

b -シンボル読出通信路における球充填限界と G-V 限界

Sphere Packing Bound and Gilbert-Varshamov Bound
for b -Symbol Read Channels

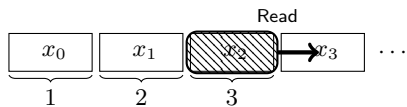
Song Seunghoan

大阪大学 基礎工学部 情報科学科 藤原研究室

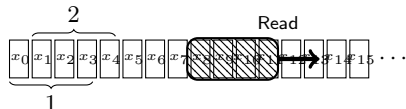
2017 年 2 月 21 日

研究背景 b -シンボル読出通信路 (b -Symbol Read Channel)

- ▶ 通信路モデル設定の背景：記憶装置において、書き込み精度が高く、読み出し精度が低いとき、複数 (b 個) のシンボルが一度に読み出される



< 従来の通信路 >



< 4-シンボル読出通信路 >

誤りがなければ、 $[x_0, x_1, \dots, x_{n-1}] \rightarrow [x_0x_1\dots x_{b-1}, x_1x_2\dots x_b, \dots, \underline{x_{n-1}x_0\dots x_{b-2}}]$

- ▶ $b = 1$ のモデル: 従来の通信路
- ▶ $b = 2$ のモデル: 2011 年, Cassuto らによって提案 (シンボルペア読出通信路)¹
- ▶ b が一般化されたモデル: 2016 年, Yaakobi らによって提案 (b -シンボル読出通信路)²

¹Y. Cassuto and M. Blaum, "Codes for symbol-pair read channels," IEEE Trans. Inf. Theory, vol. 57, no. 12, pp. 8011-8020, Dec. 2011.

²Eitan Yaakobi *et.al.*, "Constructions and Decoding of Cyclic Codes Over b -Symbol Read Channels" IEEE Trans. Inf. Theory, vol. 62, no. 4, pp. 1541- 1551, Apr. 2016.

研究目的と主な結果

符号の評価: 「伝送効率」と「誤り訂正能力」

$$\left\{ \begin{array}{l} \text{伝送効率: 符号化率 } R(C) \\ \text{誤り訂正能力: 最小距離 } d \end{array} \right\} \dots \text{トレードオフの関係がある.}$$

$b \leq 2$ の場合, 最小距離 d を固定したとき, 符号化率の上界と下界が既知である.

研究目的

- ▶ 最小距離 d を固定したとき, b -シンボル読出通信路における符号化率の上界と下界を求める.
- ▶ b -シンボル読出通信路上の符号が従来の通信路上の符号と比べて, 利点があるかを調べる.

主な結果

1. 最小距離 d を固定したとき, b -シンボル読出通信路における符号化率の上界 (球充填限界) と下界 (G-V 限界) の導出
2. G-V 限界の漸近的分析により, 以下を証明
 - 2.1 b -シンボル読出通信路上には, 従来の通信路上の符号と比べて, 高い符号化率の符号が存在する
 - 2.2 符号長 n と最小距離 d の比率に依存して, 符号化率を最適にする b が決まる

目次

- ▶ b -シンボル球充填限界・G-V 限界
 - ▶ 距離空間における球充填限界・G-V 限界
 - ▶ b -シンボル距離
 - ▶ b -シンボル球面のサイズ導出
- ▶ 漸近的 b -シンボル G-V 限界
 - ▶ 距離空間における漸近的 G-V 限界
 - ▶ 漸近的 b -シンボル G-V 限界の導出
 - ▶ b と漸近的 G-V 限界の関係性分析

距離空間における球充填限界・G-V 限界

球のサイズ $|\mathcal{B}_r(\vec{x})|$ から球充填限界・G-V 限界を導出される。

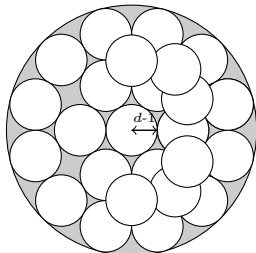
G-V 限界・球充填限界 (Gilbert-Vashamov Bound, Sphere packing bound)

距離空間 (Σ^n, D) に対して, 最小距離が $d = \min_{\vec{x}, \vec{y} \in C, \vec{x} \neq \vec{y}} D(\vec{x}, \vec{y})$ のとき,

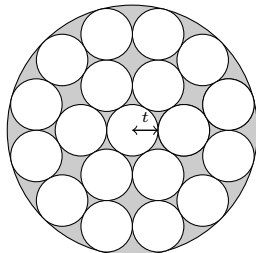
$$\underbrace{\frac{1}{n} \log_q \frac{q^n}{|\mathcal{B}_{d-1}(\vec{x})|}}_{\text{G-V 限界}} \leq R(C) \leq \underbrace{\frac{1}{n} \log_q \frac{q^n}{|\mathcal{B}_{\lfloor \frac{d-1}{2} \rfloor}(\vec{x})|}}_{\text{球充填限界}}$$

(満たす符号の存在)

(任意の符号に対して成立)



<G-V 限界>



<球充填限界>

本研究での限界導出

- ▶ 球のサイズ $|\mathcal{B}_r(\vec{x})|$ から球充填限界・G-V 限界を導出
 - ▶ 球のサイズ $|\mathcal{B}_r(\vec{x})|$ は、球の中心によらないため、 $|\mathcal{B}_r(\vec{0})|$ を求める。
- ▶ b -シンボル距離空間での球のサイズ $|\mathcal{B}_r(\vec{0})|$ の計算は、複雑
 - ▶ b -シンボル距離空間： b -シンボル読出通信路上の距離空間
- ▶ 従って、本研究では、 $\mathcal{B}_r(\vec{0})$ を計算が可能な形へ分割する。
 1. 球 $\mathcal{B}_r(\vec{0})$ を球面 $\mathcal{S}_h(\vec{0})$ に分割する ($h = 0, \dots, r$)。

$$|\mathcal{B}_r(\vec{0})| = \sum_{h=0}^r |\mathcal{S}_h(\vec{0})|$$

2. 球面 $\mathcal{S}_h(\vec{0})$ を、より分割してサイズを求める。

b -シンボル距離 (b -Symbol Distance) ²

定義 b -シンボル読出ベクトル (b -Symbol Read Vector)

$\vec{x} = [x_0, x_1, \dots, x_{n-1}] \in \Sigma^n$ に対して,

$$\pi_b(\vec{x}) \triangleq [x_0x_1\dots x_{b-1}, x_1x_2\dots x_b, \dots, x_{n-1}x_0\dots x_{b-2}]$$

$\pi_b(\vec{x})$ の i 番目の要素を $\pi_b(\vec{x})_i$ と表す.

定義 b -シンボル距離 (b -Symbol Distance)

空間 Σ^n 上の距離 D_b を以下のように定義する. $\vec{x}, \vec{y} \in \Sigma^n$ に対して

$$D_b(\vec{x}, \vec{y}) \triangleq |\{i | \pi_b(\vec{x})_i \neq \pi_b(\vec{y})_i\}|$$

- ▶ b -シンボル距離空間 (Σ^n, D_b) における球面・球を $\mathcal{S}_h^{(b)}(\vec{x})$, $\mathcal{B}_r^{(b)}(\vec{x})$ と表記

²Eitan Yaakobi *et.al.*, "Constructions and Decoding of Cyclic Codes Over b -Symbol Read Channels" IEEE Trans. Inf. Theory, vol. 62, no. 4, pp. 1541- 1551, Apr. 2016.

$\vec{0}$ からの b -シンボル距離

$|\mathcal{S}_h^{(b)}(\vec{0})|$ を求める.

$\mathcal{S}_h^{(b)}(\vec{0})$ は、 $\vec{0}$ からの b -シンボル距離が h である語の集合である.

従って、 \vec{y} の $\vec{0}$ からの距離 $w_b(\vec{y}) = D_b(\vec{0}, \vec{y})$ について考える.

定義 語のパラメータの定義

ある語 $\vec{y} \in \Sigma^n$ を以下のパラメータ w, z_k, Z, L 定義する.

- ▶ w : \vec{y} における1の数
- ▶ z_k : \vec{y} における長さ k の0のランの数
- ▶ $Z = \sum_{k=1}^{b-2} k z_k$
- ▶ $L = \sum_{k=b-1}^n z_k$

補題 $\vec{0}$ からの b -シンボル距離の計算

\vec{y} の $\vec{0}$ からの b -シンボル距離 $w_b(\vec{y}) = w + Z + (b-1)L$.

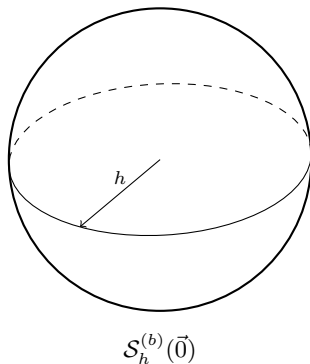
b -シンボル球面のサイズ $|\mathcal{S}_h^{(b)}(\vec{0})|$ の計算

球面のサイズの直接計算は複雑であるため、分割する。

$\mathcal{S}_h^{(b)}(\vec{0})$ の分割 (球面上の語の分類)

1. 球面 $\mathcal{S}_h^{(b)}(\vec{0})$ 上の語 \vec{y} を L と Z で分類
 - ▶ $h = w + Z + (b-1)L$ ($\vec{0}$ からの距離の計算より)
 - ▶ しかし、 L と Z では、まだサイズの計算が困難
2. 語 \vec{y} を L と $\vec{z} = (z_1, \dots, z_{b-2})$ で分類
($Z = \sum_{k=1}^{b-2} k z_k$)
 - ▶ $\mathcal{S}_h^{(b)}(L, \vec{z})$: パラメータ h, L, \vec{z} を満たす語の集合
 - ▶ $|\mathcal{S}_h^{(b)}(L, \vec{z})|$ は、計算可能。

$$|\mathcal{S}_h^{(b)}(\vec{0})| = \sum_{(L, \vec{z}) \in \mathcal{K}} |\mathcal{S}_h^{(b)}(L, \vec{z})|$$



$$|\mathcal{S}_h^{(b)}(L, \vec{z})| = \frac{n}{w} \binom{w}{L, z_1, \dots, z_{b-2}} \binom{n-h+L-1}{L-1} (q-1)^w$$

漸近的 b -シンボル G-V 限界

漸近的 G-V 限界 (Asymptotic Gilbert-Vashamov Bound) とは,

- ▶ 符号長 n を無限にしたときの G-V 限界
- ▶ 最小距離 d も n に比例して無限になるため, 比率 $\delta = d/n$ に対する下界

定理 漸近的 b -シンボル G-V 限界

fractional minimum distance $\delta (= d_b/n)$ でかつ以下の条件を満たす q 元符号 $C_\delta \subset \Sigma^n$ が存在する.

$$R(C_\delta) > 1 - H_q(\delta/b) - \delta \log_q b$$

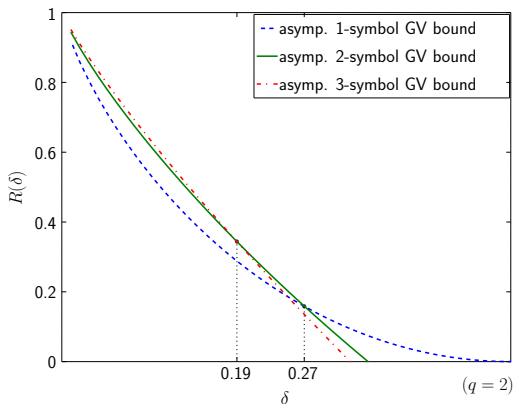
- ▶ $b = 1$ のとき, 従来の通信路上の式と一致
- ▶ $b = 2$ のとき, シンボルペア読出通信路上の式と一致

*符号化率 $R(C)$ とエントロピー $H(p)$ の定義

$$R(C) \triangleq \frac{\log |C|}{n}$$

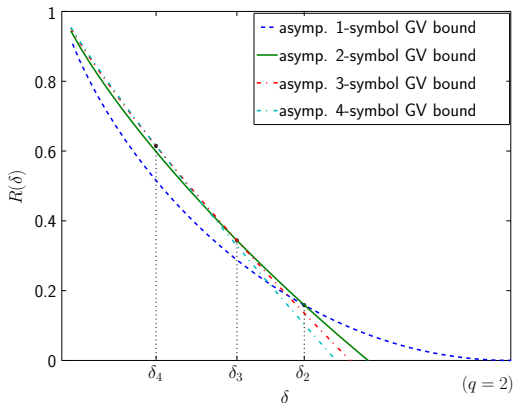
$$H_q(p) \triangleq p \log_q (q-1) + p \log_q \frac{1}{p} + (1-p) \log_q \frac{1}{1-p}$$

漸近的 b -シンボル G-V 限界の分析 1



- ▶ $\delta \in (0.27, 0.5)$ に対して, 1-シンボル限界が最適
- ▶ $\delta \in (0.19, 0.27)$ に対して, 2-シンボル限界が最適
- ▶ $\delta \in (0, 0.19)$ に対して, 3-シンボル限界が最適

漸近的 b -シンボル G-V 限界の分析 2



- ▶ $\delta \in (\delta_{b+1}, \delta_b)$ では, b -シンボル符号が最適な符号化率を達成
- ▶ b を大きくすれば, $\delta \in (0, \delta_b)$ で b -シンボル符号が最適な符号化率を達成

まとめ・今後の課題

まとめ

- ▶ b -シンボル読出通信路における「球充填限界」, 「G-V 限界」, 「漸近的 G-V 限界」の導出した
- ▶ fractional minimum distance δ に依存して, 漸近的 G-V 限界を最適にする b が決まる
- ▶ b を大きくすれば, より高い符号化率を持つ b -シンボル符号が存在する

今後の課題

1. 漸近的 G-V 限界の改善: より tight な近似が必要
2. $|\mathcal{S}_h^{(b)}(\vec{x})| = \sum_{(L, \vec{z}) \in \mathcal{K}} |\mathcal{S}_h^{(b)}(L, \vec{z})|$ の簡略化
3. b -シンボル符号の誤り訂正能力の分析: 誤り率と最小距離を考慮
4. b -シンボル完全符号の存在性
5. G-V 限界を満たす b -シンボル符号の設計

補題 : $|\mathcal{S}_h^{(b)}(L, \vec{z})|$ の計算

補題 $|\mathcal{S}_h^{(b)}(L, \vec{z})|$

$$|\mathcal{S}_h^{(b)}(L, \vec{z})| = \frac{n}{w} \binom{w}{L, z_1, \dots, z_{b-2}} \binom{n-h+L-1}{L-1} (q-1)^w$$

ただし, $w = h - Z - (b-1)L$.

- ▶ $b = 1$ のとき, ハミング距離を用いた式と一致
- ▶ $b = 2$ のとき, シンボルペア距離を用いた式と一致

証明の概略

1. $\mathcal{S}_h^{(b)}(L, \vec{z})$ の中で, 先頭の記号が 0 でないベクトル \vec{y} の数を数える
2. 1. で数えたベクトルを n 回巡回シフトし, 重複を除く

$\vec{0}$ からの b -シンボル距離

定義 語のパラメータの定義

ある語 $\vec{y} \in \Sigma^n$ を以下のパラメータ w, z_k, Z, L 定義する.

- ▶ w : \vec{y} における1の数
- ▶ z_k : \vec{y} における長さ k の0のランの数
- ▶ $Z = \sum_{k=1}^{b-2} k z_k$
- ▶ $L = \sum_{k=b-1}^n z_k$

補題 $\vec{0}$ からの b -シンボル距離の計算

\vec{y} の $\vec{0}$ からの b -シンボル距離 $w_b(\vec{y}) = w + Z + (b-1)L$.

$$\begin{array}{rcccccccccccc} \vec{y} = & & 1 & & 1 & & 1 & & 0 & & 0 & & 1 & & 1 & & 1 & & 0 & & 0 & & 0 & & 0 \\ \pi_3(\vec{y}) = & & 111 & & 110 & & 100 & & 001 & & 011 & & 111 & & 110 & & 100 & & 000 & & 000 & & 001 & & 011 \end{array}$$

$$\begin{aligned} w_3(\vec{y}) &= (\text{成分が } 000 \text{ でない } \pi_3(\vec{y}) \text{ の座標数}) \\ &= (\text{成分が } 1 \text{ である } \vec{y} \text{ の座標数}) + (\text{長さ } b-2 \text{ 以下の } 0 \text{ のランの座標数}) \\ &\quad + 2 \times (\text{長さ } b-1 \text{ 以上の } 0 \text{ のランの数}) \\ &= w + Z + (b-1)L \end{aligned}$$