修 士 論 文 の 和 文 要 旨

研究科・専攻	大学院 情報理工 学研究科　情報･ネットワーク工学専攻　博士前期課程		
氏　　名	濱田 寛也	学籍番号	1731127
論 文 題 目	(r, δ)-Locally Repairable 符号の最小距離及び次元の評価		

要　旨

分散ストレージシステムは，ノードと呼ばれる複数の小さいストレージを一つの大きなストレージとして扱うシステムである．現在様々な消失訂正符号が分散ストレージシステムで用いられているが，扱うデータの巨大化に伴いストレージの容量効率だけでなく，破損ノードの修復に対するデータの通信量，データの読み書き，読み込むノード数などのオーバーヘッドを減らすことが求められている．既存の消失訂正符号はそのような要求に対して最適化されていなかった．

Locally repairable 符号 (LRC 符号) は，このような要求を満たすために考案された符号のクラスである．局所性(r,δ) を持つ符号は，任意の $\delta-1$ 個の消失シンボルをパンクチャ符号の消失訂正により高々 r 個の他のシンボルから局所的に復元することができる．この符号のパラメータのトレードオフ関係は Singleton 限界の一般化の形で与えられている． $q > n$ のときこの上界を達成する様々な符号の構成法があるが，実装の容易さやバイナリデータを扱う分散ストレージシステムへの対応などを考えれば， $q < n$ に対して限界式とそれを達成する符号の構成法を考えることも重要である．

本論文では，既知の符号の上界を利用する手法により，局所性 (r, δ) を持つ符号の上界を，位数 q の関数として与える．本研究で得られた上界は，既知の上界より良い上界となっており，特に q が小さいとき大きく改善された上界を得ることができる．局所性 (r, δ) を持つ符号の任意のパラメータに対する上界は他に知られておらず，本論文で与える上界は現在知られている上界の中で最も良い上界である．

平成30年度　修士学位論文

(r, δ)-Locally Repairable符号の最小距離及び次元の評価

電気通信大学 大学院 情報理工学研究科
博士前期課程 情報・ネットワーク工学専攻
1731127　濱田 寛也
指導教員　八木 秀樹 准教授　川端 勉 教授

提出　平成31年1月28日

目 次

第 1 章

まえがき

分散ストレージシステムは，ノードと呼ばれる複数の小さいストレージを一つの大きなストレージとして扱うシステムである．このシステムは，データを各ノードに分散して保存することによって，ノードの破損からデータの保護をしたり秘密分散の機能を持たせたりすることができる．現在運用されている分散ストレージシステムは，主にデータの複製を複数ノードに配置するミラーリングと呼ばれる方法や，**消失訂正符号**によって冗長度を持たせる方法などによりデータの信頼性を保っている．データの複製は実装が容易であり，ノードの破損に対しても複製を読み込むだけで復元ができるため，データの信頼性を高める点から優れている方法であるが，巨大なデータの複製には大容量のストレージが必要となる．一方で消失訂正符号を運用するシステムは，一定の信頼性を保ちつつミラーリングよりもストレージの容量効率が良くなっている．現在様々な消失訂正符号が分散ストレージシステムで用いられているが，扱うデータの巨大化に伴いストレージの容量効率だけでなく，破損ノードの修復に対するデータの通信量，データの読み書き，読み込むノード数などのオーバーヘッドの削減が求められている．既存の消失訂正符号はそのような要求に対して最適化されていなかった．

符号長 n，情報記号数 k，最小 Hamming 距離 d の符号は，Singleton 限界を等号で達成するとき maximum distance separable (MDS) 符号と呼ばれ任意の $n-k=d-1$ 個以下の消失シンボルは，他の k 個のシンボルを読み取ることで全て復元可能である．よく知られた MDS 符号として Reed-Solomon 符号や繰り返し符号がある．消失訂正符号の構造は，分散ストレージシステムの構造と対応付けられる．例えば，MDS 符号を運用する分散ストレージシステムは，n 個のノードから構成され任意の $n-k$ 個以下の破損ノードを他の k 個のノードのデータを読み込むことで復元できる．文献 [1] では，MDS 符号を基に構成される新たな符号を運用する分散ストレージシステムが考えられている．この符号は Pyramid 符号と呼ばれ，MDS と比較して冗長度が増えている代わりに特定の消失パターンに対して復元過程で読み込まれるシンボル数を一定数に抑えることができる．Pyramid 符号を分散ストレージシステムで運用することで，破損ノードが少数の場合に復元のために読み込む平均ノード数を減らすことができることも確かめられている．

Pyramid 符号のように特定の消失パターンに対して，高々 r 個のシンボルから消失シンボルを復元できる符号を**局所性** r を持つ符号と呼ぶ．MDS 符号が局所性 $r = k$ を持つ符号であることは明らかである．特に r が k より小さい時，局所性を持つ符号は **locally repairable 符号** (LRC 符号) と呼ばれ，Gopalan *et al.* [2] や Papailiopoulos *et al.* [3] によって符号のクラスが確立された．これらの文献で与えられた LRC 符号は，任意の 1 個の消失シンボルを高々 r 個の他のシンボルから局所的に復元することが可能である．文献 [2] では，符号パラメータのトレードオフ関係を Singleton 限界の拡張の形で次のように与えている．

$$d \leq n - k - \left\lceil \frac{k}{r} \right\rceil + 2. \tag{1.1}$$

この上界は一般に Singleton-like 限界と呼ばれ Pyramid 符号が，この上界を等号で達成することも同様に示されている．文献 [1] で示されたように局所性 r を持つ符号を運用することで，分散ストレージシステムに良い性能がもたらされることが期待できるが，システム上で他に起こりうる様々な問題を考慮して，局所性 r を持つ符号に実用面で有用と考えられる特性を付加した符号クラスの拡張が盛んに研究されている．

局所性 (r, δ) を持つ符号は局所性 r の拡張の一つであり，Kamath *et al.* [4] によって提案された．局所性 (r, δ) を持つ符号は，任意の $\delta - 1$ 個の消失シンボルをパンクチャ符号の消失訂正により高々 r 個の他のシンボルから局所的に復元することができる．つまり，$\delta = 2$ は文献 [2] で提案された局所性 r を持つ符号である．また局所性 (r, δ) を持つ符号は，複数シンボルの局所的復元だけでなくパンクチャ符号の誤り訂正機能により局所的誤り訂正も可能である．文献 [4] では，この符号のパラメータのトレードオフ関係も Singleton-like 限界の一般化の形で与えられている．$q > n$ のときこの上界を達成する様々な符号の構成法がある．文献 [4] では，Pyramid 符号や parity-splitting と呼ばれる方法で構成される符号がバウンドを達成する最適な符号であることが示されている．文献 [10] では Reed-Solomon 符号に類似した構造を持つ最適な LRC 符号の構成法が示されている．

一方で，実装の容易さやバイナリデータを扱う分散ストレージシステムへの対応などを考えれば，位数の小さい有限体上で定義される符号に対する限界式とそれを達成する符号の構成法を考えることも重要である．符号長より小さい位数 $q < n$ に対して Singleton 限界より良い上界があることが知られているように，局所性 r を持つ符号に対して，Singleton-like 限界よりも良い上界が Cadambe and Mazumdar [5] によって与えられている．この上界は一般に C-M 限界と呼ばれ，Simplex 符号が局所性 $r = 2$ を持つ上界を等号で達成する最適な符号であることも同様に示されている．既存の代数的符号と局所性 r の関係は，文献 [6] で特に 2 元符号の場合に詳しく議論されており，C-M 限界を達成しないが最適な構造を持つ既存の代数的符号が多く挙げられている．文献 [7] では，局所性 r を持ち Singleton-like 限界を達成する最適な 2 元符号の構造の特徴づけや，C-M 限界を達成する最小距離 $d = 4$ の 2 元符号のパリティ検査行列が与えられている．文献 [8] では，2 元

LRC 符号の上界を sphere-packing と呼ばれる手法を用いて与えている．この上界は最適化問題の解として与えられており，簡単に計算可能な一部のパラメータに対しては C-M 限界より良い上界を与えることが示されている．この上界を達成する符号の構成法も同様に与えられている．

一方で，位数の小さい有限体上で定義される局所性 (r, δ) を持つ符号に対する議論は多くない．Agarawal *et al.* [9] は，局所性 (r, δ) を持つ符号の次元の上界を位数 q の関数として与えている．この上界は，任意の LRC 符号のパンクチャ符号の次元を，既知の符号の上界の凸性を利用して評価している．既知の符号の上界として Hamming 限界や Plotkin 限界を用いることで $q < n$ に対して文献 [4] で与えられた上界より良い上界が得られることもあるが，一般のパラメータに対しての比較は詳細に行われておらず，任意のパラメータに対して上界を等号で達成する最適な符号の構成法も知られていない．局所性 (r, δ) を持つ符号に対しても，C-M 限界や sphere-packing 限界が存在し，文献 [4] で示された Singleton-like 限界よりよい上界が得られることが予想できるが，局所性 r を持つ符号と異なりパンクチャ符号の最小距離を含めた議論を必要とするため，局所性 r を持つ符号に対する証明手法は容易に拡張できない．

本論文では，文献 [9] と同様に既知の符号の上界を利用する手法により，局所性 (r, δ) を持つ符号の上界を位数 q の関数として与える．この上界は，C-M 限界に近い形で与えられ一部のパラメータに対しては C-M 限界の拡張になっている．加えて本論文で得られた上界は，既知の上界 [4],[9] より良い上界となっており，特に位数 q が小さいときに大きく改善された上界を得ることができる．また，任意の位数 q を持つ有限体上で定義される局所性 (r, δ) を持つ符号に対する上界は他に知られていない．最後に Hamming 符号の直積による符号とその部分符号が上界を達成することを示す．

本論文の構成は以下の通りである．第 2 章で表記と LRC 符号の定義を行い，既存研究で与えられている上界を紹介する．第 3 章で主定理として次元と最小距離の上界，符号長の下界を与える．第 4 章で主定理から得られる結果の考察と数値計算による比較を行い，簡単な符号構成により一部のパラメータに対して上界が達成可能であることを示す．

第 2 章

準備

本章では，準備として表記の定義と LRC 符号に関する既存研究を紹介する．

2.1 表記の定義

- 整数 $x(\neq 0)$ と y に対して，$\lceil y/x \rceil$ は y/x 以上の最小の整数．$\lfloor y/x \rfloor$ は y/x 以下の最小の整数．
- 整数 $x(\neq 0)$ と y に対して，$y \bmod x$ は y の x による剰余．
- 自然数 x に対して，$[x]$ は整数の集合 $\{1, \ldots, x\}$．
- 集合 $\mathcal{S}$ に対して，$|\mathcal{S}|$ は集合の濃度．
- 素数冪 q に対して，$\mathbb{F}_q$ は位数 q の有限体．
- ベクトル $\boldsymbol{c} = (c_1, c_2, \ldots, c_n) \in \mathbb{F}_q^n$ と集合 $\mathcal{S} = \{j_1, j_2, \ldots, j_{|\mathcal{S}|}\} \subseteq [n]$ に対して，$\boldsymbol{c}_{\mathcal{S}} \triangleq (c_{j_1}, c_{j_2}, \ldots, c_{j_{|\mathcal{S}|}}) \in \mathbb{F}_q^{|\mathcal{S}|}$．
- 符号長 n の符号 $\mathcal{C}$ と部分集合 $\mathcal{S} \subseteq [n]$ に対して，$\mathcal{C}_{\mathcal{S}} \triangleq \{\boldsymbol{c}_{\mathcal{S}} \mid \boldsymbol{c} \in \mathcal{C}\}$．

$\lfloor y/x \rfloor = \lceil (y+1)/x \rceil - 1$ が成立する．

定義 1. $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$，$\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ に対して $d_H(\boldsymbol{x}, \boldsymbol{y}) \triangleq |\{i \mid a_i \neq b_i\}|$ を **Hamming 距離**と呼ぶ．さらに $d(\mathcal{C}) = \min\{d_H(\boldsymbol{a}, \boldsymbol{b}) \mid \boldsymbol{a}, \boldsymbol{b} \in \mathcal{C}, \boldsymbol{a} \neq \boldsymbol{b}\}$ を符号 $\mathcal{C}$ の**最小 Hamming 距離**と呼ぶ． □

以下これを単に最小距離と書く．有限体 $\mathbb{F}_q$ 上で定義される符号 $\mathcal{C}$ が符号長 n，次元 $k = \log_q |\mathcal{C}|$，最小距離 d であるとき $[n, k, d]_q$ 符号と呼ぶ．最小距離が d の符号は任意の $d-1$ 個のシンボルの消失，または $\lfloor (d-1)/2 \rfloor$ 個の誤りを限界距離復号により訂正可能である．したがって，与えられた他の符号パラメータに対して最小距離が大きい符号を構成することが重要である．

2.2　Locally Repairable 符号

符号語の第 $i \in [n]$ シンボルが高々 r 個の他のシンボルにアクセスすることで復元可能である時，第 i シンボルは局所性 r を持つという．本論文では全てのシンボルが局所性を持つ符号を扱う．

定義 2. $[n, k, d]_q$ 符号を $\mathcal{C}$ とする．任意の $i \in [n]$ に対して，i を要素に持つ部分集合 $S_i \subset [n]$, $|S_i| \leq r+1$ が存在して符号語の第 i シンボル c_i が $c_i = \sum_{j \in S_i \setminus \{i\}} \lambda_{ij} c_j$ $(\lambda_{ij} \neq 0)$ を満たすとき，符号 $\mathcal{C}$ は**局所性** r を持つという． □

定理 1. 局所性 r 持つ $[n, k, d]_q$ 符号の最小距離は次式を満たす．

$$d \leq n - k - \left\lceil \frac{k}{r} \right\rceil + 2. \tag{2.1}$$

この上界が Singleton 限界の一般化の形で与えられていることは $r = k$ とすれば確かめられる．局所的に復元可能なシンボル数を一般化した定義が Kamath *et al.* [4] によって与えられている．正の整数 r と δ に対して，局所性 (r, δ) を持つ符号は次のように定義される．

定義 3 (Kamath *et al.* [4])**.** $[n, k, d]_q$ 符号を $\mathcal{C}$ とする．任意の $i \in [n]$ に対して次の条件を満たす部分集合 $\mathcal{S}_i \subseteq [n]$ が存在するとき，符号 $\mathcal{C}$ は**局所性** (r, δ) を持つという．

i. $i \in \mathcal{S}_i$, $|\mathcal{S}_i| \leq r + \delta - 1$;
ii. 符号 $\mathcal{C}_{\mathcal{S}_i}$ の最小距離が δ 以上.

□

局所性 (r, δ) を持つ符号を (r, δ)-**LRC 符号**と呼ぶ．任意の $\delta - 1$ 個の消失シンボルは，その位置に対応するパンクチャ符号 $\mathcal{C}_{\mathcal{S}_i}$ の消失訂正により局所的に復元することができる．$\delta = 2$ のとき局所性 (r, δ) の定義は定義 2 と一致する．また，定義より直ちに $d \geq \delta$ であることが分かる．(r, δ)-LRC 符号の最小距離の上界が次のように与えられている．

定理 2 (Kamath *et al.* [4])**.** $[n, k, d]_q$ (r, δ)-LRC 符号の最小距離は次式を満たす．

$$d \leq n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1). \tag{2.2}$$

□

式 (2.2) が $\delta = 2$ に対して式 (1.1) と一致することが確かめられる．さらに，$r = k$ とすれば Singleton 限界と一致する．式 (2.2) は符号が定義される有限体の位数 q の関数として与えられていないが，位数 q を考慮した上界も与えられている．符号長 n 最小距離 d の

符号の達成可能な次元の最大値を与える関数を $k_{\mathrm{opt}}^{(q)}(n,d)$ とする．ただし，$n<d$ に対して $k_{\mathrm{opt}}^{(q)}(n,d)\triangleq 0$．$N\triangleq r+\delta-1$ と定義する．$(r,2)$-LRC 符号の次元の上界は次で与えられる．

定理 3 (Cadambe and Mazumdar [5])**.** $l>0$ を整数とする．$[n,k,d]_q$ $(r,2)$-LRC 符号の次元は次式を満たす．

$$k\leq \min_{l}\left\{lr+k_{\mathrm{opt}}^{(q)}(n-lN,d)\right\}. \tag{2.3}$$

□

この上界は一般に C-M 限界と呼ばれる．この式は $q<n$ のとき任意のパラメータに対して式 (1.1) より良い上界を与えることも同様に示されている．$\delta>2$ に対してパンクチャ符号の次元の上界として，既知の符号の上界を用いることで与えられた限界式が Agarwal *et al.* [9] によって与えられた．$B(q,n,\delta)$ を符号長 n，最小距離 δ の q 元符号の符号語数の上界を与え n に対して対数凸である関数とする．$B(q,0,\delta)\triangleq 1$ とする．

定理 4 (Agarwal *et al.* [9])**.** $[n,k,d]_q$ (r,δ)-LRC 符号の次元は次式を満たす．

$$k\leq \mu\log_q B(q,N,\delta). \tag{2.4}$$

ここで

$$\mu\triangleq\left\lceil\frac{n-d+1}{N}\right\rceil+1. \tag{2.5}$$

□

Singleton 限界や Plotkin 限界，Hamming 限界も対数凸関数であることが示されている (e.g. [9, Sect. 2])．したがって次の系が得られる．

系 1 (Agarawal *et al.* [9])**.** $[n,k,d]_q$ (r,δ)-LRC 符号のパラメータは次式を満たす．

$$k\leq \mu r, \tag{2.6}$$

$$k\leq \mu\log_q\frac{\delta}{\delta-\frac{q-1}{q}N}, \tag{2.7}$$

$$k\leq \mu\left\{N-\log_q\left(\sum_{i=0}^{(\delta-1)/2}\binom{N}{i}(q-1)^i\right)\right\}. \tag{2.8}$$

□

式 (2.4) は，符号のパラメータと既知の上界の選択によっては式 (2.2) よりも良い上界を与えることもあるが，一般のパラメータに対しての比較は詳細には行われていない．しかしながら，既知の上界とその凸性を利用することで計算が容易な形で上界が導出できる手法は非常に有用である．

第 3 章

Locally Repairable符号のパラメータの限界式

この章では，主定理として (r,δ)-LRC 符号の次元の上界を与える．この上界は C-M 限界に近い形で与えられ，非線形符号に対しても成立する．前節では，パンクチャ符号の次元の上界が LRC 符号の次元の上界に大きく関係することを議論した．C-M 限界はパンクチャ符号の次元の上界と局所性がともに r であるときに与えられた上界であり，$\delta > 2$ に対しては両者の値が異なるときが多く証明手法の拡張は容易でない．そこで本章では，文献 [9] の手法を参考にしてパンクチャ符号の次元の上界を与える関数を取り入れる．本章で用いる短縮符号を用いた証明手法は次元の上界と同時に最小距離の上界や符号長の下界を導出することが容易であり，LRC 符号のパラメータの限界式の導出に広く用いられている手法である．

3.1 主定理

符号長 n，次元 k の符号の達成可能な最小距離の最大値を与える関数を $d^{(q)}_{\mathrm{opt}}(n,k)$ とする．ただし $n<k$ に対して $d^{(q)}_{\mathrm{opt}}(n,d) \triangleq 0$．同様に，次元 k，最小距離 d の符号の達成可能な符号長の最小値を与える関数を $n^{(q)}_{\mathrm{opt}}(k,d)$ とする．ただし $k<d$ に対して $n^{(q)}_{\mathrm{opt}}(k,d) \triangleq d$. $k^*(a,d)$ を符号長 a，最小距離 d の符号の次元の上界を与え，符号長 a に対する凸関数とする．$k^*(a,d)$ は次の不等式を満たす．

$$k^*(a,d)+k^*(b,d) \le k^*(a-1,d)+k^*(b+1,d). \tag{3.1}$$

ただし，$0 \le a \le b$．また，$a<d$ に対して $k^*(a,d) \triangleq 0$ とする．$\log_q B(q,N,\delta)$ は k^* の条件を満たしている．主定理を以下に与える．

定理 5. $k^*(n,d)$ を上の性質を満たす関数とする．$[n,k,d]_q$ (r,δ)-LRC 符号に対して，次

の関数を定義する．

$$\lambda(x) \triangleq \left\lfloor \frac{x}{N} \right\rfloor k^*(N,\delta) + k^*(x \bmod N, \delta). \tag{3.2}$$

このとき，$[n,k,d]_q$ (r,δ)-LRC 符号のパラメータは次式を満たす．

$$n \geq \max_l \min_s \left\{ s + n^{(q)}_{\mathrm{opt}}(k - \lambda(s), d) \right\}, \tag{3.3}$$

$$k \leq \min_l \max_s \left\{ \lambda(s) + k^{(q)}_{\mathrm{opt}}(n - s, d) \right\}, \tag{3.4}$$

$$d \leq \min_l \max_s \left\{ d^{(q)}_{\mathrm{opt}}(n - s, k - \lambda(s)) \right\}. \tag{3.5}$$

ただし，l は正の整数であり，s は $(l-1)N + \delta \leq s \leq lN$ を満たす整数． □

関数 $\lambda(x)$ は $(l-1)N \leq s \leq lN$ に対して

$$\lambda(s) = (l-1)k^*(N,\delta) + k^*(s - (l-1)N, \delta) \tag{3.6}$$

とも書ける．

3.2　主定理の証明

定理 5 は以下の 2 つの補題によって証明される．

補題 1. $\mathcal{C}$ を $[n,k,d]_q$ (r,δ)-LRC 符号，l を任意の整数とする．整数 s が $(l-1)N + \delta \leq s < lN + \delta$ を満たすとき，符号長 $n-s$ の $\mathcal{C}$ の短縮符号 $\mathcal{C}^*$ が存在して $\log_q |\mathcal{C}^*| \geq k - \lambda(s)$ を満たす． □

以下の証明では $s \geq n$ のときは便宜的に符号長 0，次元 $\log_q |\mathcal{C}^*| = 0$，最小距離 $d = n$ の短縮符号を考える．

証明. 符号語 $\boldsymbol{z} \in \mathcal{C}_{\mathcal{S}}$ に対して，$\mathcal{C}(\boldsymbol{z})$ を符号語 $\boldsymbol{c} \in \mathcal{C}$ の集合 $\mathcal{S} \subseteq [n]$ 上への射影が $\boldsymbol{z}$ と一致する符号 $\mathcal{C}$ の符号語全体の集合とする．つまり，

$$\mathcal{C}(\boldsymbol{z}) \triangleq \{ \boldsymbol{c} \in \mathcal{C} \mid \boldsymbol{c}_{\mathcal{S}} = \boldsymbol{z} \}. \tag{3.7}$$

与えられた LRC 符号 $\mathcal{C}$ と整数 $l > 0$ に対して，部分符号 $\tilde{\mathcal{C}}$ を構成する Algorithm 1 を示す．

Algorithm 1 Construction of $\tilde{\mathcal{C}}$

Input: $\mathcal{C}$, l.

Output: τ, $\mathcal{I}$, $\tilde{\mathcal{C}}$.

1: $\mathcal{I}_0 \leftarrow \emptyset$, $\tilde{\mathcal{C}}^{(0)} \leftarrow \mathcal{C}$, $j \leftarrow 0$.
2: **while** $|\mathcal{I}_j| < (l-1)N + \delta$ and $|\mathcal{I}_j| < n$ **do**
3: $\quad j \leftarrow j + 1$.
4: $\quad$ choose $i \in [n] \backslash \mathcal{I}_{j-1}$ and $\mathcal{S}_i$ s.t. $|\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_i}|$ is the smallest.
5: $\quad \mathcal{I}_j \leftarrow \mathcal{I}_{j-1} \cup \mathcal{T}_i$.
6: $\quad \mathcal{T}_i \leftarrow \mathcal{S}_i \backslash \mathcal{I}_{j-1}$. ▷ $\mathcal{T}_i \neq \emptyset$.
7: $\quad$ choose $\boldsymbol{z} \in \tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_i}$ s.t. $|\tilde{\mathcal{C}}^{(j-1)}(\boldsymbol{z})|$ is the largest.
8: $\quad \tilde{\mathcal{C}}^{(j)} \leftarrow \tilde{\mathcal{C}}^{(j-1)}(\boldsymbol{z})$.
9: **end while**
10: $\tau \leftarrow j$, $\mathcal{I} \leftarrow \mathcal{I}_j$, $\tilde{\mathcal{C}} \leftarrow \tilde{\mathcal{C}}^{(j)}$.

一般性を失うことなく Algorithm 1 のステップ 4 で $i = j$ が選ばれたと仮定する．ステップ 7 において，$\tilde{\mathcal{C}}^{(j-1)}(\boldsymbol{z})$ の濃度が最大になるように $\boldsymbol{z} = \boldsymbol{z}^*$ を選んでいるから，任意の $\boldsymbol{z} \in \tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_i}$ に対して $\tilde{\mathcal{C}}^{(j-1)}(\boldsymbol{z})$ が一様に分布していると仮定したときの平均より $\tilde{\mathcal{C}}^{(j-1)}(\boldsymbol{z}^*)$ の濃度は大きい．いま，$\tilde{\mathcal{C}}^{(j-1)}$ の $\mathcal{I}_{j-1}$ に対応するシンボルは固定されているから $\boldsymbol{z} \in \tilde{\mathcal{C}}^{(j-1)}$ を $\boldsymbol{z}$ に関して和を取ると次式が成立する．

$$\sum_{\boldsymbol{z} \in \tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_i}} \left|\tilde{\mathcal{C}}^{(j-1)}(\boldsymbol{z})\right| = \left|\tilde{\mathcal{C}}^{(j-1)}\right|. \tag{3.8}$$

したがって，任意の $j \geq 1$ において次の漸化式が書ける．

$$\left|\tilde{\mathcal{C}}^{(j)}\right| = \left|\tilde{\mathcal{C}}^{(j-1)}(\boldsymbol{z}^*)\right| \geq \frac{\left|\tilde{\mathcal{C}}^{(j-1)}\right|}{\left|\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j}\right|}. \tag{3.9}$$

$\tilde{\mathcal{C}}^{(0)} = \mathcal{C}$ より漸化式を用いれば $\tilde{\mathcal{C}}$ について次式が成立する．

$$\left|\tilde{\mathcal{C}}\right| = \left|\tilde{\mathcal{C}}^{(\tau)}\right| \geq \frac{|\mathcal{C}|}{\prod_{j=1}^{\tau} \left|\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j}\right|}. \tag{3.10}$$

式 (3.10) の両辺の対数を取ると

$$\log_q \left|\tilde{\mathcal{C}}\right| \geq k - \sum_{j=1}^{\tau} \log_q \left|\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j}\right|. \tag{3.11}$$

次に $\log_q |\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{T}_j}|$ の上界を考える．$|\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j}| > 1$ のとき．$\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j}$ の任意の異なる 2 つの符号語を選ぶ．いま，$\tilde{\mathcal{C}}^{(j-1)}$ の $\mathcal{I}_{j-1}$ に対応するシンボルは固定されているから，$\mathcal{S}_j$ の $\mathcal{I}_{j-1}$ との共通部分のシンボルも固定されている．つまり，$\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j}$ の任意の異なる 2 つの符号語は，インデックス集合 $\mathcal{T}_i$ に対応するシンボルが異なり，$\mathcal{S}_j \backslash \mathcal{T}_j$ に対応するシンボルは固定されている．したがって，$|\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j}| = |\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{T}_j}|$，$d(\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{T}_j}) = d(\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j})$ である．$1 \leq j \leq \tau$ に対して $\tilde{\mathcal{C}}^{(j-1)} \subseteq \mathcal{C}$ だから $\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j} \subseteq \mathcal{C}_{\mathcal{S}_j}$ である．部分符号の最小距離は元の符号より小さくなることはないので，

$$d(\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{T}_j}) = d(\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j}) \geq d(\mathcal{C}_{\mathcal{S}_j}) \geq \delta. \tag{3.12}$$

これと $|\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j}| = |\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{T}_j}|$ より

$$\log_q \left|\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j}\right| = \log_q \left|\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{T}_j}\right| \leq k^{(q)}_{\mathrm{opt}}(|\mathcal{T}_j|, \delta). \tag{3.13}$$

$|\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j}| = 1$ のとき．$\log_q |\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j}| = \log_q |\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{T}_j}| = 0$ である．$k^{(q)}_{\mathrm{opt}}(|\mathcal{T}_j|, \delta)$ は定義より非負であるためこれを上から抑える．したがって $|\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{S}_j}|$ の大きさによらず任意の $j < \tau$ に対して次式が成立する．

$$\log_q \left|\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{T}_j}\right| \leq k^{(q)}_{\mathrm{opt}}(|\mathcal{T}_j|, \delta) \leq k^*(|\mathcal{T}_j|, \delta). \tag{3.14}$$

したがって

$$\sum_{j=1}^{\tau} \log_q \left|\tilde{\mathcal{C}}^{(j-1)}_{\mathcal{T}_j}\right| \leq \sum_{j=1}^{\tau} k^*(|\mathcal{T}_j|, \delta). \tag{3.15}$$

ここで，式 (3.15) の右辺の各 $\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_\tau$ から最も濃度が小さい集合とその次に濃度が小さい集合を選び，不等式 (3.1) を適用する．この不等式を式 (3.1) の右辺第一項の引数が 0 になるか，第二項の引数が N になるまで繰り返し用いる．不等式を適用しても第一引数の和が変わらないことに注意したい．この不等式を式 (3.15) に対して繰り返し用いれば，第一引数の総和 $\sum_{i=1}^{\tau} |\mathcal{T}_i| = |\mathcal{I}|$ は変化しないから次の不等式が得られる．

$$\sum_{j=1}^{\tau} k^*(|\mathcal{T}_j|, \delta) \leq \left\lfloor \frac{|\mathcal{I}|}{N} \right\rfloor k^*(N, \delta) + k^*(|\mathcal{I}| \bmod N, \delta). \tag{3.16}$$

式 (3.11), (3.15), (3.16) より

$$\begin{aligned} \log_q \left|\tilde{\mathcal{C}}\right| &\geq k - \sum_{j=1}^{\tau} k^*(|\mathcal{T}_j|, \delta) \\ &\geq k - \left\lfloor \frac{|\mathcal{I}|}{N} \right\rfloor k^*(N, \delta) - k^*(|\mathcal{I}| \bmod N, \delta). \end{aligned} \tag{3.17}$$

部分符号 $\tilde{\mathcal{C}}$ の $\mathcal{I}$ に対応するシンボルを取り除いて得られる符号長 $n-|\mathcal{I}|$ のパンクチャ符号 $\mathcal{C}^*$ を考える．$|\mathcal{C}^*|>1$ のとき，部分符号 $\tilde{\mathcal{C}}$ の $\mathcal{I}$ に対応するシンボルは固定されているから $\mathcal{C}^*$ は $\mathcal{C}$ の短縮符号であり，$|\tilde{\mathcal{C}}|=|\mathcal{C}^*|$ である．

Algorithm 1 から出力される部分集合 $\mathcal{I}$ の濃度は，$(l-1)N+\delta\le|\mathcal{I}|<lN+\delta$ または $|\mathcal{I}|=n$ である．(i) $(l-1)N+\delta\le|\mathcal{I}|<lN+\delta$ のとき．$|\mathcal{I}|=s$ と置き換えれば $\log_q|\mathcal{C}^*|\ge k-\lambda(s)$ が成立．(ii) $|\mathcal{I}|=n$ のとき．$|\mathcal{I}|=n<(l-1)N+\delta\le s$ となるが，式 (3.17) の右辺は，$k-\lambda(|\mathcal{I}|)$ と一致しており，$\lambda(x)$ は x に関して単調増加だから $x=s>n$ を代入すれば

$$\begin{aligned}\log_q|\mathcal{C}^*|&\ge k-\lambda(n)\\&\ge k-\lambda(s).\end{aligned}\tag{3.18}$$

□

補題 2. $\mathcal{C}$ を $[n,k,d]_q$ 符号，$\mathcal{C}^*$ を $\mathcal{C}$ の $[n',k',d']_q$ 短縮符号とする．$n'=n-s$，$k'\ge k-t$ のとき

$$n\ge s+n^{(q)}_{\mathrm{opt}}(k-t,d),\tag{3.19}$$

$$k\le t+k^{(q)}_{\mathrm{opt}}(n-s,d),\tag{3.20}$$

$$d\le d^{(q)}_{\mathrm{opt}}(n-s,k-t).\tag{3.21}$$

□

証明. 短縮符号の最小距離について，$d\le d'$ であるから

$$d\le d'\le d^{(q)}_{\mathrm{opt}}(n-s,k')\le d^{(q)}_{\mathrm{opt}}(n-s,k-t).\tag{3.22}$$

同様に，

$$n=n'+s\ge n^{(q)}_{\mathrm{opt}}(k',d')\ge n^{(q)}_{\mathrm{opt}}(k-t,d).\tag{3.23}$$

$k\le t+k'$ より，

$$k\le t+k'\le t+k^{(q)}_{\mathrm{opt}}(n-s,d')\le t+k^{(q)}_{\mathrm{opt}}(n-s,d).\tag{3.24}$$

□

補題 1 より任意の $l>0$ について短縮符号が存在し，各 l に対して補題 2 も成立する．ただし s は l によって定まる範囲内を動くことに注意する．$\lambda(s)$ は s に関して単調増加であり，その第二項は s に関して周期関数でありその周期内で $s\ge(l-1)N+\delta$ ならば単調増加，$lN<s<lN+\delta$ ならば 0 である．したがって，$k-\lambda(s)$ の最小値を与える s は $(l-1)N+\delta\le s\le lN$ の範囲にある．各 l に対して n の下界と k,d の上界を計算して最も良いものを選ぶことで定理 5 が言える．

第 4 章

定理から得られる結果

第 3 章の定理 5 で得られた上界は，一般のパラメータに対して厳密には計算可能でない関数 $n_{\mathrm{opt}}^{(q)}, k_{\mathrm{opt}}^{(q)}, d_{\mathrm{opt}}^{(q)}$ を用いて与えられているため，そのままの形では既存の上界と比較を行うのは容易でない．しかし，これらの関数を既知の上界で抑えることで計算可能な形の上界を得ることができる．本章では定理 5 の評価を緩めることで計算可能な形へ変形し，既存の上界と比較やその考察を行う．

4.1 既存の限界式との比較

与えた上界 (3.4) は任意の $\delta \geq 2$ に対して成立する．$\delta = 2$ $(N = r + 1)$ として C-M 限界と比較する．k^* に Singleton 限界を選べば次の系が得られる．

系 2. $[n, k, d]_q$ $(r, 2)$-LRC 符号の次元は次式を満たす．

$$k \leq \min_{l} \max_{s} \left\{ s - l + k_{\mathrm{opt}}^{(q)}(n - s, d) \right\}. \tag{4.1}$$

ただし，l は正の整数であり s は $(l-1)N + 2 \leq s \leq lN$ を満たす整数． □

$r = 1$ のとき $s = lN$ となり右辺は C-M 限界 (2.3) と一致する．$r > 1$ のとき右辺は C-M 限界と同じかそれより大きくなる．次に示すのは本論文で与えた上界が C-M 限界と一致するパラメータの例である．

例 1. 符号長 $n = 8$，最小距離 $d = 3$ の局所性 $(r, \delta) = (2, 2)$ を持つ 2 元 LRC 符号を考える．達成可能な次元の最大値の表 [15] によると，C-M 限界 $k \leq \min_l\{2l + k_{\mathrm{opt}}^{(q)}(8 - 3l, 3)\}$ は次のように計算できる．

$$\begin{aligned} 2 + k_{\mathrm{opt}}^{(q)}(5, 3) &= 4 \qquad (\text{for } l = 1), \\ 4 + k_{\mathrm{opt}}^{(q)}(2, 3) &= 4 \qquad (\text{for } l = 2), \\ 2l + k_{\mathrm{opt}}^{(q)}(8 - 3l, 3) &\geq 6 \qquad (\text{for all } l \geq 3). \end{aligned}$$

したがって，$k \leq 4$．一方で式 (4.1) は，$l = 1$ のとき $2 \leq s \leq 3$ より

$$1 + k_{\mathrm{opt}}^{(q)}(6, 3) = 4 \qquad (\text{for } s = 2),$$
$$2 + k_{\mathrm{opt}}^{(q)}(5, 3) = 4 \qquad (\text{for } s = 3).$$

式 (4.1) は，C-M 限界より小さい値を与えることはないので，$l \geq 2$ については計算する必要はなく $k \leq 4$ であり C-M 限界と一致する．□

一般のパラメータに対しては例 1 のような真の値を用いた比較は困難であるため，計算可能な形に置き換えることで比較を行う．$r > 1$ のとき式 (4.1) の $k_{\mathrm{opt}}^{(q)}$ に対して凸関数を適用すると，s に関する最大化項は単調増加だから $s = lN$ で最大値を取る．したがって

$$k \leq \min_{l} \left\{ lr + k^*(n - lN, d) \right\}. \tag{4.2}$$

つまり，凸関数を用いて計算可能な形に書き直した場合式 (4.1) と C-M 限界は同じ上界を与える．次に式 (3.5) が局所性 (r, δ) を持つ符号に対する Singleton-like 限界より良い上界を与えることを示す．

定理 6. 式 (3.5) は Kamath *et al.* [4] 上界式 (2.2) と同じかそれより良い上界を与える．□

証明. $l = \lceil k/r \rceil + 1$ に固定して $\lambda(s)$ 中の k^* に Singleton 限界を選ぶことで $k - \lambda(s)$ を具体的に計算する．$s \geq (l-1)N + \delta$ だから $k^*(s - (l-1)N, \delta) \neq 0$ である．$l = \lceil k/r \rceil + 1$，$k^*(N, \delta) = r$ に注意して

$$\begin{aligned} \lambda(s) &= \left\lceil \frac{k}{r} \right\rceil r + s - \left\lceil \frac{k}{r} \right\rceil N + \delta - 1 \\ &= s - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1). \end{aligned} \tag{4.3}$$

したがって

$$k - \lambda(s) \geq k + \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1) - s. \tag{4.4}$$

以上より

$$\begin{aligned} &\min_{l} \max_{s} \left\{ d_{\mathrm{opt}}^{(q)}(n - s, k - \lambda(s)) \right\} \\ &\quad \leq \max_{s} \left\{ d_{\mathrm{opt}}^{(q)}(n - s, k + \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1) - s) \right\}. \end{aligned} \tag{4.5}$$

$d_{\mathrm{opt}}^{(q)}$ の第一引数が第二引数以上であることを示す．式 (2) より

$$0 \leq d - 1$$

$$\leq n-k-\left(\left\lceil\frac{k}{r}\right\rceil-1\right)(\delta-1)$$
$$=(n-s)-\left\{k+\left(\left\lceil\frac{k}{r}\right\rceil-1\right)(\delta-1)-s\right\}. \tag{4.6}$$

式 (4.5) の右辺の $d_{\mathrm{opt}}^{(q)}$ を Singleton 限界を使って上から抑えると

$$\min_{l}\max_{s}\left\{d_{\mathrm{opt}}^{(q)}(n-s,k-\lambda(s))\right\}\leq n-k+1-\left(\left\lceil\frac{k}{r}\right\rceil-1\right)(\delta-1). \tag{4.7}$$

□

$k_{\mathrm{opt}}^{(q)}$ に対して適用する次元の上界に対する制限は無いが，適用する上界が凸関数のとき，右辺の s に関する最大化項もまた s ($(l-1)N+\delta\leq s\leq lN$) に対して凸であるから，最大値は端点 $(l-1)N+\delta$ または lN で取る．

定理 7. 式 (3.5) は Agarawal *et al.* [9] 上界式 (2.4) と同じかそれより良い上界を与える．□

証明. $k_{\mathrm{opt}}^{(q)}$ を k^* で置き換える．

$$\mu\triangleq\left\lceil\frac{n-d+1}{N}\right\rceil+1 \tag{4.8}$$

として $l=\mu$ に固定すると

$$k\leq\max_{s}\{(\mu-1)k^*(N,\delta)+k^*(s-(\mu-1)N,\delta)+k^*(n-s,d)\}. \tag{4.9}$$

$s=\mu N$ のとき右辺は $\mu k^*(N,\delta)$ で上から抑えられる．$s=(\mu-1)N+\delta$ のとき $n-s<d$ より右辺は

$$(\mu-1)k^*(N,\delta)+k^*(\delta,\delta)\leq\mu k^*(N,\delta). \tag{4.10}$$

したがって，$k\leq\mu k^*(N,\delta)$. □

4.2　数値計算による比較

本節では，与えた上界を具体的なパラメータに対して計算する．ただし本節で計算する上界は全て線形符号に対する上界である．

式 (3.4) の k^*，$k_{\mathrm{opt}}^{(q)}$ に対して同じ上界を適用して計算したものと，k^* に適用する上界を固定して $k_{\mathrm{opt}}^{(q)}$ として最も良い上界を与えるものを選んで計算したものをプロットし比較する．ただし，両者に同じ上界を適用した場合，例えば Singleton 限界を適用した場合には Singleton-like 限界のように表記し，$k_{\mathrm{opt}}^{(q)}$ として最も良い上界を与えるものを選んだ場合

には Optimum として表記する．$k_{\mathrm{opt}}^{(q)}$ に適用する限界は，Singleton 限界，Hamming 限界，Plotkin 限界，Elias 限界，Johnson 限界，Griesmer 限界の中から選んだ．$q = 2$ のとき，最小距離が奇数の $[n, k, 2d-1]_2$ 符号に対して，情報記号数が等しい拡大 $[n+1, k, 2d]_2$ 符号が存在するから両者のパラメータに対して値を計算をして良い方を上界として用いる．また，Plotkin 限界は通常 $d \leq n(1-1/q)$ に対して定義されないが，補題 2 の式 (3.20) に倣って短縮符号を考えることで上界を計算した．既存の上界は全て破線でプロットした．

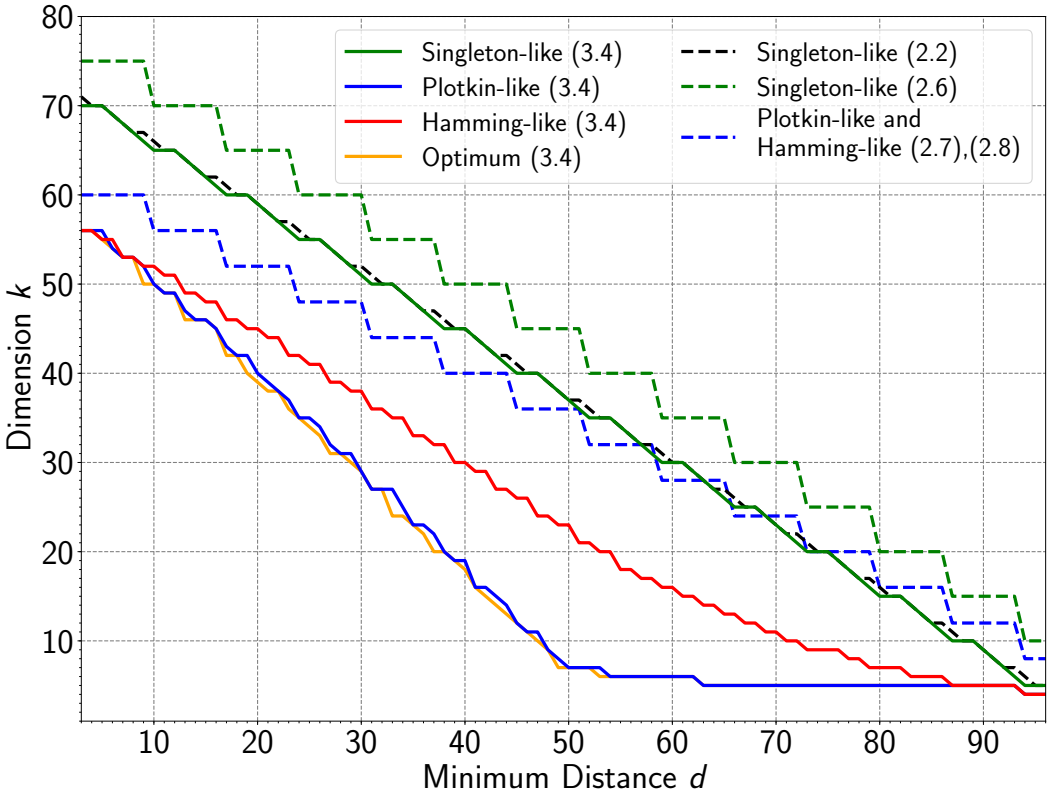

図 4.1: 上界の比較 $(n = 100,\ (r, \delta) = (5, 3),\ q = 2)$

図 4.1 は符号長を $n = 100$，局所性を $(r, \delta) = (5, 3)$ として $q = 2$ の場合に，横軸に最小距離 d，縦軸に情報記号数 k を取りプロットしたものである．この場合，式 (2.4) において $\log_q B(q, N, \delta)$ は，Plotkin 限界と Hamming 限界のどちらを適用しても同じ値を与えるために両者をまとめて青の破線でプロットした．定理 6,7 で示したように既存の上界より良い上界となっていることが確かめられる．また Optimum の k^* には Plotkin 限界を選んでいるが，$k_{\mathrm{opt}}^{(q)}$ として適切な上界を選ぶことで Plotkin-like 限界よりも良い上界が得られることも確かめられる．

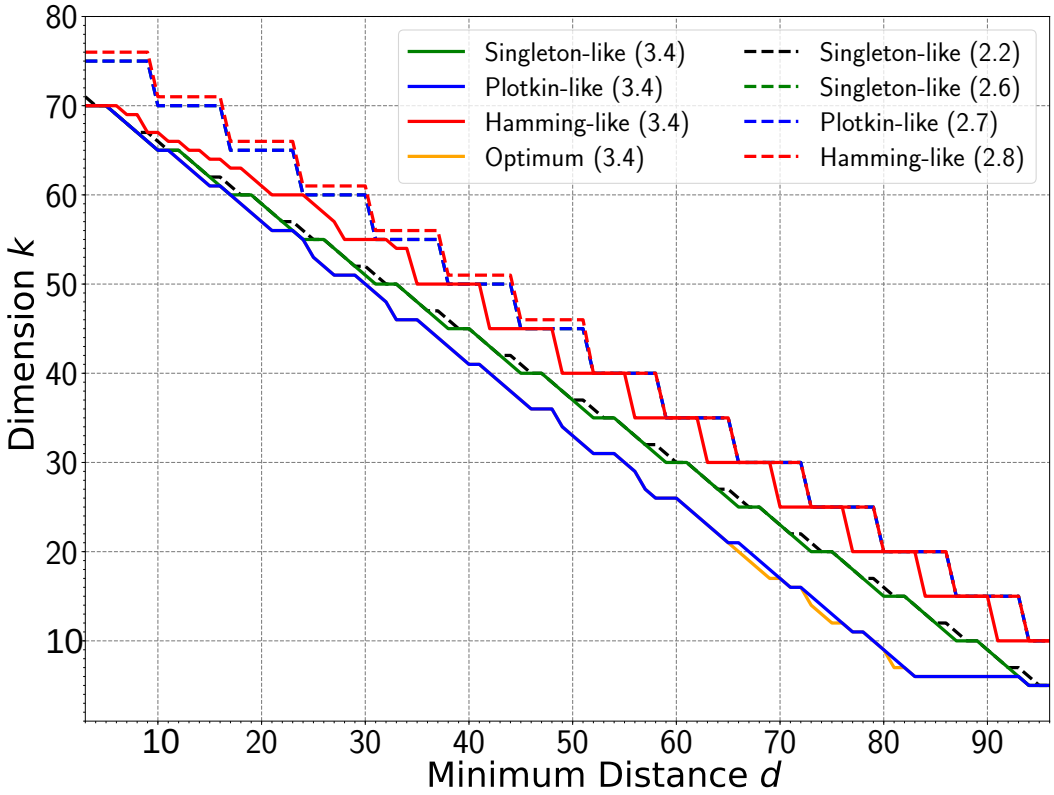

図 4.2: 上界の比較 ($n = 100$, $(r, \delta) = (5, 3)$, $q = 8$)

図 4.2 では符号長と局所性を変えずに $q = 8 > N$ とした．Singleton 限界，Singleton-like 限界は q の関数でないために図 4.1 と同じ値である．この場合も既存の上界より良い上界を与えていることが確かめられるが，最小距離が小さいときに (特に $d \leq N$ のとき) 既存の上界と一致している点がある．これは，$q > N$ のとき最小距離 $d \leq N$ を持つような Singleton-like 限界 (2.2) を達成する最適な LRC 符号の構成が可能であることを示唆していると思われる．図 4.3 は局所性を $(r, \delta) = (10, 5)$ として $q = 2$ に対してプロットしたものである．Hamming-like 限界が高レート帯，Plotkin-like 限界が低レート帯で良い上界を与えている．これは一般の符号に対する Hamming 限界と Plotkin 限界の特性と一致している．Optimum の k^* は Hamming 限界に固定して計算しているが，k^* に Plotkin 限界を選んでもこれより良い上界を得ることは無かった．図 4.1, 4.2 と異なり，$k_{\mathrm{opt}}^{(q)}$ 選ぶ限界式により上界の値が大きく変化する場合があることが分かる．

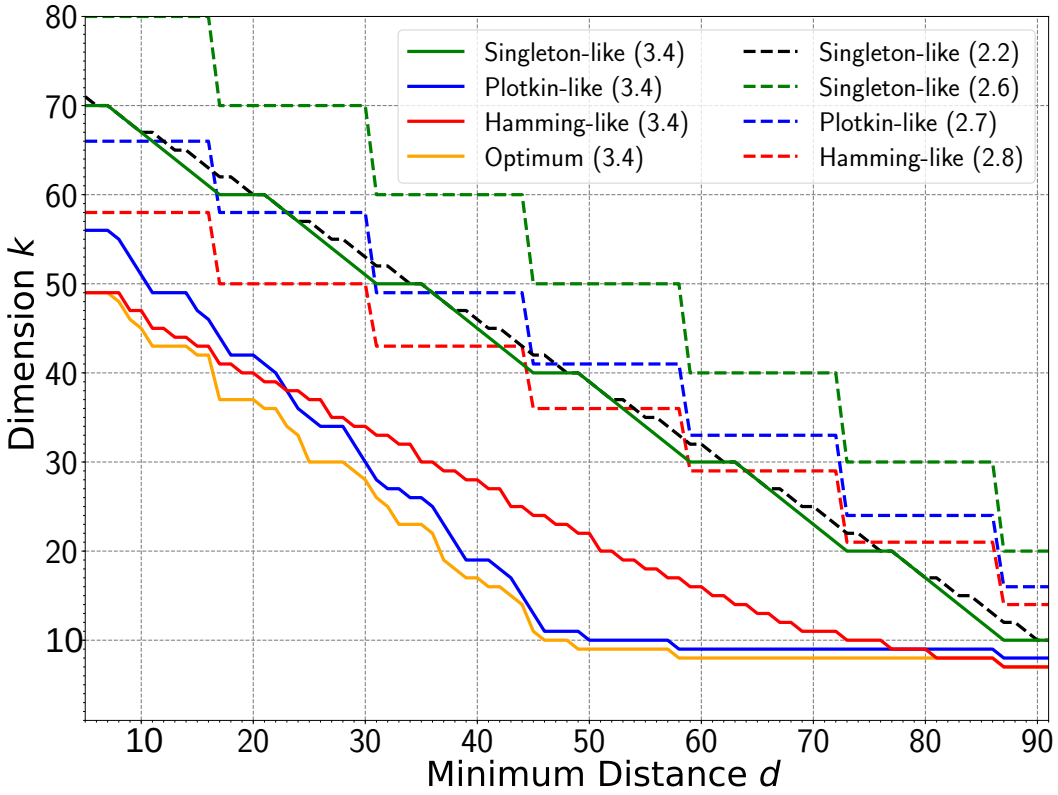

図 4.3: 上界の比較 $(n = 100,\ (r, \delta) = (10, 5),\ q = 2)$

4.3　上界を達成する符号の例

文献 [6] では，$[2^u - 1, 2^u - u - 1, 3]_2$ Hamming 符号 $(u \geq 3)$ が局所性 $r = 2^{m-1} - 1$ を持つ最適な LRC 符号であることが示されているが，定義 3 から局所性 $(r, \delta) = (2^u - 3, 3)$ 持つ LRC 符号でもあり，Hamming 限界を達成する最適な情報記号数を持つ符号である．したがって，Hamming 符号の直積による符号も局所性 $(r, \delta) = (2^u - 3, 3)$ を持ち，その情報記号数も最適となるが実際に上界と一致することを確かめる．

式 (3.4) の右辺を等号で達成する最適な 2 元符号を構成する．$[2^u - 1, 2^u - u - 1, 3]_2$ Hamming 符号の検査行列を $\tilde{H}$ とする．I を $t \times t$ 単位行列とする．次のパリティ検査行列を考える．

$$H_1 = \left(I \otimes \tilde{H} \right). \tag{4.11}$$

ここで $\otimes$ は直積を表す．

定理 8. パリティ検査行列 H_1 で与えられる符号は，符号長 $n = t(2^u - 1)$，情報記号数 $k = t(2^u - u - 1)$，最小距離 $d = 3$ の局所性 $(r, \delta) = (2^u - 3, 3)$ 持つ符号であり，式 (3.4) を等号で達成する．□

証明. Hamming 符号は，局所性 $(r,\delta)=(2^u-3,3)$ を持つから，H で与えられる符号も局所性 $(r,\delta)=(2^u-3,3)$ を持つことは自明である．このとき符号長は $n=t(2^u-1)$，情報記号数は $k=t(2^u-u-1)$ である．符号の最適性を示す．$d=3$ のとき $2<N$ に対して

$$\mu=\left\lceil\frac{n-d+1}{N}\right\rceil+1=\frac{n}{N} \tag{4.12}$$

である．定理 7 より式 (3.4) が $k\leq\mu k^*(N,\delta)$ を与えることが分かるが，実際に $k=\mu k^*(N,\delta)$ である． □

同様にして拡大 Hamming 符号や短縮 Hamming 符号の直積による符号が最適な LRC 符号であることも同様にして容易に確かめられる．最適な構造を持つ LRC 符号を修正することで，異なるパラメータを持つ LRC 符号を構成する方法が多く用いられている．ここでは，局所性を変えずに符号全体の最小距離を大きくする方法を与える．次のパリティ検査行列を考える．

$$H_2=\begin{pmatrix}I\otimes\tilde{H}\\ 1\ \cdots\ 1\end{pmatrix}. \tag{4.13}$$

このパリティ検査行列で与えられる符号は，構成法 1 で定義される符号から奇数重みの符号語をすべて取り除いた部分符号である．

定理 9. $u=3$ とする．パリティ検査行列 H_2 で与えられる符号は，符号長 $n=t(2^u-1)$，情報記号数 $k=t(2^u-u-1)-1$，最小距離 $d=4$ の局所性 $(r,\delta)=(2^u-3,3)$ 持つ符号であり，式 (3.3) を等号で達成する． □

証明. 元の符号で局所訂正できる消失パターンは全て部分符号でも訂正可能であるから，符号の局所性は少なくとも元の符号と変わらない．最小重みが奇数の 2 元線形符号は偶数重みと奇数重みの符号語を半数ずつ持つため情報記号数は $t(2^u-u-1)-1$ である．また，$d\geq4$ であるが一次従属な 4 列の組が存在するため $d=4$．式 (3.3) において $l=t-1$ と固定すれば $(t-2)N+\delta\leq s\leq(t-1)N$ である．$k'=k-\lambda(s)$ とおいて $n_{\mathrm{opt}}^{(q)}$ に対して Griesmer 限界

$$n'(k',d)\geq\sum_{i=0}^{k'-1}\left\lceil\frac{d}{2^i}\right\rceil \tag{4.14}$$

を適用したときに，式 (3.3) の右辺 $s+n'(k',d)$ を最小化する s を考える．以下 k^* には Hamming 限界を用いる．$\lambda(s)$ は s に関して単調増加だから $s=lN$ を代入すれば

$$\begin{aligned}k'&\geq k-\lambda((t-1)N)\\&=t(2^u-u-1)-1-lk^*(N,\delta)\end{aligned}$$

$$
\begin{aligned}
&= t(2^u - u - 1) - 1 - (t-1)(2^u - u - 1) \\
&= (2^u - u - 1) - 1.
\end{aligned} \tag{4.15}
$$

$u > 2$ に対して $k' > 2$．一方 $k' > 2$ のとき，Griesmer 限界に $d = 4$ を代入すれば

$$
n'(k', 4) \geq 4 + 2 + (k' - 2) \tag{4.16}
$$

と書くことができるため，$k' > 2$ に対して Griesmer 限界が与える値は k' に関して線形的に増加することが分かる．ところで k' は $\lambda(s)$ が凸関数であるから，s に対して線形的またはそれ以上の速度で減少する．したがって，Griesmer 限界が与える値も s に対して線形的またはそれ以上の速度で減少する．以上より，$s + n'(k', d)$ は $s = lN$ のときに最小値を取る．$l = t - 1$ より右辺を計算すれば

$$
\begin{aligned}
&(t-1)N + n'((2^u - u - 1) - 1, 4) \\
&\quad = (t-1)(2^u - 1) + 4 + 2 + 2^u - u - 1 - 3 \\
&\quad = t(2^u - 1) - u + 3.
\end{aligned} \tag{4.17}
$$

$u = 3$ を代入すれば最適性が示される．　□

例 2. 符号長 $n = 14$，情報記号数 $k = 7$，最小距離 $d = 4$ の局所性 $(r, \delta) = (5, 3)$ 持つ最適な符号は次の検査行列によって与えられる．

$$
H = \begin{pmatrix}
1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}. \tag{4.18}
$$

H の部分行列は Griesmer 限界を等号で達成する $[7, 3, 4]_2$ expurgated Hamming 符号の検査行列になっている．　□

第 5 章

まとめ

本論文では，(r, δ)-LRC 符号のパラメータの上界や下界を，既知の符号の上界を用いることにより与えた．得られた上界は，線形，非線形符号の両者に対して成立し任意のパラメータに対して Kamath *et al.* [4] 上界式 (2.2), Agarawal *et al.* [9] 上界式 (2.4) より良い上界を与えることを示した．本論文で与えた上界は既知の符号の上界を利用することで計算可能であり，数値計算により既知の上界の選択によって上界がどのように変化するのかを数値計算により確かめた．これらの数値計算により，特に q が小さいときに大きく改善された上界を得られることも確認できた．また，得られた上界は Hamming 符号の直積により達成可能である．

今後の課題として，本論文で得られた上界は，C-M 限界と同じように短縮符号のパラメータの関係を用いて任意の位数に対して得られる上界であるが，文献 [8] の手法のように体の位数を固定することによる厳密な評価が必要である．その際，文献 [9] で提案された既知の符号の上界を与える凸関数を用いる手法を用いることで，簡単な形で上界が与えられることが期待できる．

謝辞

本研究を進めるにあたって丁寧なご指導を頂きました八木秀樹准教授に深く感謝致します．ゼミ等でお世話になった川端勉教授，大濱靖匡教授，Santoso Bagus 助教，八木研，川端研の学生および共に研究してきた同期の皆様に感謝致します．

参考文献

[1] C. Huang and M. Chen and J. Li, "Pyramid codes: Flexible schemes to trade space for access efficiency in reliable data storage systems," in Proc. *Sixth IEEE International Symposium on Network Computing and Applications (NCA 2007)*, pp. 79–86, Cambridge, MA, USA, Jul. 2007.

[2] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, "On the locality of codeword symbols," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6925–6934, Nov. 2012.

[3] D. S. Papailiopoulos and A. G. Dimakis, "Locally repairable codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 5843–5855, Oct. 2014.

[4] G. M. Kamath, N. Prakash, V. Lalitha, and P. V. Kumar, "Codes with local regeneration and erasure correction," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4637–4660, Aug. 2014.

[5] V. R. Cadambe and A. Mazumdar, "Bounds on the size of locally recoverable codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 5787–5794, Feb. 2015.

[6] P. Huang and E. Yaakobi and H. Uchikawa and P. H. Siegel, "Binary linear locally repairable codes," *IEEE Trans. Inf. Theory*, vol. 62, no. 11, pp. 6268–6283, Nov. 2016.

[7] J. Hao and S.-T. Xia and B. Chen, "Some results on optimal locally repairable codes," in Proc. *2016 IEEE Int. Symp. Inf. Theory (ISIT)*, Barcelona, Spain, pp. 440–444, Jul. 2016.

[8] A. Wang and Z. Zhang and D. Lin, "Bounds and constructions for linear locally repairable codes over binary fields," in Proc. *2017 IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, pp. 2033–2037, Jun. 2017.

[9] A. Agarwal, A. Barg, S. Hu, A. Mazumdar, and I. Tamo, "Combinatorial alphabet-dependent bounds for locally recoverable codes," *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3481–3492, May 2018.

[10] I. Tamo and A. Barg, "A family of optimal locally recoverable codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4661–4676, Aug. 2014.

[11] N. Silberstein and A. S. Rawat and O. O. Koyluoglu and S. Vishwanath, "Optimal locally repairable codes via rank-metric codes," in Proc. *2015 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1819–1823, Hong Kong, China, Jul. 2013.

[12] A. Wang and Z. Zhang, "Repair locality with multiple erasure tolerance" *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6979–6987, Nov. 2014.

[13] A. S. Rawat, A. Mazumdar, and S. Vishwanath, "On cooperative local Repair in Distributed Storage," *EURASIP Journal on Advances in Signal Processing*, pp. 1–17, 2015.

[14] S. Kruglik and A. Frolov, "Bounds and constructions of codes with all-symbol locality and availability," in Proc. *2017 IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, pp. 1023–1027, Jun. 2017.

[15] M. Grassl, "Code Tables: Bounds on the parameters of various types of codes." [Online]. Available at http://www.codetables.de/

発表実績

i. 濱田 寛也，八木 秀樹 “符号化多項式を用いた多重局所性を持つ Locally Repairable 符号の構成法,” **信学技報**, vol. 117, no. 120, pp. 27–32, Jul. 2017. (**優秀発表賞受賞**)

ii. T. Hamada and H. Yagi, “Construction of locally repairable codes with multiple localities based on encoding polynomial,” in *Proc. 2018 RISP Int. Workshop on Nonlinear Circuits, Communication and Signal Processing (NCSP2018)*, pp. 627–630, Honolulu, USA, Mar. 2018. (**Best Student Paper Award 受賞**)

iii. T. Hamada and H. Yagi, “Construction of locally repairable codes with multiple localities based on encoding polynomial,” *IEICE Trans. Fundamentals*, vol. E101-A, no. 12, pp. 2047–2054, Dec. 2018.

iv. 濱田 寛也, 八木秀樹, “Locally Repairable 符号の次元に関する上界式の改善,” **信学技報**, vol. 118, no. 205, pp. 1–6, Aug. 2018.

v. 濱田 寛也, 八木秀樹, “(r, δ)-Locally Repairable 符号の次元の上界式の改善,” **第 41 回情報理論とその応用シンポジウム予稿集**, pp. 76–81, いわき, 福島, Dec. 2018.