sequence, such as no delay. However, it leaves less choice at the relays.

A relay can at most quantize the channel output, whereas encoding a long source sequence allows more sophisticated encoding at the relays, possibly resulting in better performance. For the one shot problem, given the relay encoders, it is straightforward to find the optimal fusion decoder. For certain special cases, given the fusion decoder, it may be possible to apply optimization techniques to finding the optimal relay encoders. In general, it is still difficult. For the distributed problems in this thesis, the simultaneous optimization of relay encoders and fusion decoder is necessary.

1.1.1 Special Case of Noiseless Channels

Much of the distributed nature of the system would be lost if the relays observed noiseless replicas of the data. Nevertheless, the noiseless observation case characterizes the asymptotic limit of very high signal to noise ratio (SNR). Suppose X is a discrete random variable and there are n relays, each of which can send at most one binary symbol. The fusion point receives n binary variables and decodes by mapping each binary n-tuple into some source symbol. Assume we have a minimum error probability criterion. Then the minimum error probability can be achieved if we could provide a distinct n-tuple from the set of relays to each of the most likely symbols and arbitrarily place each less likely uncoded symbol into the quantization bin of any coded symbol. Since each relay observes exactly the same thing, each can simply send one bit of the n-bit representation of the specified quantization code; the fusion center will then choose the most likely symbol in the specified quantization bin. For instance, Relay A can send the most significant bit, Relay B the next most significant bit, and so forth. The probability of error P_e is simply the sum of the probabilities of all the least probable symbols (the originally uncoded symbols) which are in a quantization bin with more than one symbol. For example, suppose the system consists of two relays, each allowed to send one bit. There are six data symbols with the following probabilities, $\Pr(A, B, C, D, E, F) = (\frac{6}{21}, \frac{5}{21}, \frac{4}{21}, \frac{3}{21}, \frac{2}{21}, \frac{1}{21})$. The destination receives a total of two bits. To minimize error probability, encode the four most probable symbols $A \to 0$, $B \to 1, C \to 2$, $D \to 3$. Relay A sends the most significant bit with the mapping $\{A, B\} \to 0$ and $\{C, D\} \to 1$, and Relay B sends the least significant bit with the mapping $\{A, C\} \to 0$ and $\{B, D\} \to 1$. The decoder decodes $00 \to A$, $01 \to B$, $10 \to C$, $11 \to D$ with zero error. An error occurs if E and F are observed, so $P_e = \frac{1}{7}$.

1.1.2 Special Case with Unlimited Communication from the Relays

The system also loses its distributed nature if there is no constraint on transmission rate from relay to destination. From the fusion center's perspective, the accuracy of the measurements increases as the allowed relay transmission rate increases. When each relay sends its full observation Y, the scenario becomes a standard detection problem of estimating a random variable from a finite number of independent samples. Nevertheless, this case provides a lower bound on the error probability given a finite number of relays. Furthermore, if coding is allowed at the relays, this case also provides an upper bound on the maximum mutual information of the system in the asymptotic limit of increasing bit rate. The relay can use standard data compression on its output. The compression limit is $H(Y) = -\sum_y \log \Pr(y)$. Therefore, as the relay output bits increase from 1 to $H(Y)$, the system goes from a distributed to a centralized system.

1.1.3 Example of a Gaussian Channel with One Bit Relay Outputs.

We now look at the one shot distributed detection problem of a single source symbol X. To illustrate some aspects of this problem, let's look at the following simple, albeit important, example. Two relays, call them Relay A and Relay B, each observe a binary signal independently corrupted by 0-mean Gaussian noise. Each restricted to send at most one bit to the destination. The binary symbols are equally likely, the noises are independent of the input symbol and each other, and the receivers are not allowed to communicate nor have any side information about the other receiver's observation. The observations of relays A and B are, respectively, $Y = X + N$ and $Y' = X + N'$, $x \in \{-1, +1\}$, $\Pr(X = +1) = \Pr(X = -1) = 1/2$, and N, N' are Gaussian $\mathcal{N}(0, \sigma^2)$. Since only one bit relay transmission is allowed on a single source observation, it seems reasonable that the relay should just quantize the observation space. It is well known that that a threshold test on the likelihood ratio (LR) is the optimal strategy. For increased understanding, we will show via a different argument in Chapter 2, that a threshold tests are optimal. More generally, a quantization of the likelihood ratio of each relay's observation is optimal when they are allowed more symbols; this will be shown in Chapter 4. It is well known for the binary Gaussian detection problem that if each node is allowed to send its full observation to the decoder, the optimal decision boundary for Maximum Likelihood (ML) decoding is the line $y = -y'$. However, this is impossible to achieve since each relay is only allowed to send one bit to the destination. Upon receiving the two bits, the fusion center must decide whether the original symbol is 1 or -1 with minimal error probability. This certainly appears like the classical detection problem of two independent samples of the data, so one would naturally think that the optimal solution would be for each relay to use a binary quantization with the origin separating the quantization regions. Letting $Z \in \{-1,+1\}$ and $Z' \in \{-1,+1\}$ denote the quantizer output of the relays, then $Z=-1$ for $Y<0$ and $Z=+1$ for $Y \geq 0$. Likewise for Z'. The fusion center maps $(-1,-1) \rightarrow (-1)$ and $(+1,+1) \rightarrow (+1)$, and arbitrarily decides the input for a $(-1,+1)$ or $(+1,-1)$ reception. This encoding/decoding scheme has the same error probability as if only one relay observation were used and the other relay ignored. The second relay does not help. For this specific quantizer, the reception of $(-1,+1)$ or $(+1,-1)$ carries no information and is a "don't care" map to either source symbol. To achieve better performance, it seems that $(-1,+1)$ or $(+1,-1)$ will need to carry some kind of information and having a second relay in the system should help. The importance of utilizing the $(-1,+1)$ and $(+1,-1)$ reception will become evident in the ensuing discussion.

Note that the map from the system input to the output of a relay takes the binary source input into a binary relay output. For the Gaussian problem above with a quantization at the origin, this map effectively produces a binary symmetric channel (BSC) with crossover probability $\mathcal{Q}\left(\frac{1}{\sigma}\right)$ where $\mathcal{Q}(x)$ is the complementary distribution function[1] of the zero mean, unit variance Gaussian density $q(x)$ where

$$\mathcal{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-u^2/2} du \qquad q(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad \frac{d}{dx}\mathcal{Q}(x) = -q(x).$$

Note that $1-\mathcal{Q}(x) = \mathcal{Q}(-x)$. Surprisingly, as we prove later in Chapter 3, the optimal solution to this binary input additive Gaussian noise (AGN) 1-bit problem with equal priors and 2 relays is <u>not</u> a binary quantization separated at the origin, but rather a binary quantization separated at a point strictly away from the origin with each relay using the the same quantizer. These results were first shown in [IT94]. For one of the optimal solutions, both thresholds T and T' are set equal to T^* where the relay outputs and T^* are given by

$$Z_A(y) = \begin{cases} -1 & \text{for} \quad y < T^* \\ 1 & \text{otherwise,} \end{cases} \quad ; \quad Z_B(y') = \begin{cases} -1 & \text{for} \quad y' < T^* \\ 1 & \text{otherwise,} \end{cases} \tag{1.1}$$

and

$$T^* = \frac{\sigma}{2} \ln \left[\frac{1-\mathcal{Q}\left(\frac{T^*+1}{\sigma}\right)}{1-\mathcal{Q}\left(\frac{T^*-1}{\sigma}\right)} \right]. \tag{1.2}$$

[1] A slight variant of $\mathcal{Q}(x)$, used in the communication field, is the error function $erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$, which is also called the Gauss error function.

The decision rule at the fusion point is

$$\hat{X} = \begin{cases} -1 & \text{for}(Z, Z') = (-1, -1) \\ +1 & \text{otherwise} \end{cases} . \tag{1.3}$$

The resulting probability of error is

$$P_e = \frac{1}{2}\left(1 - \left[1 - Q\left(\frac{T^* + 1}{\sigma}\right)\right]^2 + \left[1 - Q\left(\frac{T^* - 1}{\sigma}\right)\right]^2\right). \tag{1.4}$$

From the symmetry of the signals, there exist two optimal solutions. The other solution simply uses the relay quantizer at $-T^*$ and a fusion point decision mapping of $(+1, +1) \rightarrow +1$ with everything else mapped to -1. This is a totally symmetric problem with an asymmetric solution. The asymmetry takes advantage of using $(-1, +1)$ and $(+1, -1)$ to carry information. It will soon be apparent how the gains in the asymmetric solution outweigh the costs. We now provide some intuition into the plausibility of this somewhat surprising result.

Let T and T' be arbitrary thresholds for Relays A and B, respectively. The encoding vector $\vec{T} = (T, T')$ will split the observation region into 4 quadrants each corresponding to one of the 4 possible received vectors $\vec{Z} = (z, z')$ by the fusion point. Given any $\vec{T}$, the optimal decoder is easily calculated. In particular, for each quadrant, decode to the symbol which gives the minimum error probability. The following graphical argument shows how to determine the symbol for each region.

First assume that $T + T' > 0$ and observe the different decoding regions in figure 1-2. It is clear that when the fusion point receives $(+1, +1)$ which is region I, it should map it to "+1" and region IV should be mapped to "-1". Region II can be divided into 2 regions, namely , A and B as shown in figure 1-2. The error probability is the same regardless of what region B is mapped to. Refer to these types of regions as "don't care" regions. Region A should be clearly mapped to "+1" since the likelihood of "+1" is greater than "-1" for every point in region A. Therefore, region II as a whole should be mapped to "+1". The same argument applies to region III. Region D=$\{(y, y') : y \leq -T'$ and $y' \geq T'\}$ is a "don't care" region and region C=$\{(y, y') : -T' \leq y \leq T$ and $y' \geq T'\}$ should be mapped to "+1." Thus, region III as a whole should also be mapped to "+1." Now that the fusion decoder is fixed, we find the optimal T and T' that will minimize total error. Given the decoder map of (I,II,III,IV)=(+1,+1,+1,-1), the optimal $\vec{T}$ equals (T^*, T^*), this will be derived in Chapter 3. By symmetry, if $T + T' < 0$, the optimal $\vec{T}$ equals $(-T^*, -T^*)$. The optimal encoder and decoder map is unique and shown in figure 1-3. Observe that when $T + T^{'} = 0$, region A disappears and region II is just region B which is a "don't care" region. Likewise, region C disappears and region III is just region D which is also a "don't care" region. The regions can be optimally mapped into any of the following four choices $\{(+1,+1,+1,-1),(+1,-1,+1,-1),(+1,+1,-1,-1),(+1,-1,-1,-1)\}$. The third and fourth choices are equivalent to decoding on the basis of the first relay alone or the second relay alone respectively. This is equal to the error probability of just a single relay observation as seen before.

The ML error probability for the case $\vec{T} = (T^*, T^*)$ is less than the ML error probability when $\vec{T} = (0,0)$. We make two observations: First, the second relay is no longer redundant but useful in reducing error; and second, the "don't care" regions are now utilized. In other words, the $(-1,+1)$ and $(+1,-1)$ receptions are no longer "don't care" maps since they now carry information.

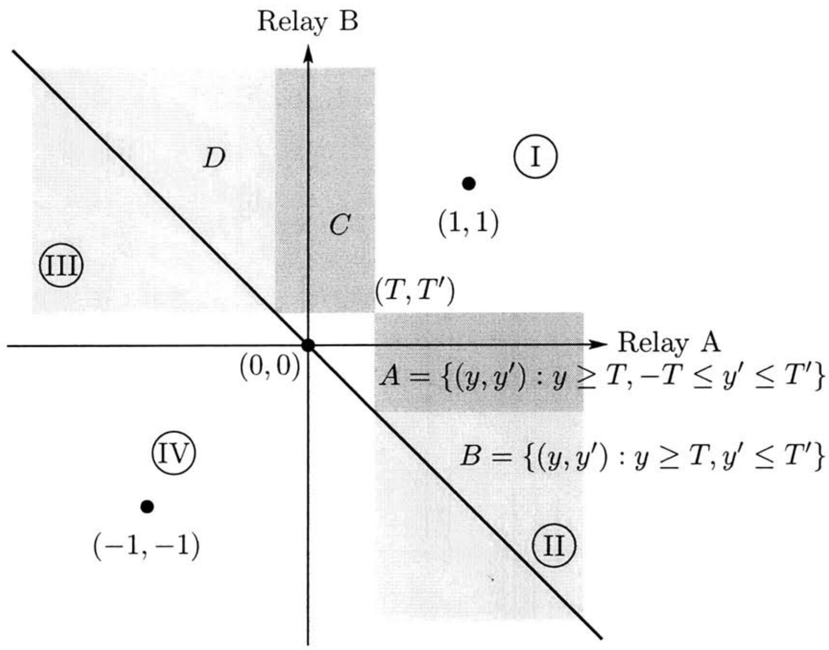


Figure 1-2: Arbitrary $(T,T^{'})$ relay decision threshold on the observation space for $T+T^{'} > 0$. Fusion center receives 1 bit from each relay which specifies the 4 regions: I= $\{(y,y') : y \geq T \text{ and } y^{'} \geq T^{'}\}$, II=$\{(y,y') : y \geq T \text{ and } y^{'} \leq T^{'}\}$, III= $\{(y,y') : y \leq T \text{ and } y^{'} \geq T^{'}\}$, IV= $\{(y,y') : y \leq T \text{ and } y^{'} \leq T^{'}\}$. The "don't care" regions of whether a 1 or -1 is assigned to it are B and D. Region A and C should be assigned to "1."

Figure 1-4 shows the error trade-offs among the different schemes. First, compare the optimal solution at $\vec{T} = (T,T)$ with the suboptimal solution at $\vec{T} = (0,0)$. Since $(-1,+1)$ and $(+1,-1)$ are "Don't Care" decoding regions for $\vec{T} = (0,0)$, choose the same decoder for $\vec{T} = (0,0)$ as for

$\vec{T} = (T^*, T^*)$ The small increase in error in region A is more than offset by an error decrease in regions B and C. Note that if T is small, the increase in error probability in region A is proportional to T^2, whereas the decrease in error probability in regions B and C is proportional to T. Therefore, for small enough T, the solution at (T, T) has a smaller error probability than that at $(0, 0)$. In Chapter 3, we will show in detail why the two relays must use the same strategy. Solving for equation 1.2, we find there exists a unique point (T^*, T^*) which gives the optimal solution.

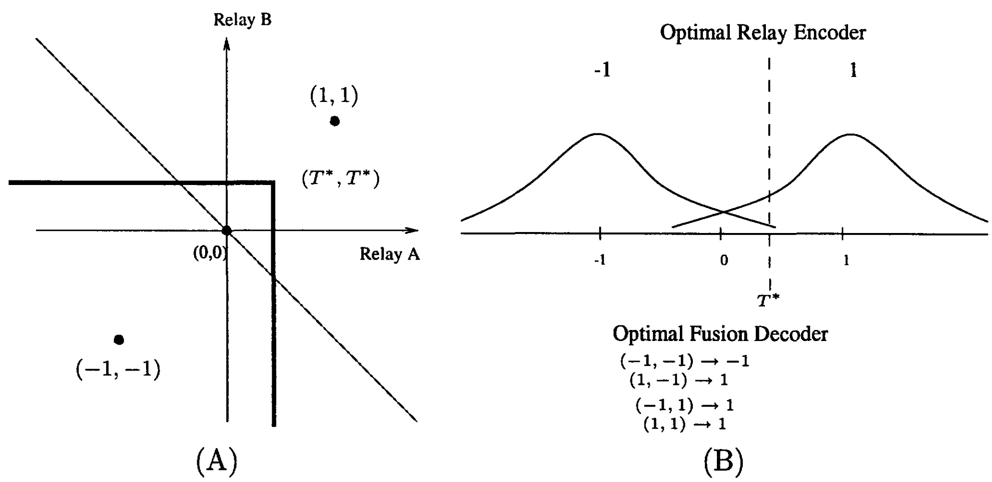


Figure 1-3: (A) The optimal solution at $\vec{T} = (T^*, T^*)$ has the decision boundary shown by the solid line. Points to the left and below the boundary map to "-1" and pints to the right and above the boundary map to "1." (B) Corresponding relay encoder and optimal fusion point.

This simple example shows that these types of distributed problems are neither trivial nor obvious. The complete results for the additive Gaussian noise channel will be established in Chapter 3.

As the number of relays increases from two to L, the probability of error P_e decreases exponentially in L for any reasonable relay quantizer and decision mapping. Continuing with our toy problem, we increase the number of relays. The nontriviality of the problem is further apparent in the case of three relays for the AGN distributed system. Suppose the noise variance is one and ML decoding is used at the fusion center. Then the error probabilities turn out to be within fractions of a percent of each other for the relay observation thresholds at $\vec{T}=(T, T', T'')=(0, 0, 0)$ with majority rule decoding and $(T, T', T'') = (0.2, 0.2, -0.8)$ using the decoder $\{000, 010, 001\} \to 0$, $\{100, 101, 011, 111\} \to 1$. The optimal result for the 3 relay case, which uses a threshold $\vec{T}=(0, 0, 0)$ and majority rule decoding, will be proved with more insight in a future paper.

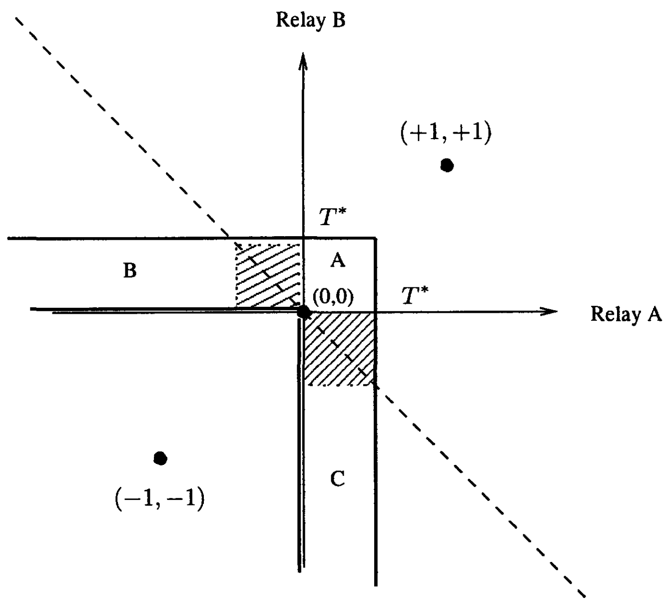


Figure 1-4: Tradeoffs in error by threshold at (T^*, T^*) compared to the obvious but suboptimal quantization at $(T, T') = (0,0)$. Shaded regions are the "don't care" regions. Region A $= \{0 \le \alpha, \alpha' \le T\}$, Region B $= \{\alpha \le -T^*$ and $0 \le \alpha' \le T^*\}$, Region C $= \{0 \le \alpha \le T$ and $\alpha' \le -T\}$. The probability of error in Region A is $P_e(A) \approx T^2$ and Region B and C are $P_e(B) = P_e(C) \approx \mathcal{Q}(0)\,T$. Thus, for small T the gain in decreased error (proportional to T) from regions B and C is worth the cost in increased error (proportional to T^2) in region A.

1.2 Previous Work and Background

The problem studied in this thesis is a sub-class of a classical problem in decentralized systems which was first posed by Tenney and Sandell in [TNS84]. The literature in the field is abundant since numerous people have worked on a number of variants of the distributed system problem. An excellent and thorough survey article [Tsi93] details the major contributions in this field of research. The book [Var97] and overview articles [VV97], [RB97] give a history and also discuss a number of variations in decentralized detection and coding.

Although more complicated networks have been addressed, parallel and serial topologies are generally considered. Almost across the board, the Bayesian and Neyman-Pearson approach is implemented. Conditional independence of sensor observations given the source is generally assumed, because the optimal solution is generally intractable otherwise. In [TA84] and [TA85], the authors proved that such problems are non-polynomial complete.

The asymptotic analysis of the system has already been well established by Tsitsikis in [Tsi88]. The complete asymptotic result and computational complexity as the number of relays tend to infinity is established. In general, any reasonable rule will drive the error probability down exponentially in the number of relays. Tsitsikis further shows that for M hypothesis, the maximum number of distinct relay rules needed is $M(M-1)/2$. Thus, for binary hypothesis, one single relay rule for each relay is sufficient for optimality as the number of relays tend to infinity.

Optimality of likelihood ratio tests at the relays and fusion point is well known. In our analysis, we will give a treatment and proof from a different viewpoint for added clarity. The problem with binary hypothesis and conditional independence assumptions (which is studied in this thesis) is the most studied in the literature. Various characteristics of the optimal solution for the binary detection problem and especially the specialized case of the additive Gaussian noise channel has been given in [IT94], [QZ02], [WS01].

[Tsi93], [PA95], and [WW90] discuss randomized strategies. They recognized the non-convexity of the optimal non-randomized strategies, so dependent randomization was necessary for some cases. We will state precisely and supply the necessary conditions for the different types of randomization in our treatment.

[TB96] considered encoding long sequences of the source and looking at the rate-distortion function and [Ooh98] gave the complete solution to the specialized case for Gaussian sources.

We by no means claim to give an exhaustive list of the field.

1.3 Outline of Thesis

Ad-hoc networks with a large numbers of nodes must use intermediate nodes to relay information between source and destination. This is a highly complicated problem. At the moment, there exists no structure to analyze the major issues in a fundamental way. This thesis attempts to understand and gain insight into a sub problem of this general problem of wireless networks.

The thesis is organized as follows. Chapter 2 describes and characterizes the Neyman-Pearson error curve for arbitrary distributions and how it is utilized as a fundamental tool in analyzing the overall distributed detection and coding problem with relays. We prove some general results for the purely discrete case which will be used in later chapters. In Chapter 3, we study the joint optimization of relay encoder and fusion decoder. For the case of two relays, each sending a single bit to the fusion center, this will aid in the understanding of the optimal solution and its structure. We provide a complete solution to the Gaussian case and also the optimal solution for arbitrary channels. Chapter 4 establishes the complete optimal solution and provides fundamental insights at the encoder and decoder for the case of 2 relays with non-binary outputs. Finally, chapter 5 concludes the thesis.

Chapter 2

The Error Curve

A significant part of this thesis deals with the detection of a binary digit through relays. Without relays, there are several approaches to the classical binary detection problem. In the maximum likelihood (ML) problem, we would like to minimize the overall error probability given equally likely inputs. More generally, in the maximum a posteriori (MAP) problem, we would like to minimize error probability for an arbitrary known input distribution. Even more generally, the minimum cost problem deals with minimizing overall cost given an arbitrary known input distribution and a cost function associated with each type of error. A final example of interest, known as the Neyman-Pearson problem , is to minimize the error probability of one type given a maximal tolerable error probability of the other type. In the classical problem without relays, the solutions to all these problems can be summarized by the *error curve*[1]. When relays are introduced into the system, we will find that the solution to all the above problems can be found from appropriately combining the error curves at the relays with an associated error curve at the fusion point.

This chapter will describe the *Error Curve*, its properties, and how it is used as a tool to both characterize the relays and the fusion process. The relay channels are described by error curves, and the final error curve at the fusion point is a function of the relay error curves.

[1]A slight variant, better known in the radar field, is called the receiver operating characteristic (ROC).

2.1 Binary Detection and the Neyman-Pearson Problem

Binary detection for both discrete-alphabet and continuous-alphabet channels is characterized by the error curves. We start by reviewing binary detection without relays. We can view this in terms of a channel with binary input X and arbitrary output Y. Let the source a priori distribution be $\Pr(X{=}0){=}p_0$ and $\Pr(X{=}1){=}p_1$. Further, assume for now that $f_{Y|X}(Y|0)$ and $f_{Y|X}(Y|1)$ are densities. The discrete channel will be dealt with later in section 2.4. The likelihood ratio (LR) is defined for the set Y to be

$$\lambda(Y) = \frac{f_{Y|X}\{Y\,|\,0\}}{f_{Y|X}\{Y\,|\,1\}} \tag{2.1}$$

For the values of y that are possible under $X{=}1$ but impossible under $X{=}0$, the likelihood ratio is $\lambda(Y) = 0$. Likewise, for the values of y that are possible under $X{=}0$ but impossible under $X{=}1$, $\lambda(Y)$ is infinite. We assume in this section that $\Pr(\lambda(Y) \leq \eta | X = 0)$ and $\Pr(\lambda(Y) \leq \eta | X = 1)$ and $\Pr(\lambda(Y) \leq \eta)$ are continuous functions of η. It follows that $\Pr(\lambda(Y) \leq \eta | X = i) = 0$ for all η and for $i = 0, 1$.

Define a *test*, $t(Y) \to \{0, 1\}$, as a function which maps each output Y to 1 or 0.

Define a *randomized test* over a finite set of tests $\{t_1, t_2, \ldots, t_n\}$ as a mapping

$$t_{\vec{\psi}}(Y) = t_i(Y) \quad \text{with probability} \quad \psi_i \tag{2.2}$$

where $\vec{\psi} = (\psi_1, \psi_2, \ldots, \psi_n)$ is the probability vector with which the tests $\{t_1, t_2, \ldots, t_n\}$ are used.

Define a *threshold test* at η as a mapping

$$t^{(\eta)}(Y) = \begin{cases} 1 & \text{if} \quad \lambda(Y) < \eta \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

A MAP test is any test which minimizes error probability for given a priori probabilities. The following lemma shows that the MAP test is a threshold test.

Lemma 2.1.1 *A threshold test at η is a MAP test for a priori probabilities p_0 and p_1 where $\eta = \frac{p_1}{p_0}$.*

Proof: A MAP test decides $\hat{X} = 1$ if $\Pr(X = 1 \mid Y = y) > \Pr(X = 0 \mid Y = y)$ and decides $\hat{X} = 0$ if $\Pr(X = 1 \mid Y = y) \leq \Pr(X = 0 \mid Y = y)$. By Bayes Rule, this is the same as deciding $\hat{X} = 1$ if $\lambda(Y) < \eta$ and $\hat{X} = 0$ if $\lambda(Y) > \eta$. More compactly, for the continuous case,

$$\lambda(Y) \underset{\hat{X}=1}{\overset{\hat{X}=0}{\gtrless}} \eta \tag{2.4}$$

where by the assumption above, $\lambda(Y)$ has no point masses. □

Having cost functions associated with the different error probabilities is a more general problem but does not change the core of the problem. The optimal test for arbitrary cost is simply another threshold test. We show that changing the cost function and a priori probabilities in the system does not alter the fundamental problem. Let C_{ij} be the cost of choosing hypothesis i given hypothesis j has occurred. Let p_1 and p_0 be the a priori probabilities, so $\eta = \frac{p_1}{p_0}$. The total cost becomes

$$\begin{aligned} C &= p_1P(0|1)C_{01} + p_0P(1|0)C_{10} + p_0P(0|0)C_{00} + p_1P(1|1)C_{11} \\ &= P(0|1)[p_1C_{01} - p_0C_{00}] + P(1|0)[p_0C_{10} - p_1C_{11}] + p_0C_{00} + p_1C_{11} \\ &= p_1^{'}P(0|1) + p_0^{'}P(1|0) + C^{'} \end{aligned} \tag{2.5}$$

where $p_1^{'} = p_1C_{01} - p_0C_{00}$ and $p_0^{'} = p_0C_{10} - p_1C_{11}$ and $C^{'} = p_0C_{00} + p_1C_{11}$. The overall cost is minimized by a threshold test on the LR with threshold η_c where

$$\eta_c = \frac{p_1^{'}}{p_0^{'}} = \frac{p_1C_{01} - p_0C_{00}}{p_0C_{10} - p_1C_{11}}. \tag{2.6}$$

Thus, for any arbitrary cost, the problem is equivalent to solving the MAP problem with the modified a priori probabilities given in equation 2.6. If the MAP problem is solved for all apriori probabilities, it also solves the arbitrary cost problem. That is, the minimum cost problem is a trivial extension of the MAP problem. In what follows we will solve the arbitrary MAP problem.

For any given η, let $\tilde{\alpha}(\eta) = \Pr(e \mid X = 1)$ and $\tilde{\beta}(\eta) = \Pr(e \mid X = 0)$ for the MAP test, i.e., the threshold test at η. Since $\Pr\{\lambda(Y) \leq \eta \mid X{=}0\}$ is continuous and increasing in η, and $\Pr\{\lambda(Y) \geq \eta \mid X{=}1\}$ is continuous and decreasing, we see that $\tilde{\beta}(\eta)$ can be expressed as a function of $\tilde{\alpha}(\eta)$. Define the *error curve* by this pair of parametric functions of the threshold η, $0 < \eta < \infty$, called $\tilde{\alpha}(\eta) = \Pr\{\lambda(Y) \geq \eta \mid X{=}1\}$ and $\tilde{\beta}(\eta) = \Pr\{\lambda(Y) < \eta \mid X{=}0\}$ both of which are

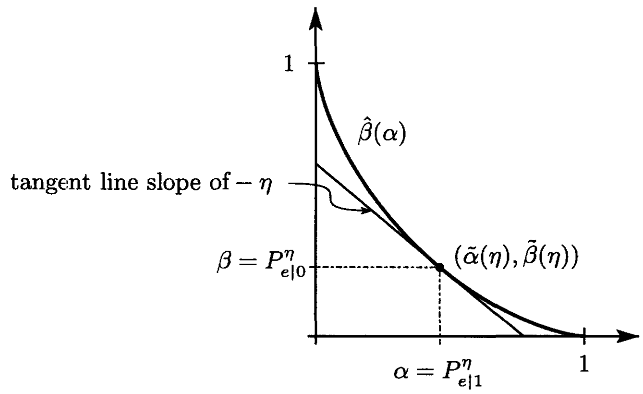


Figure 2-1: *The Error Curve.* The parametric form of the error curve is described by $(\tilde{\alpha}(\eta), \tilde{\beta}(\eta))$ and the direct form of the curve by $\beta = \hat{\beta}(\alpha)$. Each point (α, β) on the error curve results from a threshold test with threshold value $\eta = \frac{p_1}{p_0}$ which is the magnitude of the slope of its point of tangency. The overall error probability is the value of the abscissa-intercept multiplied by a priori probability p_0.

continuous functions of η. That is, $\Pr\{\lambda(Y) = \eta\}$ is zero for all η. Then the ordinate and abscissa of the error curve are respectively,

$$\tilde{\alpha}(\eta) = \Pr\{e \,|\, X{=}1\} = \Pr\{\lambda(Y){\geq}\eta \,|\, X{=}1\} \tag{2.7}$$

$$\tilde{\beta}(\eta) = \Pr\{e \,|\, X{=}0\} = \Pr\{\lambda(Y){<}\eta \,|\, X{=}0\} = \hat{\beta}(\alpha) \tag{2.8}$$

As η increases, $\tilde{\beta}(\eta)$ increases and $\tilde{\alpha}(\eta)$ decreases. The error curve is a graph plotting $\tilde{\beta}(\eta)$ on the ordinate and $\tilde{\alpha}(\eta)$ on the abscissa as η goes from 0 to ∞. An example of an error curve is shown in figure 2-1. The error curve is later shown to be convex. Note that $\tilde{\alpha}(\eta)$ is a monotonically decreasing function and $\tilde{\beta}(\eta)$ is a monotonically increasing function of η. Later we will treat the situation in which $\tilde{\alpha}(\eta)$ and $\tilde{\beta}(\eta)$ are discontinuous. We call sections of the error curve which contain discontinuous $\tilde{\alpha}(\eta)$ and $\tilde{\beta}(\eta)$ the "discrete" part of the error curve and treat it in a subsequent section. Define $\tilde{\alpha}(0){=}\lim_{\eta\to 0}\tilde{\alpha}(\eta)$ and $\tilde{\alpha}(\infty){=}\lim_{\eta\to\infty}\tilde{\alpha}(\eta)$ and $\tilde{\beta}(0){=}\lim_{\eta\to 0}\tilde{\beta}(\eta)$ and $\tilde{\beta}(\infty){=}\lim_{\eta\to\infty}\tilde{\beta}(\eta)$. When $\tilde{\alpha}(0) < 1$, there is some event with non-zero probability under $X = 1$ and with probability zero under $X = 0$. Likewise, when $\tilde{\beta}(\infty) < 1$, there is some event with probability greater than zero under $X = 0$ and with probability zero under $X = 1$. We call these types of curves *erasure-like* error curves.

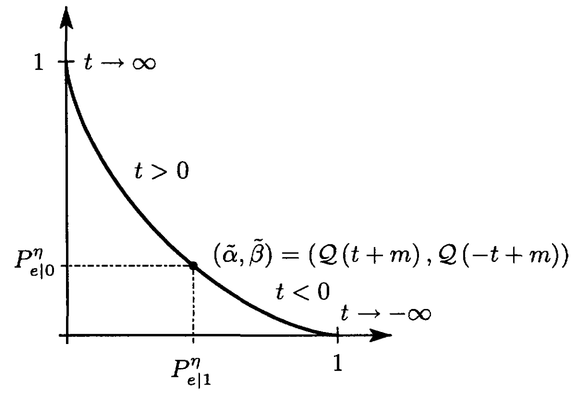


Figure 2-2: The error curve for the additive Gaussian noise channel. The parameters t and m are normalized.

2.2 The Gaussian Error Curve

Let a relay observe a binary signal at signal level m or $-m$ independently corrupted by Gaussian noise, $\mathcal{N}(0, \sigma)$. Since $Y = X + N$, then $Y = m + N$ if $X = 0$ and $Y = -m + N$ under $X = 1$. We now describe the error curve for this additive Gaussian noise channel which we call the Gaussian error curve.

The Gaussian error curve is an example of an error curve which is strictly convex and everywhere differentiable. Thus, the function $\tilde{\alpha}(\eta)$ is strictly decreasing and $\tilde{\beta}(\eta)$ strictly increasing . It is easier to represent the Gaussian error curve in parametric form, $(\alpha, \beta) = (\mathcal{Q}\left(\frac{y+m}{\sigma}\right), \mathcal{Q}\left(\frac{y-m}{\sigma}\right))$ for $m \neq 0$ and $\sigma \neq 0$ and $y \in \Re$. The LR is $\lambda(Y) = \exp\left(\frac{2my}{\sigma}\right)$. Gaussian channels can be normalized and are completely characterized by one parameter, either the signal level m or variance σ^2. To see this, let $t = \frac{y}{\sigma}$; then $(\alpha, \beta) = (\mathcal{Q}\left(t + \frac{m}{\sigma}\right), \mathcal{Q}\left(-t + \frac{m}{\sigma}\right)$ for $t \in \Re$. All that matters in the curve is the signal to noise ratio $\frac{m}{\sigma}$. Thus, without loss of generality, we can assume $\sigma = 1$ and let the mean m characterize the different Gaussian error curves. An example of a Gaussian error curve is illustrated in figure 2-2.

Furthermore, any single point (α^*, β^*) on an error curve in the region $\alpha^* + \beta^* < 1$ for $\alpha^* > 0$ and $\beta^* > 0$ completely specifies the Gaussian error curve. For any fixed point (α^*, β^*), there exists a unique t and m such that $\mathcal{Q}(t+m) = \alpha^*$ and $\mathcal{Q}(-t+m) = \beta^*$. Thus, one point completely determines the entire Gaussian error curve. From this point on, all Gaussian channels will be normalized with $\sigma = 1$, mean m, and the error curve described parametrically with parameter t by $(\tilde{\alpha}(t), \tilde{\beta}(t)) = (\mathcal{Q}(t+m), \mathcal{Q}(-t+m))$, $t \in \Re$ and $m \neq 0$. Additionally, observe that the LR is monotonic with the observation parameter y. Henceforth, for the rest of the thesis, whenever the Gaussian case is discussed, rather than give solutions as thresholds on the likelihood ratio $\lambda(Y)$, we will represent solutions as observation thresholds of the parameter t for the normalized Gaussian error curve.

2.3 The η-tangent Line

We have seen that the error curve was generated from the MAP test over all threshold values. In this section, we will show that the value of the threshold η is the magnitude of the slope of the tangent line to its respective point on the error curve. This result is known but we give a new derivation and explanation. This simpler derivation will be useful in the distributed case later. Denote the line of slope $-\eta$ through the point at $(\tilde{\alpha}(\eta), \tilde{\beta}(\eta))$ as the η-tangent line as seen in figure 2-1. We now show that every η-tangent line for an error curve lies on or below the entire error curve. We have seen that for given a priori probabilities p_0 and p_1, the threshold test at $\eta = \frac{p_1}{p_0}$ minimizes error probability. Thus, for any arbitrary test x, it must be that

$$p_0\tilde{\beta}(\eta) + p_1\tilde{\alpha}(\eta) \leq p_0\beta(x) + p_1\alpha(x) \tag{2.9}$$

$$\tilde{\beta}(\eta) + \eta\tilde{\alpha}(\eta) \leq \beta(x) + \eta\alpha(x) \tag{2.10}$$

$$\eta(\tilde{\alpha}(\eta) - \alpha(x)) \leq \beta(x) - \tilde{\beta}(\eta) \tag{2.11}$$

Suppose $\alpha(x) < \tilde{\alpha}(\eta)$, then

$$\eta \leq \frac{\beta(x) - \tilde{\beta}(\eta)}{\tilde{\alpha}(\eta) - \alpha(x)} \tag{2.12}$$

The quantity on the right of 2.12 is the magnitude of the slope of the line connecting $(\tilde{\alpha}(\eta), \tilde{\beta}(\eta))$ and $(\alpha(x), \beta(x))$, so the point $(\alpha(x), \beta(x))$ must lie above the η-tangent line. For $\alpha(x) > \alpha(\eta)$, the inequality in 2.12 is reversed, again showing $(\alpha(x), \beta(x))$ lies above the η-tangent line. Thus, all the η-tangent lines lie on or below the error curve. Furthermore, by the same argument, all the η-tangent lines lie below the point $(\tilde{\alpha}(\eta), \tilde{\beta}(\eta))$ for any test or randomized test. Finally, this argument also shows that all points of the error curve lie on the convex envelope[2] of these tangent lines. Thus, every error curve is convex.

[2]Define a *convex combination* of elements of a set $X \subset \Re^2$ as a linear combination of the form

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n$$

for some $n > 0$, where each $x_i \in X$, each $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$. Define the *convex hull*, denoted $co(X)$, to be the smallest convex set which contains X. Define the *convex envelope* as the curve which is the lower boundary of $co(X)$.

As verification for the case where $\tilde{\alpha}(\eta)$ and $\tilde{\beta}(\eta)$ are differentiable, we show that the line of slope $-\eta$ through the point $(\alpha, \beta) = (\tilde{\alpha}(\eta), \tilde{\beta}(\eta))$ is tangent to the error curve at $(\tilde{\alpha}(\eta), \tilde{\beta}(\eta))$. The overall error probability for a given p_0, p_1 and arbitrary threshold x is $q(x) = p_0\tilde{\beta}(x) + p_1\tilde{\alpha}(x)$. For $\eta = \frac{p_1}{p_0}$, the optimal test which minimizes $q(x)$ is the one where $x = \eta$. This implies that if the error curve is differentiable, then $x = \eta$ must be a stationary point of $q(x)$, so $\frac{dq(x)}{dx}|_{x=\eta} = 0$, which means that at $x = \eta$,

$$p_0 \frac{d\tilde{\beta}(x)}{dx} + p_1 \frac{d\tilde{\alpha}(x)}{dx} = 0 \tag{2.13}$$

$$\frac{d\tilde{\beta}(x)}{d\tilde{\alpha}(x)} = \frac{d\tilde{\beta}(x)/dx}{d\tilde{\alpha}(x)/dx} = -\frac{p_1}{p_0} = -\eta \quad \text{at} \quad x = \eta \tag{2.14}$$

Since the error curve is convex, this shows that the line of slope $-\eta$ through the point $(\tilde{\alpha}(\eta), \tilde{\beta}(\eta))$ lies on or below the error curve. Note that any error curve which lies strictly below another error curve will have a lower overall error probability for the MAP test at every η.

so it is called the η-tangent line. Note that the β-intercept of the η-tangent line is equal to $\beta + \eta\alpha$. The overall error probability is the value of this intercept multiplied by p_0. By virtue of the MAP test for a given ratio of a priori probabilities $\eta = \frac{p_1}{p_0}$, the tangent line to the error curve at $(\tilde{\alpha}(\eta), \tilde{\beta}(\eta))$ must have the lowest β-intercept since it is the test which minimizes the overall error probability. Therefore, the magnitude of the slope of the tangent line to $(\tilde{\alpha}(\eta), \tilde{\beta}(\eta))$ must equal η. Thus, it is called the η-tangent line. This gives another way of seeing that each η-tangent line must lie at or below the entire error curve.

2.4 The Discrete Error Curve

We have seen that when $\tilde{\alpha}(\eta)$ and $\tilde{\beta}(\eta)$ are continuous, every point on the error curve is by definition achievable by a MAP test at a specific η. The error curve is strictly convex for $\tilde{\alpha}(\eta)$ and $\tilde{\beta}(\eta)$ continuous, but not necessarily differentiable[3] at every point. This section treats the purely discrete case where $\tilde{\alpha}(\eta)$ is not continuous and the set of possible likelihood ratios is finite. That is, there is a finite set of y for which $\lambda(Y) = \eta$ has positive probability under $X = 0$ and $X = 1$. The MAP test in equation 2.4 gives the same error probability when $\lambda(Y) = \eta$ whether the decision is $\hat{X} = 0$ or $\hat{X} = 1$. However, this MAP test is no longer adequate for the Neyman-Pearson problem, since the choice of $\hat{X} = 0$ or $\hat{X} = 1$ results in different α and β. When $\lambda(Y) = \eta$ in equation 2.4, there is an ambiguity of which way to

[3]the likelihood density is not assumed to be continuous

choose the inequality sign. We have the situation of a "don't care" choice for MAP but a "do care" choice for Neyman-Pearson.

First, we take care of the ambiguous sign problem in 2.4 by defining two different MAP tests for discrete observations as shown in figure 2-3.

Define the threshold test "MAP Left" as

$$\lambda(Y) = \frac{f_{Y|X}\{y|\,0\}}{f_{Y|X}\{y|\,1\}} \underset{\hat{X}=1}{\overset{\hat{X}=0}{\gtrless}} \eta \tag{2.15}$$

and threshold test "MAP Right" as

$$\lambda(Y) = \frac{f_{Y|X}\{y|\,0\}}{f_{Y|X}\{y|\,1\}} \underset{\hat{X}=1}{\overset{\hat{X}=0}{\begin{smallmatrix}\geq\\<\end{smallmatrix}}} \eta \tag{2.16}$$

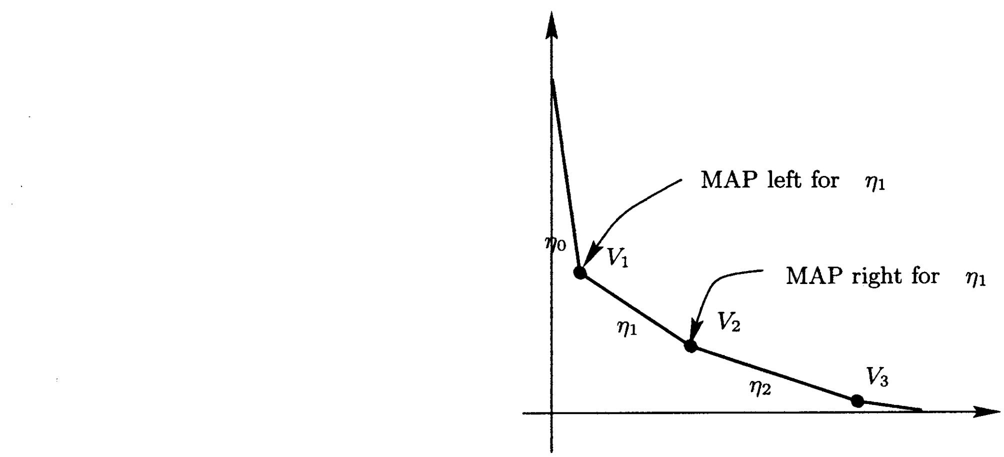


Figure 2-3: The Error Curve for the discrete case. The MAP Left test for η_1 generates the point V_1 which is also the MAP Right test for η_0; the MAP Right for η_1 generates the same point as the MAP Left test for η_2; and so forth. The derivative at a vertex point is well-defined by an interval and the values are the thresholds which determine the vertex point. For the error curve in the figure, $-\frac{d\hat{\beta}(\alpha)}{d\alpha}|_{\alpha_1} = [\eta^-_{\alpha_1}, \eta^+_{\alpha_1}] = [\eta_1, \eta_0]$, $-\frac{d\hat{\beta}(\alpha)}{d\alpha}|_{\alpha_2} = [\eta^-_{\alpha_2}, \eta^+_{\alpha_2}] = [\eta_2, \eta_1]$, and so forth.

Now, to solve the Neyman-Pearson problem, the trick is to randomize between the right and left MAP tests which give the same error probability. Then the desired α can be achieved while still retaining the same overall error probability. We will show that this is always possible. In other words, for a constraint on one error type, the solution is either a MAP test defined in equations 2.15 and 2.16, or a randomized solution between two MAP tests. This brings on the notion of a *randomized threshold test.*

Define a *randomized threshold test* at η and ψ as a mapping

$$t_{\psi}^{(\eta)}(y) = \begin{cases} 1 \ \text{ if } \ \lambda(Y) < \eta \\ 0 \ \text{ if } \ \lambda(Y) > \eta \\ 0 \ \text{ with probability } \ \psi \ \text{ if } \ \lambda(Y) = \eta \\ 1 \ \text{ with probability } \ (1-\psi) \ \text{ if } \ \lambda(Y) = \eta \end{cases} \tag{2.17}$$

Any test or randomized test is characterized by 2 error probabilities, denoted by $\alpha = \Pr\{e \mid X{=}0\}$ and $\beta = \Pr\{e \mid X{=}1\}$. For a known noise distribution, but unknown a priori probabilities with maximum acceptable error α of one kind, say $\Pr\{e \mid X{=}1\} \leq \alpha$, the *Neyman-Pearson test* is defined as a test or randomized test that minimizes the error of the other kind, namely $\Pr\{e \mid X{=}0\} = \beta$. The Neyman-Pearson test is defined independent of any a priori probabilities and is a randomized threshold test on the likelihood ratio $\lambda(Y)$.

We can now define a general error curve for channels with arbitrary LR's. An error point is any point on the error curve. Define a *vertex point* as an error point generated by a MAP left or MAP right test as in (2.15) or (2.16). Therefore, a vertex point is a point on the error curve which cannot be expressed as a convex combination of other points on the error curve. Vertex points are generated from non-randomized threshold tests; each is the unique point of tangency for some tangent line. The endpoints of linear segment are vertex points. All error points on a linear segment which are not endpoints are non-vertex points and generated by randomized threshold tests. These points will sometimes be referred to as randomized points. Thus, vertex points are a subset of the error points on the error curve.

First, generate all the vertex points with the MAP tests defined by (2.15) or (2.16). Then take the convex envelope of the vertex points by using randomized threshold tests between the adjacent vertex points. The resulting *error curve* is denoted by $\beta = \hat{\beta}(\alpha)$ for $0 \leq \alpha, \beta \leq 1$, where for any test or randomized test with error probabilities α and β, we have $\hat{\beta}(\alpha) \leq \beta$. The error curve gives the solution to the MAP test as well as the solution to the Neyman-Pearson problem. We call $\beta = \hat{\beta}(\alpha)$ the direct form of the error curve and an example is shown in figure 2-1. Note that the set of vertex points is the same for both MAP rules defined in 2.15 and 2.16 and therefore, they both generated the same error curve.

The supremum of the η-tangent lines is the *error curve* as already defined. As discussed previously, every η-tangent must lie on or beneath all points on the error curve. Therefore, the error curve is the convex envelope of the vertex points. We now show that every point on the error curve is achievable via a randomized threshold test or threshold test. It is sufficient to show achievability for just one linear segment of the error curve. Let $V_1 = (\alpha_1, \beta_1)$ and $V_2 = (\alpha_2, \beta_2)$ be MAP left and right vertex points on the error curve such that the segment of slope $-\eta = \frac{\beta_2 - \beta_1}{\alpha_2 - \alpha_1}$ passing through them is on the error curve. We show that every (α^*, β^*) which lies on this linear segment is achievable. Without loss of generality, let $\alpha_1 < \alpha_2$. Then $\alpha_1 < \alpha^* < \alpha_2$. Consider the strategy of choosing threshold test V_2 with probability $\frac{\alpha^* - \alpha_1}{\alpha_2 - \alpha_1}$ and threshold test V_1 with probability $\frac{\alpha_2 - \alpha^*}{\alpha_2 - \alpha_1}$. Then,

$$\Pr\{e \,|\, X{=}1\} = \alpha_2 \left(\frac{\alpha^* - \alpha_1}{\alpha_2 - \alpha_1} \right) + \alpha_1 \left(\frac{\alpha_2 - \alpha^*}{\alpha_2 - \alpha_1} \right) = \alpha^* \tag{2.18}$$

and similarly, since $\frac{\alpha^* - \alpha_1}{\alpha_2 - \alpha_1} = \frac{\beta^* - \beta_1}{\beta_2 - \beta_1}$ and $\frac{\alpha_2 - \alpha^*}{\alpha_2 - \alpha_1} = \frac{-\beta_2 - \beta^*}{\beta_2 - \beta_1}$, we obtain $\Pr\{e \,|\, X{=}0\} = \beta^*$. Thus, (α^*, β^*) is achievable via a randomized threshold test or time sharing between the threshold tests V_1 and V_2. This says that every point on the error curve is achievable. Note that this error curve still corresponds to a randomized threshold test in the sense that the decision is $\hat{X} = 0$ for $\lambda(Y) > \eta$ and $\hat{X} = 1$ for $\lambda(Y) < \eta$. The randomized decision occurs for $\lambda(Y) = \eta$, which is a "don't care" situation for the MAP test corresponding to η. It is this error curve that solves the Neyman-Pearson problem. Lastly, note that the error curve has at most one straight line portion of slope $\lambda(Y)$ for each output y, although, as we see later, several different choices of y might have the same LR $\lambda(Y)$.

Example: The BAC and its Error Curve. If the relay channel (the channel from source to relay input) is discrete and finite, then there is a finite set of possible likelihood ratios, so the non-randomized threshold tests consist of a finite set of single points each corresponding to one of the outputs as given in equations 2.7 and 2.8. We illustrate the error curve for discrete observations via the binary asymmetric channel (BAC) in figure 2-4(A) where $\Pr(e|0) = \epsilon_0$ and $\Pr(e|1) = \epsilon_1$. Without loss of generality, assume that $\epsilon_0 + \epsilon_1 \leq 1$ for the BAC, since the outputs could otherwise be reversed. The relay observes one of two possible values, so the likelihood ratio $\lambda(Y)$ takes on only two values, $\lambda(0) = \frac{1-\epsilon_0}{\epsilon_1}$ and $\lambda(1) = \frac{\epsilon_0}{1-\epsilon_1}$. This implies that the threshold test as given by (2.7) and (2.8) is constant over three regions of η and jumps when η equals $\lambda(0)$ and $\lambda(1)$. That is, this threshold test is constant for $0 \leq \eta < \frac{\epsilon_0}{1-\epsilon_1}$ and chooses $\hat{X}{=}0$ for all y. In this case $\tilde{\alpha}(\eta){=}1$ and $\tilde{\beta}(\eta) = 0$. Similarly, this threshold test is fixed for $\frac{\epsilon_0}{1-\epsilon_1} < \eta < \frac{1-\epsilon_0}{\epsilon_1}$ and chooses $\hat{X}{=}0$ if $y = 0$ and $\hat{X}{=}1$ if $y = 1$. At this point, $\tilde{\alpha}(\eta) = \epsilon_1$ and $\tilde{\beta}(\eta) = \epsilon_0$. Finally, the test is fixed for $\eta \geq \frac{1-\epsilon_0}{\epsilon_1}$ and chooses $\hat{X}{=}1$ for all y, which gives

$\tilde{\alpha}(\eta)$=0 and $\tilde{\beta}(\eta)$=1. Thus the vertex points resulting from the threshold test (2.7) and (2.8) consists of only three points $\{(1,0),(\epsilon_1,\epsilon_0),(0,1)\}$. The error curve is the piecewise linear curve connecting the 3 vertex points.

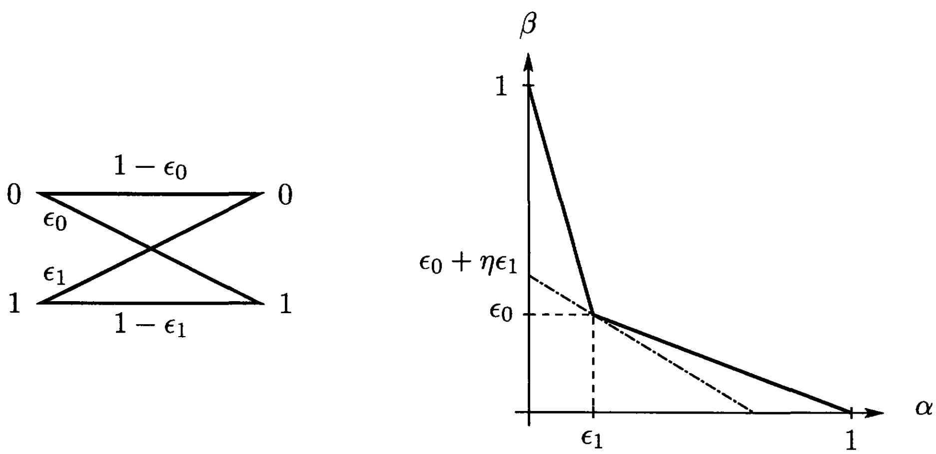


Figure 2-4: The Binary Asymmetric Channel (BAC) and its error curve.

From the perspective of the η-tangent line, for each threshold η, consider the tangent of slope $-\eta$ passing through $(\tilde{\alpha}(\eta),\tilde{\beta}(\eta))$. For example, lines of slope $-\eta$, for $0<\eta<\frac{\epsilon_0}{1-\epsilon_1}$, passing through $(\tilde{\alpha}(\eta),\tilde{\beta}(\eta))$, rotate about the point $(1,0)$. When $\eta=\frac{\epsilon_0}{1-\epsilon_1}$, the η-tangent line passes through $(1,0)$ and (ϵ_1,ϵ_0). Then as η continues to increase, the η-tangent line rotates about the point (ϵ_1,ϵ_0), until $\eta=\frac{1-\epsilon_0}{\epsilon_1}$. Then for $\eta>\frac{1-\epsilon_0}{\epsilon_1}$, it rotates about $(0,1)$.

The binary symmetric channel (BSC) and Z channels are special cases of the binary asymmetric channel (BAC), the former letting $\epsilon_0=\epsilon_1$ and the latter letting either $\epsilon_0=0$ or $\epsilon_1=0$. The same interpretation used before for the randomized threshold test applies to the error curves of these channels.

Define the left derivative of the error curve at α as the infimum of the slope of the η-tangent lines to α, i.e., the derivative approaching α from the right,
$\eta_\alpha^- = -\lim_{\alpha^*\to\alpha}\frac{\hat{\beta}(\alpha^*)-\hat{\beta}(\alpha)}{\alpha-\alpha^*}$ for $0\le\alpha^*\le\alpha\le 1$. Similarily, the right derivative at α is defined as the supremum of the slope of the η-tangent lines to α, i.e., the derivative approaching α from the left. That is, $\eta_\alpha^+ = -\lim_{a^*\to\alpha}\frac{\hat{\beta}(a^*)-\hat{\beta}(\alpha)}{a-\alpha^*}$ for $0\le\alpha\le\alpha^*\le 1$. If $\eta_\alpha^- = \eta_\alpha^+$, then the derivative of $\hat{\beta}(\alpha)$ exists at α in the ordinary sense, and this value is equal to the slope of the η-tangent at α. If $\eta_\alpha^-\neq\eta_\alpha^+$, then for all η such that $\eta_\alpha^-\le\eta\le\eta_\alpha^+$, the η-tangent to the error curve $\hat{\beta}(\alpha)$ is at α. In other words, if $\tilde{\alpha}(\eta)$ has a unique solution for a particular η, then that $\eta-$tangent line intersects the error curve at exactly one point and $\tilde{\alpha}(\eta)$ is a vertex point. If there exists more than one solution for a particular η, then the $\eta-$tangent line coincides with

a linear segment of the error curve with 2 vertex points at the endpoints.

We now summarize the relationship between (α, β) and η for every error curve. If the derivative exists at a point $(\tilde{\alpha}(\eta), \tilde{\beta}(\eta))$ on the error curve, then the slope of its tangent line $\frac{d\hat{\beta}(\alpha)}{d\alpha} = \frac{d\tilde{\beta}(\eta)/d\eta}{d\tilde{\alpha}(\eta)/d\eta} = -\eta$ as shown before. If the derivative $\frac{d\hat{\beta}(\alpha)}{d\alpha}$ does not exist in the ordinary sense at α, then for all $\eta \in [\eta_\alpha^-, \eta_\alpha^+]$, the η-tangent line intersects the error curve at α. For notational convenience, we say that $-\frac{d\hat{\beta}(\alpha)}{d\alpha} \in [\eta_\alpha^-, \eta_\alpha^+]$.

Lemma 2.4.1 *The left derivative and the right derivative exist for every point on the error curve. Furthermore, the set of non-differentiable points on $\hat{\beta}(\alpha)$ is at most countable. This is a consequence of the convexity of the function and not particular to error curves.*

Proof: Given any $s, t \in [0, 1]$ with $s < t$, let $p = \lambda s + (1 - \lambda)t$ for some λ, $0 \leq \lambda \leq 1$. The convexity of the error curve $\hat{\beta}(\alpha)$ states that $\hat{\beta}(p) \leq \lambda\hat{\beta}(s) + (1 - \lambda)\hat{\beta}(t)$. This implies that

$$\frac{\hat{\beta}(p) - \hat{\beta}(s)}{p - s} \leq \frac{\hat{\beta}(t) - \hat{\beta}(s)}{t - s} \leq \frac{\hat{\beta}(p) - \hat{\beta}(t)}{p - t} \quad \text{for} \quad s \leq p \leq t \tag{2.19}$$

It follows that $\forall s', s$ such that $s < s' < p$, we must have $\frac{\hat{\beta}(p)-\hat{\beta}(s)}{p-s} \leq \frac{\hat{\beta}(p)-\hat{\beta}(s')}{p-s'}$. Since $\lim\limits_{s' \to p} \frac{\hat{\beta}(s')-\hat{\beta}(p)}{s'-p}$ is monotonically increasing in s' and bounded above by any $t \geq p$ in equation 2.19, it must have a limit. This limit is the right derivative and its value is $-\eta_p^+$. Similarily, the left derivative is the limit approaching p from the right and its value is denoted by $-\eta_p^-$.

Now, we show that the set of non-differentiable points $D = \{\alpha \in [0, 1] : \eta_\alpha^- \neq \eta_\alpha^+\}$ is at most countable. For each α for which $\hat{\beta}(\alpha)$ is non-differentiable, the open interval $(\eta_\alpha^-, \eta_\alpha^+)$ is disjoint from $(\eta_{\alpha'}^-, \eta_{\alpha'}^+)$ for any other non-differentiable point α'. For each non-differentiable point α, let $r(\alpha)$ be an arbitrary rational in $(\eta_\alpha^-, \eta_\alpha^+)$. The rationals are disjoint and countable. Thus, the set of non-differentiable points in D is countable. $\square$

A discrete error curve is a convex piecewise-linear function. As stated before, the derivative $\frac{d\hat{\beta}(\alpha)}{d\alpha}$ does not exist in the ordinary sense at the vertex points but is well-defined by an interval. The non-vertex points of an error curve are on linear segments of the error curve where the derivative is constant. The general error curve is any arbitrary bounded convex function defined on the range $[0, 1]$ where $\hat{\beta}(0) = p$ and $\hat{\beta}(q) = 0$ for $0 < p, q \leq 1$. The set of points where the derivative does not exist is at most countable by lemma 2.4.1. We will assume that discrete channels have a finite set of LR.

Relay channels are normally determined by the environment, and the problem of interest is to find the optimal strategy at the given relay and fusion point. For understanding and insight, however, we will frequently develop classes of relay channels which exhibit certain desired properties. One of the techniques is to construct these error curves. For non-erasure discrete channels with an alphabet of N outputs, we use the convention throughout the thesis of ordering the vertex points from left to right such that $V_i < V_{i+1}$ means $\alpha_i \leq \alpha_{i+1}$ and $\beta_i \geq \beta_{i+1}$. Also, we take $V_0 = (0,1)$ and $V_{N+1} = (1,0)$. The error curve is convex so it must obey the following rules when constructing the curve from left to right. $V_j = (\alpha_j, \beta_j)$ must lie above the line through V_{j-1} and V_{j-2}, and it must lie below the line through V_{j-1} and $(1,0)$. Algebraically, for $V_j = (\alpha_j, \beta_j)$, must lie below the line $\hat{\beta}(\alpha) = \left(\frac{\beta_{j-1}}{1-\alpha_{j-1}}\right)(1-\alpha)$ and it must lie above the line $\hat{\beta}(\alpha) = \left(\frac{\beta_{j-1}-\beta_{j-2}}{\alpha_{j-1}-\alpha_{j-2}}\right)\alpha + \left(\frac{\alpha_{j-1}\beta_{j-2}-\alpha_{j-2}\beta_{j-1}}{\alpha_{j-1}-\alpha_{j-2}}\right)$. This is shown in figure 2-5. An error curve can be constructed successively after fixing the previous point. There are many ways to construct a finite discrete error curve. The only requirement is that the successive points must form a convex curve between $(0,1)$ and $(1,0)$, and thus, must satisfy the condition above.

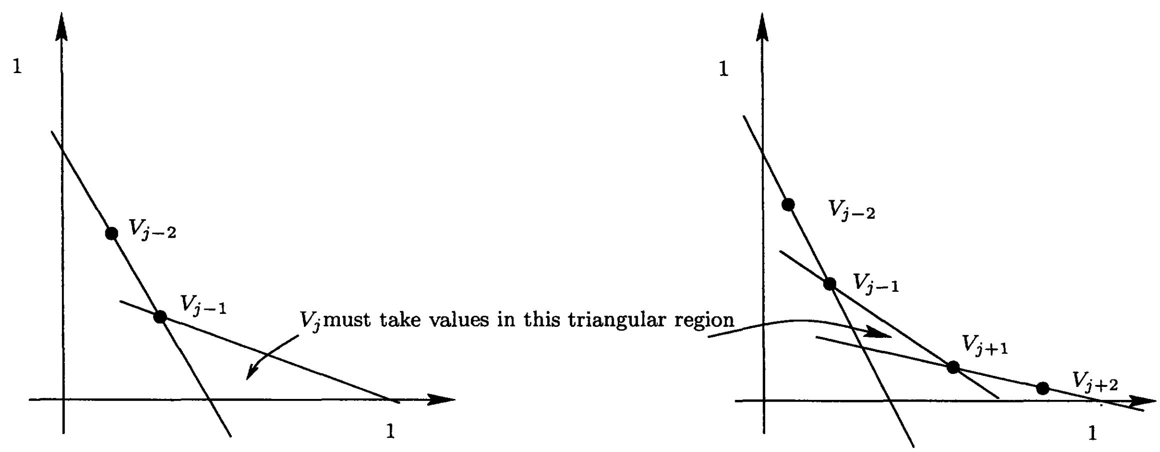


Figure 2-5: Possible values of V_j to assure convexity of the discrete error curve. This will often be useful in constructing examples of the error curve with given desired properties.

All binary input memoryless channels have an error curve. We have defined a *discrete error curve* as a system when the likelihood ratios take on a finite set of discrete values, which means that the error curve is piecewise linear and has a finite number of vertex points. Also, the function $\tilde{\alpha}(\eta)$ is a monotonically decreasing staircase function. A discrete error curve does not imply that the is necessarily discrete. Every point on a strictly convex error curve is a vertex point. A strictly convex error curve cannot be described by a finite DMC. However, any discrete error curve can be described by a continuous or a discrete memoryless channel. For example, a continuous channel with the same error curve as the BEC(ϵ) is the following. Let $Y = X + N$

with $x \in (-1, 1)$ and

$$f_N(n) = \begin{cases} \frac{1-\epsilon}{2} & \text{for} \quad |n| < \frac{1}{1-\epsilon} \\ 0 & \text{otherwise} \end{cases}$$

In our study of the distributed detection and coding problem, we will frequently consider relay error curves which are strictly convex or strictly discrete, since each brings out different fundamental insights to the fusion problem. Most importantly, the LR tells us everything we need to know about the channel and completely determines the error curve.

If a discrete error curve has N slopes (i.e., N likelihood ratios), it will have $N-1$ vertex points in addition to the endpoints $(0, 1)$ and $(1, 0)$. Recall that we take the convention of labeling the $(N-1)$ vertex points from left to right such that $\alpha_i < \alpha_j$, $\forall i < j$ with $(\alpha_0, \beta_0) = (0, 1)$ and $(\alpha_N, \beta_N) = (1, 0)$. Define the *"canonical channel"* representation for an $(N-1)$ point error curve in terms of the LR's as depicted in figure 2-6. The outputs of the channel are the LR's and ordered. The canonical channel is a binary input, N output DMC with the following transition probabilities, for $0 \leq j \leq N$, $\Pr(Y = j \mid X = 0) = \beta_j - \beta_{j+1}$ and $\Pr(Y = j \mid X = 1) = \alpha_{j+1} - \alpha_j$. The channel outputs are $(0, 1, \ldots, N-1)$ and are labeled in decreasing order of LR. The canonical channel is defined to be an "*Erasure-Like Channel*" when $\alpha_1 = 0$ or $\beta_{N-1} = 0$ or when both occur. In Erasure-Like channels, the symbols "$0''$ or "$N-1''$ or both are detected with zero error.

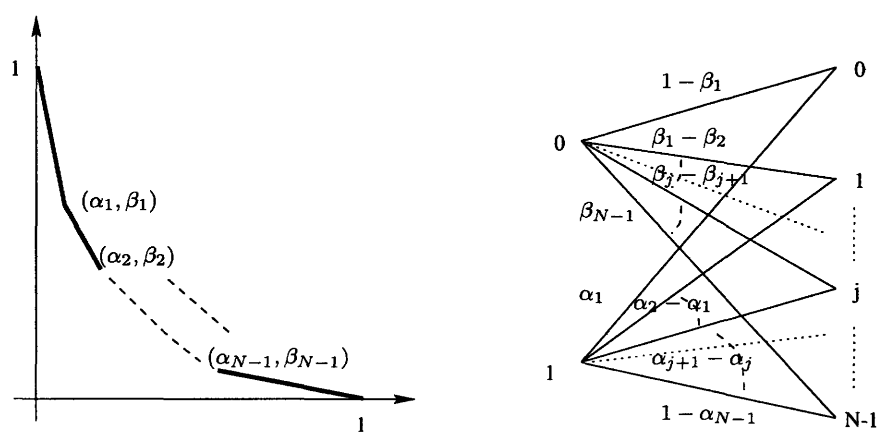


Figure 2-6: An $(N-1)$ point error curve and its canonical channel representation. The N outputs are labeled in decreasing order of their likelihood ratios.

When a channel is symmetric, the error curve is symmetric about the 45 degree line, $\alpha = \beta$. That is, for every point (α, β) on $\hat{\beta}(\alpha)$, (β, α) is also on $\hat{\beta}(\alpha)$. A binary input channel is said to be symmetric if, when the input labels are reversed, the output labels can be changed equivalently such that all channel transition probabilities are the same as they were originally.

In the examples discussed, we can reverse both input and output labels of the BSC and BEC without overall change while we cannot do so for the Z and BAC.

2.5 The Binary Asymmetric Channel and Simple Fusion

Define an *equivalent channel* as that part of the system from binary source X to relay output Z. If a relay maps its input Y into a single binary digit, then for any given mapping, i.e., any test, the equivalent channel is some given Binary Asymmetric Channel (BAC). If the relay is constrained to transmitting an M-ary symbol to the fusion point, then the equivalent channel is a binary input M-ary output DMC.

Consider a two relay system, each with a different BAC relay channel . Suppose that Relay A has its one vertex point error curve at (α, β) and Relay B at (α', β'). Assume MAP detection is implemented at the fusion point. The fusion point then receives one of four possible inputs, $\{00, 01, 10, 11\}$. Thus, there are 4 likelihood ratios and the error curve at the fusion point is shown in figure 2-7 For the input (00), $P_{Z,Z'|X}(00|0) = (1-\beta)(1-\beta')$ and $P_{Z,Z'|X}(00|1) = \alpha\alpha'$, so the LR is given by $\Lambda(00) = \left(\frac{1-\beta}{\alpha}\right)\left(\frac{1-\beta'}{\alpha'}\right)$. Similarily, $\Lambda(01) = \left(\frac{1-\beta}{\alpha}\right)\left(\frac{\beta'}{1-\alpha'}\right)$, $\Lambda(10) = \left(\frac{\beta}{1-\alpha}\right)\left(\frac{1-\beta'}{\alpha'}\right)$, and $\Lambda(11) = \left(\frac{\beta}{1-\alpha}\right)\left(\frac{\beta'}{1-\alpha'}\right)$. Without loss of generality assume that $\left(\frac{1-\beta}{\alpha}\right)\left(\frac{\beta'}{1-\alpha'}\right) \geq \left(\frac{\beta}{1-\alpha}\right)\left(\frac{1-\beta'}{\alpha'}\right)$. Otherwise, just reverse the labels of Relay A and Relay B. The fusion point receives a binary digit from each relay and must map its reception pair to either "0" or "1." MAP detection says the following. If $\eta > \left(\frac{1-\beta}{\alpha}\right)\left(\frac{1-\beta'}{\alpha'}\right)$, the fusion point maps everything to "1." The mapping of every reception to either a "0" or "1" is called a "trivial decoder." If $\left(\frac{1-\beta}{\alpha}\right)\left(\frac{\beta'}{1-\alpha'}\right) < \eta \leq \left(\frac{1-\beta}{\alpha}\right)\left(\frac{1-\beta'}{\alpha'}\right)$, then it maps $(00) \rightarrow 0$ and everything else to "1." This is called the "OR" decoder. If $\left(\frac{\beta}{1-\alpha}\right)\left(\frac{1-\beta'}{\alpha'}\right) < \eta \leq \left(\frac{1-\beta}{\alpha}\right)\left(\frac{\beta'}{1-\alpha'}\right)$ then the optimal map is $\{(00),(01)\} \rightarrow 0$ and $\{(10),(11)\} \rightarrow 1$. If $\left(\frac{\beta}{1-\alpha}\right)\left(\frac{\beta'}{1-\alpha'}\right) < \eta \leq \left(\frac{\beta}{1-\alpha}\right)\left(\frac{1-\beta'}{\alpha'}\right)$ then the fusion point maps $(11) \rightarrow 1$ and everything else to "0." This is called the "AND" decoder. Finally, if $\eta \leq \left(\frac{\beta}{1-\alpha}\right)\left(\frac{\beta'}{1-\alpha'}\right)$, the fusion center maps everything it receives to "0," which is the trivial decoder. Note that at the value η when the η-tangent coincides with 2 vertex points, there are 2 decoder maps which give the same error probability. These mappings are called *"Don't Care"* maps.

The solution to this detection problem is completely summarized in the error curve at the fusion point which consists of 3 vertex points and 4 LR. Conditional on the source, the relay outputs are

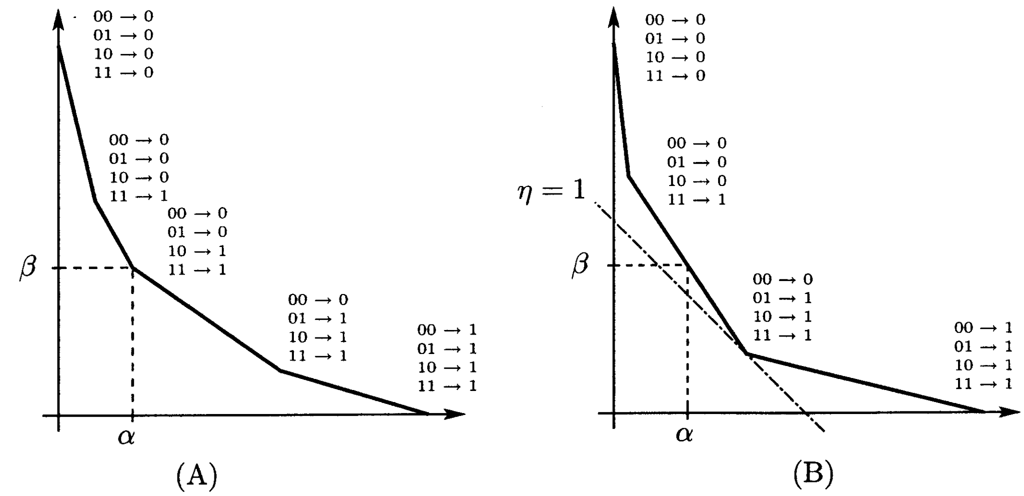


Figure 2-7: (A) The error curve at the fusion point for two different Binary Asymmetric Channel (BAC) in parallel, namely (α, β) and (α', β'). (B) Error curve at the fusion point for 2 relays using identical BAC Channel Equivalents (α, β). The optimal decoder at the fusion point is listed next to the error points. At ML, $\eta = 1$, the error point is $(2\epsilon_1 - \epsilon_1^2, \epsilon_0^2)$ and the optimal decoder is the "AND" rule.

independent, so at the fusion point, the likelihood of receiving the symbol $\vec{Z} = (Z_A, Z_B) = (i, j)$ is $\Lambda(ij) = \lambda_A(i)\lambda_B(j)$. Note that as η increases from zero to infinity, the η-tangent solution is on the error curve and then rotates around a vertex point. Each vertex point corresponds to an optimal decoding rule, the "OR," "AND," "trivial decoder," etc.

Although this two relay BAC is a centralized problem, it is the basis of the most elementary fusion problem. When each relay is restricted to a particular binary decision rule, then each relay channel is equivalent to a BAC. Thus, when 2 relays are each restricted to a particular decision rule for transmitting one bit to the fusion point, the problem reduces to this parallel BAC problem. The difficulty is determining the optimal binary decision rule among the set of possible BACs which each relay can choose from. Chapter 3 will deal with the joint optimization problem at the relays and fusion point.

If the two relays each use the same BAC (i.e., the same test), the error curve for the fusion point is shown in figure 2-7B. Since there are four possible outputs from the two relays combined, there are at most four distinct likelihood ratios. In fact, there are only three likelihood ratios since the outputs (1,0) and (0,1) have the same likelihood ratio. Observe that the error curve has three distinct straight line portions, one corresponding to $\Lambda(00)$, another to $\Lambda(11)$, and the third to $\Lambda(01)$ and $\Lambda(10)$, which must be the same. The optimal decision mapping for each LR is listed next to its respective point on the curve.

2.6 The Binary Erasure Channel and Coordinated Randomized Strategy

If the relay channel is binary, such as the BAC or Z or BSC, each relay merely forwards its observed bit to the destination. Since the relay constraint of one bit is not relevant for these channels, this system is centralized. We now look at the case when the relay channel is a binary erasure channel (BEC); this makes the relay output constraint of one bit relevant. We introduce the notion of coordinated randomized strategies and further clarify the concept of a randomized threshold test, which is a particular randomized strategy. A *coordinated randomized strategy* is defined as relays randomizing together in a coordinated way, e.g., whenever Relay A uses strategy 1 then Relay B uses strategy 2 and whenever Relay A uses strategy 3 then Relay B uses strategy 4, and so forth. The fusion point knows which strategy each relay is using at each time.

The error curve for a single BEC is shown in figure 2-8(A). For the centralized system of L relays, all transmitting their full ternary observation to the destination, an error can only occur if all relays see an erasure and thus the error curve is shown in 2-8(B).

We will show in chapter 3 that randomized tests at the relays do not help at the fusion point for MAP. Thus, all that needs to be considered is the convex hull of the fused non-random threshold tests at the relays. Finally, a randomized threshold test on the error points at the fusion point, which translates to a coordinated randomized threshold rule among the relays, is used to obtain the final error curve at the fusion point. These types of relay strategies are *Coordinated Randomized Strategies.*

For the BEC, there are only two possible threshold rules. Each of these rules correspond to a Z-channel equivalent which is illustrated in figure 2-8 with its corresponding error curve. Fusing these Z-channels, it can be seen that if both relays use the same Z-channel, an error occurs when the fusion device receives an error from both equivalent channels. Thus, the error points $(0, \epsilon^2)$ for one of the Z-channels and $(\epsilon^2, 0)$ for the other Z-channel can be achieved. If the relays use opposite Z-channels, an error can occur when the fusion center receives an error from either equivalent channel (it is impossible for both equivalent channels to make an error

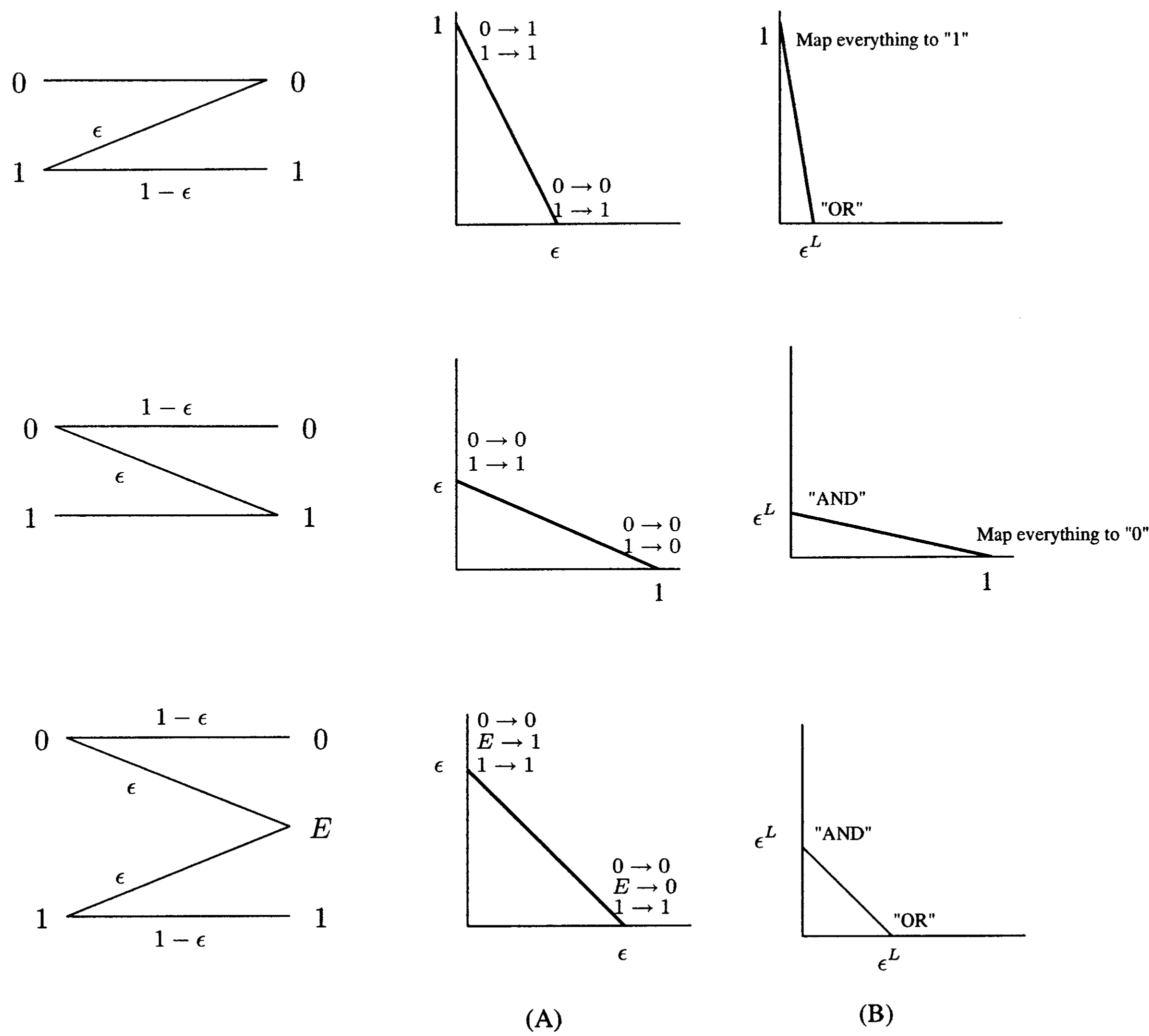


Figure 2-8: The error curve for the Z Channel and Binary Erasure Channel (BEC) for (A) 1 relay and (B) L relays. The Z Channels are the equivalent channels of the BEC when the relays can only forward 1 bit to the destination. The optimal fusion decoders are listed next to each vertex point on the error curve.

simultaneously). For this choice, the points $(0, \epsilon)$ and $(\epsilon, 0)$ can be achieved.

A distributed 1-bit relay system is allowed to coordinate the usage of these equivalent channels. Thus, if a coordinated randomized threshold rule is used among the relays, then every point between the line connecting $(0, \epsilon^2)$ and $(\epsilon^2, 0)$ can be achieved. That is, part of the time, both relays will use the equivalent Z-channel with the "OR" decoding rule at the fusion point, and the rest of the time they will simultaneously use the other Z-channel equivalent with the "AND" decoding rule. The combined error curve is exactly the error curve of the centralized BEC shown in figure 2-8. The generalization to arbitrary L relays is straightforward.

In conclusion, every point on the fused error curve for the BEC can be achieved by relays sending binary data rather than the full ternary data observation but using a coordinated randomized strategy among the relays. The BEC relay case is a very special case because there is no loss in performance for this particular distributed system. In Chapter 4, we will understand why this happens. We will also explain the class of channels and the conditions when relays can transmit less than full information and still attain the same performance as if they send their full observation to the fusion point.

Chapter 3

Fusion of Two Relays with Binary Outputs

This chapter is devoted entirely to the case of two relays, each transmitting a single binary digit to the fusion point. This will be referred to as the 2×2 system, where the numbers represent the size of the output alphabet allowed each relay. To simplify notation, we will restrict attention to the case where the two relay channels are identical. Almost all the results are applicable to non-identical relay channels and just require a simple modification of notation. We will point out the places where more significant change is required. We first describe the fusion process and the mechanics of jointly optimizing the relay and fusion point strategies. For completeness, we prove via a simple method with insight the known result that threshold tests on the relay error curve are optimal for fusion. Randomized strategies and all the special cases when randomization is optimal are discussed. This is followed by a study of the structure of ternary output relay channels to establish structure rules for the general 2×2 system. Finally, we discuss and provide the complete optimal solution for the general 2×2 system, and in particular, we generalize and establish the MAP solution for the previously discussed AGN channel.

3.1 The Error Curve and the Fusion Process

In the first part of this section, we will show how to fuse the outputs of the relays for a fixed decision rule at the relays. Recall that X is the source, Y the channel output for Relay A, Y' the channel output for Relay B, Z the output of Relay A, and Z' the output of Relay B. Given the input, we have conditional independence of the channel outputs $P(YY'|X) = P(Y|X)P(Y'|X)$, which means conditional independence at the fusion point, $P(ZZ'|X) = P(Z|X)P(Z'|X)$. When each relay is restricted to a particular binary test, the equivalent channel from source to relay output is a BAC. For two relays each restricted to a binary output, the fusion point will optimally decode the two parallel BAC's as seen in chapter 2. To accomplish this, we will derive the error curve for the two equivalent channels. The class of equivalent channels is restricted by the relay channel $\hat{\beta}(\alpha)$. Assume that the actual relay channels at the input to the relays are described by densities $f_{Y|X}(Y|0)$ and $f_{Y|X}(Y|1)$. The likelihood ratio for a single relay is defined as

$$\lambda(Y) = \frac{f_{Y|X}\{y|\,0\}}{f_{Y|X}\{y|\,1\}} \tag{3.1}$$

Suppose the relays perform threshold tests to generate the symbols Z and Z', which are forwarded to the fusion point. The output probabilities for Relay A for a given threshold test (α, β) with threshold η are

$$\begin{aligned}
\Pr\{Z(\alpha) = 0\,|\,X{=}1\} &= \Pr\{\lambda(Y){\geq}\eta\,|\,X{=}1\} = P^{\eta}_{e|1} = \alpha && (3.2)\\
\Pr\{Z(\alpha) = 1\,|\,X{=}1\} &= \Pr\{\lambda(Y){<}\eta\,|\,X{=}1\} = 1 - P^{\eta}_{e|1} = 1 - \alpha && (3.3)\\
\Pr\{Z(\alpha) = 0\,|\,X{=}0\} &= \Pr\{\lambda(Y){\geq}\eta\,|\,X{=}0\} = 1 - P^{\eta}_{e|0} = 1 - \beta && (3.4)\\
\Pr\{Z(\alpha) = 1\,|\,X{=}0\} &= \Pr\{\lambda(Y){<}\eta\,|\,X{=}0\} = P^{\eta}_{e|0} = \beta && (3.5)
\end{aligned}$$

and likewise for output Z' of Relay B. The error curve for relay input Y and output $Z(\alpha)$ is shown in figure 3-1. The magnitude of the slopes from $(0, 1)$ to (α, β) and from (α, β) to $(1, 0)$ of the error curve are the likelihood ratios of the relay outputs

$$\lambda^{(\alpha)}(0) = \frac{1 - P^{\eta}_{e|0}}{P^{\eta}_{e|1}} = \frac{1-\beta}{\alpha} \quad ; \quad \lambda^{(\alpha)}(1) = \frac{P^{\eta}_{e|0}}{1 - P^{\eta}_{e|1}} = \frac{\beta}{1-\alpha}. \tag{3.6}$$

The error curve for the other relay using a test (α', β') and threshold η is constructed the same way. The BAC equivalent channel is illustrated by the dotted line in figure 3-1. Since the relays are conditionally independent given the input, we have the following probabilities at the

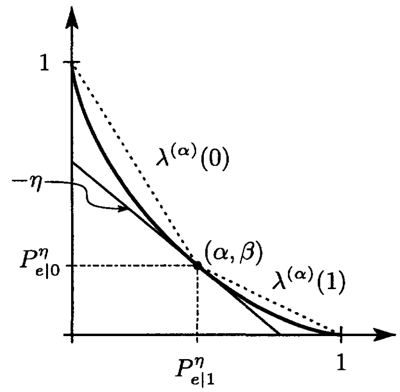


Figure 3-1: The error curve for the equivalent channel, represented by the dotted line, of a single relay output where (α, β) is the relay test. When the test is chosen as a single point (α, β) on the error curve, as shown in the figure, the relay strategy is a threshold test with threshold η where $(\alpha, \beta) = (\alpha(\eta), \beta(\eta))$.

fusion point,

$$\begin{aligned}
\Pr\{00 \mid X{=}0\} &= \left(1 - P^{\eta}_{e|0}\right)\left(1 - P^{\eta'}_{e|0}\right) = (1-\alpha)(1-\alpha') && (3.7)\\
\Pr\{00 \mid X{=}1\} &= (P^{\eta}_{e|1})(P^{\eta'}_{e|1}) = \beta\beta' && (3.8)\\
\Pr\{01 \mid X{=}0\} &= \left(1 - P^{\eta}_{e|0}\right)\left(P^{\eta'}_{e|0}\right) = (1-\alpha)(\alpha') && (3.9)\\
\Pr\{01 \mid X{=}1\} &= (P^{\eta}_{e|1})(1 - P^{\eta'}_{e|1}) = (\beta)(1-\beta') && (3.10)
\end{aligned}$$

and so forth for fusion reception of 10 and 11. The likelihood ratio for the reception of $(i, j) \in \{(0,0), (0,1), (1,0), (1,1)\}$ at the fusion point is

$$\Lambda(i,j) = \lambda_Z(i)\lambda_{Z'}(j) = \frac{\Pr\{i,j \mid X{=}0\}}{\Pr\{i,j \mid X{=}1\}} \qquad (3.11)$$

Each Likelihood Ratio $\Lambda(i, j)$ is the magnitude of one of the slopes of the fused error curve at the fusion point. Figure 3-2 shows one specific fused error curve for particular tests (α, β) for Relay A and (α', β') for Relay B. As can be seen in the figure, the error curve is formed by ordering the slopes in decreasing magnitude.

Note that α specifies $V_\alpha = (\alpha, \beta)$ a test on the error curve. Thus, for notational convenience, α, V_α, (α, β), and $(\alpha, \hat{\beta}(\alpha))$ all represent the error point on the error curve. We will attempt to use the notation with the least clutter and greatest clarity.

Define $F_{\alpha,\alpha'}(\phi)$ to be the error curve at the fusion point for 2 relays for a given choice of threshold tests α and α' at Relays A and B, respectively. Hence, $F_{\alpha,\alpha'}(\phi)$ is a double indexed function of α and α'. This indexed function is general and extends to any finite number of relays. For example, $F_{\alpha,\alpha',\alpha''}(\phi)$ represents the indexed function for 3 relays where $\alpha, \alpha', \alpha''$ are, respectively, the threshold tests of Relay A, Relay B, and Relay C. We denote $V^L_{\alpha,\alpha'} = (\alpha\alpha', \beta + \beta' - \beta\beta')$ as the *left vertex point* and $V^R_{\alpha,\alpha'} = (\alpha + \alpha' - \alpha\alpha', \beta\beta')$ the corresponding *right vertex point* of the fused curve $F_{\alpha,\alpha'}(\phi)$ as seen in figure 3-2. We will show later that choosing (α', β') as points on the relay error curve are optimal relay strategies. The final fused curve $F(\phi)$ at the fusion point is then the convex envelope of $F_{\alpha,\alpha'}(\phi)$ over all α and α' as illustrated in figure 3-2. In other words, $F(\phi) = \inf_{\alpha,\alpha'} F_{\alpha,\alpha'}(\phi)$.

There are three possibilities for the likelihood ratios of $\Lambda(0,1)$ and $\Lambda(1,0)$. When $\Lambda(0,1) > \Lambda(1,0)$, the curve will be the one shown in figure 3-2. The figure shows that the middle vertex point (α, β) then lies on the relay input error curve $\hat{\beta}(\alpha)$. When $\Lambda(0,1) < \Lambda(1,0)$, the roles of Relay A and Relay B are reversed. The center point is now (α', β') and again always lies on the relay input error curve $\hat{\beta}(\alpha)$. We will show later which of these points, (α, β) or (α', β'), will lie on $F_{\alpha\alpha'}(\phi)$. When Relay A and Relay B use identical threshold tests, i.e, when $(\alpha, \beta) = (\alpha', \beta')$, then $\Lambda(0,1) = \Lambda(1,0)$ and it can be seen from the figure that the center point is colinear with $V^L_{\alpha,\alpha'}$ and $V^R_{\alpha,\alpha'}$, and thus, not a vertex point.

Recall from chapter 2 that a randomized threshold test is represented by a single point on the error curve. All other tests must lie at or above the relay error curve. We show in the following theorem that it is sufficient to consider only relay strategies that are points on the relay error curve. We will sometimes call this choice of error points the relay encoding point. In later chapters, we will generalize the theorem and show that choosing points on the relay error curve is sufficient for optimality for multiple relays and non-binary relay outputs.

Theorem 3.1.1 *The use of points on the error curve at the relays is sufficient for optimality for the* 2×2 *system.*

Proof: Fix Relay B at an arbitrary test (α', β'). Suppose Relay A has an arbitrary non-threshold test $(\alpha, \beta + \epsilon)$ for some $\epsilon > 0$ where $\beta = \hat{\beta}(\alpha)$, i.e, when (α, β) is an error point. We will show that the fused curve of Relay B with the randomized threshold test (α, β) of Relay A lies on or below the fused curve with non-threshold test $(\alpha, \beta + \epsilon)$ of Relay A. These fused curves are illustrated in figure 3-3. The 3 vertex points of $F_{(\alpha,\beta+\epsilon),(\alpha',\beta')}(\phi)$ from left to right

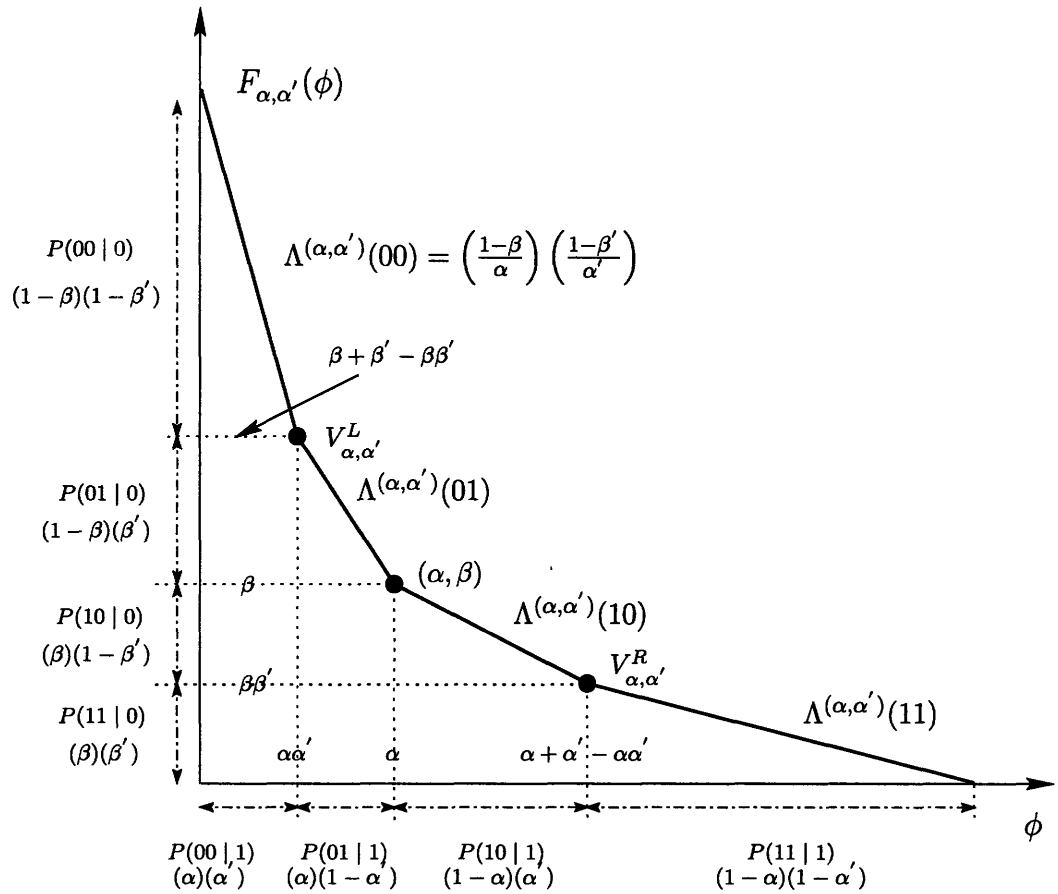


Figure 3-2: Fused error curve $F_{\alpha\alpha'}(\phi)$ for 1 relay with strategy at error point (α, β) and the other relay at (α', β'). The equivalent channels are the two relays are BAC(α) and BAC(α'). Note that the abscissa of the center point is at $\alpha\alpha' + \alpha(1-\alpha') = \alpha$ and the ordinate is at β.

are

$$V^L_{(\alpha,\beta+\epsilon),(\alpha',\beta')} = (\alpha\alpha', \beta+\beta'-\beta\beta'+\epsilon(1-\beta')) \tag{3.12a}$$

$$(\alpha, \beta+\epsilon) \quad \text{or} \quad (\alpha', \beta') \tag{3.12b}$$

$$V^R_{(\alpha,\beta+\epsilon),(\alpha',\beta')} = (\alpha+\alpha'-\alpha\alpha', \beta\beta'+\epsilon\beta'). \tag{3.12c}$$

The 3 vertex points of $F_{(\alpha,\beta),(\alpha',\beta')}(\phi)$ from left to right are

$$V^L_{(\alpha,\beta),(\alpha',\beta')} = (\alpha\alpha', \beta+\beta'-\beta\beta') \tag{3.13a}$$

$$(\alpha, \beta) \quad \text{or} \quad (\alpha', \beta') \tag{3.13b}$$

$$V^R_{(\alpha,\beta),(\alpha',\beta')} = (\alpha+\alpha'-\alpha\alpha', \beta'\beta). \tag{3.13c}$$

Comparing the equations of 3.12 with 3.13, we find that every vertex point of $F_{(\alpha,\beta),(\alpha',\beta')}(\phi)$ is on or below the corresponding vertex point of $F_{(\alpha,\beta+\epsilon),(\alpha',\beta')}(\phi)$. Thus, the ordinate of the fused curve of Relay B with a non-threshold test for Relay A is everywhere below or equal to the ordinate of the fused curve of Relay B with a threshold test for Relay A as depicted in figure 3-3. Apply the same argument to Relay B for a fixed test at Relay A. This implies that threshold tests at the relays are sufficient for optimality in fusion. □

Thus, choosing any test not on the relay error curve cannot perform any better than a test on the relay error curve. Since we are only interested in optimal solutions, from here on, only randomized threshold tests $(\alpha, \beta) = (\alpha, \hat{\beta}(\alpha))$ will be considered unless otherwise noted.

Corollary 3.1.1 *If the relay error curve $\hat{\beta}_1(\alpha)$ is less than or equal to an error curve $\hat{\beta}_2(\alpha)$, then the fusion of an arbitrary relay error curve $\hat{\beta}_3(\alpha)$ with $\hat{\beta}_1(\alpha)$ is less than or equal to the fusion of $\hat{\beta}_3(\alpha)$ with $\hat{\beta}_2(\alpha)$.*

Proof: This is a direct consequence of theorem 3.1.1. For every point $(\alpha, \hat{\beta}_2(\alpha))$, there is a corresponding point $(\alpha, \hat{\beta}_1(\alpha))$ and $\hat{\beta}_1(\alpha) < \hat{\beta}_2(\alpha)$. □

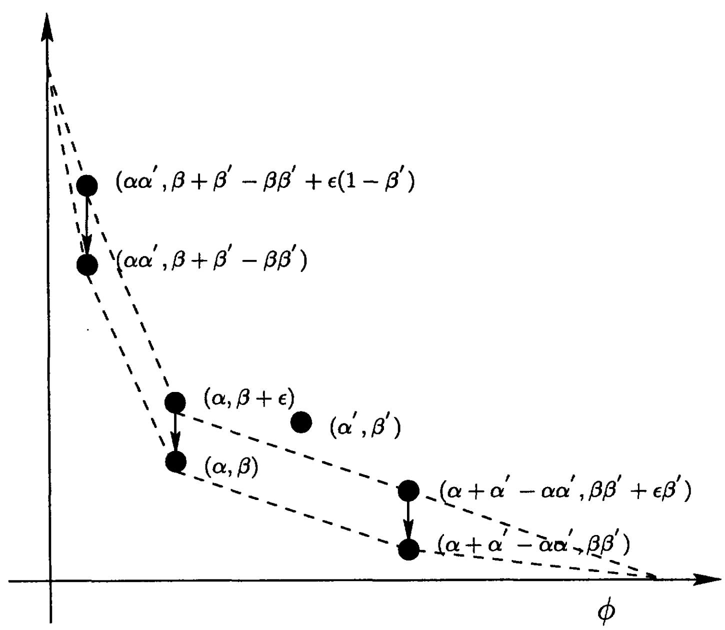


Figure 3-3: Comparison of the fused error curves of a threshold test with that of a non-threshold test. The top curve with dark vertex points is the fused error curve of non-threshold test $(\alpha, \beta + \epsilon)$ with threshold test (α', β'). The bottom curve represents the fused curve of both relays using threshold tests (α, β) and (α', β').

3.1.1 The Final Fused Curve $F(\phi)$

Recall from the previous section and figure 3-2 that the fused curve $F_{\alpha,\alpha'}(\phi)$ for a 2×2 system with given threshold tests, α and α', has at most four slopes corresponding to each of the likelihood ratios . There are at most five vertex points which include the trivial points $(0,1)$ and $(1,0)$. The non-trivial points consist of the left vertex point $V^L_{\alpha,\alpha'}$, the right vertex point $V^R_{\alpha,\alpha'}$, and perhaps a center point (α,β) or (α',β'). From now on, when we talk of vertex points, we only refer to the non-trivial points unless otherwise specified. Rather surprisingly, the following theorem shows that the two vertex points, $V^L_{\alpha,\alpha'} = (\alpha\alpha', \beta+\beta'-\beta\beta')$ and $V^R_{\alpha,\alpha'} = (\alpha+\alpha'-\alpha\alpha', \beta\beta')$ over all (α,β) and (α',β') completely determine the final fused curve $F(\phi)$. The final fused curve at the fusion point is an error curve where $\phi = P(e|1)$ and $F(\phi) = P(e|0)$.

Theorem 3.1.2 *The optimal fused error curve $F(\phi)$ is the convex envelope of the points $(0,1)$, $(1,0)$, $V^L_{\alpha,\alpha'} = (\alpha\alpha', \beta+\beta'-\beta\beta')$, and $V^R_{\alpha,\alpha'} = (\alpha+\alpha'-\alpha\alpha', \beta\beta')$ over all α, α', with $\beta = \hat{\beta}(\alpha)$ and $\beta' = \hat{\beta}(\alpha')$.*

Proof: Given α and α', $F_{\alpha,\alpha'}(\phi)$ has at most three non-trivial vertex points, namely, the left vertex point $V^L_{\alpha,\alpha'} = (\alpha\alpha', \beta+\beta'-\beta\beta')$, center point (α,β) or (α',β'), and the right vertex point $V^R_{\alpha,\alpha'} = (\alpha+\alpha'-\alpha\alpha', \beta\beta')$. We will show that the center point, i.e. $V_{\alpha'} = (\alpha',\beta')$ or $V_\alpha = (\alpha,\beta)$, of any fused curve $F_{\alpha,\alpha'}(\phi)$ is not a vertex point on the final fused curve $F(\phi)$. To see this, suppose both relays use the same threshold test (α',β'); then the center point will be $V_{\alpha'} = (\alpha',\beta')$. This center point cannot be a vertex point of $F_{\alpha',\alpha'}(\phi)$, since (α',β') lies exactly halfway in between $V^L_{\alpha',\alpha'}$ and $V^R_{\alpha',\alpha'}$ and the points $V^L_{\alpha',\alpha'}, V_{\alpha'}, V^R_{\alpha',\alpha'}$ are all colinear. Now $F(\phi)$ is the convex envelope over all $\{V^L_{\alpha,\alpha'}, V_\alpha, V_{\alpha'}, V^R_{\alpha,\alpha'}\}$, for all α, α', so neither $V_\alpha = (a,\beta)$ nor $V_{\alpha'} = (a',\beta')$ can be a vertex point on the final fused error curve $F(\phi)$. □

We call "trivial error curves" error curves that have at most one non-trivial vertex point. In the following section we will see that for all error curves with two non-trivial vertex points, V_1 and V_2, neither vertex point can be on the final fused curve $F(\phi)$. It follows that for any non-trivial relay curve, no relay error point (α,β) other than $(0,1)$ or $(1,0)$ ever lies on the final fused curve. This means that for all relay channels other than trivial ones and for all a priori probabilities other than the ones corresponding to the trivial vertex points $(0,1)$ and $(1,0)$ on $F(\phi)$, the error probability of two relays is strictly smaller than one relay.

We now look at optimal decoding strategies at the fusion point. Note that the final fused curve $F(\phi)$ is an error curve. Define $\hat{\phi}(\eta)$ as a mapping from η to ϕ. The points of $F(\phi)$ are the vertex points $(0,1), (1,0), V^L_{\alpha,\alpha'}, V^R_{\alpha,\alpha'}$ for different α, α' and linear segments with 2 vertex points as endpoints. Recall from chapter 2 that the fused curve for 2 BACs, the left vertex point $V^L_{\alpha,\alpha'}$ corresponds to the "OR" decoding rule for all η-tangents that rotate around that vertex point. For $\hat{\phi}(\eta) = (0,1)$, the decoder decodes everything to "1" (trivial decoder); if $\hat{\phi}(\eta) = V^L_{\alpha,\alpha'}$, the optimal fusion decoder is the "OR" rule; if $\hat{\phi}(\eta) = V^R_{\alpha,\alpha'}$ the "AND" rule is optimal; and finally, if $\hat{\phi}(\eta) = (1,0)$, everything is decoded to "0" (trivial decoder).

Just as for the single relay, an arbitrary cost function is not fundamentally different from the MAP problem. We have seen that for a system with relays, the final fused error curve at the fusion point can always be achieved by using randomized threshold tests at the relays. Again, given a cost function and a priori probabilities for the overall system, the solution is the solution to the MAP problem with modified a priori probabilities at the fusion point, and thus, threshold tests on the LR at the relays minimizes the overall cost. This implies that threshold tests at the relays are optimal. We will further see that even if we are only interested in ML or MAP detection for a fixed a priori distribution, the entire error curve corresponding to each relay is relevant to the solution at the fusion point.

Lemma 3.1.1 *Fusing two symmetric relay error curves will result in a symmetric fused error curve, i.e., for each ϕ, $(F(\phi), \phi)$ lies on the fused error curve $(\phi, F(\phi))$.*

Proof: Let the functions g, h be the following, $g(u,v) = uv$ and $h(u,v) = u + v - uv$. Then the left vertex point of $F_{\alpha,\alpha'}(\phi)$ is $V^L_{\alpha\alpha'} = (g(\alpha,\alpha'), h(\beta,\beta'))$ and the right vertex point $V^R_{\alpha\alpha'} = (h(\alpha,\alpha'), g(\beta,\beta'))$. Every point (α,β) on the symmetric relay error curve has a corresponding point (β,α). The vertex points of $F_{\beta\beta'}(\phi)$ are $V^L_{\beta\beta'} = (g(\beta,\beta'), h(\alpha,\alpha'))$ and $V^R_{\beta\beta'} = (h(\beta,\beta'), g(\alpha,\alpha'))$. The left and right vertex points of $F_{\alpha,\alpha'}(\phi)$ are the reflection over the line $\alpha = \beta$ of the right and left vertex points of $F_{\beta,\beta'}(\phi)$, respectively. Each point on the final fused curve has either the form $(g(\alpha,\alpha'), h(\beta,\beta'))$ or $(h(\alpha,\alpha'), g(\beta,\beta'))$ and as shown above, the points have symmetric counterpoints $(g(\beta,\beta'), h(\alpha,\alpha'))$ and $(h(\beta,\beta'), g(\alpha,\alpha'))$. Since this occurs for every point on the symmetric relay error curve, the final fused curve $F(\phi)$ must be symmetric. □

3.1.2 The Conditional Locus Curve $F_\alpha(\phi)$

Given Relay A with a fixed threshold test at (α,β), define the *conditional locus curve* $F_\alpha(\phi) = \inf_{\alpha'} F_{\alpha,\alpha'}(\phi)$ for fixed (α,β) as (α',β') varies over the relay error curve $\hat{\beta}(\alpha)$ (see figure 3-4).

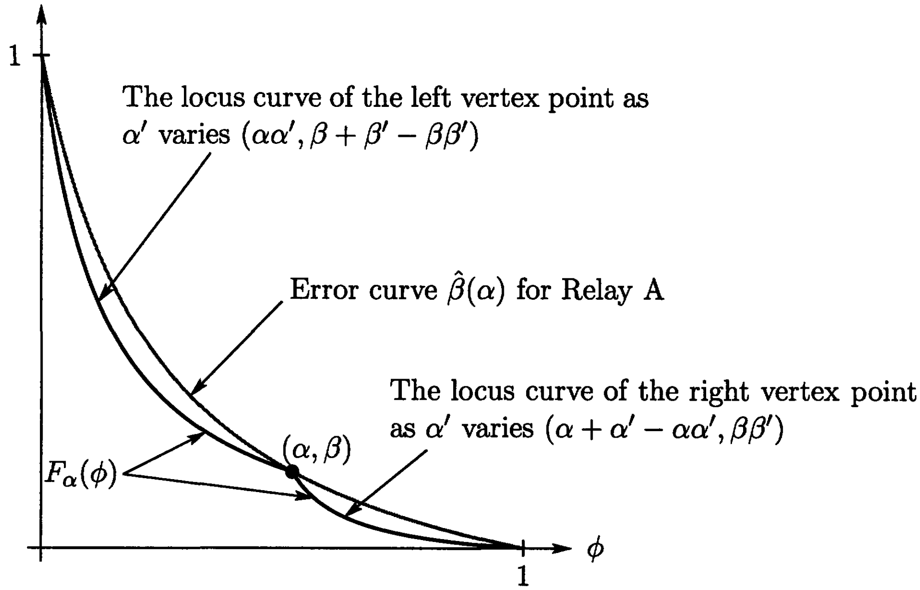


Figure 3-4: Conditional locus curve for fixed α as α' varies, $F_\alpha(\phi) = \inf_{\alpha'} F_{\alpha,\alpha'}(\phi)$. As α' varies from 0 to 1, the left vertex curve is the original curve $\hat{\beta}(\alpha)$ scaled about the point $(0,1)$ by $(\alpha, 1-\beta)$ and translated by $(0, \beta)$. The right vertex curve is the original curve scaled about the point $(0,1)$ by $(1-\alpha, \beta)$ and translated by $(\alpha, 0)$.

The conditional locus curve consists of two convex segments. The top left part of the conditional locus curve is the *conditional left locus curve* or *left vertex curve* and the bottom right part the *conditional right locus curve* or *right vertex curve.* Likewise, if Relay B had a fixed rule at α', the conditional locus curve is $F_{\alpha'}(\phi) = \inf_\alpha F_{\alpha,\alpha'}(\phi)$. Finally, the final fused error curve at the fusion point is the convex envelope of $F_\alpha(\phi)$ over all α or the convex envelope of $F_{\alpha'}(\phi)$ over all α' which is $F(\phi) = \inf_\alpha F_\alpha(\phi) = \inf_{\alpha'} F_{\alpha'}(\phi)$.

At the fusion point, the optimal decoding rule for the left vertex curve is the "OR" decoding rule. As (α', β') varies from $(0,1)$ to $(1,0)$, ϕ varies from $(0,1)$ to (α, β). The left vertex curve in figure 3-4 can be represented by

$$(\phi, F_\alpha(\phi)) = \left(\alpha\alpha', \beta + (1-\beta)\beta'\right) \quad \text{for} \quad 0 < \alpha' < 1. \tag{3.14}$$

By letting $\phi = \alpha\alpha'$, this conditional left locus curve representation for a given α in $\inf_{\alpha'} F_{\alpha,\alpha'}(\phi)$ can also be represented algebraically as

$$(\phi, F_\alpha(\phi)) = \left(\phi, \beta + (1-\beta)\hat{\beta}\left(\frac{\phi}{\alpha}\right)\right) \quad \text{for} \quad 0 < \phi < \alpha. \tag{3.15}$$

Likewise, the 'AND" rule is the optimal decoder at the fusion point for the conditional right locus curve in figure 3-4. As (α', β') varies from $(0,1)$ to $(1,0)$, ϕ varies from (α, β) to $(1,0)$ and we have

$$(\phi, F_\alpha(\phi)) = \left(\alpha + (1-\alpha)\alpha', \beta\beta'\right) \quad \text{for} \quad 0 < \alpha' < 1. \tag{3.16}$$

By letting $\phi = \alpha + (1-\alpha)\alpha'$, an alternative algebraic representation is

$$(\phi, F_\alpha(\phi)) = \left(\phi, \beta\hat{\beta}\left(\frac{\phi - \alpha}{1-\alpha}\right)\right) \quad \text{for} \quad \alpha < \phi < 1. \tag{3.17}$$

Note that it is not necessarily true that the conditional locus curve $F_\alpha(\phi)$ lies completely beneath the relay error curve $\hat{\beta}(\alpha)$.

The conditional locus curve representation and viewpoint is useful in many ways. One is to find the minimizing α' for a given α and η. Graphically, as shown in 3-5, it is clear that this minimizing α' is the α' which corresponds to the η-tangent solution to $F_\alpha(\phi)$. If the point of tangency is the left vertex curve, then α' corresponds to the η-tangent solution to $\hat{\beta}(\alpha)$ which is $(\tilde{\alpha}(\eta), \tilde{\beta}(\eta))$ scaled by $(\alpha, 1-\beta)$. In other words, the α' solution to $\frac{d}{d\phi}F_\alpha(\phi) = \eta$ of the conditional left locus curve is the α' which satisfies

$$-\frac{d\hat{\beta}(\alpha')}{d\alpha} = \frac{\alpha}{1-\beta}\eta \tag{3.18}$$

Likewise, if the η-tangent solution to $F_\alpha(\phi)$ is the right vertex curve, then the minimizing α' corresponds to the η-tangent solution to $\hat{\beta}(\alpha)$ which is $(\tilde{\alpha}(\eta), \tilde{\beta}(\eta))$ scaled by $(1-\alpha, \beta)$. The α' solution to $\frac{d}{d\phi}F_\alpha(\phi) = \eta$ of the conditional right locus curve satisfies

$$-\frac{d\hat{\beta}(\alpha')}{d\alpha} = \frac{1-\alpha}{\beta}\eta \tag{3.19}$$

We will use this extensively in the section on finding the optimal solution in this chapter.

3.1.3 Is α or α' a Vertex Point of $F_{\alpha,\alpha'}(\phi)$?

For a relay test at error point (α, β), the magnitude of the "left slope" from $(0,1)$ to point (α, β) is $\lambda^{(\alpha)}(0) = \frac{1-\beta}{\alpha}$. This represents the LR of the relay sending a "0" for a threshold test at (α, β). The magnitude of the right slope from (α, β) to $(1,0)$, $\lambda^{(\alpha)}(1) = \frac{\beta}{1-\alpha}$, is the LR of the relay sending a "1" for a threshold test at (α, β). Denote $V_i \rightarrow V_j$ as the linear segment from point V_i to point V_j. Define $\Lambda(V_i \rightarrow V_j)$ to be the magnitude of the slope of the line connecting

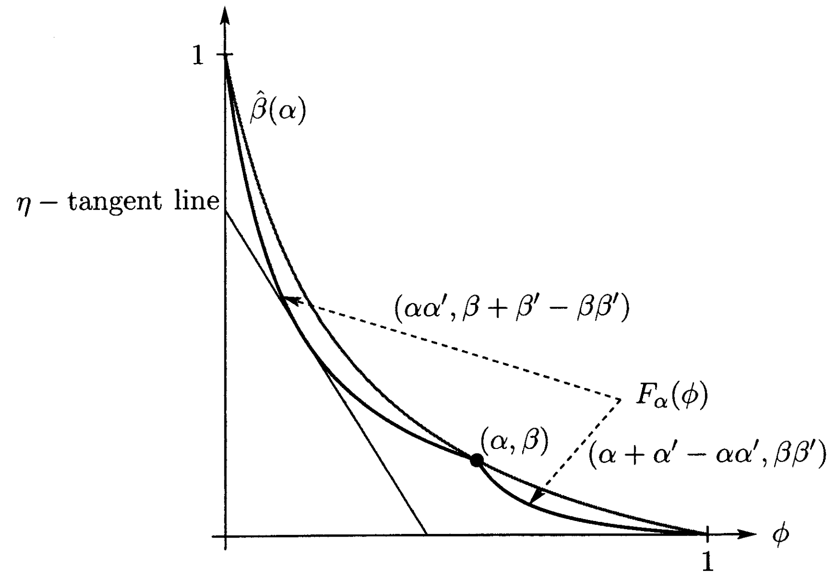


Figure 3-5: For a given η and for a fixed α, the α' which minimizes error probability is the η-tangent solution to the conditional locus curve $F_\alpha(\phi)$. The point of tangency may be on the left left vertex curve or on the right vertex curve.

point V_i and V_j. Then, $\Lambda(V_i \to V_j) = \frac{\beta_i - \beta_j}{\alpha_j - \alpha_i}$ and $V_i \to V_j \to V_k$ is the connected piecewise linear segment from point V_i to point V_j to point V_k.

Suppose that Relay A and Relay B use threshold tests $V_\alpha = (\alpha, \beta)$ and $V_{\alpha'} = (\alpha', \beta')$, respectively. At the fusion point, the optimal fused curve $F_{\alpha,\alpha'}(\phi)$ can take on only 3 possibilities, first, $(0,1) \to V^L_{\alpha,\alpha'} \to (\alpha, \beta) \to V^R_{\alpha,\alpha'} \to (1,0)$, second, $(0,1) \to V^L_{\alpha,\alpha'} \to (\alpha', \beta') \to V^R_{\alpha,\alpha'} \to (1,0)$, and third, $(0,1) \to V^L_{\alpha,\alpha'} \to V^R_{\alpha,\alpha'} \to (1,0)$. The error curve is convex so either $V_\alpha = (\alpha, \beta)$, $V_{\alpha'} = (\alpha', \beta')$, or both will be on $F_{\alpha,\alpha'}(\phi)$. More precisely, if $\Lambda(V^L_{\alpha,\alpha'} \to V_\alpha) > \Lambda(V^L_{\alpha,\alpha'} \to V_{\alpha'})$, then V_α is a vertex point and $F_{\alpha,\alpha'}(\phi)$ is the piecewise linear curve $(0,1) \to V^L_{\alpha\alpha'} \to V_\alpha \to V^R_{\alpha\alpha'} \to (1,0)$. If $\Lambda(V^L_{\alpha,\alpha'} \to V_\alpha) < \Lambda(V^L_{\alpha,\alpha'} \to V_{\alpha'})$, then $V_{\alpha'}$ is an vertex point and $F_{\alpha,\alpha'}(\phi) = (0,1) \to V^L_{\alpha\alpha'} \to V_{\alpha'} \to V^R_{\alpha\alpha'} \to (1,0)$. Observe that $\Lambda(V^L_{\alpha,\alpha'} \to V_{\alpha'}) = \Lambda(V_\alpha \to V^R_{\alpha,\alpha'})$ and $\Lambda(V^L_{\alpha,\alpha'} \to V_\alpha) = \Lambda(V_{\alpha'}) \to V^R_{\alpha,\alpha'}$. This creates a parallelogram with vertices $V^L_{\alpha,\alpha'}, V_\alpha, V_{\alpha'}, V^R_{\alpha,\alpha'}$ as depicted in figure 3-6. When $\Lambda(V^L_{\alpha,\alpha'} \to V_\alpha) = \Lambda(V^L_{\alpha,\alpha'} \to V_{\alpha'})$, then both V_α and $V_{\alpha'}$ are colinear with $V^L_{\alpha,\alpha'}$ and $V^R_{\alpha\alpha'}$ and are not vertex points. In this case, $F_{\alpha,\alpha'}(\phi)$ is the piecewise linear curve $(0,1) \to V^L_{\alpha\alpha'} \to V^R_{\alpha\alpha'} \to (1,0)$.

We will describe the conditions under which each of the 3 cases described above occurs. Note that $\Lambda(V^L_{\alpha,\alpha'} \to V_\alpha) = \lambda^{(a)}(0)\lambda^{(a')}(1) = \left(\frac{1-\beta}{\alpha}\right)\left(\frac{\beta'}{1-\alpha'}\right)$ is the LR of receiving a $(0,1)$ given that Relay A uses α and Relay B uses α'. This gives the following inequality which determines which

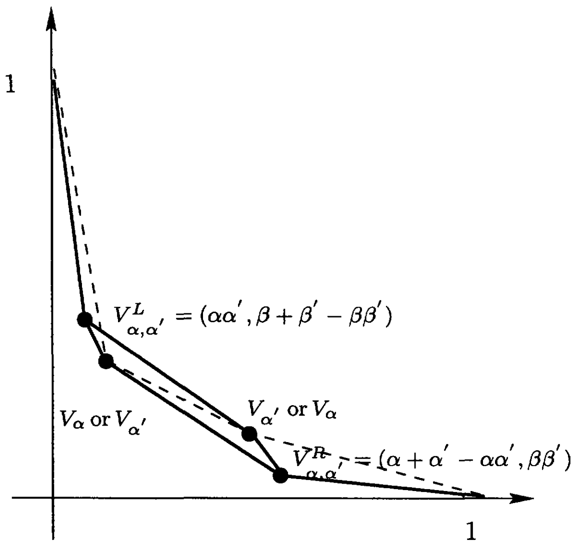


Figure 3-6: Fused curve for Relay A using threshold rule V_α and Relay B using threshold rule $V_{\alpha'}$. The dotted line represents the relay error curve $\hat{\beta}(\alpha)$.

point is the vertex point,

$$\Lambda(V^L_{\alpha,\alpha'} \to V_\alpha) = \left(\frac{1-\beta}{\alpha}\right)\left(\frac{\beta'}{1-\alpha'}\right) \underset{V_{\alpha'}}{\overset{V_\alpha}{\gtrless}} \left(\frac{\beta}{1-\alpha}\right)\left(\frac{1-\beta'}{\alpha'}\right) = \Lambda(V^L_{\alpha,\alpha'} \to V_{\alpha'}) \tag{3.20}$$

For a fixed V_α, we solve the above inequality to determine the admissible regions for $V_{\alpha'}$ such that V_α is on the fused curve $F_{\alpha,\alpha'}(\phi)$. In the opposite inequality, $V_{\alpha'}$ lies on the fused curve. At equality, both V_α and $V_{\alpha'}$ are on the fused curve. Rearranging equation 3.20 and solving for β' as a function of $V_\alpha = (\alpha, \beta)$ and α' we have

$$\beta' \underset{V_{\alpha'}}{\overset{V_\alpha}{\gtrless}} \frac{(1-\alpha')}{(1-\alpha') + \alpha'\left(\frac{1-\alpha}{\beta}\right)\left(\frac{1-\beta}{\alpha}\right)} \tag{3.21}$$

For $\alpha' \geq \alpha$, at equality for equation 3.20, $\hat{\beta}(\alpha')$ is a non-linear convex curve from the point V_α to $(1,0)$. When $V_{\alpha'}$ lies on this curve, both V_α and $V_{\alpha'}$ lie on $F_{\alpha,\alpha'}(\phi)$ but neither are vertex points. For a given V_α, figure 3-7 shows the regions of $V_{\alpha'}$ in the 3 different cases. Note that when $V_{\alpha'}$ lies on the left dotted line (it is colinear with $(0,1)$ and V_α), then $V_{\alpha'}$ is a vertex point of $F_{\alpha,\alpha'}(\phi)$; if $V_{\alpha'}$ lies on the right dotted line (colinear with V_α and $(1,0)$), then V_α is the vertex point of $F_{\alpha,\alpha'}(\phi)$. It follows that the boundary curve in which both V_α and $V_{\alpha'}$ are vertex points must lie inside the triangle of permissible values of $V_{\alpha'}$. When $V_{\alpha'}$ is below this boundary curve, $V_{\alpha'}$ is a vertex point of $F_{\alpha,\alpha'}(\phi)$ and when $V_{\alpha'}$ is above this curve, V_α is a vertex point of $F_{\alpha,\alpha'}(\phi)$.

In general, if the relay error curve is symmetric and if Relay A and Relay B use symmetric points, $(\alpha, \beta) = (a, b)$ and $(\alpha', \beta') = (b, a)$, then both α and α' are on the fused curve $F_{\alpha,\alpha'}(\phi)$ but neither is a vertex point.

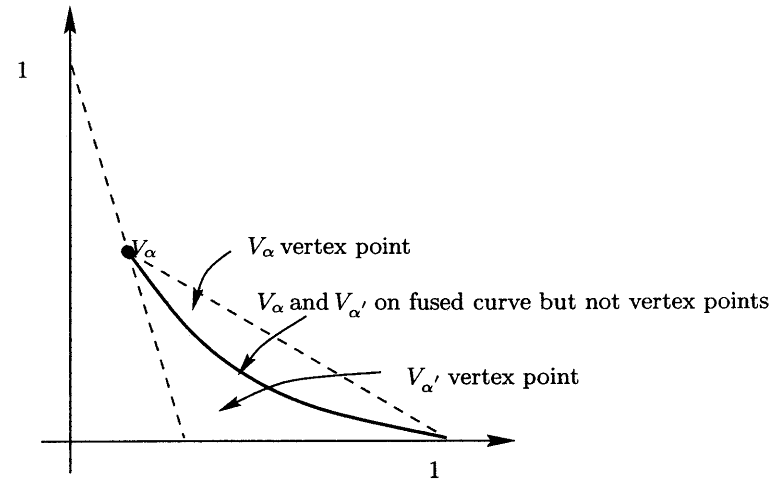


Figure 3-7: The above curve shows for each possible value of $V_{\alpha'}$, relative to a given V_α, whether V_α or $V_{\alpha'}$, or both are on $F_{\alpha,\alpha'}(\phi)$.

For the Gaussian relay channel, the encoding point furthest away from the middle point of the Gaussian error curve $\alpha = \beta = \mathcal{Q}(m)$ is the vertex point of the fused curve $F_{\alpha,\alpha'}(\phi)$. This result will be proven in the last section of this chapter.

3.2 Randomized Strategies and Discrete Memoryless Relay Channels

Every discrete memoryless channel has a piecewise linear error curve which we refer to as a discrete error curve. The canonical channel for a piecewise linear error curve is a binary input DMC with outputs labeled in decreasing LR. The discrete relay error curve merits some special attention, since it gives insight into the nature of the fusion process. Moreover, every continuous relay error curve can be approximated as closely as desired by a discrete relay error curve. On the flip side, a piecewise linear error curve can also be approximated by a strictly convex error curve which, for most situations of concern, still captures the characteristics of the discrete case. Examples and counter-examples from both piecewise linear and strictly convex error curves are mutually beneficial to our understanding of this fusion problem.

Randomized strategies can never improve MAP error probability because averaging over a set of error probabilities can never be smaller than the smallest error probability in the set. Any randomized strategy at the relays can be taken from the viewpoint of a special case of randomized strategies at the fusion point. This concept will be clarified in the following discussion. Randomized strategies are sometimes needed in the Neyman-Pearson problem (by achieving points on the linear segments of the error curve) problem), but are never needed in the MAP or Minimum Cost problems.

As discussed earlier, non-vertex points (the linear regions of the error curve) on the error curve can be viewed as randomized threshold tests between two vertex points. A randomized rule is also equivalent to time sharing among different strategies. In a system with relays, the overall performance can be viewed at the fusion point. When the relays use randomized strategies, the MAP performance depends on whether the strategies are coordinated and whether the fusion point knows the randomized choice. These distinctions are clarified in detail below. We will show the surprising result in this section that it is never necessary to randomize at the relays unless there is a need for a coordinated randomized strategy to achieve all the points on the final fused error curve. In other words, the vertex points of the relay error curve $\hat{\beta}(\alpha)$ are sufficient to achieve every vertex point on the final fused curve $F(\phi)$. It follows that a discrete relay error curve will yield a discrete final fused curve. We further show, except for one very special case, that independent randomization by both relays is strictly suboptimal.

A randomized strategy at the relay means that a relay randomizes between 2 or more vertex points. In all cases, the fusion point and the relays are assumed to know the set of randomizing points and the randomizing distribution. There are 3 types of randomized strategies which can by typified by the following cases.

(1) *Coordinated Randomized Strategy* - a randomized strategy on the final fused curve $F(\phi)$ which corresponds to the relays randomizing between 2 fixed rules on the relay error curve $\hat{\beta}(\alpha)$ in a coordinated fashion. For any two points ϕ_1 and ϕ_2 on $F(\phi)$ and any μ, $0 < \mu < 1$, this strategy achieves the value $F(\mu\phi_1 + (1-\mu)\phi_2) = \mu F(\phi_1) + (1-\mu)F(\phi_2)$. The implementation is as follows. Assume ϕ_1 is achieved by the tests α_1 and α_1' on the two relays and ϕ_2 by tests α_2 and α_2'. Then with probability μ, Relay A uses α_1 and Relay B uses α_1', and the fusion point uses the optimal test to achieve $F(\phi_1)$. With probability $1-\mu$, Relay A uses α_2 and Relay B uses α_2', and the fusion point uses the optimal test to achieve $F(\phi_2)$. In the end, the fusion point has used $F(\phi_1)$ with probability μ and $F(\phi_2)$ with probability $1-\mu$. This randomized test is coordinated because the fusion point and

both relays observe the outcome of the random variable μ and exactly what relay tests are being used for each outcome.

(2) *Independent relay randomization where the fusion point has knowledge of which sample rule has been selected by each relay.* The choice made by the first relay is independent of the choice of the second relay, and the fusion point has knowledge of which choice the relays make at each time. That is, the fusion point always knows the relay encoding points used by each relay, and therefore, can always implement the optimal decoder ("AND," "OR," or Trivial Decoder).

(3) *Independent relay randomization where the fusion point has* **no** *knowledge of which sample rule is selected by each relay.* The choice made by the first relay is independent of the choice of the second relay. However, the fusion point has no knowledge of which choice the relays make at each time. That is, the fusion point knows the distribution and set of choices, but not the sample value being chosen. Without knowledge of the sample rule, the fusion point does not know whether the optimal decoding rule is the left vertex ("OR") or right vertex ("AND") point or the trivial decoder, so the fusion point must implement some convex combination of all the possible opitmal decoders associated with the set of relay choices. From the fusion point perspective, this situation is equivalent to just knowing the relay encoding points at the average value of their randomized vertex points.

The third type of randomization has worse performance than the second type since the fusion point can always ignore the extra information about the relay choices. These 3 different types of randomized strategies listed above are illustrated with the following BEC example with ML decoding. Theorem 3.2.2 specifies the conditions for all the special cases when each of the 3 types of randomization is optimal.

Example: *The BEC with ML decoding.* Recall from chapter 2 that the final fused curve $F(\phi) = (0,1) \to (0,\epsilon^2) \to (\epsilon^2,0) \to (1,0)$ as in figure 2-8. We examine the case when the relays randomize equally between $(0,\epsilon)$ and $(\epsilon,0)$.

(1) *Coordinated randomized strategy* - To achieve any point on the extended error curve of $F(\phi)$, e.g., the linear segment between $(0,\epsilon^2)$ and $(\epsilon^2,0)$, Relay A and B will use the same rule simultaneously. They will both use $(0,\epsilon)$ which leads to the fusion error curve $(0,1) \to (0,\epsilon^2) \to (1,0)$ or they both use $(\epsilon,0)$ which leads to the fusion error curve $(0,1) \to (\epsilon^2,0) \to (1,0)$. Note that each point on the line between $(0,\epsilon^2)$ and $(\epsilon^2,0)$ can be achieved by a coordinated

randomized strategy between $(0,\epsilon)$ and $(\epsilon,0)$ with the appropriate distribution. The resulting error probability for each such strategy is $P_e = \frac{\epsilon^2}{2}$.

(2) *Independent Relay Randomization - The fusion point knows which choice is made by each relay.* Relays A and B independently randomize with equal probability between $(0,\epsilon)$ and $(\epsilon,0)$. The relays will be using the same threshold tests $\{((0,\epsilon),(0,\epsilon))\}$ and $\{((\epsilon,0),(\epsilon,0))\}$ with probability 1/2 and different threshold tests $\{((0,\epsilon),(\epsilon,0))\}$ and $\{((\epsilon,0),(0,\epsilon))\}$ with probability 1/2. The overall error probability $P_e = \frac{1}{4}(2P_e(\text{same rule}) + 2P_e(\text{different rule})) = \frac{1}{2}\left(\frac{\epsilon^2}{2} + \frac{\epsilon}{2}\right) = \frac{\epsilon(1+\epsilon)}{4}$.

(3) *Independent Relay Randomization - The fusion point does not know which choice is selected by each relay.* Relay A and B again independently randomize equally between $(0,\epsilon)$ and $(\epsilon,0)$. This is equivalent to both relays simply using the error point at $(\frac{\epsilon}{2},\frac{\epsilon}{2})$, which is a BSC with crossover probability $\frac{\epsilon}{2}$. The error probability is $P_e = \frac{\epsilon}{2}$.

As expected, the error probability of (1) is optimal and lower than the error probability in (2) which in turn is lower than the error probability of (3).

We now show that only the vertex points of the relay error curve $\hat{\beta}(\alpha)$ are necessary for MAP optimality at the fusion point. This means that vertex points on the relay error curve are sufficient to generate the entire final fused error curve $F(\phi)$, and thus, independent randomized strategies at the relays are never needed.

From theorem 3.1.2, the final fused curve $F(\phi)$ is the convex envelope of all the fused left and right vertex points generated by every possible error point on the relay error curve. Recall that relay vertex points are the non-randomized threshold tests at the relays. We will show in the following theorem that all fused points lie in the convex hull of all the fused relay vertex points. This implies that all the vertex points of $F(\phi)$ are determined by the vertex points on the relay error curve. The rest of the error points on the final fused error curve $F(\phi)$ are achieved through a coordinated randomized strategy of the adjacent vertex points of $F(\phi)$.

Theorem 3.2.1 *Each vertex point of the final fused curve $F(\phi)$ is either $V^L_{\alpha,\alpha'}$ or $V^R_{\alpha,\alpha'}$ for some vertex points α and α' of the relay error curve $\hat{\beta}(\alpha)$.*

Proof: Consider each relay with a strategy on its error curve (theorem 3.1.1), either a vertex point or a randomized point. For the the randomized case, if the theorem is true when the fusion point has knowledge of the relay choice, it must also hold true when the fusion point has no knowledge of the choice. Thus, we will prove the theorem for the former.

Suppose Relay A uses V_i with probability μ and V_j with probability $1 - \mu$ and independently, Relay B uses V_k with probability $\mu^{'}$ and V_l with probability $1-\mu^{'}$. Assume that all of the vertex points are distinct ($V_i \neq V_j$ and $V_k \neq V_l$). At a given η, let $P_e^{\eta}(V_r, V_s)$ be the error probability at the fusion point for a MAP test given that Relay A uses error point V_r and Relay B error point V_s. Note that $P_e^{\eta}(V_r, V_s)$ is the η-tangent solution to $F_{\alpha_r,\alpha_s}(\phi)$ and will be either a left vertex point V_{rs}^L or a right vertex point V_{rs}^R or both (theorem 3.1.2). At the fusion point, the total error probability for a MAP test at η is

$$P_e^{\eta} = \mu\mu^{'}P_e^{\eta}(V_i, V_k)+\mu(1-\mu^{'})P_e^{\eta}(V_i, V_l)+(1-\mu)\mu^{'}P_e^{\eta}(V_j, V_k)+(1-\mu)(1-\mu^{'})P_e^{\eta}(V_j, V_l) \quad (3.22)$$

which is a convex combination of the 4 fused vertex points $\{V_{ik}, V_{il}, V_{jk}, V_{jl}\}$. The overall error probability cannot be smaller than the minimum of the error probabilities of the 4 fused relay vertex points and thus,

$$P_e^{\eta} \geq \min\{P_e^{\eta}(V_i, V_k), P_e^{\eta}(V_i, V_l), P_e^{\eta}(V_j, V_k), P_e^{\eta}(V_j, V_l)\} \quad (3.23)$$

We have just shown that each MAP test at the fusion point can be achieved by each relay using a vertex point (non-randomized threshold test). This implies that randomization is never necessary for MAP optimization. □

We show in Appendix A that when μ and $\mu^{'}$ are not 0 or 1 (randomizing in the strict sense), the inequality in 3.23 is strict. This is done by showing that the fused points $\{V_{ik}, V_{il}, V_{jk}, V_{jl}\}$, with some choice of left or right vertex point, cannot all be colinear on the final fused curve, thus, demonstrating that the error probabilities cannot all be equal.

Corollary 3.2.1 *If the input to each relay is the output of a binary input finite output DMC, then the optimal fused error curve $F(\phi)$ is piecewise linear.*

Proof: From theorem 3.2.1, the vertex points of the relay error curve $\hat{\beta}(\alpha)$ determine the final fused error curve $F(\phi)$. Since there are a finite number of vertex points on the relay error curve, then the number of vertex points on the final fused error curve $F(\phi)$ is also finite. This implies that $F(\phi)$ is piecewise linear. □

Therefore, fusing piecewise linear relay error curves results in a piecewise linear final fused curve. However, a strictly convex relay error curve will not necessarily have a strictly convex final fused curve. The additive Gaussian noise channel discussed at the end of this chapter is an example.

Recall that theorem 3.1.1 stated that the optimal rule at each relay is an error point for MAP decoding at the fusion point. We now give some added insight and additional perspective why this must be true.

We will show that it is strictly suboptimal if Relay A uses some arbitrary test $V_p = (\alpha_p, \beta_p)$ **not** on the relay error curve and Relay B uses a point V_l on the error curve. Recall that V_p cannot lie below the relay error curve, only at or above the error curve and must satisfy $\alpha_p + \beta_p \leq 1$. Then V_p can be written as a convex combination of three distinct vertex points, say V_i, V_j, V_k chosen from the error curve. Thus, for some $0 < \lambda, \mu < 1$,

$$V_p = \mu[\lambda V_i + (1-\lambda)V_j] + (1-\mu)V_k \tag{3.24}$$

Note that λ and μ cannot both be 0 or 1; otherwise, V_p will be a vertex point which violates the assumption. Now fuse V_p with V_l and the left vertex point becomes

$$V_{p,l}^L = \mu[\lambda V_{i,l}^L + (1-\lambda)V_{j,l}^L] + (1-\mu)V_{k,l}^L \tag{3.25}$$

which is a convex combination of 3 distinct vertex points on the fused curve $F(\phi)$. These vertex points are not colinear since $\Lambda(V_{il}^L \to V_{jl}^L) \neq \Lambda(V_{jl}^L \to V_{kl}^L)$. Therefore, $V_{p,l}^L$ cannot possibly be optimal since the convex combination of three vertex points lies in the convex hull of the three points and thus, cannot lie on its convex envelope . Thus, it cannot be a vertex point of $F(\phi)$. The same argument shows that the right vertex point $V_{p,l}^R$ cannot be optimal.

In general, independent relay randomization is strictly suboptimal. However, it is of some interest to find if independent randomized rules at the relays are ever optimal. It turns out that there are just two special cases. There is one special case of optimality when both relays randomize and one special case of optimality when only one relay uses a randomized rule. All other cases of independent randomization are strictly suboptimal. For completeness, we discuss all the special cases when randomized strategies are optimal in the following theorem.

Theorem 3.2.2 *All the special cases when randomized strategies at the relays are optimal.*

(1) **Coordinated Randomized Strategy.** *If the fused points of the coordinated randomized relay strategies correspond to adjacent vertex points of $F(\phi)$, then the coordinated*

randomized strategy of those vertex points on the relay error curve is optimal.

(2) **Independent relay randomization when the fusion point has knowledge of the relay choices.** *Performance is strictly suboptimal, except in the following cases: (A) If both relays independently randomize between the same 2 vertex points on the relay error curve, and the resulting fused vertex points (with some choice of left or right vertex) all lie on a linear portion of $F(\phi)$. (B) When one relay rule has a fixed rule at a vertex point, the other relay independently randomizes between 2 vertex points, and the fused vertex points (with some choice of left or right vertex point) are adjacent vertex points on $F(\phi)$.*

(3) **Independent relay randomization and the fusion point has <u>no</u> knowledge of the relay choices.** *Performance is strictly suboptimal, except in the following case: When one relay is fixed at a vertex point, the other relay independently randomizes between 2 vertex points, and the fused vertex points are adjacent vertex points on $F(\phi)$ and are of the same type (either both left vertex points or both right vertex points).*

Proof: (1) Randomizing between any two adjacent vertex points on $F(\phi)$ is optimal because the average value is on $F(\phi)$. If V_{ij} and V_{kl} are adjacent vertex points of $F(\phi)$, then any point on the error curve, $\mu V_{ij} + (1-\mu)V_{kl}$ for $0 < \mu < 1$ is achieved by the following coordinated randomized strategy. With probability μ, Relay A uses V_i while Relay B uses V_j, and with probability $(1-\mu)$, Relay A uses V_k while Relay B uses V_l. The fusion point chooses the optimal decoding rule corresponding to each vertex point of $F(\phi)$. □

Proof: (2A) The adjacent vertex points on $F(\phi)$ can be either a left vertex or right vertex point. When the fusion point knows the relay selection, it can choose the fused left or right vertex point for optimal decoding. Like the proof in theorem 3.2.1, the total error probability is a convex combination of the error probabilities of 3 vertex points (if both relays randomize between the same 2 vertex points) and 4 vertex points (if the combined set of vertex points being randomized is 3 or 4). We will show in Appendix A that if each relay randomization is between the same two vertex points, there is a possibility of optimality. Appendix A also shows that when both relays independently randomize on a combined set of greater than two vertex points, performance is strictly suboptimal.

(2B) Suppose Relay A has fixed threshold test at V_i and the other Relay B independently randomizes between vertex points V_j and V_k. If V_{ij} and V_{ik} are adjacent vertex points on $F(\phi)$, with some choice of left vertex or right vertex, then the independent relay strategy of Relay B becomes the same as the coordinated random strategy, and thus is optimal. $\square$

Proof: (3) To prove that independently randomization by both relays is strictly suboptimal, we will show that the fusion of any two arbitrary non-vertex points q and q' on the relay error curve will lie inside the convex hull of the fused points generated from the relay vertex points neighboring q and q'. Note that these fused relay vertex points may or may not be on the final fused curve $F(\phi)$. Any two arbitrary points q and q' chosen on the linear segment of an error curve can be written as a convex combination of two vertex points. Without loss of generality, let $q = \mu V_i + (1-\mu)V_j$ and $q' = \mu' V_k + (1-\mu')V_l$. Assume V_i, V_j, V_k, V_l are distinct vertex points. We will show that $V^L_{q,q'}$ and $V^R_{q,q'}$ lies inside the convex envelope of $\{V^L_{ik}, V^L_{il}, V^L_{jk}, V^L_{jl}\}$ and $\{V^R_{ik}, V^R_{il}, V^R_{jk}, V^R_{jl}\}$, respectively. The left vertex point is a convex combination of 4 left vertex points and can be expressed in 2 ways as

$$\text{(i)}\quad V^L_{q,q'} = \mu'[\mu V^L_{ik} + (1-\mu)V^L_{jk}] + (1-\mu')[\mu V^L_{il} + (1-\mu)V^L_{kl}] \tag{3.26}$$

$$\text{(ii)}\quad V^L_{q,q'} = \mu[\mu' V^L_{ik} + (1-\mu')V^L_{il}] + (1-\mu)[\mu' V^L_{jk} + (1-\mu')V^L_{jl}]. \tag{3.27}$$

The fused vertex point $V^L_{q,q'}$ is simply a strict convex combination of the points $\{V^L_{ik}, V^L_{jk}, V^L_{il}, V^L_{kl}\}$ which cannot be colinear (Appendix A). The two cases in 3.26 and 3.27 can be seen geometrically in figure 3-8 for the left vertex point. The same argument applies for right vertex points. Simply relabel "left" to "right" in equations 3.26 and 3.27.

If V_i, V_j, V_k, V_l are not distinct, the above argument is slightly modified. Suppose there are only two distinct vertex points $\{V_i, V_j\}$. Then $q = \mu V_i + (1-\mu)V_j$ and $q' = \mu' V_i + (1-\mu')V_j$. Then the left vertex point $V^L_{qq'}$ and right vertex point $V^R_{qq'}$ will lie, respectively, in the convex hull of the fused points $\{V^L_{ii}, V^L_{ij}, V^L_{jj}\}$ and $\{V^R_{ii}, V^R_{ij}, V^R_{jj}\}$. If there are three distinct vertex points $\{V_i, V_j, V_k\}$ be vertex points, then $q = \mu V_i + (1-\mu)V_j$ and $q' = \mu' V_j + (1-\mu')V_k$ The left and right vertex point will lie, respectively, in the convex hull of $\{V^L_{ij}, V^L_{ik}, V^L_{jj}, V^L_{jk}\}$ and $\{V^R_{ij}, V^R_{ik}, V^R_{jj}, V^R_{jk}\}$. Appendix A shows it is impossible for all the vertex points in these sets to be colinear. Thus, both relays independently randomizing is strictly suboptimal.

We now prove the special case of optimality. As discussed before, the relay randomizing points must be adjacent vertex points on the relay error curve $\hat{\beta}(\alpha)$ and the fused vertex points on $F(\phi)$ must be both right or both left. Therefore, if $V^L_{i,k}$ and $V^L_{i+1,k}$ are adjacent vertex points on $F(\phi)$, then the randomized rule $q = \mu V_i + (1-\mu)V_{i+1}$ at Relay A fused with $q' = V_k$ at Relay B will achieve $\mu V^L_{i,k} + (1-\mu)V^L_{i+1,k}$ on $F(\phi)$ for $0 \le \mu \le 1$. This sole Relay A randomization is equivalent to a coordinated randomized strategy of both relays between $V^L_{i,k}$ and $V^L_{i+1,k}$. The

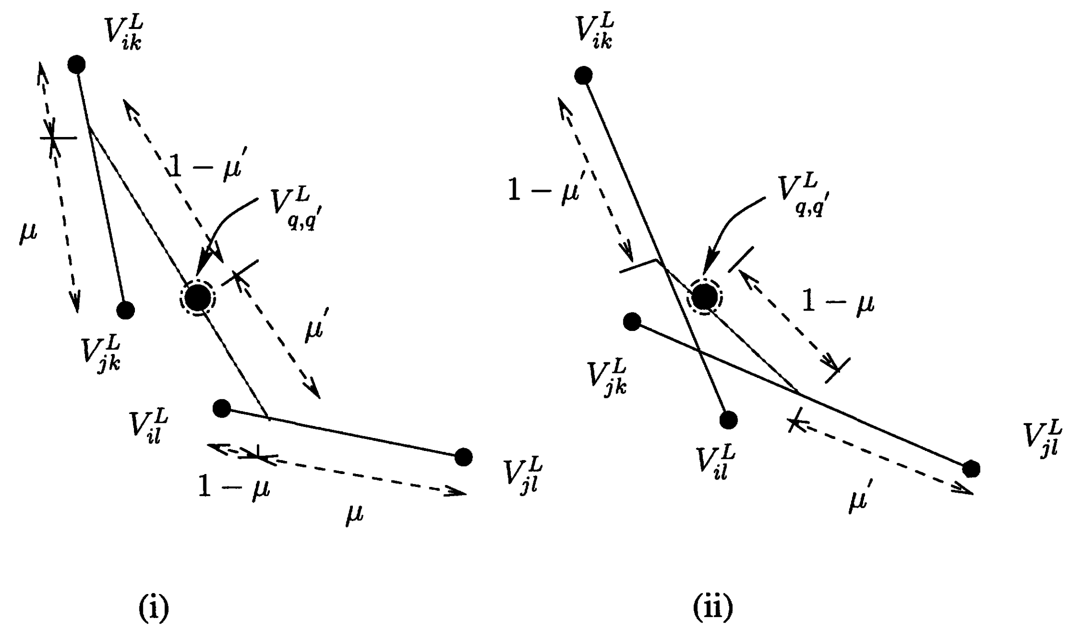


Figure 3-8: Fused left vertex point of independent randomized relay strategies where fusion point does not know sample value. Relay A uses non-vertex point $q = \mu V_i + (1-\mu)V_j$ and Relay B the non-vertex point $q' = \mu' V_k + (1-\mu')V_l$. The left vertex point $V_{q,q'}^L$ is the convex combination of the fusion of at most 4 vertex points, $\{V_{ik}^L, V_{il}^L, V_{jk}^L, V_{jl}^L\}$ and lies inside the convex hull of these vertex points. Same situation for the right vertex point.

same applies to the right vertex points. □

Note that for the special cases when randomization is not sub-optimal, the individual relay randomizing points might be anywhere on the error curve for the coordinated randomized strategy and independent randomization when the fusion point knows the choice. However, for independent relay randomization when the fusion point does not know the choice, the relay randomization must be between adjacent vertex points on the relay error curve to even have a chance at optimality (theorem 3.1.1). Thus, it follows that if every point on the relay error curve is a vertex point, a.k.a. the strictly convex error curve, then independent randomized strategy of the relays is always suboptimal.

When the relays independently randomize between 2 vertex points and the fusion point **knows** the choice, the average encoding point on the relay error curve throws information away. In other words, knowing the relays encoding point average value is not equivalent to knowing the 2 vertex points they are randomizing from. The reason is as follows: since the fusion point knows the vertex points used at every time instant, it can optimally choose the fused right or left vertex point and the associated decoding rule. However, if the fusion point does **not know** the relay choices, the fusion point having only knowledge of average value of the relay randomized points is equivalent to having knowledge of the 2 vertex points they are randomizing from.

It is never necessary to randomize at the relays unless there is a need for a coordinated randomized strategy to achieve the non-vertex points on $F(\phi)$. We have shown all the special cases when it is optimal for independent relay randomized strategies. In particular, it is optimal for one relay to independently randomize occurs when the situation is equivalent to a coordinated randomized strategy that achieves a point on the extended error curve of the final fused curve. The other very special case occurs when both relays independently randomize between the same 2 relay vertex points and the fusion point knows the relay selection. Other than these very special cases, we have also shown that it is otherwise never optimal for both relays to independently randomize.

3.3 Structure of $F(\phi)$ for Discrete Channels

3.3.1 The Two Relay Ternary Alphabet Channel

This section studies the structure of the optimal fused curve $F(\phi)$ at the fusion point when the relay error curve is restricted to 2 vertex points, $V_1 = (\alpha_1, \beta_1)$ and $V_2 = (\alpha_2, \beta_2)$. This means that V_1 is not colinear with $(0, 1)$ and V_2, and also V_2 is not colinear with $(1, 0)$ and V_1. Otherwise, the relay error curve would reduce to the trivial or BAC curve. We also assume no erasure-like channels, which means $\alpha_i \neq 0, \beta_i \neq 0, i = 1, 2$. For the next 2 sections on structure, the results only apply to identical relay channels. There is no simple change of notation or extension to non-identical channels.

Let $V_i < V_j < V_k$ mean that $\alpha_i < \alpha_j < \alpha_k$ and $\beta_k < \beta_j < \beta_i$. Recall that $V_i \to V_j$ denotes the linear segment from points V_i to point V_j and $\Lambda(V_i \to V_j)$ is defined as the magnitude of the slope of the line connecting point V_i and V_j. Then, $\Lambda(V_i \to V_j) = \frac{\beta_i - \beta_j}{\alpha_j - \alpha_i}$ and represents the LR of the output determined by the points V_i and V_j. When the superscripts R or L are left off a fused vertex point, this implies that the statement applies to both L and R.

As discussed earlier in Chapter 2, the LR tells us everything we need to know about the channel in order to solve this distributed optimization problem. The canonical system representation for a two-vertex relay error curve is a DMC with a binary input and ternary output. This is shown in figure 3-9 for a relay error curve with vertices $V_1 = (\alpha_1, \beta_1)$ and $V_2 = (\alpha_2, \beta_2)$.

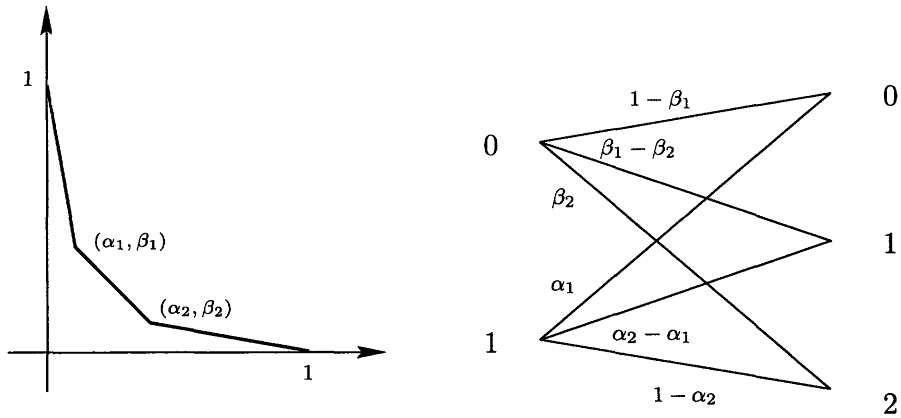


Figure 3-9: Canonical DMC for a error curve with 2 vertex points.

We will study the structure of the optimal fused curve for this simplest non-trivial relay error curve. All the structure rules generated for the 2 vertex error curve apply to any pair of points on an arbitrary relay error curve because the relays either choose the same error point or different error points.

Thus, as demonstrated in the previous section, the *left locus curve of* V_1 is equal to the entire relay error curve scaled by $(\alpha, 1-\beta)$ and translated by $(0, \beta_1)$. The *right locus curve of* V_1 is equal to the entire relay error curve scaled by $(1-\alpha, \beta)$ and translated by $(\alpha_1, 0)$. Since the error curve is convex, all the locus curves are convex. Specifically, the relay error curve with vertex points V_1, V_2 has 4 locus curves with 2 points on each curve, for a total of 8 points (not distinct as some points may be the same). The left and right locus points of V_1 are respectively

$$\begin{aligned} \{[\alpha_1^2, 2\beta_1 - \beta_1^2], [\alpha_1\alpha_2, \beta_1 + \beta_2(1-\beta_1)]\} &= \{V_{11}^L, V_{12}^L\} \\ \{[2\alpha_1 - \alpha_1^2, \beta_1^2], [\alpha_1 + \alpha_2(1-\alpha_1), \beta_1\beta_2]\} &= \{V_{11}^R, V_{12}^R\} \end{aligned} \tag{3.28}$$

The left and right locus points for V_2 are, respectively,

$$\begin{aligned} \{[\alpha_1^2, 2\beta_1 - \beta_1^2], [\alpha_1\alpha_2, \beta_1 + \beta_2(1-\beta_1)]\} &= \{V_{12}^L, V_{22}^L\} \\ \{[2\alpha_1 - \alpha_1^2, \beta_1^2], [\alpha_1 + \alpha_2(1-\alpha_1), \beta_1\beta_2]\} &= \{V_{12}^R, V_{22}^R\} \end{aligned} \tag{3.29}$$

The final fused curve is the convex envelope of the conditional locus curves of V_1 and V_2 in 3.28 and 3.29. There are 8 fused points of which two are repeated, namely V_{12}^L and V_{12}^R. Hence, there are a total of 6 distinct fused points, namely $\{V_{11}^L, V_{12}^L, V_{22}^L, V_{11}^R, V_{12}^R, V_{22}^R\}$, to consider for the convex envelope.

Lemma 3.3.1 *The left most part of the curve $F(\phi)$ is always $(0,1) \to V_{11}^L \to V_{12}^L$ and the right most part is always $V_{12}^R \to V_{22}^R \to (1,0)$.*

Proof: $F(\phi)$ is the convex envelope of all possible convex combinations of vertex points. Construct the fused curve $F(\phi)$ from the 6 possible vertex points. Starting at $(0,1)$, the one which produces the steepest slope from $(0,1)$ is V_{11}^L and it is equal to $\left(\frac{1-\beta_1}{\alpha_1}\right)\left(\frac{1-\beta_1}{\alpha_1}\right)$. The next steepest slope is $\Lambda(V_{11}^L \to V_{12}^L) = \left(\frac{1-\beta_1}{\alpha_1}\right)\left(\frac{\beta_1-\beta_2}{\alpha_2-\alpha_1}\right)$. The next steepest slope will need a comparison of $\Lambda(V_{12}^L \to V_{11}^R)$ with $\Lambda(V_{12}^L \to V_{22}^L)$. Therefore, the beginning or left most part of $F(\phi)$ is the piecewise linear segment $(0,1) \to V_{11}^L \to V_{12}^L$. The right most part of $F(\phi)$ consists of slopes with the smallest magnitude. Looking at the construction of $F(\phi)$ from right to left, starting with $(1,0)$, the slope with the smallest magnitude is $\left(\frac{\beta_1}{1-\alpha_1}\right)\left(\frac{\beta_2}{1-\alpha_2}\right)$ which is $\Lambda(V_{22}^R \to (1,0))$. From the point V_{22}^R the next slope with smallest magnitude is $\Lambda(V_{12}^R \to V_{22}^R) = \left(\frac{\beta_1-\beta_2}{\alpha_2-\alpha_1}\right)\left(\frac{\beta_2}{1-\alpha_2}\right)$. From V_{12}^R, a comparison will need to be made. Thus, the right most part of $F(\phi)$ is the piecewise linear segment $V_{12}^R \to V_{22}^R \to (1,0)$. □

Lemma 3.3.2 *V_{12}^L satisfies $V_{11}^L < V_{12}^L < V_{22}^L$ and the points form a convex curve.*

Proof: We will first show that (1) V_{12}^L lies in between V_{11}^L and V_{22}^L and then (2) show that $\Lambda(V_{11}^L \to V_{12}^L) \geq \Lambda(V_{12}^L \to V_{22}^L)$. (1) Since $\alpha_1 < \alpha_2$, this implies that $\alpha_1^2 < \alpha_1\alpha_2 < \alpha_2^2$. (2) $\Lambda(V_{11}^L \to V_{12}^L) = \left(\frac{1-\beta_1}{\alpha_1}\right)\left(\frac{\beta_1-\beta_2}{\alpha_2-\alpha_1}\right)$ and $\Lambda(V_{12}^L \to V_{22}^L) = \left(\frac{1-\beta_2}{\alpha_2}\right)\left(\frac{\beta_1-\beta_2}{\alpha_2-\alpha_1}\right)$. Due to the convexity of the relay error curve, we have $\frac{1-\beta_i}{\alpha_i} > \frac{1-\beta_j}{\alpha_j}$ for all $i < j$. This implies that $\Lambda(V_{11}^L \to V_{12}^L) > \Lambda(V_{12}^L \to V_{22}^L)$. □

Lemma 3.3.3 *V_{12}^L satisfies $V_{11}^L < V_{12}^L < V_{11}^R$ and the points form a convex curve.*

Proof: As illustrated in figure 3-10 V_{12}^L lies between V_{11}^L and V_1. Now V_1 is between and colinear with V_{11}^L, V_{11}^R, so it follows that V_{12}^L must lie in between V_{11}^L and V_{11}^R. Since

$$\Lambda(V_{11}^L \to V_{12}^L) = \left(\frac{1-\beta_1}{\alpha_1}\right)\left(\frac{\beta_1-\beta_2}{\alpha_2-\alpha_1}\right) > \Lambda(V_{12}^L \to V_1) = \left(\frac{1-\beta_1}{\alpha_1}\right)\left(\frac{\beta_2}{1-\alpha_2}\right) \tag{3.30}$$

and $\Lambda(V_{12}^L \to V_1) = \Lambda(V_{12}^L \to V_{12}^R)$, $V_{11}^L \to V_{12}^L \to V_{11}^R$ is convex. □

Lemma 3.3.4 *$V_{12}^L < V_{22}^L < V_{12}^R$ and $V_{12}^L < V_{11}^R < V_{12}^R$.*

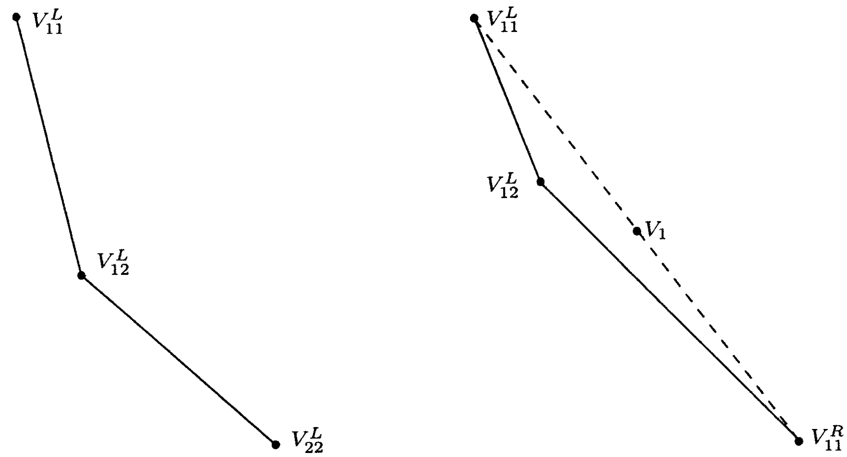


Figure 3-10: Partial structure of 2 point error curve.

Proof: We need to show the following set of inequalities

$$(i) \qquad \alpha_1\alpha_2 < \alpha_2^2 \overset{(a)}{<} \alpha_1 + \alpha_2 - \alpha_1\alpha_2 \tag{3.31}$$

$$(ii) \qquad \beta_1 + \beta_2 - \beta_1\beta_2 \overset{(b)}{>} 2\beta_2 - \beta_2^2 \overset{(c)}{>} \beta_1\beta_2 \tag{3.32}$$

Since $\alpha_1 < \alpha_2$ and $\beta_1 > \beta_2$,

$$(i) \qquad (a) \quad \alpha_1 + \alpha_2 - \alpha_1\alpha_2 - \alpha_2^2 = (\alpha_1 + \alpha_2)(1 - \alpha_2) > 0 \tag{3.33}$$

$$(ii) \qquad (b) \quad \beta_2[(1 - \beta_1) + (1 - \beta_2)] > 0 \tag{3.34}$$

$$(c) \quad \beta_1 + \beta_2 - \beta_1\beta_2 - 2\beta_2 - \beta_2^2 = (1 - \beta_2)(\beta_1 - \beta_2) > 0 \tag{3.35}$$

If one interchanges α to β everywhere above, one finds that $V_{12}^L < V_{11}^R < V_{12}^R$. $\square$

Lemma 3.3.5 *Either V_{11}^R or V_{22}^L or both will be on the final fused curve $F(\phi)$.*

Proof: Recall that $\{V_{12}^L, V_1, V_2, V_{12}^R\}$ always form a parallelogram with V_1 and V_2 at opposite corners and V_{12}^L and V_{12}^R at opposite corners. Both V_1 and V_2 are between V_{12}^L and V_{12}^R. Assume first that $V_{12}^L \rightarrow V_1 \rightarrow V_{12}^R$ form a convex envelope. It follows that $V_{11}^L \rightarrow V_{12}^L \rightarrow V_1 \rightarrow V_{12}^R$ forms a convex envelope . Recall that $\{V_{11}^L, V_1, V_{11}^R\}$ are colinear. It follows that V_{11}^R must lie below the line connecting V_1 and V_{12}^R. Thus, $V_{11}^L \rightarrow V_{12}^L \rightarrow V_{11}^R \rightarrow V_{12}^R$ form a convex envelope. Similarily, if we assume that $V_{12}^L \rightarrow V_2 \rightarrow V_{12}^R$ form a convex envelope, it follows that $V_{12}^L \rightarrow V_{22}^L \rightarrow V_{12}^R \rightarrow V_{22}^R$ forms a convex envelope. Therefore, in both cases, either V_{11}^R or V_{22}^L or both will lie on $F(\phi)$. $\square$

Theorem 3.3.1 *Given a two point relay error curve* $(0,1) \to V_1 \to V_2 \to (1,0)$ *with two relays each transmitting a binary symbol, there exists four types of error curves* $F(\phi)$ *with the following vertex points*[1],

$$(0,1) \to V_{11}^L \to V_{12}^L \to \left\{ \begin{array}{c} V_{11}^R \\ V_{22}^L \\ V_{11}^R \to V_{22}^L \\ V_{22}^L \to V_{11}^R \end{array} \right\} \to V_{12}^R \to V_{22}^R \to (1,0) \tag{3.36}$$

Proof: Combining the above lemmas proves the theorem. □

Corollary 3.3.1 *(Erasure-Like Channels)*

(A) If $V_1 = (0,q)$ *and* $V_2 = (p,0)$*, then the two point relay error curve is* $(0,1) \to (0,q) \to (p,0) \to (1,0)$*. Its final fused curve is* $F(\phi) = (0,1) \to (0,q^2) \to (p^2,0) \to (1,0)$*, which is* $(0,1) \to V_{11}^R \to V_{22}^L \to (1,0)$.

(B) If $V_1 = (\alpha,\beta)$ *and* $V_2 = (p,0)$*, then the right most part of* $F(\phi)$ *is* $V_{22}^L = (p^2,0) \to (1,0)$*. From equation 3.36 the final fused curve can be either of the following*

(i) $F(\phi) = (0,1) \to V_{11}^L \to V_{12}^L \to V_{22}^L \to (1,0)$

(ii) $F(\phi) = (0,1) \to V_{11}^L \to V_{12}^L \to V_{11}^R \to V_{22}^L \to (1,0)$.

Similar results for $V_1 = (0,q)$ *and* $V_2 = (\alpha,\beta)$.

Theorem 3.3.1 and corollary 3.3.1 can also be understood graphically from the viewpoint of section 3.1.2. The final fused curve is the convex envelope of the conditional locus curves of V_1 and V_2. Note that the vertex points of the convex envelope can alternate between left and right.

[1]Note that it is possible that the line segments in each of the 4 cases $\{V_{12}^L \to V_{11}^R \to V_{22}^L, V_{12}^L \to V_{22}^L \to V_{11}^R, V_{11}^R \to V_{22}^L \to V_{12}^R, V_{22}^L \to V_{11}^R \to V_{12}^R\}$ can be colinear. Since the error curve $F(\phi)$ is specified by vertex points, these cases are included in 3.36.

3.3.2 The General Two Relay Discrete Output Channel

The following structure rules apply to any discrete relay error curve with J vertex points. We examine the structure of $F(\phi)$ for two relays with binary outputs each.

Lemma 3.3.6 *For a relay error curve with J vertex points, the left most part of the final fused curve $F(\phi)$ is always $(0,1) \to V_{11}^L \to V_{12}^L$ and the right most part of the fused curve $F(\phi)$ is always $V_{J,J-1}^R \to V_{J,J}^R \to (1,0)$.*

Proof: With a slight modification of the last point as J on the relay error curve the proof is virtually identical to that of lemma 3.3.1. □

Corollary 3.3.2 *(Erasure-Like Channels) If $V_1 = (0,q)$, then the left-most part of $F(\phi)$ is $(0,1) \to V_{11}^R = (0,q^2)$. If $V_1 = (p,0)$, then the right most part of $F(\phi)$ is $V_{J,J}^L = (p^2,0) \to (1,0)$.*

All the possible left vertex points and right vertex points can be represented as elements in 2 $J \times J$ matrices, called respectively, the *Left Convex Envelope Matrix* C^L and the *Right Convex Envelope Matrix* C^R. Since Relays A and B are the same, $V_{ij} = V_{ji}$. Therefore, either the upper right or lower left diagonal triangle of each Convex Envelope Matrix is sufficient for all the fused vertex representations. We choose the convention that $i \le j$ for the element indices, which means we use the upper right triangular part of the Convex Envelope Matrix. Conditional on a vertex point on $F(\phi)$, the lemmas above and below limit the next possible adjacent vertex point. When the LR comparisons are made, the path through the matrix determines the convex envelope for each decoding rule associated with the matrix. Note that once the path has passed a point, it cannot return to any point before that. The matrix of vertex points and the possible paths for the convex envelope are depicted in figure 3-11. Finally, $F(\phi)$ is the convex envelope of the left vertex curve generated by C^L and the the right vertex curve generated from C^R.

The following lemma shows that in the vertex point matrix, the convex envelope can only travel one entry from left to right or one entry down in C^L and C^R. In other words, the path of the vertex points of the convex envelope cannot skip any vertex points when the path is vertical or horizontal.

Lemma 3.3.7 *(1) $\Lambda(V_{ij} \to V_{i,j+1}) > \Lambda(V_{ij} \to V_{i,j+k})$, $\forall k > 1$. and*
(2) $\Lambda(V_{ij} \to V_{i+1,j}) > \Lambda(V_{ij} \to V_{i+k,j})$, $\forall k > 1$.

Proof: (1) Since $\Lambda(V^L_{ij} \to V^L_{ik}) = \left(\frac{1-\beta_i}{\alpha_i}\right)\left(\frac{\beta_j-\beta_k}{\alpha_k-\alpha_j}\right)$ and $\Lambda(V^R_{ij} \to V^R_{ik}) = \left(\frac{\beta_i}{1-\alpha_i}\right)\left(\frac{\beta_j-\beta_k}{\alpha_k-\alpha_j}\right)$, the assertion follows from the convexity of the error curve, $\frac{\beta_{j+k}-\beta_j}{\alpha_{j+k}-\alpha_j} \leq \frac{\beta_{j+1}-\beta_j}{\alpha_{j+1}-\alpha_j}$ for all $k > 1$.
(2) Follows from interchangeability of relays. □

This next lemma shows that the vertex point index in C^L and C^R cannot increase simultaneously for both relays. In other words, the convex envelope can only travel one element horizontal right or one element vertical down in C^L and C^R. It cannot travel diagonal down to the right.

Lemma 3.3.8 *(1)* $\Lambda(V^L_{ij} \to V^L_{i,j+1}) > \Lambda(V^L_{ij} \to V^L_{i+1,j+1})$ *and*
(2) $\Lambda(V^R_{ij} \to V^R_{i,j+1}) > \Lambda(V^R_{ij} \to V^R_{i+1,j+1})$.

Proof: (1) Computing the slopes from one vertex point to the next vertex point in the left convex envelope matrix,

$$\Lambda(V^L_{ij} \to V^L_{i,j+1}) = \left(\frac{1-\beta_i}{\alpha_i}\right)\left(\frac{\beta_j-\beta_{j+1}}{\alpha_{j+1}-\alpha_j}\right) \tag{3.37}$$

$$\Lambda(V^L_{ij} \to V^L_{i+1,j+1}) = \left(\frac{(1-\beta_j)+(1-\beta_{j+1})}{\alpha_j+\alpha_{j+1}}\right)\left(\frac{\beta_j-\beta_{j+1}}{\alpha_{j+1}-\alpha_j}\right) \tag{3.38}$$

Comparing 3.37 and 3.38, we need to show $\frac{(1-\beta_j)+(1-\beta_{j+1})}{\alpha_j+\alpha_{j+1}} < \frac{1-\beta_j}{\alpha_j}$. Note that if $a, b, c, d > 0$ and if $\frac{a}{b} > \frac{c}{d}$, then $\frac{c}{d} < \frac{a+c}{b+d} < \frac{a}{b}$. Let $a = 1-\beta_j, b = \alpha_j, c = 1-\beta_{j+1}, d = \alpha_{j+1}$ and the assertion is proved.

For (2), the right convex envelope matrix,

$$\Lambda(V^R_{ij} \to V^R_{i,j+1}) = \left(\frac{\beta_i}{1-\alpha_i}\right)\left(\frac{\beta_j-\beta_{j+1}}{\alpha_{j+1}-\alpha_j}\right) \tag{3.39}$$

$$\Lambda(V^R_{ij} \to V^R_{i+1,j+1}) = \left(\frac{\beta_i}{1-\alpha_i}\right)\left(\frac{\beta_j-\beta_{j+1}}{\alpha_{j+1}-\alpha_j}\right). \tag{3.40}$$

We just need to show $\frac{\beta_j+\beta_{j+1}}{(1-\alpha_j)+(1-\alpha_{j+1})} < \frac{\beta_j}{1-\alpha_j}$. Same argument as in (1). Just let $a = \beta_j, b = 1-\alpha_j, c = \beta_{j+1}, d = 1-\alpha_{j+1}$. □

The following rules summarize the lemmas above.

<u>Allowed</u>:

(a) One element to the right. $V_{ij} \to V_{i,j+1}$

(b) One element down. $V_{ij} \rightarrow V_{i+1,j}$

(c) Diagonal up and right. $V_{ij} \rightarrow V_{i-l,j+k}$, for $k, l \geq 1$.

(d) Diagonal down and left. $V_{ij} \rightarrow V_{i+l,j-k}$, for $k, l \geq 1$.

<u>Not allowed</u>:

(a) Skipping horizontally or vertically

(b) Diagonal down and right.

(c) Diagonal up and left.

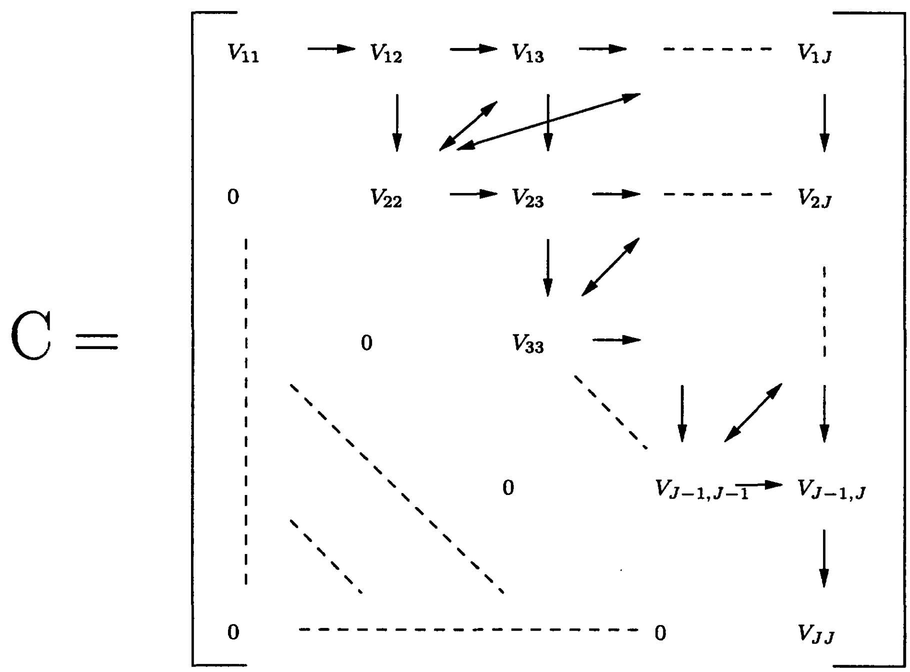


Figure 3-11: Convex Envelope Matrix of allowed successive point for the right vertex points and left vertex points of a discrete error curve. The convex envelope of the piecewise linear left curve and right piecewise linear curve is the final fused curve $F(\phi)$.

As a specific example, for the relay error curve with 2 vertex points, the dimension of the convex envelope matrix is 2×2. The path is fixed, namely it is $V_{11}^L \rightarrow V_{12}^L \rightarrow V_{22}^L$ for the left convex envelope matrix and $V_{11}^R \rightarrow V_{12}^R \rightarrow V_{22}^R$ for the right convex envelope matrix. Finally, the convex envelope of both curves results in the final fused curve $F(\phi)$. In the 2×2 case, the only points that need to be compared (to determine which point lies on the convex envelope for $F(\phi)$) are the points V_{22}^L and V_{11}^R. The comparisons result in 4 different types of error curves as seen in theorem 3.3.1.

An upper-bound on the total number of vertex point computations required to obtain the final fused curve is the number of upper diagonal elements in the left and right convex envelope matrix $\left[\frac{J(J+1)}{2}\right]2 = J(J+1)$.

3.3.3 An Example on Identical/Non-Identical Relay Strategies

We may be inclined to believe that for a symmetric relay error curve and ML detection at the fusion center, the relays will use identical threshold tests. We have already seen that relays use identical but asymmetric strategies for the Gaussian relay channel at ML. We will shortly show that in fact, the optimal solution for the Additive Gaussian Noise case requires that both relays use identical threshold tests on every part of the final fused error curve $F(\phi)$. However, this is not true in general. Recall from previous discussion, for the discrete $J \times J$ case (each relay error curve has J vertex points), the structure of $F(\phi)$ is piecewise linear. Furthermore, the adjacent vertex points $\{V_{11}^L, V_{12}^L\}$ and adjacent vertex points $\{V_{J,J-1}^R, V_{J,J}^R\}$ are always on the final fused error curve $F(\phi)$. At $\eta = \Lambda(V_{11}^L, V_{12}^L) = \left(\frac{1-\beta_1}{\alpha_1}\right)\left(\frac{\beta_1-\beta_2}{\alpha_2-\alpha_1}\right)$ the η-tangent line goes through the two vertex points $\{V_{11}^L, V_{12}^L\}$ where one vertex point corresponds to identical relay threshold tests and the other to the relays using different threshold tests. This means that the optimal solution is not unique and the relays can use the same threshold test, different threshold tests, or some linear combination of both and arrive at the same error probability. Do relays use identical strategies for ML decoding and a symmetric relay channel? Again this is not true. We provide the following example.

Example: (*Symmetric Erasure-Like Channel*) A binary input, 4 symbol output symmetric relay channel and its error curve is shown in figure 3-12. The vertex points are $V_1 = (0, p), V_2 = (e, e), V_3 = (p, 0)$. The relays forward one bit to the fusion point. There are 3 types of relay strategies: (i) non-identical relay tests, (ii) identical but asymmetric relay tests, and (iii) identical and symmetric relay tests. We show below that the optimal relay strategy for ML decoding is non-identical relay tests for values of p and e such that $p > \frac{e}{p-e}$.

Suppose that p and e are allowed to vary. For the implicit output ordering, to assure convexity of the error curve and an erasure-like channel, the condition $0 < p < 1$ and $p > 2e$ must be satisfied. The final fused curve is symmetric from lemma 3.1.1. From the structure rules, there are 2 final fused curves. When $p > \frac{e}{p-e}$, the final fused curve is $(0,1) \rightarrow V_{11}^R \rightarrow V_{23}^L \rightarrow V_{12}^R \rightarrow V_{33}^L \rightarrow (1,0)$ as shown in figure 3-13 (B). We see that for $\frac{ep}{p^2-e} < \eta < 1$, the the optimal solution is V_{12}^R. If $1 < \eta < \frac{p^2-e}{ep}$, then V_{23}^L is optimal. When $p < \frac{e}{p-e}$, the final fused curve is

$(0,1) \to (0,p^2) \to (p^2,0) \to (1,0)$, which is shown in figure 3-13 (A). This is the error curve of the BEC. The optimal solutions are V_{11}^R for $\eta > 1$ and V_{33}^L for $\eta < 1$. This is the identical and asymmetric relay strategy. It is somewhat surprising that the symmetric and identical relay strategy V_{22} is never optimal.

As p decreases and approaches $2e$, the channel looks and behaves increasingly like a BEC. Thus, the optimal solution is like that of the BEC, identical relay tests $(\alpha, \alpha') = (0,0)$ and $(\alpha, \alpha') = (p,p)$ with asymmetric points on the relay error curve. In other words, when $2e < p < \frac{e}{p-e}$, the gain of zero-error detection when the symbol "1" or "4" is received by each relay is significant, so the best strategy for each relay is a Z-channel equivalent, similar to the BEC example in section 2.6. When $p > \frac{e}{p-e}$, the gain of the zero-error detection is less significant compared to the symbol output of a "2" or "3" by each relay. The optimal strategy (V_{23}^L or V_{12}^R) is then a combination of a Z-channel and BSC equivalent. As seen from the final fused curve, an interesting fact for the symmetric erasure-like channel is that the relay strategy of two BSC equivalent channels is never optimal.

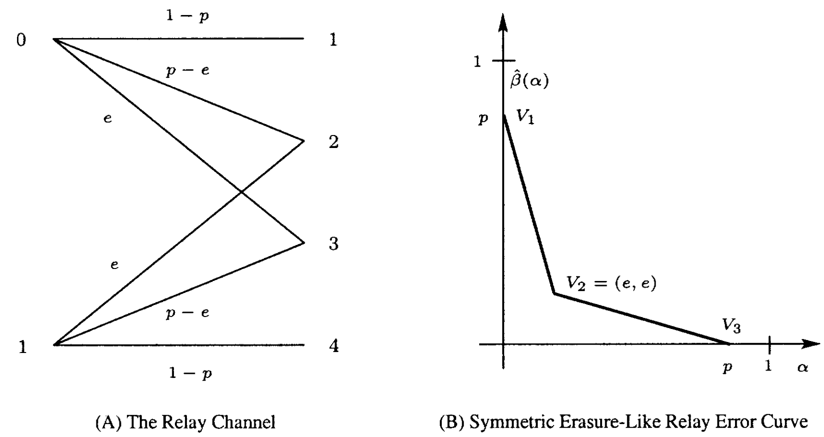
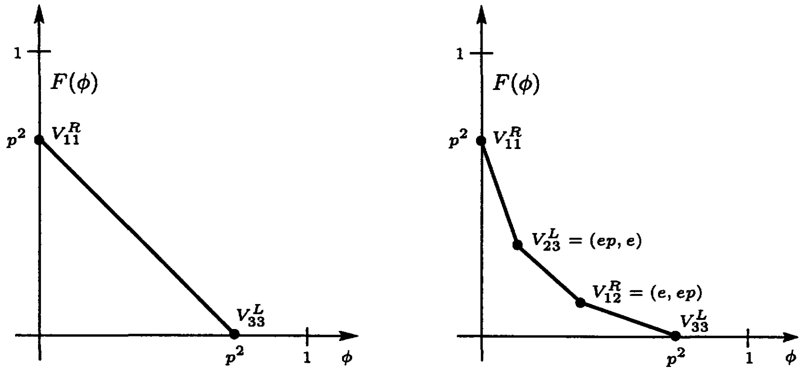


Figure 3-12: (A) Symmetric erasure-like relay channel and (B) the corresponding relay error curve.

We might further be inclined to believe that identical relay strategies for symmetric and strictly convex relay error curves and ML decoding may be optimal. This again is not true by a slight modification of the example just given by approximating the error curve in figure 3-12(B) with a strictly convex curve. A discrete error curve can be approximated arbitrarily closely by a continuous strictly convex curve which still preserves the same essential characteristics. We will revisit this example in Chapter 4.



(A) Final Fused Curve when $p < \frac{e}{p-e}$ (B) Final Fused Curve when $p > \frac{e}{p-e}$.

Figure 3-13: Symmetric erasure-like relay channel. (A) If $p < \frac{e}{p-e}$, then the identical relay tests V_{33}^L and V_{11}^R are optimal. (B) For $p > \frac{e}{p-e}$ and $\frac{ep}{p^2-e} < \eta < \frac{p^2-e}{ep}$ then the non-identical relay tests, V_{23}^L and V_{12}^R, are optimal. The symmetric and identical relay test V_{22} which corresponds to the relay mapping $\{1,2\} \to 0$ and $\{3,4\} \to 1$ is never optimal.

3.4 The Optimal MAP Solution to the 2×2 System

We have now shown that the final error curve for the 2×2 system is the convex hull of the left vertex points and the right vertex points of all possible fused curves where the relay strategies are taken over the entire relay error curve. These left and right vertex points correspond, respectively, to the (OR) and (AND) decoding rules at the fusion point. This final optimal fused curve is an error curve at the fusion point. For any given η, the optimal relay strategies are the relay threshold tests, (α, β) and (α', β'), corresponding to the the η-tangent solution to the fusion point error curve $F(\phi)$. The tangency point will be either a left vertex point, right vertex point, trivial point, or a linear segment with vertex points at the end of the segments. The overall error probability is the ordinate intercept of the η-tangent scaled by p_0 (or p_0' if there is a cost function other than the cost of one for a wrong decision). For given relay threshold tests at (α, β) and (α', β') the error probability of the left vertex point is

$$P_e^L(\alpha, \alpha') = p_1 \alpha \alpha' + p_0(\beta + \beta' - \beta\beta') \tag{3.41}$$

and the right vertex point is

$$P_e^R(\alpha, \alpha') = p_1(\alpha + \alpha' - \alpha\alpha') + p_0 \beta\beta'. \tag{3.42}$$

Call P_e^L the error probability for the *left vertex curve* or *left decoder curve* and P_e^R the error probability for the *right vertex curve* or *right decoder curve*. Both $P_e^L(\alpha, \alpha')$ and $P_e^R(\alpha, \alpha')$ are continuous functions of 2 variables, α and $\alpha^{'}$, since β and β' are functions of α and $\alpha^{'}$, respectively. We provide geometric insight into the behavior of the solutions as well as the optimal relay tests strategies. We also show how to determine the optimal solutions in a geometric and insightful manner (rather than taking the convex hull of all possible relay strategies in theorem 3.1.2).

We first study the strictly convex relay error curve $\hat{\beta}(\alpha)$. This means that every point on the relay error curve is a vertex point. The results derived hold in general but the conditions and arguments need to be slightly modified for discrete and more general error curves.

Since P_e for a fixed η is a bounded continuous function of two variables, $0 \leq \alpha \leq 1$ and $0 \leq \alpha^{'} \leq 1$, defined over a closed and bounded region $[0,1] \times [0,1]$, there exists a global maximum and minimum. The minimum and maximum is located at a boundary point or an interior point where $\frac{\partial P_e}{\partial \alpha} = \frac{\partial P_e}{\partial \alpha^{'}} = 0$ or where $\frac{\partial P_e}{\partial \alpha}, \frac{\partial P_e}{\partial \alpha^{'}}$ does not exist. If $(\alpha, \alpha^{'})$ is a stationary point, which means that both partial derivatives are equal to zero, then we have the following from calculus of variations:

(1) $\frac{\partial^2 P_e}{\partial \alpha^2} < 0$ and $\frac{\partial^2 P_e}{\partial \alpha^2}\frac{\partial^2 P_e}{\partial (\alpha^{'})^2} - (\frac{\partial^2 P_e}{\partial \alpha \partial \alpha^{'}})^2 > 0$, implies that $(\alpha, \alpha^{'})$ is a relative maximum;

(2) $\frac{\partial^2 P_e}{\partial \alpha^2} > 0$ and $\frac{\partial^2 P_e}{\partial \alpha^2}\frac{\partial^2 P_e}{\partial (\alpha^{'})^2} - (\frac{\partial^2 P_e}{\partial \alpha \partial \alpha^{'}})^2 > 0$, implies that $(\alpha, \alpha^{'})$ is a relative minimum;

(3) $\frac{\partial^2 P_e}{\partial \alpha^2}\frac{\partial^2 P_e}{\partial (\alpha^{'})^2} - (\frac{\partial^2 P_e}{\partial \alpha \partial \alpha^{'}})^2 < 0$, implies that $(\alpha, \alpha^{'})$ is a saddle point;

(4) $\frac{\partial^2 P_e}{\partial \alpha^2}\frac{\partial^2 P_e}{\partial (\alpha^{'})^2} - (\frac{\partial^2 P_e}{\partial \alpha \partial \alpha^{'}})^2 = 0$, implies that $(\alpha, \alpha^{'})$ is inconclusive.

For an everywhere differentiable relay error curve, the second derivative of the error probability function with respect to α and $\alpha^{'}$ is always positive. We will show that $\frac{\partial^2 P_e}{\partial \alpha^2} > 0$ in theorem 3.4.1. This says that all the stationary points of P_e are local minimum of saddle points. For a discrete relay error curve, we will later show that all the stationary points (α, α') for which α and α' are vertex points are local minima.

Denote $\eta^{++} = -\lim_{\phi \to 0} \frac{dF(\phi)}{d\phi}$ and $\eta^{--} = -\lim_{\phi \to 1} \frac{dF(\phi)}{d\phi}$. Then for $\eta < \eta^{--}$, the optimal decoder is the trivial decoding rule which maps everything to "0." For $\eta > \eta^{++}$, the optimal decoder is the trivial decoder which maps everything to "1." For the trivial decoder, the performance is the same regardless of the number of relays and the relay encoding rate. Observe that for any non-trivial relay channel, for any particular $\eta \in (\eta^{--}, \eta^{++})$, the MAP performance of two relays strictly decreases the error probability of one relay. This follows from theorem

3.1.2 which says V_α and $V_{\alpha'}$ are never on the final fused curve.

Recall that an *erasure-like channel* occurs when one or both of the following occurs: (i) $\hat{\beta}(0) < 1$; (ii) for some $0 < p < 1$, $\hat{\beta}(\alpha) = 0$ for $p < \alpha \leq 1$. In other words, the error curve for an erasure-like channel will either start at $(0, v)$, end at $(w, 0)$ for some $0 < v, w < 1$, or both. For these erasure-like channels, the optimal solutions may lie at these endpoints and must be checked. For non-trivial channels which are not erasure-like, suppose one relay has a threshold test at $(0,1)$ or $(1,0)$. Then regardless of what that relay observes, it always sends a "1" or "0" to the fusion center. The fusion center receives no information from this particular relay and might as well ignore it. Hence, relay tests at either point $(0,1)$ or $(1,0)$ are suboptimal unless η is so large ($\eta > \eta^{++}$) or so small $\eta < \eta^{--}$ that the fusion point uses the trivial decoder, in which case it does not care what either relay sees. Therefore, for non-erasure like error curves, when $(\eta^{--})^2 < \eta < (\eta^{++})^2$, the minimum does not occur at the boundary, $(\alpha, \alpha') \in \{(0,0), (0,1), (1,0), (1,1)\}$, but at an interior point. In other words, if $\eta^{--} = \infty$ the minimum does not occur at $(0,1)$; or if $\eta^{++} = 0$, the minimum does not occur at $(1,0)$. From the necessary optimality conditions, the optimal solution on the left vertex point curve must be a stationary point and satisfy the extremal condition that all the partial derivatives of P_e with respect to α and α' are zero. Taking partial derivatives of equations 3.41 and 3.42, we have

$$\frac{\partial P_e^L}{\partial \alpha} = p_1 \alpha' + p_0 \frac{d\hat{\beta}(\alpha)}{d\alpha}(1 - \beta') = 0 \tag{3.43}$$

$$\frac{\partial P_e^L}{\partial \alpha'} = p_1 \alpha + p_0 \frac{d\hat{\beta}(\alpha)}{d\alpha}(1 - \beta) = 0 \tag{3.44}$$

and similarily for the right vertex point,

$$\frac{\partial P_e^R}{\partial \alpha} = p_1(1 - \alpha') + p_0 \frac{d\hat{\beta}(\alpha)}{d\alpha}(\beta') = 0 \tag{3.45}$$

$$\frac{\partial P_e^R}{\partial \alpha'} = p_1(1 - \alpha) + p_0 \frac{d\hat{\beta}(\alpha)}{d\alpha}(\beta) = 0 \tag{3.46}$$

which reduces to the following necessary conditions for the left vertex curve

$$\left(\frac{d\hat{\beta}(\alpha)}{d\alpha}\right)\left(\frac{1-\beta'}{\alpha'}\right) = -\eta \quad \text{and} \quad \left(\frac{d\hat{\beta}(\alpha')}{d\alpha'}\right)\left(\frac{1-\beta}{\alpha}\right) = -\eta \tag{3.47}$$

and for the right vertex point curve,

$$\left(\frac{d\hat{\beta}(\alpha)}{d\alpha}\right)\left(\frac{\beta'}{1-\alpha'}\right) = -\eta \quad \text{and} \quad \left(\frac{d\hat{\beta}(\alpha')}{d\alpha'}\right)\left(\frac{\beta}{1-\alpha}\right) = -\eta. \tag{3.48}$$

Geometrically, the first equation of 3.47 for the left vertex curve above is the product of the tangent slope at α multiplied by the likelihood of Relay B receiving a "0." Recall that $\Lambda^{(\alpha')}(0) = \frac{1-\beta'}{\alpha'}$ is the value of the magnitude of the slope from (α', β') to $(0, 1)$. Likewise, the second equation of 3.47 for the left vertex curve above is merely the product of the magnitude of the slope at α' multiplied by the likelihood of Relay A receiving a "0." Recall $\Lambda^{(\alpha)}(0) = \frac{1-\beta}{\alpha}$ is the value of the slope from (α, β) to $(0, 1)$. Similar geometric interpretation for the equations in 3.48 of the right curve. The right vertex curve above is the product of the tangent slope at α multiplied by the likelihood of Relay B receiving a "1," where $\Lambda^{(\alpha')}(1) = \frac{\beta'}{1-\alpha'}$ is the value of the magnitude of the slope from (α', β') to $(1, 0)$.

3.4.1 Joint Optimality Conditions

A more insightful method to derive the joint optimality conditions of 3.47 and 3.48 is to use the conditional locus curve representations $F_\alpha(\phi)$ in section 3.1.2. For a given η and α, the α' which minimizes the P_e for the left decoding rule is the η-tangent solution to the left vertex curve, which is the conditional left locus curve of $F_\alpha(\phi)$. Likewise, for the given η and α', the α which minimizes the P_e must also be the η-tangent solution to the conditional left locus curve of $F_{\alpha'}(\phi)$. If (a, α') is an optimal solution, then α must be optimal given α' and vice versa, α' must be optimal given α. That is, both conditions of optimality must be satisfied. Thus, from equation 3.18, any stationary point solution $(\alpha, \alpha') = ((\alpha, \beta), (\alpha', \beta'))$ of P_e using the left vertex point decoding rule must satisfy the joint optimality conditions

$$-\frac{d\hat{\beta}(\alpha')}{d\alpha} = \frac{\alpha}{1-\beta}\eta \qquad \text{and} \qquad -\frac{d\hat{\beta}(\alpha)}{d\alpha} = \frac{\alpha'}{1-\beta'}\eta \tag{3.49}$$

Similarily, from equation 3.19, any stationary point solution (α, α') of P_e using the right vertex point decoding rule must satisfy the joint optimality conditions

$$-\frac{d\hat{\beta}(\alpha')}{d\alpha} = \frac{1-\alpha}{\beta}\eta \qquad \text{and} \qquad -\frac{d\hat{\beta}(\alpha)}{d\alpha} = \frac{1-\alpha'}{\beta'}\eta \tag{3.50}$$

This argument using conditional locus curves also shows that stationary points of P_e must be a minimum or saddle point. It also further shows that if vertex point solutions of a discrete relay error curve are chosen to satisfy the joint optimality conditions, these stationary points must be local minimum. The reason is as follows. Any small perturbation from a vertex point will be on a linear segment (randomized point), which can only increase P_e. Thus, the vertex solutions of a discrete relay error curve which satisfy the joint optimality solutions of 3.49 and 3.50 are local minimum. Finally, the P_e for the global minimum for the left vertex curve must be compared to the global minimum for the right vertex curve for the overall global minimum solution.

3.4.2 The Strictly Convex Relay Error Curve

This section treats the strictly convex relay error curve which is everywhere differentiable. Strict convexity implies that all stationary points (α, α') are distinct. Note that a stationary point (α, α') represents two points on the relay error curve $(\alpha, \alpha') = ((\alpha, \beta), (\alpha', \beta'))$. Equations 3.49 and 3.50 give some insight into where to look for the optimal solutions on the relay error curve. We show in the following lemma that all left vertex curve solutions lie on the right/lower part of the relay error curve $\hat{\beta}(\alpha)$, and all the right vertex curve solutions lie on the left/upper part of the relay error curve.

Lemma 3.4.1 *For any given η, the (α, α') that minimizes $F^L_{\alpha,\alpha'}(\phi)$ must satisfy $\alpha, \alpha' > \tilde{\alpha}(\eta)$. The (α, α') that minimizes $F^R_{\alpha,\alpha'}(\phi)$ must satisfy $\alpha, \alpha' < \tilde{\alpha}(\eta)$. For a strictly convex relay error curve, the stationary points are unique for each vertex curve.*

Proof: To see this, note that for a fixed η, the necessary optimal conditions for the left locus from equation 3.49 say that a stationary point $(\alpha, \alpha') = ((\alpha, \beta), (\alpha', \beta'))$ must satisfy both $\mid \frac{d\hat{\beta}(\alpha)}{d\alpha} \mid = \frac{\alpha'}{1-\beta'}\eta < \eta$ and $\mid \frac{d\hat{\beta}(\alpha')}{d\alpha'} \mid = \frac{\alpha}{1-\beta}\eta < \eta$. This implies that all the solutions (α, α') to the left decoding rule are located in the right/lower part of the relay error curve, in the range $\alpha, \alpha' \in [\tilde{\alpha}(\eta), 1]$ where $\eta = -\frac{d\hat{\beta}(\alpha)}{d\alpha}$ for some α. Likewise, for the right decoder curve, from equation 3.50, $\mid \frac{d\hat{\beta}(\alpha)}{d\alpha} \mid = \frac{1-\alpha'}{\beta'}\eta > \eta$ and $\mid \frac{d\hat{\beta}(\alpha')}{d\alpha'} \mid = \frac{1-\alpha}{\beta}\eta > \eta$ which implies that all the optimal solutions (α, α') occur in the top/left half of the relay error curve, in the range $\alpha, \alpha' \in [0, \tilde{\alpha}(\eta)]$. Uniqueness of the stationary point solutions follows from the fact that the η-tangent solutions are unique for each η for a strictly convex error curve. □

Note that this is not inconsistent with the general structure rules of the previous section. The first two vertex points on the final fused curve for a discrete system are V^L_{11} and V^L_{12}. In this case, η^{++} is just large enough that the right/lower part of the relay error curve contains all the vertex points for $\eta < \eta^{++}$.

This lemma is also consistent with the BEC example as we have seen before where the relay error curve is $(0,1) \to (0,\epsilon) \to (\epsilon,0) \to (1,0)$ and the final fused curve is $(0,1) \to (0,\epsilon^2) \to (\epsilon^2,0) \to (1,0)$. The left vertex solution is the point $(0,\epsilon^2)$ and corresponds to both relays using the same threshold test $(\epsilon,0)$, which is the stationary point $(\alpha,\alpha') = (\epsilon,\epsilon)$. This optimal solution is located in the right/lower part of the relay error curve. The right vertex solution is the point $(\epsilon^2,0)$ and the result of both relays using the same threshold test $(0,\epsilon)$ which is the stationary point $(\alpha,\alpha') = (0,0)$. This stationary point solution is located in the left/upper half of the relay error curve.

Theorem 3.4.1 *(Nesting property of stationary solutions for a strictly convex everywhere differentiable relay error curve.) At every η, 3.47 has exactly one solution of the form $\alpha = \alpha'$; all other pairs of solutions (α,α') are of the form $\alpha \neq \alpha'$ and will occur in strictly nested pairs around the $\alpha = \alpha'$ solution. The same holds true at every η for 3.48. Finally, all the stationary solutions are either local minima or saddle points.*

Proof: All stationary points are disjoint for the strictly convex relay error curve because there is exactly one point of tangency to the fused curve for each η. If $\alpha = \alpha'$, then the pair of left vertex optimal equations are satisfied with $-\frac{d\hat{\beta}(\alpha)}{d\alpha} = \eta\frac{\alpha}{1-\beta}$ and the right vertex optimal equations are satisfied with $-\frac{d\hat{\beta}(\alpha)}{d\alpha} = \eta\frac{1-\alpha}{\beta}$. Since $-\frac{d\hat{\beta}(\alpha)}{d\alpha}$ is monotonically decreasing from ∞ to 0 and $\frac{\alpha}{1-\beta}\eta$ is monotonically increasing from 0 to 1 as a function of α, they can only intersect at one point. See figure 3-14. Thus, there exists only one solution of the form $\alpha = \alpha'$, which we call the "identical solution," for the left decoding rule. The same argument applies for the right decoding rule, for which the conditions $-\frac{d\hat{\beta}(\alpha)}{d\alpha} = \eta\frac{1-\alpha'}{\beta'}$ and $-\frac{d\hat{\beta}(\alpha')}{d\alpha} = \eta\frac{1-\alpha}{\beta}$ must be satisfied. Since $\frac{1-\alpha}{\beta}\eta$ is monotonically increasing from 1 to infinity, again there exists one point of intersection which implies exactly one solution of the form $\alpha = \alpha'$.

Next, we prove the second part of the lemma, which says that if additional stationary points (α,α') exist, they must be nested. These stationary points are called the "non-identical solutions." The nesting property is a direct consequence of the monotonicity of the functions $\frac{d\hat{\beta}(\alpha)}{d\alpha}$, $\frac{\alpha}{1-\beta}$, and $\frac{1-\alpha}{\beta}$ which can be understood graphically in figure 3-14. Let the solution where $\alpha = \alpha'$ be denoted by α_0^* on the left vertex curve.

Assume without loss of generality that $\alpha_{-1}^* < \alpha_1^*$. We will show that $\alpha_{-1}^* < \alpha_0^* < \alpha_1^*$ for any solution $(\alpha_{-1}^*,\alpha_1^*)$ other than (α_0^*,α_0^*). Suppose there are two additional solution pairs, $(\alpha_{-1}^*,\alpha_1^*)$ and $(\alpha_{-2}^*,\alpha_2^*)$. without loss of generality, let $\alpha_{-2}^* < \alpha_{-1}^*$. The necessary conditions

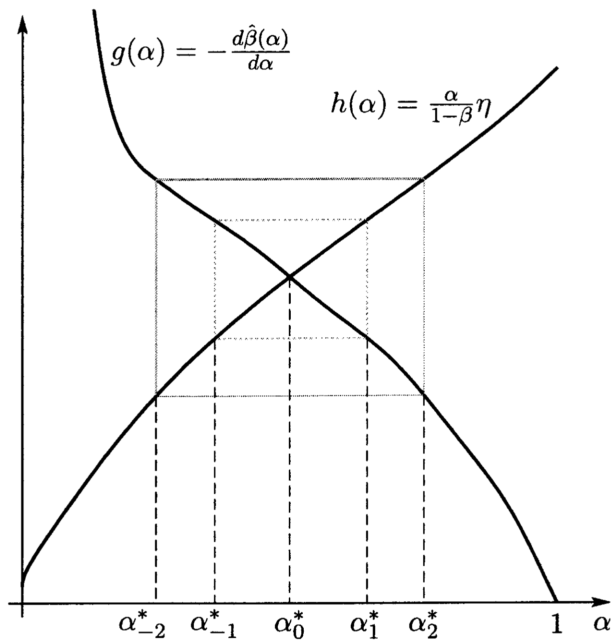


Figure 3-14: The 3 stationary point solutions for the left vertex curve at (α_0^*, α_0^*), $(\alpha_{-1}^*, \alpha_1^*)$, and $(\alpha_{-2}^*, \alpha_2^*)$ all satisfy the necessary optimal conditions $g(\alpha) = h(\alpha^{'})$ and $g(\alpha^{'}) = h(\alpha)$. There exists exactly one identical relay solution (α_0^*, α_0^*). All other solutions must successively nest around (α_0^*, α_0^*) and form a "box" as seen in the figure.

for optimality in equation 3.47 require that the solution (α_0^*, α_0^*) solution that $-\frac{d\hat{\beta}(\alpha)}{d\alpha}|_{\alpha_0^*} = \frac{\alpha_0^*}{1-\beta_0^*}$ and for the $(\alpha_{-1}^*, \alpha_1^*)$ satisfies $-\frac{d\hat{\beta}(\alpha)}{d\alpha}|_{\alpha_{-1}^*} = \frac{\alpha_1^*}{1-\beta_1^*}$ and $-\frac{d\hat{\beta}(\alpha)}{d\alpha}|_{\alpha_1^*} = \frac{\alpha_{-1}^*}{1-\beta_{-1}^*}$. Since $\alpha_{-1}^* < \alpha_0^*$, then $-\frac{d\hat{\beta}(\alpha)}{d\alpha}|_{\alpha_{-1}^*} > -\frac{d\hat{\beta}(\alpha)}{d\alpha}|_{\alpha_0^*}$ which implies $\frac{\alpha_1^*}{1-\beta_1^*} > \frac{\alpha_0^*}{1-\beta_0^*}$ and it follows that $\alpha_1^* > \alpha_0^*$. Thus, the solution pair $(\alpha_{-1}^*, \alpha_1^*)$ must nest around α_0^*.

We next show that another solution pair $(\alpha_{-2}^*, \alpha_2^*)$ nests around $(\alpha_{-1}^*, \alpha_1^*)$. Without loss of generality, assume $\alpha_{-2}^* < \alpha_{-1}^* < \alpha_0^*$. The necessary optimality conditions in equation 3.47 require $-\frac{d\hat{\beta}(\alpha)}{d\alpha}|_{\alpha_{-2}^*} = \frac{\alpha_2^*}{1-\beta_2^*}$ and $-\frac{d\hat{\beta}(\alpha)}{d\alpha}|_{\alpha_2^*} = \frac{\alpha_{-2}^*}{1-\beta_{-2}^*}$. Since $\alpha_{-2}^* < \alpha_{-1}^*$, then $-\frac{d\hat{\beta}(\alpha)}{d\alpha}|_{\alpha_{-2}^*} > -\frac{d\hat{\beta}(\alpha)}{d\alpha}|_{\alpha_{-1}^*}$ which implies $\frac{\alpha_2^*}{1-\beta_2^*} > \frac{\alpha_1^*}{1-\beta_1^*}$ and it follows that $\alpha_2^* > \alpha_1^*$. Thus, there exists exactly one solution where the relays use identical rules $\alpha = \alpha^{'}$ and all other solutions are of the form $\alpha \neq \alpha^{'}$ which must successively nest around the identical solution.

Finally, to show that no stationary point is a maximum, we show that $\frac{\partial^2 P_e}{\partial \alpha^2} > 0$ for both locus curves. The second partial derivative of the left vertex curve is $\frac{\partial^2 P_e}{\partial \alpha^2} = (1 - \beta^{'})\frac{d^2\hat{\beta}(\alpha)}{d\alpha^2}$ and of the right vertex curve $\frac{\partial^2 P_e}{\partial \alpha^2} = \beta^{'}\frac{d^2\hat{\beta}(\alpha)}{d\alpha^2}$. By convexity of the relay error curve, $\frac{d^2\hat{\beta}(\alpha)}{d\alpha^2} > 0$ which implies $\frac{\partial^2 P_e}{\partial \alpha^2} > 0$ for both left and right decoder curves. This implies that all stationary point solutions must be local minima or saddle points. □

After finding the minimum for the left decoding rule and the minimum for the right decoding rule, we must compare their error probabilities to determine the global minimum.

We now investigate further the optimality conditions of 3.47 and 3.48 for a stationary point. First consider the left decoder curve. For simplicity, let $g(\alpha) = -\frac{d\hat{\beta}(\alpha)}{d\alpha}$ and $h(\alpha) = \frac{\alpha}{1-\beta}\eta$ Then the 2 necessary optimality conditions in equations 3.47 and 3.48 can be restated as $g(\alpha) = h(\alpha')$ and $g(\alpha') = h(\alpha)$. For the strictly convex and everywhere differentiable error curve, the functions g and h are both 1-1 with α so their inverse functions, g^{-1} and h^{-1} exist. An iterative procedure is shown in figure 3-15 The optimality conditions of g and h can be combined and expressed as $\alpha = g^{-1}h(\alpha')$ and $\alpha' = g^{-1}h(\alpha)$. It can further be combined to $\alpha = g^{-1}(h(g^{-1}(h(\alpha)))) = f(\alpha)$ where the function f represents the composite function $f = g^{-1}hg^{-1}h$. Then, all stationary points (α, α') are fixed points of $\alpha = f(\alpha)$ and $\alpha' = f(\alpha')$. Note that $g^{-1}h(\alpha) = \tilde{\alpha}(\frac{\alpha}{1-\beta})$ for the left vertex curve and $g^{-1}h(\alpha) = \tilde{\alpha}(\frac{1-\alpha}{\beta})$ for the right vertex curve. Hence, $\alpha' = f(\alpha') = \tilde{\alpha}\left(\tilde{\alpha}(\frac{\alpha}{1-\beta})\right)$. A fixed point solution α to $\alpha = f(\alpha)$ has the corresponding α' to the solution pair (α, α') defined by $\alpha' = g^{-1}h(\alpha) = \tilde{\alpha}(\frac{\alpha}{1-\beta})$. Similarily a fixed point solution α' has corresponding α defined by $g^{-1}h(\alpha)$. The equations can also be combined in the other way to $\alpha = h^{-1}gh^{-1}g(\alpha) = k(\alpha)$ where k is the composite function $k = h^{-1}gh^{-1}g$. Note that the functions f and k are monotonically increasing and are continuous functions of α and α'. We show that f is monotonically increasing by showing that if $\alpha > \alpha_0 \implies f(\alpha) > f(\alpha_0)$. If $\alpha > \alpha_0$, then $h(\alpha) < h(\alpha_0) \implies g^{-1}h(\alpha) < g^{-1}h(\alpha_0) \implies hg^{-1}h(\alpha) > hg^{-1}h(\alpha_0) \implies g^{-1}hg^{-1}h(\alpha) > g^{-1}hg^{-1}h(\alpha_0) \implies f(\alpha) > f(\alpha_0)$.

Now suppose that $f(x)$ is an arbitrary monotonic increasing continuous function and that $f(x)$ intersects the line $y = x$ at least once. We now look at the fixed points of f. As seen in the figure 3-16, if for all x, $f(x) - x$ is monotonically decreasing, then this implies $f'(x) - 1 < 0$, which can be restated as $f'(x) < 1$, and then this further implies $f(x)$ has one *stable fixed point* at x^*. This means that starting at any arbitrary point x_1 and using the iterative procedure, $x_2 = f(x_1), x_3 = f(x_2), \ldots, x_n = f(x_{n-1})$, the sequence x_n converges. A stable fixed point x^* of the algorithm means that the algorithm converges to the fixed point solution $\lim_{n\to\infty} x_n = x^*$. On the other hand, if $\forall x$, $f(x) - x$ is monotonically increasing, this implies $f'(x) > 1$, then the algorithm has an unstable fixed point at x^*, which is noted as $f(x)$ has one *unstable fixed point* at x^*. This means that starting at any arbitrary point x_1 and using the iterative procedure, $x_n = f(x_{n-1})$, the sequence x_n diverges as $n \to \infty$.

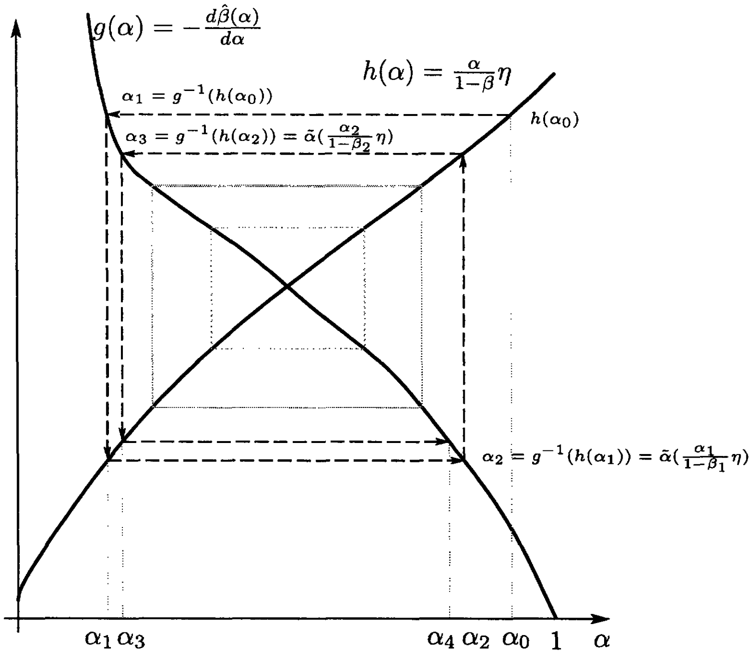
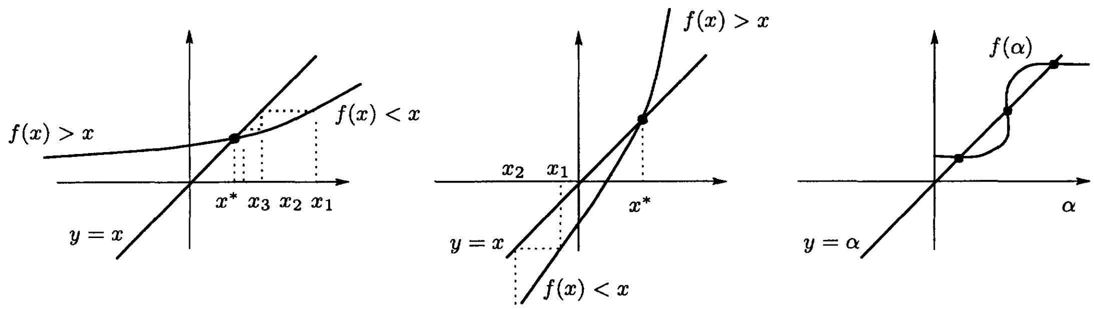


Figure 3-15: (Continuation of the example in figure 3-14)The stationary point solutions will alternate between local minima and saddle points. For this particular example of the left vertex solution, the iterative procedure starts at α_0, then to $\alpha_1 = g^{-1}h(\alpha_0) = \tilde{\alpha}(\frac{\alpha_0}{1-\beta_0}\eta)$, then to $\alpha_2 = g^{-1}h(\alpha_1) = \tilde{\alpha}(\frac{\alpha_1}{1-\beta_1}\eta)$, and continue to spiral into to a local minimum solution. In the figure, $(\alpha^*_{-2}, \alpha^*_2)$ is a local minimum, $(\alpha^*_{-1}, \alpha^*_1)$ is a saddle point, and (α^*_0, α^*_0) is a local min. If the iterative procedure started in between the nested boxes, $\alpha^*_{-2} < \alpha_0 < \alpha^*_{-1}$, it would spiral out to the stationary point $(\alpha^*_{-2}, \alpha^*_2)$, which is the outer nested box. If the iterative algorithm started in between the nested box $\alpha^*_0 < \alpha_1 < \alpha^*_1$ and the point of intersection (α^*_0, α^*_0), it would spiral into (α^*_0, α^*_0).

Now back to the composite function f. From theorem 3.4.1, $f(\alpha)$ has at least one fixed point,α^*, in the range $\alpha^* \in (0,1)$. Isolated fixed points of f cross the line $y = \alpha$ at a slope less than one (stable fixed point) or greater than one (unstable fixed point). In Appendix B, we will show by example that the composite function $f(x)$ can equal 1 on an interval which implies that relay error curves exists that have countably infinite stationary points. Since the composite functions f and k are continuous functions of α and α', and the stationary points are either local minima or saddle points, these solutions must alternate as fixed points of f. In fact, the local minimum solutions are stable fixed points and cross the $y = x$ line at slope less than one (convergence to α^* must occur for any α close to α^*); the saddle points are unstable fixed points and cross the $y = x$ line at slope greater than one (small perturbation for α^* must move away).

We now examine these fixed point solutions of $f(\alpha)$ in more detail in figure 3-15:. This is the same viewpoint of figure 3-16 but in expanded detail. The solutions must be successively nested and alternate between local min and saddle point. To find the stable fixed point solutions, we implement the same iterative procedure just discussed, $\alpha_{n+1} = f(\alpha_n) = g^{-1}hg^{-1}h(\alpha_n)$, but in 4



(A) stable fixed point at x^* $\quad$ (B) unstable fixed point at x^* $\quad$ (C) composite function f

Figure 3-16: The fixed point solution to $f(x) = x$ at x^* in (A) is *stable* and corresponds to a local min and in (B) it is *unstable* and corresponds to a saddle point. (C) The composite function $f = g^{-1}hg^{-1}h$ where $g(\alpha) = -\frac{d\hat{\beta}(\alpha)}{d\alpha}$ and $h(\alpha) = \frac{\alpha}{1-\beta}\eta$ (left curve) and $h(\alpha) = \frac{1-\alpha}{\beta}\eta$ (right curve). Every α and α' in the stationary point solution (α, α') of P_e is a fixed point of f. Observe that the stable and unstable fixed points of f must alternate since f is continuous and monotonically increasing. Note that $f(\alpha)$ can have intervals which coincide with $y = \alpha$.

steps instead of one. As shown in the figure, start at α_1, the first point is $g(\alpha_1)$; travel vertically down to $h(\alpha_1)$; then right across to $\alpha_2 = g^{-1}h(\alpha_1)$; then straight up to $h(\alpha_2) = g^{-1}h(\alpha_1)$; then left across to $\alpha_3 = g^{-1}h(\alpha_2) = g^{-1}hg^{-1}h(\alpha_1)$, and so forth. The figure shows local min at (α_0^*, α_0^*) and $(\alpha_{-2}^*, \alpha_2^*)$ and saddle point at $(\alpha_{-1}^*, \alpha_1^*)$. The iterative procedure either spirals into a local min solution, out to a local min solution, spirals around a "box," or remains stuck at (α_0^*, α_0^*). It will stay on the "box" of a saddle point if it starts anywhere on that "box." If it starts anywhere near a local min, the error probability decreases at every iteration and converges into that local min "box" or point.

Lemma 3.4.2 *(Sufficient condition for identical relay strategies.) For a given relay error curve and fixed η, let (1) the left solution iteration $\alpha_{i+1} = \tilde{\alpha}\left(\tilde{\alpha}\left(\frac{\alpha_i}{1-\beta_i}\eta\right)\right)$ and (2) the right solution iteration $\alpha_{i+1} = \tilde{\alpha}\left(\tilde{\alpha}\left(\frac{1-\alpha_i}{\beta_i}\eta\right)\right)$ and the optimal identical relay strategy at $\alpha = \alpha' = \alpha_L^*$ for the left curve and $\alpha = \alpha' = \alpha_R^*$ for the right curve. Then for the chosen η, a sufficient condition for identical relay strategies on the fused error curve $F(\phi)$ occurs if for both (1) and (2), we have $\forall \alpha_i < \alpha^*$ implies $\alpha_i < \alpha_j < \alpha^*$ for all $j > i$; and if $\forall \alpha_i > \alpha^*$ implies $\alpha^* < \alpha_j < \alpha_i$ for all $j > i$. In other words, if for any initial value α_1, we have* $\lim_{n\to\infty} \alpha_n = \alpha_L^*$ *and* $\lim_{n\to\infty} \alpha_n = \alpha_R^*$.

Proof: If for any starting point α_1 on the relay error curve, the left iterative algorithm $\alpha_{n+1} = f(\alpha_n)$ spirals into the intersection of $g(\alpha) = -\frac{d\hat{\beta}(\alpha)}{d\alpha}$ and $h(\alpha) = \frac{\alpha}{1-\beta}\eta$ for the left vertex curve and the right iterative algorithm $\alpha_{n+1} = f(\alpha_n)$ spirals into the intersection of $g(\alpha) = -\frac{d\hat{\beta}(\alpha)}{d\alpha}$ and $h(\alpha) = \frac{1-\alpha}{\beta}\eta$ for the right vertex curve, then there exists only one fixed point solution to f. In other words, if for all α, for both curves, each sequence of the set of points of α_n has a limit, $\lim_{n\to\infty} \alpha_n = \alpha_L^*$ and $\lim_{n\to\infty} \alpha_n = \alpha_R^*$, then that limit is the only fixed point solution α^*. for the left and right curves, respectively. Since there is only one solution, it must be the identical solution from theorem 3.4.1. □

Corollary 3.4.1 *A trivial extension of lemma 3.4.2 says that if the lemma is true for all η, then the solution to the entire final fused error curve has identical relay strategies.*

3.4.3 The Discrete Relay Error Curve

Most of the results of the previous section for the strictly continuous everywhere differentiable error curve hold but need to be slightly modified for discrete error curves. From the previous section, local minimum and saddle points alternate and must occur at vertex points for the strictly convex error curve. Recall that the derivative at the vertex points of discrete error curves does not exist mathematically in the usual sense, although it is well-defined by an interval, $\frac{d\hat{\beta}(\alpha)}{d\alpha}|_p \in [\eta_p^-, \eta_p^+]$ for a vertex point p. Thus, all stationary points must satisfy the following modified joint optimality conditions. For the left vertex curve, (α, α') is a stationary point of $P_e^L(\alpha, \alpha'; \eta)$ if α and α' satisfy

$$\frac{\alpha'}{1-\beta'}\eta \in [\eta_\alpha^-, \eta_\alpha^+] \quad \text{and} \quad \frac{\alpha}{1-\beta}\eta \in [\eta_{\alpha'}^-, \eta_{\alpha'}^+] \tag{3.51}$$

Similarly, for the right vertex curve, (α, α') is a stationary point of $P_e^R(\alpha, \alpha'; \eta)$ if α and α' satisfy

$$\frac{1-\alpha'}{\beta'}\eta \in [\eta_\alpha^-, \eta_\alpha^+] \quad \text{and} \quad \frac{1-\alpha}{\beta}\eta \in [\eta_{\alpha'}^-, \eta_{\alpha'}^+]. \tag{3.52}$$

From the discussion of the joint optimality conditions in section 3.4.1, stationary point solutions which are vertex points are local minima. Thus, between any 2 nested stationary vertex points, there exists a nested saddle point solution (which is a randomized threshold test). This must happen because the error probability is a continuous function of α and α'. In order for the P_e function to have several local minima, there must be a saddle point or local maximum in between the local minima. From the conditional locus curves, it can be seen from the figure that conditional on α, the α' which maximizes the error probability must occur at the boundary, $\{0, 1\}$. Therefore, the local minima and saddle point solutions are alternately nested.

Lemma 3.4.3 *(Locations of optimal solutions on the relay error curve). For any given η, the*

(α, α') that minimizes $F^L_{\alpha,\alpha'}(\phi)$ must satisfy $\alpha, \alpha' \geq \tilde{\alpha}(\eta)$. The (α, α') that minimizes $F^R_{\alpha,\alpha'}(\phi)$ must satisfy $\alpha, \alpha' \leq \tilde{\alpha}(\eta)$. At most one vertex point, specifically, the vertex point $a = \tilde{\alpha}(\eta)$, on the relay error curve can be a part of a left vertex stationary point as well as a right vertex stationary point.

Proof: Similar to proof of lemma 3.4.1. Since the η-tangent solutions can have at most two vertex point solutions and each vertex point can have a range of η as its η-tangent, the point $a = \tilde{\alpha}(\eta)$ can be a part of the left and right stationary point solutions, and it is the only such point. □

We now examine the solution structure for the discrete error curve. Since $\frac{d\hat{\beta}(\alpha)}{d\alpha}$ is defined to be an interval for vertex points, $g(\alpha) = -\frac{d\hat{\beta}(\alpha)}{d\alpha}$ is a staircase function and monotonically decreasing. Since $\frac{\alpha}{1-\beta}$ and $\frac{1-\alpha}{\beta}$ are monotonically increasing functions of α, the solution structure is essentially the same as in the strictly convex case.

Theorem 3.4.2 *(Nesting property of stationary points). At every η, 3.51 has exactly one solution of the form $\alpha = \alpha'$; all other pairs of solutions (α, α') are of the form $\alpha \neq \alpha'$ and will occur in nested pairs around the $\alpha = \alpha'$ solution. The nesting condition is not strict, which means that the the stationary points need not be distinct. An error point, α or α' can be part of at most 2 stationary point solutions. For example, a set of solutions can be (α_0, α_0), (α_0, α_1), (α_{-1}, α_1), (α_{-1}, α_2), and so forth. The same holds true at every η for 3.52. Finally, all the stationary solutions which are vertex points are local minima.*

Proof: The nesting part of the theorem is essentially the same proof as theorem 3.4.1. The η-tangent line solution can have at most 2 vertex point solutions. Thus, if α_i^* is part of a stationary point solution, say (α_i^*, α_j^*), then it can at most be a part of another solution, (α_i^*, α_k^*), where $\frac{\alpha_i^*}{1-\beta_i^*} = \eta_j^- = \eta_k^+$, assuming $j < k$. □

There is no discrete counterpart to corollary 3.5.1. From the structure rules, for a discrete relay error curve with J vertex points which is not erasure-like, the left part of every final fused curve of every discrete system is $(0,1) \to (V^L_{11}) \to (V^L_{12})$ and the rightmost part is $(V^R_{J-1,J}) \to (V^R_{J,J}) \to (1,0)$. Thus, there exists some η such that the optimal fused relay rule is non-identical strategies. Only the erasure-like relay error curve $(0,1) \to (0,q) \to (p,0) \to (1,0)$ has identical relay strategies for all η. In conclusion, other than the 2-point erasure-like error curve just described, there exists no discrete error curve where the optimal solution is identical relay strategies everywhere on the final fused curve.

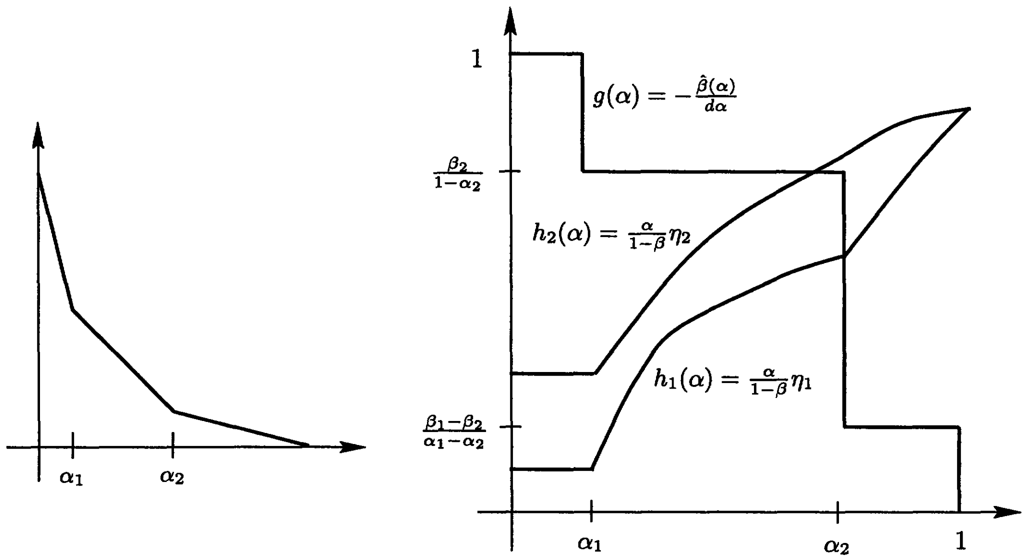


Figure 3-17: The local minima solutions of the discrete error curve with 2 vertex points (α_1, β_1) and (α_2, β_2) for the left decoder rule. At η_1: there is one point of intersection of $g(\alpha)$ and $h_1(\alpha)$, which is with the vertical segment of $g(\alpha)$. There is no "box" anywhere, so the optimal solution must be (α_2, α_2) or V_{22}^L. At η_2: the point of intersection of $g(\alpha)$ and $h_1(\alpha)$ is with the horizontal segment of $g(\alpha)$. This means the identical solution is a randomized point between V_1 and V_2 and thus cannot be optimal. There exists a box created by α_1 and α_2 so the optimal solution is V_{12}^L .

We first look at the point of intersection of $g(\alpha)$ and $h(\alpha)$ for the left decoding rule. This is the stationary point solution when both relays use the same rule, the α such that $g(\alpha) = h(\alpha)$. Unlike the strictly convex case, there does not exist a composite function f since g^{-1} and h^{-1} do not exist. However, the solution still exists as seen from figure 3-17.

Suppose for some η_2, the identical stationary point solution (α_0^*, α_0^*) is a non-vertex point, then $h_2(\alpha) = \frac{\alpha}{1-\beta}\eta_2$ intersects $-\frac{d\hat{\beta}(\alpha)}{d\alpha}$ on a horizontal segment of $g(\alpha)$ (see figure 3-17). This means that the stationary point for the left decoding rule where both relays use identical threshold tests is a randomized test for both relays. From theorem 3.2.1 this cannot be optimal, so it must be a saddle point. Hence, the global minimum for the left decoding rule must occur when the relays use strictly different threshold tests. For the example in figure 3-17, the global minimum must be V_{12}^L. Indeed, the stationary point (α_1, α_2) forms a "box" and is a vertex stationary point.

Suppose for some η_1, the intersection of $g(\alpha)$ and $h(\alpha)$ occurs at a vertical segment of $g(\alpha)$, then the identical stationary point solution (α_0^*, α_0^*) is a vertex point and a local minimum. This is demonstrated by the curve of $h_1(\alpha) = \frac{\alpha}{1-\beta}\eta_1$ in figure 3-17. The identical stationary point is $(\alpha_0^*, \alpha_0^*) = (\alpha_2, \alpha_2)$ and nested stationary point is $(\alpha_1, 1)$.

For the general error curve, the joint optimality conditions of equation 3.49 for solutions of the left decoding rule imply the following geometrical conditions which are depicted in figure 3-18:

(I) A vertex point α^* is a stationary point $(\alpha, \alpha') = (\alpha^*, \alpha^*)$ if $\alpha^* = \tilde{\alpha}\left(\frac{\alpha^*}{1-\beta^*}\right)$. This means that there exists no part of the relay error curve below the line of slope $\frac{\alpha^*}{1-\beta^*}$ passing through α^*.

(II) A pair of points (α_i^*, β_i^*) and (α_j^*, β_j^*) for $i \neq j$, is a stationary point $(\alpha, \alpha') = (\alpha_i^*, \alpha_j^*)$ if both of the following 2 conditions below are satisfied. These conditions guarantee the necessary joint optimal conditions $\alpha_i^* = \tilde{\alpha}\left(\frac{\alpha_j^*}{1-\beta_j^*}\right)$ and $\alpha_j^* = \tilde{\alpha}\left(\frac{\alpha_i^*}{1-\beta_i^*}\right)$.

(1) $\frac{\alpha_i^*}{1-\beta_i^*} \leq \frac{\beta_i^*-\beta_j^*}{\alpha_j^*-\alpha_i^*} \leq \frac{\alpha_j^*}{1-\beta_j^*}$

(2) No part of the Relay error curve $\hat{\beta}(\alpha)$ lies below the line of slope $\frac{\alpha_i^*}{1-\beta_i^*}$ through (α_i^*, β_i^*), and no part of the Relay error curve lies below the line of slope $\frac{\alpha_j^*}{1-\beta_j^*}$ through (α_j^*, β_j^*).

The above conditions say the following about relay error curves. If there are 2 solution pairs, say (α_2, α_2) and (α_1, α_3), then conditions (I) and (II) imply that the following must hold:

(i) $\frac{\alpha_2}{1-\beta_2} \leq \frac{\beta_1-\beta_2}{\alpha_2-\alpha_1} \leq \frac{\alpha_3}{1-\beta_3}$

(ii) $\frac{\alpha_1}{1-\beta_1} \leq \frac{\beta_2-\beta_3}{\alpha_3-\alpha_2} \leq \frac{\alpha_2}{1-\beta_2}$

(iii) $\frac{\alpha_3}{1-\beta_3} \leq \frac{1-\beta_1}{\alpha_1} \leq 1$

(iv) $0 \leq \frac{\beta_3}{1-\alpha_3} \leq \frac{\alpha_1}{1-\beta_1}$

These conditions are shown graphically in figure 3-18. Combining the conditions of (i)-(iv), we arrive at the following nested inequality

$$0 \leq \frac{\beta_3}{1-\alpha_3} \leq \frac{\alpha_1}{1-\beta_1} \leq \frac{\beta_2-\beta_3}{\alpha_3-\alpha_2} \leq \frac{\alpha_2}{1-\beta_2} \leq \frac{\beta_1-\beta_2}{\alpha_2-\alpha_1} \leq \frac{\alpha_3}{1-\beta_3} \leq \frac{1-\beta_1}{\alpha_1} \leq 1 \tag{3.53}$$

This string of inequalities can easily be extended for more nested solutions. There can exist an infinite number of local minima solutions for each locus curve. We will demonstrate this by constructing a discrete error curve with countably infinite number of stationary points for $P_e^L(\alpha, \alpha')$ in Appendix B. Similar optimality conditions exist for the right vertex curve.

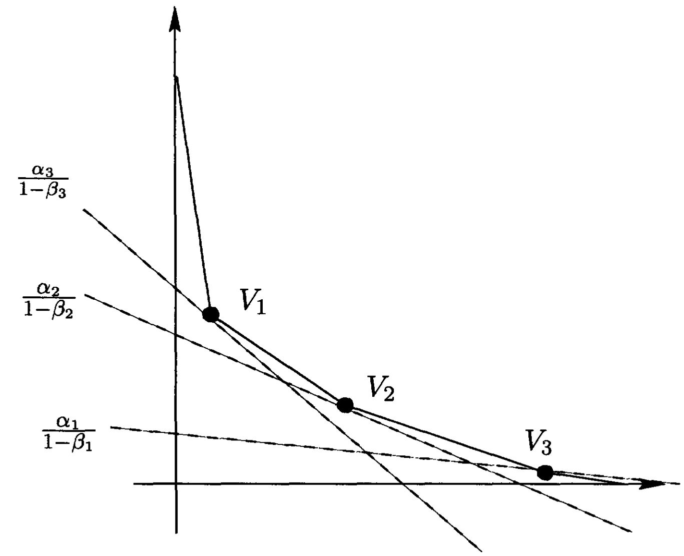


Figure 3-18: Discrete example of necessary conditions for optimality for the left decoding rule. Local minimum solutions at V_{22}^L and V_{13}^L.

Note that, given any η, an error curve can be constructed with given desired properties. For example, we can always construct a relay error curve such that the global optimal solution occurs when the relays use different threshold tests. By the same token, we can also construct, for a given η, a relay error curve such that the relays use identical strategies for global optimality.

There are many special cases which have interesting quirks that we are not going to treat them all here. The theory for each of these special cases can be worked out from what has been developed in this chapter.

3.5 Additive Gaussian Noise Channel (AGNC)

The ML solution to the AGNC for the 2×2 system was discussed in chapter 1. There are 2 optimal solutions, one corresponding to the "OR" decoding rule and the other to the "AND" decoding rule at the fusion point. Any coordinated randomized strategy among these two solutions results in the same error probability. We now prove the complete result for the Gaussian Relay Channel in the following lemma and theorem.

Lemma 3.5.1 *Given two points (α, β) and (α', β') on the Gaussian error curve, the point farthest from the 45 degree line is a vertex point of the fused curve $F_{\alpha,\alpha'}(\phi)$. In other words, if $|\alpha - \beta| > |\alpha' - \beta'|$, then V_α is on $F_{\alpha,\alpha'}(\phi)$ and V_α for the reversal of the inequality.*

Proof: From the LR, we have α or $\alpha^{'}$ as a vertex point if

$$\left(\frac{1-\beta}{\alpha}\right)\left(\frac{\beta^{'}}{1-\alpha^{'}}\right) \underset{\alpha'}{\overset{\alpha}{\gtrless}} \left(\frac{\beta}{1-\alpha}\right)\left(\frac{1-\beta^{'}}{\alpha^{'}}\right) \tag{3.54}$$

Both points are on $F_{\alpha,\alpha'}(\phi)$ if equality holds above. Expressing 3.54 in terms of Gaussian probabilities,

$$\left(\frac{\mathcal{Q}\left(t-m\right)}{\mathcal{Q}\left(t+m\right)}\right)\left(\frac{\mathcal{Q}\left(-t^{'}+m\right)}{\mathcal{Q}\left(-t^{'}-m\right)}\right) \underset{\alpha'}{\overset{\alpha}{\gtrless}} \left(\frac{\mathcal{Q}\left(-t+m\right)}{\mathcal{Q}\left(-t-m\right)}\right)\left(\frac{\mathcal{Q}\left(t^{'}-m\right)}{\mathcal{Q}\left(t^{'}+m\right)}\right)$$

and rearranging,

$$R(t) = \frac{\mathcal{Q}\left(t+m\right)}{\mathcal{Q}\left(t-m\right)}\frac{\mathcal{Q}\left(-t+m\right)}{\mathcal{Q}\left(-t-m\right)} \underset{\alpha}{\overset{\alpha'}{\gtrless}} \frac{\mathcal{Q}\left(t^{'}+m\right)}{\mathcal{Q}\left(t^{'}-m\right)}\frac{\mathcal{Q}\left(-t^{'}+m\right)}{\mathcal{Q}\left(-t^{'}-m\right)} = R(t^{'}). \tag{3.55}$$

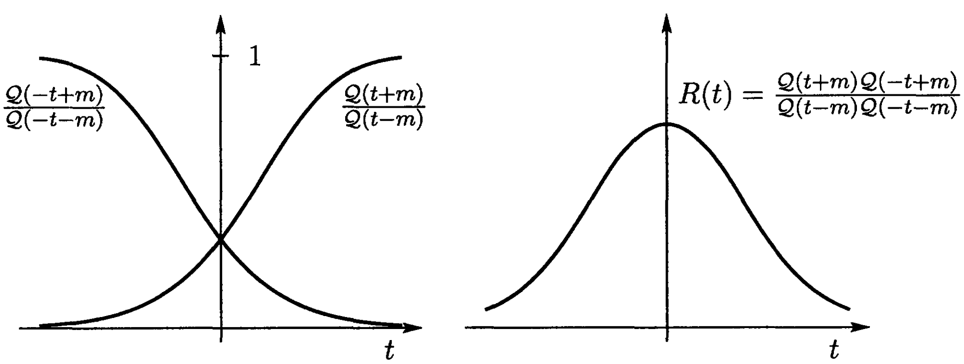


Figure 3-19: Functions of ratios of the Gaussian error function used for deciding which relay vertex point will be on the final fused curve.

Note that $R(t) = R(-t)$. We will show that for all m, $R(t)$ is monotonically increasing as t varies from negative infinity to the origin (which also means that $R(t)$ is monotonically decreasing as t varies from the origin to infinity) by showing that $\frac{d}{dt}R(t) = 0$ only at $t = 0$. By doing so, the lemma is proved.

$$\frac{d}{dt}R(t) = R(t)\left[-\frac{q(t+m)}{\mathcal{Q}\left(t+m\right)} - \frac{q(-t-m)}{\mathcal{Q}\left(-t-m\right)} + \frac{q(-t+m)}{\mathcal{Q}\left(-t+m\right)} + \frac{q(t-m)}{\mathcal{Q}\left(t-m\right)}\right] \tag{3.56}$$

Letting $L(x) = \frac{q(x)}{\mathcal{Q}(x)}$,

$$\frac{dR(t)}{dt} = -R(t)\left[L(t+m) + L(-t-m) - L(-t+m) - L(t-m)\right] \tag{3.57}$$

Now, $\lim_{t \to \pm\infty} \frac{d}{dt} R(t) = 0$ and $R'(0) = 0$. We need to show that $\forall m, t$, $\frac{d}{dt} R(t) \neq 0$ except at $t = 0, \forall m$. Without loss of generality, suppose t and m are the same sign. Then $|t+m| > |t-m|$. Let $a = |t-m|$ and $a+b = |t+m|$. If t and m are the opposite sign, merely reverse the indices by letting $a = |t+m|$ and $a+b = |t-m|$. Since $|t+m| < |t-m|$ the argument is the same. We now show that $\frac{d}{dt} R(t) = -R(t) \left[L(a+b) + L(-a-b) - L(a) - L(-a)\right]$ is equal to zero only at the origin. To do this, we will show that the midpoint of the line connecting $L(a+b)$ and $L(-a-b)$ will always lie above the midpoint of the line connecting $L(a)$ and $L(-a)$ for all $a, b \neq 0$. This is depicted in figure 3-20. We show in lemma 3.5.2 that $L(x)$ is a monotonically increasing

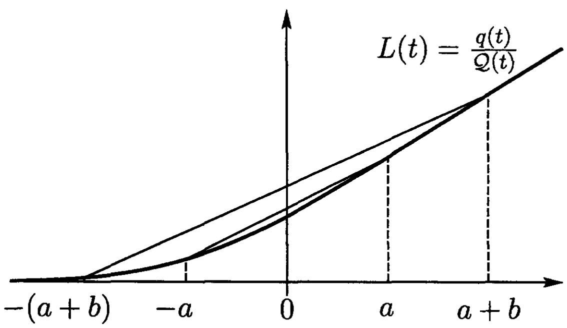


Figure 3-20: The chords do not intersect except at $t = 0$. Therefore, $\frac{1}{2}(L(a+b) - L(-a-b)) > \frac{1}{2}(L(a) - L(-a))$ except when $a = 0$ or $b = 0$.

function and that $L'(v) < L'w)$, $\forall v < 0$ and $\forall w > 0$. If $L'(v) < L'(w)$, $\forall v < 0$ and $\forall w > 0$, then the midpoint of the chord connecting $L(-a-b)$ and $L(a+b)$ is strictly above the midpoint of the chord connecting $L(-a)$ and $L(a)$. From the MVT (Mean Value Theorem), there exists a $v \in [-a-b, -a]$ and $w \in [a, a+b]$ such that $L'(w) = \frac{L(a+b)-L(a)}{b}$ and $L'(v) = \frac{L(-a)-L(-a-b)}{b}$. Since $\forall v < 0$ and $\forall w > 0$, $L'(v) < L'(w)$, this implies that $L(a+b) + L(-a-b) > L(a) + L(-a)$. $\square$

Theorem 3.5.1 *For the 2×2 additive Gaussian noise channel, the optimal relay strategy for every cost function is unique and identical for both relays. That is, every point on the final optimal fused error curve $F(\phi)$ corresponds to both relays using identical threshold tests.*

Proof: It is more convenient to use a parametric representation for the Gaussian relay error curve. Let $\tilde{\alpha}(t) = P(e|1) = \mathcal{Q}(t+m)$ and $\tilde{\beta}(t) = P(e|0) = \mathcal{Q}(-t+m)$ for $t \in (-\infty, \infty)$.

We prove the lemma for the the left decoder curve and then show that the right decoder curve follows the same argument. Calculating expressions for $\frac{d\hat{\beta}(\alpha)}{d\alpha}$ and $\eta \frac{\alpha}{1-\beta}$ for the necessary

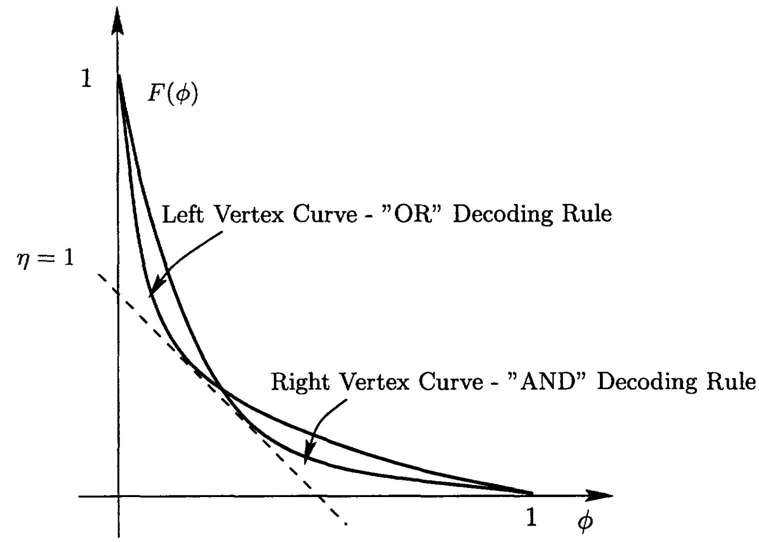


Figure 3-21: The final fused error curve for the additive Gaussian noise channel. $F(\phi)$ is the convex envelope of the left vertex curve and the right vertex curve. Only the ML ($\eta = 1$) solution has 2 solutions, the "OR" and the "AND" fusion decoding rules, and a randomized region between them. All other MAP and cost functions have one unique solution and identical relay threshold tests.

optimality condition equations in equation 3.47 for the Gaussian case, we have respectively,

$$\frac{d\hat{\beta}(\alpha)}{d\alpha} = \frac{\frac{d\tilde{\beta}(t)}{dt}}{\frac{d\tilde{\alpha}(t)}{dt}} = \frac{\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(-t+m)^2}{2}\right)}{-\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(t+m)^2}{2}\right)} = -\exp(2mt) \tag{3.58}$$

and

$$\eta\frac{\alpha}{1-\beta} = \eta\frac{\mathcal{Q}(t+m)}{1-\mathcal{Q}(-t+m)} = \eta\frac{\mathcal{Q}(t+m)}{\mathcal{Q}(t-m)} \tag{3.59}$$

Let t be the normalized parameter for Relay A, so its threshold decisions will be $\alpha = \alpha(t)$, and let t' be that of Relay B whose decisions will be $\alpha' = \alpha(t')$. Substituting in equations 3.58 and 3.59 for the left vertex curve, the local extrema conditions of equation 3.47 must satisfy $-\frac{d\hat{\beta}(\alpha)}{d\alpha} = \eta\frac{\alpha'}{1-\beta'}$ and $-\frac{d\hat{\beta}(\alpha')}{d\alpha'} = \eta\frac{\alpha}{1-\beta}$ which becomes

$$\exp(2mt) = \frac{\mathcal{Q}\left(t'+m\right)}{\mathcal{Q}(t'-m)}\eta \quad \text{and} \quad \exp(2mt') = \frac{\mathcal{Q}(t+m)}{\mathcal{Q}(t-m)}\eta \tag{3.60}$$

and reduces to

$$t = \frac{1}{2m}\ln\frac{\mathcal{Q}\left(t'+m\right)}{\mathcal{Q}(t'-m)}\eta \quad \text{and} \quad t' = \frac{1}{2m}\ln\frac{\mathcal{Q}(t+m)}{\mathcal{Q}(t-m)}\eta. \tag{3.61}$$

Let $G(t) = \frac{1}{2m}\ln\frac{Q(t+m)}{Q(t-m)} + \frac{1}{2m}\ln\eta$. Then equation 3.61 can be restated as $t = G(t')$ and $t' = G(t)$ which can be combined to $t = G(G(t))$ or $t' = G(G(t'))$. All solutions to equations 3.61 are also fixed points of $t = G(G(t))$. Furthermore, a fixed point solution t to $t = G(G(t))$ has the corresponding t' to the solution pair (t, t') defined in equation 3.61. The stationary points are the fixed points to $t = G(G(t))$ and we will now show for the Gaussian case that there exists exactly one fixed point solution. If there exists only one solution, then it must be of the form $t = t'$ as proven in lemma 3.4.2.

As shown previously, there is only one stable[2] solution t^* to the fixed point equation $t = G(G(t))$ iff $\forall t < t^*,\quad G(G(t)) > t^*$ and $\forall t > t^*,\quad G(G(t)) < t^*$. In other words, if $G(G(t)) - t$ is monotonically increasing, then there is only one stable solution. Equivalently, if $\frac{d}{dt}G(G(t)) < 1$ for all t, then there is only one fixed point solution.[3] We will show that there is only one stable solution by showing

$$\frac{d}{dt}G(G(t)) = G'(G(t))G'(t) = G'(t')G'(t) < 1. \tag{3.62}$$

Since

$$\frac{d}{dt}G(t) = G'(t) = \frac{1}{2m}\left[\frac{q(t+m)}{Q(t+m)} - \frac{q(t-m)}{Q(t-m)}\right] \tag{3.63}$$

Using equation 3.63 in equation 3.62, we need to show that

$$\frac{d}{dt}G(G(t)) = \frac{1}{2m}\left[\frac{q(t'+m)}{Q(t'+m)} - \frac{q(t'-m)}{Q(t'-m)}\right]\frac{1}{2m}\left[\frac{q(t+m)}{Q(t+m)} - \frac{q(t-m)}{Q(t-m)}\right] < 1 \tag{3.64}$$

The left side of equation 3.64 is the product of two functions evaluated at different points. Note that $\frac{1}{2m}\left[\frac{q(t+m)}{Q(t+m)} - \frac{q(t-m)}{Q(t-m)}\right]$ is merely the slope between two points of the function $\frac{q(t)}{Q(t)}$ evaluated at $t+m$ and $t-m$.

The Mean Value Theorem (MVT) says that if $y = f(x)$ is continuous and differentiable in the interval $[a, b]$, then there exists at least one $x \in [a, b]$ such that $f'(x) = \frac{f(b)-f(a)}{b-a}$. If a function f satisfies $f'(x) < 1$ everywhere, then it follows that $\frac{f(b)-f(a)}{b-a} < 1$ for all $b \neq \alpha$. Since m and y can be arbitrary real numbers and $\frac{q(t)}{Q(t)}$ is continuous and differentiable everywhere, it is sufficient to show that $\frac{d}{dt}\left(\frac{q(t)}{Q(t)}\right) < 1$ for all t. This is shown in lemma 3.5.2. Thus, the left decoding rule ("OR") corresponding to the left vertex curve has only one unique solution which implies that the relay strategies are identical and the solution is the global minimum.

[2]A stable point of the algorithm means that starting at any arbitrary value and iterating, the algorithm will converge to the solution.

[3]Note that the converse is not necessarily true – having only one fixed point solution does not imply that the derivative is less than one everywhere.

The argument is the same for the right decoding rule ("AND"). The necessary local extrema conditions of the right vertex curve must be a stationary point given in equation 3.48 which must satisfy $-\frac{d\hat{\beta}(\alpha)}{d\alpha} = \eta\frac{1-\alpha'}{\beta'}$ and $-\frac{d\hat{\beta}(\alpha')}{d\alpha'} = \eta\frac{1-\alpha}{\beta}$. This reduces to

$$t = \frac{1}{2m}\ln\frac{\mathcal{Q}\left(-t'-m\right)}{\mathcal{Q}\left(-t'+m\right)}\eta \quad \text{and} \quad t' = \frac{1}{2m}\ln\frac{\mathcal{Q}\left(-t-m\right)}{\mathcal{Q}\left(-t+m\right)}\eta \tag{3.65}$$

Like the technique for the left vertex solution, this can be combined to $G(t) = \frac{1}{2m}\ln\frac{\mathcal{Q}(-t-m)}{\mathcal{Q}(-t+m)} + \frac{1}{2m}\ln\eta$ which has derivative

$$h'(t) = \frac{1}{2m}\left[\frac{q(-t-m)}{\mathcal{Q}\left(-t-m\right)} - \frac{q(-t+m)}{\mathcal{Q}\left(-t+m\right)}\right] \tag{3.66}$$

Follow the same argument as for the left decoder rule. Since there is only one fixed point solution, it must be identical and the global minimum. $\square$

Lemma 3.5.2 *The ratio of the Gaussian density to its cumulative distribution function $L(x) = q(x)/\mathcal{Q}(x)$ is monotonically increasing and its derivative $L'(x) = \frac{d}{dx}L(x)$ is bounded in the following regions, $0 \leq L'(x) \leq \frac{2}{\pi}$ for $x \leq 0$ and $\frac{2}{\pi} \leq L'(x) \leq 1$ for $x \geq 0$.*

Proof: First, since $\mathcal{Q}(x), q(x) > 0$, we have $L(x) > 0, \forall x \in \Re$. For the first part of the lemma, we show that $L(x)$ is monotonically increasing by showing that $W(x) = \frac{1}{L(x)} = \frac{\mathcal{Q}(x)}{q(x)}$ is monotonically decreasing. Note that $\lim_{x\to\infty} W(x) = 0$ and $\lim_{x\to-\infty} W(x) = \infty$. If, in addition, $W(x)$ is convex, then $W(x)$ is monotonically increasing. We will show $W(x)$ is convex by showing that $W''(x) > 0$ for all x.

$$\begin{aligned} W(x) &= e^{\frac{x^2}{2}}\int_x^\infty e^{\frac{-u^2}{2}}du \\ &= \int_x^\infty e^{\frac{-(u^2-x^2)}{2}}du \\ &= \int_x^\infty e^{\frac{-(u+x)(u-x)}{2}}du \quad \text{let} \quad y = u - x \\ &= \int_0^\infty e^{\frac{-y^2}{2}}e^{-xy}dy \end{aligned} \tag{3.67}$$

Taking the second derivative of $W(x)$, we have

$$\frac{d^2W(x)}{dx^2} = \int_0^\infty y^2 e^{\frac{-y^2}{2}}e^{-xy}dy \tag{3.68}$$

which is positive for all x. This implies $W(x) = \frac{Q(x)}{q(x)}$ is convex [4].

Now for the second part of the lemma. We first prove that $\frac{2}{\pi} \leq L'(x) \leq 1$ for $x \geq 0$. The derivative of $L(x)$ is $\frac{dL(x)}{dx} = \frac{-xq(x)}{Q(t)} + \frac{q^2(x)}{Q(t)^2} = L(x)(L(x) - x)$. There are many known bounds on $L(x)$ for $x > 0$. We will use the following bound of the ratio of the Gaussian density to its distribution function for $x > 0$,

$$\frac{x}{2} + \frac{1}{2}\sqrt{x^2 + \frac{8}{\pi}} \leq \frac{q(x)}{Q(x)} \leq \frac{x}{2} + \frac{1}{2}\sqrt{x^2 + 4} \tag{3.69}$$

For all $x \geq 0$, use of the lower bound in 3.69 gives $L(x) - x \geq x[\frac{1}{2}\sqrt{x^2 + \frac{8}{\pi}} - \frac{1}{2}] \geq 0$. We have for $x \geq 0$, $\frac{dL(x)}{dx}$ is upper-bounded by

$$\begin{aligned} \frac{dL(x)}{dx} &= L(x)(L(x) - x) \\ &\leq \left(\frac{x}{2} + \frac{1}{2}\sqrt{x^2+4}\right)\left(-\frac{x}{2} + \frac{1}{2}\sqrt{x^2+4}\right) \\ &= -\frac{x^2}{4} + \frac{x^2}{4} + 1 = 1 \end{aligned} \tag{3.70}$$

and lower-bounded by

$$\begin{aligned} \frac{dL(x)}{dx} &= L(x)(L(x) - x) \\ &\geq \left(\frac{x}{2} + \frac{1}{2}\sqrt{x^2+\frac{8}{\pi}}\right)\left(-\frac{x}{2} + \frac{1}{2}\sqrt{x^2+\frac{8}{\pi}}\right) \\ &= -\frac{x^2}{4} + \frac{x^2}{4} + \frac{2}{\pi} = \frac{2}{\pi} \end{aligned} \tag{3.71}$$

Combining both lower and upper bounds, we have $\frac{2}{\pi} \leq L'(x) \leq 1$ for $x > 0$.

Now for $x < 0$, we establish convexity by showing that $\frac{d^2L(x)}{dx^2} > 0$ in this region. We prove by contradiction. Note that $L(0) = \sqrt{\frac{2}{\pi}}$, $L'(0) = \frac{2}{\pi}$, and $L''(0) > 0$. First,

$$\begin{aligned} L''(x) &= L(x)\left(2L^2(x) - 3xL(x) + x^2 - 1\right) \\ &\leq L(x)\left((\sqrt{2}L(x) - x)^2 - 1\right) \\ &= L(x)\left(\sqrt{2}L(x) - x + 1\right)\left(\sqrt{2}L(x) - x - 1\right) \end{aligned} \tag{3.72}$$

[4]Note that if a function $W(x)$ is convex , this does not imply that $1/W(x)$ is concave.

Suppose that there exists some support set such that $L''(x) \leq 0$. That means those x must satisfy $\sqrt{2}L(x) - x - 1 < 0$. Since $\sqrt{2}L(x) - x - 1$ is a continuous function, there must exist some x_1 such that $L(x)$ is convex for $x \in [x_1, 0]$. This means that $\sqrt{2}L(x) - x - 1 > 0$ in $[x_1, 0]$. Choose x_1 to be the largest x such that $\sqrt{2}L(x) - x - 1 = 0$. Since $L(x)$ is convex for $x \in [x_1, 0]$, $L(x_1) \geq L(0) + x_1 \frac{d}{dx} L(x) \mid_{x=0} = \sqrt{\frac{2}{\pi}} + x_1 \frac{2}{\pi}$. Now, $\sqrt{2}L(x_1) - x_1 - 1 \geq \frac{2}{\sqrt{\pi}} + x_1 \frac{2\sqrt{2}}{\pi} - x_1 - 1 > 0$. A contradiction!

Thus, $L'' > 0$ for all $x \leq 0$, so L must be convex for $x < 0$. This implies that the derivative must be monotonically increasing from $\lim_{x \to -\infty} L'(x)$ to $L'(0)$ which gives $0 \geq L'(x) \leq \frac{2}{\pi}$. $\square$

3.6 Appendix A

We show in this appendix that the inequality in 3.23 is strict by showing that $\{V_{ik}, V_{il}, V_{jk}, V_{jl},\}$ cannot all lie on a straight line. We prove this for the case when the relays use independent randomization and the fusion point knows the relay selection. Then the assertion must also be true when the fusion point does not know the relay selection.

It is sufficient to prove the result for randomizing at each relay between 2 vertex points. The results hold for relay randomization with more than 2 vertex points since a randomization will be a convex combination of those vertex points. Without loss of generality let Relay A randomize between 2 vertex points with probability μ and Relay B between 2 vertex points with probability μ', where $0 < \mu, \mu' < 1$. Let $\{V_i, V_j, V_k, V_l\}$ be distinct vertex points of the relay error curve. Assume without loss of generality that $i < j$ and $k < l$ and $i < k$. There are 3 cases to consider. The relays randomize from a combined set of either 2 relay vertex points, 3 relay vertex points, or 4 relay vertex points. These cases are listed individually below.

(1) Relay A and B independently randomize between 2 vertex points, say V_i and V_j. The fused vertex points are $\{V_{ii}, V_{ij}, V_{jj}\}$.

(2) Relay A and B independently randomize from a combined set of 3 vertex points, $\{V_i, V_j, V_l\}$. Without loss of generality assume V_j is the common vertex point shared. The fused vertex points are $\{V_{ij}, V_{il}, V_{jj}, V_{jl}\}$.

(3) Relay A and B independently randomize from a combined set of 4 vertex points, $\{V_i, V_j, V_k, V_l\}$. Assume Without loss of generality Relay A randomizes between V_i and V_j and Relay B randomizes between V_k and V_l. The fused vertex points are $\{V_{ik}, V_{il}, V_{jk}, V_{jl}\}$.

We prove the following 3 assertions corresponding to the 3 cases listed above.

(1) $\{V_{ii}, V_{ij}, V_{jj}\}$ can be colinear if they are not all left or all right vertex points. $\{P_e(V_i, V_i), P_e(V_i, V_j), P_e(V_j, V_j)\}$ can all have the same error probability. This implies that two relays independently randomizing between the same two vertex points can be optimal.

(2) $\{P_e(V_i, V_j), P_e(V_i, V_k), P_e(V_j, V_j), P_e(V_j, V_k)\}$ cannot all have the same error probability.

(3) $\{P_e(V_i, V_k), P_e(V_i, V_l), P_e(V_j, V_k), P_e(V_j, V_l)\}$ cannot all have the same error probability.

We first show, for all 3 cases, that if all of the fused points in a set are all left vertex points or all right vertex points, then they cannot be colinear. We will show this for the left vertex points. The proof for the right vertex points is similar.

(1) Calculating the slopes of $V_{ii}^L \to V_{ij}^L$ and $V_{ij}^L \to V_{jj}^L$ and checking for equality, we get $\left(\frac{1-\beta_i}{\alpha_i}\right)\left(\frac{\beta_i-\beta_j}{\alpha_j-\alpha_i}\right) \neq \left(\frac{1-\beta_j}{\alpha_j}\right)\left(\frac{\beta_i-\beta_j}{\alpha_j-\alpha_i}\right)$, which implies that $\Lambda\left(V_{ii}^L \to V_{ij}^L\right) \neq \Lambda\left(V_{jj}^L \to V_{ij}^L\right)$. Thus, $V_{ii}^L, V_{ij}^L, V_{jj}^L$ cannot be colinear.

(2) Since $\left(\frac{1-\beta_i}{\alpha_i}\right)\left(\frac{\beta_j-\beta_k}{\alpha_k-\alpha_j}\right) \neq \left(\frac{1-\beta_j}{\alpha_j}\right)\left(\frac{\beta_j-\beta_k}{\alpha_k-\alpha_j}\right)$, this implies that $\Lambda\left(V_{ij}^L \to V_{ik}^L\right) \neq \Lambda\left(V_{jj}^L \to V_{jk}^L\right)$. Thus, $V_{ij}^L, V_{ik}^L, V_{jj}^L, V_{jk}^L$ are not colinear.

(3) Since $\left(\frac{1-\beta_i}{\alpha_i}\right)\left(\frac{\beta_k-\beta_l}{\alpha_k-\alpha_l}\right) \neq \left(\frac{1-\beta_j}{\alpha_j}\right)\left(\frac{\beta_k-\beta_l}{\alpha_k-\alpha_l}\right)$, this implies that $\Lambda\left(V_{ik}^L \to V_{il}^L\right) \neq \Lambda\left(V_{jk}^L \to V_{jl}^L\right)$. Thus, $V_{ik}^L, V_{il}^L, V_{jk}^L, V_{jl}^L$ are not colinear.

We now prove the 3 assertions by showing whether the fused vertex points in each set can be colinear.

(1) Section 3.4 shows that the relay error curve with 2 vertex points can be constructed such that either $(V_{11}^R \to V_{22}^L \to V_{12}^R)$ or $(V_{22}^L \to V_{11}^R \to V_{12}^R)$ is colinear.

(2) Since $V_i < V_j < V_l$, if we start with any of the 4 fused vertex points in the set $\{V_{ij}, V_{il}, V_{jj}, V_{jl}\}$, and arbitrarily assign it either a left or right vertex point, then according to lemma 3.4.1 the 3 other points must be of the same kind. For example, suppose V_{ij} were a right vertex point, then this implies that V_{il} must also be a right vertex point by lemma 3.4.1. In other words, if V_{ij}^L is on $F(\phi)$ at a given η, then V_{il}^R cannot be on $F(\phi)$. Continuing with the same argument with V_{il}^R, we have $V_{il}^R \implies V_{jl}^R$, and then $V_{jl}^R \implies V_{jj}^R$. Therefore, by lemma 3.4.1, all 4 points must be of the same type and if so, then they cannot be colinear as just shown above. This happens to all 8 different possibilities of initially choosing a vertex point to be either right or left.

(3) There are exactly 3 types of ordering, namely $V_i \to V_j \to V_k \to V_l$ and $V_i \to V_k \to V_j \to V_l$ and or $V_i \to V_k \to V_l \to V_j$. The 4 vertex points are $\{V_{ik}, V_{il}, V_{jk}, V_{jl}\}$. Start with any vertex point and arbitrarily assign it a left or right vertex point. Then use the same argument as in (2). It turn out that all 4 vertex points must be of the same type (all left or all right). If they are all the same type, then they cannot be colinear. This is true for all 24 possibilities.

3.7 Appendix B

We demonstrate by construction the existence of a relay error curve which has an infinite number of stationary vertex points. We will show this for the left vertex point which is the "OR" decoding rule and $\eta = 1$. The point α represents the point (α, β) on the relay error curve. A stationary point (α_{-j}, α_j) implies that the pair of points $((\alpha_{-j}, \beta_{-j}), (\alpha_j, \beta_j))$ on the relay error curve satisfies $\alpha_j = \tilde{\alpha}(\frac{\alpha_{-j}}{1-\beta_{-j}})$ and $\alpha_{-j} = \tilde{\alpha}(\frac{\alpha_j}{1-\beta_j})$. We will construct a relay error curve with countably infinite stationary vertex points. These stationary points nest out from the identical solution stationary point (α_0, α_0) as described in theorem 3.4.1 and satisfy the conditions of equation 3.53 for $\eta = 1$. The construction process follows the construction conditions of figure 3-18 and described as follows.

(1) Start the construction of the relay error curve with the 3 points $(\alpha_{-1}, \beta_{-1}) = (0.1, 0.2)$, $(\alpha_0, \beta_0) = (0.25, 0.079)$, $(\alpha_1, \beta_1) = (0.8, 0.06)$. Then (α_{-1}, α_1) and (α_0, α_0) are stationary points.

(2) Let the point B be the β-intercept of the line through α_{-1} with slope $\frac{\alpha_1}{1-\beta_1}$ as shown in figure 3-22. Let C be the midpoint between B and $(\alpha_{-1}, \beta_{-1})$. Let m_1 be the slope of the line through $(0, 1)$ and C.

(3) Let D be the α-intercept of the line of slope $1/m_1$ through the point (α_1, β_1). Then $\alpha_1 < D < 1$ and $\{\alpha_1,\ D,\ (1, 0)\}$ form a triangular region. Let (α_2, β_2) be the center of mass of this triangular region. This vertex point will be on the relay error curve being constructed so it will be referred to as α_2 as shown in figure 3-22.

(4) Let $m_2 = \frac{\alpha_2}{1-\beta_2}$. The line of slope m_2 through the point $(\alpha_{-1}, \beta_{-1})$ will intersect the line of slope m_1 through $(0, 1)$. Call this intersection point E (shown in figure 3-23).

(5) Choose the midpoint between C and E and call this point $(\alpha_{-2}, \beta_{-2})$. Then the set of points $\{\alpha_{-2}, \alpha_{-1}, \alpha_0, \alpha_1, \alpha_2\}$ form a convex curve (figure 3-23). Now, (α_{-2}, α_2) is a stationary point since by construction $\alpha_2 = \tilde{\alpha}\left(\frac{\alpha_{-2}}{1-\beta_{-2}}\right)$ and $\alpha_{-2} = \tilde{\alpha}\left(\frac{\alpha_2}{1-\beta_2}\right)$.

(6) Continue the same procedure using the vertex points α_{-2} and α_2 to construct the next set of points α_{-3} and α_3 on the relay error curve where (α_{-3}, α_3) is the next nested stationary point. Iterate on α_{-3} and α_3 to construct the next nested stationary point (α_{-4}, α_4), and so forth.

The iterative procedure described above constructs a relay error curve with a countably infinite number of stationary points for the left vertex point.

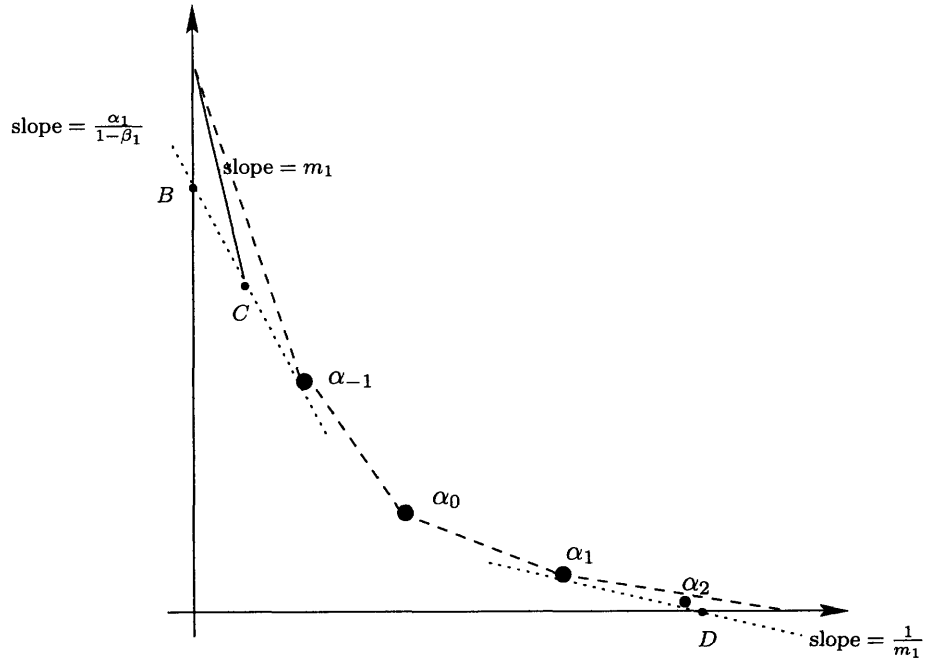


Figure 3-22: The construction of a relay error curve with an infinite number of stationary points. Given initial points $\{\alpha_{-1}, \alpha_0, \alpha_1\}$, where (α_{-1}, α_1) and (α_0, α_0) are stationary points, construct α_2, where α_2 is the center of mass of the triangular region formed by $(\alpha_1, D, (1, 0))$.

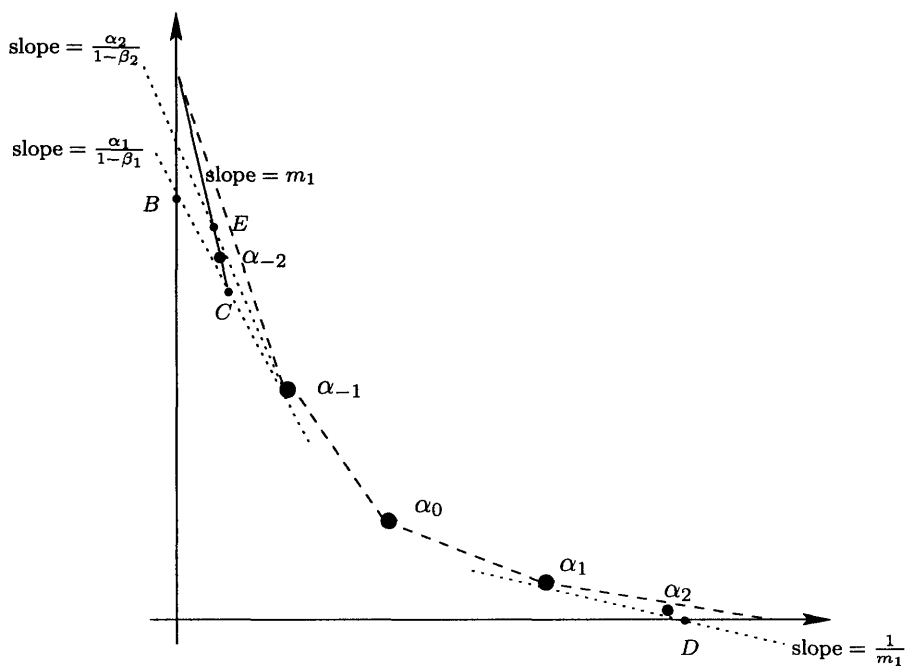


Figure 3-23: The construction of a relay error curve with an infinite number of stationary points continued. Construct vertex point α_{-2}, the midpoint between E and C, then (α_{-2}, α_2) is a stationary point.

Chapter 4

Two Relays with Arbitrary Output Alphabets

In this chapter, we continue with a system of 2 relays but remove the restriction of one bit per relay. Unless otherwise noted, all the results are general and apply to non-identical relay channels which are described by relay error curves $\hat{\beta}(\alpha)$ and $\hat{\beta}'(\alpha')$. We first study the situation when one of the relays is the fusion center and the other relay is restricted to a binary output. We use the insights from this problem to provide a complete optimal solution when both relays are constrained to send a symbol from an alphabet of arbitrary size. Suppose there is a limit of M symbols for relay A and M' for Relay B (assume M and M' are less than or equal to the the number of vertex points on their respective relay error curves). The fusion point receives one of at most MM' distinct pairs of symbols. From conditional independence of the outputs given the inputs, the likelihood ratio of receiving a symbol pair (i, j) is the product of the likelihood ratios of receiving each individual symbol, $\Lambda(i, j) = \Lambda(i)\Lambda(j)$. We have shown in Chapter 2 the well-known result that a sufficient condition for optimality with $M = M' = 2$ is for each relay to use a threshold test, i.e., a binary quantization of its LR. We will show, via a simple argument with detailed explanation, the known result that it is sufficient for optimality if the strategies of Relays A and B are interval quantizations of their respective LR. That is, with an alphabet of M symbols, the relay quantizes the LR into M intervals and associates each interval with one of the alphabet letters. Any such quantization will be represented by a vector $\bar{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{M-1})$ where α_i represents the boundary between i and $(i + 1)$. Specifically, we will show that, conditional on Relay A, the optimal encoding points for Relay B, $\bar{\alpha}' = (\alpha'_1, \alpha'_2, \ldots, \alpha'_{M'-1})$, lie on the Relay B error curve; and vice versa, conditional on

Relay B, the optimal encoding points for Relay A, $\bar{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{M-1})$ lie on the Relay A error curve. In this chapter, we will determine the necessary optimality conditions and give an iterative algorithm to find the optimal encoding points for Relays A and B. Lastly, we determine the conditions when the encoding rate of the relays can be reduced without a deterioration in performance. This case will only arise on the discrete portions of the relay error curves and under very specific conditions.

4.1 Binary-Ternary Case

Suppose first that the fusion center is co-located with one of the relays, say Relay B. This scenario is equivalent to the distributed system model we have been studying except that Relay B sees what Relay A is sending to the fusion center, makes a decision based both on its own observation and what the other relay is sending to the fusion center, and then sends its decision to the fusion point. This scenario is also equivalent to a system where the fusion point makes a decision after receiving full information from Relay B and limited information from Relay A (figure 4-1(A) and (B)). In this section, we will prove that in fact the fusion point does not require full information from Relay B. Specifically, if Relay A uses an alphabet of size M, then an alphabet of size $M+1$ from Relay B is sufficient for optimality. That is, the fusion center just needs one more alphabet symbol from Relay B than from Relay A. Furthermore, we will determine joint optimality conditions for both relays. We will show that figure 4-1(B) is equivalent to 4-1(C) by proving that figure 4-1(A) is equivalent to 4-1(C). It then follows that all three systems illustrated in figure 4-1 are equivalent.

For simplicity, first consider a two relay system where Relay A has an arbitrary fixed test at (α, β) which is not necessarily on the relay error curve; it broadcasts its binary choice to Relay B, which acts as the fusion point (figure 4-2 and figure 4-1(A)). From the equivalent parallel structures of figures 4-1(A) and (C), what should Relay B send to the fusion point? We will show that as far as the fusion point is concerned, all it needs is at most ternary information, which translates to a 3 level quantization of the LR from Relay B. How can it possibly be that the fusion center does not need the full information from Relay B? That is, how is transmitting a ternary symbol from Relay B equivalent to sending its entire observation information?

Suppose that the priors are fixed with $\eta = \frac{p_1}{p_0}$. We now look at the situation at Relay B when it has observed the output from Relay A but before it pays any attention to its own output from channel 2. Suppose that Relay A uses an arbitrary test, (α, β) not necessarily on its error

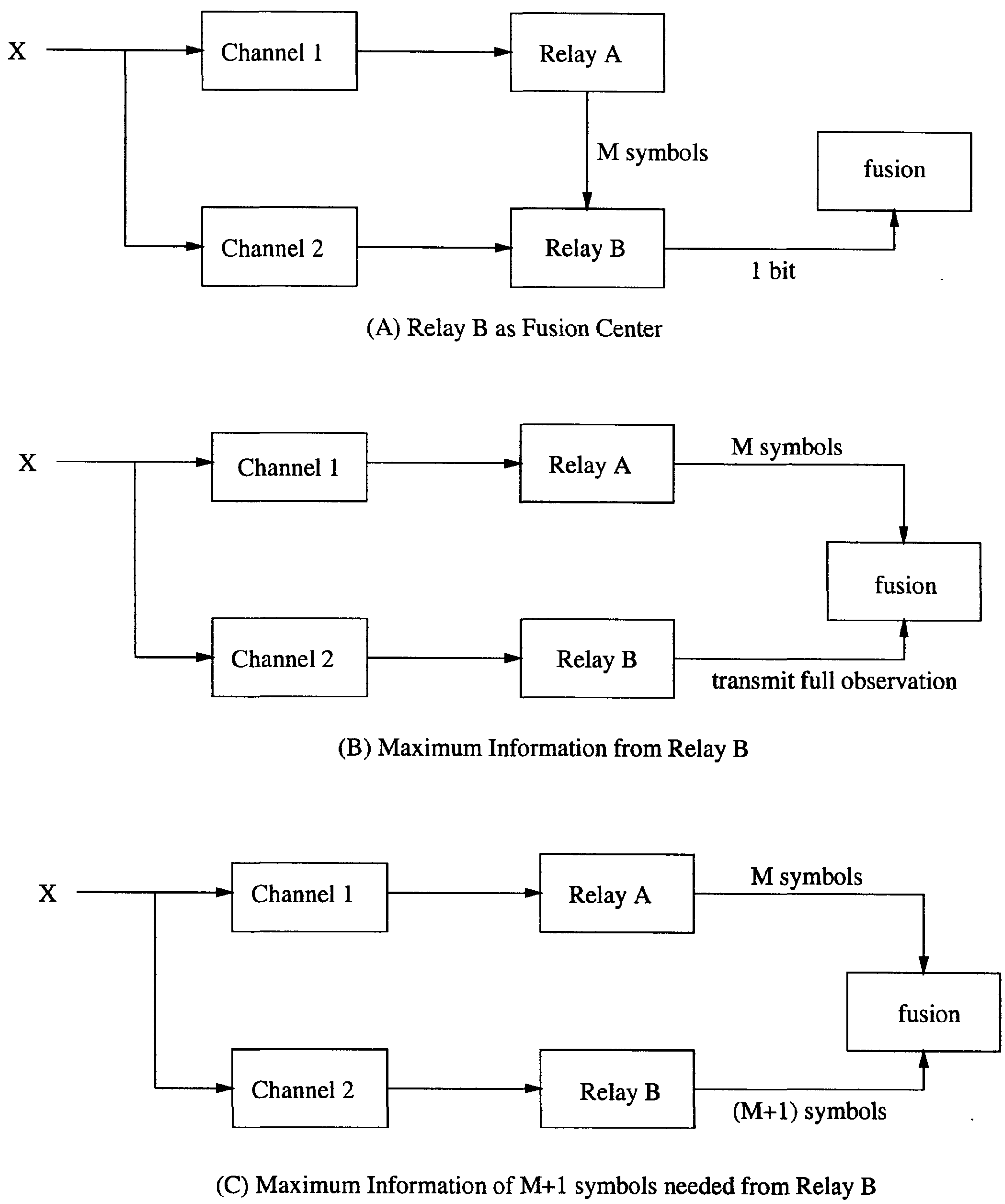


Figure 4-1: Equivalent performance of all 3 systems (A), (B), and (C) when Relay A is restricted to M symbols. Channel 1 and channel 2 are arbitrary and can be different.

curve. Then $\Pr(Z_A = 0|X = 1) = \alpha$ and $\Pr(Z_A = 1|X = 0) = \beta$.

From the perspective of Relay B, if it receives a "0" from relay A, then the new ratio of a priori probabilities changes to $\eta_0 = \frac{\Pr(X=1|Z_A=0)}{\Pr(X=0|Z_A=0)} = \frac{\Pr(Z_A=0|X=1)p_1}{\Pr(Z_A=0|X=0)p_0} = \frac{\alpha}{1-\beta}\eta$. Relay B makes the optimal decision based on this new ratio of a priori probabilities, (call this the *ratio of conditional a priori probabilities*), and forwards its one bit decision to the fusion center. Similarly, if Relay B receives a "1" from relay A, then the *ratio of conditional a priori probabilities* becomes $\eta_1 =$

$\frac{\Pr(X=1|Z_A=1)}{\Pr(X=0|Z_A=1)} = \frac{\Pr(Z_A|X=1)p_1}{\Pr(Z_A|X=0)p_0} = \frac{1-\alpha}{\beta}\eta$. Note that the ratio of conditional a priori probabilities is equal to the inverse of the received symbol multiplied by η. Based on either conditional a priori information, Relay B makes a binary decision by comparing the LR for its channel observation with either η_0 or η_1. This translates to having two threshold tests. Define (α_1', β_1') as the threshold test for η_1 (when Relay B receives a "1" from Relay A) and (α_2', β_2') as the threshold test for η_0.

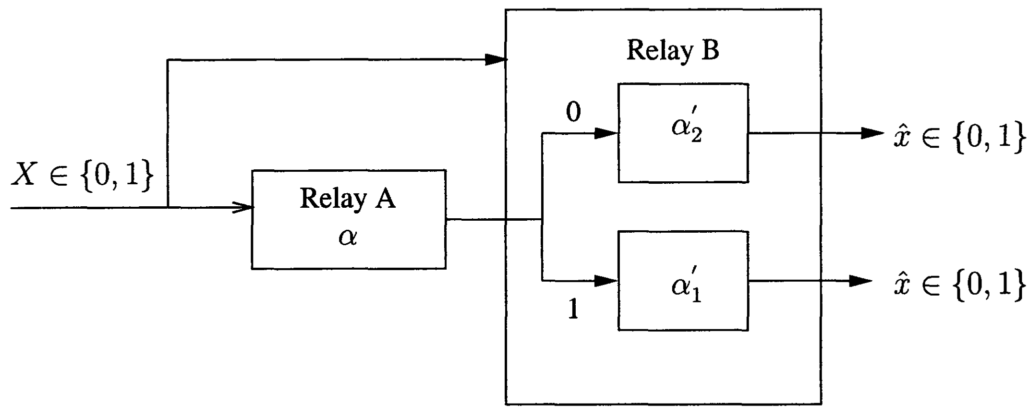


Figure 4-2: Relay A sends its one bit observation to Relay B which acts as the fusion center. Relay B forms a threshold test at α_2' if it receives a "0," and a threshold test at α_1' if a "1" is received.

We next show the equivalence of figure 4-1(A) and (C) for the binary-ternary case. That is, for this 2×3 system, the necessary two threshold tests from Relay B in figure 4-1(A) can be represented by a single ternary symbol from Relay B in figure 4-1(C).

In the situation of figure 4-1(C), Relay B does not observe the output of Relay A, but it could still send two binary signals, one based on the test (α_1', β_1') and the other on (α_2', β_2'). The fusion point will implement what Relay B did in figure 4-2. It sees the output from Relay A and then observes the corresponding binary symbol from Relay B to make its final decision. Specifically, if the fusion point sees a "0" from Relay A, it needs the result of the threshold test (α_2', β_2') at Relay B. The fusion point decodes a "0" if the observation of Relay B is less than α_2' and a "1" if greater than α_2'. If the fusion point sees a "1" from Relay A, it needs the results of the threshold test (α_1', β_1') at Relay B; it decodes a "0" if the observation of Relay B is less than α_1' and a "1" if greater than α_1'.

We can now understand the equivalence of figure 4-1(A) and 4-1(C) for this 2×3 system. From the standpoint of the fusion point, it will decode a "0" or "1" depending on what it receives from Relay A and what side of the threshold (either α_1' or α_2') the observation of Relay B lies on. This can be achieved with a 3 level quantizer using the optimal encoding points from Relay B.

Specifically, the mapping $[0, \alpha'_1] \to 0$ and $(\alpha'_1, \alpha'_2] \to 1$ and $(\alpha'_2, 1] \to 2$ at the output of Relay B answers both threshold questions at the fusion point. In other words, this ternary symbol from Relay B tells the fusion point the results of its threshold tests at (α'_1, β'_1) and (α'_2, β'_2). That is, if "0" is received from Relay A, the ternary symbol indicates whether the observation of Relay B is less than or greater than α'_2. Similarily, if "1" is received, the ternary symbol indicates whether the observation of Relay B is less than or greater than α'_1. This is summarized in the following lemma.

Lemma 4.1.1 *Given η and given an arbitrary binary test at Relay A, a ternary test at Relay B suffices for an optimal final decision. The ternary test at B is a quantization $\bar{\alpha}' = (\alpha'_1, \alpha'_2)$ of the Relay B error curve where α'_1 is the optimal test conditional on "1" from Relay A and α'_2 is the optimal test conditional on "0" from Relay A.*

For the remainder of this section, we assume that the relay error curve is strictly convex, so that every η-tangent to the error curve has a unique tangent point. The general curve will be treated in the last section and just requires some slight modifications of the treatment in this section. Thus, (α'_1, β'_1) is uniquely determined from $\eta_1 = \frac{1-\alpha}{\beta}\eta$, with $\alpha'_1 = \tilde{\alpha}(\eta_1)$. Similarily, (α'_2, β'_2) is uniquely determined from $\eta_0 = \frac{\alpha}{1-\beta}\eta$ with $\alpha'_2 = \tilde{\alpha}(\eta_0)$. Since $\eta_0 < \eta < \eta_1$, $\alpha'_1 < \alpha'_2$. Finding α'_1, α'_2 can be done graphically since $-\frac{\alpha}{1-\beta}$ is the inverse of the slope from (α, β) to the point $(0, 1)$ and $-\frac{1-\alpha}{\beta}$ is the inverse of the slope from (α, β) to the point $(1, 0)$.

The symbol outputs $\bar{Z} = (z_1, z_2)$ from the relays can be represented by a 2×3 matrix at the fusion point. Let the rows represent the symbol outputs from Relay A and the columns represent the symbol outputs from Relay B. Given η, given a choice of α at Relay A, and given $\bar{\alpha}' = (\alpha_1, \alpha_2)$ at Relay B, let the 2×3 *Decoder Matrix* $D(\alpha, \bar{\alpha}')$ represent the optimal fusion decoding rule. That is, each entry $D_{ij} = \hat{x}(i, j)$ is the optimal fusion decoder map given that symbol i from Relay A and symbol j from Relay B was received.

Denote $(\alpha, \arg\min_{\bar{\alpha}'} P_e^{(\eta)}(\alpha, \bar{\alpha}'))$ as a *conditional solution pair*, such that given η and arbitrary test (α, β) at Relay A, the optimal encoding points at Relay B is $\bar{\alpha}' = \arg\min_{\bar{\alpha}'} P_e^{(\eta)}(\alpha, \bar{\alpha}')$ where $P_e^{(\eta)}(\alpha, \bar{\alpha}')$ is the probability of error given η and $(\alpha, \bar{\alpha}')$.

From the discussion of the 2×3 case, for every $(\alpha, \arg\min_{\bar{\alpha}'} P_e^{(\eta)}(\alpha, \bar{\alpha}'))$, the optimal decoder matrix at the fusion point is unique and is

$$D(\bar{\alpha}, \bar{\alpha}') = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \tag{4.1}$$

At this point, we know the optimal decoding rule and how to choose the optimal $\bar{\alpha}'$ conditional on any given α and η. That is, we know how to determine $\bar{\alpha}'$ in $(\alpha, \arg\min_{\bar{\alpha}'} P_e^{(\eta)}(\alpha, \bar{\alpha}'))$. However, for overall optimality, optimization over Relay A's set of possible tests is necessary. We now show that the optimal α for a given $\bar{\alpha}'$ is a threshold test determined by the ratio of conditional a priori probabilities given that a "1" was received by Relay B.

From the optimal decoder matrix $D(\alpha, \bar{\alpha}')$ in (4.1), when Relay B transmits a "0" or "2," the fusion point uses the trivial decoder, ignoring Relay A. However, if Relay B sends a "1," the fusion point makes a decision based on Relay A's output. What is the optimal $''\alpha''$ for Relay A? We implement the same idea just discussed but with the added condition that the decoder matrix of 4.1 must be implemented, since the optimal MAP decision for every 2×3 system uses this unique decoding rule. Interchange the roles of Relay A and B in figure 4-1(A). Relay A is now the fusion point and receives a ternary symbol from Relay B. Based on its own observation, the given η, and the decoder matrix 4.1, Relay A must make an optimal choice and forward its decision to the fusion point. When Relay A receives a "0" from Relay B, it decodes a "0"; when a "2" is received from Relay B, it decodes a "1"; finally, when Relay A receives a "1" from Relay B, it makes an optimal decision based on the conditional ratio of a priori probabilities $\eta_1 = \frac{\Pr(X=1|Z_B=1)}{\Pr(X=0|Z_B=1)} = \frac{\Pr(Z_B|X=1)p_1}{\Pr(Z_B|X=0)p_0} = \frac{\alpha_2'-\alpha_1'}{\beta_1'-\beta_2'}\eta$; this is a threshold test. Figure 4-1(A) is equivalent to figure 4-1(B) with the the labels of Relay A and B interchanged and constrained to binary and ternary symbols, respectively. Thus, the optimal Relay A encoding point for a given $\bar{\alpha}'$ is $\alpha = \tilde{\alpha}\left(\frac{\alpha_2'-\alpha_1'}{\beta_1'-\beta_2'}\eta\right)$, and hence, $(\arg\min_{\alpha} P_e^{(\eta)}(\alpha, \bar{\alpha}'), \bar{\alpha}')$ is a *conditional solution pair*, the α is optimal given $\bar{\alpha}'$. For this pair and newly chosen α, the old $(\bar{\alpha}')$ chosen optimally based on the original α need no longer be a conditional solution pair. The global optimal $(\alpha, \bar{\alpha}')$ must satisfy both conditional solution pairs and be a stationary point, $(\arg\min_{\alpha} P_e^{(\eta)}(\alpha, \bar{\alpha}'), \arg\min_{\bar{\alpha}'} P_e^{(\eta)}(\alpha, \bar{\alpha}'))$.

We now have the joint necessary conditions the optimal $(\alpha, \bar{\alpha}')$ must satisfy for a given η. We also now have an iterative algorithm to find a local minimum for the MAP test at η. Start with any arbitrary threshold test α for Relay A, find the optimal $\bar{\alpha}'$ for Relay B, then conditional on this $\bar{\alpha}'$, find the optimal α, and so forth. That is, iterate back and forth between the 2 conditional solution pairs, $(\arg\min_{\alpha} P_e^{(\eta)}(\alpha, \bar{\alpha}'), \bar{\alpha}')$ and $(\alpha, \arg\min_{\bar{\alpha}'} P_e^{(\eta)}(\alpha, \bar{\alpha}'))$, until convergence to a $(\alpha, \bar{\alpha}')$. Convergence must occur since $P_e^{(\eta)}$ decreases at each step. Furthermore, if the error probability $P_e^{(\eta)}(\alpha, \bar{\alpha}')$ as a function of $(\alpha, \bar{\alpha}')$ has only one stationary point, e.g., the AGN channel, this iterative algorithm will converge to the global optimum solution; otherwise convergence may be to a local minimum.

Let entry $D_{ij} = \Lambda^{(\alpha,\bar{\alpha}')}(i,j) = \Lambda^{(\alpha)}(i)\Lambda^{(\bar{\alpha}')}(j)$ of the *Likelihood Ratio Matrix* $\Lambda^{(\bar{\alpha},\bar{\alpha}')}$ be the likelihood ratio of receiving symbol i from Relay A and symbol j from Relay B.

$$\Lambda^{(\alpha,\bar{\alpha}')} = \begin{bmatrix} \left(\frac{1-\beta}{\alpha}\right)\left(\frac{1-\beta_1'}{\alpha_1'}\right) & \left(\frac{1-\beta}{\alpha}\right)\left(\frac{\beta_1'-\beta_2'}{\alpha_2'-\alpha_1'}\right) & \left(\frac{1-\beta}{\alpha}\right)\left(\frac{\beta_2'}{1-\alpha_2'}\right) \\ \left(\frac{\beta}{1-\alpha}\right)\left(\frac{1-\beta_1'}{\alpha_1'}\right) & \left(\frac{\beta}{1-\alpha}\right)\left(\frac{\beta_1'-\beta_2'}{\alpha_2'-\alpha_1'}\right) & \left(\frac{\beta}{1-\alpha}\right)\left(\frac{\beta_2'}{1-\alpha_2'}\right) \end{bmatrix} \tag{4.2}$$

Note that each row and each column of 4.2 is monotonically decreasing. One method to find the entire final fused curve $F(\phi)$ is to implement the optimization procedure just discussed for every η. In other words, $F(\phi)$ is constructed point by point for every η.

For some η and the $(\alpha, \bar{\alpha}')$ which satisfies both conditional solution pairs, the fused curve $F_{\alpha,(\alpha_1',\alpha_2')}(\phi)$ has at most 6 distinct slopes. From (4.2) and (4.1), for all relay encoding vectors $(\alpha, \bar{\alpha}') = (\arg\min_{\alpha} P_e^{(\eta)}(\alpha, \bar{\alpha}'), \arg\min_{\bar{\alpha}'} P_e^{(\eta)}(\alpha, \bar{\alpha}'))$, the likelihood ratios $\Lambda(00)$, $\Lambda(01)$, $\Lambda(10)$ are all greater than η and $\Lambda(11)$, $\Lambda(02)$, $\Lambda(12)$ are all less than η. Thus, the η-tangent line is tangent at the middle vertex point of $F_{\alpha,(\alpha_1',\alpha_2')}(\phi)$. This middle vertex point is $(\phi, F_{\alpha,(\alpha_1',\alpha_2')}(\phi)) = (\alpha\alpha' + \alpha(\alpha_2' - \alpha_1') + (1-\alpha)\alpha_1', (1-\beta)\beta_2' + \beta(\beta_1' - \beta_2') + \beta\beta_2')$. Since $\phi = \Pr(e|1)$, the decoder matrix in 4.1 says that ϕ is the sum of the denominator of entries D_{11}, D_{12}, and D_{21} of the LR matrix in 4.2. Similarily, $F_{\alpha,(\alpha_1',\alpha_2')}(\phi) = \Pr(e|0)$ and is the sum of the numerator of entries D_{13}, D_{21}, and D_{23} of the LR matrix in 4.2. Ordering the LR slopes for the fused curve $F_{\alpha,(\alpha_1',\alpha_2')}(\phi)$, parallelogram structures exist created by $\{\Lambda(01), \Lambda(10)\}$ and $\{\Lambda(11), \Lambda(02)\}$. This is similar to the parallelogram structure of the 2×2 system discussed in Chapter 3.

The following theorem summarizes the 2×3 system discussed in this section.

Theorem 4.1.1 *Consider the 2×3 system with strictly convex relay error curves. For a fixed η and α, the optimal $\bar{\alpha}' = (\alpha_1', \alpha_2')$ is such that $\alpha_1' = \tilde{\alpha}\left(\frac{1-\alpha}{\beta}\eta\right)$ and $\alpha_2' = \tilde{\alpha}\left(\frac{\alpha}{1-\beta}\eta\right)$. For a given $\bar{\alpha}' = (\alpha_1', \alpha_2')$, the optimal $\alpha = \tilde{\alpha}\left(\frac{\alpha_2'-\alpha_1'}{\beta_1'-\beta_2'}\eta\right)$. The optimal decoding rule at the fusion point is given by the decoder matrix of 4.1.*

4.2 Generalization to Larger Output Alphabets

The ideas of the 2×3 system are now generalized in several directions in the following theorem. Relay A can transmit a symbol from an arbitrary sized alphabet. For fixed a priori probabilities, each symbol from Relay A produces a different ratio of conditional a priori probabilities for Relay B. Each of these ratios determines an encoding point of Relay B. Thus, the cardinality of the set of encoding points of Relay B is equal to the cardinality of Relay A's alphabet size. This implies that the alphabet size (number of quantization regions of the LR) of Relay B is just one more than the alphabet size of Relay A.

Any M level quantization of a relay output corresponds to a M-ary DMC which has an error curve that lies on or above the relay error curve. The points on that error curve are ordered and form a convex curve. In other words, any M-ary encoding at Relay A can be associated with $M-1$ likelihood ratios ordered so that $\frac{\alpha_1}{1-\beta_1} < \frac{\alpha_2-\alpha_1}{\beta_1-\beta_2} < \cdots < \frac{1-\alpha_{M-1}}{\beta_{M-1}}$.

Theorem 4.2.1 *Suppose the encoding at Relay A is constrained to M levels. Then Relay B needs at most $M+1$ output quantization levels for optimality. Furthermore, conditional on Relay A quantization vector $\bar{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{M-1})$, the optimal encoding vector of Relay B is* $\bar{\alpha}' = (\alpha_1', \alpha_2', \ldots, \alpha_M') = \left(\tilde{\alpha}'\left(\frac{1-\alpha_{M-1}}{\beta_{M-1}}\eta\right), \tilde{\alpha}'\left(\frac{\alpha_{M-1}-\alpha_{M-2}}{\beta_{M-1}-\beta_{M-2}}\eta\right), \ldots, \tilde{\alpha}'\left(\frac{\alpha_{j+1}-\alpha_j}{\beta_{j+1}-\beta_j}\eta\right), \ldots, \tilde{\alpha}'\left(\frac{\alpha_1}{1-\beta_1}\eta\right)\right)$.

Proof: Let η and Relay A encoding vector $\bar{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{M-1})$ be arbitrary. If quantization symbol $j \in \{1, 2, \ldots, M\}$ is received, the ratio of conditional a priori probabilities for Relay B is $\eta_j = \frac{\Pr(X=1|Z_1=j)}{\Pr(X=0|Z_1=j)}\eta$. It follows that

$$\eta_j = \begin{cases} \frac{\alpha_1}{1-\beta_1}\eta & \text{for} \quad j=1 \\ \frac{\alpha_{j+1}-\alpha_j}{\beta_j-\beta_{j+1}}\eta & \text{for} \quad 2 \le j \le M-1 \\ \frac{1-\alpha_M}{\beta_M}\eta & \text{for} \quad j=M \end{cases} \tag{4.3}$$

For each ratio of conditional a priori probability η_j, Relay B will make an optimal binary decision. These optimal tests, (α_j', β_j') for $j \in \{1, 2 \ldots M\}$ are simply threshold tests and are the η_j-tangent solutions to the Relay B error curve $\hat{\beta}'(\alpha)$. Since the encoding points of Relay A are ordered, the convexity of the relay error curve assures that the encoding points of Relay

B are also ordered (in reverse). Thus, the encoding points of Relay B are

$$\alpha'_j = \begin{cases} \tilde{\alpha}'\left(\frac{\alpha_1}{1-\beta_1}\eta\right) & \text{for} \quad j = 1 \\ \tilde{\alpha}'\left(\frac{\alpha_{j+1}-\alpha_j}{\beta_j-\beta_{j+1}}\eta\right) & \text{for} \quad 2 \le j \le M-1 \\ \tilde{\alpha}'\left(\frac{1-\alpha_M}{\beta_M}\eta\right) & \text{for} \quad j = M \end{cases} \tag{4.4}$$

Since there are M new ratios of conditional a priori probabilities, there are M threshold tests, which translate to $M + 1$ quantizations of the LR for Relay B, or equivalently, M encoding points on $\hat{\beta}'(\alpha)$. □

The optimal tests (α'_j, β'_j) are the η_j-tangent solutions to the conditional ratio of a priori probabilities $\eta_j = \frac{\Pr(X=1|Z_1=j)}{\Pr(X=0|Z_1=j)}\eta$, $0 < j \le M$. Therefore, $\alpha'_j = \tilde{\alpha}(\eta_j)$ for $0 \le j \le M$. Still assuming that the error curve is strictly convex, the optimal tests (α'_j, β'_j) are all unique and satisfy

$$-\frac{d\hat{\beta}(\alpha'_j)}{d\alpha} = \eta_j \quad \text{for} \quad 0 \le j \le M \tag{4.5}$$

After finding the new quantizations of the LR for Relay B and fusing them with the M vertex points of Relay A, the error probability is determined by looking at the β-intercept of the η-tangent line and then scaling by p_0.

In general, if the encoding points of one relay are known, the optimal encoding points of the other relay can be easily determined. Relay A does not necessarily need to have all its encoding points on its relay error curve, but we will show it is sufficient for optimality in the next section. Theorem 4.2.1 generalizes easily to an arbitrary number of relays with an arbitrary number of quantizations of the LR.

Theorem 4.2.1 shows that one of the relays needs at most a maximum of one extra quantization level of its LR than the other relay. This implies that for an arbitrary number of alphabet symbols allowed per relay, the only cases that need to be considered are $M \times M$ and $M \times (M+1)$. Assume that $M + 1$ is less than or equal to the rate needed for a relay to transmit its full observation to the fusion point. Section 4.3 assumes that the relay error curve is strictly convex, i.e., η and $\tilde{\alpha}(\eta)$ are a 1-1 bijection. The arbitrary relay error curve merely requires a slight modification of the strictly convex case and is treated in section 4.4.

4.3 The Strictly Convex Relay Error Curve

4.3.1 The $(M \times (M+1))$ System

This section is under the assumption of a system with strictly convex relay error curves. Suppose Relay A is constrained to an alphabet of M symbols and Relay B to an alphabet of M' symbols. Without loss of generality, assume $M < M'$, where M' might be infinite. As shown in theorem 4.2.1, Relay B can be restricted to $(M+1)$ symbols without loss in performance. For every $\bar{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_{M-1})$, choose the optimal $\bar{\alpha}' = (\alpha_1, \alpha_2, \dots, \alpha_M)$ by equation 4.4 to get the conditional solution pair $(\bar{\alpha}, \arg\min_{\bar{\alpha}'} P_e^{(\eta)}(\bar{\alpha}, \bar{\alpha}'))$. We will shortly show how to find the optimal $\bar{\alpha}$ given $\bar{\alpha}'$; in other words, how to find the conditional solution pair $(\arg\min_{\bar{\alpha}} P_e^{(\eta)}(\bar{\alpha}, \bar{\alpha}'), \bar{\alpha}')$. For a given η, every $(\bar{\alpha}, \bar{\alpha}')$ which simultaneously satisfies both conditional solution pairs is a stationary point of $P_e^{(\eta)}(\bar{\alpha}, \bar{\alpha}')$ and a local minimum. For every η, we are interested in finding the optimal encoding vectors $(\bar{\alpha}, \bar{\alpha}') = (\arg\min_{\bar{\alpha}} P_e^{(\eta)}(\bar{\alpha}, \bar{\alpha}'), \arg\min_{\bar{\alpha}'} P_e^{(\eta)}(\bar{\alpha}, \bar{\alpha}'))$ which minimize error probability, i.e., $\arg\min_{\bar{\alpha}, \bar{\alpha}'} P_e^{(\eta)}(\bar{\alpha}, \bar{\alpha}')$ which is the global minimum.

We will shortly see that the optimal Relay A encoder mapping is a quantization of its LR, just like Relay B. The regions are chosen by $(M-1)$ encoding points on the error curve. Since α uniquely determines β on the error curve, we will use the convention that the encoding symbols $(0, 1, \dots, M-1)$ will specify the quantization regions, $[0, \alpha_1] \to 0$, $(\alpha_1, \alpha_2] \to 1 \dots (\alpha_{M-2}, \alpha_{M-1}) \to (M-2)$, and $(\alpha_{M-1}, 1] \to (M-1)$. For the received symbol pair (i, j) at the fusion point, let i, where $0 \le i \le M-1$, be the symbol from Relay A and j, where $0 \le j \le M$, be the symbol from Relay B.

Since $\frac{\alpha_1}{1-\beta_1} < \frac{\alpha_2 - \alpha_1}{\beta_1 - \beta_2} < \frac{\alpha_3 - \alpha_2}{\beta_2 - \beta_3} < \cdots < \frac{1 - \alpha_{M-1}}{\beta_{M-1}}$, the optimal encoding points of $\bar{\alpha}'$ are determined in reverse order from equation 4.4. In other words, α_1 will determine threshold point α'_M. Likewise, α_1 and α_2 will determine threshold point α'_{M-1}; α_2 and α_3 will determine threshold point α'_{M-2}; and so forth. Finally, α_{M-1} will determine threshold point α'_1. The optimal decoding rule at the fusion point is the following: When the fusion point receives a "0" from Relay A (which means that Relay A has an observation whose LR is in the quantization region $[0, \alpha_1)$, the ratio of conditional a priori probabilities places the optimal threshold at α'_M for Relay B. Thus, the decoder maps the reception of $\bar{Z} = (0, M) \to 1$ and $\bar{Z} = (0, j) \to 0$ for $0 \le j \le M-1$. When the decoder receives symbol i or LR observation in the region $[\alpha_i, \alpha_{i+1}]$ from Relay A, the optimal Relay B threshold decision is α'_{M-i} and the decoder maps $\bar{Z} = (i, j) \to 1$ for $j \ge M - i$ and $\bar{Z} = (i, j) \to 0$ for $j \le M - i - 1$. Thus, given $\bar{\alpha}$, the optimal decoder is fixed.

The LR matrix $\Lambda^{(\bar{\alpha},\bar{\alpha}')}$ for an $M \times (M+1)$ system for relay encoder mappings of $\bar{\alpha}$ and $\bar{\alpha}'$ at the relays is an $M \times (M+1)$ dimension matrix. Let the rows be the output of Relay A and the columns output of Relay B. Then the entry (i,j) of the LR matrix is the LR of receiving the symbol $''i''$ from Relay A and symbol $''j''$ from Relay B. Since the output of the relays are independent conditional on the input, the LR of the output is the product of the LR of the individual relay outputs, i.e., $\Lambda^{(\bar{\alpha},\bar{\alpha}')}(i,j) = \Lambda^{(\bar{\alpha})}(i)\Lambda^{(\bar{\alpha}')}(j)$.

$$\Lambda^{(\bar{\alpha},\bar{\alpha}')} = \begin{bmatrix} \left(\frac{1-\beta}{\alpha}\right)\left(\frac{1-\beta_1'}{\alpha_1'}\right) & \left(\frac{1-\beta}{\alpha}\right)\left(\frac{\beta_1'-\beta_2'}{\alpha_2'-\alpha_1'}\right) & \cdots & \left(\frac{1-\beta}{\alpha}\right)\left(\frac{\beta_M'}{1-\alpha_M'}\right) \\ \left(\frac{\beta_1-\beta_2}{\alpha_2-\alpha_1}\right)\left(\frac{1-\beta_1'}{\alpha_1'}\right) & \left(\frac{\beta_1-\beta_2}{\alpha_2-\alpha_1}\right)\left(\frac{\beta_1'-\beta_2'}{\alpha_2'-\alpha_1'}\right) & \cdots & \left(\frac{\beta_1-\beta_2}{\alpha_2-\alpha_1}\right)\left(\frac{\beta_M'}{1-\alpha_M'}\right) \\ \vdots & \vdots & \vdots & \vdots \\ \left(\frac{\beta_{M-1}}{1-\alpha_{M-1}}\right)\left(\frac{1-\beta_1'}{\alpha_1'}\right) & \cdots & \cdots & \left(\frac{\beta_{M-1}}{1-\alpha_{M-1}}\right)\left(\frac{\beta_M'}{1-\alpha_M'}\right) \end{bmatrix} \quad (4.6)$$

The elements for each row and column of the LR matrix are monotonically decreasing since $\forall k > i$, $\Lambda^{(\bar{\alpha},\bar{\alpha}')}(i,j) > \Lambda^{(\bar{\alpha},\bar{\alpha}')}(k,j)$ and $\Lambda^{(\bar{\alpha},\bar{\alpha}')}(j,i) > \Lambda^{(\bar{\alpha},\bar{\alpha}')}(j,k)$. Therefore, by our labeling convention for the relay outputs ($\alpha_i < \alpha_j, \forall i < j$), the LR for symbol i is greater than the LR for symbol j for $i < j$. It follows that the optimal decoder must be monotonic in the following sense. At the decoder, for each row k , there exists a column j_k such that $\hat{x}(k,j) = 0$, $\forall j \leq j_k$ and $\hat{x}(k,j) = 1$, $\forall j > j_k$. In other words, (k, j_k) is the last entry on the k-th row which is decoded to $''0''$ and the next entry (k, j_k+1) will be decoded to $''1''$. A monotonic decoder means that $j_{k+1} \leq j_k$ for all k. A strict monotonic decoder means that $j_{k+1} < j_k$ for all k. An example is given in equation 4.7. The top left entries of the LR ratio matrix are mapped to $''0''$ and the bottom right mapped to $''1''$. We can express this mapping by a decoder matrix $D(\bar{\alpha},\bar{\alpha}')$. From what was discussed previously, the optimal decoder at the fusion point for every $M \times (M+1)$ system solution pair $(\alpha, \arg\min_{\bar{\alpha}'} P_e^{(\eta)}(\alpha,\bar{\alpha}'))$ must be

$$\hat{x}(i,j) = \begin{cases} 0 & \text{for} \quad i+j \leq M-1 \\ 1 & \text{otherwise} \end{cases} \quad (4.7)$$

This decoding rule, upon reception of symbol pair $\bar{Z} = (z_1, z_2)$, can be expressed by the following

$M \times (M+1)$ decoder matrix

$$D(\bar{\alpha}, \bar{\alpha}') = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & \iddots & \iddots & 1 & 1 \\ \vdots & \iddots & \iddots & \iddots & \iddots & \vdots \\ 0 & 0 & 1 & \iddots & \iddots & \vdots \\ 0 & 1 & 1 & \cdots & \cdots & 1 \end{bmatrix} \tag{4.8}$$

This is the only possible form for the decoder matrix of an $M \times (M+1)$ optimal solution $(\bar{\alpha}, \bar{\alpha}')$. For each symbol from Relay A, the optimal solution for Relay B says that each row must be different and monotonic. For the optimal solution to be satisfied, this means there are no degrees of freedom left in the decoder matrix D. Thus, there are no identical rows and no identical columns in D. This is due to the strict convexity of both relay error curves and distinct encoding points of Relay A, which implies distinct encoding points for Relay B. Thus, for each distinct symbol of Relay A, there is a corresponding distinct threshold test at Relay B.

From a simple extension of the 2×3 case to the$M \times (M+1)$ case , we argue that for a given $\bar{\alpha}'$ and the specified Decoder Matrix, the optimal $\bar{\alpha}$ is determined by the ratio of conditional probabilities from the symbols $1, ..., M-1$ of the Relay B. The decoder matrix in 4.8 says that when Relay B transmits the symbol "0", the fusion point always decodes "0" and when Relay B transmits symbol "M", the fusion point always decodes "1". Thus, the two end symbols of $\bar{\alpha}'$, "0" and "M", are never considered in determining the optimal $\bar{\alpha}$ conditional on $\bar{\alpha}'$. By theorem 4.2.1, for optimality, each element in $\bar{\alpha}$ is determined by its associated ratio of conditional a priori probabilities determined by each symbol from Relay B for symbols $1, ..., M-1$. The encoding points of Relay A are determined by the symbols of Relay B in reverse order, e.g., symbol "1" from Relay B will determine α_{M-1} of Relay A, symbol "2" from Relay B will determine α_{M-2}, $\ldots$, symbol $M-2$ from Relay B will determine α_2, and symbol $M-1$ will determine α_1. Call the associated pair of optimal $\bar{\alpha}$ given $\bar{\alpha}'$ in $(\arg\min_{\bar{\alpha}} P_e^{(\eta)}(\bar{\alpha}, \bar{\alpha}'), \bar{\alpha}')$ a conditional solution pair. The joint optimality conditions are $\alpha_n = \tilde{\alpha}\left(\frac{\alpha'_{M-n+1}-\alpha'_{M-n}}{\beta'_{M-n}-\beta'_{M-n+1}}\eta\right)$, $1 \le n \le M-1$, and $\alpha'_n = \tilde{\alpha}\left(\frac{\alpha_{M-n+1}-\alpha_{M-n}}{\beta_{M-n}-\beta_{M-n+1}}\eta\right)$, $1 \le n \le M$, with $\alpha_0 = 0$. These conditions are also expressed by equations 4.11 and 4.12.

We will now demonstrate the same joint optimality condition results via a different method which provide some extra insights. Each entry of the LR matrix is specifically given as the ratio of a numerator and denominator pair which are used separately to derive the error probability from the Decoder Matrix. For example, if an entry D_{ij} is decoded to a "0" from the decoder matrix, the error probability for that entry is the denominator term of that entry $\Lambda(i,j)$ in the LR Matrix. Similarily, if an entry is decoded to a "1", then the numerator of the same entry given in the LR Matrix is added to the overall error probability. For any η and $\bar{\alpha}$, equations 4.4 gives an optimal choice for encoding vector $\bar{\alpha}'$ and 4.8 the optimal decoding matrix D to minimize error probability. Thus, in looking for a joint $\bar{\alpha}$ and $\bar{\alpha}'$ to minimize error probability, we can restrict attention to the decoding matrix in equation 4.8. For any given solution pair $(\alpha, \arg\min_{\bar{\alpha}'} P_e^{(\eta)}(\alpha, \bar{\alpha}'))$, minimum error probability is then given by $P_e = p_1\phi + p_0 F_{\bar{\alpha},\bar{\alpha}'}(\phi)$ where $\phi = P(e|1)$ and $F_{\bar{\alpha},\bar{\alpha}'}(\phi) = P(e|0)$. If the error probability for each entry in 4.8 were summed up, we get the following expressions,

$$\phi = \alpha_1\alpha'_M + (\alpha_2 - \alpha_1)\alpha'_{M-1} + \cdots + (\alpha_{M-1} - \alpha_{M-2})\alpha'_2 + (1 - \alpha_{M-1})\alpha'_1 \tag{4.9}$$

$$F_{\bar{\alpha},\bar{\alpha}'}(\phi) = (1 - \beta_1)\beta'_M + (\beta_1 - \beta_2)\beta'_{M-1} + \cdots + (\beta_{M-2} - \beta_{M-1})\beta'_2 + \beta_{M-1}\beta'_1 \tag{4.10}$$

These points actually specify the middle vertex point of the fused error curve $F_{\bar{\alpha},\bar{\alpha}'}(\phi)$ for any given $\bar{\alpha}$, $\bar{\alpha}'$ which has at most $M(M+1)$ slopes.

For a given η, the necessary optimal conditions for (α, α') are obtained by taking the partial derivative of $P_e = p_1\phi + p_0 F_{\bar{\alpha},\bar{\alpha}'}(\phi)$ (ϕ and $F_{\bar{\alpha},\bar{\alpha}'}(\phi)$ are given in equations 4.9 and 4.10) with respect to each relay decision point α_i and α'_j and setting them each equal to zero. That is, $\frac{\partial P_e}{\partial \alpha_n} = 0$ and $\frac{\partial P_e}{\partial \alpha'_k} = 0$, for $1 \le n \le M-1$ and $1 \le k \le M$. For Relay A, the necessary conditions for optimality required of each element of $\bar{\alpha} = (\alpha_1, \ldots \alpha_{M-1})$ are

$$-\frac{\partial\hat{\beta}(\alpha)}{\partial\alpha_n} = \frac{\alpha'_{M-n+1} - \alpha'_{M-n}}{\beta'_{M-n} - \beta'_{M-n+1}}\eta \quad \text{for} \quad 1 \le n \le M-1 \tag{4.11}$$

and for Relay B, the necessary optimal conditions required of each element of $\bar{\alpha}' = (\alpha'_1, \ldots \alpha'_M)$ are

$$-\frac{\partial\hat{\beta}'(\alpha')}{\partial\alpha'_n} = \begin{cases} \frac{1-\alpha_{M-1}}{\beta_{M-1}}\eta & \text{for} \quad n = 1 \\ \frac{\alpha_{M-n+1}-\alpha_{M-n}}{\beta_{M-n}-\beta_{M-n+1}}\eta & \text{for} \quad 1 < n < M \\ \frac{\alpha_1}{1-\beta_1}\eta & \text{for} \quad n = M \end{cases} \tag{4.12}$$

These joint optimality conditions are as expected from our earlier derivation.

We have now verified that the optimal encoding points for both relays are interval quantizations of their respective likelihood ratios for an $M \times (M+1)$ system. For any given η, an iterative procedure can be implemented to find the local extrema points $(\bar{\alpha}, \bar{\alpha}')$ which satisfy the necessary optimal conditions in equations 4.11 and 4.12. First, find the optimal $\bar{\alpha}'$ conditional on $\bar{\alpha}$, then for the next iteration, find the optimal $\bar{\alpha}$ conditional on $\bar{\alpha}'$, and so forth. At each iteration, the error probability can only decrease. Thus, the algorithm must converge. It will converge to a stationary point which is a local minimum which is not necessarily the global minimum. However, the stationary points do satisfy the necessary condition on an optimal solution.

The following theorem summarizes the $M \times (M+1)$ system discussed in this section.

Theorem 4.3.1 *For strictly convex relay error curves, if Relay A is constrained to M output symbols, then the optimal Relay B output is $M+1$ symbols. The optimal relay encoding points have joint optimality conditions determined by the ratio of conditional a priori probabilities given by equations 4.11 and 4.12. This optimal $M \times (M+1)$ relay system has a unique decoding rule at the fusion point given by the Decoder Matrix in equation 4.8.*

4.3.2 The $(M \times M)$ System

We now examine the $M \times M$ system for strictly convex relay error curves and start by showing the relationship between the $M \times (M+1)$ and $M \times M$ systems. The following lemma shows that as the number of quantizations allowed per relay increases, the error probability strictly decreases for a strictly convex relay error curve.

Lemma 4.3.1 *For a strictly convex relay error curve, the optimal $M \times (M+1)$ system has strictly lower error probability than the optimal $M \times M$ system which in turn has strictly lower error probability than the optimal $(M-1) \times M$ system. For an arbitrary relay error curve, the optimal $M \times (M+1)$ system has error probability lower or equal to that of the $M \times M$ system which has error probability lower or equal to that of the $(M-1) \times M$ system. In other words, extra encoding rate at the relays always helps when η and $\tilde{\alpha}(\eta)$ have a 1-1 bijection. It may not help error probability if η and $\tilde{\alpha}(\eta)$ are not a bijection.*

Proof: For an $M \times (M+1)$ system, let the encoding vector for Relay A be $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{M-1})$ where $\alpha_{M-1} = \tilde{\alpha}\left(\frac{\alpha_1'}{1-\beta_1'}\eta\right)$. Recall that $\left(\frac{\alpha_1'}{1-\beta_1'}\eta\right)$ is the ratio of conditional a priori probabilities conditioned on symbol 1 being received by the fusion point from Relay B. Now, an additional symbol for Relay A allows the fusion point to make the optimal MAP decision conditional on receiving symbol 1 from Relay B. Since the relay curve is strictly convex, the optimal decision for this $(M+1)\times(M+1)$ system is strictly better than the decision made on the $M\times(M+1)$ case, which always decodes symbol 1 from Relay B into "0." The fusion point can use the same decision rule as in the $M\times(M+1)$ case for all other symbols from Relay B, making the error probability strictly smaller in the $(M+1)\times(M+1)$ case than in the $M\times(M+1)$ case.

We have shown in theorem 4.3.1 that for the strictly convex relay error curve, the optimal solution must have $M\times(M+1)$ symbols if one relay is constrained to M symbols. This implies that the error probability of an $M\times M$ system must strictly increase from a system of $M\times(M+1)$. $\square$

Let $\bar{\alpha}=\alpha_1<\alpha_2<\cdots<\alpha_{M-1}$ be the encoding vector for Relay A and $\bar{\alpha'}=\alpha'_1<\alpha'_2<\cdots<\alpha'_{M-1}$ be the encoding vector for Relay B and let $D(\bar{\alpha},\bar{\alpha'})$ be the corresponding decoder matrix. Each column of D must be monotonic in the sense that $D_{i,j}\leq D_{i+1,j}$ and each row in i must be monotonic in the sense that $D_{i,j}\leq D_{i,j+1}$. Note that if two rows, say row i and $i+1$ are identical, then the fusion point makes the same decision whether i or $i+1$ is received from Relay A. This means that the fusion point makes the same decision independent of whether symbol i or $i+1$ is received from Relay A, which means that α_i can be eliminated from $\bar{\alpha}$, refining the system from $M\times M$ to $(M-1)\times M$.

Lemma 4.3.1 says that $\bar{\alpha}$ and $\bar{\alpha'}$ cannot be optimal if two rows of D are identical. By the same argument, they cannot be optimal if two columns are identical. Combining the monotonicity of D with the distinct columns and rows, we see that there are only two possible choices of an $M\times M$ optimal decoder matrix, seen in equations 4.14 and 4.20. We call the form of these matrices the "Left Decoding Rule" and the "Right Decoding Rule," which are explicitly expressed below. No other decoder matrix other than Left and Right can be used at optimality. For given η, the decoding rule (left or right) specifies how to generate the conditional solution pairs $(\bar{\alpha},\arg\min_{\bar{\alpha}'}P_e^{(\eta)}(\bar{\alpha},\bar{\alpha}'))$ and $(\arg\min_{\bar{\alpha}}P_e^{(\eta)}(\bar{\alpha},\bar{\alpha}'),\bar{\alpha}')$. Each solution pair $(\bar{\alpha},\bar{\alpha}')$ is a stationary point if it satisfies both conditional solution pairs. Each solution pair also corresponds to a vertex point on the fused curve $F_{\bar{\alpha},\bar{\alpha}'}(\phi)$. These vertex points are called the "left decoding point" and "the right decoding point." Finally, the optimal solution is the minimum error probability of the optimal left decoding rule or the optimal right decoding rule.

The likelihood ratio matrix $\Lambda^{(\bar{\alpha},\bar{\alpha}')}$ of an $(M \times M)$ system given encoding vector $(\bar{\alpha},\bar{\alpha}')$ is

$$\Lambda^{(\bar{\alpha},\bar{\alpha}')} = \begin{bmatrix} \left(\frac{1-\beta}{\alpha}\right)\left(\frac{1-\beta_1'}{\alpha_1'}\right) & \left(\frac{1-\beta}{\alpha}\right)\left(\frac{\beta_1'-\beta_2'}{\alpha_2'-\alpha_1'}\right) & \cdots & \left(\frac{1-\beta}{\alpha}\right)\left(\frac{\beta_{M-1}'}{1-\alpha_{M-1}'}\right) \\ \left(\frac{\beta_1-\beta_2}{\alpha_2-\alpha_1}\right)\left(\frac{1-\beta_1'}{\alpha_1'}\right) & \left(\frac{\beta_1-\beta_2}{\alpha_2-\alpha_1}\right)\left(\frac{\beta_1'-\beta_2'}{\alpha_2'-\alpha_1'}\right) & \cdots & \left(\frac{\beta_1-\beta_2}{\alpha_2-\alpha_1}\right)\left(\frac{\beta_{M-1}'}{1-\alpha_{M-1}'}\right) \\ \vdots & \vdots & \vdots & \vdots \\ \left(\frac{\beta_{M-1}-\beta_{M-2}}{\alpha_{M-2}-\alpha_{M-1}}\right)\left(\frac{1-\beta_1'}{\alpha_1'}\right) & \cdots & \cdots & \left(\frac{\beta_{M-1}-\beta_{M-2}}{\alpha_{M-2}-\alpha_{M-1}}\right)\left(\frac{1-\beta_1'}{\alpha_1'}\right) \\ \left(\frac{\beta_{M-1}}{1-\alpha_{M-1}}\right)\left(\frac{1-\beta_1'}{\alpha_1'}\right) & \cdots & \cdots & \left(\frac{\beta_{M-1}}{1-\alpha_{M-1}}\right)\left(\frac{\beta_{M-1}'}{1-\alpha_{M-1}'}\right) \end{bmatrix} \tag{4.13}$$

Again, the monotonicity of the LR results in the monotonicity of D.

Case: $M \times M$ (Left Decoding Rule)

At the fusion point, the optimal decoder for the *Left Decoding Rule* is defined as

$$\hat{x}(i,j) = \begin{cases} 0 & \text{for} \quad i+j \leq M-2 \\ 1 & \text{otherwise} \end{cases} \tag{4.14}$$

and represented in matrix form called the *Left Decoder Matrix*

$$D^L(\bar{\alpha},\bar{\alpha}') = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & \iddots & \iddots & 1 & 1 \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ 0 & 1 & \iddots & \iddots & 1 \\ 1 & 1 & \cdots & \cdots & 1 \end{bmatrix} \tag{4.15}$$

When the last quantization region $[\alpha_{M-1}, 1]$ or symbol M is received from either relay, the fusion point always decodes a $''1''$, regardless of what the other relay says. The Left Decoding Rule says for each relay to use each of the first $M-1$ symbols, $\{0, \ldots, M-2\}$ from the other relay to generate its optimal encoding points from the ratio of conditional a priori probabilities. Thus, using a similar argument for the $M \times (M+1)$ system, the joint necessary optimal conditions for encoding points of Relays A and B are given by equations 4.19 and 4.18, respectively. The joint optimality conditions can also be expressed as $\alpha_n = \tilde{\alpha}\left(\frac{\alpha_{M-n}'-\alpha_{M-n-1}'}{\beta_{M-n-1}'-\beta_{M-n}'}\eta\right)$ and $\alpha_n' =$

$\tilde{\alpha}\left(\frac{\alpha_{M-n}-\alpha_{M-n-1}}{\beta_{M-n-1}-\beta_{M-n}}\eta\right)$, with $\alpha_0 = 0$, for $1 \leq n \leq M-1$.

We now use the alternative method to arrive at the same results. Again, using the same arguments for the $M \times (M+1)$ system, from the LR matrix in 4.13 and Decoder Matrix in 4.15, the corresponding middle locus point $(\phi, F_{\bar{\alpha},\bar{\alpha}'}(\phi))$ on the error curve to optimize over is

$$\phi = \alpha_1\alpha'_{M-1} + (\alpha_2 - \alpha_1)\alpha'_{M-2} + \cdots + (\alpha_{M-1} - \alpha_{M-2})\alpha'_1 \tag{4.16}$$

$$F_{\bar{\alpha},\bar{\alpha}'}(\phi) = (1-\beta_1)\beta'_{M-1} + (\beta_1 - \beta_2)\beta'_{M-2} + \cdots + (\beta_{M-2} - \beta_{M-1})\beta'_1 + \beta_{M-1} \tag{4.17}$$

Taking partial derivatives of $P_e^{(\eta)}(\bar{\alpha}, \bar{\alpha}') = p_1\phi + p_0 F_{\bar{\alpha},\bar{\alpha}'}(\phi)$ with respect to each encoding point, α and α' and setting it to zero, we get the necessary optimal conditions for the Left Decoding Rule. The encoding points of Relay B must satisfy

$$-\frac{\partial\hat{\beta}'(\alpha')}{\partial\alpha'_n} = \begin{cases} \frac{\alpha_{M-n}-\alpha_{M-n-1}}{\beta_{M-n-1}-\beta_{M-n}}\eta & \text{for} \quad 1 \leq n < M-1 \\ \frac{\alpha_1}{1-\beta_1}\eta & \text{for} \quad n = M-1 \end{cases} \tag{4.18}$$

and the encoding points for Relay A must satisfy

$$-\frac{\partial\hat{\beta}(\alpha)}{\partial\alpha_n} = \begin{cases} \frac{\alpha'_{M-n}-\alpha'_{M-n-1}}{\beta'_{M-n-1}-\beta'_{M-n}}\eta & \text{for} \quad 1 \leq n < M-1 \\ \frac{\alpha'_1}{1-\beta'_1}\eta & \text{for} \quad n = M-1 \end{cases} \tag{4.19}$$

As expected, the form of these optimality conditions is exactly the same as implementing theorem 4.2.1 for both relays, where each relay uses each of the first $M-1$ symbols of the other relay to determine each of the ratios of conditional a priori probabilities.

Case: $M \times M$ (Right Decoding Rule)

The *Right Decoding Rule* at the fusion point is defined as

$$\hat{x}(i,j) = \begin{cases} 0 & \text{for} \quad i+j \leq M-1 \\ 1 & \text{otherwise} \end{cases} \tag{4.20}$$

and represented by a fusion decoder matrix upon reception of $\bar{Z} = (z_1, z_2)$ as the *Right Decoder Matrix*

$$D^R(\bar{\alpha}, \bar{\alpha}') = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \iddots & \iddots & 1 \\ 0 & 0 & \iddots & \iddots & \vdots \\ 0 & 1 & 1 & \cdots & 1 \end{bmatrix} \tag{4.21}$$

Note again that all rows and columns are distinct, which arises from the monotonicity of the optimal decoder. This decoder matrix says that when the first quantization region $[0, \alpha_1]$ or symbol "0" is received from either relay, the fusion point always decodes a "0", regardless of what the other relay says. Again, using the similar argument for the $M \times (M+1)$ system and essentially the same argument as the left decoding rule, the joint necessary optimal conditions for encoding points of Relays A and B are given by $\bar{\alpha}_n = \tilde{\alpha}\left(\frac{\alpha'_{M-n+1} - \alpha'_{M-n}}{\beta'_{M-n} - \beta'_{M-n+1}}\eta\right)$ and $\bar{\alpha}'_n = \tilde{\alpha}\left(\frac{\alpha'_{M-n+1} - \alpha'_{M-n}}{\beta'_{M-n} - \beta'_{M-n+1}}\eta\right)$, with $\alpha_0 = 0$, for $1 \leq n \leq M-1$. This is also expressed in equations 4.25 and 4.24.

The corresponding middle locus point on the error curve to optimize over is

$$\phi = \alpha_1 + (\alpha_2 - \alpha_1)\alpha'_{M-1} + \cdots + (1 - \alpha_{M-1})\alpha'_1 \tag{4.22}$$

$$F_{\bar{\alpha},\bar{\alpha}'}(\phi) = (\beta_1 - \beta_2)\beta'_{M-1} + (\beta_2 - \beta_3)\beta'_{M-2} + \cdots + (\beta_{M-2} - \beta_{M-1})\beta'_2 + \beta_{M-1}\beta'_1 \tag{4.23}$$

Again, taking partial derivatives of P_e with respect to each encoding point and setting it to zero, the necessary optimal conditions for the right decoding rule for Relay B are

$$-\frac{\partial \hat{\beta}'(\alpha')}{\partial \alpha'_n} = \begin{cases} \frac{1-\alpha_{M-1}}{\beta_{M-1}}\eta & \text{for} \quad n = 1 \\ \frac{\alpha_{M-n+1} - \alpha_{M-n}}{\beta_{M-n} - \beta_{M-n+1}}\eta & \text{for} \quad 1 \leq n \leq M-1 \end{cases} \tag{4.24}$$

and for Relay A

$$-\frac{\partial \hat{\beta}(\alpha)}{\partial \alpha_n} = \begin{cases} \frac{1-\alpha'_{M-1}}{\beta'_{M-1}}\eta & \text{for} \quad n = 1 \\ \frac{\alpha'_{M-n+1} - \alpha'_{M-n}}{\beta'_{M-n} - \beta'_{M-n+1}}\eta & \text{for} \quad 1 \leq n \leq M-1 \end{cases} \tag{4.25}$$

Similar to the left decoding rule, the right decoding rule says to use the last $M-1$ symbols $(1, \ldots, M-1)$ for both relays and implement theorem 4.2.1.

The difference between the $M \times (M+1)$ and $M \times M$ systems lies in how the end symbols are treated. There is one end symbol in $M \times M$ which will not have an optimal threshold point for the other relay when it is transmitted. The left decoding rule says it is the last symbol. The right rule says it is the first symbol. For the $M \times (M+1)$ system, the relay restricted to $M+1$ symbols has both end symbols, 1 and $M+1$, which do not have an associated threshold test from the other relay. The iterative algorithm described for the $M \times (M+1)$ case to find local optimal solutions also works for the $M \times M$ system.

For relays with arbitrary output alphabet size constraints, we have shown for each relay that the necessary conditions for optimality require the encoding point solutions to be the η-tangents to the relay error curve. Thus, we have provided a simple explanation and understanding of why the relay encoding points must lie on its error curve, verifying the known result that a quantization of its LR is sufficient for optimality.

The fundamental ideas and main results of Chapter 4 are summarized in the following two theorems. The final fused curve $F(\phi)$ can be constructed by finding the optimal solution vertex point for each η by theorem 4.3.2. The alternative and less insightful way is to take the convex hull of the locus points varied over all $(\bar{\alpha}, \bar{\alpha}')$ in theorem 4.3.3.

Theorem 4.3.2 *If both relays have strictly convex relay error curves and are constrained to M alphabet symbol outputs then the optimal $M \times M$ system decoder is either the Left Decoder Rule (4.14) or the Right Decoder Rule (4.20). The optimal relay encoding points have joint optimality conditions determined by the ratio of conditional a priori probabilities given by, respectively, equations 4.19 and 4.18 or equations 4.25 and 4.24.*

Theorem 4.3.3 *Given Relay A with error curve $\hat{\beta}(\alpha)$ and Relay B error curve $\hat{\beta}'(\alpha)$, both strictly convex,*

(A) The final fused curve $F(\phi)$ of every $M \times (M+1)$ system is the convex hull of the middle decoder point (or middle vertex point defined by equations 4.9 and 4.10) of every fused curve $F_{\bar{\alpha},\bar{\alpha}'}(\phi)$ taken over all $\bar{\alpha}$ and $\bar{\alpha}'$.

(B) The final fused curve $F(\phi)$ of every $M \times M$ system is the convex hull of the left decoder point (equations 4.16 and 4.17) and the right decoder point (equations 4.22 and 4.23) of $F_{\bar{\alpha},\bar{\alpha}'}(\phi)$ taken over all $\bar{\alpha}$ and $\bar{\alpha}'$.

Proof: For all $M \times (M+1)$ systems, every η-tangent solution to $F(\phi)$ is the middle decoder point of $F_{\bar{\alpha},\bar{\alpha}'}(\phi)$ for some $\bar{\alpha}$ and $\bar{\alpha}'$ a stationary point $(\bar{\alpha},\bar{\alpha}')$. A stationary point must simultaneously satisfy both conditional solution pairs $(\bar{\alpha}, \arg\min_{\bar{\alpha}'} P_e^{(\eta)}(\bar{\alpha},\bar{\alpha}'))$ and $(\arg\min_{\bar{\alpha}} P_e^{(\eta)}(\bar{\alpha},\bar{\alpha}'),\bar{\alpha}')$. It follows that the convex hull of the middle decoder point of $F_{\bar{\alpha},\bar{\alpha}'}(\phi)$ described by equations 4.9 and 4.10 over all $\bar{\alpha}$ and $\bar{\alpha}'$ will give $F(\phi)$. Note that the middle vertex point may not exist for some $\bar{\alpha}$ and $\bar{\alpha}'$, but that does not matter since we are taking the convex hull of the equations 4.9 and 4.10, which are well defined for every choice of $\bar{\alpha}$ and $\bar{\alpha}'$. The same argument applies to the $M \times M$ system. Here, every η-tangent solution to $F(\phi)$ corresponds to either the left decoder point or right decoder point. By the same argument, taking the convex hull of 4.16 and 4.17 and 4.22 and 4.23 over all $\bar{\alpha}$, $\bar{\alpha}'$ will result in $F(\phi)$. $\square$

4.3.3 Additive Gaussian Noise Channel

Since the AGN Channel produces a strictly convex relay error curve, it is used as an example in this section to demonstrate the issues of the optimal solution of constrained relay output alphabets just discussed. The solution to the additive Gaussian noise channel system with 2 relays, multiple symbol relay outputs, the solutions to ML decoding is assumed in the figures. From the previous discussion, the LR of the Gaussian error curve is monotonic in the observation space, so a quantization of the LR can be viewed as a quantization of the observation space. For the 2×3 system, figure 4-3 shows the optimal encoding points for Relay A and B and the decision boundary associated with those relay encoding points. The decoder maps all points to the left and bottom of the decision boundary to "-1" and all points to the right and top of the decision boundary to "1". The solution is symmetric which is not surprising. Symmetry of the relays allows another solution where the decision boundary of figure 4-3(B) is reflected across the $t = -t'$ line. The roles of Relay A and B are reversed. From lemma 4.3.1, η and $\tilde{\alpha}(\eta)$ are in 1-1 correspondence for the Gaussian error curve. Thus, as the number of symbols increase, the performance will be strictly better. For the 3×3 system shown in figure 4-4, the optimal encoding points for both relays are the same and asymmetric. This is not too surprising from our analysis of the 2×2 case. From our discussion of the $M \times M$ system, there should be 2 solutions - the left decoding rule solution and the right decoding rule solution. Figure 4-4 shows the solution to the left decoding rule. The right decoding rule has its optimal encoding points which are symmetric to the left encoding rule on the relay error curve (solutions flipped across the $\alpha = \beta$ line, i.e. (β_1, α_1) and (β_2, α_2)). Again, the decision boundary for the right decoding rule is the decision boundary of Figure 4-4 (B) flipped across the line $t = -t'$. Moving up to the 3×4 system, we are back to one unique and symmetric solution with different encoding points for each relay. The solution and associated decision boundary is depicted in figure 4-5. Again, the second solution reverses the roles of Relays A and B. In summary the additive Gaussian noise channel always has two solutions (the decision boundary has a mirror image across the line $t=t'$), but the reasons are different for the $M \times M$ and $M \times (M+1)$ system. The $M \times (M+1)$ system will have different encoding points for each relay which are symmetric on the Gaussian relay error curve. There is one decoding rule. The roles of Relay A and B can be interchanged for the second solution. The $M \times M$ system will have the same encoding points for both relays which are asymmetric on the Gaussian relay error curve. The mirror image of the encoding points across the line $\alpha = \beta$ on the Gaussian error curve is the other solution. There are two asymmetric decoding rules - left and right - associated with each solution.

It is not surprising that as the number of symbols allowed per relay increases, the decision border approximates more closely to the maximum likelihood diagonal line and converges to the solution when the relays send full information. In general, this type of convergence to the centralized MAP decision is true for all systems. The rate of convergence of the error probability goes to 0 as roughly M^2.

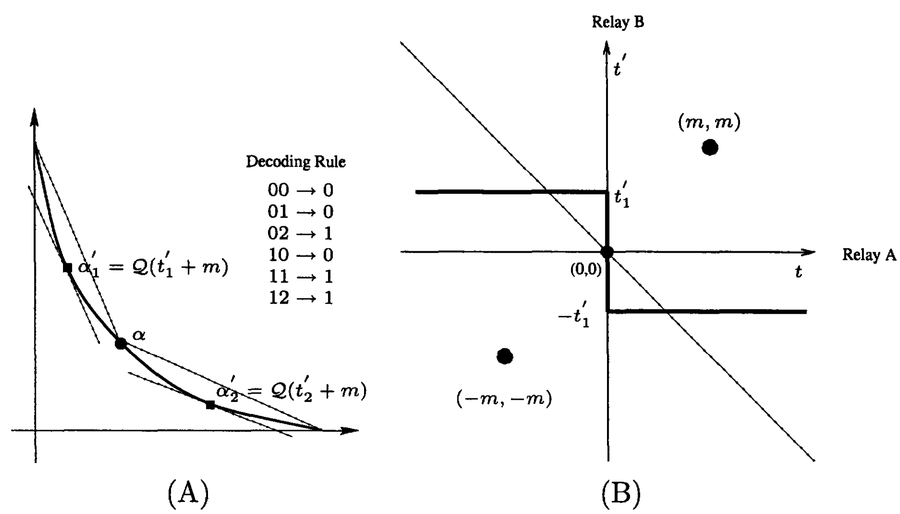


Figure 4-3: The 2×3 system for the Additive Gaussian Noise Channel and ML decoding ($\eta = 1$). (A) Optimal Relay B encoding points $\bar{\alpha}' = (\alpha_1', \alpha_2') = \left(\mathcal{Q}\left(t_1' + m\right), \mathcal{Q}\left(-t_2' + m\right)\right)$ conditional on Relay A encoding point at $\alpha = \beta = \mathcal{Q}(m)$. (B) Corresponding decision boundary for the optimal relay encoding strategies $(\alpha, \bar{\alpha}')$.

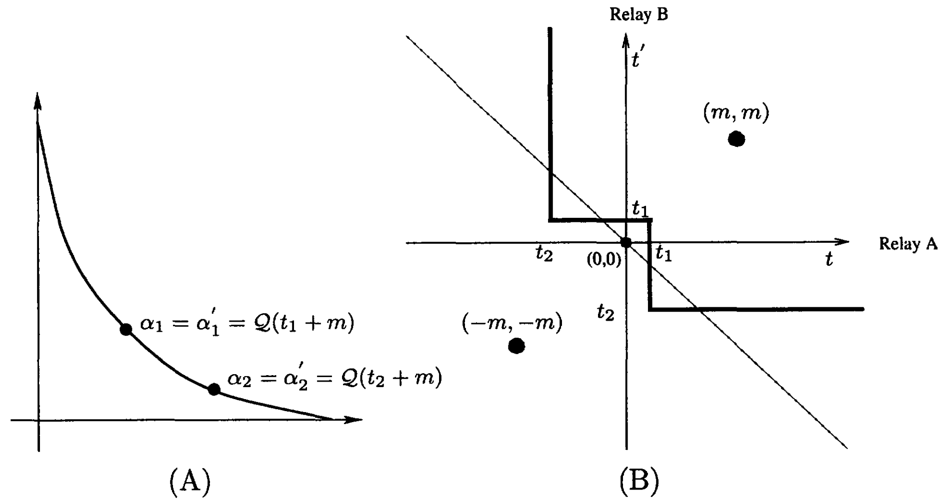


Figure 4-4: The 3×3 system for Additive Gaussian Noise Channel and ML decoding at the fusion point. (A) Optimal Relay B encoding vector $\bar{\alpha}' = (\alpha_1', a_2') = (\mathcal{Q}(t_1 + m), \mathcal{Q}(t_2 + m))$ conditional on Relay A encoding vector $\bar{\alpha} = (\alpha_1, a_2) = (\mathcal{Q}(t_1 + m), \mathcal{Q}(t_2 + m))$ for the left decoding rule. (B) Associated decision boundary for optimal encoding points $(\bar{\alpha}, \bar{\alpha}')$ for the left decoding rule. Note that the encoding points are the same for the relays $\bar{\alpha} = bar\alpha'$ Thresholds for both relays are the same.

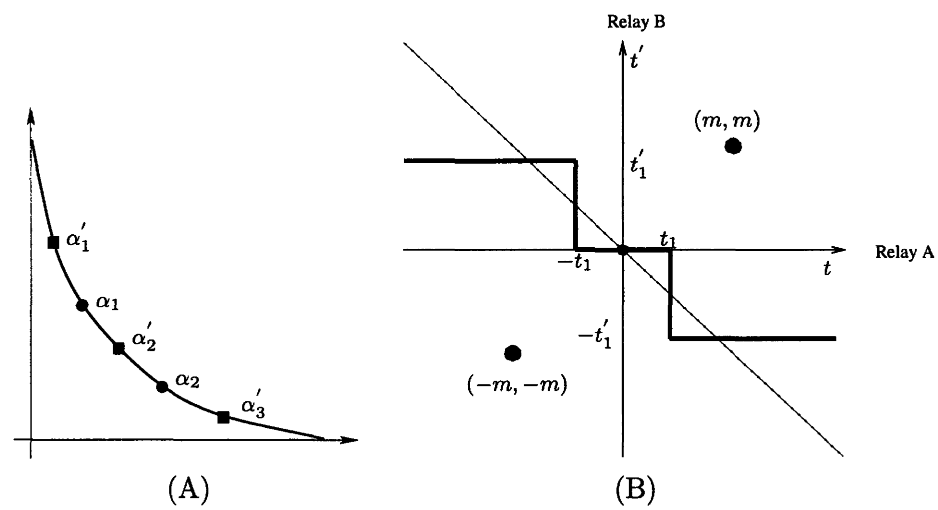


Figure 4-5: The 3×4 system for Additive Gaussian Noise Channel and ML decoding. (A) Optimal Relay B encoding vector $\bar{\alpha}' = (\alpha_1', a_2', a_3') = \left(\mathcal{Q}\left(t_1' + m\right), \mathcal{Q}(m), \mathcal{Q}\left(-t_1' + m\right)\right)$ conditional on Relay A encoding vector $\bar{\alpha} = (\alpha_1, a_2) = (\mathcal{Q}(t_1 + m), \mathcal{Q}(-t_1 + m))$ for $\eta = 1$. (B) Decision boundary for optimal encoding vectors$(\bar{\alpha}, \bar{\alpha}')$.

4.3.4 The 2×2 Case Revisited

The 2×2 case was analyzed in detail in Chapter 3. It can now be understood from the perspective of Chapter 4 since this 2×2 system is the smallest $M \times M$ case. Recall that for each η, the optimal solution is either a left vertex or right vertex point. The "OR" rule is the "Left Decoding Rule" and corresponds to a 2×2 Left Decoder Matrix $D(\alpha, \alpha) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ where the ratio of conditional a priori probabilities is $\eta_0 = \frac{\alpha}{1-\beta}\eta$. The "AND" rule is the "Right Decoding Rule" which corresponds to a 2×2 Right Decoder Matrix $D(\alpha, \alpha) = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ where the ratio of conditional a priori probabilities is $\eta_1 = \frac{1-\alpha}{\beta}\eta$. Both are consistent with the treatment in Chapter 3. The error probability of the optimal solutions to both decoder rules must be compared for global optimality. Equivalently, the type of the vertex point solution, either left or right, of the η-tangent line to the final fused curve must be known.

Recall that for every binary channel (α, β), there exists an η such that for this particular η, using one relay is equivalent to using 2 relays. For all other η, the error probability of two relays is strictly less than the error probability of one relay. Let the single vertex point (α, β) specify the binary channel. If both relays have tests at (α, β), then at the fusion point, the optimal decoder matrix for given $\eta = \frac{\beta(1-\beta)}{\alpha(1-\alpha)}$ is

$$D(\alpha, \alpha) = \begin{bmatrix} 0 & \{0,1\} \\ \{0,1\} & 1 \end{bmatrix} \tag{4.26}$$

where $\{0, 1\}$ means "don't care" – that choosing either symbol results in the same error probability. Suppose the "0" symbol is chosen for both "don't care" symbols in 4.26. Then the optimal decoding rule is the "OR" or "Left Decoder Rule" and corresponds to a left vertex point of the final fused curve $F(\phi)$. Likewise, suppose the "1" symbol is chosen for both "don't care" symbols in 4.26, then the "AND" or "Right Decoder Rule" is optimal and this corresponds to a right vertex point of the final fused curve. However, if we choose the "don't care" symbols in the following way, $D_{12} = 0$ and $D_{21} = 1$ (opposite of each other) then the columns are identical and the decoder matrix can be reduced to a (2×1) matrix. This implies that the fusion point is only using the threshold test of Relay A at (α, β) and ignoring Relay B. Likewise, if $D_{12} = 1$ and $D_{21} = 0$ then the rows are identical and the decoder matrix can be reduced to a (1×2) and only the Relay B threshold test at (α, β) is used. The optimal tests corresponding to the decoder matrix in 4.26 correspond to points on the straight line portion of slope $\eta = \frac{\beta(1-\beta)}{\alpha(1-\alpha)}$ between the left and right vertex point. The β-intercept of all three decoding rules is the same, so they all have the same error probability, as expected.

We now analyze this situation using the ideas of this chapter. Conditional on Relay A with encoding point at (α, β), the ratio of conditional a priori probabilities for Relay B are $\eta_0 = \frac{\alpha}{1-\beta}\eta = \frac{\beta}{1-\alpha}$ and $\eta_1 = \frac{1-\alpha}{\beta}\eta = \frac{1-\beta}{\alpha}$. Note that $\tilde{\alpha}(\eta_0) = \tilde{\alpha}\left(\frac{\beta}{1-\alpha}\right)$ has two solutions, (α, β) and $(1, 0)$. Also, $\tilde{\alpha}(\eta_1) = \tilde{\alpha}\left(\frac{1-\beta}{\alpha}\right)$ has two solutions, (α, β) and $(0, 1)$. The choice of encoding point solution will determine the form of the decoder matrix. Suppose the Left Decoding Rule is implemented. Then Relay B must choose the $\tilde{\alpha}(\eta_0)$ solution. If $\tilde{\alpha}(\eta_0) = (\alpha, \beta)$ were chosen, then the decoder matrix is the standard Left Decoder matrix $D(\alpha, \alpha) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. If Relay B chooses the trivial solution $\tilde{\alpha}(\eta_0) = (1, 0)$, then the Left Decoder Matrix becomes $D(\alpha, 1) = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$ which can be reduced to $D(\alpha, 1) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Only Relay A's threshold test is used and Relay B ignored, as expected.

We have seen that carefully choosing the "don't care" symbols in the decoder matrix by carefully choosing the optimal encoding point solution to the ratio of conditional a priori probabilities can result in reduced dimension of the Decoder Matrix, and thus, lower symbol rate transmission at the relays without loss in performance. The next section addresses this topic.

4.4 The Arbitrary Relay Error Curve

This section addresses the arbitrary relay error curve where $\tilde{\alpha}(\eta)$ and η are not in $1-1$ correspondence. There are exactly two issues which need to be addressed which do not arise in the strictly convex error curve. First, the η-tangent line may intersect the relay error curve at more than one point, and hence there may be 2 vertex point solutions to a particular η-tangent line. Second, one vertex point will be the η-tangent solution for a range of η. This means that the η-tangent solution to different ratios of conditional a priori probabilities can map into the same vertex point on the relay error curve.

For the arbitrary relay error curve, the form of Decoder Matrix is still monotonic but no longer strict. The decoding matrices which were optimal before are now just sufficient for optimality. When there are identical rows or columns, the decoder matrix can be reduced to a lower dimension decoder matrix by combining all identical rows and columns. From the property of decoder matrices, the only optimal forms that the lower dimension decoder matrix can reduce to are still the $M \times (M+1)$ and $M \times M$ decoder matrices for some integer M.

4.4.1 The Decoder Matrix and System Reduction

Identical rows in the Decoder Matrix indicates that those quantization regions of Relay A can be combined, which means that the decision points that separate those rows can be eliminated. Likewise, identical columns means that those regions of Relay B can be combined and the corresponding encoding points which separated these columns can be eliminated. Furthermore, if there are some encoding points which are the same, then the regions which these points separate can also be combined into one region. For example, if there exists some α_i in $\bar{\alpha}$ such that there exists at least one α_j for $i \neq j$ where $\alpha_i = \alpha_j$, then rows i through $j+1$ will be identical and can all be combined into one row. Therefore, any identical encoding points or any identical rows/columns in the decoder matrix means that the dimension of the decoder matrix can be reduced. For a strictly convex error curve, lemma 4.3.1 implies that a reduced dimension decoder matrix must be strictly suboptimal since an $(M-1) \times M$ system is strictly worse than an $M \times M$ system which in turn is strictly worse than an $M \times (M+1)$ system. In other words, if the decoder matrix is of dimension less than $M \times (M+1)$, then extra encoding points can always be added to make the decoder matrix of dimension $M \times (M+1)$, which will strictly decrease error probability. Thus, the 1-1 correspondence between η and $\bar{\alpha}(\eta)$ guarantees that all the elements of the optimal $\bar{\alpha}$ and $\bar{\alpha}'$ are unique and no two rows and no two columns of the decoder matrix $D(\bar{\alpha}, \bar{\alpha}')$ are identical. This implies that the form of the decoder matrix must be strictly monotonic, which says that the matrix in 4.8 is the only form that the decoder matrix can have for all optimal solution pairs $(\bar{\alpha}, \bar{\alpha}')$ for an $M \times (M+1)$ system. Likewise, the matrices in 4.15 and 4.21 are the only forms that the decoder matrix can have for all optimal solution pairs $(\bar{\alpha}, \bar{\alpha}')$ for an $M \times M$ system.

The reduction process of reducing rows and columns to a minimum in the Decoder Matrix can only be possible if there are (i) "don't care" symbol entries or (ii) identical rows or columns.

A situation of a (i) "don't care" symbol entry occurs when the encoding points of the conditional solution pair associated with the "don't care" entry has two solutions, a trivial and non-trivial encoding point. To reduce matrix dimension, choose the trivial encoding point solution. This will result in either two identical columns or two identical rows in the decoder matrix. A situation of (ii) identical rows or identical columns occurs when the encoding points of the conditional solution pair associated with those rows or columns are not unique. That is, there exist different ratios of conditional a priori probabilities which map to the same vertex point. In other words, $\tilde{\alpha}(\eta_i) = \tilde{\alpha}(\eta_j)$ for some $i \neq j$.

We now address the iterative algorithm described previously for finding the stationary encoding points $(\bar{\alpha}, \bar{\alpha}')$ for the strictly convex error curve. Observe that this algorithm does not work as well on discrete parts of the relay error curve. The initial start value is critical because convergence can be to a trivial point. Once the dimension of a Decoder Matrix is reduced, the algorithm does not allow the matrix to return to its original larger dimension, unless the algorithm is modified. We modify the iterative algorithm as follows. If at each step, the number of output symbols for a relay is less than the maximum allowable, then increase the number of encoding points at the next iteration step by using all the symbols of the other relay for the ratio of conditional a priori probabilities.

4.4.2 The Maximum Rate Reduction Process

We are now interested in the problem of determining the smallest encoding alphabet size needed per relay to do the best possible job for any MAP detection, with or without relay symbol constraints. We have seen from lemma 4.3.1 that if a relay error curve is strictly convex, minimum error probability is achieved only when both relays send full information, which requires an infinite number of bits. Hence, any extra encoding rate allowed at the relays will strictly decrease error probability. Suppose that both relay error curves have the same N vertex points. If both relays send their full information or an alphabet size of $(N+1)$ symbols per relay, which is the centralized problem, then the best MAP detection can be achieved. It might appear that sending anything less will increase error. It turns out that sometimes the relays can send less than an alphabet size of $(N+1)$ symbols and still achieve the minimal error probability. A simple example of this situation is provided below.

Example: Consider the binary input and four output relay channel in figure 4-6 for both relays. Suppose $p_0 = p_1$ and we are interested in minimizing error probability. The fusion point achieves its best performance if Relay A and B each send their full information which requires 4 symbols as shown in figure 4-6 (I). However, suppose Relay B transmits 2 symbols to the fusion point by mapping its outputs $\{0, 1\} \rightarrow 0$ and $\{2, 3\} \rightarrow 1$ which is a binary channel equivalent. Suppose Relay A transmits 3 symbols by mapping its output $\{0\} \rightarrow 0$ and $\{1, 2\} \rightarrow 1$ and

$\{3\} \rightarrow 2$ which is a ternary channel equivalent as shown in figure 4-6 (II). Surprisingly, the error probability is exactly the same.

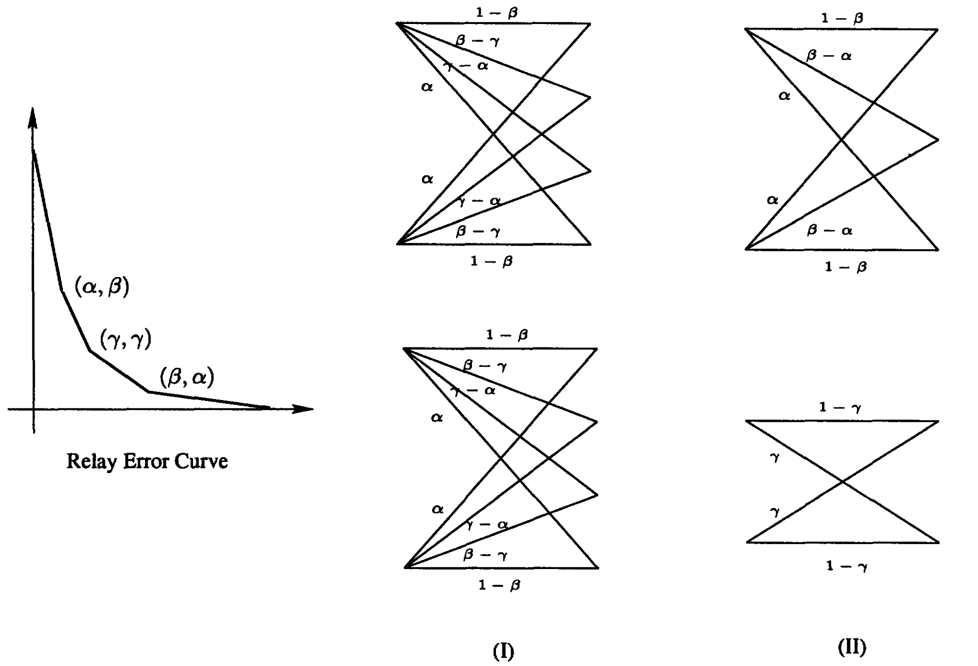


Figure 4-6: ML decoding ($\eta = 1$) for a symmetric relay error curve. If the vertex points $(\alpha, \beta), (\gamma, \gamma), (\beta, \alpha)$ form a convex curve, then the parallel channel in (I) has the same error probability as the parallel channel in (II). In other words, Relay A and B transmitting their full information of 4 symbols each to the fusion point is equivalent to Relay A transmitting 3 symbols and Relay B transmitting 2 symbols to the fusion point.

Definition 1 *: An $(M-1)$ point error curve is defined to be symmetric with respect to η if it satisfies all of the following 2 conditions for a given η: (i) $\frac{\alpha_1}{1-\beta_1}\eta = \frac{\beta_{M-1}}{1-\alpha_{M-1}}$ and (ii)$\frac{\alpha_{j+1}-\alpha_j}{\beta_j-\beta_{j+1}}\eta = \frac{\beta_{M-j-1}-\beta_{M-j}}{\alpha_{M-j}-\alpha_{M-j-1}}$ for all $1 \leq j \leq M-2$.*

The greatest symbol reduction is a symmetric relay error curve. The number of relay output symbols can be reduced by half. If both relays have the same error curve, then from symmetry of the relays (the labels of Relay A and Relay B can be interchanged), the decoder matrix generated from both relays using all the points on the relay error curve must have anti-diagonal symmetry.

If both relays have the same symmetric finite discrete relay error curves with respect to a particular η, then the minimum number of encoding points for each relay to preserve the error probability at η is half the number of vertex points on the relay error curve. The relay encoding points are $\tilde{\alpha} = (\alpha_1, \alpha_3, \ldots)$ and $\tilde{\alpha}' = (\alpha_2, \alpha_4, \ldots)$. In this particular example, each encoding point chosen is the η-tangent solution to two different ratios of conditional a priori probabilities. This allows for the maximum reduction of rows and columns in D.

The reduction procedure described in section 4.4.1 and 4.4.2 can be applied to any optimal $M \times M$ or $M \times (M+1)$ system. In summary, first find all the global optimal encoding points of the system. The set of elements in the stationary point solutions will be a subset of vertex points of the relay error curve. Take all the optimal encoding points for one relay and order them by α. If a trivial point solution exists, choose it. If an encoding point is part of two or more solutions, then choose it. Now create the new decoder matrix D from the chosen encoding points and reduce the dimension of D by combining identical rows and columns. Note that if there is no intersection of vertex point solutions to the set of ratios of conditional a priori probabilities or if there are no trivial point solutions, then all the encoding points are distinct and non-trivial, and no reduction in the dimension of the decoder matrix is possible.

For maximal reduction, the encoding points should be chosen to maximize the number of identical columns and rows of D. This is achieved by strategically mapping as many different ratios of conditional a priori probabilities into the same encoding points for each relay. In other words, maximal reduction is achieved when the set of encoding points is minimzed for each relay.

We then examined the case of non-binary relay outputs and two relays. The problem was solved through the perspective of one relay as fusion point. We first investigated the 2×3 case in detail and then generalized to arbitrary relay outputs. The structure for the necessary conditions of optimality and the decoding matrix were established. We found that all constrained systems reduce to an $M \times (M + 1)$ or $M \times M$ system. We established the joint optimality conditions and specified the optimal decoding structure for each system.

We have only looked at the simplest non-trivial forms of distributed information and detection networks, but have developed an essentially complete set of results about this sub-class of problems. The set of counter-examples presented have results we once hoped were true is as impressive as the set of results that are true. This is an area of research in which intuition is developed slowly and great patience and care are required.

5.1 Future Directions for Research

We have developed preliminary results about a number of extensions to the problems here. Many of the results extend to the case of more than two relays, but these results are not yet in a sufficiently mature state to express them clearly. Another interesting extension is when the relays employ coding on a sequence of successive binary symbols from the source. It turns out that significant reductions in error probability are possible by first using symbol by symbol encoding at the relays, followed by Slepian-Wolf encoding of the relay outputs.

Bibliography

[CT91] Thomas Cover and Joy Thomas, *Elements of Information Theory*, Athena Scientific, Belmont, MA, 1991.

[Gal68] Robert G. Gallager, *Information Theory and Reliable Communication*, John Wiley & Sons, New York, 1968.

[IT94] William W. Irving and John N. Tsitsiklis, *Some Properties of Optimal Thresholds in Decentralized Detection*, IEEE Transactions on Automatic Control **39** (1994), no. 4, 835–838.

[Ooh98] Y. Oohama, *The Rate-Distortion Function for the Quadratic Gaussian CEO Problem*, IEEE Transactions on Information Theory **44** (1998), 1057–1070.

[PA95] J.D. Papastavrou and M. Athans, *The Team ROC Curve in a Binary Hypothesis Testing Environment*, IEEE Transactions on Aerospace and Electronic Systems **31** (1995), 96–105.

[QZ02] R.D. Wesel Qian Zhang, P.K. Varshney, *Optimal Bi-Level Quantization of i.i.d. Sensor Observations for Binary Hypothesis Testing*, IEEE Transactions on Information Theory **48** (2002), 2105 - 2111.

[RB97] H.V. Poor R.S. Blum, S.A. Kassam, *Distributed Detection with Multiple Sensors I. Advanced Topics*, Proceedings of the IEEE **85** (1997), 64–79.

[TA84] John N. Tsitsiklis and Michael Athans, *Convergence and Asymptotic Agreement in Distributed Decision Problems*, IEEE Transactions on Automatic Control **29** (1984), 42–50.

[TA85] ———, *On the Complexity of decentralized decision making and detection problems*, IEEE Transactions on Automatic Control **30** (1985), 440–446.

[TB96] H. Viswanathan T. Berger, Zhen Zhang, *The CEO Problem*, IEEE Transactions on Information Theory **42** (1996), 887–902.

[TNS84] R.R. Teney and Jr. N.R. Sandell, *Detection with Distributed Sensors*, IEEE Trans. Aero. Electron. Syst. **AES-17** (1984), no. 4, 501–510.

[Tsi88] John N. Tsitsiklis, *Decentralized Detection by a Large Number of Sensors*, Mathematics of Control, Signals and Systems **1** (1988), no. 2, 167–182.

[Tsi93] ———, *Decentralized Detection*, in Advances in Signal Processing **2** (1993), 297–344.

[Var97] Pramod K. Varshney, *Distributed Detection and Data Fusion*, Springer-Verlag, 1997.

[VV97] R. Viswanathan and P.K. Varshney, *Distributed Detection with Multiple Sensors I. Fundamentals*, Proceedings of the IEEE **85** (1997), 54–63.

[WS01] T.W R.D. Wesel Wei Shi, T.W. Sun, *Quasi-Convexity and Optimal Bbinary Fusion for Distributed Detection with Identical Sensors in Generalized Gaussian Noise*, IEEE Transactions on Information Theory **47** (2001), 446–450.

[WW90] D.J. Warren and P.K. Willet, *Randomization in Distributed Detection*, Proc. 29th Conf. Decision Control, Honolulu **31** (1990), 1151–1152.

Index