誤り訂正符号における 誤りの単調性を利用した訂正能力分析

安永憲司

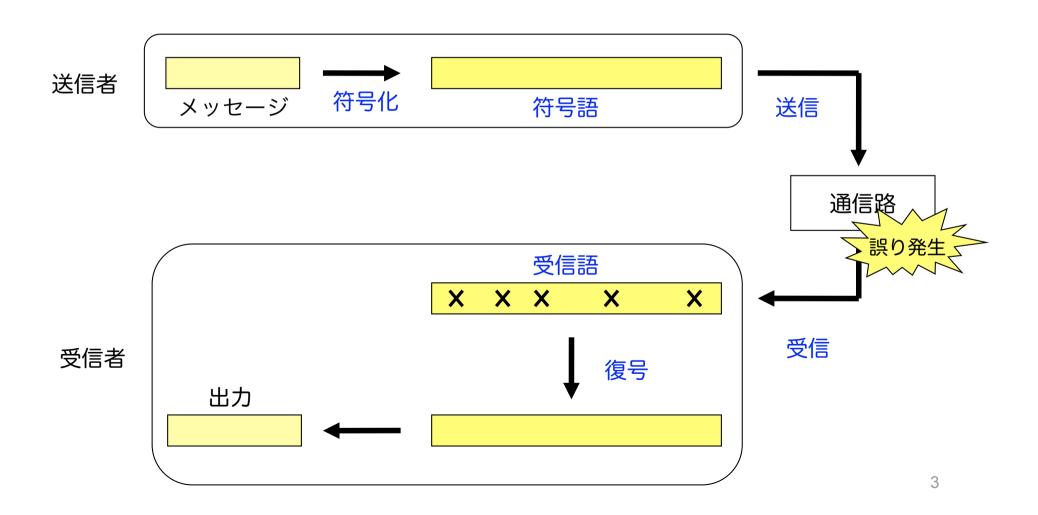
東京工業大学 大学院情報理工学研究科 数理·計算機科学専攻 GCOE特任助教 2008年10月22日

発表の流れ

- 誤り訂正符号
- 本研究で扱う問題(誤り訂正能力)
- 既存の結果・これまでの研究成果
- 扱う問題についてもう少し詳しく
- 誤りの単調性
- 研究成果の詳細
- **まとめ**

誤り訂正符号

■ 送信メッセージに冗長性をもたせることで通信路で発生した誤りを 訂正することが可能



誤り訂正符号の例

- 3回繰り返し符号
 - メッセージの各ビットを3回ずつ繰り返して送る

- 1ビットの誤り(0と1が反転)ならば訂正できる受信語が010111ならば000111に復号
- 2ビット以下の誤りのとき、訂正できない場合もある
 - □ 受信語が 010101 ならば 000111 に復号
 - □ 受信語が 011111 ならば 111111 or 000111

用語の定義

- 符号: 符号語の集合 (線形空間をなす → 線形符号)
- (*n*, *k*) 線形符号 *C* :符号長 *n*, メッセージ長(次元)*k* の線形符号
 - $C \subseteq \{0, 1\}^n, |C| = 2^k$
- 符号の最小距離 d:異なる符号語間の最小ハミング距離

$$d = \min_{\substack{\mathbf{c}_1, \mathbf{c}_2 \in C \\ \mathbf{c}_1 \neq \mathbf{c}_2}} d(\mathbf{c}_1, \mathbf{c}_2) = \min_{\mathbf{c} \in C \setminus \{0\}} w(\mathbf{c})$$
 線形符号の場合

- □ *d*(*x*, *y*): *x* と *y* のハミング距離
- □ w(x):xのハミング重み

本研究では2元線形符号を扱う

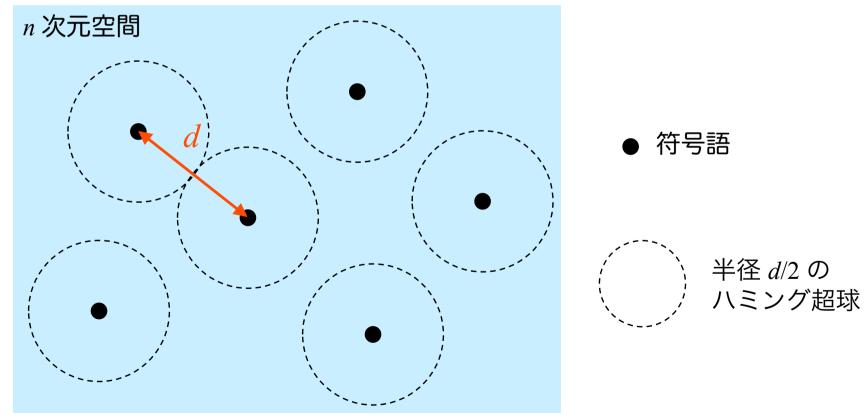
本研究で扱う問題(誤り訂正能力)

- 受信語 $y = c + e \in \{0,1\}^n$
 - c∈ C: 送信符号語
 - $e \in \{0,1\}^n$: 誤りベクトル
 - □ w(e) = 発生した誤りのビット数

$$w(e) < d/2$$
 ⇒ 必ず訂正可能 $w(e) \ge d/2$ ⇒ ??

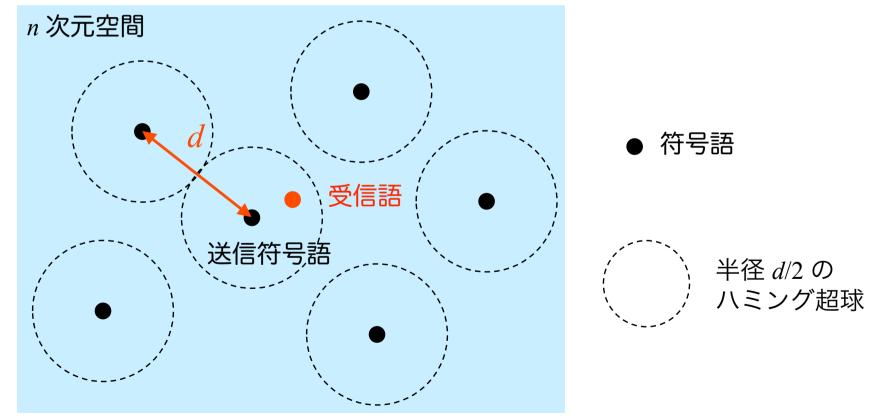
重み d/2 未満の誤りが必ず訂正可能な理由

符号の最小距離 d: すべての符号語間の最小ハミング距離



重み d/2 未満の誤りが必ず訂正可能な理由

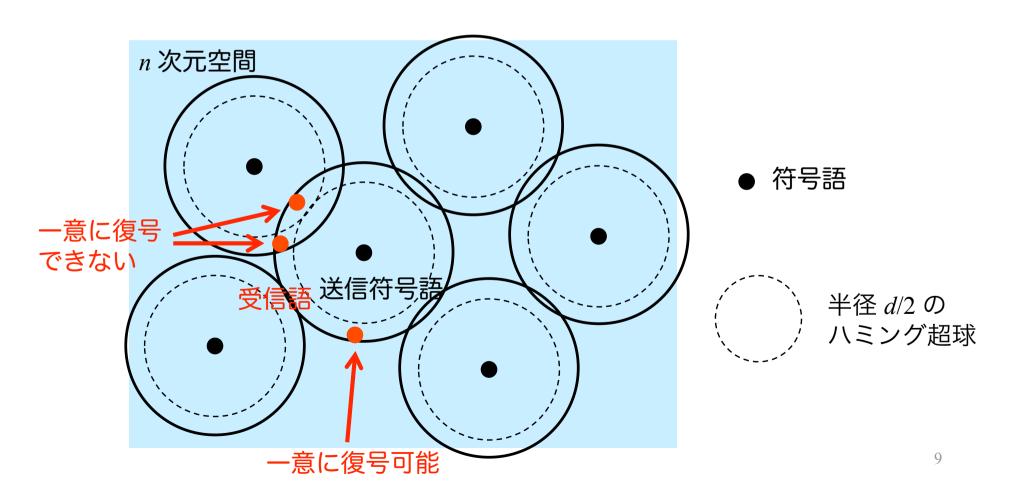
誤りの重みが d/2 未満のとき、 受信語は送信符号語のハミング超球に含まれる ⇒ 一意に復号可能



重みが d/2 以上の誤りが発生する場合

誤りの重みが d/2 より少し大きい場合 $(d/2+\alpha$ 以下の場合)

半径 $d/2+\alpha$ のハミング超球が重ならない点に 受信語があれば一意に復号できる

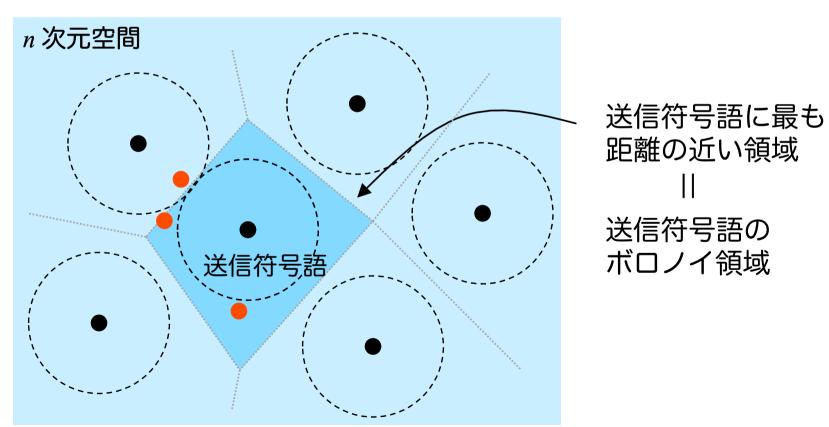


本研究で考える復号法(最小距離復号法)

- 受信語から距離が最小の符号語に復号
- 2元対称通信路で最適(=復号誤り率を最小にする)復号法
 - 2元対称通信路
 - □ 各ビットごとに一定確率 p (0 ≤ p < 1/2)で0と1が反転する通信路</p>
 - □ 符号理論で最も研究されている通信路モデル

本研究で考える復号法(最小距離復号法)

- 受信語から距離が最小の符号語に復号
- 送信符号語に最も距離の近い領域 = 訂正可能な受信語の領域



本研究で扱う問題(誤り訂正能力)

- 受信語 $y = c + e \in \{0,1\}^n$
 - $c \in C$: 送信符号語, $e \in \{0,1\}^n$: 誤りベクトル

最小距離復号法を行う場合

w(e) < d/2 ⇒ 必ず訂正可能

 $w(e) \ge d/2$ \Rightarrow 受信語が送信符号語のボロノイ領域に 入っていれば訂正可能

線形符号では、どの符号語のボロノイ領域も同じ形

ightarrow 訂正可能な誤りはどの送信符号語でも同じ (c=0) を考える)

本研究では、 重み d/2 以上の訂正可能な誤りベクトルの数について研究

既存の結果

- 一般の符号に対して
 - 重み i (i ≥ d/2) の訂正不可能誤りベクトルの数の上界を導出
 [Poltyrev 1994], [Helleseth, Kløve 1997], [Helleseth, Kløve, Levenshtein 2005]
- 1次 Reed-Muller 符号に対して
 - n = 32 について、すべての重みの訂正可能誤りベクトルの数を計算 [Berlekamp, Welch 1972]
 - 重み d/2 の訂正可能誤りベクトルの数を導出 [Wu 1998]
- その他の符号に対して
 - 2重誤り訂正 BCH 符号 [Charpin 1994]
 - 3重誤り訂正 BCH 符号 [Charpin, Helleseth, Zinoviev 2006]
 - $n \le 128, 29 \le n k \le 42$ の Reed-Muller 符号・BCH 符号について計算 [Maeda, Fujiwara 2001]

これまでの研究成果

■ 一般の符号に対して

(成果 1) 重み d/2 の訂正不可能誤りベクトルの数の下界を導出 (成果 2) 重み d/2+1 以上への拡張

■ 1次 Reed-Muller 符号に対して

(成果3) 重み d/2 の訂正可能誤りベクトルの数について別証明

(成果4) 重み d/2+1 の訂正可能誤りベクトルの数を導出

いずれの結果も誤りの単調性を利用

*上記の研究成果は大阪大学藤原融先生との共同研究であり、 まとめて論文として投稿中

以降の発表の流れ

- 扱う問題についてもう少し詳しく
 - 訂正可能・不可能な誤り
 - 扱う問題の別の見方
 - □ 訂正可能な誤り = コセットリーダー
- 誤りの単調性
 - ベクトル間のカバー関係・極小訂正不可能誤り・Larger Half
- 研究成果の詳細
 - 結果の紹介
 - 証明概要
- **■** まとめ

訂正可能・不可能な誤り

- \blacksquare 訂正可能誤り $E^0(C)$ = 最小距離復号で訂正可能な誤り
 - $E_i^0(C) = \{ v \in E^0(C) : w(v) = i \}$
- 訂正不可能誤り $E^1(C) = \{0,1\}^n \setminus E^0(C)$
 - $E_i^1(C) = \{ v \in E^1(C) : w(v) = i \}$
 - $|E_i^0(C)| + |E_i^1(C)| = \binom{n}{i}$, $|E_i^1(C)| = 0$ for i < d/2
 - ullet 2 元対称通信路で最適な復号をしたときの復号誤り率 P_{error}

$$P_{error} = \sum_{i=0}^{n} p^{i} (1-p)^{n-i} |E_{i}^{1}(C)|$$

本研究では $|E_i^{-1}(C)|$ for $i \ge d/2$ を求めることが目標

扱う問題の別の見方

- 符号によるコセット分割
 - (n, k) 線形符号 C によって {0, 1}ⁿ は 2^{n-k} 個のコセットに分割

- シンドローム復号
 - 最小距離復号の一つ
 - y が入力されたとき, $y \in D_i$ ならば $y + v_i$ を出力
 - コセットリーダー = 訂正可能な誤り

コセット分割の例

■ (5, 2) 線形符号 C = { 00000, 11100, 00111, 11011 }

```
D_1 = \{ 00000 \ 11100 \ 00111 \}
                                  11011 }
D_2 = \{ 00001 \ 11101 \ 00110 \ 11010 \}
D_3 = \{ 00010 \ 11110 \ 00101 \ 11001 \}
D_{\Delta} = \{ 00100 \ 11000 \ 00011 \ 111111 \}
D_5 = \{ 01000 \ 10100 \ 10111 \ 10011 \}
D_6 = \{ 10000 \quad 01100 \quad 10111 \}
                                   01011 }
D_7 = \{ 01001 \ 10101 \ 01110 \ 10010 \}
D_8 = \{ 01010 \ 10110 \ 01101 \ 10001 \}
                     \{0, 1\}^5
```

コセット分割の例

■ (5, 2) 線形符号 C = { 00000, 11100, 00111, 11011 }

$$\begin{array}{c} D_1 = \{ \begin{array}{c} 00000 \\ D_2 = \{ \begin{array}{c} 00001 \\ 00001 \\ \end{array} \end{array} & \begin{array}{c} 11100 & 00111 & 11011 \\ 11101 & 00110 & 11010 \\ \end{array} \} \\ D_3 = \{ \begin{array}{c} 00010 \\ 00100 \\ \end{array} & \begin{array}{c} 11110 & 00110 & 11010 \\ \end{array} \} \\ D_4 = \{ \begin{array}{c} 00100 \\ 01000 \\ \end{array} & \begin{array}{c} 11100 & 00011 & 11001 \\ \end{array} \} \\ D_5 = \{ \begin{array}{c} 01000 \\ 10000 \\ \end{array} & \begin{array}{c} 10100 & 10111 & 10011 \\ \end{array} \} \\ D_6 = \{ \begin{array}{c} 10000 \\ 01001 \\ \end{array} & \begin{array}{c} 01100 & 10111 & 01011 \\ \end{array} \} \\ D_8 = \{ \begin{array}{c} 01001 \\ \end{array} & \begin{array}{c} E^0(C) \end{array} & \begin{array}{c} E^1(C) \end{array} \end{array}$$

コセット分割の例

■ (5, 2) 線形符号 C = { 00000, 11100, 00111, 11011 }

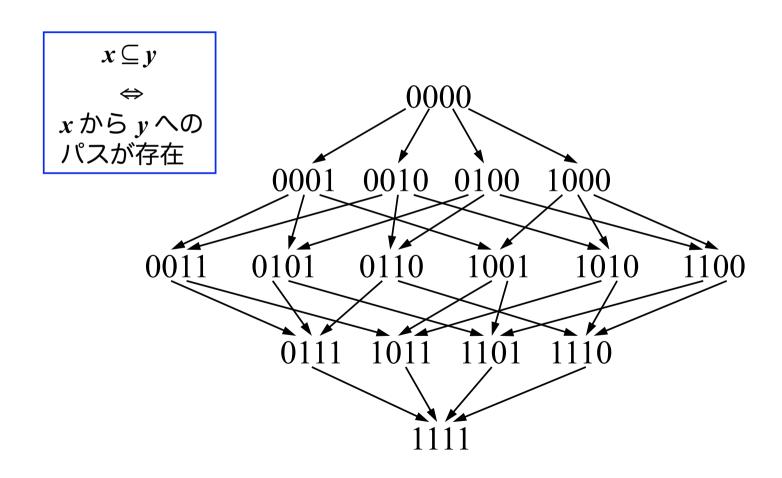
同じコセットに最小重みのベクトルが複数存在することもある

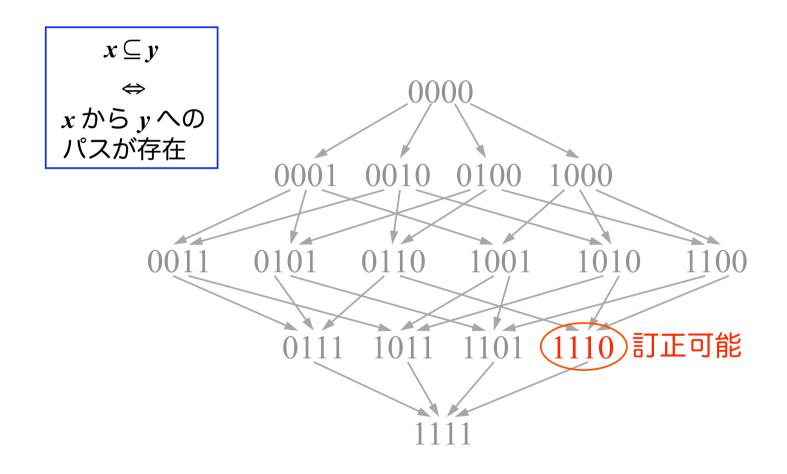
- 最小距離復号では訂正可能誤りに選択の余地がある (同じコセットに最小重みのベクトルが複数) (受信語の最近に複数の符号語)
 - ⇒ 辞書順で最小の誤りを訂正
 - ⇒ 誤りが単調性を持つ [Peterson, Weldon 1972]
- 誤りの単調性:

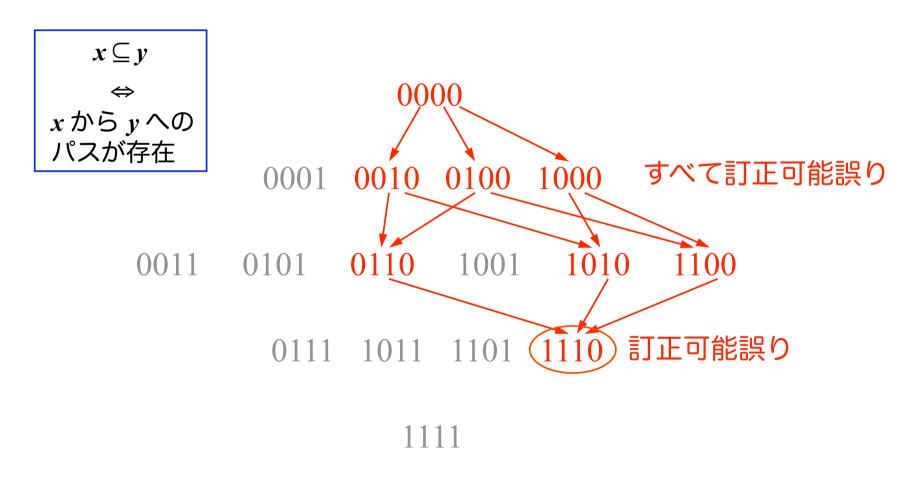
```
x が訂正可能 \Rightarrow x にカバーされる誤りもすべて訂正可能
```

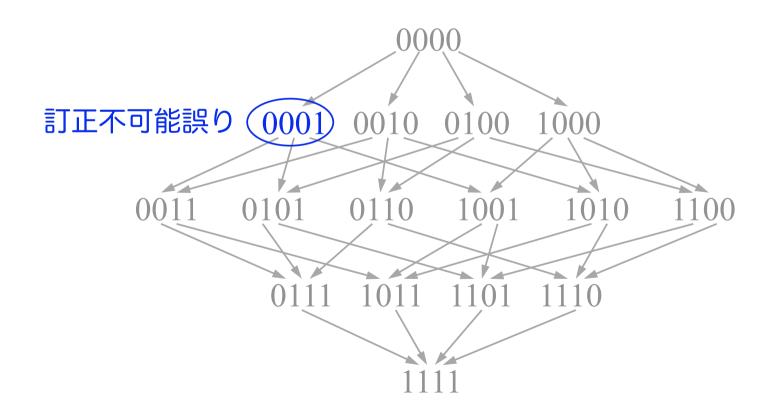
x が訂正不可能 \Rightarrow x をカバーする誤りもすべて訂正不可能

- $x = (x_1, ..., x_n)$ のサポート: $S(x) = \{i : x_i \neq 0\}$
- x が y にカバーされる \Leftrightarrow $S(x) \subseteq S(y)$ \Leftrightarrow $x_i \le y_i$ for all i
 - 以降,簡単のため x⊆yと表記



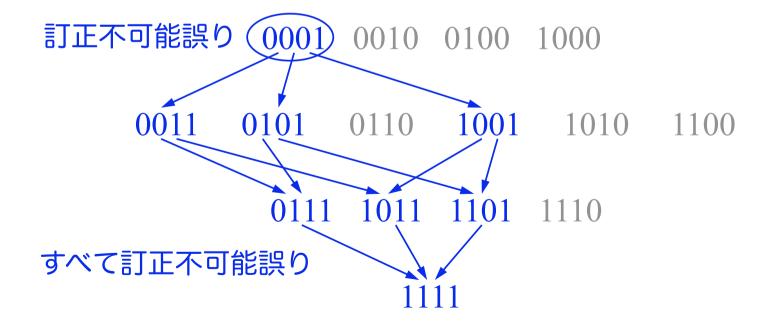






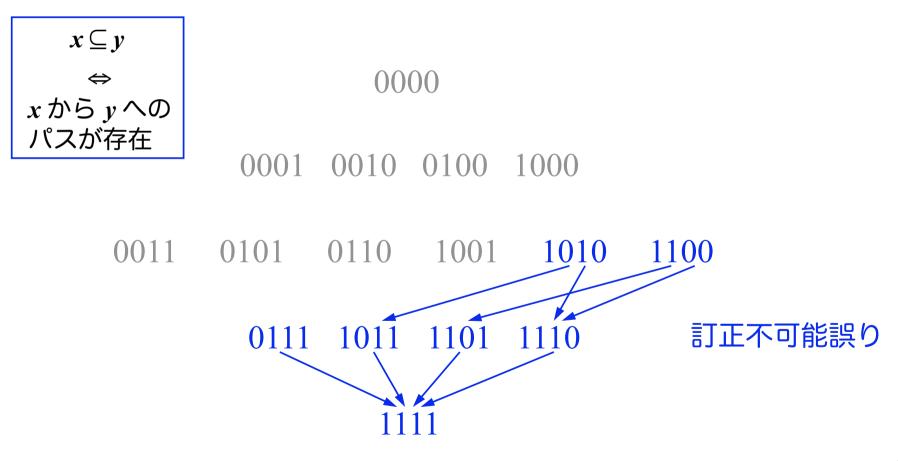
x が訂正可能 $\Rightarrow v \subseteq x$ である v もすべて訂正可能 y が訂正不可能 $\Rightarrow y \subseteq v$ である v もすべて訂正不可能

0000



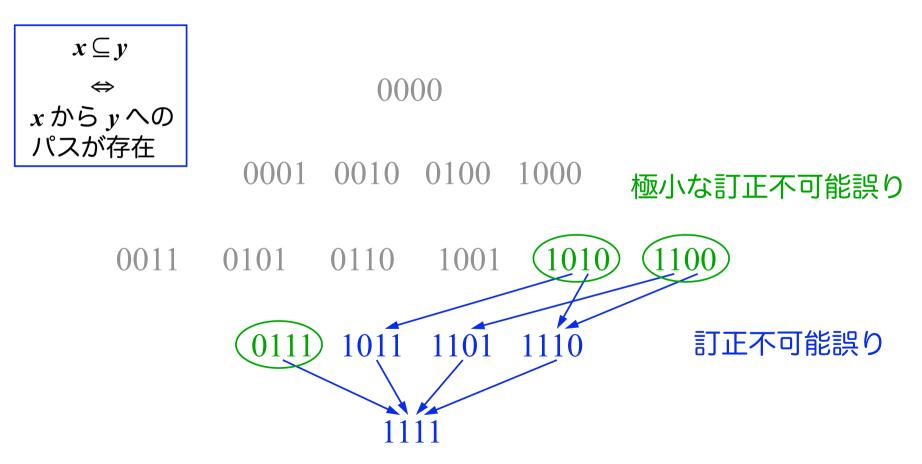
単調性があるとき

- 訂正不可能誤りは $M^1(C)$ によって特徴付けられる
 - M¹(C):カバー(⊆)に関して極小な訂正不可能誤り
 - $M^1(C)$ が決まれば訂正不可能誤りは一意に決まる



単調性があるとき

- 訂正不可能誤りは $M^1(C)$ によって特徴付けられる
 - M¹(C):カバー(⊆)に関して極小な訂正不可能誤り
 - $M^1(C)$ が決まれば訂正不可能誤りは一意に決まる



Larger Half

- 符号語 cの Larger Half; LH(c)
 - $M^1(C)$ を特徴付けるために導入[Helleseth et al. 2005]
 - LH の直感的定義 $LH(c) = \{ v \in \{0,1\}^n : c \text{ によって訂正不可能誤りだとわかる ベクトルの中でカバーに関して極小なもの} \}$
 - 重要な性質

• LH の組み合わせ的構成法

$$(1) \mathbf{v} \subseteq \mathbf{c}$$

$$\mathbf{v} \in LH(\mathbf{c}) \iff (2) \mathbf{w}(\mathbf{c}) \le 2\mathbf{w}(\mathbf{v}) \le \mathbf{w}(\mathbf{c}) + 2$$

$$(3) \begin{cases} l(\mathbf{v}) = l(\mathbf{c}) & \text{if } 2\mathbf{w}(\mathbf{v}) = \mathbf{w}(\mathbf{c}) \\ l(\mathbf{v}) > l(\mathbf{c}) & \text{if } 2\mathbf{w}(\mathbf{v}) = \mathbf{w}(\mathbf{c}) + 2 \end{cases}$$

$$(3) \begin{cases} l(\mathbf{v}) > l(\mathbf{c}) & \text{if } 2\mathbf{w}(\mathbf{v}) = \mathbf{w}(\mathbf{c}) + 2 \end{cases}$$

これまでの研究成果

■ 一般の符号に対して

(成果 1) 重み d/2 の訂正不可能誤りベクトルの数の下界を導出 (成果 2) 重み d/2+1 以上への拡張

■ 1次 Reed-Muller 符号に対して

(成果3) 重み d/2 の訂正可能誤りベクトルの数について別証明

(成果4) 重み d/2+1 の訂正可能誤りベクトルの数を導出

成果1・2・4について以降で紹介

(成果1): 結果(dが偶数の場合)

$$d$$
 が偶数であり $\frac{1}{2} \binom{d}{d/2} > \left[\frac{|C_d| - 1}{2} \right]$ であるとき
$$\frac{1}{2} \binom{d}{d/2} |C_d| - \left[\frac{|C_d| - 1}{2} \right] |C_d| \le \left| E_{d/2}^1(C) \right| \le \frac{1}{2} \binom{d}{d/2} |C_d|$$

 $C_w = \{ C$ で重み w の符号語 $\}$

上界は [Helleseth et al .2005] から

- $n \to \infty$ で $|C_d| / {d \choose d/2} \to 0$ なら上界・下界が漸近的に一致
 - Reed-Muller 符号やランダム線形符号では漸近的に一致

(成果1) : 結果 (d が奇数の場合)

$$d$$
 が奇数であり $\frac{1}{2} \binom{d}{(d+1)/2} > \left[\frac{|C_d|}{2} \right] + \left[\frac{|C_{d+1}|-1}{2} \right]$ であるとき
$$\frac{1}{2} \binom{d}{(d+1)/2} \left(|C_d| + |C_{d+1}| \right) - \left(\left[\frac{|C_d|}{2} \right] + \left[\frac{|C_{d+1}|-1}{2} \right] \right) |C_{d+1}|$$

$$\leq \left| E_{(d+1)/2}^1(C) \right| \leq \frac{1}{2} \binom{d}{(d+1)/2} \left(|C_d| + |C_{d+1}| \right)$$

 $C_w = \{ C$ で重み w の符号語 $\}$

上界は [Helleseth et al .2005] から

- $n \to \infty$ で $|C_{d+1}| / \binom{d}{(d+1)/2} \to 0$ なら上界・下界が漸近的に一致
 - ランダム線形符号では漸近的に一致

(成果1): 証明概要

- $ullet \left| E^1_{\lceil d/2 \rceil}(C) \right|$ を求めたい
- 次の関係が成立

$$M^{1}_{\lceil d/2 \rceil}(C) = LH_{\lceil d/2 \rceil}(C \setminus \{\mathbf{0}\}) = E^{1}_{\lceil d/2 \rceil}(C)$$

[証明]

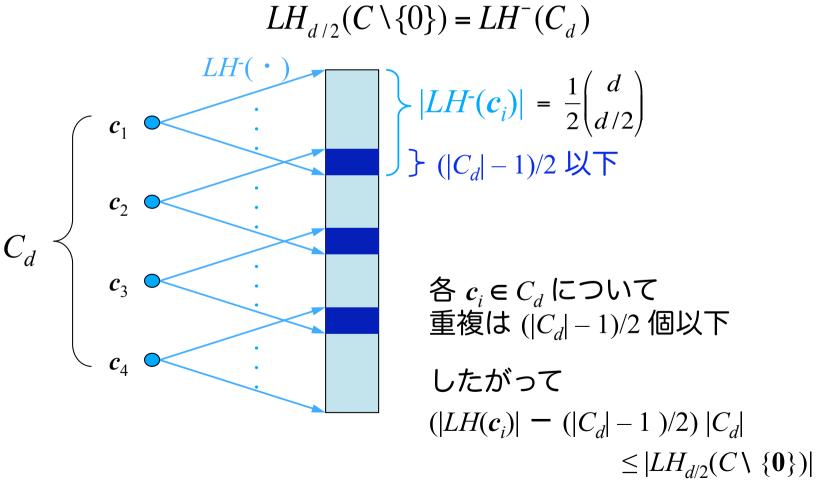
- $M^1(C) \subseteq LH(C \setminus \{\mathbf{0}\}) \subseteq E^1(C)$
- 重み [d/2] は $E^1(C)$ の中で最小の重みであり、その重みをもつ誤りはその他の訂正不可能誤りにカバーされない

$$\Rightarrow M^1_{\lceil d/2 \rceil}(C) = E^1_{\lceil d/2 \rceil}(C)$$

■ $\left|LH_{[d/2]}(C\setminus\{\mathbf{0}\})\right|$ の下界を考えることで $\left|E^1_{[d/2]}(C)\right|$ の下界を 導出

(成果1): 証明概要(dが偶数の場合)

■ |*LH*_{d/2}(*C*\{**0**})| の下界を求める



Reed-Muller 符号への適用

- 符号長 2^m の r 次 Reed-Muller 符号
 - $d = 2^{m-r}$
 - $|C_d| \le (2^{m+1}-2)^r$
- 条件 $\frac{1}{2} \binom{d}{d/2} > \left[\frac{|C_d|-1}{2} \right]$ は r 固定・ $m \to \infty$ で満たされる

条件を満たす r, m

■ また、 $m \to \infty$ のとき

$$|C_d| / {d \choose d/2} \le \frac{(2^{m+1} - 2)^r}{2^{2^{m-r}}} \le 2^{(m+1)r - 2^{m-r}} \to 0$$

なので上界・下界は漸近的に一致

r	m
1	≥ 4
2	≥6
3	≥8
4	≥ 10
5	≥ 11
6	≥ 13
	<u> </u>

ランダム線形符号への適用

- 生成行列 (nk ビット) を確率 2^{-nk} でとってくるランダム 線形符号(のアンサンブル)
 - レート R = k/n をあらかじめ決める
 - n→∞としたときの平均を考える
- dはGilbert-Varshamov bound上にある

■ 重み分布は2項分布にしたがう

$$\left|C_d\right| \approx (2^k - 1) \binom{n}{d} 2^{-n} \approx 2^{n(H(\delta) - 1 + R)} \approx 1, \quad \left|C_{d+1}\right| \approx \left|C_d\right|$$

ランダム線形符号への適用

- 条件は d が偶数のとき $\frac{1}{2} \binom{d}{d/2} > \left[\frac{|C_d|-1}{2} \right] \approx 0$ d が奇数のとき $\frac{1}{2} \binom{d}{(d+1)/2} > \left[\frac{|C_d|}{2} \right] + \left[\frac{|C_{d+1}|-1}{2} \right] \approx 1$ であり、 $d \approx \delta_{\mathrm{GV}} n$ なので満たされる
- $|C_d| / \binom{d}{d/2} \to 0, |C_{d+1}| / \binom{d}{(d+1)/2} \to 0$

なので上界・下界は漸近的に一致

(成果2):結果

成果1と同様の議論から

 $|LH_i(C)| \le |E_i^1(C)|$ であるため下界を与えている

大きな i に対して

- 下界のための条件が厳しい
- 弱い下界である
 - あくまで $|LH_i(C)|$ に対する下界であり、 i が大きいと $|LH_i(C)|$ と $|E_i^{-1}(C)|$ の差が広がる

1次 Reed-Muller 符号 RM_m

- $(2^m, m+1)$ 符号で最小距離 $d=2^{m-1}=n/2$
 - 次元 k = m+1 は小さいが、最小距離が n/2 と非常に大きい
 - non-trivial な符号の中では構造が非常にシンプル
 - □ ある論文では、RM_m は符号理論で最も研究されてきた符号と紹介
- 各符号語は m 変数の線形ブール関数と一対一に対応
 - r 次 Reed-Muller 符号の符号語は r 次ブール関数に対応
- RM_m の重み *i* の訂正可能誤りの数
 - ⇒ 非線形性が i のブール関数の数
 - 関数 f の非線形性: f が線形関数からどのくらい離れているか
 - 関数の非線形性は、暗号システム(対称鍵暗号、ストリーム暗号) の安全性指標として重要

(成果4) : 結果

 $m \ge 5$ の 1 次 Reed-Muller 符号 $(n = 2^m)$ に対し

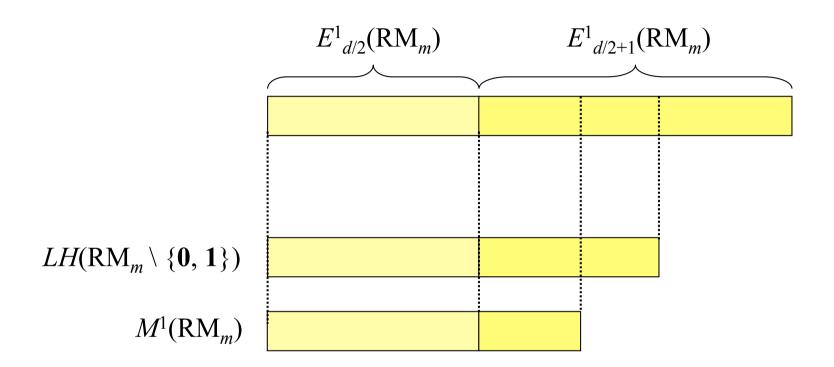
$$|E_{d/2+1}^{1}(RM_{m})| = 4(2^{m} - 1)(2^{m-3} + 1)\begin{pmatrix} 2^{m-1} \\ 2^{m-2} + 1 \end{pmatrix} - (4^{m-2} + 3)\begin{pmatrix} 2^{m} \\ 3 \end{pmatrix}$$

■ 重み d/2+1 の訂正可能な誤りベクトルの数は

$$|E_{d/2+1}^{0}(RM_{m})| + |E_{d/2+1}^{1}(RM_{m})| = \begin{pmatrix} 2^{m} \\ 2^{m-2} + 1 \end{pmatrix}$$

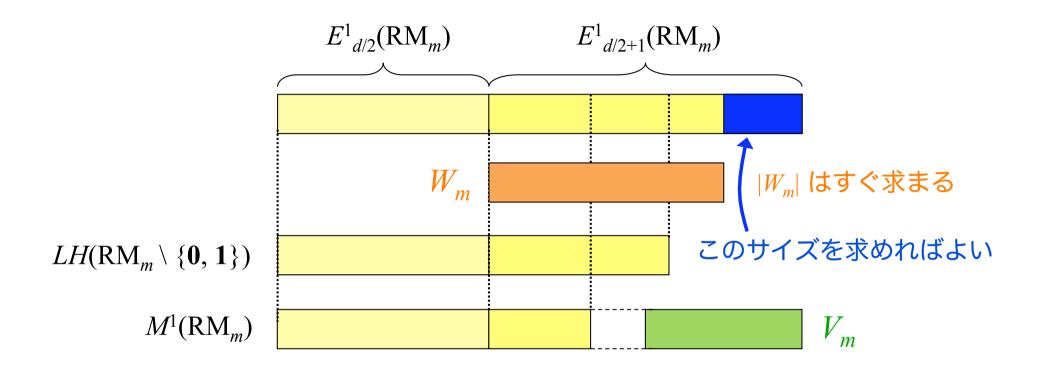
から導出可能

(成果4): 証明概要



■ $M^1(C) \subseteq LH(C \setminus \{0\}) \subseteq E^1(C)$ の関係を RM_m について調べると上記の関係が成立

(成果4): 証明概要



- $W_m = \{ v \in \{0, 1\}^n : v \subseteq c \text{ for } c \in RM_m \setminus \{0,1\}, w(v) = d/2+1 \}$ を考える
- に含まれる訂正不可能誤りは極小でない
 - ⇒ 重み d/2 の訂正不可能誤りに重み 1 のベクトルを足した形
 - \Rightarrow そのようなベクトル集合 V_m を構成し $|V_m \setminus W_m|$ を求める

成果4の結果の考察

(成果4) 訂正可能な重み d/2+1 の誤りベクトルの数

■ 数値例(符号長 2^m)

\overline{m}	n	k	訂正可能誤り数	訂正不可能誤り数
5	32	6	21,288,320	6,760,480
6	64	7	1.378×10^{15}	1.238×10^{12}
7	128	8	4.299×10^{30}	1.535×10^{22}
8	256	9	5.625×10^{61}	7.938×10^{41}
9	512	10	1.329×10^{124}	7.605×10^{80}

■ *m* = 9 のとき、 訂正不可能な誤りは 10⁴⁴ 個に 1 個の割合

まとめ

- 誤り訂正符号の誤りの単調性を利用した訂正能力分析
 - 重み d/2 以上の訂正可能な誤りベクトルの数について研究
 - 誤りの単調性 (Larger Half) を利用

■ 研究成果

- 一般の符号に対して (成果1)重み d/2 の訂正不可能誤りベクトルの数の下界を導出 (成果2)重み d/2+1 以上への拡張
- 1次 Reed-Muller 符号に対して (成果3)重み d/2 の訂正可能誤りベクトルの数について別証明 (成果4)重み d/2+1 の訂正可能誤りベクトルの数を導出

■ 今後の研究の方向

- 一般の符号における d/2+1 以上の場合の下界の改善
- その他の符号(2次以上 Reed-Muller 符号・BCH符号)への適用

参考文献 (1/2)

[Berlekamp, Welch 1972]

E.R. Berlekamp, L.R. Welch, "Weight distributions of the cosets of the (32,6) Reed-Muller code," *IEEE Trans. Inf. Theory*, 1972.

[Charpin 1994]

P. Charpin, "Weight distributions of cosets of two-error-correcting binary BCH codes, extended or not", *IEEE Trans. Inf. Theory*, vol. 40, no. 5, pp. 1425–1442, Sept. 1994.

[Charpin, Helleseth, Zinoviev 2006]

P. Charpin, T. Helleseth, and V.A. Zinoviev, "The coset distribution of triple-error-correcting binary primitive BCH codes," *IEEE Tran. Inf. Theory*, vol. 52, no. 4, pp. 1727–1732, Apr. 2006.

[Helleseth, Klove 1997]

T. Helleseth, T. Kløve, "The Newton radius of codes," *IEEE Trans. Inf. Theory*, 1997.

[Helleseth, Kove, Levenshtein 2005]

T. Helleseth, T. Kløve, and V. Levenshtein, "Error-correction capability of binary linear codes," *IEEE Trans. Inf. Theory*, 2005.

参考文献 (2/2)

[Peterson, Weldon 1972]

W.W. Peterson and E.J. Weldon, Jr., Error-Correcting Codes, 2nd Edition, MIT Press, 1972.

[Poltyrev 1994]

G. Poltyrev, "Bounds on the decoding error probability of binary linear codes via their spectra," *IEEE Trans. Inf. Theory,* 1994.

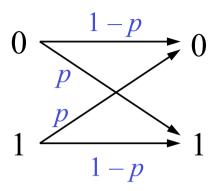
[Wu 1998]

C.K. Wu, "On distribution of Boolean functions with nonlinearity ≤ 2 n -2", *Australasian Journal of Combinatorics*, vol. 17, pp. 51–59, Mar. 1998.

通信路モデル

2元対称通信路

- 各ビット毎に 0 と 1 を一定確率で反転
- 離散通信路
 - 受信語 $y \in \{0,1\}^n$



加法的白色ガウス雑音(AWGN)通信路

- 各ビット毎に白色ガウス雑音を付加
- 連続通信路
 - 受信語 $y \in \mathbb{R}^n$

