Optimal Ultra-Small Block-Codes for Binary Input DMCs

Hsuan-Yin Lin

Joint work with Prof. Stefan M. Moser and Prof. Po-Ning Chen
Channel Coding Theory

- Four important parameters of a code:
  - **rate** $R$: how many messages can be sent per channel use
    \[ R \triangleq \frac{\log M}{n} \]
  - **blocklength** $n$: the larger the better performance, but more delay
  - **number of codewords** $M$
  - **error probability** $P_e$: depends strongly on the rate, coding scheme, and blocklength
## Introduction

<table>
<thead>
<tr>
<th>Channel coding theorem</th>
<th>Rate $R &lt; C$ (sufficiently large $n$)</th>
<th>Error $P_e \rightarrow 0$ (arbitrary small)</th>
<th>Code (linear code exists)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Shannon, Gallager, Berlekamp, part II, IC67]</td>
<td>$n \gg M$ (zero rate)</td>
<td>lower and upper bounds (average prob.)</td>
<td>fair weak flip codes (error exponent)</td>
</tr>
<tr>
<td>Our approach</td>
<td>fixed $M$ for any $n$</td>
<td>find the smallest one (average prob.)</td>
<td>give optimal code structure</td>
</tr>
</tbody>
</table>

- **Introduction**

- **Table**

- **Rate**
- **Blocklength $n$**
- **Error $P_e$**
- **Code**

- **Channel coding theorem**: $R < C$ (sufficiently large $n$)
- **Error** $P_e \rightarrow 0$ (arbitrary small)
- **Code**: random coding

- **[Shannon, Gallager, Berlekamp, part II, IC67]**
  - Rate: $n \gg M$ (zero rate)
  - Error: lower and upper bounds (average prob.)
  - Code: fair weak flip codes (error exponent)

- **Our approach**
  - Rate: fixed $M$
  - Blocklength: fixed for any $n$
  - Error: find the smallest one (average prob.)
  - Code: give optimal code structure

Hsuan-Yin Lin, National Chiao Tung University (NCTU), Taiwan
Channel Model: Binary Asymmetric Channel (BAC)

- To simplify the problem, we focus on binary input DMCs

For symmetry, let $0 \leq \epsilon_0 \leq \epsilon_1 \leq 1$, and $\epsilon_0 \leq 1 - \epsilon_0$.
The Analysis Region on BAC

We analyze region $\Omega$: 

\[ \epsilon_1 \]
\[ 1 \]
\[ \frac{1}{2} \]
\[ \Omega \]

\[ \epsilon_0 \]
\[ \frac{1}{2} \]
\[ 1 \]
Channel Models: BSC & ZC

• Special cases:
  – if $\epsilon_0 = \epsilon_1$: Binary Symmetric Channel (BSC)
  – if $\epsilon_0 = 0$: Z-channel (ZC)
  – if $\epsilon_0 = 1 - \epsilon_1$: completely noisy channel ($X \perp\!\!\!\!\perp Y$)
The Analysis Region of BAC: Special Cases

The region $\Omega$:

- $\epsilon_0 + \epsilon_1 = 1$ (completely noisy)
- $\epsilon_0 - \epsilon_1 = 0$ (BSC)
- $\epsilon_0 = 0$ (Z-channel)
Channel Model: BEC

- The other binary input DMC: **Binary Erasure Channel (BEC)**
How to Define the Codes

- Let $M =$ number of codewords, $n =$ length of a codeword

- Usually we write a code as

\[ C(M,n) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{pmatrix} \]

which is called an $(M, n)$ code
Column-Based Analysis

- To address the optimal codes of a DMC, we now have a good approach based on the column-wise viewpoint.

- What’s the difference from row-wise viewpoint?

\[
\mathcal{C}^{(M,n)} = \begin{pmatrix}
\cdots
\vdots
\cdots
\end{pmatrix} = \begin{pmatrix}
c_1 \\
c_2 \\
\cdots \\
c_n
\end{pmatrix}
\]
Motivation: An Example on BSC

- Two arbitrary given codebooks, \((M, n) = (4, 4)\):

\[
\mathcal{C}_1 = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1
\end{pmatrix}, \quad d_{\text{min}} = 1
\]

\[
\mathcal{C}_2 = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1
\end{pmatrix}, \quad d_{\text{min}} = 2
\]

- Which one has smaller average error probability by ML decoder?*

\[
P_e \leq (M - 1) \exp \left[ -d_{\text{min}} \log \frac{1}{\sqrt{4\epsilon(1 - \epsilon)}} \right]
\]

\[\Rightarrow\] compare the exact error, the better one is the first!

Outline

- Introduction and channel models
- Weak Flip Codes and Hadamard Codes
- Main result:
  - Optimal codes for Binary Input DMCs
    * $\ell$-optimal codes for BAC with $M = 2$
    * Optimal codes for ZC with $M \leq 4$
    * Optimal codes with $M \leq 3$, linear optimal codes with $M = 4$
      for BSC
    * Optimal codes with $M \leq 3$, linear optimal codes with $M = 4$
      for BEC
  - Comparisons of exact error probability with some known bounds
- Conclusion and future work
Flip Codes

- For $M = 2$, we define the following code:

$$C_t^{(2,n)} \triangleq \begin{pmatrix} x \bar{x} \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 & 1 & \cdots & 1 \\ 1 & \cdots & 1 & 0 & \cdots & 0 \end{pmatrix}$$

flip code of type $t$
Flip Codes

• For $M = 2$, we define the following code:

$$C_t^{(2,n)} \triangleq \begin{pmatrix} x \\ \bar{x} \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 & 1 & \cdots & 1 \\ 1 & \cdots & 1 & 0 & \cdots & 0 \end{pmatrix}$$

flip code of type $t$

• A flip code of type $t$ is given by a codebook that consists of $(n - t)$ columns $c_1^{(2)}$ and $t$ columns $c_2^{(2)}$, where

$$\begin{cases} c_1^{(2)} \triangleq \begin{pmatrix} 0 \\ 1 \end{pmatrix}, & c_2^{(2)} \triangleq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{cases}$$
Weak Flip Code

- For $M = 3$ or $M = 4$: A weak flip code of type $(t_2, t_3)$, $c_{t_2, t_3}^{(M,n)}$ is defined by a code that consists of

  - $t_1 (= n - t_2 - t_3)$ columns $c_1^{(M)}$
  - $t_2$ columns $c_2^{(M)}$
  - $t_3$ columns $c_3^{(M)}$

where...
Weak Flip Code

for $M = 3$,

$$\left\{ \begin{array}{l}
c_1^{(3)} \triangleq \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad c_2^{(3)} \triangleq \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad c_3^{(3)} \triangleq \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \\
\end{array} \right.$$ 

for $M = 4$,

$$\left\{ \begin{array}{l}
c_1^{(4)} \triangleq \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad c_2^{(4)} \triangleq \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad c_3^{(4)} \triangleq \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \\
\end{array} \right.$$
Weak Flip Code for General $M$

- **Weak flip column**: its first component is 0 and its Hamming weight equals to $\lfloor \frac{M}{2} \rfloor$ or $\lceil \frac{M}{2} \rceil$

- **Weak flip code**: constructed only by weak flip columns
  - **Fair weak flip code**: by using an equal number of all possible weak flip columns
Weak Flip Code for General $M$

• **Weak flip column**: its first component is 0 and its Hamming weight equals to $\left\lfloor \frac{M}{2} \right\rfloor$ or $\left\lceil \frac{M}{2} \right\rceil$

• **Weak flip code**: constructed only by weak flip columns
  – **Fair weak flip code**: by using an equal number of all possible weak flip columns

• If an $(M, n)$ binary code is linear, then each column has Hamming weight $\frac{M}{2}$, i.e., the code is a weak flip code!
Plotkin Bound

• **Plotkin Bound**: An upper bound for the minimum Hamming distance of any given \((M, n)\) binary block code.

• **Theorem**: Plotkin bound can be achieved, provided that Hadamard matrices exist, *i.e.*, achieved by **Hadamard codes**†

---

**Plotkin Bound**

- **Plotkin Bound**: An upper bound for the minimum Hamming distance of any given \((M,n)\) binary block code.

- **Theorem**: Plotkin bound can be achieved, provided that Hadamard matrices exist, *i.e.*, achieved by Hadamard codes†

- Fair weak flip codes also achieve the Plotkin bound

---

Plotkin Bound

- **Plotkin Bound**: An upper bound for the minimum Hamming distance of any given \((M, n)\) binary block code.

- **Theorem**: Plotkin bound can be achieved, provided that Hadamard matrices exist, *i.e.*, achieved by **Hadamard codes**

- Fair weak flip codes **also achieve the Plotkin bound**

- For \((M, n)\) such that Hadamard codes exist, fair weak flip codes can be seen as a **subset** of Hadamard codes

---

Fair Codes and Hadamard Codes

- A weak flip code but not fair with $M = 4$, blocklength $n = 7$:

$$
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 & 1 & 1 \\
1 & 1 & 0 & 1 & 1 & 0 & 1 \\
\end{pmatrix}
$$
Fair Codes and Hadamard Codes

• A weak flip code but not fair with $M = 4$, blocklength $n = 7$:

\[
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 & 1 & 1 \\
1 & 1 & 0 & 1 & 1 & 0 & 1
\end{pmatrix}
\]

• A fair weak flip code with $M = 3$, blocklength $n = 9$:

\[
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\
1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1
\end{pmatrix}
\]

– Using candidate columns fairly, $n \mod 3 = 0$
Outline

• Introduction and channel models

• Weak Flip Codes and Hadamard Codes

• Main result:
  – Optimal codes for Binary Input DMCs
    * \(\ell\)-optimal codes for BAC with \(M = 2\)
    * Optimal codes for ZC with \(M \leq 4\)
    * Optimal codes with \(M \leq 3\), linear optimal codes with \(M = 4\) for BSC
    * Optimal codes with \(M \leq 3\), linear optimal codes with \(M = 4\) for BEC
  – Comparisons of exact error probability with some known bounds

• Conclusion and future work
Main Results (BAC)

• The optimal codes for $M = 2$ are always flip codes of type $t$

• Recall $t =$ Hamming weight of the first codeword of a flip code of type $t$
Main Results (BAC)

- The optimal codes for $M = 2$ are always flip codes of type $t$

- Recall $t = \text{Hamming weight of the first codeword}$ of a flip code of type $t$

- Note: On BAC the exact choice of $t$ is not obvious and depends strongly on $n$, $\epsilon_0$, and $\epsilon_1$
The Optimal Codes with $M = 2$ on BAC

$n = 7$

$\epsilon_0 = 0.25$

$\epsilon_1 = 0.6$

Hsuan-Yin Lin, National Chiao Tung University (NCTU), Taiwan
The Optimal Codes with $M = 2$ on BAC

$n = 9$

$\epsilon_0 = 0.25$

$\epsilon_1 = 0.6$

Hsuan-Yin Lin, National Chiao Tung University (NCTU), Taiwan
The Optimal Codes with $M = 2$ on BAC

$n = 11$

$\epsilon_0 = 0.25$, $\epsilon_1 = 0.6$

Hsuan-Yin Lin, National Chiao Tung University (NCTU), Taiwan
Optimal Decoding Rule

- $t =$ number of 1’s in the first codeword
- $d =$ Hamming distance of a received $y$ to the first codeword
Optimal Decoding Rule

- $t =$ number of 1's in the first codeword
- $d =$ Hamming distance of a received $y$ to the first codeword

We can prove that for any fixed $t, n, \epsilon_0, \epsilon_1$, the ML decoding consists of a threshold $\ell$ such that

\[
\begin{cases}
0 \leq d \leq \ell & \implies g(y) = x \\
\ell + 1 \leq d \leq n & \implies g(y) = \overline{x}
\end{cases}
\]

(note that $t \leq \ell \leq \lfloor \frac{n-1}{2} \rfloor$)
Optimal Flip Codes Depend on \((\epsilon_0, \epsilon_1)\)

\[\epsilon_1 = 0.4, n = 7\]

\(\ell = 2 \text{ to } \ell = 3 \text{ for } t = 1\)

\(\ell = 2 \text{ to } \ell = 3 \text{ for } t = 0\)

\[\begin{align*}
\text{error probability} & \\
\epsilon_0 & \\
\text{Hsuan-Yin Lin, National Chiao Tung University (NCTU), Taiwan}
\end{align*}\]
Fix Decision Rule

- The optimal code $t$ depends strongly on $(\epsilon_0, \epsilon_1), n, \ell$
- The optimal $\ell$ depends strongly on $(\epsilon_0, \epsilon_1), n, t$

- Both $\ell$ and the optimal choice of $t$ vary strongly with different $(\epsilon_0, \epsilon_1)$
Fix Decision Rule

- The optimal code $t$ depends strongly on $(\epsilon_0, \epsilon_1), n, \ell$
- The optimal $\ell$ depends strongly on $(\epsilon_0, \epsilon_1), n, t$

- Both $\ell$ and the optimal choice of $t$ vary strongly with different $(\epsilon_0, \epsilon_1)$

- To simplify our problem use a sub-optimal decoder: fix $\ell$, try to find the best choice of $t$
The Best Codes When Fixed Decision Rule

$n = 7, \ell = 2$

$t = 0$

$t = 1$

$t = 2$

Hsuan-Yin Lin, National Chiao Tung University (NCTU), Taiwan
The Optimal Codes and Fixed Decision Rule

$n = 7$

$t = 0$
$t = 1$
$t = 2$
$t = 3$
Main Results (ZC)

- The optimal code with $M = 2$ is the flip code of type 0, $C^{(2,n)}_0$

- An optimal code with $M = 3$ or $M = 4$ is the weak flip code of type $(t^*, 0)$, $C^{(M,n)}_{t^*,0}$ with

$$t^* \triangleq \left\lceil \frac{n}{2} \right\rceil$$

- $M = 3$: $t_2 + t_3 = t^*$
- $M = 4$, $n$ odd, $t^*$ can be $\left\lceil \frac{n}{2} \right\rceil$
Main Results (ZC)

• The optimal code with $M = 2$ is the flip code of type 0, $\mathcal{C}_{0}^{(2,n)}$

• An optimal code with $M = 3$ or $M = 4$ is the weak flip code of type $(t^*, 0)$, $\mathcal{C}_{t^*,0}^{(M,n)}$ with

$$t^* \triangleq \left\lceil \frac{n}{2} \right\rceil$$

- $M = 3$: $t_2 + t_3 = t^*$
- $M = 4$, $n$ odd, $t^*$ can be $\left\lfloor \frac{n}{2} \right\rfloor$

• These codes are optimal through all $\epsilon_1$ on ZC
Recursive Construction (ZC)

- An optimal code of blocklength $n$ for $M = 3$ on ZC is
  \[
  \begin{pmatrix}
  0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
  0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\
  1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
  \end{pmatrix}
  \]

- An optimal code of blocklength $n$ for $M = 4$ on ZC is
  \[
  \begin{pmatrix}
  0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
  0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\
  1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
  1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
  \end{pmatrix}
  \]
Outline

• Introduction and channel models
• Weak Flip Codes and Hadamard Codes
• Main result:
  – Optimal codes for Binary Input DMCs
    * $\ell$-optimal codes for BAC with $M = 2$
    * Optimal codes for ZC with $M = 3, 4$
    * Optimal codes with $M \leq 3$, linear optimal codes with $M = 4$
      for BSC
    * Optimal codes with $M \leq 3$, linear optimal codes with $M = 4$
      for BEC
  – Comparisons of exact error probability with some known bounds
• Conclusion and future work
Main Results (BSC & BEC)

- Optimal codes for $M = 2$ are flip codes of any type $t$

- An optimal code of blocklength $n$ for $M = 3$ on BSC is

$$
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\
1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\
\end{pmatrix}
\ldots
$$

- An optimal code of blocklength $n$ for $M = 3$ on BEC is

$$
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\
1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\
\end{pmatrix}
\ldots
$$
Main Results (BSC & BEC)

• A linear optimal code of blocklength $n$ for $M = 4$ on BSC is

\[
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & \ldots \\
1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\
1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\
\end{pmatrix}
\]

• A linear optimal code of blocklength $n$ for $M = 4$ on BEC is

\[
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & \ldots \\
1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\
\end{pmatrix}
\]
A Comparison of Exact Error and Other Bounds (ZC)

ZC \( M = 3, \epsilon_1 = 0.3 \)

Error Probability vs. Blocklength

- Gallager upper bound
- SGB* up. b. for fair code
- SGB up. b. for optimal code
- Optimal (exact)
- SGB low. b. for optimal code

* SGB: Shannon–Gallager–Berlekamp, IC67

Hsuan-Yin Lin, National Chiao Tung University (NCTU), Taiwan
A Comparison of Exact Error and Other Bounds (BSC)

BSC $M = 3, \epsilon = 0.3$

* PPV: Polyanskiy–Poor–Verdú, IT10
A Comparison of Exact Error and Other Bounds (BEC)

BEC $M = 3, \delta = 0.3$
Conclusion

- We have shown the global optimal codes for some certain ultra-small $M$ on binary input DMCs.

- New way of looking at codes: *column-wise analysis*

- Analyze the relation between weak flip codes and Hadamard codes
Future Work

• Finish the proof of $M = 4, 5, 6$

• Investigate some known bounds for finite blocklength $n$
  – compare to the exact performance of best codes with arbitrary $M$
Thank you for your attention!