平成２４年度（２０１２年度）　修士論文

逆シャノン定理による
通信路の再現に関する研究

電気通信大学大学院 情報システム学研究科

情報ネットワーク学専攻

1152026　長井　大地

指導教員

小川　朋宏　准教授

長岡　浩司　教 授

山口　和彦　准教授

提出日　平成２５年（２０１３年）２月２２日

目 次

第1章　序論

1.1　研究の背景

2002 年に Bennett ら [1] によって逆シャノン定理が示された．これは，任意に与えられた通信路の通信路容量を越える符号化レートを持つ適切な符号と恒等通信路を用いる事により，その通信路を再現出来ることを証明したものである．しかし，逆シャノン定理について実用的な符号を用いたときの理論的な考察はなされていない．また，実用的な符号を用いた例も，これまでのところ報告されていなかった．

実用的な符号の代表例として BCH 符号がある．BCH 符号は，Bose と Ray-Chaudhurui (1960)[2] および独立に Hocquenghem (1959)[4] により発見され，彼らにちなんで命名されている．BCH 符号は代数的符号の一種であり，誤り訂正が効率的に行なえ，復号が容易であるという利点を持っている．さらに冗長度をある程度自由に設定できるのも利点である．

上記の 2 点から BCH 符号を用いた通信路の再現の例は，実用的な符号を用いた逆シャノン定理の理論的背景を理解することに大きく貢献するものである．

1.2　本研究の概要と結果

本研究では，実用的な符号である BCH 符号を用いて逆シャノン定理の検証を行った．逆シャノン定理における通信路の再現とは，任意に与えられた n ビットの通信路を k ビットの恒等通信路を用いて再現することである．ここで，k/n は符号化レートと呼ばれ，小さい方が望ましい．逆シャノン定理は，最適な符号を使用することで，通信路の再現に必要な符号化レートの下限がほぼ通信路容量であることを保証している．そこで，本研究

ではBCH符号の符号化レートと通信路の再現成功確率の関係を調べた．その結果，BCH符号のクラスでも，逆シャノン定理の予想に基づいた通信路の再現が可能であることがわかった．しかも，通信路容量よりも大きい符号化レートで，通信路の再現成功確率が0から1に変わるしきい値的な挙動が確認できた．

1.3　本論文の構成

第2章では，本論文の基礎概念となる通信路符号化定理についてその概略を説明する．第3章では，次章の準備としてレート歪み理論の詳細を解説し，これを用いて，第4章では，逆シャノン定理の目的を述べ，証明を行う．第5章では，シミュレーションにおいて使用するBCH符号について．第6章では，行ったシミュレーションの手順とアルゴリズムについて説明する．さらに，第7章では前章で説明した手順に基づいた数値実験の結果を述べ，それに基づき考察を行う．最後に，最終章では，本論文を総括し，本研究で得られた内容と今後の課題を提示する．

第2章　通信路符号化定理

通信路符号化定理とは，シャノンが 1948 年の論文において示したもので，シャノンの第二定理とも呼ばれる．この定理が示したことは，雑音のある通信路に対して，通信路容量を定義し，これより遅い符号化レートで情報を送信する限り，符号器・復号器を効率的に設計した場合，漸近的に誤り確率を 0 にできることである．通信路符号化定理は情報理論において重要な定理である．本章では，情報理論の基礎的な事項と通信路符号化定理について述べる．本章での説明は Cover[3] を参考にした．

2.1　情報量について

まず，これからの議論に必要となる様々な情報量の定義を行う．

定義 1 (エントロピー). アルファベット $\mathcal{X}$ 上に値を取る離散確率変数 X について，X の確率分布を $p_X(\cdot)$ とする．この時，X のエントロピー $H(X)$ を以下で定義する．

$$\begin{aligned} H(X) &:= -\sum_{x\in\mathcal{X}} p_X(x)\log p_X(x) \\ &= -\mathbb{E}[\log p(X)] \end{aligned}$$

ここで，$\mathbb{E}[\cdot]$ は期待値を表している．

この時，エントロピーについて以下の定理が成り立つ．

定理 1.

$$H(X) \geq 0$$

定理 2. $H(X)$ は P_X についての上に凸な関数である．

次に複数の確率変数に対するエントロピーを定義する．

定義 2 (同時エントロピー)**.** アルファベット $\mathcal{X}$ 上の確率変数 X と $\mathcal{Y}$ 上の確率変数 Y について，同時確率分布を $p_{XY}(x,y)$ とする．この時，X と Y の同時エントロピー $H(X,Y)$ を以下で定義する．

$$\begin{aligned} H(X,Y) &:= -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p_{XY}(x,y)\log p_{XY}(x,y) \\ &= -\mathbb{E}\log p(X,Y) \end{aligned}$$

定義 3 (条件付きエントロピー)**.** 確率変数 (X,Y) が $p(x,y)$ に従って発生しているとき，条件付きエントロピー $H(Y|X)$ を以下で定義する．

$$\begin{aligned} H(Y|X) &:= \sum_{x\in\mathcal{X}} p(x)H(Y|X=x) \\ &= -\mathbb{E}\log p(Y|X) \end{aligned}$$

この時，同時エントロピーと条件付きエントロピーの間には次の定理が成り立つ．

定理 3 (チェイン則)**.**

$$\begin{aligned} H(X,Y) &= H(X)+H(Y|X) \\ &= H(Y)+H(X|Y) \end{aligned}$$

定理 3 を繰り返し用いることで以下の定理が成り立つ，

定理 4 (エントロピーのチェイン則)**.** $X_1,X_2,\cdots,X_n$ が $p(x_1,x_2,\cdots,x_n)$ に従っている時，

$$H(X_1,X_2,\cdots,X_n)=\sum_{i=1}^{n} H(X_i|X_{i-1},\cdots,X_1)$$

次に，２つの分布間の近さを測るダイバージェンス（divergence）と呼ばれる情報量を定義する．ダイバージェンスは KL 情報量（Kullback-Leibler Information）とも呼ばれる．$\mathcal{X}$ 上に定義されるあらゆる分布全体の集合を $\mathcal{P}_{\mathcal{X}}$ とおく．

定義 4 (ダイバージェンス)**.** 任意の 2 つの分布 $p, q \in \mathcal{P}_{\mathcal{X}}$ について

$$\begin{aligned} D(p||q) &= \sum_{x\in\mathcal{X}} p(x)\log\frac{p(x)}{q(x)} && (2.1)\\ &= \mathbb{E}_p \log\frac{p(X)}{q(X)} && (2.2)\end{aligned}$$

を p と q のダイバージェンス，あるいは KL 情報量と呼ぶ.

定理 5 (対数和の不等式)**.** 非負の数 $a_1, a_2, \cdots, a_n$ と $b_1, b_2, \cdots, b_n$ に対して

$$\sum_{i=1}^{n} a_i \log\frac{a_i}{b_i} \geq \left(\sum_{i=1}^{n} a_i\right)\log\frac{\sum_{i-1}^{n} a_i}{\sum_{i=1}^{n} b_i} \tag{2.3}$$

等号成立の必要十分条件は，全ての i に対して a_i/b_i が定数になること.

ダイバージェンスに定理 5 を適用することで以下の定理が成り立つ.

定理 6 (ダイバージェンスの非負性)**.** 任意の 2 つの分布 $p, q \in \mathcal{P}_{\mathcal{X}}$ に対して

$$D(p||q) \geq 0 \tag{2.4}$$

次に，2 つの確率変数 X, Y に対する相互情報量 $I(X;Y)$ を定義する.

定義 5 (相互情報量)**.**

$$\begin{aligned} I(X;Y) &:= \sum_{x,y} p(x,y)\log\frac{p(x,y)}{p(x)p(y)}\\ &= D(p(x,y)||p(x)p(y))\\ &= \sum_{x,y} p(x,y)\log\frac{p(x|y)}{p(x)}\\ &= -\sum_{x} p(x)\log p(x) - \left(-\sum_{x,y} p(x,y)\log p(x|y)\right)\\ &= H(X) - H(X|Y)\end{aligned}$$

また，ダイバージェンスの非負性より，相互情報量は以下の性質を持つ

定理 7.

$$I(X;Y) \geq 0$$

ただし，等号成立の必要十分条件は X と Y が独立であること

相互情報量の凸性について以下の定理が成り立つ．

定理 8. (X,Y) が $p(x,y) = p(x)p(y|x)$ に従って発生しているとする．X,Y に対する相互情報量 $I(X;Y)$ は， $p(y|x)$ を固定すると，$p(x)$ に関して上に凸な関数である．また，$p(x)$ を固定すると，$p(y|x)$ に関して下に凸な関数である．

最後に，後の議論で用いる不等式を紹介する．

定理 9 (Jensen の不等式)**.** 上に凸な関数 $f(x)$ と確率変数 X に対し

$$\mathbb{E}f(X) \geq f(\mathbb{E}X) \tag{2.5}$$

定理 10 (情報処理不等式)**.** 3 つの確率変数 X,Y,Z がマルコフ鎖 $X \to Y \to Z$ をなしている時，

$$I(X;Y) \geq I(X;Z) \tag{2.6}$$

2.2　通信路について

通信路を通してデータを送信すると，何らかの雑音が混入されて，送信したデータが歪んでしまうことがある．そのため，送信した記号 x に対して，受信される記号 y は x で条件付けられた確率法則に従う，すなわち通信路を条件付き確率　$W(y|x) = \Pr(Y = y|X = x)$ として考えることとする．

$$x \to \boxed{\text{channel } W(y|x)} \to y$$

図 2.1: 通信路のモデル

一般に, 通信路にはある集合 $\mathcal{X}$ から選ばれた文字 $x \in \mathcal{X}$ が入力され，ある集合 $\mathcal{Y}$ に含ま

れる文字 $y \in \mathcal{Y}$ が出力される．このとき $\mathcal{X}$ を通信路の入力アルファベット，$\mathcal{Y}$ を出力アルファベットとよぶ．本論文では，単純な通信路モデルとして以下の過程をおく．

1. 入出力アルファベット $\mathcal{X}, \mathcal{Y}$ は有限集合である．(離散)

2. 任意の $n \geq 1$ に対して，入力系列を $x^n = (x_1 x_2 \cdots x_n)$，出力系列を $y^n = (y_1 y_2 \cdots y_n)$ とした時，条件付き確率 $W^n(y^n|x^n)$ は

$$W^n(y^n|x^n) = \prod_{i=1}^{n} W(y_i|x_i)$$

で与えられる．

以上の条件を満たす通信路を定常離散無記憶通信路あるいは SDMC(Stationary Discrete Memolyless Channel) と呼ぶ．単に DMC とよぶことも多い．なお無記憶というのは時刻ごとの伝送が互いに独立であることを表している．本論文では以後，通信路を DMC として議論を進めていく．

2.3 通信路符号化について

通信路を通して信頼性のある通信を行なうために，図 2.2 に示すように,，符号器 (encoder) と復号器 (decoder) を用いて通信路符号化を行なう. 伝えたいメッセージは符号語に符号化されてから通信路に送られ, 復号器では受信語から送られてきたメッセージが推定される．ここで，通信路には雑音の少ないものもあるし，多いものもある．同じ数のビット誤りを訂正する符号も片方の通信路には十分であっても, もう一方には役に立たないことがある．そのため,「何ビットを訂正できるか」ということは必ずしも信頼性を良く表しているとは言えない事に留意しておく．まず，送信したいメッセージの集合を

$$S_n \longrightarrow \boxed{\text{encoder}} \xrightarrow{X^n} \boxed{\text{channnel}} \xrightarrow{Y^n} \boxed{\text{decoder}} \longrightarrow \hat{S}_n$$

図 2.2: 通信路符号化モデル

$\mathcal{M}_n = \{1, 2, \cdots, M_n\}$（メッセージ集合）とする．ここでメッセージ S_n は等しい確率で

$\mathcal{M}_n$ 上に値を取る確率変数と仮定する．すなわち，任意の $s \in \mathcal{M}_n$ に対し，

$$\Pr(S_n = s) = \frac{1}{M_n}$$

通信符号器では写像 $\phi_n : \mathcal{M}_n \to \mathcal{X}^n$ を用いてメッセージ $s \in \mathcal{M}_n$ を長さ n の系列 $\phi_n(s) = x^n(s) \in \mathcal{X}^n$ に変換する．この操作を符号化と呼び，メッセージ s を符号化した $x^n(s) = \phi_n(s)$ を符号語，この符号語の集合 $C_n = \{x^n(1), x^n(2), \cdots, x^n(M_n)\}$ をコードブック（符号）と呼ぶ．この時，通信路の入力 X^n は任意の $x \in \mathcal{M}_n$ に対して

$$\Pr(X^n = x^n(s)) = \frac{1}{M_n} \tag{2.7}$$

が成り立つ．一方，通信復号器では写像 $\psi_n : \mathcal{Y}^n \to \mathcal{M}_n$ を用いて受信系列 $y^n \in \mathcal{Y}^n$ をメッセージ $\hat{s}$ に変換する．この操作を復号化と呼び，復号写像 ψ_n は復号して $\hat{s} \notin \mathcal{M}_n$ となった場合，$\hat{s} = 0$ を出力するようにしている．以上が，通信路符号化モデルの概要である．

次に符号化レート（伝送レート）を定義する．符号化レート R とは通信路 1 回使用辺りに伝送できるビット数（ビット/通信路使用回数）を表している．

定義 6 (符号化レート)**.**

$$R := \frac{\log M_n}{n} \tag{2.8}$$

復号誤り確率は，送信したメッセージ S と復号したメッセージ $\hat{S}$ が異なる事象の確率 $\Pr(\hat{S} \neq S)$ の事である．一般に復号誤り確率は，通信路モデル $W_{Y|X}$ だけでなく，用いている符号・復号器（ϕ_n, ψ_n）に依存して定まる．以下では，(2.7) を用いて平均復号誤り確率 $\mathrm{Pe}(\phi_n, \psi_n)$ を定義する．

定義 7 (平均復号誤り確率)**.**

$$\mathrm{Pe}(\phi_n, \psi_n) := \frac{1}{M_n} \sum_{s \in \mathcal{M}_n} \Pr(\psi_n(\phi_n(s)) \neq s) \tag{2.9}$$

符号化レート R と平均復号誤り確率 $\mathrm{Pe}(\phi_n, \psi_n)$ は一般的にトレードオフの関係がある．例えば，通信路を使う回数 n を増やせば増やすほど $\mathrm{Pe}(\phi_n, \psi_n)$ が小さくなるが，R が大きくなる．逆に n を減らすと R が大きくなる．そこで通信路符号化では，平均復号誤り確率 $\mathrm{Pe}(\phi_n, \psi_n)$ を一定の値以下に抑えたもとで，符号化レート R を出来るだけ大きくすることを目的とする．問題を定式化するために，以下の定義を行う．

定義 8 (R が達成可能)**.** 任意の $n \geq 1$ に対し，ある符号・復号関数の組 (ϕ_n, ψ_n) が存在し，

$$\lim_{n\to\infty} \mathrm{Pe}(\phi_n, \psi_n) = 0$$
$$\varliminf_{n\to\infty} \frac{1}{n} \log M_n \geq R$$

の両方を満たすとき，符号化レート R は達成可能 (achievable) であるという．

さらに，達成可能な R の上限を通信路容量 $C(W)$ と呼び，以下で定義する．

定義 9 (通信路容量)**.**

$$C(W) := \sup\{R | R \text{ は達成可能である }\} \tag{2.10}$$

通信路容量 $C(W)$ は，通信路 W 固有の量である．符号化レート R が $C(W)$ よりも小さければ，R が達成可能となるような符号・復号器が存在し，$C(W)$ よりも大きければ R が達成可能となるような符号・復号器が存在しない事を意味している．

ここで，SDMC である $W(y|x)$ とその入力アルファベット $\mathcal{X}$ 上の確率分布 $P(x)$ に関して，同時確率は以下のように定まる．

$$P_{XY}(x, y) = P_X(x)W(y|x) \tag{2.11}$$

そこで相互情報量を以下のようにおく

$$
\begin{aligned}
I(X;Y) &= \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P(x,y)\log\frac{P(x,y)}{P(x)P(y)} && (2.12)\\
&= \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_x(x)Q(y|x)\log\frac{W(y|x)}{P(y)} && (2.13)\\
&=: I(P_X,W) && (2.14)
\end{aligned}
$$

シャノンは通信路容量を用いて以下の定理を示している．

定理 11 (通信路符号化定理)**.** 定常離散無記憶通信路 $(\mathcal{X}, W_{Y|X}, \mathcal{Y})$ に対して，通信路容量は以下のように与えられる．

$$C(W) = \max_{P_X} I(P_X, W) \tag{2.15}$$

この定理は，誤りを抑えた符号化レート R の上限を相互情報量という別の量で表している．それだけではなく，可能な全ての X の分布 P_X に関する相互情報量の最大化問題を解けば，通信路容量が求まることを示している．これらの点で非常に重要な定理といえる．しかし，この定理はどうやって「良い符号」を作るかを教えていない．それを作ることが符号理論の大きな課題の一つとなっている．

第3章　レート歪み理論

レート歪み理論は，一定の歪みを許しながら符号化レートをどこまで抑えられるかを示した理論である．実際，カメラやビデオ等の映像機器やオーディオ機器では，人間の視覚や聴覚に違和感を感じさせない程度の歪を許すことで符号化レートを抑えている．本章では，レート歪み理論の有歪み情報源符号化定理について，説明と定理の証明を述べる．本章で述べる証明手順は逆シャノン定理の証明でも使用する．本章での説明は Cover [3] を参考にした．

3.1　レート歪み理論の概要

一般に，情報源系列 $\mathcal{X}$ とそれに対応する復元系列は同じアルファベット上の系列であるとは限らない．そこで以下では，復元アルファベットを $\hat{\mathcal{X}}$ と書くこととする．なお，$\mathcal{X}, \hat{\mathcal{X}}$ は有限であると仮定する．

準備として，任意の $n > 1$ に対して，正整数 M_n を与えておく．更に $\mathcal{M}_n = \{1.\cdots, M_n\}$ とおき，レート歪み符号 (ϕ_n, ψ_n) を以下で定義する．

$$\phi_n : \mathcal{X}^n \to \mathcal{M}_n \tag{3.1}$$

$$\psi_n : \mathcal{M}_n \to \hat{\mathcal{X}}^n \tag{3.2}$$

この時，符号の 1 記号あたりのレートを $R = (1/n)\log(M_n)$ とする．

以上を用いて，次のような状況を考える．まず，任意の入力 $x^n \in \mathcal{X}^n$ を符号器 ϕ_n をで符号化する．次に，この時の出力を $u = \phi_n(x^n), u^i \in \mathcal{M}_n$ とし，u を恒等通信路を用いて送信する．最後に，u を復号器 ψ_n で復号し，その出力を $\hat{x^n} = \psi_n(u), \hat{x^n} \in \hat{\mathcal{X}}^n$ とする．次に歪み測度を定義する．

$$X^n \longrightarrow \boxed{\phi_n} \xrightarrow{U} \boxed{\text{ID}} \xrightarrow{U} \boxed{\psi_n} \longrightarrow \hat{X}^n$$

図 3.1: レート歪み理論の符号機と復号機

定義 10 (歪み測度と最大歪み測度)**.**

$$d : X \times \hat{X} \to [0, \infty)$$

を歪み測度と定義する.

次に，一つの歪み測度 d を長さ n の歪み測度 d_n に拡張する.

定義 11 (長さ n の歪み測度)**.** $n \geq 1$ と任意の $\mathbf{x}^n, \hat{\mathbf{x}}^n$ に対して

$$d(\mathbf{x}^n, \hat{\mathbf{x}}^n) = \frac{1}{n} \sum_{i=1}^{n} d(x_i, \hat{x}_i)$$

この d_n を用いて，歪みの期待値を定義する.

定義 12 (歪みの期待値)**.**

$$\mathbb{E}[d_n(X^n, \hat{X}^n)] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[d(x_i, \hat{x}_i)]$$

符号 (ϕ_n, ψ_n) を用いた場合に生じうる歪み $\Delta_n(\phi_n, \psi_n)$ は，情報系列 X^n に対する期待値で与えられる.

定義 13 (平均歪み)**.**

$$\Delta_n(\phi_n, \psi_n) := \mathbb{E}[d_n(X^n, \psi_n(\phi_n(X^n)))]$$

この値 $\Delta_n(\phi_n, \psi_n)$ を符号 (ϕ_n, ψ_n) の平均歪みと呼ぶ.

レート歪み理論の目的は，任意の $D > 0$ に対し，$\Delta_n(\phi_n, \psi_n) \leq D$ を満たしながら符号化レート R を出来るだけ小さくすることである．そこで以下を定義する.

定義 14 ((R, D) が達成可能 (achievable))**.** ある正整数の列 $\{M_n, n \geq 1\}$ とレート歪み符

号 $\{(\phi_n, \psi_n), n \geq 1\}$ が存在し，

$$\lim_{n\to\infty} \frac{1}{n} \log M_n \leq R \tag{3.3}$$

$$\overline{\lim_{n\to\infty}} \Delta_n(\phi_n, \psi_n) \leq D \tag{3.4}$$

が満たされる事を言う．

また，任意の $D > 0$ に対して

$$R(D) := \inf\{R : (R, D) \text{ が達成可能 }\}$$

とする．この時，定義 14 より，$R(D)$ は D に関して単調減少であることがわかる．

ここで Information Rate Distortion 関数 $R^{(I)}(D)$ を定義する．

定義 15 (Information Rate Distortion 関数)**.**

$$R^{(I)}(D) := \min_{p(\hat{x}|x):\sum_{(x,\hat{x})} p(x)p(\hat{x}|x)d(x,\hat{x})\leq D} I(X;\hat{X})$$

以上を用いて，有歪み情報源符号化定理を述べる．

定理 12 (有歪み情報源符号化定理)**.** 定常情報源 X と任意の歪み測度 d に対して

$$R(D) = R^{(I)}(D)$$

が成り立つ．

3.2 有歪み情報源符号化定理の証明

有歪み情報源符号化定理は，定常情報源 X と任意の歪み測度 d に対して

$$R(D) \leq R^{(I)}(D) \quad \text{（順定理）}$$

$$R(D) \geq R^{(I)}(D) \quad \text{（逆定理）}$$

を示せれば，$R(D) = R^{(I)}(D)$ が成り立つため，これらを証明していく．

3.2.1　有歪み情報源符号化の順定理

証明にあたり，以下を定義する．

定義 16 (経験分布 (タイプ))**.** 系列 $x^n = x_1x_2\cdots x_n \in \mathcal{X}^n$ に対して，以下で定義される $\mathcal{X}$ 上の分布 P_{x^n} を経験分布（タイプ）とよぶ

$$\forall a \in \mathcal{X}, P_{x^n}(a) := \frac{1}{n}N(a|x^n)$$

ただし，$N(a|x^n)$ は a が x^n 中に現れている回数である．

定義 17 (同時タイプ)**.** 系列 $x^n = x_1x_2\cdots x_n \in \mathcal{X}^n, \hat{x}^n = \hat{x}_1\hat{x}_2\cdots\hat{x}_n \in \hat{\mathcal{X}}^n$ の同時タイプとは

$$\forall a \in \mathcal{X}, \forall b \in \hat{\mathcal{X}}, P_{x^n\hat{x}^n}(a,b) := \frac{1}{n}N(a,b|x^n,\hat{x}^n)$$

ただし，$N(a,b|x^n,\hat{x}^n)$ は (a,b) が $(x^n,\hat{x}^n)$ 中に現れている回数である．

定義 18 (条件付きタイプ)**.** 系列 $x^n = x_1x_2\cdots x_n \in \mathcal{X}^n, \hat{x}^n = \hat{x}_1\hat{x}_2\cdots\hat{x}_n \in \hat{\mathcal{X}}^n$ の条件付きタイプ V とは

$$\forall a \in \mathcal{X}, \forall b \in \hat{\mathcal{X}}, V(b|a) = \frac{N(a,b|x^n,\hat{x}^n)}{N(a|x^n)}$$

定義 19 (V-shell)**.** 確率遷移行列 $V : \mathcal{X} \to \mathcal{Y}$ が与えられた時，$x^k \in \mathcal{X}^k$ に対し，条件付きタイプ V を有する $y^k \in \mathcal{Y}^k$ の集合を x の V-shell と呼び，$T_V(x^k)$ と書く．

定義 20 (強典型系列 (strongly typical))**.** $\mathcal{X}$ 上の任意の分布 P と定数 $\varepsilon > 0$ に対し，$x^n \in \mathcal{X}^n$ の時

$$\forall a \in \mathcal{X}, p(a) > 0 \text{ に対し}, |P_{x^n}(a) - p(a)| < \frac{\varepsilon}{|\mathcal{X}|}$$

かつ

$$p(a) = 0 \text{ ならば}, N(a|x^n) = 0$$

を満たすとき，x^n は強典型系列であると言う．

定義 21 (同時強典型系列 (jointly strongly typical)). 任意の定数 $\varepsilon > 0$ に対し，$(x^n, \hat{x}^n) \in \mathcal{X}^n \times \hat{\mathcal{X}}^n$ の時

$$\forall (a,b) \in \mathcal{X} \times \hat{\mathcal{X}}, p(a,b) > 0 \text{ に対し}, |P_{x^n \hat{x}^n}(a,b) - p(a,b)| < \frac{\varepsilon}{|\mathcal{X}||\hat{\mathcal{X}}|}$$

かつ

$$p(a,b) = 0 \text{ ならば}, N(a,b|x^n, \hat{x}^n) = 0$$

を満たすとき，$(x^n, \hat{x}^n)$ は同時強典型系列であると言う.

また，強典型系列集合 $A^X_{n,\varepsilon}$ と同時強典型系列集合 $A^{X,\hat{X}}_{n,\varepsilon}$ を

$$\begin{aligned} A^X_{n,\varepsilon} &:= \{x^n | x^n \text{が強典型系列}\} \\ A^{X,\hat{X}}_{n,\varepsilon} &:= \{(x^n, \hat{x}^n) | (x^n, \hat{x^n}) \text{ が同時強典型系列}\} \end{aligned}$$

とする．この時以下の定理が成り立つ.

定理 13. $X_i \sim^{i.i.d} p(x)$ の時，$\lim_{n\to\infty} \Pr(A^X_{n,\varepsilon}) = 1$

定理 14. $(X_i, \hat{X}_i) \sim^{i.i.d} p(x, \hat{x})$ の時，$\lim_{n\to\infty} \Pr(A^{X,\hat{X}}_{n,\varepsilon}) = 1$

定理 15. $\hat{X}_i \sim^{i.i.d} p(\hat{x})$ の時，$x^n \in A^X_{n,\varepsilon}$ ならば，

$$e^{-n(I(X;Y)+\varepsilon)} \leq \Pr((x^n, \hat{X}^n) \in A^{X,\hat{X}}_{n,\varepsilon}) \leq e^{-n(I(X;Y)-\varepsilon)}$$

以上を用いて，順定理の証明を行う．まず，符号化，復号化を以下のように作成する.

符号化 (ϕ_n) の作成

始めに $D > 0$ に対して $\mathbb{E}[d(X, \hat{X})] \leq D$ を満たす同時分布 $p_{X\hat{X}}$ を一つ選ぶ．ここで，n 次分布 $p_{\hat{X}^n}$ は $p_{X\hat{X}}$ の $\hat{X}^n$ に関する周辺分布の積で与えられる.

そして，任意の $n \geq 1$ と M_n に対して $C_n = \{\hat{x}^n(1), \hat{x}^n(2), \cdots, \hat{x}^n(M_n)\}$ を $p_{\hat{X}^n}$ に従ってランダムに発生させる (ランダムコーディング).

次に，情報源系列 X^n に対し，ある $i = \{1, 2, \cdots, M_n\}$ が存在し，$(X^n, \hat{x}^n(i)) \in A^X_{n,\varepsilon}$

を満たすならば，i を送信する．複数存在した場合は最小の i を選ぶ．もし満たさなければ，常に 1 を送信する．

復号化 (ψ_n) の作成

受け取った番号 i から対応する $\hat{x}^n(i)$ を出力する．

作成した符号 (ϕ_n, ψ_n) の平均歪み $\Delta_n(\phi_n, \psi_n)$ が D 以下になるような条件を求める．

平均歪みの評価

符号 (ϕ_n, ψ_n) の平均歪み $\Delta_n(\phi_n, \psi_n)$ について考える．

$J(C_n) := \{x^n | \exists i = \{1, 2, \cdots, M_n\} \quad (x^n, \hat{x}^n(i)) \in A_{n,\varepsilon}^{X,\hat{X}}\}$ とすると，

$$\Delta_n(\phi_n, \psi_n) \quad = \quad \sum_{x^n \in \mathcal{X}^n} p(x^n) d(x^n, \psi_n(\phi_n(x^n)))$$

ここで x^n について場合分けを行う．

$$\begin{aligned}
\Delta_n(\phi_n, \psi_n) \quad &= \quad \sum_{x^n \notin A_{n,\varepsilon}^X} p(x^n) d(x^n, \psi_n(\phi_n(x^n))) && (3.5) \\
&+ \quad \sum_{x^n \in A_{n,\varepsilon}^X \cap x^n \in J(C_n)} p(x^n) d(x^n, \psi_n(\phi_n(x^n))) && (3.6) \\
&+ \quad \sum_{x^n \in A_{n,\varepsilon}^X \cap x^n \notin J(C_n)} p(x^n) d(x^n, \psi_n(\phi_n(x^n))) && (3.7)
\end{aligned}$$

(3.5) $\sim$ (3.7) についてそれぞれ評価していく．

・ $x^n \notin A_{n,\varepsilon}^X$ の場合 (3.5)

$\lim_{n\to\infty} \Pr(x^n \notin A_{n,\varepsilon}^X) = 0$ より，

$$\lim_{n\to\infty} \sum_{x^n \notin A_{n,\varepsilon}^X} p(x^n) d(x^n, \psi_n(\phi_n(x^n))) = 0 \qquad (3.8)$$

・ $x^n \in A_{n,\varepsilon}^X \cap x^n \in J(C_n)$ の場合 (3.6)

歪みの期待値は

$$\mathbb{E}[d(x,\hat{x})] = \sum_{(a,b)\in\mathcal{X}\times\hat{\mathcal{X}}} p(a,b)d(a,b)$$

それに対し，経験分布の期待値は

$$d(x^n,\hat{x}^n) = \frac{1}{n}\sum_{(a,b)\in\mathcal{X}\times\hat{\mathcal{X}}} N(a,b|x^n,\hat{x}^n)d(a,b)$$

で与えられる．歪みの期待値と経験分布の期待値を比べると

$$|\mathbb{E}[d(x,\hat{x})] - d(x^n,\hat{x}^n)| = \left|\sum_a\sum_b d(a,b)\{p(a,b) - \frac{1}{n}N(a,b|x^n,\hat{x}^n)\}\right| \tag{3.9}$$

三角不等式より

$$\begin{aligned} |\mathbb{E}[d(x,\hat{x})] - d(x^n,\hat{x}^n)| &\leq \sum_a\sum_b \left|d(a,b)\{p(a,b) - \frac{1}{n}N(a,b|x^n,\hat{x}^n)\}\right| && (3.10)\\ &= \sum_a\sum_b d(a,b)\left|\{p(a,b) - \frac{1}{n}N(a,b|x^n,\hat{x}^n)\}\right| && (3.11)\end{aligned}$$

同時強典型系列の定義（定義 21）より

$$\begin{aligned} (3.11) &< \sum_x\sum_{\hat{x}} d(x,\hat{x})\frac{\varepsilon}{|\mathcal{X}||\hat{\mathcal{X}}|} && (3.12)\\ &\leq \varepsilon\cdot d_{max} && (3.13)\end{aligned}$$

以上より，歪み測度の最大値を D_{max} とすると，以下の式が与えられる．

$$\begin{aligned} d(x^n,\hat{x}^n) &= \mathbb{E}[d(X,\hat{X})] + d(x^n,\hat{x}^n) - \mathbb{E}[d(X,\hat{X})]\\ &\leq \mathbb{E}[d(X,\hat{X})] + \left|d(x^n,\hat{x}^n) - \mathbb{E}[d(X,\hat{X})]\right|\\ &\leq D + \varepsilon\cdot D_{max}\end{aligned}$$

結果として (3.8) の場合の平均歪みは

$$\begin{aligned}\sum_{x^n\in A^X_{n,\varepsilon}\cap x^n\in J(C_n)} p(x^n)d(x^n,\psi_n(\phi_n(x^n))) &\leq \sum_{x^n\in A^X_{n,\varepsilon}\cap x^n\in J(C_n)} p(x^n)\{D+\varepsilon\cdot D_{max}\}\\ &\leq D+\varepsilon\cdot D_{max} \end{aligned} \tag{3.14}$$

・ $x^n \in A^X_{n,\varepsilon} \cap x^n \notin J(C_n)$ の場合 (3.7)

$\mathrm{Pe(C_n)} = \sum_{x^n\in A^X_{n,\varepsilon}\cap x^n\notin J(C_n)} p(x^n)$ とすると

$$\sum_{x^n\in A^X_{n,\varepsilon}\cap x^n\notin J(C_n)} p(x^n)d(x^n,\psi_n(\phi_n(x^n))) \leq \mathrm{Pe(C_n)}\cdot D_{max} \tag{3.15}$$

(3.8), (3.14), (3.15) より，ある符号 (ϕ_n, ψ_n) の平均歪みは

$$\sum_{x^n\in\mathcal{X}^n} p(x^n)d(x^n,\psi_n(\phi_n(x^n))) \leq \delta + D + \varepsilon\cdot D_{max} + \mathrm{Pe(C_n)}\cdot D_{max} \tag{3.16}$$

ただし，δ, ε は任意の正の実数である．以上より，平均ひずみの評価は $\mathrm{Pe(C_n)}$ に帰着された．

(3.16) は，ランダムに作った符号 (ϕ_n, ψ_n) の平均歪みなので，順定理を証明するためにランダム符号全体の平均を求める．

ランダム符号全体の平均

1{ 条件式 } を，条件式を満たすならば 1，満たさなければ 0 を出力する関数とする．

$$\begin{aligned}\mathbb{E}[\mathrm{Pe(C_n)}] &= \sum_{C_n} p(C_n)\cdot \mathrm{Pe(C_n)}\\ &= \sum_{C_n} p(C_n) \sum_{x^n\in A^{X,\hat{X}}_{n,\varepsilon}} p(x^n)\cdot 1\{x^n\notin J(C_n)\}\\ &= \sum_{x^n\in A^{X,\hat{X}}_{n,\varepsilon}} p(x^n) \sum_{C_n} p(C_n)\cdot 1\{x^n\notin J(C_n)\}\end{aligned} \tag{3.17}$$

$\sum_{C_n} p(C_n)\cdot 1\{x^n \in J(C_n)\}$ について，x^n を一つ固定して考える．

$\hat{X}^n(i)(i=1,2,\cdots,M_n)$ は $p_{\hat{X}^n}$ に従って独立かつ同一に発生させた事に注意すると

$$\begin{aligned}
\sum_{C_n} p(C_n)\cdot 1\{x^n \in J(C_n)\} &= \sum_{\hat{x}^n(1)\in\hat{\mathcal{X}}^n}\sum_{\hat{x}^n(2)\in\hat{\mathcal{X}}^n}\cdots\sum_{\hat{x}^n(M_n)\in\hat{\mathcal{X}}^n} \\
&\quad \times p(\hat{x}^n(1))p(\hat{x}^n(2))\cdots p(\hat{x}^n(M_n))1\{x^n \notin J(C_n)\} \\
&= \Pr\{x^n \notin J(\hat{x}^n(1),\hat{x}^n(2),\cdots,\hat{x}^n(M_n))\} \\
&= \Pr\{(x^n,\hat{x}^n(1)) \notin A_{n,\varepsilon}^{X,\hat{X}} \cap\cdots\cap (x^n,\hat{x}^n(M_n)) \notin A_{n,\varepsilon}^{X,\hat{X}}\} \\
&= \Pr\{(x_n,\hat{X}^n(1)) \in A_{n,\varepsilon}^{X,\hat{X}}\}^{M_n}
\end{aligned}$$

以上と定理 15 より以下の式が成り立つ．

$$\begin{aligned}
\mathbb{E}[\mathrm{Pe(C_n)}] &= \sum_{x^n\in A_{n,\varepsilon}^{X,\hat{X}}} p(x^n)\Pr\{(x^n,\hat{x}^n) \notin A_{n,\varepsilon}^{X,\hat{X}}\}^{M_n} \\
&\le \sum_{x^n\in A_{n,\varepsilon}^{X,\hat{X}}} p(x^n)(1-e^{-n(I(X;\hat{X})+\varepsilon_1)})^{M_n} \qquad (3.18)
\end{aligned}$$

ここで，後の議論に必要な定理を示す．

定理 16. 任意の実数 $0\le A\le 1, 0\le B\le 1$ と正整数 m に対して以下の式が成り立つ

$$(1-A\cdot B)^m \le 1-A+e^{-B\cdot m}$$

定理 16 の変数を $A=1, B=e^{-n(I(X;\hat{X})+\varepsilon_1)}, m=M_n$ と置き換えると

$$(3.18) \le \sum_{x^n\in A_{n,\varepsilon}^{X,\hat{X}}} p(x^n)\cdot\exp(-M_n\cdot\exp(-n(I(X;\hat{X})+\varepsilon_1)))$$

$M_n = e^{nR}$ とすると

$$\mathbb{E}[\mathrm{Pe(C_n)}] \le \exp(\exp(-n(I(X;\hat{X})-R+\varepsilon_1))) \qquad (3.19)$$

以上より (3.16) に (3.19) を代入すると，ある符号 $(\hat{\phi}_n, \hat{\psi}_n)$ に対する平均歪みは

$$\Delta_n(\hat{\psi}_n, \hat{\phi}_n) \leq \delta + D + \varepsilon \cdot D_{max} + D_{max} \cdot \exp(\exp(-n(I(X;\hat{X}) - R + \varepsilon_1)))$$

$R > I(X;\hat{X}) + \varepsilon_1$ を満たす R を考えると

$$\lim_{n\to\infty} \exp(\exp(-n(I(X;\hat{X}) - R + \varepsilon_1))) = 0$$

$D_{max} > 0$ は定数なので，十分大きな n に対して

$$D_{max} \cdot \exp(\exp(-n(I(X;\hat{X}) - R + \varepsilon_1))) \leq \varepsilon$$

と変形できる．以上より

$$\Delta_n(\hat{\psi}_n, \hat{\phi}_n) \leq D + \delta + 2\varepsilon \cdot D_{max}$$

ここで，$\varepsilon, \delta > 0$ は任意に小さく取れるので，ある符号 $(\hat{\psi}_n, \hat{\phi}_n)$ が存在して

$$\begin{cases} \lim_{n\to\infty} \Delta_n(\hat{\psi}_n, \hat{\phi}_n) \leq D + \delta \\ \lim_{n\to\infty} \frac{1}{n} \log M_n = R \end{cases}$$

を満たす．よって $R > I(X;\hat{X})$ ならば $(R, D+\delta)$ は達成可能であることが導かれた．ここで

$$\begin{cases} R^{(I)}(D) = \inf\{R | R > I(X;\hat{X})\} \\ R(D+\delta) = \inf\{R | (R, D+\delta) \text{ は達成可能 }\} \end{cases}$$

を比較すると，包含関係より

$$\forall \delta > 0, R^{(I)}(D) \geq R(D+\delta)$$

が導かれる．$R(D+\delta)$ は下に凸かつ連続な関数なので，連続性より

$$R^{(I)}(D) \geq \lim_{\delta \downarrow 0} R(D+\delta) = R(D)$$

結果として，以下の定理が得られた．

定理 17 (有歪み情報源符号化の順定理)**.** 任意の $D > 0$ と定常情報源 X に対して

$$R(D) \leq R^{(I)}(D)$$

が成り立つ．

次に有歪み情報源符号化逆定理の証明を行う．

3.2.2 有歪み情報源符号化の逆定理

本節では，以下の定理の証明を行う．

定理 18 (有歪み情報源符号化の逆定理)**.** 任意の $D > 0$ と定常情報源 X に対して

$$R(D) \leq R^{(I)}(D)$$

が成り立つ．

定理 18 の証明を行う前に議論で必要となる定理を示す．

定理 19. $R^{(I)}(D)$ は D に関して単調非増加でかつ，下に凸 (concave) の関数である．

証明 1.（単調非増加性の証明）D を大きくすることで $I(X;\hat{X})$ を最小にする $p(x,\hat{x})$ の範囲を広げることができるため，D に対して非増加である．
（下に凸の照明）2 つの異なる平均歪み D_1, D_2 を設定した時，任意の $0 < \lambda < 1$ に対して

$$R^{(I)}(\lambda \cdot D_1 + (1-\lambda)D_2) \leq \lambda \cdot R^{(I)}(D_1) + (1-\lambda)R^{(I)}(D_2)$$

が成立すれば，下に凸となる．

$R^{(I)}(D_1), R^{(I)}(D_2)$ を与える $p(x,\hat{x})$ をそれぞれ $p_1(x,\hat{x}), p_2(x,\hat{x})$ で表し，以下を定義する．

$$\begin{aligned} I(p_1) &:= R^{(I)}(D_1) = \min_{\sum p_1(x,\hat{x})d(x,\hat{x}) \le D_1} I(X;\hat{X}) \\ I(p_2) &:= R^{(I)}(D_2) = \min_{\sum p_2(x,\hat{x})d(x,\hat{x}) \le D_2} I(X;\hat{X}) \\ p_\lambda &:= \lambda \cdot p(x,\hat{x}) + (1-\lambda)p_2(x,\hat{x}) \\ D_\lambda &:= \lambda \cdot D_1 + (1-\lambda)D_2 \end{aligned}$$

この時，p_λ についての平均歪みを求めると

$$\bar{d}(p_\lambda) = \sum_x \sum_{\hat{x}} d(x,\hat{x})\{\lambda \cdot p_1(x,\hat{x}) + (1-\lambda)p_2(x,\hat{x})\} \tag{3.20}$$

ここで定義より

$$D_1 = \sum_x \sum_{\hat{x}} p_1(x,\hat{x})d(x,\hat{x}), D_2 = \sum_x \sum_{\hat{x}} p_2(x,\hat{x})d(x,\hat{x})$$

となるので，(3.20) に代入すると

$$\begin{aligned} \bar{d}(p_\lambda) &= \lambda \cdot D_1 + (1-\lambda)D_2 \\ &= D_\lambda \end{aligned}$$

が成り立つ．よって

$$R^{(I)}(D_\lambda) \le I(p_\lambda)$$

$I(X;\hat{X})$ は $p(\hat{x}|x)$ に対して下に凸の関数より

$$\begin{aligned} R^{(I)}(D_\lambda) &\le I(p_\lambda) \\ &\le \lambda \cdot I(p_1) + (1-\lambda)I(p_2) \\ &= \lambda \cdot R(D_1) + (1-\lambda)R(D_2) \end{aligned}$$

が得られ，下に凸性が示された．

この定理を用いて，逆定理（定理 18）の証明を行う．

逆定理の証明

(R, D) が達成可能である $(2^{nR}, n)$ レート歪み符号 (ϕ'_n, ψ'_n) が存在するとし，$\hat{X}^n = \psi_n(\phi_n(X^n))$ とおく．この時，エントロピーの最小値より

$$
\begin{aligned}
nR &\geq H(\phi_n(X^n)) && (3.21)\\
&\geq H(\phi'_n(X^n)) - H(\phi'_n(X^n)|X^n) && (3.22)\\
&= I(X; \phi'_n(X^n)) && (3.23)\\
&\geq I(X; \hat{X}) && (3.24)\\
&= H(X^n) - H(X^n|\hat{X}^n) && (3.25)\\
&= \sum_{i=1}^{n} H(X_i) - H(X^n|\hat{X}^n) && (3.26)\\
&= \sum_{i=1}^{n} H(X_i) - \sum_{i}^{n} H(X_i|\hat{X}^n, X_{i-1}, \cdots, X_1) && (3.27)\\
&\geq \sum_{i=1}^{n} H(X_i) - \sum_{i=1}^{n} H(X_i|\hat{X}_i) && (3.28)\\
&= \sum_{i=1}^{n} I(X_i; \hat{X}_i) && (3.29)\\
&\geq \sum_{i=1}^{n} R^{(I)}(\mathbb{E}[d(X_i, \hat{X}_i)]) && (3.30)\\
&= n\left(\sum_{i=1}^{n} \frac{1}{n} R^{(I)}(\mathbb{E}[d(X_i, \hat{X}_i)])\right) && (3.31)\\
&\geq nR^{(I)}\left(\sum_{i=1}^{n} \frac{1}{n} \mathbb{E}[d(X_i, \hat{X}_i)]\right) && (3.32)\\
&= nR^{(I)}\left(\mathbb{E}[d(X_i, \hat{X}_i)]\right) && (3.33)
\end{aligned}
$$

なお，(3.22) はエントロピーの非負性，(3.24) は情報処理不等式，(3.26) に関しては X_i がそれぞれ独立のため，(3.27) はエントロピーのチェイン則，(3.28) は $H(X_i|\hat{X}^n X^{i-1}) \leq$

$H(X_i|\hat{X}_i)$ より，(3.30) は $R^{(I)}(D)$ の定義，(3.32) は Jensen の不等式よりそれぞれ成り立つ．ここで (R, D) が達成可能であるので，任意の $\delta > 0$ に対して n を十分に大きく取ると

$$\mathbb{E}[d(X_i, \hat{X}_i)] \leq D + \delta$$

が成り立ち，また $R^{(I)}(D)$ は D に関して単調減少なので

$$R \geq R^{(I)}(D + \delta)$$

が成り立つ．$R^{(I)}(D)$ は D に関して連続で $\delta > 0$ は任意なので

$$R \geq R^{(I)}(D)$$

以上より，$R(D)$ の定義から

$$R(D) \geq R^{(I)}(D)$$

となり定理 18 が証明できた．

結果として，定理 17 と定理 18 を示したことにより有歪み情報源符号化定理が正しいことが導きだせた．本章で用いた順定理，逆定理の証明手順は，後に出てくる逆シャノン定理の証明でも用いる．

さて，このレート歪み関数 $R(D)$ における相互情報量の最小化問題は，前章で述べた通信路容量 $C(W)$ における相互情報量の最大化とちょうど対を成す関係にある．通信路容量の場合，通信路 $W_{Y|X}$ が与えられていて，P_X を媒介変数として $I(X;Y)$ を最大化させている．レート歪み関数の場合は，情報源の分布 P_X を固定して，$P_{\hat{X}|X}$ を媒介変数として $I(\hat{X}|X)$ を最小化させている．また，レート歪み関数では歪みに関する制約が追加されている．

第4章　逆シャノン定理

前述の通信路符号化定理の対をなす符号化の問題として，雑音のない通信路を用いて雑音のある通信路を再現できるかという問題がある．それが可能であることを示したのが逆シャノン定理 [1] である．本章では，逆シャノン定理の目的を述べ，前章のレート歪み理論と同様の証明方法で逆シャノン定理を証明する．

4.1　逆シャノン定理の目的

通信路符号化定理とは，通信路 W の通信路容量 $C(W)$ を定め，符号化レート R が $R < C(W)$ ならば誤り確率 $\mathrm{Pe}(\phi_n, \psi_n)$ を漸近的に 0 にできる符号・復号器 (ϕ_n, ψ_n) が存在する事を示した定理である．この事はすなわち，$R < C(W)$ ならば任意の与えられた通信路 W を恒等通信路として扱えるということである．それに対して逆シャノン定理 [1] は，$R > C(W)$ ならば，恒等通信路を用いることにより，漸近的に誤りなく任意の与えられた通信路 W を再現可能である事を示している．本節では，任意の与えられた通信路 W の再現方法について述べる．

$$x^n \to \boxed{\text{channel } W^n(y|x)} \to y^n$$

図 4.1: 再現したい通信路

任意の x^n に対して，恒等通信路を k 回用いて，通信路容量 $C(W)$ の通信路 W を再現する状況を考える．本章で用いる符号・復号器 (ϕ_n, ψ_n) を以下で定義する．

$$\phi_n : \mathcal{X}^n \to \mathcal{M}_n$$

$$\psi_n : \mathcal{M}_n \to \mathcal{Y}^n$$

ここで，$\mathcal{M}_n = \{1, 2, \cdots, M_n\}$，$\mathcal{X} = \mathcal{Y} = \{0, 1\}$ とする．

まず，任意の入力 $x^n \in \mathcal{X}^n$ を符号器 ϕ_n を用いて符号化する．次に，この時に出力された値を $u_i^k = \phi_n(x^n), u_i^k \in \mathcal{M}_n$ $(i = 1, 2, \cdots, M_n)$ とし，これを恒等通信路で送信する．最後に u_i^k を復号器 ψ_n を用いて $y^n = \psi_n(\phi_n(x^n)), y^n \in \mathcal{Y}^n$ を復号する．

$$X^n \longrightarrow \boxed{\phi_n} \xrightarrow{U_i^k} \boxed{\mathrm{ID}} \xrightarrow{U_i^k} \boxed{\psi_n} \longrightarrow Y^n$$

図 4.2: 逆シャノン定理による通信路 W のシミュレーション

入力 x^n を固定した場合，シミュレーションによる出力 y^n の経験分布は，任意の $a \in \mathcal{X}, b \in \mathcal{Y}$ に対して，条件付きタイプ $V(b|a)$ で書ける．この経験分布が通信路 W に近くあってほしい．そのため，

$$\max_{a,b} |V(b|a) - W(b|a)| \leq \frac{\varepsilon}{|\mathcal{X}||\mathcal{Y}|} \tag{4.1}$$

が成り立つような符号・復号器を用意すれば良い．なお，このことは x^n が入力された時，条件付き確率 $W(y|x)$ に対する $T_W(x^n)$（定義 19）の要素 $y^n \in T_W(x^n)$ を出力することとほぼ等しい．ここで (4.1) の左辺を，本論文では ”ずれ率 ” と記述する．以下でずれ率を定義する．

定義 22 (ずれ率)**.** 再現したい通信路 W とシミュレーション (ϕ_n, ψ_n) に対し，$x^n \in \mathcal{X}$ のとき，ずれ率 $d(x^n, \phi_n, \psi_n)$ を

$$d(x^n, \phi_n, \psi_n) := \max_{a,b} |V(b|a) - W(b|a)|$$

とする．ここで $V(b|a) = N(a, b|x^n, y^n)/N(a|x^n)$ である．

逆シャノン定理の目的は，ずれ率を一定の値以下に抑えたもとで，符号化レート R をできるだけ小さくすることである．そこで以下を定義する．

定義 23 (P_X について R が達成可能)**.** ある正整数 M_n と $(\phi_n, \psi_n)(n = 1, 2, \cdots)$ が存在し，

$$\lim_{n\to\infty} \mathbb{E}_{X^n}[d(X^n, \phi_n, \psi_n)] = 0$$

かつ

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log M_n \leq R$$

が満たされるとき，P_X について R が達成可能であるという．

また，R の下限について以下を定義する．

$$C_R(P_X, W) = \inf\{R | P_X \text{について } R \text{が達成可能}\}$$

以上を用いて，逆シャノン定理を述べる．

定理 20 (逆シャノン定理)**.**

$$C_R(P_X, W) = I(P_X, W)$$

4.2 逆シャノン定理の証明

本節では逆シャノン定理の証明を行う．逆シャノン定理は

$$C_R(P_X, W) \leq I(P_X, W) \quad \text{（順定理）}$$
$$C_R(P_X, W) \geq I(P_X, W) \quad \text{（逆定理）}$$

を示せば成り立つため，これらを証明していく．

まず，逆シャノン定理の順定理を求める．レート歪み理論で定義した強典型系列集合 $A^X_{n,\varepsilon}$ と同時強典型系列集合 $A^{X,Y}_{n,\varepsilon}$ を用いて，任意の通信路 W を再現する．通信路 W は確率遷移行列 $W : \mathcal{X} \to \mathcal{Y}$ で与えられる． この時 n を大きくすると，通信路 W は，任意の入力 $x^n \in A^X_{n,\varepsilon}$ に対し，$(x^n, y^n) \in A^{X,Y}_{n,\varepsilon}$ となる y^n を出力するはずである．そこで，以下の符号・復号化により通信路 W を再現する．

符号化 (ϕ_n) の作成

互いに独立でかつ同じ分布 p_{Y^n} を持つ M_n 個の確率変数の列 $Y^n(1), Y^n(2), \cdots, Y^n(M_n)$

を符号語とする（ランダムコーディング）．このときの分布 p_{Y^n} は p_{X^n} の周辺化で得られる．

$$p_{Y^n}(y^n) = \sum_{x^n} p_{X^n}(x^n) W(y^n|x^n)$$

情報系列 $\mathbf{X}^n$ に対し，ある $i = \{1, 2, \cdots, M_n\}$ が存在して，$(\mathbf{X}^n, Y^n(i)) \in A_{n,\varepsilon}^{X,Y}$ を満たすならば，i を送信する．複数存在した場合は最小の i を選ぶ．もし満たさなければ，常に 1 を送信する．

復号化 (ψ_n) の作成

受け取った番号 i から対応する $y^n(i)$ を出力する．

すなわち，上記の方法での符号化の場合，

$$\mathcal{S} = \{i | (x^n, Y^n(i)) \in A_{n,\varepsilon}^{X,Y}\}$$

とおくと，符号器 ϕ_n，復号器 ψ_n は，それぞれ $x^n \in \mathcal{X}^n$ と $i \in \mathcal{M}_n$ に対して

$$\begin{aligned} \phi_n(x^n) &= \begin{cases} \min \mathcal{S} & \mathcal{S} \neq \emptyset \\ 1 & \text{otherwise} \end{cases} \\ \psi_n(i) &= Y^n(i) \end{aligned}$$

として与えられる．

ずれ率の期待値の評価

上記の方法で作った符号・復号器 (ϕ_n, ψ_n) のずれ率 $d(x^n, \phi_n, \psi_n)$ の期待値を評価する．$J(C_n) := \{x^n | \exists i = \{1, 2, \cdots, M_n\} (x^n, y^n) \in A_{n,\varepsilon}^{X,Y}\}$ とすると，

$$\mathbb{E}_{X^n}[d(X^n, \phi_n, \psi_n)] = \sum_{x^n \in \mathcal{X}^n} p(x^n) d(x^n, \phi_n, \psi_n) \tag{4.2}$$

ここで x^n について場合分けを行う

$$\begin{aligned}\mathbb{E}_{X^n}[d(X^n,\phi_n,\psi_n)] &= \sum_{x^n\notin A^X_{n,\varepsilon}} p(x^n)d(x^n,\phi_n,\psi_n) &(4.3)\\ &+ \sum_{x^n\in A^X_{n,\varepsilon}\cap x^n\in J(C_n)} p(x^n)d(x^n,\phi_n,\psi_n) &(4.4)\\ &+ \sum_{x^n\in A^X_{n,\varepsilon}\cap x^n\notin J(C_n)} p(x^n)d(x^n,\phi_n,\psi_n) &(4.5)\end{aligned}$$

(4.3) ∼ (4.5) についてそれぞれ評価していく.

・ $x^n \notin A^X_{n,\varepsilon}$ の場合 (4.3)

$\lim_{n\to\infty}\Pr(x^n \notin A^X_{n,\varepsilon}) = 0$ より,

$$\lim_{n\to\infty}\sum_{x^n\notin A^X_{n,\varepsilon}} p(x^n)d(x^n,\phi_n,\psi_n) = 0 \tag{4.6}$$

が成り立つ.

・ $x^n \in A^X_{n,\varepsilon} \cap x^n \in J(C_n)$ の場合 (4.4)

強典型系列の定義（定義 21）より

$$\begin{aligned}d(x^n,\phi_n,\psi_n) &\le \frac{\varepsilon}{|\mathcal{X}||\mathcal{Y}|}\\ &\le \varepsilon\end{aligned}$$

このことから

$$\sum_{x^n\in A^X_{n,\varepsilon}\cap x^n\in J(C_n)} p(x^n)d(x^n,\phi_n,\psi_n) \le \varepsilon \tag{4.7}$$

が成り立つ.

・ $x^n \in A^X_{n,\varepsilon} \cap x^n \notin J(C_n)$ の場合 (4.5)

$\mathrm{Pe(C_n)} := \sum_{x^n\in A^X_{n,\varepsilon}\cap x^n\notin J(C_n)} p(x^n)$,$d_{max}(\phi_n,\psi_n) := \max_{x^n} d(x^n,\phi_n,\psi_n)$ とすると,

$$\sum_{x^n\in A^X_{n,\varepsilon}\cap x^n\notin J(C_n)} p(x^n)d(x^n,\phi_n,\psi_n) \le \mathrm{Pe}(C_n)\cdot d_{max}(\phi_n,\psi_n) \tag{4.8}$$

(4.6) ∼ (4.8) の結果より，ある符号 (ϕ_n, ψ_n) のずれ率の期待値は

$$\sum_{x^n \in \mathcal{X}^n} p(x^n) d(x^n, \phi_n, \psi_n) \leq \varepsilon + \mathrm{Pe(C_n)} \cdot d_{max}(\phi_n, \psi_n) \tag{4.9}$$

ただし，ε は任意の正の実数である．よって，平均ひずみの評価は $\mathrm{Pe(C_n)}$ に帰着された．(4.9) は，ランダムに作った符号 (ϕ_n, ψ_n) のずれ率の期待値なので，順定理を証明するためにランダム符号全体の平均を求める．

ここで，前章で証明した有歪み情報源符号化と同様の証明方法を用いることにより，符号化レート R の条件が $R > I(P_X; W)$ ならば，P_X について R は達成可能であることが導かれる．以上のことと包含関係より以下の定理が得られる．

定理 21 (逆シャノン定理の順定理)**.** 任意の与えられた通信路 W と P_X に対して

$$C_R(P_X, W) \leq I(P_X; W)$$

が成り立つ．

また，逆シャノン定理の逆定理についても有歪み情報源符号化の逆定理と同様の手順で証明でき，

定理 22 (逆シャノン定理の逆定理)**.** 任意の与えられた通信路 W と P_X に対して

$$C_R(P_X, W) \geq I(P_X; W)$$

が成り立つ．

結果として，定理 21 と定理 22 を示したことにより逆シャノン定理が正しいことが導き出せた．ここで，$C_R(P_X, W)$ の最大値に関して以下のことが言える．

$$\max C_R(P_X, W) = \max I(P_X, W) \tag{4.10}$$

逆シャノン定理では通信路 W が与えられているため，通信路容量 $C(W) = \max_{P_X} I(P_X, W)$

がわかる．以上より

$$\max_{P_X} C_R(P_X, W) = \max_{P_X} I(P_X, W) \tag{4.11}$$

$$= C(W) \tag{4.12}$$

が成り立つ．

第5章　BCH符号

本章では，本研究の実験で使用する BCH 符号の定義，符号化，復号化の方法について述べる．BCH 符号とは巡回符号とよばれるものの 1 種であり，最もよく研究されている誤り訂正符号の 1 つである．なお，符号化を構成するアルファベットは 0,1 すなわち 2 元符号のみを考えることにする．

5.1　巡回符号の定義

線型符号の実用上重要な符号の一つとして巡回符号がある．以下で巡回符号を定義する．

定義 24 (巡回符号). 線型符号 C_n が次の性質を持つとき巡回符号という．

$$\forall(c_0c_1\cdots c_{n-1}) \in C_n \Rightarrow (c_{n-1}c_0\cdots c_{n-2}) \in C_n \tag{5.1}$$

巡回符号の場合は生成行列でその性質を考えると表現が煩雑になるため，多項式表現を定義する．

定義 25 (符号の多項式表現). 符号 $(c_0c_1\cdots c_{n-1}) \in \mathbb{F}^n$ に $\mathbb{F}$ 上の多項式

$$f(x) = c_0 + c_1x + c_2x^n + \cdots + c_{n-1}x^{n-1} \tag{5.2}$$

を対応させ，$(c_0c_1\cdots c_{n-1})$ の多項式表現と呼ぶことにする．

ここで $p(x) = x^n - 1 (= x^n + 1)$ とおくと，$f(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}$ に対して，

$$\begin{aligned} xf(x) &= a_0 x + a_1 x^2 + \cdots + a_{n-1} x^n & (5.3) \\ &= a_0 x + a_1 x^2 + \cdots + a_{n-2} x^{n-1} + a_{n-1} + a_{n-1}(x^n - 1) & (5.4) \end{aligned}$$

$p(x)$ で mod を取ると

$$(5.4) \equiv a_{n-1} + a_0 x + a_1 x^2 + \cdots + a_{n-2} x^{n-1} \pmod{x^n - 1} \quad (5.5)$$

が成り立つ．従って $(c_0 c_1 \cdots c_{n-1})$ を $(c_{n-1} c_0 \cdots c_{n-2})$ を変換する操作は，$p(x)$ による剰余をなす環 $\mathcal{F}[x]/p(x)$ 上では，多項式 $f(x)$ に x を掛けることに他ならない．この時，符号 C_n が巡回符号であるとは $f(x), g(x) \in C_n, \alpha \in \mathbb{F}$ に対して，以下の条件を満たすことと等しい．

1. $f(x) + g(x) \in C_n$
2. $\alpha f(x) \in C_n$
3. $xf(x) (\mod x^n - 1) \in C_n$

なお，条件 $1, 2$ は線型符号であることの条件である．この時，巡回符号について以下の定理が成り立つ

定理 23. C_n を $\mathbb{F}[x]/(x^n - 1)$ 上の巡回符号とし，$g(x)$ を C_n の 0 でない最小次数の符号語とすると，C_n の任意の符号語 $f(x)$ は適当な多項式 $a(x)$ により，

$$f(x) = a(x) g(x) \quad (5.6)$$

と書ける．この時，C_n はイデアル (ideal) であると言う．

この定理より巡回符号 C_n はある多項式 $g(x)$ の倍数となっていることがわかる．この多項式 $g(x)$ は巡回符号の多項式の中で最高次の項の係数が 1 のものを選べば良い．この多項式 $g(x)$ を巡回符号の生成多項式と呼ぶ．また，線型符号について以下を定義する．

定義 26 (生成行列). C_n を k 次元の符号語長 n の線型符号とすると，C_n の符号語の中に k 個の線形独立なベクトルが存在する．それらの符号語を行として並べた $n \times k$ の行列を C_n の生成行列 G とよぶ．この時，以下の式が成り立つ．

$$C_n = \{G\mathbf{u} | \mathbf{u} \in \mathbb{F}^k\}$$

定義 27 (パリティ検査行列). C_n を k 次元の符号語長 n の線型符号とする．この符号は $n \times (n-k)$ 行列 H を用いて

$$C_n = \{\mathbf{x} \in \mathbb{F}^n | \mathbf{x}H = \mathbf{0}\}$$

と書くことができる．この行列 H を C_n のパリティ検査行列という．

最後に重要な式として Vandermonde の行列式について述べる．

定理 24 (Vandermonde の行列式). 体 K 上の任意の元 $a_1 \cdots a_t$ に対して $t \times t$ 行列を

$$A_t = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_t \\ a_1^2 & a_2^2 & \cdots & a_t^2 \\ a_1^{t-1} & a_2^{t-1} & \cdots & a_t^{t-1} \end{pmatrix} \tag{5.7}$$

とすると

$$\det A = \prod_{1 \le i < j \le t} (a_i - a_j) \tag{5.8}$$

5.2 BCH符号の定義

巡回符号の中で実用上重要な符号が Bose と Ray-Chaudhurui(1960)，それとは別に Hocquenghem(1959) により発見された．彼らに因んで BCH 符号と呼ばれる．BCH 符号は，誤り訂正が効率的に行なえ，復号も容易である．それに加えて，冗長度をある程度自由に設定できるという利点を持っている．以下で BCH 符号の定義を行う．

定義 28 (BCH 符号)**.** m を任意の整数とし，n を 2^m-1 の約数であるとする．α を $GF(2^m)$ における 1 の原始 n 乗根とすし，$g(x)$ を x^n-1 を割り切る $(n-k)$ 次多項式とする．$g(x)=0$ の根 $\beta_1,\cdots,\beta_{n-k}$ の中に $\alpha^{\alpha},\alpha^{\alpha+1},\cdots,\alpha^{\alpha+\delta-1}$ なる連続した δ 個の 1 の n 乗根が存在したとすると，$g(x)$ を生成多項式とする巡回符号の最小距離は少なくとも $\delta+1$ 以上になる．

ただし，最小距離とはハミング距離の最小値である．上の定理を満たす符号を，長さ n，計画距離 $\delta+1$ の BCH 符号と呼び，特に $n=2^m+1$ の場合，原始 BCH 符号と呼ぶ．[6] また，BCH 符号のパリティ検査行列について以下の定理が成り立つ．

定理 25. 符号長 n，設計距離 δ の q 元 *BCH* 符号のパリティ検査行列は

$$H=\begin{pmatrix} 1 & \beta^{l} & (\beta^{l})^2 & \cdots & (\beta^{l})^{n-1} \\ 1 & \beta^{l+1} & (\beta^{l+1})^2 & \cdots & (\beta^{l+1})^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \beta^{l+\delta-2} & (\beta^{l+\delta-2})^2 & \cdots & (\beta^{l+\delta-2})^{n-1} \end{pmatrix} \tag{5.9}$$

で与えられる．

また，符号長 n，メッセージ長 k，最小距離 d の BCH 符号を (n,k,d)-BCH 符号と呼ぶ．その際の符号化レートは以下のとおりである．

$$R=\frac{\log 2^k}{n}=\frac{k}{n} \tag{5.10}$$

BCH 符号を用いる利点として以下の点が挙げられる．

1. 符号語長やメッセージ長を自由に設定でき，さらにそのような符号の設計方法が明示できる．
2. 保証されたハミング距離の下界に対して，保証される個数の誤りを訂正する具体的なアルゴリズムが知られている．

本研究では，主に 1 の理由から BCH 符号を採用した．

5.3 BCH符号の符号化方法

$C_n = \{c_1^n, c_2^n, \cdots c_{2^k}^n\}$ を $g(x)$ を生成多項式とする BCH 符号とし，$\deg\ g(x) = n - k$ とする．今，k ビットのメッセージ $(a_0 a_1 \cdots a_{k-1})$ を符号化することを考える．$a(x) = a_0 + a_1 x + \cdots + a_{k-1} x^{k-1}$ とおき，$a(x)$ を情報多項式 （information polynomial) とよぶ．符号化方法は様々だが，本研究では以下の方法を用いて符号化した．

$$x^{n-k} a(x) \equiv r(x) \pmod{g(x)} \tag{5.11}$$

となる $r(x)$ を求め，$c(x) = x^{n-k} a(x) + b(x)$ を符号語として送信することにより符号化を行う．この方法で符号化を行うと符号語 $c(x) = (c_0 \cdots c_{n-k-1} c_{n-k} \cdots c_{n-1})$ において，後半部分の $c_{n-k} \cdots c_{n-1}$ がメッセージと一致する．

5.4 シンドロームについて

BCH 符号で符号化された符号語 $c(x)$ に誤りベクトルを多項式で表した $e(x)$ が加わり，$b(x) = c(x) + e(x)$ が受信されたとする．この時，$c(x) = a(x)g(x)$ と書けたことに注意すると以下のようになる．

$$b(x) = a(x)g(x) + e(x) \tag{5.12}$$

$g(x)$ の根 α を用いることにより，

$$b(\alpha) \equiv e(\alpha) \equiv s(\alpha) \tag{5.13}$$

なる（$n-k-1$）次以下の多項式 $s(\alpha)$ を $b(\alpha)$ のシンドロームとよぶ．したがって，$s(\alpha)$ と $b(\alpha)$ の対応が分かれば，$s(\alpha)$ から誤り多項式 $e(\alpha)$ を見つけることが出来る．また，v 個の誤りが位置 $j_1, j_2, \cdots, j_v$ に生じている時，シンドロームの第 i 番目の要素 s_i（$i = 1, 2, \ldots, 2t$）は $s_i = r(x^i)$ と定義する．

BCH 符号の復号アルゴリズムでは，符号語の第 j（$j = 0 \sim n-1$) 要素に誤りが生じ

た時に α^{-j} が根となるような誤り位置多項式を導入し，更に，受信後から計算されるシンドロームを用いて誤り位置多項式を求めることによって誤りの生じた位置を推定する．

定義 29 (誤り位置多項式)**.** $v(v \leq t)$ 個の誤りが $j_1, j_2, \cdots, j_v$ の位置に生じているものとする．この時，$GF(2^m)$ 上の v 多項式

$$\sigma(z) = (1-\alpha^{j_1}z)(1-\alpha^{j_2}z)\cdots(1-\alpha^{j_v}z) \tag{5.14}$$

$$= 1+\sigma_1 z+\cdots\sigma_v z^v \tag{5.15}$$

を誤り位置多項式と呼ぶ.

この時，シンドローム $s_i(i=1,2,\cdots,2t)$ と誤り位置多項式の係数 $\sigma_k(k=0,1,\cdots,v)$ の関係は以下のようになる．

$$\begin{bmatrix} s_1 & s_2 & \cdots & s_v \\ s_2 & s_3 & \cdots & s_{v+1} \\ \vdots & \vdots & \ddots & a\vdots \\ s_v & s_{v+1} & \cdots & s_{2v-1} \end{bmatrix} \begin{bmatrix} \sigma_v \\ \sigma_v - 1 \\ \cdots \\ \sigma_1 \end{bmatrix} = \begin{bmatrix} s_{v+1} \\ s_{v+2} \\ \cdots \\ s_{2v} \end{bmatrix} \tag{5.16}$$

5.5 BCH 符号の復号アルゴリズム

本節では，BCH 符号の復号アルゴリズムを紹介する．
$v \leq t$ ならば，受信後からシンドローム $s_i(i=1,2,\cdots,2v)$ を計算でき，式 (5.16) の連立方程式を解くことによって誤り位置多項式の係数 $\sigma_k(k=1,2\cdots,v)$ をすべて求めることができる．そのアルゴリズムを以下で示す.

BCH符号の復号アルゴリズム

1. 受信語からシンドローム $s_1, s_2, \cdots, s_{2t}$ を計算する．
2. シンドロームが全て0でならば，誤りなしと判定して終了する．
3. シンドロームに非零のものがあれば，シンドロームから方程式（5.16）を解く．
4. 誤り位置多項式から誤り位置を決定する．
5. 受信語における誤り位置をビット反転して復号結果とする．

ただし，本研究では，逆シャノン定理を検証する事が目的なので，通常の誤り訂正符号の使い方とは異なる目的でBCH符号を使用している．具体的に述べると，ランダムな系列に対して任意のハミング距離にある符号語を求めている．そのため，一般の復号方法とは異なる復号方法を用いて復号している．

第6章　BCH符号による通信路の再現

逆シャノン定理を検証するために，BCH 符号を用いて通信路を再現する実験を行った．本章では，本研究で行った実験の手順とアルゴリズムについて述べる．

6.1　再現する通信路

本研究では，任意の x^n を符号・復号化することで，通信路を模倣する．

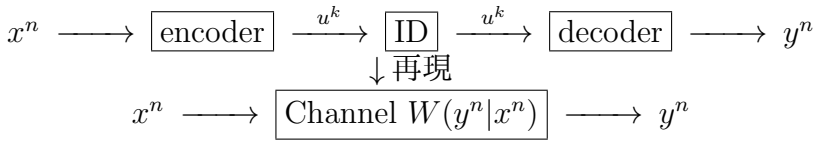

図 6.1: RST における通信路の再現

逆シャノン定理 [1] では，符号化レート $R = k/n$ が模倣したい通信路 W の通信路容量 $C(W)$ よりも大きい $(R > C(W))$ ならば，典型的な $x^n \in A^X_{n,\varepsilon}$ に対して，x^n が入力された時の任意の通信路 $W(y^n|x^n)$ の出力 y^n と同じタイプ $A^{X,Y}_{n,\varepsilon}$ を持つ c^n を出力できる符号・復号器の組 (ϕ_n, ψ_n) が存在する事を示している．本研究では，図 6.2 の通信路の再現を行った．

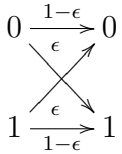

図 6.2: ビット誤り確率 ϵ の二元対称通信路

図 6.2 は文字が入力された時にビット誤りが起きる確率が ϵ，すなわち $W(0|1)W(1|0) =$

ϵ，$W(0|0) = W(1|1) = 1 - \epsilon$ の通信路である．これをビット誤り確率 ϵ の二元対称通信路，または BSC (Binary Symmetric Channnel) という．また，2 元対称通信路の通信路容量 C は $1 - H(\epsilon)$ で求められる．

2 元対称通信路は $W(0|1) = W(1|0) = \epsilon$，$W(0|0) = W(1|1) = 1-\epsilon$ のため，$\frac{1}{n}N(0,0|x^n,c^n) = \frac{1}{n}N(1,1|x^n,c^n) \approx 1-\epsilon$ かつ $\frac{1}{n}N(0,1|x^n,c^n) = \frac{1}{n}N(1,0|x^n,c^n) \approx \epsilon$ となる c^n を出力するような符号・複合器 (ϕ_n, ψ_n) を作れれば，通信路を再現できたと言える．ここで実験のために，ハミング距離 (Hamming distance) を定義する．

定義 30 (ハミング距離)**.** 任意の $x, y \in \{0,1\}$ に対してハミング距離 $d_H(x,y)$ は

$$d_H(x,y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{else} \end{cases}$$

また，任意の長さ n の系列 $x^n, y^n \in \{0,1\}^n$ に対して以下が成り立つ．

$$d_H(x^n, y^n) = \sum_{i=1}^{n} d_H(x_i, y_i)$$

特に零ベクトルと x^n のハミング距離 $d_H(0, x^n)$ を x^n のハミング重み $w_H(x^n)$ と言う．

2 元対称通信路は，ビット誤りを起こす確率が ϵ のため，$d_H(x^n, y^n) \approx n \cdot \epsilon$ である．そのため，$d_H(x^n, c^n) \approx n \cdot \epsilon$ となる c^n を出力できる符号・復号器 (ϕ_n, ψ_n) があれば再現できる．

6.2　提案アルゴリズムについて

本節では，(n,k,d)-BCH 符号を用いて，逆シャノン定理を検証するアルゴリズムについて述べる．このアルゴリズムは，誤り確率 ϵ の 2 元対称通信路を再現するために，(n,k,d)-BCH 符号のレート R がどの程度必要なのかを調べている．

RST 検証アルゴリズム

1. n, ϵ を決める.
2. k, d を決める.($C_n = \{c_1^n, c_2^n, \cdots, c_{2^k}^n\}$ が定まる.)
3. $x^n \in \{0,1\}^n$ をランダムに作成する.
4. $d_H(x^n, c_i^n)$ が $n \cdot \epsilon$ に一番近い符号語 c_i^n を選ぶ.($i = 1 \sim 2^k$)
5. $d_H(x^n, c_i^n) \approx n \cdot \epsilon$ なら成功とする.
6. 3. ~6. を N 回繰り返す.
7. $R = \frac{k}{n}$ での成功率(成功数/N) を計算する.
8. k を変えて 2. ~ 7. を実行する.

RST 検証アルゴリズムでは,まず n, ϵ, k, d を決定する.これにより,(n, k, d)-BCH 符号の符号語 $C_n = \{c_1^n, c_2^n, \cdots, c_{2^k}^n\}$ が定まる.次に,入力 $x^n = \{0,1\}^n$ をランダムに作成し,x^n とのハミング距離が $n \cdot \epsilon$ に最も近い符号語 c_i^n を一つ選ぶ.再現したい通信路の入力 x^n と出力 y^n 間のハミング距離 $d_H(x^n, y^n)$ は $n \cdot \epsilon$ 程度だと考えられるので,$d_H(x^n, c_i^n) \approx n \cdot \epsilon$ ならば,再現に成功したと見なす.そして,x^n をランダムに作り,ハミング距離が $n \cdot \epsilon$ に近い符号語を探索する事を N 回繰り返し,成功率を計算する.上記の計算を k を変える事でレート R と成功率の関係を調べた.($R = k/n$ のため,n を固定した場合,k を変えることと R を変えることは同義である.)

ここで (n, k, d)-BCH 符号は誤り訂正符号なので,x^n をそのまま復号すると,任意のハミング距離にある符号語を見つけられないという問題がある.しかし,全探索を行うと計算量が $O(2^k)$ となり,プログラムの効率が悪くなってしまう.そのため,本研究では BCH 符号の性質を利用し,目的の符号語を効率良く探索するプログラムを加えた.そのアルゴリズムを以下で述べる.

6.2.1 符号語の探索方法

まず,準備として (n, k, d)-BCH 符号の中で,ハミング重み $w_H(c^n)$ が小さい符号語を以下のアルゴリズムで探索する.

最小重みの符号語探索アルゴリズム

1. $u^k = 100\cdots000$ を符号化する.（$C_n(u^k)$ を求める.）
2. ハミング重み $w_H(C_n(u^k))$ を計算する.
3. $w_H(C_n(u^k)) = d$ なら終了する.
4. u^k を 1 ビット右へシフトする.
5. 2. $\sim$ 4 を繰り返し，最小重みの符号語 C_{min} を見つける.

最小重みの符号語探索アルゴリズムは，$u^k = 10\cdots00$ を 1 ビット右へシフトさせ，$u^k = 00\cdots01$ までの全ての符号語についてハミング重みを調べている．また，(n, k, d)-BCH 符号の場合，$00\cdots00$ は符号語なので，符号語の最小重みは符号語間最小距離 d より大きい．そのため，ハミング重みが d の符号語が見つかれば，アルゴリズムを終了させる．

(n, k, d)-BCH 符号は線型符号の 1 種のため，C_n が BCH 符号であるとき，

$$\forall \mathbf{x}, \mathbf{y} \in C_n, \mathbf{x} \oplus \mathbf{y} \in C_n \tag{6.1}$$

が成り立つ.

x^n とのハミング距離 $n \cdot \epsilon$ の符号語を探索をするアルゴリズムは，式 (6.1) と定義 24 を利用している．具体的に述べると，始点となる符号語を一つ求め，最小符号語アルゴリズムで求めた C_{min} を巡回させていき，符号語に加えることで目的の符号語に近づけている.

まず，始点となる符号語を決める．この時，適当な符号語を選ぶよりも，可能な限り x^n にハミング距離が近い符号語を選ぶ方が効率が良いと考えられる．(n, k, d)-BCH 符号において，ある符号語からのハミング距離が t 以下の系列 x^n に対しては誤りの訂正は容易である．これは， x^n から一番ハミング距離が近い符号語を見つけることは容易である事を示している．しかし，t を越える系列に対してはその限りではない．実際，そのような系列に対して，最も近い符号語を求める効率の良い方法は見出されていない．そこで，本研究では，入力された x^n に対し誤り訂正を行う．訂正できれば，訂正した符号語を用いて探索を行う．しかし，訂正出来ない場合，$x^n = x_1x_2\cdots x_n$ の $n-k$ ビット目から n ビット目（$x_{n-k}x_{n-k+1}\cdots x_n$）を符号化し，その符号語 c_n を始点とする．このようにす

る事で，$n-k$ ビット目から n ビット目までは x^n と c^n は同じ文字であるため，x^n とのハミング距離が $n-k$ 以下の符号語を始点にする事が出来る．

以上で求めた符号語 c^n と C_{min} を用いて，x^n とのハミング距離が $n\cdot\epsilon$ となる符号語を探索する．そのためのアルゴリズムを以下で示す．

距離 $n\cdot\epsilon$ の符号語探索アルゴリズム

1．$C'=c^n\oplus C_{min}$ を求める．

2．x^n とのハミング距離 $d_H(x^n,C')$ を計算する．

3．$d_H(x^n,C')$ の値が $d_H(x^n,c^n)$ よりも $n\cdot\epsilon$ に近ければ，$c^n:=C'$ とする．($d_H(x^n,C')=n\cdot\epsilon$ なら終了する．)

4．C_{min} を 1 ビット右へシフトする．

5．1. $\sim$ 4. を繰り返す．(C_{min} が最初の符号語に戻ったら終了)

上記のアルゴリズムについて述べる．まず，線形符号の性質（式（6.1））を利用し，最初に求めた c^n と C_{min} を排他的論理和で足しあわせることで，符号語 C' を作成する．x^n とのそれぞれのハミング距離 $d_H(x^n,c^n)$，$d_H(x^n,C')$ を比べ，C' の方が $n\cdot\epsilon$ に近ければ始点を $c^n:=C'$ で更新する．もし近くなければ，c^n はそのままで，C_{min} を 1 ビット右にシフトさせる．例として，$C_{min}=10001$ ならば，1 ビット右へシフトした後の符号語は $C'_{min}=11000$ となる．1 ビット右へシフトした符号語 C'_{min} と c_n を排他的論理和で足し，ハミング距離を比べる．$n\cdot\epsilon$ に近づかなければ C'_{min} を 1 ビット右にシフトさせる．以上の処理を C'_{min} が C_{min} に戻るまで繰り返す事で，始点とした符号語を目的の符号語に近づけている．

6.2.2 提案アルゴリズム

RST 検証アルゴリズムに最小重み符号語と距離 $n\cdot\epsilon$ の符号語を探索するアルゴリズムを加えた提案アルゴリズムを以下で示す．

提案アルゴリズム

1. n, ϵ を決める.
2. k, d を決める.（$C_n = \{c_1^n, c_2^n, \cdots, c_{2^k}^n\}$ が定まる.）
3. 最小重みの符号語探索アルゴリズムを使用する.（C_{min} を探索する.）
4. $x^n \in \{0,1\}^n$ をランダムに作成する.
5. 距離 $n \cdot \epsilon$ の符号語探索アルゴリズムを使用し，目的の符号語 c_i^n を探索する.
6. $d_H(x^n, c_i^n) \approx n \cdot \epsilon$ なら成功とし，成功数をカウントする.
7. 4. ∼6. を N 回繰り返す.
8. $R = k/n$ での成功率（成功数/N) を計算する.
9. k を変えて 1. ∼8. を繰り返す.

提案アルゴリズムで求めた，レート R と成功率の関係の結果を次節で述べる.

第7章　結果

逆シャノン定理の検証のため，BCH 符号の符号化レートと目的の通信路の再現成功確率との関係を第 6 章で提案したアルゴリズムを用いて調べた．本章では，その結果を述べ，それに基づく考察を行う．

7.1　シミュレーションパラメータ

シミュレーション結果で用いたパラメータを表 7.1 にまとめた．本プログラムを作成する際に [7] を参考にした．実測機のスペック，プログラムの言語は以下のとおりである．

プロセッサ　Intel Core i5

実装メモリ　8.00GB

OS　Windows 7 64bit

プログラム　C 言語

表 7.1: シミュレーションパラメータ

符号語長	情報ビット数	計画距離
n	k	d
レート	通信路容量	通信路の誤り確率
R	C	ϵ

7.2　シミュレーション結果

本研究では，第 6 章で述べたアルゴリズムを用いて実験を行った．準備として (n, k, d)-BCH 符号 $C_n = \{c_1^n, c_2^n \cdots c_{2^k}^n\}$ と 2 元対称通信路の誤り確率 ϵ を固定しておく．まず，確

率 $p(1) = 1/2, p(0) = 1/2$ に従ってランダムに $x^n \in \{0,1\}^n$ を作成した．次に x^n とのハミング距離が $n \cdot \epsilon$ に最も近い符号語 $c_i^n (i = 1, 2, \cdots, 2^k)$ を探索した．そして，ハミング距離 $d_H(x^n, c_i^n)$ が

$$0.99 \cdot n \cdot \epsilon \leq d_H(x^n, c_i^n) \leq 1.01 \cdot n \cdot \epsilon$$

の範囲に収まれば成功，収まらなければ失敗とした．最後に x^n を 10000 回ランダムに作成し，同様の事を繰り返し，成功確率（成功した回数/ 10000）を求めた．上記のことを n, ϵ を固定したまま，k を変えていき，レート $R = k/n$ ごとの成功確率を調べた．その結果を図 7.1 ～ 7.4 に示す．ここで，図の x 軸はレート R，y 軸は成功確率 （success probability），simulation とはシミュレーション結果，C は誤り確率 ϵ の 2 元対称通信路の通信路容量を表している．

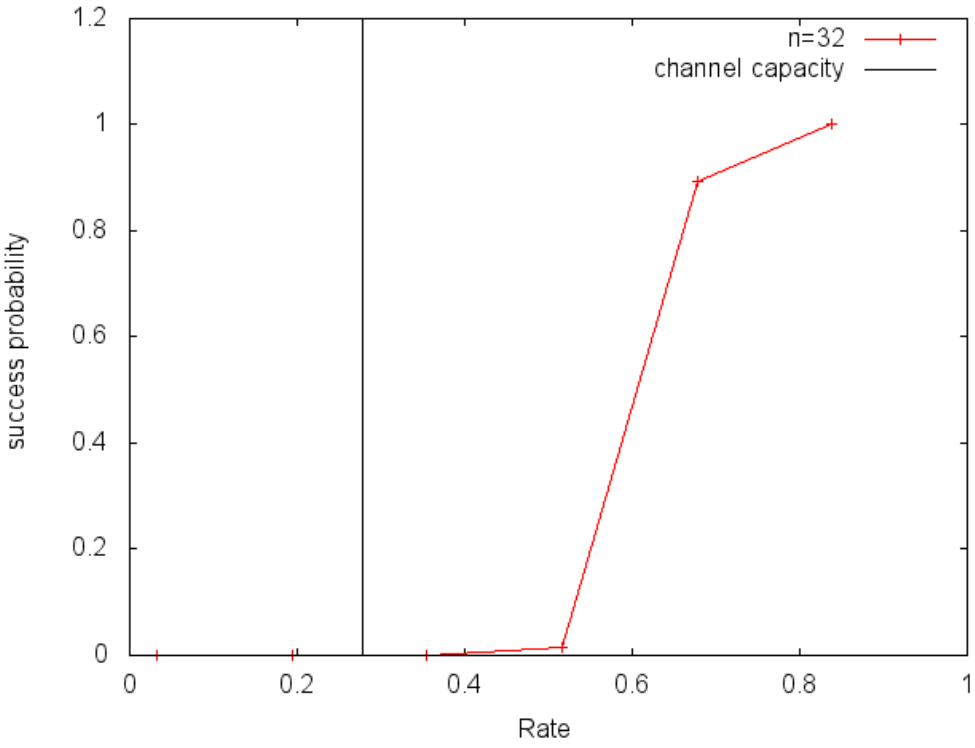

図 7.1: $n = 32, \epsilon = 0.2$ の場合のレートと成功確率の関係

図 7.1 ～ 7.4 は，ϵ を固定した場合に n によってレートと成功確率の関係がどのように変化するのかを確認したものである．なお，図 7.1 と図 7.2 のプロット数が少ない理由は，BCH 符号の場合，符号長 n が小さいと存在する符号語の数が少ないためである．(図 7.1

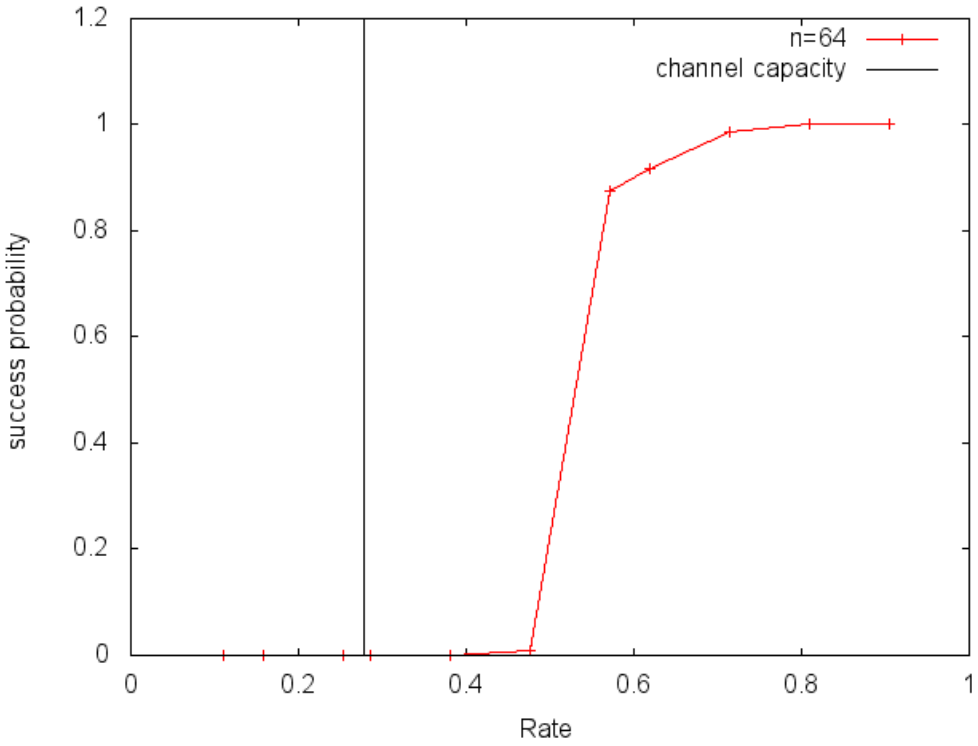

図 7.2: $n = 64, \epsilon = 0.2$ の場合のレートと成功確率の関係

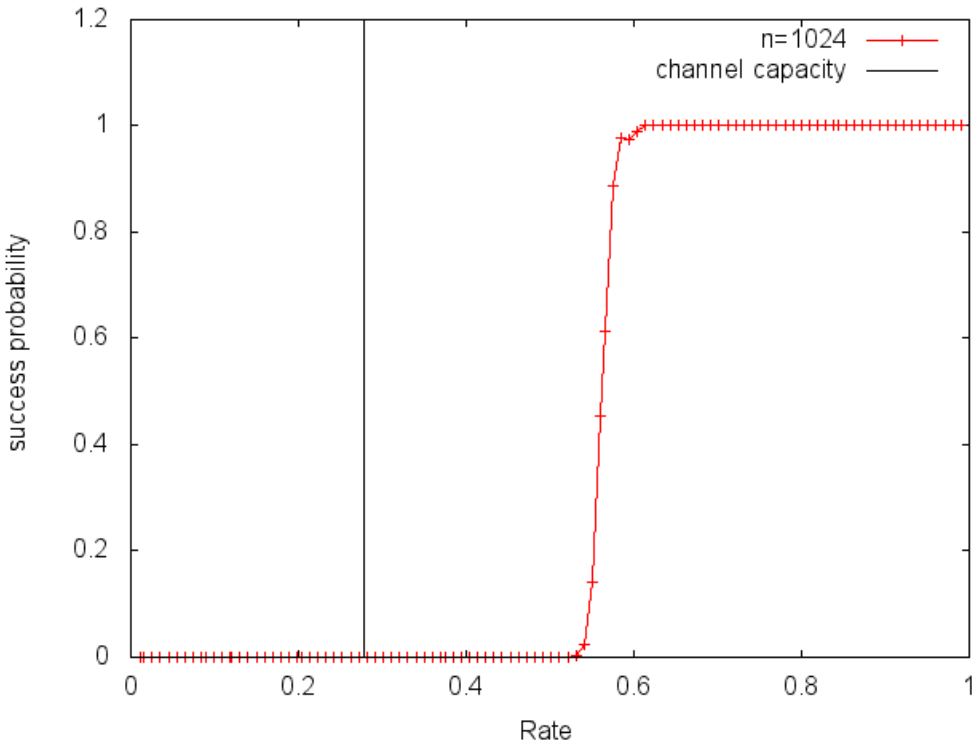

図 7.3: $n = 1024, \epsilon = 0.2$ の場合のレートと成功確率の関係

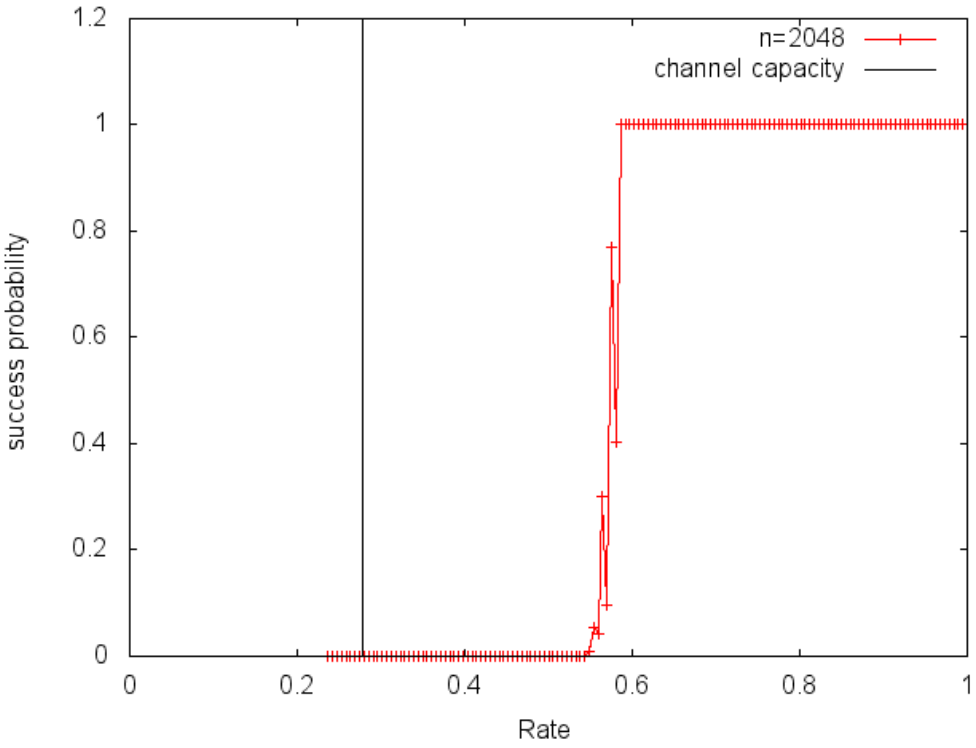

図 7.4: $n = 2048, \epsilon = 0.2$ の場合のレートと成功確率の関係

のプロット数は 6 点，図 7.2 のプロット数は 11 点）また，成功確率が 0 であるレートの最大値を R_{0max}，成功確率が 1 であるレートの最小値を R_{1min} とする．この時，各 n と R_{0max}, R_{1min} の関係を表 7.2 で示す．

表 7.2: n 毎の R_{1min} と R_{0max} の差

n	R_{0max}	R_{1min}	$R_{1min} - R_{0max}$	R_{0max} と R_{1min} の中間点
32	0.194	0.839	0.645	0.517
64	0.286	0.810	0.524	0.548
1024	0.501	0.614	0.113	0.558
2048	0.538	0.586	0.048	0.562

本来，符号語長は十分大きい方が望ましいが，プログラムで探索を行なっているため，大きいと時間がかかってしまうという問題がある．表 7.2 を見ると，符号語長 1024 で $R_{1min} - R_{0max}$ の差が十分小さくなり，R_{0max} と R_{1min} の中間点も符号長 2048 の場合と差異は少ない．そこで，符号語長を 1024 に固定し，通信路の誤り確率 ϵ を変化させ，同様の実験を行った．その結果を図 7.5 ～ 7.11 で示す．

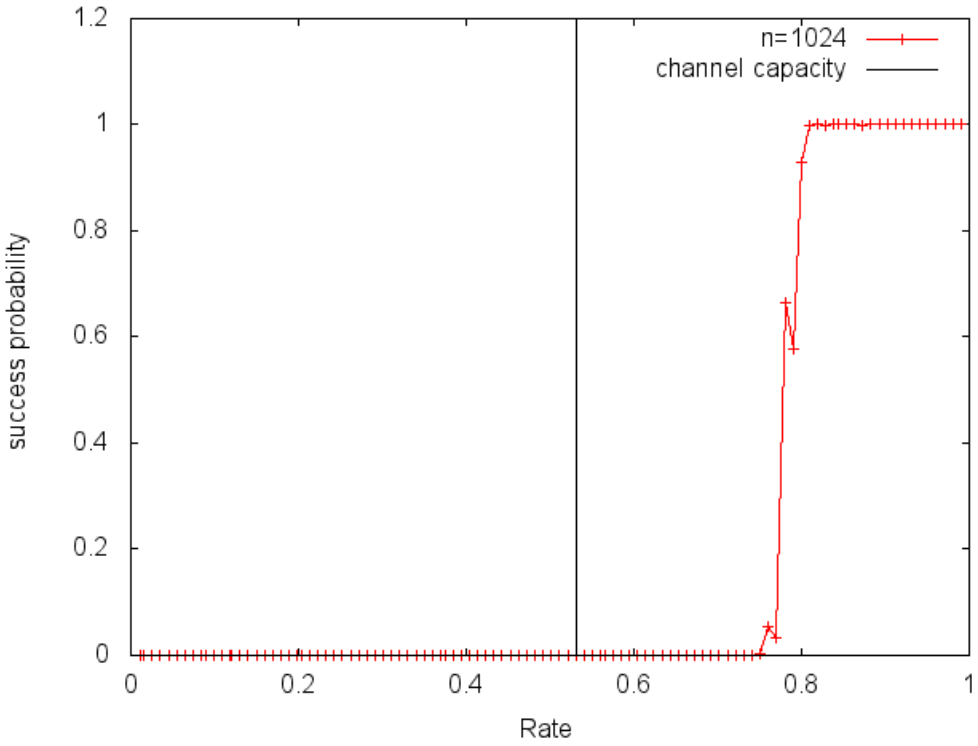

図 7.5: $n = 1024, \epsilon = 0.1$ の場合のレートと成功確率の関係

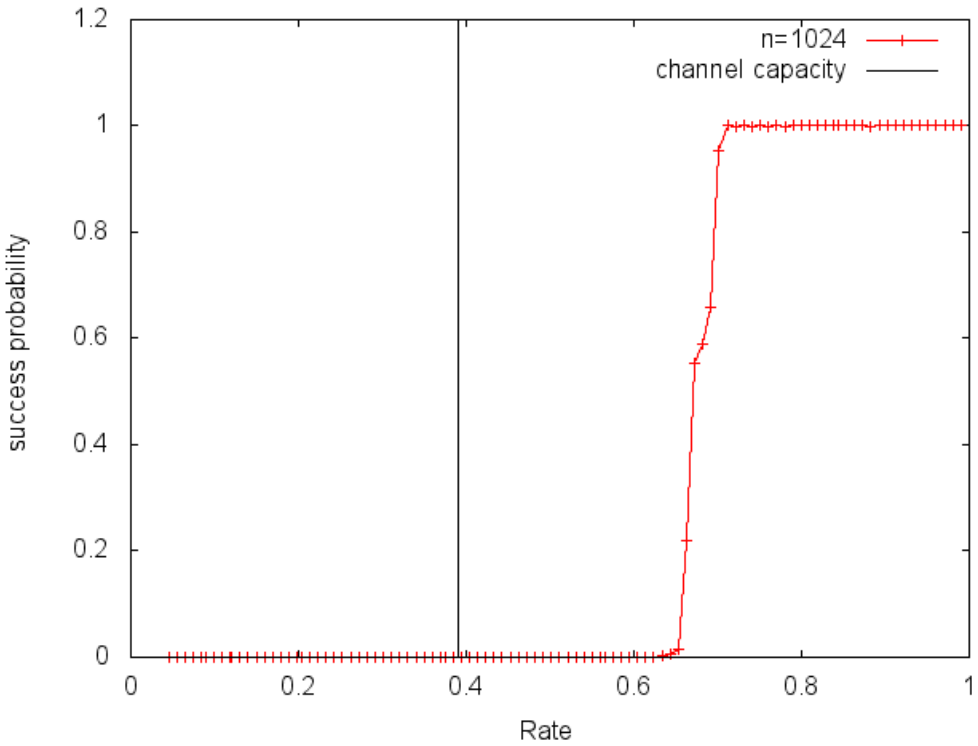

図 7.6: $n = 1024, \epsilon = 0.15$ の場合のレートと成功確率の関係

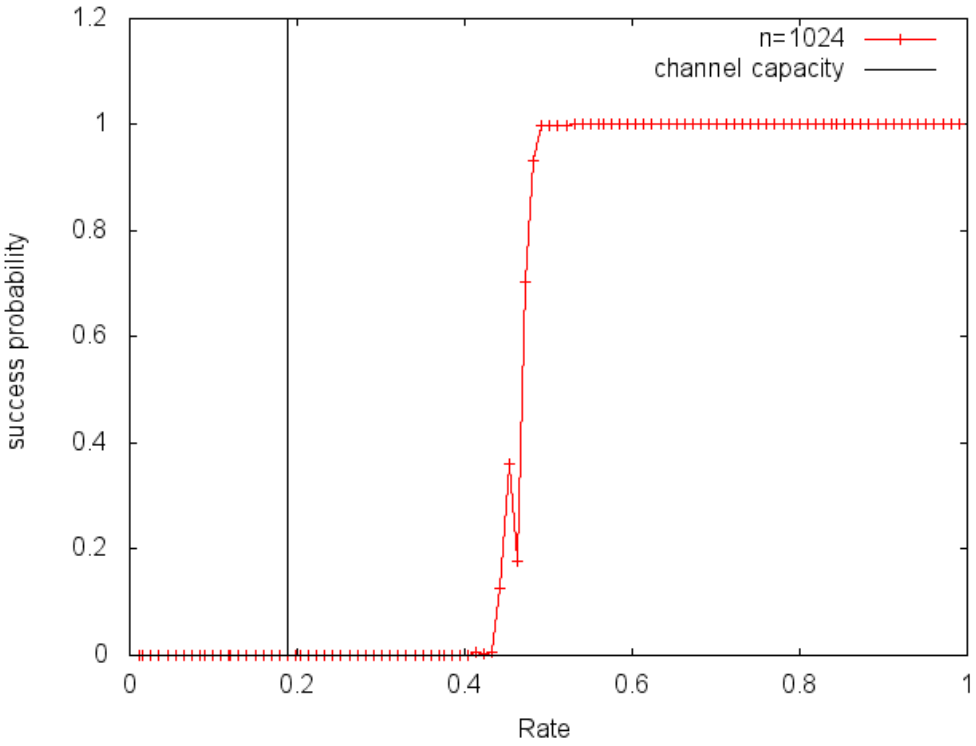

図 7.7: $n = 1024, \epsilon = 0.25$ の場合のレートと成功確率の関係

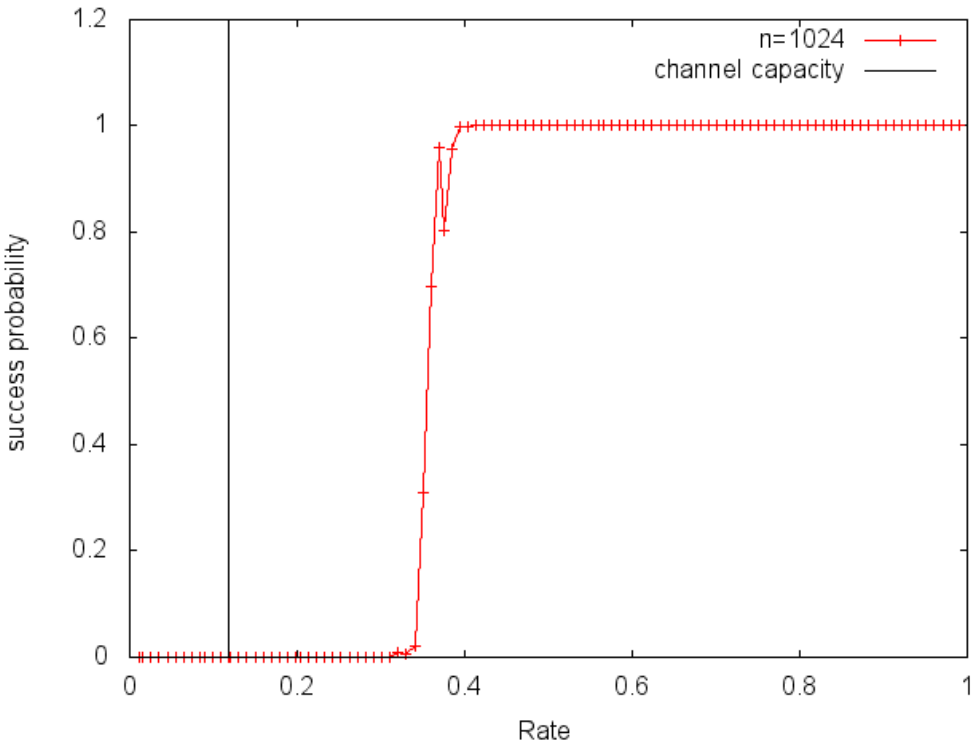

図 7.8: $n = 1024, \epsilon = 0.3$ の場合のレートと成功確率の関係

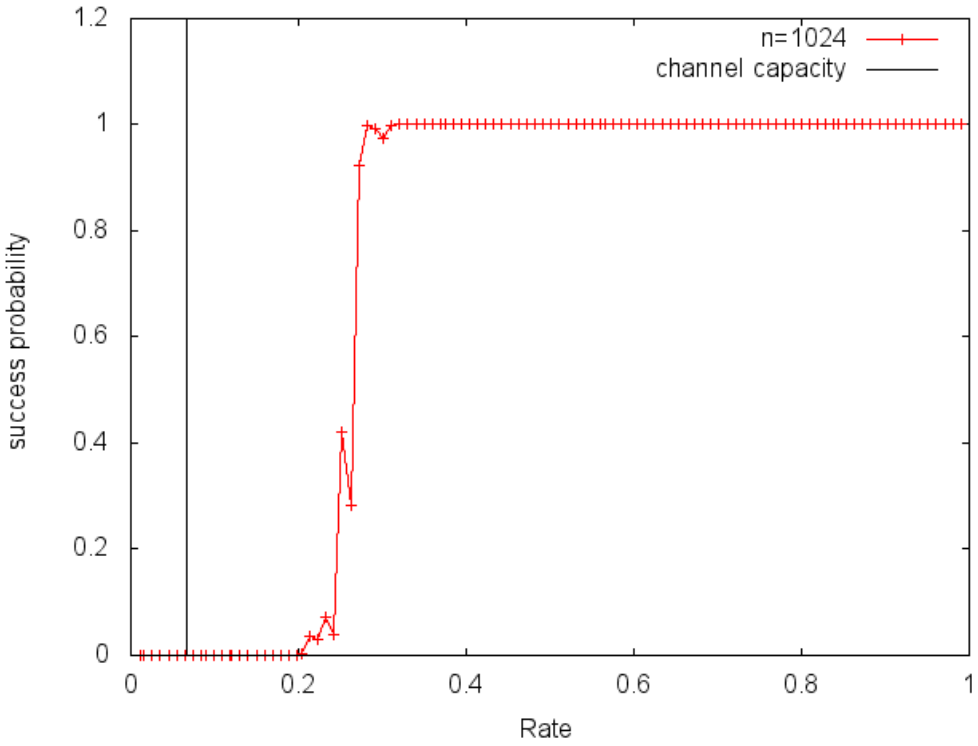

図 7.9: $n = 1024, \epsilon = 0.35$ の場合のレートと成功確率の関係

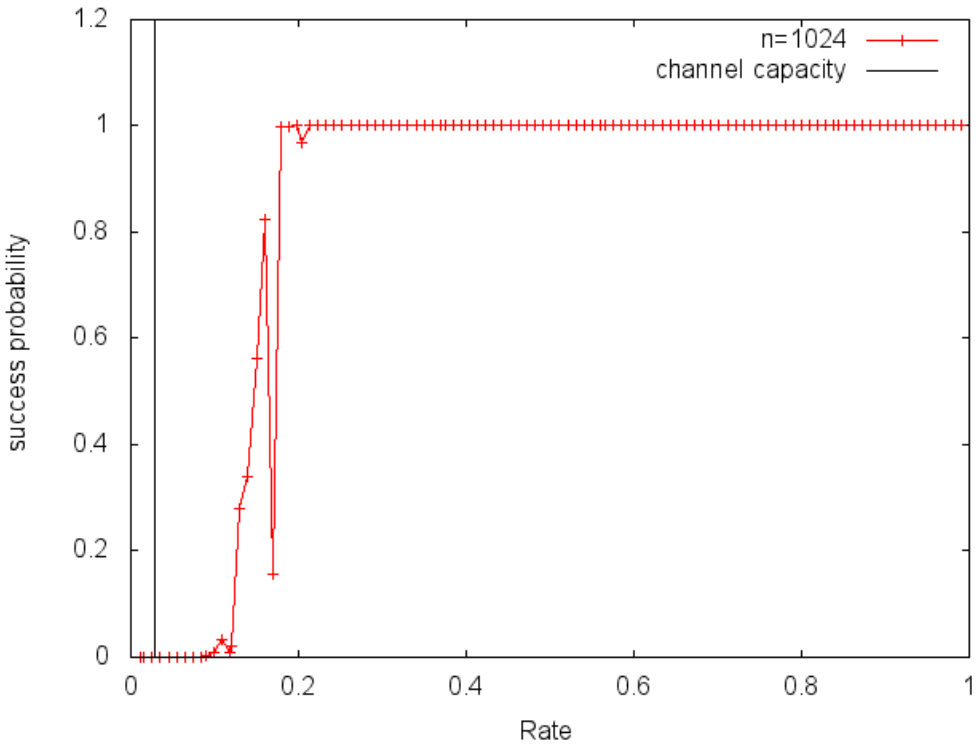

図 7.10: $n = 1024, \epsilon = 0.4$ の場合のレートと成功確率の関係

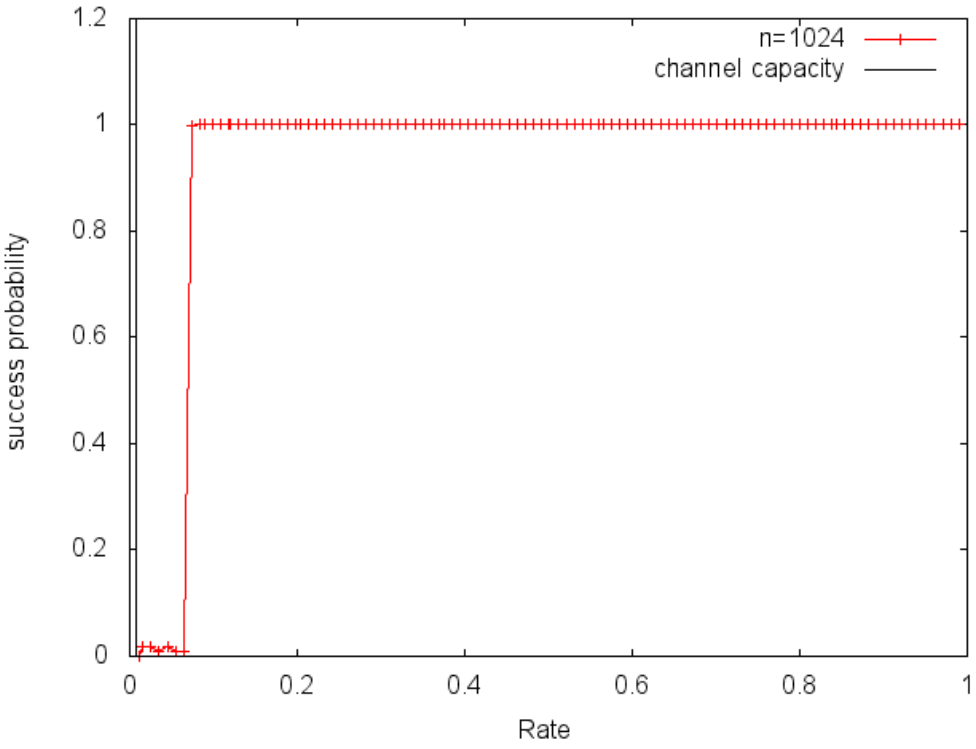

図 7.11: $n = 1024, \epsilon = 0.45$ の場合のレートと成功確率の関係

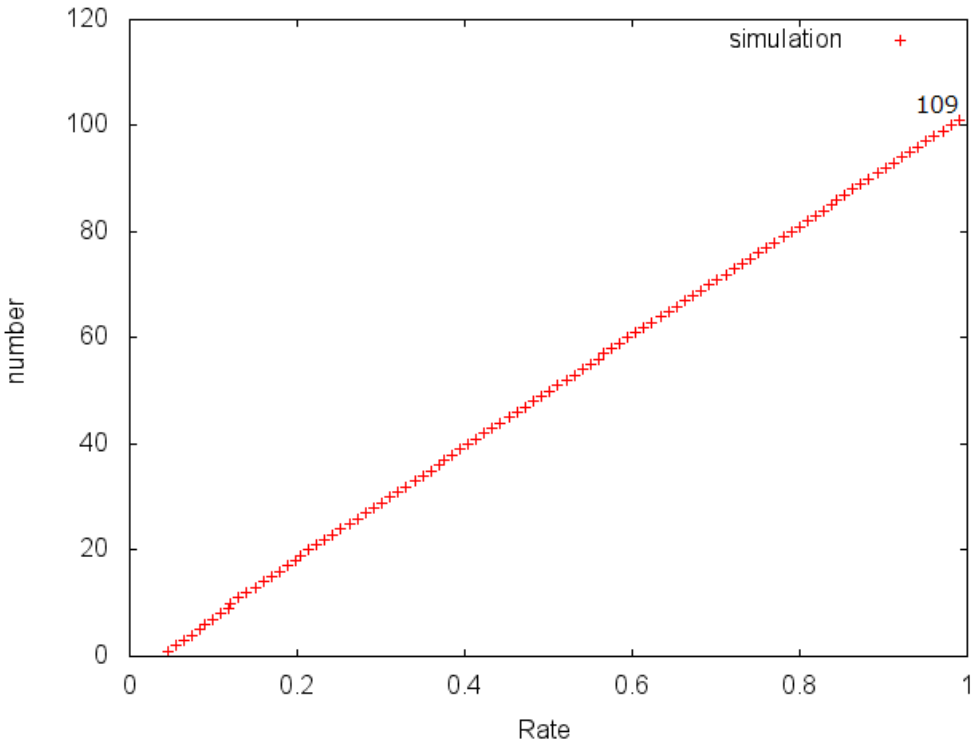

図 7.12: $n = 1024$ の場合のレートとプロット数の関係

図 7.12 では，$n = 1024$ の場合のレートとプロットした点の個数を示している．縦軸はそのレートまでのプロット数，横軸はレートである．すなわちこの図は，レートの値が 0.010753 から 0.990225 までをプロットしており，合計プロット数は 109 個であり，プロットの間隔があまり偏っていない事を示している．

最後に，図 7.1 ～ 7.11 の結果から $n = 1024$ を固定し，ϵ を変えた場合の R_{0max}, R_{1min} と誤り確率 ϵ の二元対称通信路の通信路容量 $C(Q)$ の関係を表 7.3 にまとめた．

表 7.3: ϵ 毎の R_{1min},R_{0max},C の関係

ϵ	C	R_{0max}	R_{1min}	R_{0max} と R_{1min} の中間点
0.1	0.531	0.731	0.819	0.775
0.15	0.390	0.624	0.712	0.668
0.2	0.278	0.501	0.614	0.558
0.25	0.189	0.394	0.531	0.463
0.3	0.119	0.272	0.413	0.343
0.35	0.066	0.169	0.321	0.245
0.4	0.029	0.054	0.198	0.126
0.45	0.007	0.011	0.084	0.048

7.3 考察

図 7.1 ～ 7.4 より，n が増大するとともに，R_{0max} と R_{1min} の差 $R_{1min} - R_{0max}$ が短くなっているのがわかる．(表 7.2) そのため，n を更に増やしていくと $R_{1min} - R_{0max}$ が 0 に近づいていくと考えられる．

図 7.1 ～ 7.11 より，以下の 3 点がわかる．

まず，BCH 符号のクラスでも，通信路容量 C 以上の符号化レートをとることで，任意の誤り確率 ϵ を持つ 2 元対称通信路を再現できることがわかる．

次に，それぞれの図において R_{0max} から R_{1max} の間のレートの成功確率が不安定であることがわかる．この理由は 2 つ考えられる．1 つ目は x^n の発生回数が 10000 回と少ないため，ばらつきができたためと考えられる．これに関しては発生回数を増やすと改善されると考えられるが，提案アルゴリズムを用いると時間がかかってしまうため，提案

アルゴリズムの改良が必要である．もう 1 つは，本研究では (n, k, d)-BCH 符号を全探索をしていないので，任意のハミング距離に符号語が存在するが，その符号語を探索できていないためと考えられる．これに関しては，符号語探索アルゴリズムを改良する必要があるが，現在効率の良い探索方法は見つかっていないため，今後も研究が必要である．また，符号語探索アルゴリズムを改良することができれば，各レートにおける通信路再現成功確率も上がると考えられるため， R_{0max}，R_{1min} 及び R_{0max} と R_{1min} の中間点はそれぞれ通信路容量 $C(W)$ に近づくと考えられる．

最後に図 7.1 ∼ 7.11 からわかることは，通信路 W の再現に必要な符号化レート R が，通信路容量 $C(W)$ よりもかなり大きいことである．これは BCH 符号のクラスが，逆シャノン定理における最適な符号語のクラスよりも小さいためと考えられる．

第8章　結論と今後の課題

8.1　結論

本研究では，符号化レートをコントロールしやすい (n, k, d)-BCH 符号を用いて，逆シャノン定理の検証を行った．すなわち，符号化レート R と任意に与えられた通信路 W の再現成功確率の関係を数値実験により確認した．数値実験の結果は，第 7 章の図 7.1 ～ 7.11 に示されている．ここから，以下の 2 点がわかる．

まず，通信路容量 $C(W)$ よりも大きい符号化レートをとることで，再現成功確率を 1 にすることが可能であることを示せた．これは BCH 符号のクラスでも逆シャノン定理に基づいた通信路の再現が可能である事を意味している．

次に，BCH 符号の場合，通信路 W を再現するためには，通信路容量 $C(W)$ よりもかなり大きい符号化レート R が必要であることがわかった．これについては 2 つの理由が考えられる．1 つ目の理由として，BCH 符号のクラスが，逆シャノン定理における最適な符号のクラスよりも小さいためであると考えられる．2 つ目の理由として，符号語を全探索していないため，正確に距離 $n \cdot \epsilon$ の符号語を発見できていないためと予想される．それゆえ，符号化探索アルゴリズムを改良できれば，通信路の再現に必要な符号化レートが通信路容量に近づくと考えられる．

8.2　今後の課題

本研究の課題は大きく分けて 2 つある．

第 7 章の図 7.1 ～ 7.11 には，通信路の再現成功確率が 0 から 1 に変わる符号化レートのしきい値が見られる．今後の課題の 1 つは，この符号化レートのしきい値の精度と確

度の向上である．そのため，第 6 章の最後に提示した提案アルゴリズムを改善する必要がある．精度について述べると，本研究では符号語長 1024 の場合で実験を行ったが，符号語長やサンプル数を増やすことで，大数の法則によりしきい値の精度が上昇すると考えられる．しかし，提案アルゴリズムでは時間がかかってしまい，実現困難である．そのため，アルゴリズムの高速化が必要である．一方，確度について述べると，符号語を全探索すれば，距離 $n \cdot \epsilon$ にある符号語を発見できるため，確度が上昇すると考えられる．しかし，全探索は探索時間が $O(2^k)$ となるため，現実的ではない．とはいえ，提案アルゴリズムでは最適な符号語を発見できていない可能性がある．そこで，全探索をせずに最適な符号を発見する探索アルゴリズムの開発が求められる．

もう 1 つの課題は，BCH 符号以外での適用である．一般的な符号あるいは新たに構成する符号で，符号化レートと通信路再現成功確率の関係を調べ，逆シャノン定理に最適な符号を発見もしくは構成することが望まれる．

通信路は暗号では信頼出来る第三者に変わるリソースと考えられるが，逆シャノン定理により再現される通信路は，暗号で必要とされる意味での通信路とはほど遠い．しかし逆シャノン定理により，遠方の二人が相関を持つ確率分布 $P(x, y)$ に従う擬似乱数をりようしたいときに，通信量を $I(X;Y)$ まで減らすことが可能となる．

参考文献

[1] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem," *IEEE Trans. Inform. Theory*, vol. 48, pp. 2637-2655, 2002.

[2] R. C. Bose and D. K. Ray-Chaudhuri, "On A Class of Error Correcting Binary Group Codes," Information and Control, vol. 3, pp. 68-79, 1960.

[3] T. M. Cover and J. A. Thomas, "Elements of Information Theory second edition," Wiley-Interscience publication, June 2006.

[4] A. Hocquenghem, "Codes correcteurs d'erreurs," Chiffres, vol. 2, pp. 147-156, 1959.

[5] 今井 秀樹, 『符号理論』 電子情報通信学会, 1994 年.

[6] 神保 雅一, 藤原 良, 『符号と暗号の数理』 共立出版株式会社, 1993 年.

[7] M. Z. Robert, "The Error Correcting Codes (ECC) Page," http://www.eccpage.com/ (2013 年 1 月現在)

謝辞

本研究を遂行するに当たり，終始適切な助言を賜り，また丁寧に指導して下さった小川朋宏准教授に深く感謝いたします．

長岡浩司教授には，ゼミの際に的確な助言をしてくださり，Cover ゼミの際にも難しい箇所もわかりやすく教えて頂きました．ありがとうございます．

鈴木淳助教には，Cover ゼミにお付き合いして頂いただけでなく，ゼミの空間が居心地の良い空間になるよう働きかけて頂きました．心より感謝いたします．

博士の久保卓也氏には，研究生活をサポートして頂き，多くの刺激と示唆を得ることが出来ました．感謝の意を表します．そして，ネットワーク基礎学講座の院生，学部生には多大な協力と励ましを頂きました．本当にありがとうございました．