CUTOFF RATE FOR FIXED-COMPOSITION CODING OVER ENERGY CONSTRAINED AWGN CHANNELS

A THESIS
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL
AND ELECTRONICS ENGINEERING
AND THE INSTITUTE OF ENGINEERING AND SCIENCES
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

By
Nihat Cem OĞUZ
February, 1990

CUTOFF RATE FOR FIXED-COMPOSITION CODING OVER ENERGY CONSTRAINED AWGN CHANNELS

A THESIS
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING
AND THE INSTITUTE OF ENGINEERING AND SCIENCES
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

By
Nihat Cem Oğuz
February 1990

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

Associate Prof. Dr. Erdal Arıkan (Principal Advisor)

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

Assistant Prof. Dr. Enis Çetin

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

Assistant Prof. Dr. Gürhan Şaplakoğlu

Approved for the Institute of Engineering and Sciences:

Prof. Dr. Mehmet Baray
Director of Institude of Engineering and Sciences

ABSTRACT

CUTOFF RATE FOR FIXED-COMPOSITION CODING OVER ENERGY CONSTRAINED AWGN CHANNELS

Nihat Cem Oğuz
M.S. in Electrical and Electronics Engineering
Supervisor: Assoc. Prof. Dr. Erdal Arıkan
February, 1990

Shannon showed that, under an energy constraint, the ensemble of shell constrained codes optimizes the cutoff rate for AWGN channels. Unfortunately, this ensemble is not very practical since its input alphabet is the entire real line. In this thesis, we consider the ensemble of fixed-composition codes which satisfy the shell constraint and have a finite input alphabet.

For a certain four-letter symmetric input alphabet, the cutoff rates for ensembles of fixed-composition codes of blocklengths up to 40 are computed for the AWGN channel at various signal-to-noise ratios. Also an asymptotic analysis of these cutoff rates is carried out as blocklenghth tends to infinity.

These results are compared with the cutoff rates optimized over the independent-letters code ensemble, which is the ensemble ordinarily used in practice. The results of this comparison show that, for relatively moderate signal-to-noise ratios, it is possible to achieve cutoff rates within 1-2% of the optimum value by using fixed-composition codes; whereas, with independent-letters codes, one can get at most within 9-10% of the optimum value. Thus, fixed-composition codes can provide significant improvements in cutoff rate in practice, especially for moderate to high signal-to-noise ratios.

Key words: fixed-composition codes, permutation codes, cutoff rate, energy constrained AWGN channels.

ÖZET

ENERJİ KISITLI AWGN KANALLARDA SABİT BİLEŞİMLİ KODLAMA İÇİN KESİLİM HIZI

Nihat Cem Oğuz
Elektrik ve Elektronik Mühendisliği Bölümü Yüksek Lisans
Tez Yöneticisi: Doç. Dr. Erdal Arıkan
Şubat, 1990

Shannon, enerjinin kısıtlı olduğu durumlarda, kabuk kısıtlı kodlar topluluğunun AWGN kanallar için kesilim hızını en iyileştirdiğini göstermiştir. Ne var ki, bu topluluk, kod alfabesi bütün gerçel sayılar kümesi olduğundan, pek uygulanabilir değildir. Bu tez çalışmasında, kabuk kısıtlamasını sağlayan ve sonlu bir kod alfabesi üzerinde tanımlı sabit bileşim kodlar topluluğu ele alınır.

Dört harfli simetrik bir kod alfabesi seçilerek, çeşitli sinyal-gürültü oranlarında, AWGN kanallar ve 40'a kadar çeşitli blok uzunlukları için, sabit bileşim kodlar topluluklarının kesilim hızları hesaplanır. Bu kesilim hızlarının, blok uzunlukluğu sonsuza giderken aldıkları asimtotik değerler de hesaplanır.

Bu sonuçlar, pratikte kullanılan bağımsız harfli kodlar topluluğu üzerinden en iyileştirilen kesilim hızlarıyla karşılaştırılır. Bu karşılaştırmanın sonuçları, bağımsız harfli kodlar ile en iyi kesilim hızının en fazla %90-91'i elde edilebilirken, göreceli olarak orta sinyal-gürültü oranları için, sabit bileşim kodları kullanarak en iyi değerin %98-99'unu elde etmenin olası olduğunu gösterir. Böylece, sabit bileşim kodlar, özellikle orta ve yüksek sinyal-gürültü oranlarında, kesilim hızında önemli gelişmeler sağlayabilir.

Anahtar sözcükler: sabit bileşim kodlar, permütasyon kodları, kesilim hızı, enerji kısıtlı AWGN kanallar.

ACKNOWLEDGEMENT

I would like to express my deep gratitude to Assoc. Prof. Dr. Erdal Arıkan for his invaluable guidance and suggestions during the development of this study.

Contents

1 INTRODUCTION 1

1.1 Fixed-Composition and Independent-Letters Codes 1

1.2 What is Coding Gain? 2

1.3 Background and Motivation 3

1.4 Summary of Results 6

2 COMPARISON OF R_0 FOR FIXED-COMPOSITION AND INDEPENDENT-LETTERS ENSEMBLES 7

2.1 Mathematical Preliminaries 7

2.2 Comparison of $R^{\star}_{0,ilc}$ and $R_{0,fcc}$ 9

2.3 Discussion of Results 19

3 CONCLUSION AND FURTHER RESEARCH TOPICS 22

A Optimization of $R_{0,ilc}$ under Energy Constraint 23

A.1 Minimization of $e^{-R_{0,ilc}}$ 23

A.2 $Q^{\star}$ and $R^{\star}_{0,ilc}$ as Functions of E 24

B Computation of $R_{0,fcc}$ 29

B.1 Algorithm 1: Enumeration in Lexicographical Order 29

B.2 Algorithm 2: Enumerating Joint-Composition Classes 30

B.3 Codes Implementing Algorithms 1 and 2 32

B.4 Numerical Results of $R_{0,fcc}$ Computations 36

C Asymptotic Analysis of $R_{0,fcc}$ **42**

C.1 Computation of $R_{0,fcc}(\infty)$ 42

C.2 Analysis of the Error Term 45

C.2.1 Error due to Approximating $|\mathcal{F}_{N,Q^*}|$ and $|V|$ 45

C.2.2 Error due to Approximating the Summation by the Dominant Term 46

References **48**

Chapter 1

INTRODUCTION

Shannon [Sha48] proved that under power limitations, there is, associated with any physical channel, an upperbound, called *channel capacity*, to the rates at which reliable communication over the channel can be achieved. At rates above channel capacity, the communication system suffers a high probability of error no matter how much effort is made to design the system cleverly. For years, it has been of interest to build systems that can communicate reliably at higher and higher rates to bridge the gap between the channel capacity and the rates achieved in practice and so will be the case for years. This thesis work is another effort in that direction.

In this thesis work, the performances of *fixed-composition* and *independent-letters codes* are compared for the important example of discrete-time, memoryless, *additive white gaussian noise (AWGN) channel*. For this channel, the input and output are related at any time instant (channel use) j by

$$r_j = s_j + n_j \tag{1.1}$$

where the input s_j is an arbitrary real number, the noise term n_j is a zero mean, gaussian random variable with variance σ^2, and r_j is the channel output. We show that, in case of energy constraints at the channel input, significant coding gains are practically achievable by using fixed-composition codes, especially in moderate to high signal-to-noise ratio cases, as anticipated in [Gal86].

1.1 Fixed-Composition and Independent-Letters Codes

A fixed-composition code of blocklength N is a code each codeword of which contains each code letter a_i the same number of times, n_i times, where the code letters come from a finite set, $\mathcal{A} = \{a_1, a_2, \ldots, a_K\}$, called the *code alphabet*. Obviously,

$$n_1 + n_2 + \cdots + n_K = N. \tag{1.2}$$

Normalizing the frequency of occurance of each code letter by the blocklength, we get a probability distribution $Q = \{q_l : l = 1, 2, \ldots, K\}$ on the code alphabet. That is, defining

$$q_l \triangleq \frac{n_l}{N} \quad ; \, l = 1, 2, \ldots, K, \tag{1.3}$$

we have

$$q_l \geq 0 \quad ; \, l = 1, 2, \ldots, K, \text{ and } \sum_{l=1}^{K} q_l = 1. \tag{1.4}$$

Hence having fixed the code alphabet $\mathcal{A}$, the parameters N and Q define a fixed-composition code, as does the set of letter frequencies alone. Throughout this text, pair (N, Q) denotes the parameters of such a fixed-composition code. Without loss of generality we assume that none of the letter probabilities is zero.

On the other hand, an independent-letters code is such a code that each codeword component is assigned the code letter a_l with probability q_l, independent of all other component assignments both in the same codeword and in other codewords. Therefore, for independent-letters codes, the blocklength and the probability distribution over the code alphabet are independent parameters. For fixed-composition codes, observe that, given the probability distribution, the blocklength can take certain values so as to make sure that $q_l N$ is an integer for all l.

In applications, one encounters various channel input constraints. Among these are, for example, runlength constraints in magnetic recording applications, charge constraints in DC free communication lines, spectral constraints in telephone lines, average or peak power constraints, energy constraints, etc. The theory indicates that, under input constraints, one may achieve significant coding gains by using fixed-composition codes rather than codes that are not restricted in this manner [Sha59]. Thus, we are motivated to compare fixed-composition codes with independent-letters codes in particular. Here, we have to explain what we mean by coding gain.

1.2 What is Coding Gain?

We measure the coding gain by the *improvement in cutoff rate* R_0 of *sequential decoding*. That is, fixed-composition and independent-letters ensembles are compared with respect to their cutoff rates. This makes sense when *trellis coding* along with *sequential decoding* is considered since sequential decoding can be used successfully for all rates below the cutoff rate. In other words, it is possible to build sequential decoders that can correctly recover the message with probabilities approaching one as much as desired by increasing the constraint span L of the trellis code provided that the communication rate is bounded by R_0. More important than that, increasing L does not result in an extra computational cost. The significance of R_0 lies mainly in this fact, i.e. in its being the computational cutoff rate of the sequential decoding. For a detailed discussion of why R_0 is taken as the quantity of primary interest, one may refer to [WoJ65, p.440] and [WoK66].

In fact, theoretically both trellis and block codes exhibit an error performance that improves exponentially with L; but, whether realizable decoders for large L exist or not is the basic question. In this regard, fixed-composition trellis codes are more promising since there exist sequential decoders that can successfully decode such codes for large L and communication rates below cutoff rate. Within the scope of this work, however, no effort is made on specific aspects of trellis coding and sequential decoding parts of the problem. Only the cutoff rates of the two ensembles are compared.

1.3 Background and Motivation

Let $\mathcal{C}_A$, $\mathcal{C}_B$ and $\mathcal{C}_C$ be three block codes over $\mathcal{R}$ each having M equiprobable codewords of blocklength N for the AWGN channel and satisfy the *shell*, *sphere* and *average power constraints* respectively. That is, each codeword $\mathsf{s} = (s_1, s_2, \ldots, s_N) \in \mathcal{R}^N$ in $\mathcal{C}_A$ and $\mathcal{C}_B$ satisfies the constraints

$$\| \mathsf{s} \|^2 \triangleq \sum_{j=1}^{N} s_j^2 = NE \tag{1.5}$$

and

$$\| \mathsf{s} \|^2 \leq NE \tag{1.6}$$

respectively, and $\mathcal{C}_C$ satisfies

$$\sum_{\mathsf{s} \in \mathcal{C}_C} P(\mathsf{s}) \| \mathsf{s} \|^2 = \sum_{m=1}^{M} \frac{1}{M} \sum_{j=1}^{N} s_{mj}^2 \leq NE \tag{1.7}$$

for some positive constant E (joules/ch.use). Observe that the first two codes can be recognized respectively as two sets of M points *on the surface of* and *on or inside* an N-dimensional euclidean sphere of radius $\sqrt{NE}$; that is why these are said to satisfy shell and sphere constraints respectively.

Now consider random coding over the corresponding three code ensembles $\{\mathcal{C}_A\}$, $\{\mathcal{C}_B\}$ and $\{\mathcal{C}_C\}$. Shannon [Sha59] showed that the ensemble average of the probability of maximum likelihood decoding error for $\{\mathcal{C}_A\}$ is smaller than those for $\{\mathcal{C}_B\}$ and $\{\mathcal{C}_C\}$. This fact can be justified heuristically by observing that $\{\mathcal{C}_A\}$ is a subset of $\{\mathcal{C}_B\}$ which in turn is a subset of $\{\mathcal{C}_C\}$ and, $\{\mathcal{C}_B\}$ and $\{\mathcal{C}_C\}$ contain some very poor codes that are not contained in $\{C_A\}$. Therefore, if a code is to satisfy an energy constraint, it is desirable to have all codewords satisfy the constraint with equality. Since fixed-composition codes fulfill this requirement, they are expected to be beneficial. However, at this point, one has to make sure that it is really worth trying fixed-composition codes, i.e., there is a significant improvement which fixed-composition codes promise to provide so that one can undertake the additional difficulties in encoding and decoding fixed-composition codes.

Consider block coding over the AWGN channel described in the previous section and let N be the blocklength. Suppose that inputs to the channel are generated at a rate R bits per channel use. Then there exist $M = 2^{NR}$ distinct messages to send through the channel and one has to associate a distinct codeword $s_m = (s_{m1}, s_{m2}, \ldots, s_{mN})$ to each message m. Shannon [Sha59] showed that one can find at least one set of M codewords $\{\mathbf{s}_m\}$, constrained only in energy by

$$\| \mathbf{s}_m \|^2 \triangleq \sum_{j=1}^{N} s_{mj}^2 \leq NE \quad ; m = 1, 2, \ldots, M, \tag{1.8}$$

so that the probability of maximum likelihood decoding error is bounded by

$$P_{error} < 2^{-N(R_0^\star - R)} \quad , 0 < R < R_0^\star \tag{1.9}$$

where

$$R_0^\star \triangleq \frac{\log_2 e}{2}\left[1 + \frac{A}{2} - \sqrt{1 + \frac{A^2}{4}}\right] + \frac{1}{2}\log_2\left[\frac{1}{2}\left(1 + \sqrt{1 + \frac{A^2}{4}}\right)\right] \text{ bits/ch.use}, \tag{1.10}$$

and

$$A \triangleq \frac{E}{\sigma^2} \tag{1.11}$$

is the signal-to-noise ratio [WoJ65, pp.309-311], [Gal68,pp.333-343].

It is also shown in [Sha59] and [Gal65] that the cutoff rate for $\{C_A\}$ is equal to $R_0^\star$. Recall that the fixed-composition code with parameter (N, Q) is a code over a finite quantization of the real line whereas, in deriving $R_0^\star$, Shannon and Gallager considered codes of arbitrary blocklengths with code letters being arbitrary real numbers. Noting also that fixed-composition codes satisfy the shell constraint, it follows that cutoff rate for fixed-composition ensemble approaches $R_0^\star$ as the quantization is made finer.

On the other hand, Gallager [Gal86] considered random coding for the AWGN channel under the shell constraint and showed that, in the limit of large signal-to-noise ratio,

$$C - R_0^\star = (1 - \frac{\log_2 e}{2}) \approx 0.28 \text{ bits/ch.use} \tag{1.12}$$

where

$$C = \frac{1}{2}\log_2(1 + A) \text{ bits/ch.use} \tag{1.13}$$

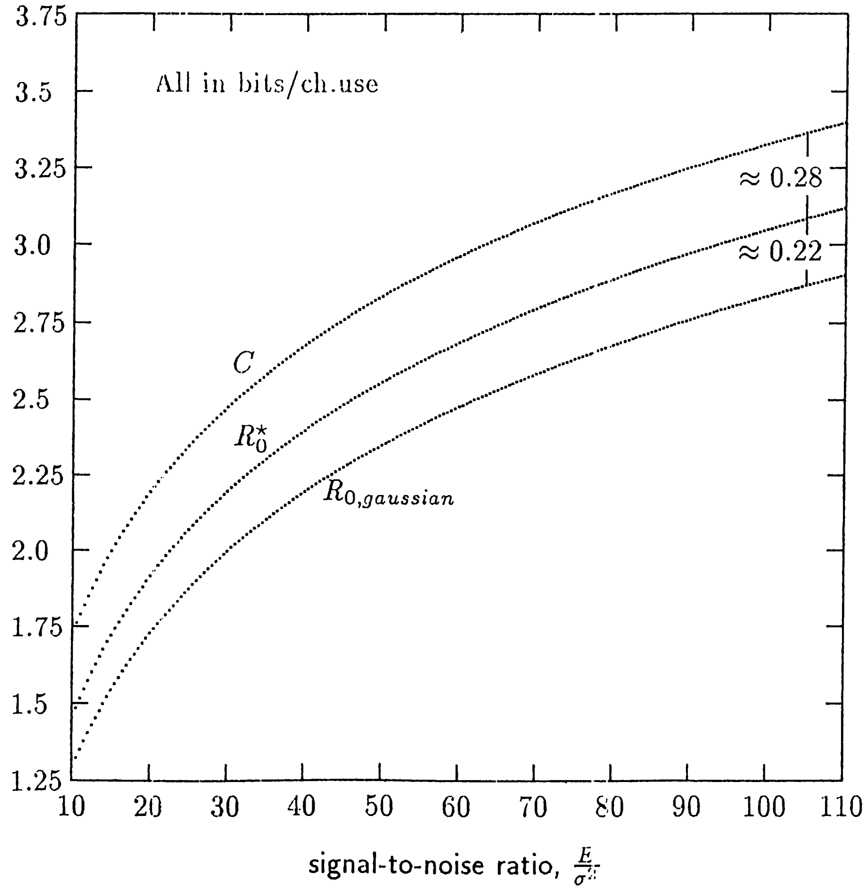

Figure 1.1: C, R_0^* and $R_{0,gaussian}$ over AWGN channel.

is the capacity of AWGN channel.

Now consider the independent-letters code ensemble in which the code letters are selected independently from a zero mean gaussian distribution with variance E. The cutoff rate for this ensemble is given by

$$R_{0,gaussian} = \frac{1}{2}\log_2(1 + \frac{A}{2}) \text{ bits/ch.use.} \tag{1.14}$$

It can be shown that, in the limit of large signal-to-noise ratio, we have

$$R_0^* - R_{0,gaussian} = \frac{1}{2}(\log_2 e - 1) \approx 0.22 \text{ bits/ch.use,} \tag{1.15}$$

which is a significantly large gap. On the other hand, for low signal-to-noise ratios, we have

$$R_0^* \approx R_{0,gaussian} \approx C/2 \approx A/4 \tag{1.16}$$

which shows that no coding gain can be achieved for low signal-to-noise ratios.

Although choosing code letters from a gaussian distribution does not maximize the cutoff rate of independent-letters ensembles and there exist already better codes achieving higher cutoff rates[1], this result together with the previous observations suggest that some benefit may result from using fixed-composition codes especially at moderate to high signal-to-noise ratios. These cutoff rates are shown in Figure 1.1 to clarify the above discussion.

1.4 Summary of Results

As stated before, it is theoretically expected to achieve some coding gains by using fixed-composition codes rather than independent-letters codes. In this thesis work, our original contribution is to show that significant improvements in cutoff rate can be achieved in practice by using fixed-composition codes. Showing this requires computation of cutoff rates of various fixed-composition ensembles-a task involving certain computational difficulties which are discussed in Appendix B. The results of these computations indicate that, for certain fixed, finite code alphabets, it is possible to bridge the gap between the cutoff rate for optimum[2] independent-letters ensembles and R_0^* by up to 94.5% using fixed-composition codes of blocklength 40. These results together with those of asymptotic analysis of the cutoff rate for fixed-composition ensembles as blocklength tends to infinity are summarized and discussed in Chapter 2. The optimization of cutoff rate for independent-letters ensemble and the mathematical details of this asymptotic analysis are discussed in Appendices A and C, respectively. Finally, we conclude in Chapter 3 by discussion of further research topics.

[1]Suppose $E = 0.55$ and $\sigma^2 = 0.25$. Then $R_{0,gaussian} = 0.535$. But, our results show that a cutoff rate of 0.543 can be achieved by using an independent-letters code over a finite code alphabet (see Table 2.6).

[2]Here, the optimality of independent-letters ensembles is in the restricted sense that we optimize the cutoff rate over all probability distributions on a *finite* code alphabet, not on an unquantized one.

Chapter 2

COMPARISON OF R_0 FOR FIXED-COMPOSITION AND INDEPENDENT-LETTERS ENSEMBLES

In this chapter, the cutoff rates for the ensembles of fixed-composition and independent-letters codes over the energy constrained AWGN channel are compared for a particular finite code alphabet. Here, the cutoff rate for the ensemble of independent-letters codes is optimized over all probability distributions on the code alphabet.

2.1 Mathematical Preliminaries

Let $\mathcal{A} = \{a_l : l = 1, 2, \ldots, K\}$ be the code alphabet and $Q = \{q_l : l = 1, 2, \ldots, K\}$ be an associated probability distribution. Suppose that $\mathcal{A}$ and Q satisfy the energy constraint

$$\sum_{l=1}^{K} q_l a_l^2 \leq E \tag{2.1}$$

for some $E > 0$.

First, consider the ensemble of independent-letters codes over $\mathcal{A}$ containing $M = 2^{NR}$ codewords of blocklength N in which any codeword component s_j is assigned the code letter a_l with probability q_l independently as stated in Section 1.1. Observe that the union of all codes in this ensemble is $\mathcal{A}^N$. The cutoff rate for this ensemble is

$$R_{0,ilc} = -\frac{1}{N} \log_2 \sum_{\mathbf{s} \in \mathcal{A}^N} \sum_{\mathbf{s}' \in \mathcal{A}^N} P(\mathbf{s}) P(\mathbf{s}') e^{-d^2(\mathbf{s},\mathbf{s}')/8\sigma^2} \tag{2.2}$$

where

$$P(\mathbf{s}) = \prod_{j=1}^{N} P(s_j) \tag{2.3}$$

is the probability of codeword s and

$$d(\mathbf{s},\mathbf{s}') \triangleq \| \mathbf{s} - \mathbf{s}' \| = \left(\sum_{j=1}^{N} (s_j - s'_j)^2 \right)^{1/2} \tag{2.4}$$

is the euclidean distance between s and s'. Since codeword components are assigned code letters independently, this expression reduces to [WoJ65, p.316]

$$R_{0,ilc}(Q) = -\log_2 \sum_{l=1}^{K} \sum_{h=1}^{K} q_l q_h e^{-|a_l - a_h|^2/8\sigma^2}. \tag{2.5}$$

On the other hand, the cutoff rate for the ensemble of fixed-composition codes over $\mathcal{A}$ containing M codewords each with composition (N, Q) is given similarly by

$$R_{0,fcc} = -\frac{1}{N} \log_2 \sum_{\mathbf{s}\in\mathcal{F}_{N,Q}} \sum_{\mathbf{s}'\in\mathcal{F}_{N,Q}} P(\mathbf{s})P(\mathbf{s}')e^{-d^2(\mathbf{s},\mathbf{s}')/8\sigma^2} \tag{2.6}$$

where $\mathcal{F}_{N,Q}$ is the set of all (N, Q)-composition codewords and $P(\mathbf{s}) = 1/|\mathcal{F}_{N,Q}|$ for all $\mathbf{s} \in \mathcal{F}_{N,Q}$. Here, $|\mathcal{F}_{N,Q}|$ is the cardinality of $\mathcal{F}_{N,Q}$ and is given by

$$|\mathcal{F}_{N,Q}| = \frac{N!}{\prod_{l=1}^{K}(q_l N)!}. \tag{2.7}$$

Now, let $\mathbf{s}_i$ and $\mathbf{s}_k$ be two codewords in $\mathcal{F}_{N,Q}$ and $\mathbf{s}'_i$ be a permutation of $\mathbf{s}_i$. Then observe that there exists a codeword $\mathbf{s}'_k$, the same permutation of $\mathbf{s}_k$, such that $d(\mathbf{s}'_i, \mathbf{s}'_k) = d(\mathbf{s}_i, \mathbf{s}_k)$. Therefore, we have

$$R_{0,fcc} = -\frac{1}{N} \log_2 \sum_{\mathbf{s}\in\mathcal{F}_{N,Q}} P(\mathbf{s}) \left(\sum_{\mathbf{s}'\in\mathcal{F}_{N,Q}} P(\mathbf{s}')e^{-d^2(\mathbf{s},\mathbf{s}')/8\sigma^2} \right) \tag{2.8}$$

where the inner summation is the same constant for all $\mathbf{s} \in \mathcal{F}_{N,Q}$. Finally, it follows from this observation that

$$R_{0,fcc}(N, Q) = -\frac{1}{N} \log_2 \frac{1}{|\mathcal{F}_{N,Q}|} \sum_{\mathbf{s}\in\mathcal{F}_{N,Q}} e^{-d^2(\mathbf{s},\mathbf{s}_r)/8\sigma^2} \tag{2.9}$$

where $\mathbf{s}_r \in \mathcal{F}_{N,Q}$ is a fixed but arbitrary reference codeword.

In Appendix A, we discuss the optimization of $R_{0,ilc}$ over Q under an energy constraint. There, we show that the optimum probability distribution and the corresponding cutoff rate, denoted respectively by $Q^\star$ and $R^\star_{0,ilc}$, can be expressed as functions of E by

$$q_l^\star(E) = \beta_{l1}E + \beta_{l0} \quad ; \; l = 1, 2, \ldots, K, \tag{2.10}$$

and

$$R^\star_{0,ilc}(E) = -\log_2(\alpha_2 E^2 + \alpha_1 E + \alpha_0) \tag{2.11}$$

N	$\lvert\mathcal{F}_{N,Q}\rvert$
10	25200
20	8147739600
30	$3.885697753 \times 10^{15}$
40	$2.187409495 \times 10^{21}$

Table 2.1: $|\mathcal{F}_{N,Q}|$ for a four-letter alphabet with $Q = \{0.2, 0.3, 0.3, 0.2\}$.

where $\{\beta_{li}\}$, α_0, α_1 and α_2 are functions of the code alphabet $\mathcal{A}$ and the noise variance σ^2, and E is to satisfy $E_{min} \leq E \leq E_{sat}$ for some $E_{sat} > E_{min} > 0$ so that (2.1) is satisfied with equality (see Appendix A). Therefore, it is reasonable to compare $R^{\star}_{0,ilc}(E)$ and $R_{0,fcc}(N, Q^{\star}(E))$ at each $E \in [E_{min}, E_{sat}]$ such that $Nq^{\star}_l(E)$ is an integer, which is indeed the main aim of this work.

The computation of $R_{0,fcc}(N, Q^{\star})$ for finite N involves the enumeration of all $|\mathcal{F}_{N,Q^{\star}}|$ codewords in $\mathcal{F}_{N,Q^{\star}}$. Unfortunately, the complexity of this enumeration task is exponential in N. To have an idea about how fast the complexity increases, take the numerical results in Table 2.1 for a four-letter alphabet. For larger alphabet sizes, the complexity is even higher. Despite these huge numbers, cutoff rates are computed for various probability distributions on a four-letter alphabet and blocklength being equal to 40. The details of this computation task are discussed in Appendix B. This problem of computational complexity leads us to study the asymptotic behavior of $R_{0,fcc}(N, Q^{\star})$ as N tends to infinity which we discuss in Appendix C.

We are now in a position to summarize and discuss the results of comparison of $R^{\star}_{0,ilc}$ and $R_{0,fcc}(N, Q^{\star})$ for $N = 40$ and ∞ where the code alphabet is fixed to be the four-letter symmetric alphabet $\mathcal{A}_4 = \{\pm 0.5, \pm 1.5\}$ and the noise variance σ^2 runs from 0.05 to 0.4 in steps of 0.05.

2.2 Comparison of $R^{\star}_{0,ilc}$ and $R_{0,fcc}$

The results of Appendix A show that the probability distribution $Q^{\star}$ which maximizes $R_{0,ilc}$ over $\mathcal{A}_4$ is symmetric, i.e.

$$q^{\star}_1 = q^{\star}_4, \;\; q^{\star}_2 = q^{\star}_3, \;\; \text{and hence } q^{\star}_2 = 0.5 - q^{\star}_1 \tag{2.12}$$

as one should expect due to the symmetry of the code alphabet. It is also shown in Appendix A that, regardless of the value of σ^2,

$$q^{\star}_1 = -0.0625 + 0.25E. \tag{2.13}$$

From these observations, it follows that

$$E = \sum_{l=1}^{4} q_l^\star a_l^2 \tag{2.14}$$

$$= 2 \cdot (-0.0625 + 0.25E) \cdot 1.5^2 + 2 \cdot (0.5625 - 0.25E) \cdot 0.5^2 \tag{2.15}$$

as desired. Now suppose that we wish to compute $R_{0,fcc}(40, Q^\star)$. Then the letter probabilities have to be multiples of 0.025. Notice that the choice of $E = 0.35 + 0.1k, k = 0, 1, \ldots, 17$ yields all such nontrivial probability distributions on $\mathcal{A}_4$[1]. The cutoff rates $R_{0,ilc}^\star$ and $R_{0,fcc}$ for $\mathcal{A}_4$ are compared in Tables 2.2-2.9 for $N = 40$ and $\sigma^2 = 0.05, 0.10, \ldots, 0.40$. In these tables, E is a free parameter running from E_{min} to E_{sat} in steps of 0.10. These tables also include $R_0^\star$ (1.10) in order to show the extent to which the fixed-composition code improves the cutoff rate. As a measure of this quantity, the percentage improvement factor defined as

$$\eta = \frac{R_{0,fcc} - R_{0,ilc}^\star}{R_0^\star - R_{0,ilc}^\star} \times 100 \tag{2.16}$$

is also included in these tables. These results are also depicted in Figures 2.1-2.8 together with the asymptotic values that $R_{0,fcc}(N, Q^\star)$ takes as N tends to infinity. In these figures, $R_{0,fcc}(N, Q^\star)$ values for all possible $N \leq 40$ are depicted. The circles show $R_{0,fcc}(40, Q^\star)$ and the diamonds below the circles correspond to $R_{0,fcc}(N, Q^\star)$ for $N < 40$. For example, in Figure 2.1, the diamonds at $E = 1.05$ correspond to blocklengths of 10, 20 and 30, respectively starting from the one at the bottom.

[1] Refer to Appendix A for the values of E_{min} and E_{sat} for $\mathcal{A}_4$.

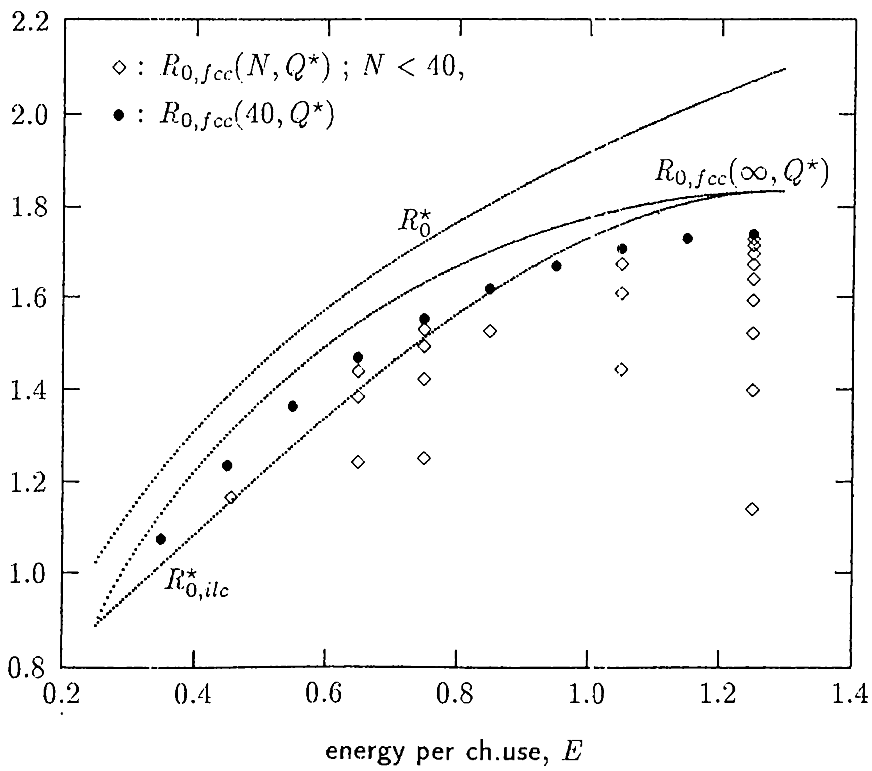

Figure 2.1: $R^\star_{0,ilc}$ and $R_{0,fcc}(N, Q^\star)$ for $\sigma^2 = 0.05$.

E	$q_1^\star$	$R^\star_{0,ilc}$	$R_{0,fcc}$	$R_0^\star$	η
0.35	0.025	1.019047	1.074402	1.227390	26.6
0.45	0.050	1.149965	1.234435	1.386135	35.8
0.55	0.075	1.277097	1.362302	1.516461	35.6
0.65	0.100	1.398178	1.466158	1.626947	29.7
0.75	0.125	1.510515	1.550317	1.722812	18.7
0.85	0.150	1.611058	1.617367	1.807462	3.2
0.95	0.175	1.696540	1.668961	1.883242	-14.8
1.05	0.200	1.763726	1.706168	1.951830	-30.6
1.15	0.225	1.809738	1.729655	2.014472	-39.1
1.25	0.250	1.832420	1.739781	2.072114	-38.6

Table 2.2: Comparison of $R^\star_{0,ilc}$ and $R_{0,fcc}$ for $\mathcal{A}_4$ and $\sigma^2 = 0.05$.

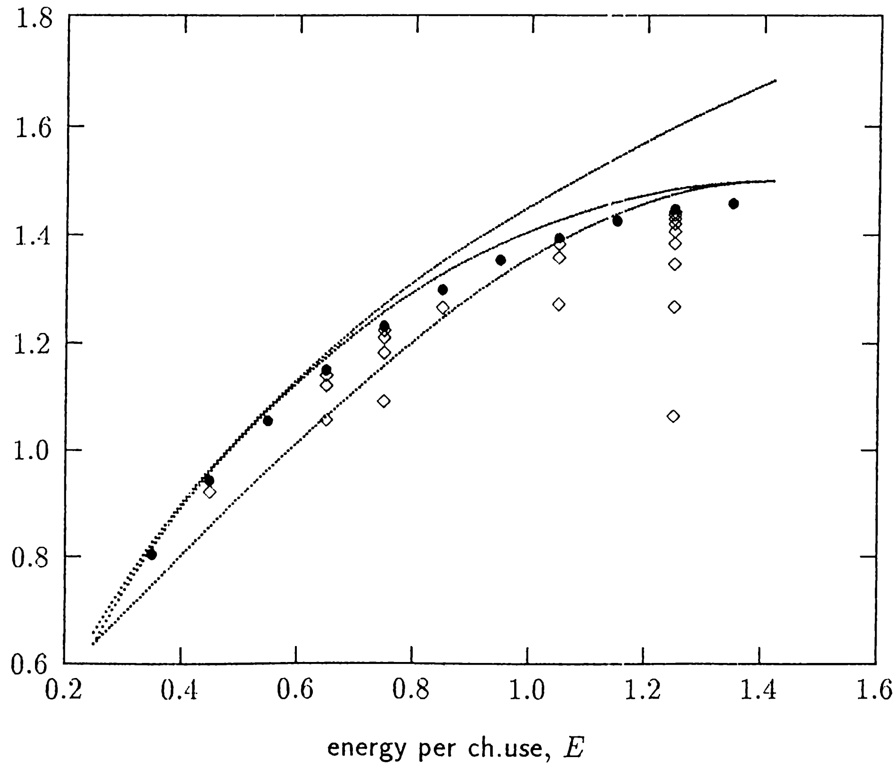

Figure 2.2: $R^{\star}_{0,ilc}$ and $R_{0,fcc}(N, Q^{\star})$ for $\sigma^2 = 0.10$.

E	$q_1^{\star}$	$R^{\star}_{0,ilc}$	$R_{0,fcc}$	$R_0^{\star}$	η
0.35	0.025	0.747307	0.803548	0.825997	71.5
0.45	0.050	0.856146	0.940396	0.964115	78.0
0.55	0.075	0.961877	1.053851	1.080826	77.3
0.65	0.100	1.063106	1.149289	1.181692	72.7
0.75	0.125	1.158227	1.229599	1.270414	63.6
0.85	0.150	1.245454	1.296544	1.349558	49.1
0.95	0.175	1.322875	1.351262	1.420966	28.9
1.05	0.200	1.388544	1.394495	1.486001	6.1
1.15	0.225	1.440605	1.426704	1.545698	-13.2
1.25	0.250	1.477440	1.448128	1.600861	-23.7
1.35	0.275	1.497820	1.458811	1.652127	-25.3

Table 2.3: Comparison of $R^{\star}_{0,ilc}$ and $R_{0,fcc}$ for $\mathcal{A}_4$ and $\sigma^2 = 0.1$.

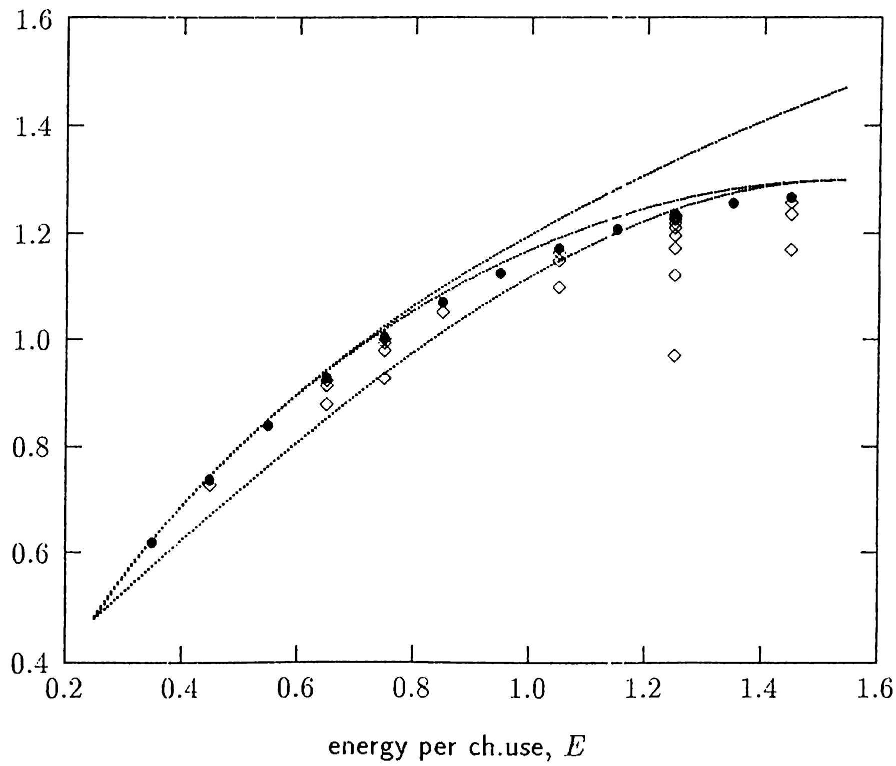

Figure 2.3: $R^\star_{0,ilc}$ and $R_{0,fcc}(N, Q^\star)$ for $\sigma^2 = 0.15$.

E	$q_1^\star$	$R^\star_{0,ilc}$	$R_{0,fcc}$	$R_0^\star$	η
0.35	0.025	0.575721	0.618477	0.625949	85.1
0.45	0.050	0.670384	0.737019	0.746369	87.7
0.55	0.075	0.762523	0.838728	0.850821	86.3
0.65	0.100	0.851193	0.926770	0.942769	82.5
0.75	0.125	0.935327	1.002984	1.024744	75.7
0.85	0.150	1.013746	1.068544	1.098620	64.6
0.95	0.175	1.085192	1.124231	1.165808	48.4
1.05	0.200	1.148362	1.170561	1.227390	28.1
1.15	0.225	1.201970	1.207860	1.284213	7.2
1.25	0.250	1.244815	1.236293	1.336948	-9.2
1.35	0.275	1.275859	1.255888	1.386135	-18.1
1.45	0.300	1.294304	1.266531	1.432216	-20.1

Table 2.4: Comparison of $R^\star_{0,ilc}$ and $R_{0,fcc}$ for $\mathcal{A}_4$ and $\sigma^2 = 0.15$.

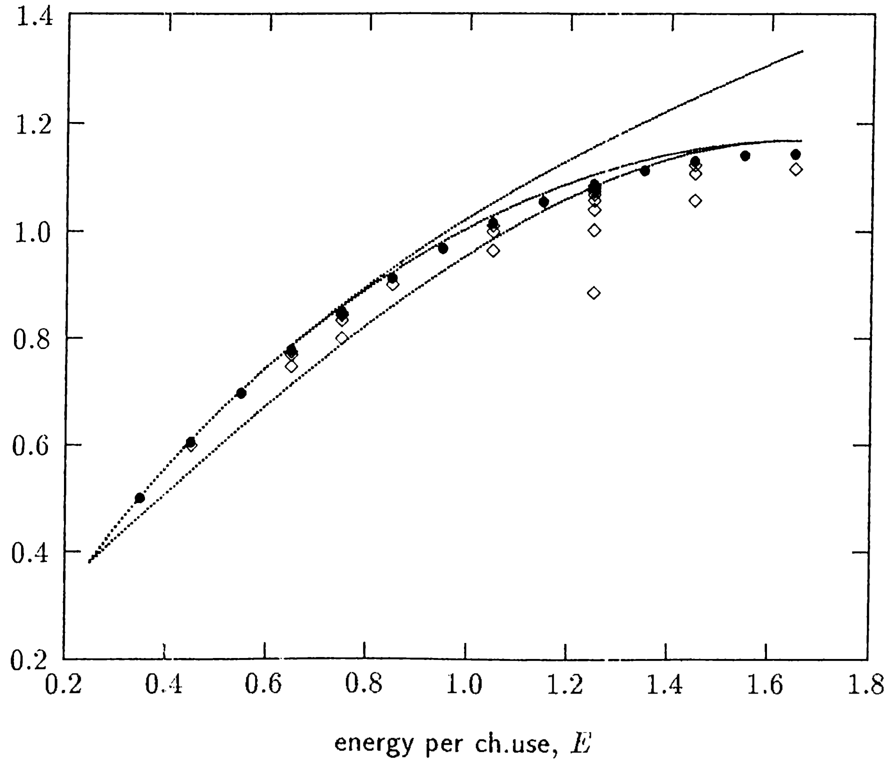

Figure 2.4: $R^{\star}_{0,ilc}$ and $R_{0,fcc}(N,Q^{\star})$ for $\sigma^2 = 0.20$.

E	$q_1^\star$	$R^\star_{0,ilc}$	$R_{0,fcc}$	$R_0^\star$	η
0.35	0.025	0.467204	0.500028	0.503806	89.7
0.45	0.050	0.551395	0.604260	0.609554	90.9
0.55	0.075	0.633489	0.695999	0.703306	89.5
0.65	0.100	0.712809	0.777156	0.787200	86.5
0.75	0.125	0.788595	0.848917	0.862939	81.1
0.85	0.150	0.860022	0.912074	0.931865	72.5
0.95	0.175	0.926204	0.967165	0.995041	59.5
1.05	0.200	0.986220	1.014555	1.053312	42.2
1.15	0.225	1.039142	1.054471	1.107359	22.5
1.25	0.250	1.084073	1.087032	1.157735	4.0
1.35	0.275	1.120188	1.112252	1.204896	-9.4
1.45	0.300	1.146776	1.130046	1.249217	-16.3
1.55	0.325	1.163291	1.140216	1.291016	-18.1
1.65	0.350	1.169378	1.142430	1.330560	-16.7

Table 2.5: Comparison of $R^\star_{0,ilc}$ and $R_{0,fcc}$ for $\mathcal{A}_4$ and $\sigma^2 = 0.2$.

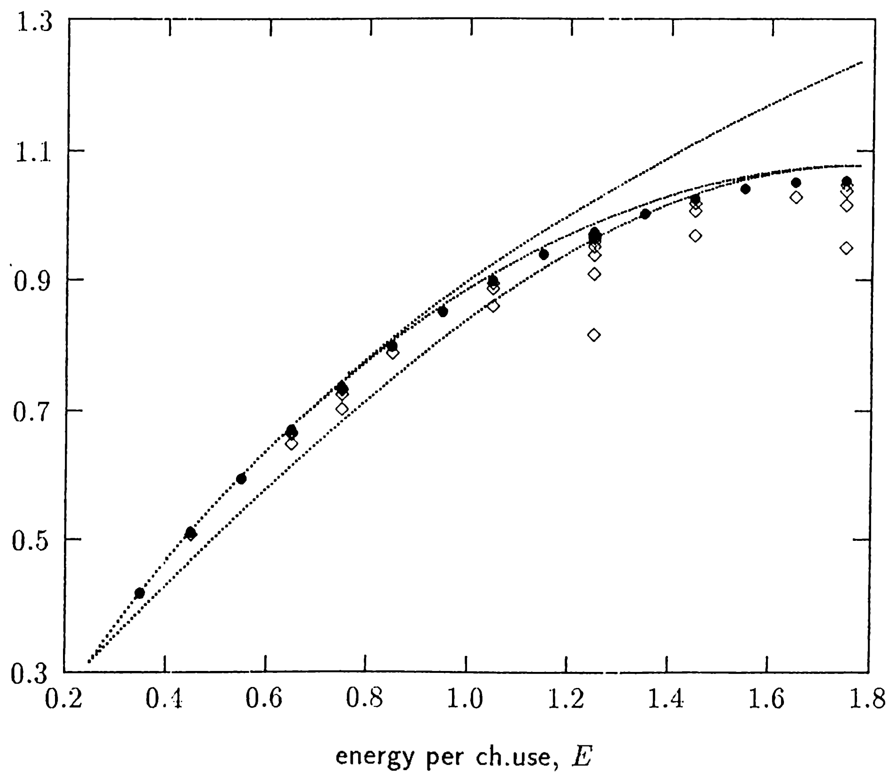

Figure 2.5: $R^{\star}_{0,ilc}$ and $R_{0,fcc}(N, Q^{\star})$ for $\sigma^2 = 0.25$.

E	$q_1^{\star}$	$R^{\star}_{0,ilc}$	$R_{0,fcc}$	$R_0^{\star}$	η
0.35	0.025	0.393180	0.418942	0.421267	91.7
0.45	0.050	0.468994	0.511527	0.514992	92.5
0.55	0.075	0.543042	0.594624	0.599559	91.3
0.65	0.100	0.614822	0.669408	0.676304	88.8
0.75	0.125	0.683776	0.736654	0.746369	84.5
0.85	0.150	0.749301	0.796897	0.810710	77.5
0.95	0.175	0.810752	0.850512	0.870118	67.0
1.05	0.200	0.867457	0.897753	0.925247	52.4
1.15	0.225	0.918730	0.938782	0.976638	34.6
1.25	0.250	0.963894	0.973680	1.024744	16.1
1.35	0.275	1.002301	1.002458	1.069942	0.2
1.45	0.300	1.033363	1.025047	1.112554	-10.5
1.55	0.325	1.056577	1.041303	1.152849	-15.9
1.65	0.350	1.071550	1.050983	1.191061	-17.2
1.75	0.375	1.078020	1.053719	1.227390	-16.3

Table 2.6: Comparison of $R^{\star}_{0,ilc}$ and $R_{0,fcc}$ for $\mathcal{A}_4$ and $\sigma^2 = 0.25$.

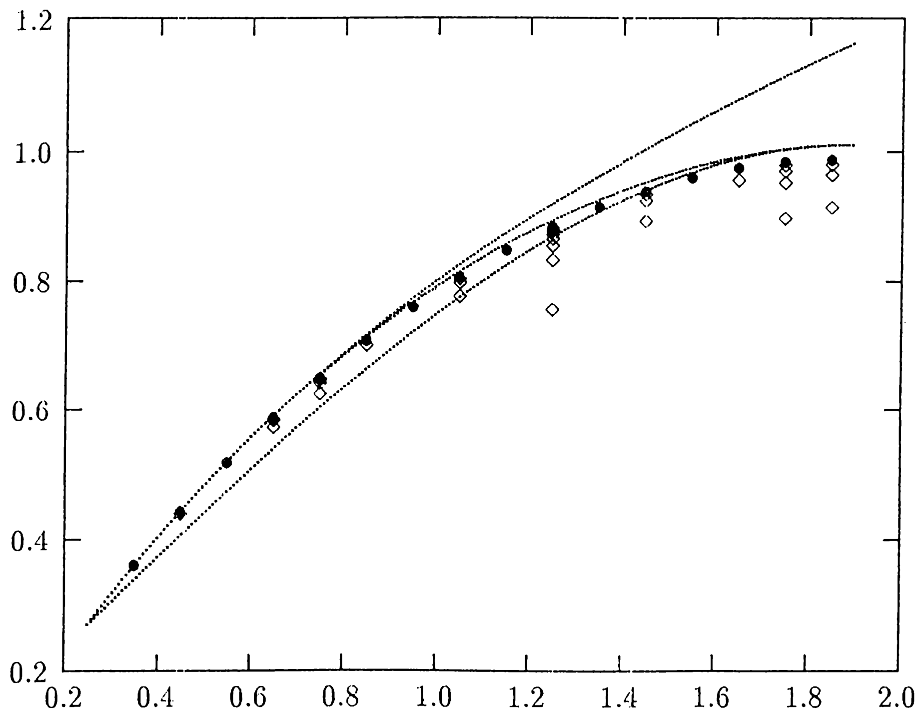

Figure 2.6: $R^{\star}_{0,ilc}$ and $R_{0,fcc}(N, Q^{\star})$ for $\sigma^2 = 0.30$.

E	$q_1^{\star}$	$R^{\star}_{0,ilc}$	$R_{0,fcc}$	$R_0^{\star}$	η
0.35	0.025	0.339533	0.360186	0.361775	92.9
0.45	0.050	0.408435	0.443185	0.445636	93.4
0.55	0.075	0.475836	0.518819	0.522369	92.4
0.65	0.100	0.541354	0.587829	0.592828	90.3
0.75	0.125	0.604568	0.650731	0.657785	86.7
0.85	0.150	0.665028	0.707888	0.717922	81.0
0.95	0.175	0.722250	0.759561	0.773825	72.3
1.05	0.200	0.775731	0.805931	0.825997	60.1
1.15	0.225	0.824954	0.847112	0.874868	44.4
1.25	0.250	0.869402	0.883161	0.920803	26.8
1.35	0.275	0.908567	0.914083	0.964115	9.9
1.45	0.300	0.941974	0.939828	1.005073	-3.4
1.55	0.325	0.969189	0.960288	1.043910	-11.9
1.65	0.350	0.989844	0.975285	1.080826	-16.0
1.75	0.375	1.003648	0.984556	1.115997	-17.0
1.85	0.400	1.010399	0.987713	1.149574	-16.3

Table 2.7: Comparison of $R^{\star}_{0,ilc}$ and $R_{0,fcc}$ for $\mathcal{A}_4$ and $\sigma^2 = 0.3$.

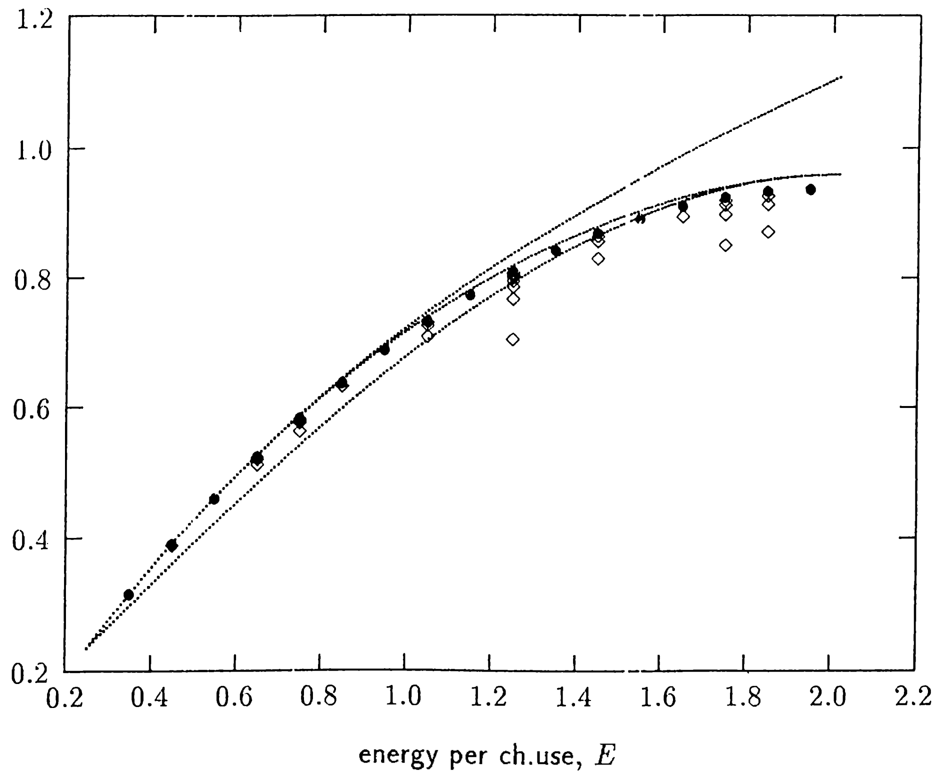

Figure 2.7: $R^{\star}_{0,ilc}$ and $R_{0,fcc}(N, Q^{\star})$ for $\sigma^2 = 0.35$.

E	$q_1^\star$	$R^{\star}_{0,ilc}$	$R_{0,fcc}$	$R_0^\star$	η
0.35	0.025	0.298858	0.315733	0.316893	93.6
0.45	0.050	0.361959	0.390774	0.392600	94.0
0.55	0.075	0.423773	0.459979	0.462649	93.1
0.65	0.100	0.484004	0.523830	0.527600	91.4
0.75	0.125	0.542328	0.582674	0.587986	88.4
0.85	0.150	0.588400	0.636764	0.644293	86.5
0.95	0.175	0.651851	0.686283	0.696959	76.3
1.05	0.200	0.702296	0.731359	0.746369	65.9
1.15	0.225	0.749337	0.772074	0.792862	52.2
1.25	0.250	0.792574	0.808469	0.836734	36.0
1.35	0.275	0.831609	0.840546	0.878242	19.2
1.45	0.300	0.866056	0.868267	0.917612	4.3
1.55	0.325	0.895556	0.891552	0.955040	-6.7
1.65	0.350	0.919784	0.910271	0.990700	-13.4
1.75	0.375	0.938460	0.924236	1.024744	-16.5
1.85	0.400	0.951362	0.933173	1.057305	-17.2
1.95	0.425	0.958332	0.936686	1.088504	-16.6

Table 2.8: Comparison of $R^{\star}_{0,ilc}$ and $R_{0,fcc}$ for $\mathcal{A}_4$ and $\sigma^2 = 0.35$.

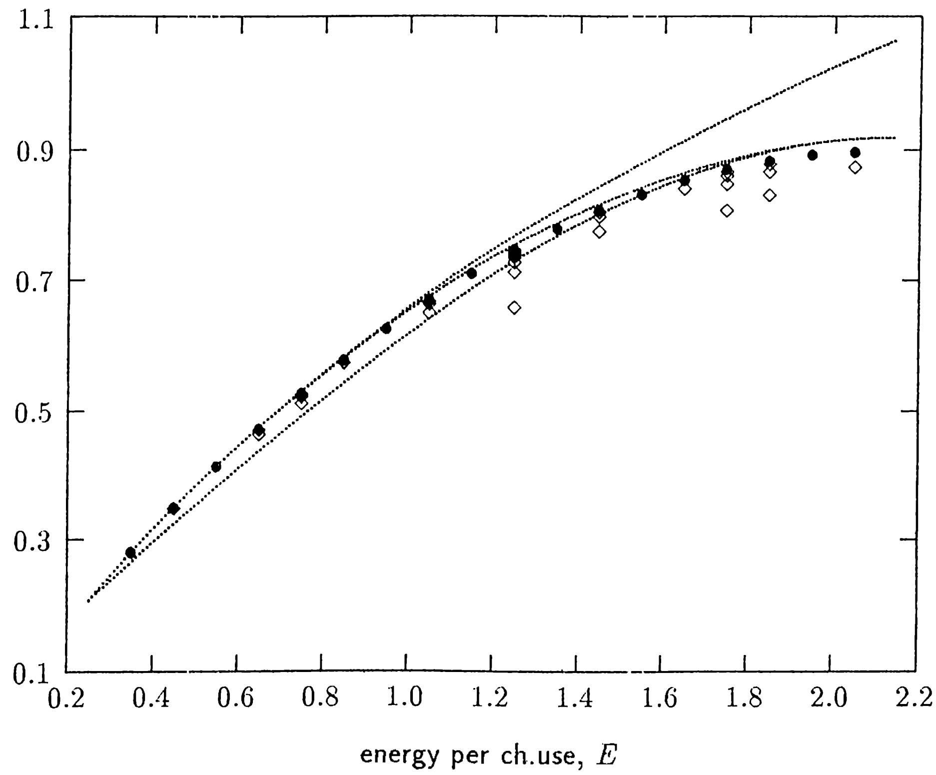

Figure 2.8: $R^{\star}_{0,ilc}$ and $R_{0,fcc}(N, Q^{\star})$ for $\sigma^2 = 0.40$.

E	$q_1^{\star}$	$R^{\star}_{0,ilc}$	$R_{0,fcc}$	$R_0^{\star}$	η
0.35	0.025	0.266945	0.280964	0.281849	94.1
0.45	0.050	0.325114	0.349336	0.350747	94.5
0.55	0.075	0.382171	0.412995	0.415071	93.7
0.65	0.100	0.437882	0.472264	0.475199	92.1
0.75	0.125	0.491995	0.527382	0.531503	89.6
0.85	0.150	0.544238	0.578528	0.584336	85.5
0.95	0.175	0.594324	0.625831	0.634025	79.4
1.05	0.200	0.641955	0.669384	0.680868	70.5
1.15	0.225	0.686819	0.709246	0.725130	58.5
1.25	0.250	0.728601	0.745446	0.767051	43.8
1.35	0.275	0.766985	0.777986	0.806841	27.6
1.45	0.300	0.801662	0.806836	0.844690	12.0
1.55	0.325	0.832333	0.831937	0.880763	-0.8
1.65	0.350	0.858721	0.853196	0.915210	-9.8
1.75	0.375	0.880575	0.870473	0.948162	-14.9
1.85	0.400	0.897680	0.883576	0.979737	-17.2
1.95	0.425	0.909863	0.892232	1.010040	-17.6
2.05	0.450	0.916995	0.896039	1.039164	-17.2

Table 2.9: Comparison of $R^{\star}_{0,ilc}$ and $R_{0,fcc}$ for $\mathcal{A}_4$ and $\sigma^2 = 0.4$.

2.3 Discussion of Results

Now, recall that $[E_{min}, E_{sat}]$ is the interval for E on which the energy constraint (2.1) is satisfied with equality. Therefore, as stated in Chapter 1, we expect that fixed-composition codes provide coding gains for $E \in [E_{min}, E_{sat}]$. The results are in accordance with our expectation (see Figures 2.1-2.8), i.e.,

$$R_{0,fcc}(\infty, Q^\star) > R^\star_{0,ilc} \text{ for } E_{min} < E < E_{sat} \tag{2.17}$$

and

$$R_{0,fcc}(\infty, Q^\star) = R^\star_{0,ilc} \text{ at } E = E_{min} \text{ and } E_{sat}. \tag{2.18}$$

Here, we leave (2.18) as a conjecture the proof of which needs further work. But, since E_{min} and E_{sat} are the boundary points of the region on which the energy constraint is satisfied with equality, it is quite normal that one expects no coding gain at these energy values.

The trend common to Figures 2.1-2.8 indicates that for $E > E_{sat}$ we have

$$R_{0,fcc}(\infty, Q^\star(E)) = R^\star_{0,ilc}(E) = R^\star_{0,ilc}(E_{sat}) \tag{2.19}$$

justifying the use of label 'saturation' for the situation. For $E < E_{min}$, to argue in a similar way is difficult; because, in this case, the size of the code alphabet K is to be decreased and, hence, everything changes.

For fixed σ^2, as E gets closer to E_{sat}, $R^\star_{0,ilc}$ starts beating $R_{0,fcc}(40)$. Having noted above that $R^\star_{0,ilc}$ and $R_{0,fcc}$ become asymptotically equal at E_{sat} as N tends to infinity, we should increase the size of the code alphabet in order to change the picture. This result is in accordance with the general statement that the cutoff rate for an ensemble of codes over a finite code alphabet saturates as signal-to-noise ratio increases and one should increase the alphabet size to achieve higher cutoff rates for large signal-to-noise ratios.

Observe that for fixed σ^2, the percentage improvement factor peaks around $E = 0.45$ and then decreases monotonically. This is because $R_{0,fcc}$ and $R^\star_{0,ilc}$ saturate for large E whereas $R^\star_0$ increases monotonically.

The results indicate that fixed-composition codes fare significantly better than independent-letters codes: Improvements from 35.8% (Table 2.2) up to 94.5% (Table 2.9) in the percentage improvement are achievable by using fixed-composition codes over $\mathcal{A}_4$ with parameter $(40, Q^\star)$. Observe that in going from $\sigma^2 = 0.4$ to 0.05, the percentage

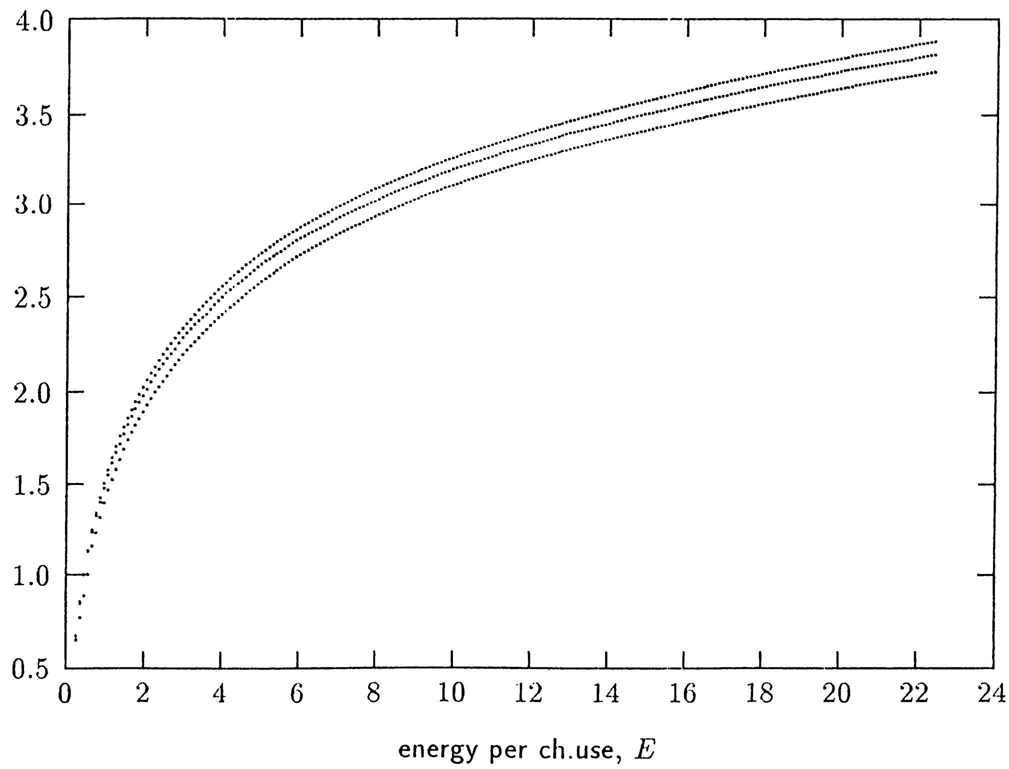

Figure 2.9: $R_0^\star$, $env\{R_{0,fcc}(\infty, Q^\star)\}$ and $env\{R_{0,ilc}^\star\}$ for $\sigma^2 = 0.1$.

improvement factor decreases monotonically for all E values, in other words, $(40, Q^\star)$-composition codes over $\mathcal{A}_4$ start doing worse as σ^2 decreases. Another observation in the same direction is that the crossover E after which $R_{0,ilc} > R_{0,fcc}(40, Q^\star)$ decreases as σ^2 decreases. So, similarly for small σ^2 values, i.e. for large signal-to-noise ratios, we should increase the alphabet size to achieve a further improvement provided that the gap between $R_0^\star$ and $R_{0,ilc}^\star$ is significantly large. Observe that as σ^2 increases, $R_0^\star$, $R_{0,fcc}$ and $R_{0,ilc}^\star$ get closer and closer to each other at E close to E_{min}. An abrupt change of this kind in the behavior of these cutoff rates can be recognized in going from $\sigma^2 = 0.05$ to 0.10.

The above discussion is made for $N = 40$. But, Figures 2.1-2.8 also reveal that as σ^2 increases (signal-to-noise ratio decreases) even smaller values of N provides coding gains.

As the need of increasing the size of the code alphabet arises for large signal-to-noise ratios, in Figure 2.9, we compare the envelopes of $R_{0,ilc}(E)$ and $R_{0,fcc}(\infty, Q^\star(E))$ for $\mathcal{A}_4, \mathcal{A}_5, \ldots, \mathcal{A}_{21}$, $\sigma^2 = 0.1$, and $E \in [E_{min}(K = 4), E_{sat}(K = 16)]$. This figure shows that using code alphabets of the particular form defined by (A.23), it is possible to bridge the gap between $R_{0,ilc}^\star$ and $R_0^\star$ by 56% for high signal-to-noise ratios. To obtain closer cutoff rates to $R_0^\star$, we should definitely use code alphabets which are finer quantizations of the real line.

Finally, we finish this chapter by observing that, for relatively medium signal-to-noise ratios at which it is still reasonable to use the four-letter symmetric alphabet $\mathcal{A}_4$, we can achieve cutoff rates within approximately 1% of R_0^* by using fixed-composition codes of blocklength 40. This result has a practical significance as will be discussed in Chapter 3.

Chapter 3

CONCLUSION AND FURTHER RESEARCH TOPICS

The results discussed in Chapter 2 prove the basic claim stated in Chapter 1: For medium to high signal-to-noise ratios, one can achieve significant coding gains by using fixed-composition codes rather than codes selected from an independent-letters ensemble even when the selection is done from an optimum distribution. It is shown that for moderate signal-to-noise ratios it is possible to achieve cutoff rates within approximately 1% of R_0^* by using fixed-composition codes of blocklength 40 over a four-letter symmetric code alphabet.

This is an important result as it is stated in Chapter 1 that fixed-composition codes are expected to achieve cutoff rates getting closer and closer to R_0^* as the quantization is made finer. On the other hand, a blocklength of 40 is a reasonable one for practical purposes. Arıkan [Arı89] has recently proposed a method for constructing fixed-composition trellis codes with smallest possible degree which is independent of the blocklength.

Finally, we conclude by pointing out two topics that may be of interest for further research. Firstly, for relatively larger values of signal-to-noise ratio, the need of increasing the size of the code alphabet arises as the results of Chapter 2 indicate. One may seek ways of computing cutoff rates of fixed-composition codes over code alphabets of sizes larger than 4. Secondly, as also pointed out in [Arı89], the sequential decoding of fixed-composition codes needs to be investigated further. Namely, the problem stems from the memory introduced by the fixed-composition constraint; hence, optimum metrics for sequential decoding require excessive computation.

Appendix A

Optimization of $R_{0,ilc}$ under Energy Constraint

Suppose that the code alphabet $\mathcal{A}$ of size K is fixed. Rewriting the expression for $R_{0,ilc}(Q)$ (2.5) as [WoJ65, p.354]

$$R_{0,ilc}(Q) = -\log_2 \sum_{l=1}^{K} \sum_{h=1}^{K} q_l b_{lh} q_h \tag{A.1}$$

where

$$b_{lh} \triangleq e^{-d_{lh}^2/8\sigma^2} = b_{hl}, \tag{A.2}$$

and

$$d_{lh} \triangleq |a_l - a_h| = d_{hl}, \tag{A.3}$$

our objective is to find the probability distribution Q^* for which $R_{0,ilc}$ is maximum, subject to an energy constraint. This is same as minimizing

$$e^{-R_{0,ilc}} = \sum_{l=1}^{K} \sum_{h=1}^{K} q_l b_{lh} q_h \tag{A.4}$$

over all valid probability distributions Q on $\mathcal{A}$ such that $\mathcal{A}$ and Q satisfy the energy constraint (2.1).

A.1 Minimization of $e^{-R_{0,ilc}}$

Let $2\lambda_0$ and $2\lambda_1$ be Lagrange multipliers. Then we have

$$\frac{\partial}{\partial q_l}\left[\sum_{l=1}^{K}\sum_{h=1}^{K} q_l b_{lh} q_h - 2\lambda_0 \sum_{l=1}^{K} q_l - 2\lambda_1 \sum_{l=1}^{K} q_l a_l^2\right] = 2\left[\sum_{h=1}^{K} b_{lh} q_h - \lambda_0 - \lambda_1 a_l^2\right] \quad ; \quad l = 1, 2, \ldots, K. \tag{A.5}$$

Setting each partial derivative equal to zero yields the following set of K inhomogeneous linear equations:

$$\sum_{h=1}^{K} b_{lh} q_h = \lambda_0 + \lambda_1 a_l^2 \quad ; \; l = 1, 2, \ldots, K. \tag{A.6}$$

Now, suppose that not only the code alphabet but also the noise variance σ^2 is fixed. Then these linear equations can be solved for $\{q_l^\star\}$ in terms of λ_0 and λ_1 which can be determined using the constraints $\sum_{l=1}^{K} q_l^\star = 1$ and $\sum_{l=1}^{K} q_l^\star a_l^2 = E$. Whenever the $\{q_l^\star\}$ are all non-negative, they maximize $R_{0,ilc}$ with energy constraint satisfied with equality and we have

$$\sum_{l=1}^{K}\sum_{h=1}^{K} q_l^\star b_{lh} q_h^\star = \sum_{l=1}^{K} q_l^\star(\lambda_0 + \lambda_1 a_l^2) = \lambda_0 + \lambda_1 E, \tag{A.7}$$

and hence

$$R_{0,ilc}^\star = -\log_2(\lambda_0 + \lambda_1 E) \text{ bits/ch.use} \tag{A.8}$$

But, observe that, for large values of E, it may happen that no valid probability distribution $\{q_l^\star\}$ solving (A.6) and at the same time satisfying the energy constraint with equality exists. This corresponds to the case of having the energy constraint inactive, or equivalently $\lambda_1 = 0$. So, solving (A.6) for $\{q_l^{sat}\}$ with λ_1 set equal to zero, we have

$$R_{0,ilc}^{sat} = -\log_2 \lambda_0. \tag{A.9}$$

This solution holds whenever $E > E_{sat}$ where

$$E_{sat} = \sum_{l=1}^{K} q_l^{sat} a_l^2; \tag{A.10}$$

that is why associated quantities are labeled with 'sat' standing for 'saturation'.

This completes the optimization of $R_{0,ilc}$ under an energy constraint. In the following section, we express $Q^\star$ and $R_{0,ilc}^\star$ as functions of E for $E \leq E_{sat}$.

A.2 $Q^\star$ and $R_{0,ilc}^\star$ as Functions of E

Recall that we have all b_{lh} determined since $\mathcal{A}$ and σ^2 are fixed. Therefore, we can solve the problem explicitly. Fortunately, the solution has a simple form as the $\{q_l^\star\}$ are linear functions of E, and $R_{0,ilc}^\star$ is given by the logarithm of a quadratic function of E as expressed in (2.11).

Let $B^{-1} = [b'_{lh}]_{l,h=1}^{K}$ be the inverse of the matrix $B = [b_{lh}]_{l,h=1}^{K}$. Then from (A.6)

$$q_l^\star = \lambda_0 \left(\sum_{h=1}^{K} b'_{lh}\right) + \lambda_1 \left(\sum_{h=1}^{K} b'_{lh} a_h^2\right) \quad ; \; l = 1, 2, \ldots, K. \tag{A.11}$$

Imposing the two constraints, we have the following system of two linear equations in unknowns λ_0 and λ_1:

$$\sum_{l=1}^{K} q_l^\star = X\lambda_0 + Y\lambda_1 = 1 \tag{A.12}$$

and

$$\sum_{l=1}^{K} q_l^\star a_l^2 = Y\lambda_0 + Z\lambda_1 = E \tag{A.13}$$

where

$$X \triangleq \sum_{l=1}^{K} X_l \triangleq \sum_{l=1}^{K}\sum_{h=1}^{K} b'_{lh} \tag{A.14}$$

$$Y \triangleq \sum_{l=1}^{K} Y_l \triangleq \sum_{l=1}^{K}\sum_{h=1}^{K} b'_{lh} a_h^2 \tag{A.15}$$

and

$$Z = \sum_{l=1}^{K}\sum_{h=1}^{K} a_l^2 b'_{lh} a_h^2. \tag{A.16}$$

Solving (A.12) and (A.13) simultaneously, we have

$$\lambda_0 = \frac{Z}{XZ - Y^2} + \frac{-Y}{XZ - Y^2}E \tag{A.17}$$

and

$$\lambda_1 = \frac{-Y}{XZ - Y^2} + \frac{X}{XZ - Y^2}E \tag{A.18}$$

which yields

$$q_l^\star = \underbrace{\frac{X_l Z - Y_l Y}{XZ - Y^2}}_{\beta_{l0}} + \underbrace{\frac{-X_l Y + Y_l X}{XZ - Y^2}}_{\beta_{l1}} E \quad ; \; l = 1, 2, \ldots, K. \tag{A.19}$$

Observe that the constraint that $\{q_l^\star\}$ is a probability distribution implies an allowable range for E, i.e. assuming $0 \le q_l^\star \le 0.5$, E has to satisfy $E_{min} \le E \le E_{max}$ where

$$E_{min} = \max_{1 \le l \le K} \left[\frac{0.5 \cdot 1\{\beta_{l1} < 0\} - \beta_{l0}}{\beta_{l1}}\right], \tag{A.20}$$

$$E_{max} = \min_{1 \le l \le K} \left[\frac{0.5 \cdot 1\{\beta_{l1} > 0\} - \beta_{l0}}{\beta_{l1}}\right], \tag{A.21}$$

and $1\{.\}$ is the indicator function which takes the value 1 or 0 according to whether its argument is logically *true* or *false* respectively. Here, we assume that the code alphabets are restricted to be symmetric around the origin so that $0 \le q_l^\star \le 0.5$. Now, observe that $E_{min} \le E_{sat} \le E_{max}$; that is because E_{min} and E_{max} correspond to the cases of using only the lowest and highest energy code letters with non-zero probabilities, and obviously, $\{q_l^{sat}\}$ that yield E_{sat} is somewhere between the two extremes.

Finally, from (A.8), (A.17) and (A.18), we have the following result:

$$R_{0,ilc}(E) = -\log_2\left(\underbrace{\frac{X}{XZ - Y^2}}_{\alpha_2} E^2 + \underbrace{\frac{-2Y}{XZ - Y^2}}_{\alpha_1} E + \underbrace{\frac{Z}{XZ - Y^2}}_{\alpha_0}\right), \quad E_{min} \le E \le E_{sat}. \tag{A.22}$$

Now, consider a particular class of code alphabets consisting of K equispaced code letters symmetrically located around the origin with the distance between the adjacent letters equal to one. That is, consider the code alphabets of the form

$$\mathcal{A}_K = \{a_l : a_l = l - \frac{K+1}{2} \quad , \; l = 1, 2, \ldots, K\}. \tag{A.23}$$

Then, the optimum probability distribution $Q^\star$ is symmetric, i.e.,

$$q_l^\star = q_{K+1-l}^\star \quad , \; l = 1, 2, \ldots, K. \tag{A.24}$$

The results of optimization of $R_{0,ilc}$ for $\mathcal{A}_K$, $K = 4, 5, 6, 7$ and 8 are summarized in Tables A.1 and A.2. Observe that for $\mathcal{A}_4$, we have

$$E_{min} = 0.5 \cdot (-0.5)^2 + 0.5 \cdot (0.5)^2 = 0.25, \tag{A.25}$$

$$E_{max} = 0.5 \cdot (-1.5)^2 + 0.5 \cdot (1.5)^2 = 2.25, \tag{A.26}$$

and

$$q_1^\star = \beta_{10} + \beta_{11} E_{min} = 0, \tag{A.27}$$

$$q_1^\star = \beta_{10} + \beta_{11} E_{max} = 0.5 \tag{A.28}$$

imply $\beta_{10} = -0.0625$ and $\beta_{11} = 0.25$ regardless of the value of σ^2. However, this is not the case for larger K.

K	σ^2	E_{min}	E_{sat}	α_0	α_1	α_2
4	0.05	0.25	1.2928	0.6810	-0.6198	0.2397
4	0.10	0.25	1.4185	0.7807	-0.6029	0.2125
4	0.15	0.25	1.5404	0.8595	-0.5756	0.1868
4	0.20	0.25	1.6577	0.8926	-0.5405	0.1630
4	0.25	0.25	1.7751	0.9202	-0.5032	0.1417
4	0.30	0.25	1.8944	0.9387	-0.4672	0.1233
4	0.35	0.25	2.0158	0.9516	-0.4340	0.1077
4	0.40	0.25	2.1390	0.9609	-0.4040	0.0944
5	0.05	0.6093	2.0664	0.5305	-0.2947	0.0713
5	0.10	0.6480	2.2511	0.6383	-0.3106	0.0690
5	0.15	0.7023	2.4415	0.7113	-0.3085	0.0632
5	0.20	0.7651	2.6443	0.7572	-0.2948	0.0557
6	0.05	1.0711	3.0073	0.4375	-0.1651	0.0275
6	0.10	1.1384	3.2470	0.5419	-0.1833	0.0282
6	0.15	1.2077	3.4768	0.6170	-0.1905	0.0274
6	0.20	1.2749	3.7019	0.6694	-0.1904	0.0257
6	0.25	1.3421	3.9275	0.7077	-0.1865	0.0237
7	0.05	1.6340	4.1152	0.3732	-0.1023	0.0124
7	0.10	1.7346	4.4117	0.4707	-0.1175	0.0133
7	0.15	1.8404	4.6963	0.5423	-0.1249	0.0133
7	0.20	1.9529	4.9858	0.5931	-0.1267	0.0127
8	0.05	2.2976	5.3900	0.3258	-0.0679	0.0063
8	0.10	2.4308	5.7429	0.4162	-0.0799	0.0070
8	0.15	2.5622	6.0746	0.4840	-0.0867	0.0071
8	0.20	2.6915	6.4009	0.5339	-0.0897	0.0070
8	0.25	2.8214	6.7291	0.5722	-0.0904	0.0067

Table A.1: Results of Optimization of $R_{0,ilc}$ for $\mathcal{A}_4$ to $\mathcal{A}_8$.

For $E_{min} \leq E \leq E_{sat}$, the optimum cutoff rate for the independent-letters code ensemble is given as a function of E by

$$R^{\star}_{0,ilc}(E) = -\log_2(\alpha_2 E^2 + \alpha_1 E + \alpha_0) \text{ bits/ch.use.}$$

K	σ^2	β_{10}	β_{11}	β_{20}	β_{21}	β_{30}	β_{31}
4	$\star$	-0.0625	0.25	-	-	-	-
5	0.05	-0.0879	0.1442	0.3515	-0.0768	-	-
5	0.10	-0.0972	0.1500	0.3888	-0.1000	-	-
5	0.15	-0.1113	0.1585	0.4452	-0.1339	-	-
5	0.20	-0.1290	0.1686	0.5160	-0.1744	-	-
6	0.05	-0.0973	0.0909	0.2295	-0.0226	-	-
6	0.10	-0.1096	0.0963	0.2664	-0.0389	-	-
6	0.15	-0.1234	0.1022	0.3077	-0.0565	-	-
6	0.20	-0.1376	0.1079	0.3502	-0.0737	-	-
6	0.25	-0.1525	0.1136	0.3951	-0.0909	-	-
7	0.05	-0.0997	0.0610	0.1536	-0.0036	0.2826	-0.0347
7	0.10	-0.1141	0.0658	0.1902	-0.0157	0.2660	-0.0292
7	0.15	-0.1310	0.0712	0.2366	-0.0306	0.2328	-0.0185
7	0.20	-0.1510	0.0773	0.2950	-0.0485	0.1788	-0.0018
8	0.05	-0.0987	0.0429	0.1038	0.0033	0.2182	-0.0177
8	0.10	-0.1140	0.0469	0.1374	-0.0053	0.2093	-0.0154
8	0.15	-0.1311	0.0512	0.1781	-0.0155	0.1899	-0.0105
8	0.20	-0.1497	0.0556	0.2245	-0.0266	0.1621	-0.0038
8	0.25	-0.1700	0.0603	0.2773	-0.0387	0.1257	0.0046

Table A.2: Results of Optimization of $R_{0,ilc}$ for $\mathcal{A}_4$ to $\mathcal{A}_8$ continued.

And letter probabilities that optimize $R_{0,ilc}$ are given by

$$q_l^\star = \beta_{l1} E + \beta_{l0} \quad ; \; l = 1, 2, \ldots, K.$$

Appendix B

Computation of $R_{0,fcc}$

In this appendix, we discuss two enumeration algorithms used in computing $R_{0,fcc}(N, Q)$. The first of them enumerates all codewords in $\mathcal{F}_{N,Q}$ in a lexicographical order, whereas the second divides $\mathcal{F}_{N,Q}$ into subclasses of codewords at equal distances to a fixed reference codeword and enumerates these subclasses.

B.1 Algorithm 1: Enumeration in Lexicographical Order

Define a lexicographical order on the code letters so that

$$a_1 < a_2 < \cdots < a_K.$$

The elements of $\mathcal{F}_{N,Q}$ listed with respect to this lexicographical order start with

$$\underbrace{a_1 a_1 \ldots a_1}_{q_1 N} \underbrace{a_2 a_2 \ldots a_2}_{q_2 N} a_3 \ldots a_{K-1} \underbrace{a_K a_K \ldots a_K}_{q_K N}.$$

The following algorithm enumerates all codewords in the above order [PaW79, p.108]. Let $\mathbf{s} = (s_1, s_2, \ldots, s_N)$ be the current input to the algorithm.

1. Find the largest i such that $s_{i-1} < s_i$.
2. Find the largest j such that $s_{i-1} < s_j$.
3. Interchange s_{i-1} and s_j.
4. Reverse the order of the digits $s_i s_{i+1} \ldots s_N$.

Interchanging s_{i-1} and s_j yields a codeword that comes after $\mathbf{s}$ in the list, but not necessarily the immediate successor of $\mathbf{s}$. Despite this, the first codeword after $\mathbf{s}$ has to have s_j in $(i-1)$-st position; because, s_j is the smallest code symbol which is larger than s_{i-1} and lies to the right of s_{i-1}. This can be seen by observing that $s_i s_{i+1} \ldots s_N$ satisfy

$$s_i \geq s_{i+1} \geq \cdots \geq s_N$$

since i is the largest index such that $s_{i-1} < s_i$. Also after interchanging s_{i-1} and s_j, we have

$$s_i \geq s_{i+1} \geq \ldots \geq s_{j-1} \geq s_{i-1} \geq s_{j+1} \geq \ldots \geq s_N;$$

because j is the largest index such that $s_{i-1} < s_j$ and hence, $s_{i-1} \geq s_{j+1}$. Therefore, reversing the order of the digits from i to N in the fourth step yields the smallest possible ordering of these digits and hence, the immediate successor of s in the list. In Section B.3 a code implementing this algorithm is given. Unfortunately, this algorithm is not fast enough to run through huge ensembles. To overcome this difficulty, Algorithm 2, discussed next, takes advantage of the symmetries inherent in a fixed-composition code.

B.2 Algorithm 2: Enumerating Joint-Composition Classes

Let the fixed reference codeword $\mathbf{s}_r$ be the first codeword in the lexicographical order defined in the previous subsection. Then comparing any codeword $\mathbf{s}$ with $\mathbf{s}_r$, consider the *joint-composition matrix*, $W = [w_{ij}]_{i,j=1}^{K}$, with w_{ij}'s defined as

$$w_{ij} = \sum_{m=I_i+1}^{I_{i+1}} 1\{s_m = a_j\} \tag{B.1}$$

where

$$I_i = N \sum_{m=1}^{i-1} q_m. \tag{B.2}$$

In other words, w_{ij} is the number of a_j's in the subsequence $s_{I_i+1} s_{I_i+2} \ldots s_{I_{i+1}}$ of s which corresponds to the portion of $\mathbf{s}_r$ that is reserved for the code letter a_i. Observe that j-th column sum and i-th row sum of W are equal to $q_j N$ and $q_i N$ respectively. That is,

$$\sum_{i=1}^{K} w_{ij} = \sum_{i=1}^{K} \sum_{m=I_i+1}^{I_{i+1}} 1\{s_m = a_j\} = \sum_{m=1}^{N} 1\{s_m = a_j\} = q_j N, \tag{B.3}$$

$$\sum_{j=1}^{K} w_{ij} = \sum_{j=1}^{K} \sum_{m=I_i+1}^{I_{i+1}} 1\{s_m = a_j\} = \sum_{m=I_i+1}^{I_{i+1}} \underbrace{\sum_{j=1}^{K} 1\{s_m = a_j\}}_{1} = I_{i+1} - I_i = q_i N. \tag{B.4}$$

Also, observe that corresponding to any codeword $\mathbf{s} \in \mathcal{F}_{N,Q}$ there exists only one joint-composition matrix whereas many codewords correspond to the same joint-composition matrix and all codewords in the same joint-composition class are at the same euclidean distance to the reference codeword $\mathbf{s}_r$. Thus, one should expect some computational savings by enumerating all joint-composition matrices instead of all codewords in $\mathcal{F}_{N,Q}$. It is an easy task to show that

$$R_{0,fcc}(N,Q) = -\frac{1}{N}\log_2 \sum_{W} \frac{|W|}{|\mathcal{F}_{N,Q}|} \cdot e^{-d_W^2(\mathbf{s}_r)/8\sigma^2} \tag{B.5}$$

where the summation is over all joint-composition matrices and $d_W(\mathbf{s}_r)$ is the euclidean distance of any one of $|W|$ codewords in the joint-composition class represented by the matrix W to the reference codeword $\mathbf{s}_r$, and

$$|W| = \prod_{i=1}^{K} \frac{(q_i N)!}{\prod_{j=1}^{K} w_{ij}!}. \tag{B.6}$$

So, one has to enumerate at most

$$\frac{|\mathcal{F}_{N,Q}|}{\min_W\{|W|\}}$$

joint-composition matrices, which certainly indicates a computational saving. An implementation of this idea of enumerating joint-composition matrices is given in Section B.3. It can be extended so as to compute $R_{0,fcc}$ for code alphabets of larger size.

B.3 Codes Implementing Algorithms 1 and 2

`cw` is the global array of length `N` that represents the current input codeword to the algorithm. The code finds the immediate successor of `cw`. If the input codeword is the last one in the list, then it will remain unchanged.

```
/* Implementation of Algorithm 1 */
void find_next_codeword()
{
    int i,j,m,x;

    for (i = N-1; i > 0; --i)
    {
         if (cw[i] > cw[i-1])
         {
              x = cw[i-1];
              for (j = N-1; j >= i; --j)
              {
                   if (cw[j] > x)
                   {
                        cw[i-1] = cw[j];
                        cw[j] = x;
                        for (m = 0; m <= (N-i-1)/2; ++m)
                        {
                             x = cw[N-1-m];
                             cw[N-1-m] = cw[i+m];
                             cw[i+m] = x;
                        }
                        break;
                   }
              }
              break;
         }
    }
}
```

```
/* Implementation of Algorithm 2 */
#include <math.h>
#include <stdio.h>

#define max3(a,b,c) (((a>b)?a:b)>c)?((a>b)?a:b):c
#define min(a,b) (a<b)?a:b

int n1,n2,n3,n4;            /* symbol frequencies */
int v11,v12,v13,v14;        /* elements of the */
int v21,v22,v23,v24;        /* joint-composition */
int v31,v32,v33,v34;        /* matrix */
int v41,v42,v43,v44;
int block_length;
double no_of_cws;
double N_0;
double R_0;

double fact(n)
int n;
{
    int i;
    double r = 1.0;

    if (n <= 1)
        return(r);
    else
    {
        for (i = 2; i <= n; ++i)
            r *= i;
        return(r);
    }
}

double no_of_cws_in_class()
{
    double x;

    x = fact(n1) * fact(n2) * fact(n3) * fact(n4);
    x = x / (fact(v11) * fact(v12) * fact(v13) * fact(v14));
    x = x / (fact(v21) * fact(v22) * fact(v23) * fact(v24));
```

```
/* Implementation of Algorithm 2 */
#include <math.h>
#include <stdio.h>

#define max3(a,b,c) (((a>b)?a:b)>c)?((a>b)?a:b):c
#define min(a,b) (a<b)?a:b

int n1,n2,n3,n4;          /* symbol frequencies */
int v11,v12,v13,v14;      /* elements of the */
int v21,v22,v23,v24;      /* joint-composition */
int v31,v32,v33,v34;      /* matrix */
int v41,v42,v43,v44;
int block_length;
double no_of_cws;
double N_0;
double R_0;

double fact(n)
int n;
{
    int i;
    double r = 1.0;

    if (n <= 1)
        return(r);
    else
    {
        for (i = 2; i <= n; ++i)
            r *= i;
        return(r);
    }
}

double no_of_cws_in_class()
{
    double x;

    x = fact(n1) * fact(n2) * fact(n3) * fact(n4);
    x = x / (fact(v11) * fact(v12) * fact(v13) * fact(v14));
    x = x / (fact(v21) * fact(v22) * fact(v23) * fact(v24));
```

```
    x = x / (fact(v31) * fact(v32) * fact(v33) * fact(v34));
    x = x / (fact(v41) * fact(v42) * fact(v43) * fact(v44));
    return(x);
}

main()
{
    int a11,b11,a12,b12,a21,b21,a22,b22;
    double sum;
    double exponent,nocwic;

    N_0 = 0.4;
    n1 = n4 = 2;
    n2 = n3 = 3;
    block_length = n1 + n2 + n3 + n4;
    no_of_cws = fact(block_length) / (fact(n1)*fact(n2)*fact(n3)
                                                        *fact(n4));

    sum = 0.0;
    for (w11 = 0; w11 <= n1; ++w11)
    for (w12 = 0; w12 <= n1-w11; ++w12)
    for (w21 = 0; w21 <= n1-w11; ++w21)
    for (w13 = 0; w13 <= n1-(w11+w12); ++w13)
    for (w31 = 0; w31 <= n1-(w11+w21); ++w31)
    {
         w14 = n1-(w11+w12+w13);
         w41 = n1-(w11+w21+w31);
         a11 = max3(0,n2-w21-(n3-w13+n4-w14),n2-w12-(n3-w31+n4
                                                         -w41));
         b11 = min(n2-w21,n2-w12);
         for (w22 = a11; w22 <= b11; ++w22)
         {
         a12 = max3(0,n2-(w21+w22)-(n4-w14),n3-w13-(n3-w31+n4
                                                        -w41));
         b12 = min(n2-(w21+w22),n3-w13);
         for (w23 = a12; w23 <= b12; ++w23)
         {
         a21 = max3(0,n3-w31-(n3-(w13+w23))-(n4-w14),n2-(w12
                                            +w22)-(n4-w41));
         b21 = min(n3-w31,n2-(w12+w22));
```

```
        for (w32 = a21; w32 <= b21; ++w32)
        {
        w24 = n2-w21-w22-w23;
        w42 = n2-w12-w22-w32;
        a22 = max3(0,n3-(w31+w32)-(n4-(w14+w24)),n3-(w13+w23)
                                             -(n4-(w41+w42)));
        b22 = min(n3-(w31+w32),n3-(w13+w23));
        for (w33 = a22; w33 <= b22; ++w33)
        {
             w34 = n3-(w31+w32+w33);
             w43 = n3-(w13+w23+w33);
             w44 = n4-(w14+w24+w34);

             nocwic = no_of_cws_in_class();
             exponent = w12+w21+w23+w32+w34+w43+4.0*(w13+w31
                                  +w24+w42)+ 9.0*(w14+w41);
             sum += nocwic * exp(-exponent / (4 * N_0));
        }
        }
        }
        }
        }
        R_0 = -log(sum / no_of_cws) / (block_length * log(2.0));
}
```

N	T_1	T_2	$\|\{W\}\|$
10	30 sec	1 sec	626
20	-	19 sec	28469
30	-	5 min 21 sec	412460
40	-	48 min 33 sec	3228457

Table B.1: Time complexities of Algorithms 1 and 2.

The time complexities of the two algorithms discussed above are compared in Table B.1 for $(N, \{0.2, 0.3, 0.3, 0.2\})$-composition ensemble. T_1 and T_2 denote the run-times of the two programs that compute $R_{0,fcc}$ by using Algorithms 1 and 2 respectively, and $|\{W\}|$ denotes the number of distinct joint-composition classes in $\mathcal{F}_{N,Q}$. One can see the significant computational savings by comparing the number of distinct joint-composition classes with the $|\mathcal{F}_{N,Q}|$ values in Table 2.1.

B.4 Numerical Results of $R_{0,fcc}$ Computations

Although $R_{0,fcc}(40, Q^*)$ values with a precision of six significant digits are given in Tables 2.2-2.9, in this section, we tabulate all $R_{0,fcc}(N, Q^*)$ data for $N \leq 40$ with a precision of nine significant digits so as to summarize all $R_{0,fcc}$ computations carried out in this thesis work. Tables B.2-B.6 also include $R^*_{0,ilc}$ data for the sake of immediate comparison.

σ^2	E	$R^*_{0,ilc}$	$R_{0,fcc}/N$
0.05	0.35	1.019047432	1.074402093/40
	0.45	1.149965011	1.164629119/20
			1.234434911/40
	0.55	1.277097450	1.362301899/40
	0.65	1.398177755	1.239642421/10
			1.381937417/20
			1.437533477/30
			1.466158048/40
	0.75	1.510515202	1.248391546/ 8
			1.420158914/16
			1.490509220/24
			1.527778842/32
			1.550316832/40
	0.85	1.611057941	1.525034620/20
			1.617367432/40
	0.95	1.696540129	1.668961268/40
	1.05	1.763726129	1.442304898/10
			1.608509997/20
			1.673136883/30
			1.706167864/40
	1.15	1.809738485	1.729654586/40
	1.25	1.832419971	1.139006247/ 4
			1.398111503/ 8
			1.520190130/12
			1.591739429/16
			1.638602131/20
			1.671463505/24
			1.695609341/28
			1.713973933/32
			1.728322480/36
			1.739781356/40
0.10	0.35	0.747307416	0.803547950/40
	0.45	0.856145947	0.918123041/20

σ^2	E	$R^*_{0,ilc}$	$R_{0,fcc}/N$
0.10	0.45	0.856145947	0.940396283/40
	0.55	0.961877244	1.053850651/40
	0.65	1.063105786	1.054909688/10
			1.120844952/20
			1.140319953/30
			1.149289105/40
	0.75	1.158226836	1.090187475/ 8
			1.181589273/16
			1.209471304/24
			1.222280721/32
			1.229599424/40
	0.85	1.245453701	1.263279162/20
			1.296544070/40
	0.95	1.322874600	1.351261740/40
	1.05	1.388543689	1.271224065/10
			1.356856598/20
			1.382429259/30
			1.394494739/40
	1.15	1.440605117	1.426703877/40
	1.25	1.477440354	1.064231357/ 4
			1.265595392/ 8
			1.344304411/12
			1.383507145/16
			1.406137134/20
			1.420643403/24
			1.430686457/28
			1.438047484/32
			1.443678135/36
			1.448127949/40
	1.35	1.497819503	1.458810605/40
0.15	0.35	0.575720551	0.618477354/40
	0.45	0.670383731	0.727043639/20
			0.737019353/40

Table B.2: Numerical results of $R_{0,fcc}$ computations.

σ^2	E	$R^{\star}_{0,ilc}$	$R_{0,fcc}/N$
0.15	0.55	0.762523089	0.838728166/40
	0.65	0.851193312	0.877116024/10
			0.912086273/20
			0.922058198/30
			0.926770330/40
	0.75	0.935326622	0.925283759/ 8
			0.977133811/16
			0.991976320/24
			0.998931952/32
			1.002984224/40
	0.85	1.013746032	1.049760284/20
			1.068544145/40
	0.95	1.085191582	1.124230574/40
	1.05	1.148361673	1.097191536/10
			1.147926657/20
			1.163174770/30
			1.170561117/40
	1.15	1.201969901	1.207859632/40
	1.25	1.244815142	0.969490581/ 4
			1.119925139/ 8
			1.171538701/12
			1.195852549/16
			1.209771620/20
			1.218785704/24
			1.225110604/28
			1.229799383/32
			1.233416616/36
			1.236292975/40
	1.35	1.275859349	1.255887704/40
	1.45	1.294304287	1.168916321/10
			1.235734001/20
			1.256455480/30
			1.266531410/40

σ^2	E	$R^{\star}_{0,ilc}$	$R_{0,fcc}/N$
0.20	0.35	0.467203935	0.500027999/40
	0.45	0.551394827	0.598576967/20
			0.604259559/40
	0.55	0.633489363	0.695999383/40
	0.65	0.712808548	0.746179819/10
			0.767857497/20
			0.774143148/30
			0.777155774/40
	0.75	0.788595405	0.798637219/ 8
			0.832065429/16
			0.841665412/24
			0.846236021/32
			0.848916894/40
	0.85	0.860022161	0.899387490/20
			0.912073506/40
	0.95	0.926203917	0.967164946/40
	1.05	0.986219829	0.963553178/10
			0.998549618/20
			1.009304606/30
			1.014554582/40
	1.15	1.039142273	1.054471317/40
	1.25	1.084073463	0.886484928/ 4
			1.002938218/ 8
			1.040131211/12
			1.057529281/16
			1.067589111/20
			1.074161012/24
			1.078797170/28
			1.082244991/32
			1.084910115/36
			1.087032149/40
	1.35	1.120187751	1.112252372/40
	1.45	1.146776429	1.056958091/10

Table B.3: Numerical results of $R_{0,fcc}$ computations continued.

σ^2	E	$R^\star_{0,ilc}$	$R_{0,fcc}/N$	σ^2	E	$R^\star_{0,ilc}$	$R_{0,fcc}/N$
0.20	1.45	1.146776429	1.106779879/20	0.25	1.25	0.963893557	0.967282673/28
			1.122390341/30				0.969960610/32
			1.130045936/40				0.972031344/36
	1.55	1.163290663	1.140215720/40				0.973680440/40
	1.65	1.169377567	1.114459671/20		1.35	1.002300664	1.002457564/40
			1.142430324/40		1.45	1.033363115	0.967276417/10
0.25	0.35	0.393179776	0.418941694/40				1.006560544/20
	0.45	0.468993614	0.507834979/20				1.018955892/30
			0.511526853/40				1.025047260/40
	0.55	0.543042316	0.594623821/40		1.55	1.056577378	1.041303237/40
	0.65	0.614821813	0.648158197/10		1.65	1.071550289	1.028327482/20
			0.662938851/20				1.050983171/40
			0.667302804/30		1.75	1.078020353	0.948684398/ 8
			0.669407538/40				1.015412852/16
	0.75	0.683775999	0.701264557/ 8				1.036985610/24
			0.724667530/16				1.047497063/32
			0.731471922/24				1.053719246/40
			0.734734911/32	0.30	0.35	0.339532702	0.360185665/40
			0.736653620/40		0.45	0.408434500	0.440587558/20
	0.85	0.749300889	0.787651917/20				0.443184964/40
			0.796897179/40		0.55	0.475835857	0.518818751/40
	0.95	0.810752243	0.850511755/40		0.65	0.541353518	0.572348875/10
	1.05	0.867457204	0.859580359/10				0.583070402/20
			0.885675078/20				0.586277668/30
			0.893785603/30				0.587829384/40
			0.897752623/40		0.75	0.604568290	0.624459385/ 8
	1.15	0.918730275	0.938781538/40				0.641760884/16
	1.25	0.963893557	0.816107155/ 4				0.646844683/24
			0.908979172/ 8				0.649290728/32
			0.937453132/12				0.650730627/40
			0.950816393/16		0.85	0.665027547	0.700846413/20
			0.958590107/20				0.707887875/40
			0.963684128/24		0.95	0.722249701	0.759561370/40

Table B.4: Numerical results of $R_{0,fcc}$ computations continued.

σ^2	E	$R^\star_{0,ilc}$	$R_{0,fcc}/N$
0.30	1.05	0.775730916	0.776152693/10
			0.796479014/20
			0.802825904/30
			0.805931180/40
	1.15	0.824954297	0.847111968/40
	1.25	0.869401567	0.755850769/ 4
			0.831568973/ 8
			0.854208412/12
			0.864869359/16
			0.871085771/20
			0.875162055/24
			0.878041942/28
			0.880184950/32
			0.881841882/36
			0.883161294/40
	1.35	0.908567060	0.914083259/40
	1.45	0.941973553	0.892818922/10
			0.924758914/20
			0.934863380/30
			0.939828117/40
	1.55	0.969189184	0.960287651/40
	1.65	0.989844385	0.956548207/20
			0.975284996/40
	1.75	1.003647616	0.896220898/ 8
			0.952706909/16
			0.970641998/24
			0.979380328/32
			0.984555786/40
	1.85	1.010398565	0.913902764/10
			0.963921901/20
			0.979905507/30
			0.987713095/40
0.35	0.35	0.298858491	0.315732969/40

σ^2	E	$R^\star_{0,ilc}$	$R_{0,fcc}/N$
0.35	0.45	0.361958759	0.388846374/20
			0.390774359/40
	0.55	0.423772903	0.459979235/40
	0.65	0.484003690	0.512065646/10
			0.520189588/20
			0.522641762/30
			0.523830279/40
	0.75	0.542328342	0.562426255/ 8
			0.575721707/16
			0.579658971/24
			0.581556607/32
			0.582674283/40
	0.85	0.598400106	0.631233512/20
			0.636764147/40
	0.95	0.651850976	0.686283062/40
	1.05	0.702295735	0.707485722/10
			0.723773935/20
			0.728867621/30
			0.731359128/40
	1.15	0.749337451	0.772074252/40
	1.25	0.792574473	0.703579674/ 4
			0.766360029/ 8
			0.784818535/12
			0.793526383/16
			0.798606495/20
			0.801937057/24
			0.804289422/28
			0.806039427/32
			0.807392199/36
			0.808469211/40
	1.35	0.831608884	0.840545831/40
	1.45	0.866056228	0.829307574/10
			0.855780714/20

Table B.5: Numerical results of $R_{0,fcc}$ computations continued.

σ^2	E	$R^{\star}_{0,ilc}$	$R_{0,fcc}/N$
0.35	1.45	0.866056228	0.864155238/30
			0.868266598/40
	1.55	0.895556165	0.891551523/40
	1.65	0.919783600	0.894579958/20
			0.910271497/40
	1.75	0.938459645	0.849412698/ 8
			0.897470893/16
			0.912546644/24
			0.919889179/32
			0.924236314/40
	1.85	0.951361727	0.870559112/10
			0.913161762/20
			0.926606429/30
			0.933173273/40
	1.95	0.958332120	0.936686174/40
0.40	0.35	0.266945353	0.280963785/40
	0.45	0.325113759	0.347848083/20
			0.349335875/40
	0.55	0.382170968	0.412994978/40
	0.65	0.437882425	0.463034102/10
			0.469394785/20
			0.471326719/30
			0.472264150/40
	0.75	0.491994971	0.511326278/ 8
			0.521847047/16
			0.524980276/24
			0.526491864/32
			0.527382439/40
	0.85	0.544237849	0.574079384/20
			0.578527764/40
	0.95	0.594324421	0.625830822/40
	1.05	0.641954689	0.649851996/10
			0.663177939/20

σ^2	E	$R^{\star}_{0,ilc}$	$R_{0,fcc}/N$
0.40	1.05	0.641954689	0.667345921/30
			0.669383822/40
	1.15	0.686818698	0.709245919/40
	1.25	0.728600875	0.657726084/ 4
			0.710485519/ 8
			0.725810812/12
			0.733044456/16
			0.737263280/20
			0.740027913/24
			0.741979816/28
			0.743431468/32
			0.744553343/36
			0.745446356/40
	1.35	0.766985289	0.777985777/10
	1.45	0.801661772	0.774114452/10
			0.796360601/20
			0.803388766/30
			0.806835902/40
	1.55	0.832332748	0.831937429/40
	1.65	0.858720547	0.839929133/20
			0.853195706/40
	1.75	0.880574911	0.806714518/ 8
			0.847791937/16
			0.860573072/24
			0.866793029/32
			0.870473153/40
	1.85	0.897680297	0.830265918/10
			0.866625560/20
			0.878015520/30
			0.883576208/40
	1.95	0.909862598	0.892231742/40
	2.05	0.916994853	0.873400932/20
			0.896038891/40

Table B.6: Numerical results of $R_{0,fcc}$ computations continued.

Appendix C

Asymptotic Analysis of $R_{0,fcc}$

Consider the K-letter symmetric code alphabet $\mathcal{A}_K$ defined by (A.23) and the associated probability distribution $Q^\star$ which maximizes the cutoff rate for the ensemble of independent-letters codes over $\mathcal{A}_K$. The computation of cutoff rates for the ensemble of $(N, Q^\star)$-composition codes was discussed in Appendix B, and the results were given for the particular code alphabet $\mathcal{A}_4$ and for $N = 40$ in Chapter 2. Unfortunately, it was not possible to go beyond blocklengths of 40 and alphabet sizes of 4 due to the exponentially increasing complexity of the problem. Therefore, one may wonder which values $R_{0,fcc}(N, Q^\star)$ would take as N tends to infinity. Here, having fixed the code alphabet and hence determined the optimal probability distribution $Q^\star$, we suppress $Q^\star$ to simplify the notation. We define the asymptotic value of $R_{0,fcc}$ by

$$R_{0,fcc}(\infty) = \lim_{N\to\infty} R_{0,fcc}(N). \tag{C.1}$$

In the following section, the computation of $R_{0,fcc}(\infty)$ is discussed and the error term $\varepsilon_N = R_{0,fcc}(\infty) - R_{0,fcc}(N)$ is analyzed in Section C.2.

C.1 Computation of $R_{0,fcc}(\infty)$

From (B.5), we have

$$e^{-NR_{0,fcc}} = \sum_V \frac{|V|}{|\mathcal{F}_{N,Q^\star}|} e^{-d_V^2(\mathbf{s}_r)/8\sigma^2}. \tag{C.2}$$

where V is the normalized version of the joint-composition matrix W and is defined by

$$w_{ij} \longrightarrow v_{ij} \triangleq \frac{w_{ij}}{q_i^\star N} \quad ; \; i,j = 1,2,\ldots,K. \tag{C.3}$$

Then, from (B.3) and (B.4), we have

$$\sum_{j=1}^{K} v_{ij} = 1, \tag{C.4}$$

$$\sum_{i=1}^{K} q_i^\star v_{ij} = q_j^\star. \tag{C.5}$$

Observe that v_{ij}'s can be regarded as transition probabilities on a discrete channel with K input and K output letters; that is, v_{ij} can be regarded as the probability of receiving letter j at the channel output given that letter i is sent.

Using the Stirling formula for the factorials, we can approximate $|\mathcal{F}_{N,Q^\star}|$ and $|V|$ for large N as

$$|\mathcal{F}_{N,Q^\star}| \approx e^{NH(Q^\star)} \tag{C.6}$$

and

$$|V| \approx e^{NH(V|Q^\star)} \tag{C.7}$$

where $H(Q^\star)$ and $H(V|Q^\star)$ are unconditional and conditional entropy functions given by

$$H(Q^\star) = -\sum_{i=1}^{K} q_i^\star \ln q_i^\star \tag{C.8}$$

and

$$H(V|Q^\star) = -\sum_{i=1}^{K}\sum_{j=1}^{K} q_i^\star v_{ij} \ln v_{ij}. \tag{C.9}$$

Leaving the verification of this result to be discussed in Section C.2, observe that we have on the other hand

$$\frac{d_V^2(\mathbf{s}_r)}{8\sigma^2} = \sum_{i=1}^{K}\sum_{j=1}^{K} v_{ij}\frac{d_{ij}^2}{8\sigma^2} = N\sum_{i=1}^{K}\sum_{j=1}^{K} q_i^\star v_{ij}\frac{d_{ij}^2}{8\sigma^2} = NE(d^2/8\sigma^2). \tag{C.10}$$

Therefore, combining (C.6), (C.7) and (C.10), (C.2) reduces to

$$e^{-NR_{0,fcc}} \approx \sum_{V} e^{-Nf(V)} \tag{C.11}$$

where

$$f(V) = H(Q^\star) - H(V|Q^\star) + E(d^2/8\sigma^2), \tag{C.12}$$

$$= \sum_{i=1}^{K} q_i^\star \left[\ln\frac{1}{q_i^\star} + \sum_{j=1}^{K} v_{ij}\ln\left(v_{ij}e^{d_{ij}^2/8\sigma^2}\right)\right]. \tag{C.13}$$

Form (C.4), it follows that

$$f(V) = \sum_{i=1}^{K} q_i^\star \sum_{j=1}^{K} v_{ij}\ln\left(\frac{v_{ij}e^{d_{ij}^2/8\sigma^2}}{q_j^\star}\right). \tag{C.14}$$

Notice that $f(V)$ is convex cup in v_{ij}'s; because,

$$\frac{\partial^2 f}{\partial (v_{ij})^2} = \frac{q_i^\star}{v_{ij}} > 0 \quad ; \; i,j = 1,2,\ldots,K. \tag{C.15}$$

So, it follows from (C.11) that only one term in the summation becomes dominant as N gets large. This term is the one that corresponds to the joint-composition matrix $V_N^\star$ which minimizes $f(V)$ over all joint-composition matrices of $\mathcal{F}_{N,Q^\star}$. Observe that the entries of $V_N^\star$ are multiples of $1/q_i^\star N$. That is, for large N, the solution $V_N^\star \in \mathcal{Q}^{K\times K}$; but, as N tends to infinity the solution becomes more and more likely to be in $\mathcal{R}^{K\times K}$. Since $R_{0,fcc}(\infty)$ is the quantity of interest, the solution $V^\star(=V_\infty^\star) \in \mathcal{R}^{K\times K}$. Hence, to find $R_{0,fcc}(\infty)$ we have to minimize $f(V)$ over the set of normalized joint-composition matrices $V \in \mathcal{R}^{K\times K}$ subject to the constraints (C.4), (C.5) and $v_{ij} \geq 0$. Then, $R_{0,fcc}(\infty)$ is given by

$$R_{0,fcc}(\infty) = \min_V\{f(V)\} = f(V^\star). \tag{C.16}$$

Minimization of $f(V)$: Let $\{\lambda_{0i}\}$ and $\{\lambda_{1i}\}$, $i = 1, 2, \ldots, K$, be two sets of Lagrange multipliers and define

$$F(V, \lambda_{0i}, \lambda_{1j}) = f(V) - \lambda_{0i}\sum_{j=1}^{K} v_{ij} - \lambda_{1j}\sum_{i=1}^{K} q_i^\star v_{ij} \quad ; \; i,j = 1,2,\ldots,K. \tag{C.17}$$

Then, taking partial derivatives of F with respect to v_{ij}'s and equating these to zero we have

$$\frac{\partial F}{\partial v_{ij}} = q_i^\star \ln\left(\frac{v_{ij}e^{d_{ij}^2/8\sigma^2}}{q_i^\star}\right) + q_i^\star - \lambda_{0i} - \lambda_{1j}q_i^\star = 0 \quad ; \; i,j = 1,2,\ldots,K, \tag{C.18}$$

which implies

$$v_{ij} = q_i^\star e^{-d_{ij}^2/8\sigma^2} e^{(\lambda_{0i}+\lambda_{1j}q_i^\star - q_i^\star)/q_i^\star}. \tag{C.19}$$

Let $\mu_i = e^{\lambda_{0i}/q_i^\star}$ and $\nu_j = e^{\lambda_{1j}}$, then

$$v_{ij} = q_i^\star e^{-(1+d_{ij}^2/8\sigma^2)}\mu_i\nu_j. \tag{C.20}$$

Imposing the constraints we have

$$\sum_{j=1}^{K} v_{ij} = \sum_{j=1}^{K} q_i^\star e^{-(1+d_{ij}^2/8\sigma^2)}\mu_i\nu_j = 1 \tag{C.21}$$

and

$$\sum_{i=1}^{K} q_i^\star v_{ij} = \sum_{i=1}^{K} q_i^{\star 2} e^{-(1+d_{ij}^2/8\sigma^2)}\mu_i\nu_j = q_j^\star. \tag{C.22}$$

(C.21) implies

$$\mu_i = \left(q_i^\star \sum_{j=1}^{K} \nu_j e^{-(1+d_{ij}^2/8\sigma^2)}\right)^{-1} \tag{C.23}$$

Finally, combining (C.22) and (C.23) we have

$$q_j^\star = \sum_{i=1}^{K} q_i^\star \left(\frac{\nu_j e^{-d_{ij}^2/8\sigma^2}}{\sum_{k=1}^{K}\nu_k e^{-d_{ik}^2/8\sigma^2}}\right) \quad ; \; j = 1,2,\ldots,K. \tag{C.24}$$

Solving this equation for ν_j iteratively using

$$\nu_j^{(n+1)} = \frac{q_j^\star}{\sum_{i=1}^K \left[q_i^\star e^{-d_{ij}^2/8\sigma^2} \left(\sum_{k=1}^K \nu_k^{(n)} e^{-d_{ik}^2/8\sigma^2} \right)^{-1} \right]} \quad ; \; j = 1, 2, \ldots, K, \tag{C.25}$$

we have the solution. (Observe that regardless of the initial values of ν_j's the iteration converges to the same solution up to a scaling factor.) Having found ν_j's and hence μ_i's, we have the joint-composition matrix $V^\star$ that minimizes $f(V)$ and the solution for $R_{0,fcc}(\infty)$ follows from (C.14) and (C.20)

$$R_{0,fcc}(\infty) = \frac{1}{e \ln 2} \left[\sum_{i=1}^K q_i^{\star 2} \mu_i \sum_{j=1}^K \nu_j e^{-d_{ij}^2/8\sigma^2} \ln \left(\frac{\mu_i \nu_j}{e} \right) \right] \text{ bits/ch.use.} \tag{C.26}$$

C.2 Analysis of the Error Term

Having discussed the computation of $R_{0,fcc}(\infty)$, in this section the behavior of the error term $\varepsilon_N = R_{0,fcc}(\infty) - R_{0,fcc}(N)$ is analyzed. Observe that in approximating $R_{0,fcc}(N)$ for large N there are two types of errors-one originating from the approximations made for $|\mathcal{F}_{N,Q^\star}|$ and $|V|$, and the other originating from approximating the summation over all joint-composition matrices V with a single dominant term. Both of these errors are discussed in the following subsections.

C.2.1 Error due to Approximating $|\mathcal{F}_{N,Q^\star}|$ and $|V|$

First recall the Stirling formula for the factorials [Gal68, p.530].

$$n! = \sqrt{2\pi n} \left(\frac{n}{e} \right)^n e^{\epsilon_n} \tag{C.27}$$

where ϵ_n is decreasing with n and satisfies $0 < \epsilon_n < 1/12n$. Therefore, from (2.7) we have

$$|\mathcal{F}_{N,Q^\star}| = \frac{N!}{\prod_{i=1}^K (q_i^\star N)!} = \frac{\sqrt{2\pi N}(N/e)^N e^{\epsilon_N}}{\prod_{i=1}^K \sqrt{2\pi q_i^\star N} (q_i^\star N/e)^{q_i^\star N} e^{\epsilon_{q_i^\star N}}} \tag{C.28}$$

$$= \left((2\pi N)^{K-1} \prod_{i=1}^K q_i^\star \right)^{-1/2} \cdot \underbrace{\left(\prod_{i=1}^K (q_i^\star)^{-q_i^\star N} \right)}_{e^{NH(Q^\star)}} \cdot e^{\epsilon_N - \sum_{i=1}^K \epsilon_{q_i^\star N}} \tag{C.29}$$

$$= e^{N[H(Q^\star) + \Delta_1(N, Q^\star)]} \tag{C.30}$$

where

$$\Delta_1(N, Q^\star) = \frac{1}{N} \left[\epsilon_N - \sum_{i=1}^K \epsilon_{q_i^\star N} - \frac{K-1}{2} \ln(2\pi N) - \frac{1}{2} \ln \left(\prod_{i=1}^K q_i^\star \right) \right]. \tag{C.31}$$

Observe that $\exists M \in \mathcal{R}^+$ such that $|\Delta_1(N,Q^\star)| < M \ln N/N$, in other words, $\Delta_1(N,Q^\star) = O(\ln N/N)$ [DeB81]; because, $(\epsilon_N - \sum_{i=1}^{K} \epsilon_{q_i^\star N})/N$ goes to zero with $1/N^2$ and is dominated by $\ln N/N$ term. Hence, it follows that

$$|\mathcal{F}_{N,Q^\star}| = e^{N[H(Q^\star)+O(\ln N/N)]} \tag{C.32}$$

$$\approx e^{NH(Q^\star)} \text{ for large } N. \tag{C.33}$$

Similarly, from (B.6) we can show that

$$|V| = e^{N[H(V|Q^\star)+\Delta_2(N,Q^\star,V)]} \tag{C.34}$$

where

$$\Delta_2(N,Q^\star,V) = \frac{1}{N}\left[\sum_{i=1}^{K} \epsilon_{q_i^\star N} - \sum_{i=1}^{K}\sum_{j=1}^{K} \epsilon_{v_{ij}q_i^\star N} - \frac{K(K-1)}{2}\ln(2\pi N) - \frac{K-1}{2}\ln\left(\prod_{i=1}^{K} q_i^\star\right) - \frac{1}{2}\ln\left(\prod_{i=1}^{K}\prod_{j=1}^{K} v_{ij}\right)\right] \tag{C.35}$$

$$= O(\frac{\ln N}{N}). \tag{C.36}$$

Hence, it follows that

$$|V| = e^{N[H(V|Q^\star)+O(\ln N/N)]} \tag{C.37}$$

$$\approx e^{NH(V|Q^\star)} \text{ for large } N. \tag{C.38}$$

Combining these results, we have

$$\frac{|V|}{|\mathcal{F}_{N,Q^\star}|} = e^{-N[H(Q^\star)-H(V|Q^\star)+\Delta(N,Q^\star,V)]} \tag{C.39}$$

where

$$\Delta(N,Q^\star,V) = \Delta_1(N,Q^\star) - \Delta_2(N,Q^\star,V) \tag{C.40}$$

$$= \frac{1}{N}\left[\epsilon_N - 2\sum_{i=1}^{K} \epsilon_{q_i^\star N} + \sum_{i=1}^{K}\sum_{j=1}^{K} \epsilon_{v_{ij}q_i^\star N} + \frac{(K-1)^2}{2}\ln(2\pi N) + \frac{K-2}{2}\ln\left(\prod_{i=1}^{K} q_i^\star\right) + \frac{1}{2}\ln\left(\prod_{i=1}^{K}\prod_{j=1}^{K} v_{ij}\right)\right] \tag{C.41}$$

$$= O(\frac{\ln N}{N}), \tag{C.42}$$

and (C.11) follows. Here, one point to note is that products including v_{ij} factors are over non-zero v_{ij}'s since 0! is defined to be 1.

C.2.2 Error due to Approximating the Summation by the Dominant Term

Observe that to be exact we have to consider

$$e^{-NR_{0,fcc}} = \sum_{V} e^{-N[f(V)+\Delta(N,Q^\star,V)]}; \tag{C.43}$$

but, since $\Delta(N, Q^\star, V)$ term is shown to be $O(\ln N/N)$ we can neglect it and thus treat the rest of the problem as in Section C.1. Therefore, forgetting about $\Delta(N, Q^\star, V)$ term the following very rough bounds on $e^{-NR_{0,fcc}(N)}$ for large N follow from (C.11)

$$e^{-Nf(V^\star)} < e^{-NR_{0,fcc}(N)} < (N+1)^{K^2} e^{-Nf(V^\star)} = e^{-N[f(V^\star)-K^2\ln(N+1)/N]} \qquad (C.44)$$

which imply

$$R_{0,fcc}(\infty) > R_{0,fcc}(N) > R_{0,fcc}(\infty) - \frac{K^2}{N}\ln(N+1). \qquad (C.45)$$

Therefore, it follows for large N that

$$0 < \varepsilon_N = R_{0,fcc}(\infty) - R_{0,fcc}(N) = O\left(\frac{\ln N}{N}\right) \qquad (C.46)$$

which is same as saying that the error term goes to zero with $\ln N/N$. Observe that even the rough bound of $(N+1)^{K^2}$ on the number of distinct joint-composition matrices leads to an $O(\ln N/N)$ term which is of the same order with the error term due to approximating $|\mathcal{F}_{N,Q^\star}|$ and $|V|$. Therefore, we can conclude that there is no need to use a better estimate for the number of distinct joint-composition matrices; because otherwise, there will still be an $O(\ln N/N)$ term remaining even if the better estimate leads to an error of order less than $\ln N/N$.

References

[Arı89] Arıkan, E., 'Trellis coding for high signal-to-noise ratio gaussian noise channels,' *Conference Record, IEEE Military Communication Conference*, vol. 1, pp.196-199, Boston, Massachusetts, Oct. 1989.

[DeB81] De Bruijn, N.G., *Asymptotic Methods in Analysis.* New York: Dover Publications, Inc., 1981.

[Gal65] Gallager, R.G., 'A simple derivation of the coding theorem and some applications,' *IEEE Trans. Inform. Theory*, vol. IT-11, pp.3-18, Jan. 1965.

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

[Gal86] Gallager, R.G., 'Coding bounds for high signal-to-noise ratio gaussian noise channels,' *Abstract of Papers, IEEE Int. Symp. Inform. Theory*, pp.66-67, Ann Arbor, Michigan, Oct. 1986.

[PaW79] Page, E.S. and Wilson, L.B., *An Introduction to Computational Combinatorics.* Cambridge: Cambridge University Press, 1979.

[Sha48] Shannon, C.E., 'A mathematical theory of communication,' *Bell Syst. Tech. J.*, vol.27, pp.379-423, July 1948.

[Sha59] Shannon, C.E., 'Probability of error for optimal codes in a gaussian channel,' *Bell Syst. Tech. J.*, vol.38, no.3, pp.611-656, May 1959.

[WoJ65] Wozencraft, J.M. and Jacobs, I.M., *Principles of Communication Engineering.* New York: John Wiley & Sons, Inc., 1965.

[WoK66] Wozencraft, J.M. and Kennedy, R.S., 'Modulation and demodulation for probabilistic coding,' *IEEE Trans. Inform. Theory*, vol. IT-12, pp.291-297, July 1966.