PROBLEMS IN CODED COMMUNICATIONS: PERFORMANCE BOUNDS AND POLAR CODING

Eran Hof
Supervised by
Igal Sason and Shlomo Shamai
Outline

• Performance bounds
  – Non-binary linear block codes under ML decoding
  – Binary and non-binary linear block codes under generalized decoding rules

• Polar coding
  – Wire-tap channel
  – Parallel channels with arbitrary input-permutation
Performance analysis of coded communication systems

- Upper and lower bounds on the decoding error probability are interesting as both theoretical and engineering tools
- Union bounds are useless at rates exceeding the channel cut-off rate
- Modern coding schemes perform reliably at rates close to the capacity
A bounding tour

- **Gallager**
  - Fully random block code ensembles
  - Informative at all rates below capacity
- **Duman and Salehi**
  - Structured codes and code ensembles
  - DS2 and generalizations for various memoryless channels
  - Facilitates the derivation of previously reported bounds
- **Shulman and Feder**
  - Structured codes and code ensembles
  - Coincides with Gallager’s random coding bound
  - Non-binary adaptations
- **Sphere Packing lower bounds**
New results


• **Chapter 3**: “Performance Bounds for Erasure, List and Decision Feedback Schemes with Linear Block Codes,” *IEEE Trans. on Information Theory*, vol. 56, no. 8, pp. 3754–3778, August 2010.
Non-binary linear block codes under ML decoding

• Message independence proposition for memoryless symmetric channels
• Adaptation of Gallager’s’65 technique for structured codes or ensembles
  – Adapt the DS2 bound for non-binary codes
  – Derivation of particular bounds which are easy to calculate and provide some insight to the problem.
• Applications for the non-binary regular LDPC ensembles of Gallager
  – A derivation of the exact complete composition spectra
  – A comparison to the ultimate code performance (SP59, ISP)
  – Transmissions over various memoryless channel models
Memoryless symmetric channels

Definition (discrete memoryless channels)

Input-alphabet $\mathcal{X}$
Output alphabet $\mathcal{Y}$
transition probability $p$

$$\exists \mathcal{T} : \mathcal{Y} \times \mathcal{X} \rightarrow \mathcal{Y}$$

1. $\mathcal{T}(\cdot, x) : \mathcal{Y} \rightarrow \mathcal{Y}$ is bijective
2. $p(y|x_1) = p(\mathcal{T}(y, x_2 - x_1)|x_2)$
Particular cases

- Continuous-output alphabet
- MBIOS channels

\[ \mathcal{T}(y, x) = \begin{cases} 
  y & x = 0 \\
  -y & x = 1 
\end{cases} \]

- Coding schemes with a random coset mechanism

\[ \mathcal{T}((y, v), x) = (y, v - x), \quad y \in \mathcal{Y}, \; x, v \in \mathcal{X} \]
Message independence

Proposition

– A1. Linear block codes
– A2. Memoryless symmetric channels

The block error probability under ML decoding is independent on the transmitted message.

Example

Gallager’s symmetry does not guarantee message independence
The DS2 bound

\[ P_{e|m} \leq \left( \sum_{y \in \mathcal{Y}^n} G_n^m(y)p_n(y|x_m) \right)^{1-\rho} \]

\[ \cdot \left\{ \sum_{m' \neq m} \sum_{y \in \mathcal{Y}^n} G_n^m(y)^{1-\frac{1}{\rho}} p_n(y|x_m) \left( \frac{p_n(y|x_{m'})}{p_n(y|x_m)} \right)^\lambda \right\}^\rho \]

Theorem (Non-binary Shulman-Feder)

\[ P_e \leq q^{-nE_r \left( R + \frac{1}{n} \log_q \alpha(C) \right)} \]

\[ \alpha(C) \triangleq \max_{t \in \mathcal{T}^*} \left\{ \frac{E[|C_t|]}{q^{-n(1-R)} \binom{n}{t}} \right\} \]
Ensemble symmetry

\[ \mathbb{E}[|C_t|] = P(n - t_0) \binom{n}{t} \]

Theorem

\[ P_e \leq A(\rho)^{n(1-\rho)} \left( \sum_l \left( P(l) \binom{n}{l} B(\rho)^{n-l} C(\rho)^l \right) \right)^\rho \]

\[ A(\rho) \triangleq \sum_{y \in \mathcal{Y}} \left( \frac{1}{q} \sum_{x \in \mathcal{X}} p(y|x) \frac{1}{1+\rho} \right)^{1+\rho} \]

\[ B(\rho) \triangleq \sum_{y \in \mathcal{Y}} \left( \frac{1}{q} \sum_{x \in \mathcal{X}} p(y|x) \frac{1}{1+\rho} \right)^{\rho^{-1}} \left( \frac{1}{q} \sum_{x \in \mathcal{X}} p(y|x) \frac{2}{1+\rho} \right) \]

\[ C(\rho) \triangleq qA(\rho) - B(\rho) \]
Example
Example
Binary and non-binary linear block codes under generalized decoding rules

• Generalized decoding rules
  – Forney’s optimal decoding rule with erasure and list decoding
  – LR test decoding
  – Fixed-size list decoding

• Message independence properties for linear block codes

• Performance bounds
  – Gallager type (DS2)
  – Error exponents
  – Finite block length analysis
Exponential Error Bounds for Erasure, List, and Decision Feedback Schemes

G. David Forney, Jr., Member, IEEE

Abstract—By an extension of Gallager’s bounding methods, exponential error bounds applicable to coding schemes involving erasures, variable-size lists, and decision feedback are obtained. The bounds are everywhere the tightest known.

Introduction

One of the central problems of coding theory is the question of the error probability obtainable by coding on memoryless channels. Coding theory’s first result, and still its most important, was Shannon’s proof that every memoryless channel has a capacity $C$ such that arbitrarily small error probabilities can be obtained if and only if the code rate $R$ is less than $C$.

Many years of continuing attempts to make more precise statements about error probabilities culminated in Gallager’s elegant derivation of an exponential upper bound on attainable error probabilities, and in a nearly identical lower bound.

(a)  
(b)  
(c)
Exponential Error Bounds for Erasure, List, and Decision Feedback Schemes

G. DAVID FORNEY, JR., MEMBER, IEEE

Abstract—By an extension of Gallager's bounding methods, exponential error bounds applicable to coding schemes involving erasures, variable-size lists, and decision feedback are obtained. The bounds are everywhere the tightest known.

INTRODUCTION

ONE of the central problems of coding theory is the question of the error probability obtainable by coding on memoryless channels. Coding theory's first result, and still its most important, was Shannon's [17] proof that every memoryless channel has a capacity $C$ such that arbitrarily small error probabilities can be obtained if and only if the code rate $R$ is less than $C$. Many years of continuing attempts to make more precise statements about error probabilities culminated in Gallager's elegant derivation [18] of an exponential upper bound on attainable error probabilities, and in a nearly identical lower bound [19].

(a)

\[ \Lambda_m \cap \Lambda_k = \emptyset \]

\[ \bigcup_m \Lambda_m = \mathcal{Y}^n \]
Generalized decoding

Exponential Error Bounds for Erasure, List, and Decision Feedback Schemes

Abstract—By an extension of Gallager's bounding methods, exponential error bounds applicable to coding schemes involving erasures, variable-size lists, and decision feedback are obtained. The bounds are everywhere the tightest known.

Introduction

ONE of the central problems of coding theory is the question of the error probability obtainable by coding on memoryless channels. Coding theory's first result, and still its most important, was Shannon's\textsuperscript{1} proof that every memoryless channel has a capacity \( C \), such that arbitrarily small error probabilities can be obtained if and only if the code rate \( R \) is less than \( C \). Many years of continuing attempts to make more precise statements about error probabilities\textsuperscript{19} culminated in Gallager's elegant derivation\textsuperscript{15} of an exponential upper bound on attainable error probabilities, and in a nearly identical bound bound\textsuperscript{14}.

\[ \Lambda_m \cap \Lambda_k = \emptyset \]
\[ \bigcup_m \Lambda_m = \mathcal{Y}^n \]

\[ \Lambda_m \cap \Lambda_k = \emptyset \]
\[ \bigcup_m \Lambda_m \subseteq \mathcal{Y}^n \]
Exponential Error Bounds for Erasure, List, and Decision Feedback Schemes

G. DAVID FORNEY, JR., MEMBER, IEEE

Abstract—By an extension of Gallager’s bounding methods, exponential error bounds applicable to coding schemes involving erasures, variable-size lists, and decision feedback are obtained. The bounds are everywhere the tightest known.

INTRODUCTION

ONE of the central problems of coding theory is the question of the error probability obtainable by coding on memoryless channels. Coding theory’s first result, and still its most important, was Shannon’s proof that every memoryless channel has a capacity C, such that arbitrarily small error probabilities can be obtained if and only if the code rate R is less than C. Many years of continuing attempts to make more precise statements about error probabilities culminated in Gallager’s elegant derivation of an exponential upper bound on attainable error probabilities, and in a nearly identical lower bound.

\[ \Lambda_m \cap \Lambda_k = \emptyset \]
\[ \bigcup_m \Lambda_m = \mathcal{Y}^n \]

\[ \Lambda_m \cap \Lambda_k = \emptyset \]
\[ \bigcup_m \Lambda_m \subseteq \mathcal{Y}^n \]

\[ \Lambda_m \cap \Lambda_k \neq \emptyset \]
\[ \bigcup_m \Lambda_m \subseteq \mathcal{Y}^n \]
Error probabilities

Block error probability

\[ P_{e|m} = \sum_{y \in \Lambda_m^c} p(y|x_m) \]
Error probabilities

Block error probability

\[ P_{e|m} = \sum_{y \in \Lambda_m^c} p(y|x_m) \]

Undetected error probability

\[ P_{ue|m} = \sum_{m' \neq m} \sum_{y \in \Lambda_{m'}} p(y|x_m) \]
Error probabilities

Block error probability

\[ P_{e|m} = \sum_{y \in \Lambda_{m}^{c}} p(y|x_m) \]

Undetected error probability

\[ P_{ue|m} = \sum_{m' \neq m} \sum_{y \in \Lambda_{m'}} p(y|x_m) \]

Erasure probability

\[ P_{x|m} = P_{e|m} - P_{ue|m} \]
Error probabilities

Block error probability

\[ P_{e|m} = \sum_{y \in \Lambda_m^c} p(y | x_m) \]

Undetected error probability

\[ P_{ue|m} = \sum_{m' \neq m} \sum_{y \in \Lambda_{m'}} p(y | x_m) \]

Expected number of incorrect codewords
Optimal decoding

\[ \Lambda_m = \left\{ y \in \mathcal{Y}^n : \frac{\Pr(y, x_m)}{\sum_{m' \neq m} \Pr(y, x_{m'})} \geq e^{nT} \right\} \]

Optimal generalized decoding of convolutional codes

• Via the BCJR algorithm

\[ \frac{\Pr(y, x_m)}{\sum_{m' \neq m} \Pr(y, x_{m'})} = \frac{\Pr(y, x_m)}{\Pr(y) - \Pr(y, x_m)} \]

• Via the VA
  – path and branch metrics: standard add-compare-select
  – recursive evaluation of the denominator via the generalized metrics
    • Multiply by the branch metric (addition)
    • Add the not survived metrics

• Channel state information at the RX
Example
Optimal decoding

\[ \Lambda_m = \left\{ y \in \mathcal{Y}^n : \frac{\Pr(y, x_m)}{\sum_{m' \neq m} \Pr(y, x_{m'})} \geq e^{nT} \right\} \]

Error exponents for feedback schemes
Message independence for generalized decoding

**Proposition**
A1. Linear block codes
A2. Memoryless symmetric channels

The block error probability and undetected error probability under Forney’s optimal decoding rule are independent of the transmitted message.

**For list decoding:** the block error probability and the average number of incorrect codewords are independent of the transmitted message.

Similar propositions for LR test and fixed-size list-decoding.
Block codes under optimal decoding

**Proposition:**

\[
P_{e|m} \leq e^{nsT} D_B(m, G^m_n, s, \rho)
\]

\[
P_{ue} \leq e^{n(s-1)T} \frac{1}{M} \sum_{m=1}^{M} D_B(m, G^m_n, s, \rho), \quad s \geq 0, \quad 0 \leq \rho \leq 1
\]

\[
D_B(m, G^m_n, s, \rho) = \left( \sum_y G^m_n(y)p(y|x_m) \right)^{1-\rho}
\]

\[
\left( \sum_{m' \neq m} \sum_y p(y|x_{m'}) G^m_n(y)^{1-\frac{1}{\rho}} \left( \frac{p(y|x_{m'})}{p(y|x_m)} \right)^{\frac{s}{\rho}} \right)^{\rho}
\]
Block codes under optimal decoding

Proposition:

\[ P_{e|m} \leq e^{nsT} D_B(m, G_n^m, s, \rho) \]
\[ P_{ue} \leq e^{n(s-1)T} \frac{1}{M} \sum_{m=1}^{M} D_B(m, G_n^m, s, \rho), \quad s \geq 0, \quad 0 \leq \rho \leq 1 \]

Corollary (random coding):

\[ P_e \leq e^{-nE_1(R,T)} \]
\[ P_{ue} \leq e^{-nE_2(R,T)} \]
Linear block codes
Memoryless symmetric channels
Optimal decoding

\[ P_e \leq e^{-n \left( E(\rho, R, C) - \frac{\rho T}{1+\rho} \right) } , \quad P_{ue} \leq e^{-n \left( E(\rho, R, C) + \frac{T}{1+\rho} \right) } \]

\[ E(\rho, R, C) \triangleq E_0(\rho) - \rho \left( R + \frac{\ln(\alpha(C))}{n} \right) \]

\[ E_0(\rho) \triangleq -\ln \left( \sum_y \left( \frac{1}{2} p(y|0)^{\frac{1}{1+\rho}} + \frac{1}{2} p(y|1)^{\frac{1}{1+\rho}} \right)^{1+\rho} \right) \]

\[ \alpha(C) \triangleq \max_{1 \leq i \leq n} \frac{|C_i|}{2^{-(n-k)} \binom{n}{i}} \]
Example
Performance of Feedback schemes

(a) Lower bounds on the expected rates

(b) Upper bounds on the error probability
Polar codes

- Arikan, IEEE IT July 2009
- Arikan and Telatar, ISIT 2009
- Korada and Urbanke, IEEE IT, April 2010
Recursive definition of polar codes
Recursive definition of polar codes
Combined channels

\[ W_N(y|w) \]

Information bits:

\[ w_i, \quad i \in \mathcal{A} \]

Pre-determined bits

\[ w_i, \quad i \notin \mathcal{A} \]
Split channels

Inputs

outputs

$W_N^{(1)}(y | w_1)$
Split channels

Outputs

- $w_1$
- $w_2$
- $w_{N/2} - 1$
- $w_{N/2}$
- $w_{N/2} + 1$
- $w_{N/2} + 2$
- $w_{N-1}$
- $w_N$

Inputs

$W_{N/2}$

outputs

- $y_1$
- $y_2$

$W_N^{(1)}(y|w_1)$

$W_N^{(2)}(y, w_1|w_2)$
The ‘magic’ of channel polarization

Magic no. 1:
The split channels polarize to perfect or nothing

Magic no. 2:
There are enough “good” indices
The symmetric capacity is achievable=index ordering is non-trivial and non-consecutive

Magic no. 3:
Successive cancellation decoding
Chapter 5: Polar coding for the wiretap channel

Theorem: For symmetric and degraded channels

Achievable rates \((R, R_e)\)

\[ P_e \to 0 \]

\[ R_e \leq \lim_{N \to \infty} \frac{1}{N} H(U|Z) \]

Theorem: For symmetric and degraded channels

\[
\left\{ (R, R_e) : \begin{array}{l}
0 \leq R \leq C(P_Y|X) \\
0 \leq R_e \leq R \\
R_e \leq C(P_Y|X) - C(P_Z|X)
\end{array} \right\}
\]
Polar secrecy encoding

Index sets

\( A \) \( \cap \) \( N \) \( \cap \) \( B \)

fixed bit
information bit
Random bit
information bit
Random bit
A secrecy achieving property

Theorem (physically-degraded and symmetric wire-tap channels)

\[ R < C(P_{Y|X}) - C(P_{Z|X}) \]

\[ \exists \mathcal{A}, \mathcal{N} \quad R \leq \frac{1}{N} |\mathcal{A}_N| \]

\[ \lim_{n \to \infty} \frac{1}{N} H(U|Z) \geq R \]

\[ P_e \to 0 \]
Secrecy achieving polar coding

Index sets
Good for both \( \mathcal{N} \)
Good only for the legitimate \( \mathcal{A} \)
The rest \( \mathcal{B} \)

1. A Polarization argument for the set sizes
2. Successive decoding for the legitimate user
3. Analysis of the equivocation rate – based on Fano’s inequality
Secret and private messages with polar coding

Secret bit
Private bit
Secret bit
Secret bit
Private bit

W_{N/2}
W_{N/2}
Chapter 6: Polar coding for Arbitrarily Permuted Parallel Channels

Achievable rates:

\[
\frac{1}{n} \log_2 M \geq R - \delta
\]
\[
P_{e}^{(\pi)}(n) \leq \delta, \forall \pi : [S] \rightarrow [S]
\]

Theorem (Willems and Gorokhov):

\[
C_{\Pi} = \sum_{s=1}^{S} C_{s}
\]
Monotone index sets

Corollary:

Binary -input memoryless degraded symmetric channels.

\[ C_1 \geq C_2 \geq \cdots \geq C_S. \]

Rates:

\[ 0 \leq R_s \leq C_s \]

There exists “good” monotone index sets:

\[ \mathcal{A}^{(S)} \subseteq \mathcal{A}^{(S-1)} \subseteq \cdots \subseteq \mathcal{A}^{(1)} \]

\[ |\mathcal{A}^{(s)}| \geq NR_s \]
Monotone index sets

Polarized for the worse channel: \( C_S, \ R_S, \ \mathcal{A}^{(S)} \)

What is good for the worse, is good for the almost worse

Polarization gives the rest of the “good” indices:

\[ \mathcal{A}^{(S)} \subseteq \mathcal{A}^{(S-1)} \]

Repeat this,...Finally:

\[ \mathcal{A}^{(S)} \subseteq \mathcal{A}^{(S-1)} \subseteq \cdots \subseteq \mathcal{A}^{(1)} \]

\[ |\mathcal{A}^{(s)}| \geq NR_s \]
Parallel polar coding for 2 channels
Parallel polar coding for 2 channels

The decoding process for the stronger channel generates the “predetermined and fixed” bits for the degraded channel.
Parallel polar coding for 3 channels

Channel 1

A repetition code for the good indices of channel 1 which are not good for channel 2

Channel 2

Channel 3

A parity check code for the good indices of channel 2 which are not good for channel 3
Encoding and decoding for 3 channels

\[ x_1 = u_{1,1} G_n \left( A_n^{(3)} \right) + u_{1,2} G_n \left( A_n^{(2)} \setminus A_n^{(3)} \right) + u_r G_n \left( A_n^{(1)} \setminus A_n^{(2)} \right) + b G_b \left( [n] \setminus A_n^{(1)} \right) \]

\[ x_2 = u_{2,1} G_n \left( A_n^{(3)} \right) + u_{2,2} G_n \left( A_n^{(2)} \setminus A_n^{(3)} \right) + u_r G_n \left( A_n^{(1)} \setminus A_n^{(2)} \right) + b G_n \left( [n] \setminus A_n^{(1)} \right) \]

\[ x_3 = u_3 G_n \left( A_n^{(3)} \right) + (u_{1,2} + u_{2,2}) G_n \left( A_n^{(2)} \setminus A_n^{(3)} \right) + u_r G_n \left( A_n^{(1)} \setminus A_n^{(2)} \right) + b G_n \left( [n] \setminus A_n^{(1)} \right) \]

<table>
<thead>
<tr>
<th>Channel ( P_1 )</th>
<th>Channel ( P_2 )</th>
<th>Channel ( P_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmitted Codeword</td>
<td>Decoded Information</td>
<td>Transmitted Codeword</td>
</tr>
<tr>
<td>( x_1 )</td>
<td>( u_{1,1}, u_{1,2}, u_r )</td>
<td>( x_2 )</td>
</tr>
<tr>
<td>( x_2 )</td>
<td>( u_{2,1}, u_{2,2}, u_r )</td>
<td>( x_3 )</td>
</tr>
<tr>
<td>( x_3 )</td>
<td>( u_3, u_{1,2} + u_{2,2}, u_r )</td>
<td>( x_1 )</td>
</tr>
<tr>
<td>( x_1 )</td>
<td>( u_{1,1}, u_{1,2} )</td>
<td>( x_2 )</td>
</tr>
</tbody>
</table>
The general case

• Theorem (MDS codes):
  For an MDS code of dimension $k$ every $k$ symbols completely characterize the codeword

• Combined MDS & Polar codes for the general degraded case

• The proposed technique does not achieve the capacity in the non-degraded case
Non-degraded parallel channels

- The worst Bhattacharyya parameter whole the split channels
- A parallel polarization based on the corresponding erasure channels

\[ S - \sum_{s=1}^{S} B(P_s) \]

- Upper and (improved) lower bounds based on a compound setting
Upper and lower bounds

• A lower bound for the compound setting

\[ C\left(\{P_s\}_{s \in S}\right) \geq 1 - \frac{1}{2^k} \sum_{\sigma \in \{0,1\}^k} \max_{s \in S} B(P_s^\sigma) \]

• A lower bounds for the parallel transmission scheme

\[ C(P_{sS}) + S - 1 - \frac{1}{2^k} \sum_{s \in [S-1]} \sum_{\sigma \in \{0,1\}^k} \max_{i \in \{s, \ldots, S\}} B(P_{si}^\sigma). \]

• An Upper bound

• A BSC+BEC Example (0.982 bits per channel use)
Summary

• Performance bounds
  – Non-binary linear block codes under ML decoding
  – Binary and non-binary linear block codes under generalized decoding rules

• Polar coding
  – Wire-tap channel
  – Parallel channels with arbitrary input-permutation