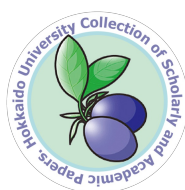




HOKKAIDO UNIVERSITY

Title	2005年度 混沌系工学特論講義ノート
Author(s)	井上, 純一; Inoue, Jun-ichi
Description	とりあげた各トピックスに関する最新情報は著者のホームページを参照ください。 http://www005.upp.so-net.ne.jp/j_inoue/index.html 他
Issue Date	2005-11-30T05:17:27Z
Doc URL	https://hdl.handle.net/2115/828
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/2.1/jp/
Type	learning object
File Information	konton05_slide5.pdf, 第5回講義スライド





混沌系工学特論 #5

情報科学研究科 井上純一

URL : http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/

Mirror : http://www005.upp.so-net.ne.jp/j_inoue/index.html

平成17年11月14日 第5回講義

デジタルデータの転送と復元再考

$$P(\{\sigma\}) = \frac{\exp\left(J \sum_{\langle ij \rangle} \sigma_i \sigma_j\right)}{\sum_{\sigma} \exp\left(J \sum_{\langle ij \rangle} \sigma_i \sigma_j\right)}$$

画像の復元では画像の特質を
事前分布で表現できる

デジタル画像

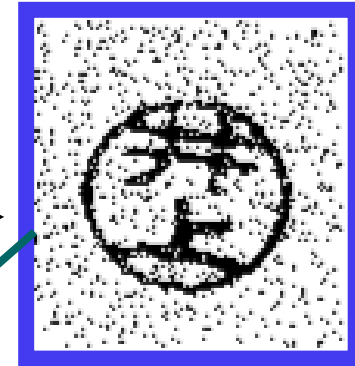


確率モデルで表現

伝送路

符号

劣化されたデジタル画像



$\{\tau\}$

ビット列

011000101001000001001

復元されたデジタル画像



復号

$\{\sigma\}$

ノイズの乗ったビット列

100111001010001001001

復号

011000110010001001001

復元されたビット列

ビット列の場合は？

$\{\xi\}$

雑音のある状況でのデータ送信

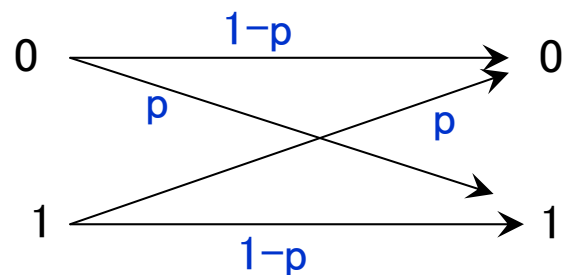
誤り訂正符号

誤り率 p の2元対称通信路

ビット反転率

ただ1度だけでなく、何回か同じ記号を繰り返し送信する

受信者側は**多数決**で送信された記号を確定する



2元対称通信路

5回送信した場合 : 00101 $\Rightarrow$ 0、01111 $\Rightarrow$ 1 等

送信回数 $n = 5$ のときの誤り確率は

$$\begin{aligned} f_e^{(5)}(p) &= {}_5C_3 p^3 (1-p)^2 + {}_5C_4 p^4 (1-p) + p^5 \\ &= 10 p^3 (1-p)^2 + 5 p^4 (1-p) + p^5 \end{aligned}$$

3ビットの誤りが5ビットの何番目にくるかという場合の数

繰り返し回数 n を大きくするとどうなるか？

伝送速度と誤り確率

$$R = \frac{\log M}{n} = \frac{1}{n}$$

伝送速度(レート)

送信記号の種類の数
1, 0 ならば2

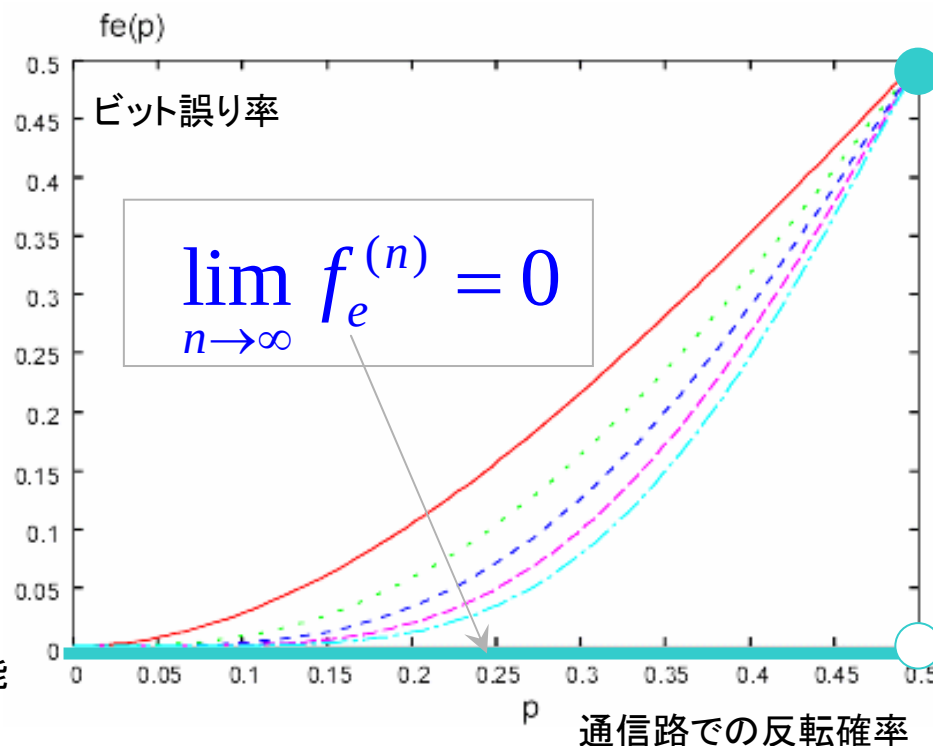
繰り返し回数を限りなく多くとると、
誤り確率は限りなくゼロに近づくが、
同時に伝送速度もゼロになってしまう

⇒ 実用的ではない。しかし

$$R < C$$

通信路容量

ならば、限りなく小さな誤り確率での情報伝送が可能



通信路容量

$$C = \max_{P_X} I(X; Y)$$
$$I(X; Y) = H(Y) - H(Y | X)$$

通信路容量の定義

X : 入力
 Y : 出力

相互情報量 : 入力 X を知ったときに、出力 Y に対して得られる知識の増加分

↓ 入力の確率分布に関して、相互情報量を最大化したものが通信路容量

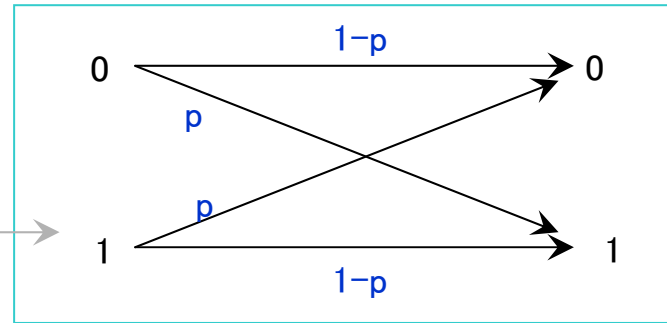
通信路容量は、通信路が伝送できる最大の情報量という意味を持つ

通信路容量の計算例

誤りの無い2元対称通信路

$$\begin{aligned}P_{Y|X}(0|0) &= P_{Y|X}(1|1) = 1-p \\P_{Y|X}(0|1) &= P_{Y|X}(1|0) = p\end{aligned}$$

条件付き確率での表現



グラフ表現

$$H(Y | X) = -p \log p - (1-p) \log(1-p) \quad \text{入力分布によらない}$$

$$C = I(X; Y) = \max_{P_X} \{H(Y) - H(Y | X)\} = 1 - h(p)$$

$$= 1 + p \log p + (1-p) \log(1-p) \quad \text{BSCの通信路容量}$$

同様にガウス通信路に対して

通信路容量は

$$C = \max_p I(X; Y) = \frac{1}{2} \log \left(1 + \frac{a_0^2}{a^2} \right)$$

通信路符号化定理

通信路符号化定理

- i) $R < C$ なる任意の速度 $R = \log M / n$ に対し
任意に小さな復号誤り率 p_E の符号が存在する。
- ii) $R > C$ なる R に対し、任意に小さな p_E の符号は存在しない。

ゼロでなくてもよい。しかし、符号長 n を無限大とすることは必要

情報源の記号数 $M = 2^{nR}$

この定理を「ランダム符号」と呼ばれる符号に対して証明していく

ランダム符号

情報源の記号： $S_1, S_2, \dots, S_M$ （等確率で生成される）

符号化：

$S_1, S_2, \dots, S_M$ の各々にランダムに0,1を n 個並べた系列を割りあてる

$S_1 \leftarrow 0010 \cdots 000$

$S_2 \leftarrow 0100 \cdots 001$

$S_3 \leftarrow 0010 \cdots 100$

.....

$S_M \leftarrow 1000 \cdots 010$

符号長 n

各ビットに $\frac{1}{2}$ の確率で 0,1 を割り振る

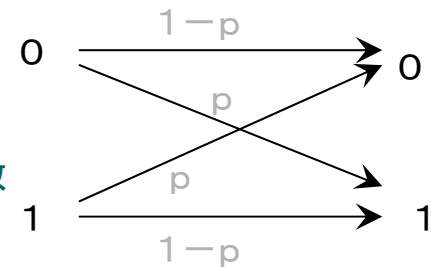
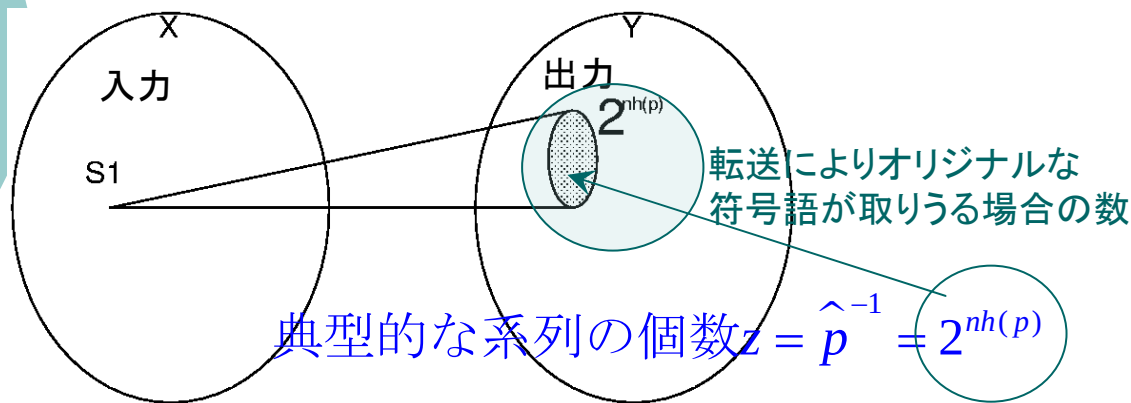
$$M = 2^{nR} < 2^n$$

$R \leq 1$ として議論を進める

これらを2元対称通信路を介して転送する状況下で定理を証明していく

通信路符号化定理の証明#1

一つの符号語が転送により拡大する大きさの評価



符号語は2元対称通信路を介して転送される

ある符号語 $S_1 = 000 \dots 000$ は通信路の雑音により $010 \dots 101$ として受信される
全 n ビットのなかで誤りの生じるビット数はおおよそ np と見積もられる

$$\frac{(0100 \dots 011), (0111 \dots 100), \dots, (0101 \dots 111)}{\begin{array}{ccc} np \text{ ビットの誤り} & np \text{ ビットの誤り} & np \text{ ビットの誤り} \end{array}}$$

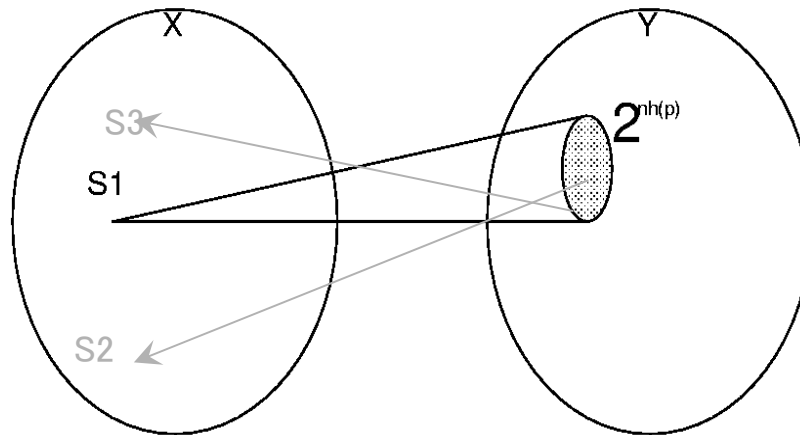
通信路出力の典型的な系列

典型的な系列のなかの一つが現れる確率

$$\hat{p} = p^{np} \cdot (1-p)^{n(1-p)} = 2^{np \log p + n(1-p) \log(1-p)} = 2^{-nh(p)} \quad \text{2値エントロピー関数}$$

通信路符号化定理の証明#2

復号誤り率の評価



復号誤りが生じるのは...

受信される可能性のある $2^{nh(p)}$ 個の各々が、実際に送信された符号語 S_1 以外の $S_2, S_3, \dots, S_M$ の $(M-1)$ 個のいずれかに間違って復号されるとき

$$p_E = \frac{(M-1)}{2^n} \times 2^{nh(p)} \approx \frac{M}{2^n} \cdot 2^{nh(p)} = 2^{n(R-1+h(p))} = 2^{n(R-C)} \rightarrow 0 \quad (C=1-h(p) > R)$$

(M-1)個のいずれかが選ばれる確率 受信される可能性のある系列の個数

i) の証明終わ

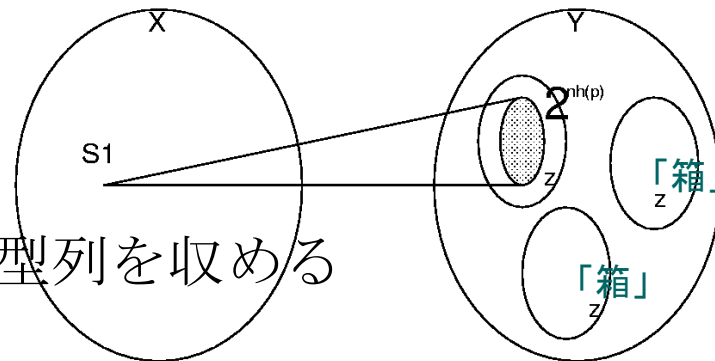
通信路符号化定理の証明#3

ii) の証明

受信系列の全ての可能性 2^n

1つの符号語に対する通信路出力の典型列を収める

「箱」の数 $= z = 2^n / M$



2元対称通信路により実際に得られる典型列の数 $2^{nh(p)}$

は全てこの「箱」に収まらなければならない

$$\frac{2^n}{M} > 2^{nh(p)} \Rightarrow 2^{n(C-R)} > 1$$

これは $R > C$ では満たされない

ii)の証明終わり

スピンモデルを用いた符号/復号化

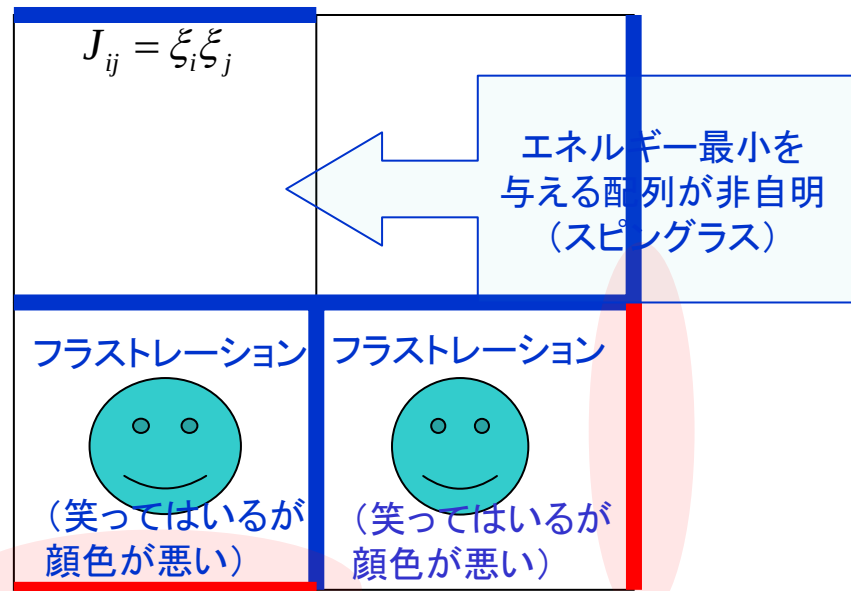
ビット列を送るのではなく、結合 $J_{ij} = \xi_i \xi_j$ を送信

エネルギー関数 $H = -\sum_{ij} \xi_i \xi_j \sigma_i \sigma_j$ エネルギー関数の最小状態を選ぶことで復号する

ノイズが無い場合

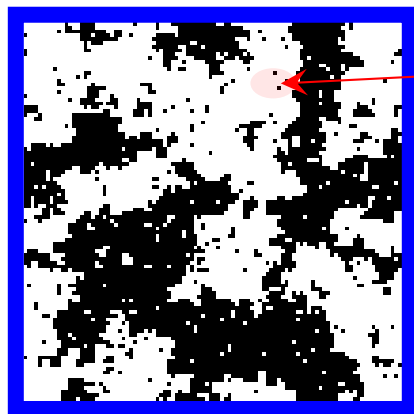


ノイズのある場合



ノイズにより反転したボンド

パリティと画素の同時送信



$$J_{ij} = \xi_i \xi_j$$

画素だけではなく、隣接画素対も同時に送信する(正方格子で $2N$ 個ある)

尤度

$$P(\{J\}, \{\tau\} | \{\sigma\}) = \frac{e^{\beta_J \sum_{ij} J_{ij} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i}}{[2 \cosh(\beta_J)]^{N_B} [2 \cosh(h)]^N}$$

パリティ送信部分

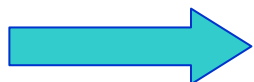
画素送信部分

事後分布とエネルギー関数

$$P(\{\sigma\} | \{\tau\}) = \frac{e^{-H_{\text{eff}}}}{\sum_{\{\sigma\}} e^{-H_{\text{eff}}}}$$

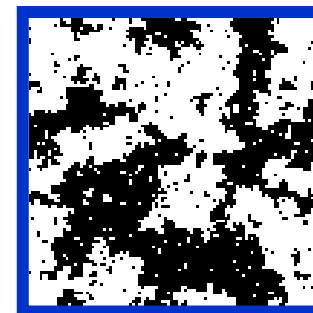
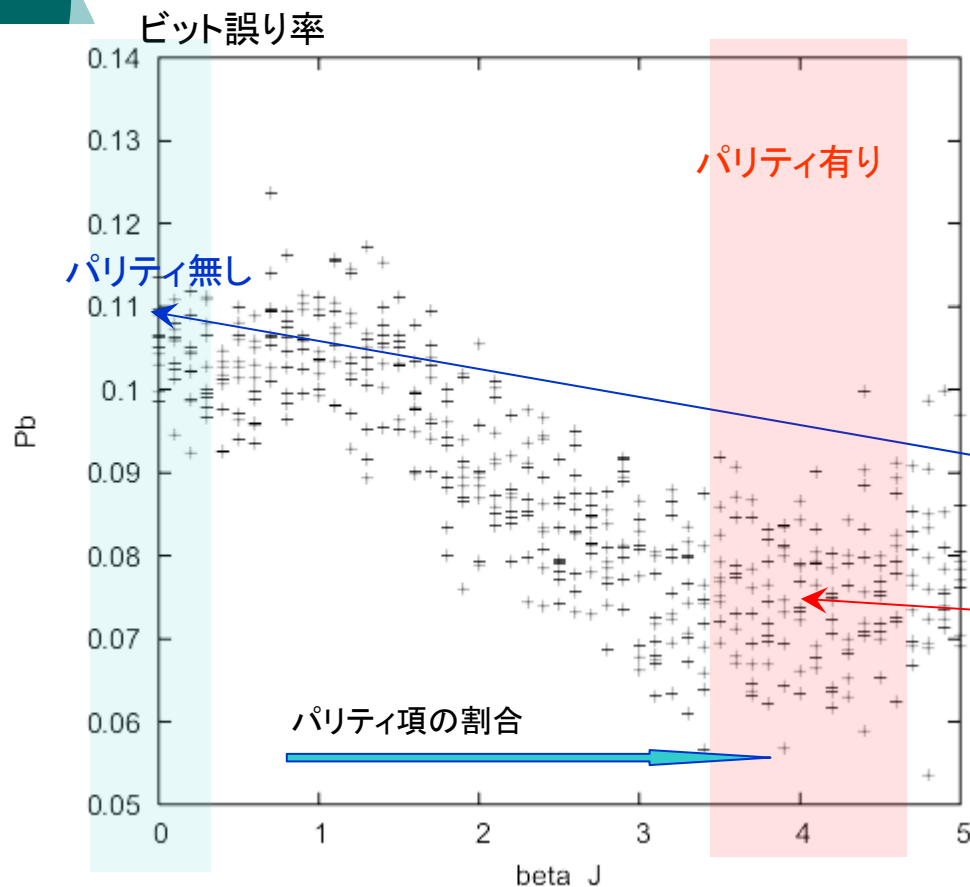
$$H_{\text{eff}} = -\beta_J \sum_{ij} J_{ij} \sigma_i \sigma_j - J \sum_{ij} \sigma_i \sigma_j - h \sum_i \tau_i \sigma_j$$

事前分布の部分

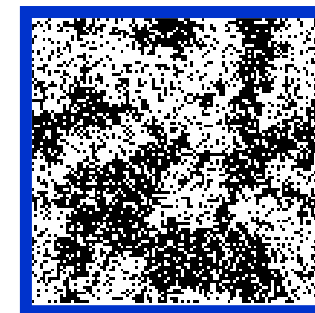


MAP復元、MPM復元を行う

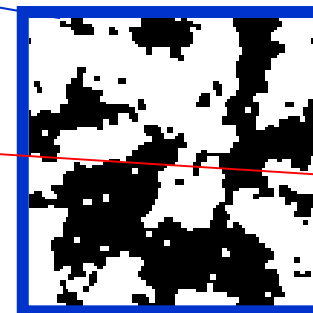
復元結果



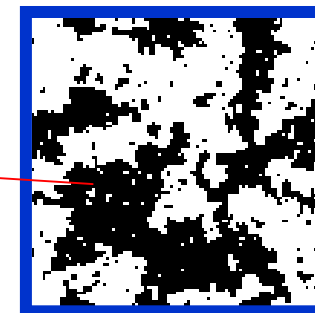
原画像



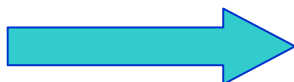
20%劣化画像



パリティ無しMPM復号



パリティ有りMPM復号



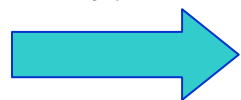
冗長性を付加することにより、より効果的な復元が実現される

Sourlas符号

N個のビットの中からp個をピックアップして送信する

$$J_{i_1 \dots i_p} = \xi_{i_1} \cdots \xi_{i_p} \quad \text{オリジナル・ビットのp個の積を送る}$$

速度

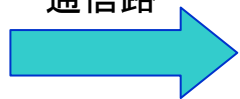


$$R = \frac{N}{N C_p} \simeq \frac{p!}{N^{p-1}}$$

ガウス通信路で送信

$$P(\{J\} | \{\xi\}) = \left(\frac{N^{p-1}}{J^2 \pi p!} \right) \exp \left[-\frac{N^{p-1}}{J^2 p!} \sum_{i_1 < \dots < i_p} \left(J_{i_1 \dots i_p} - \frac{J_0 p!}{N^{p-1}} \xi_{i_1} \cdots \xi_{i_p} \right)^2 \right]$$

通信路



$$C = \frac{1}{2} \log \left(1 + \frac{a_0^2}{a^2} \right) \simeq \frac{J_0^2 p!}{J^2 N^{p-1} \log 2}$$

シャノンの限界式

$$\frac{R}{C} = \left(\frac{J_0}{J} \right)^2 \frac{1}{\log 2}$$

平均場理論による性能評価

事後確率

$$P(\{\sigma\}|\{J\}) = \frac{e^{\beta_J \sum_{i_1 < \dots < i_p} J_{i_1 \dots i_p} \sigma_{i_1} \dots \sigma_{i_p}}}{\sum_{\{\sigma\}} e^{\beta_J \sum_{i_1 < \dots < i_p} J_{i_1 \dots i_p} \sigma_{i_1} \dots \sigma_{i_p}}}$$

エネルギー関数

$$H_{\text{eff}} = -\beta_J \sum_{i_1 < \dots < i_p} J_{i_1 \dots i_p} \sigma_{i_1} \dots \sigma_{i_p}$$

レプリカ法から得られる状態方程式

$$m = [\langle \sigma_i \rangle] = \int_{-\infty}^{\infty} Dx \tanh(G)$$

磁化

$$q = [\langle \sigma_i \rangle^2] = \int_{-\infty}^{\infty} Dx \tanh^2(G)$$

スピングラス秩序変数

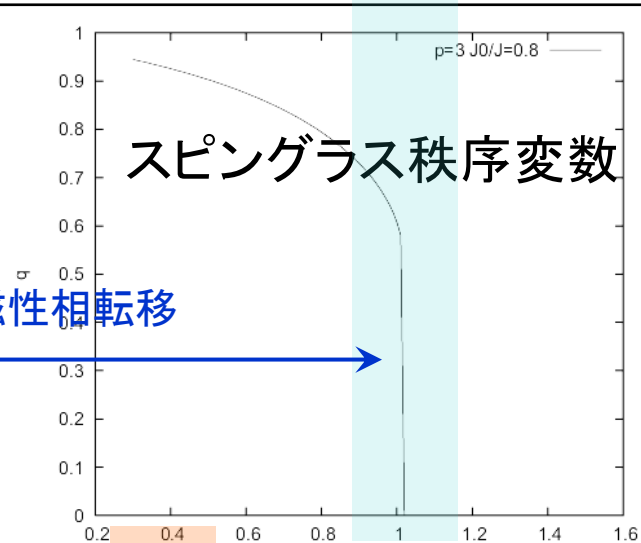
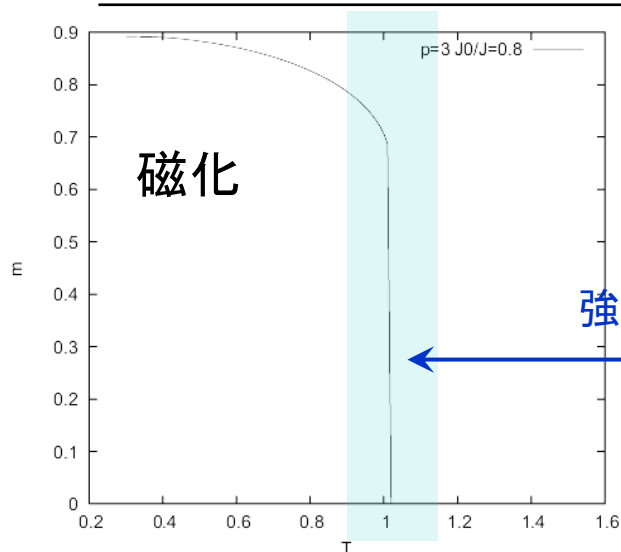
$$G = x \sqrt{\frac{p\beta_J^2 J^2 q^{p-1}}{2}} + \beta_J J_0 p m^{p-1}$$

セルフ・コンシステントに解く

$$P_b = \frac{1}{2} \left(1 - \left[\xi_1 \operatorname{sgn}(\langle \sigma \rangle_i) \right] \right) = \frac{1}{2} \left(1 - \int_{-\infty}^{\infty} Dx \operatorname{sgn}(G) \right)$$

ビット誤り率

解析結果 #1

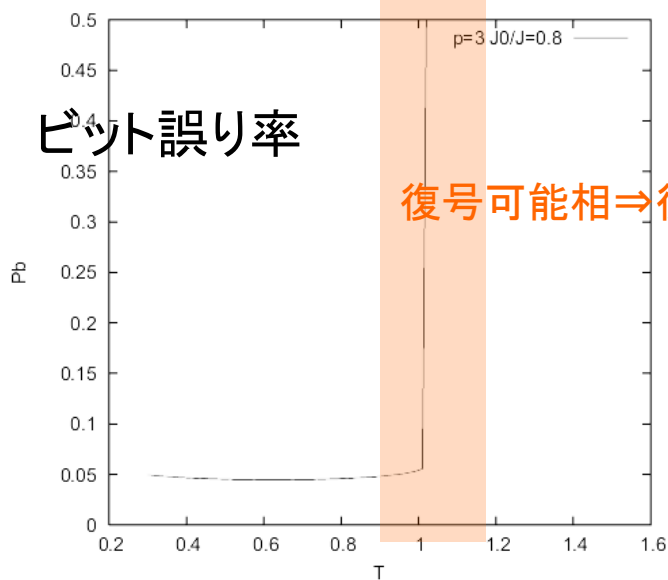


強磁性⇒常磁性相転移



$$p = 3$$
$$\frac{J_0}{J} = 0.8$$

シグナルノイズ比



復号可能相⇒復号不可能相転移

解析結果 #2

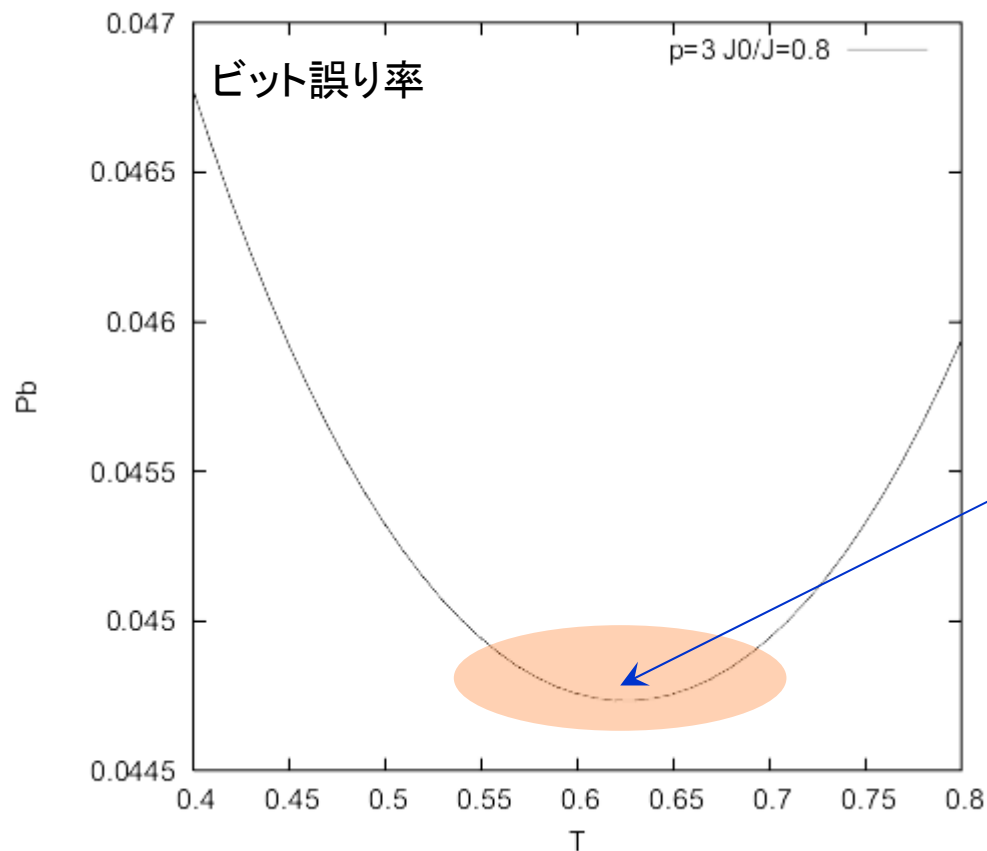
温度に関する最良性

$$p = 3$$

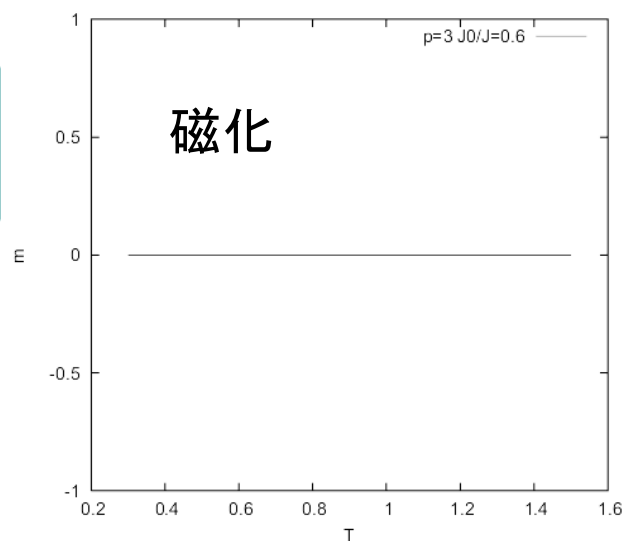
$$J_0/J = 0.8$$

西森温度

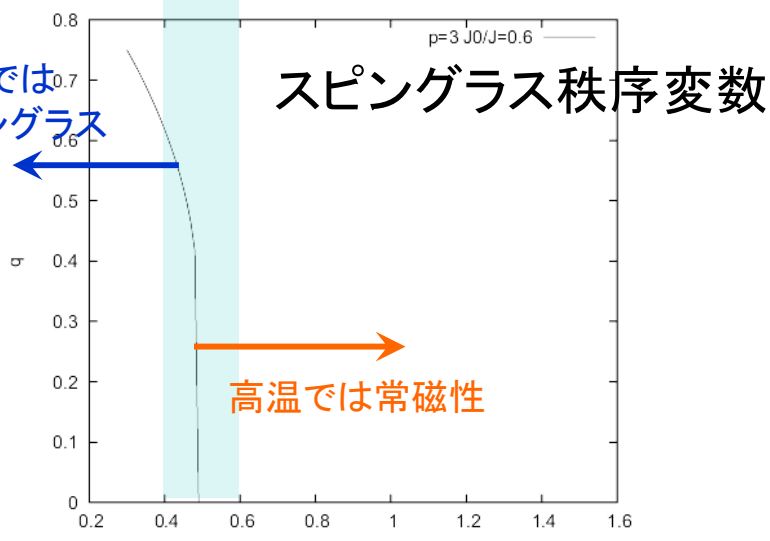
$$T = \frac{J^2}{2J_0^2}$$



解析結果 #3



低温では
スピングラス

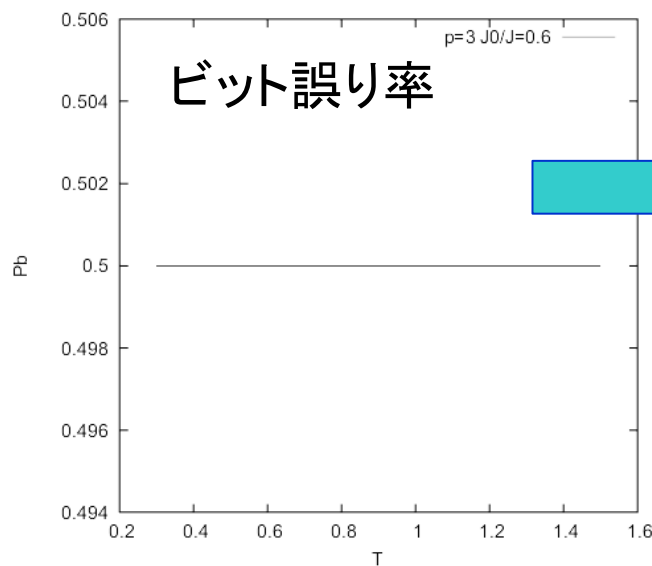


高温では常磁性

$$p = 3$$

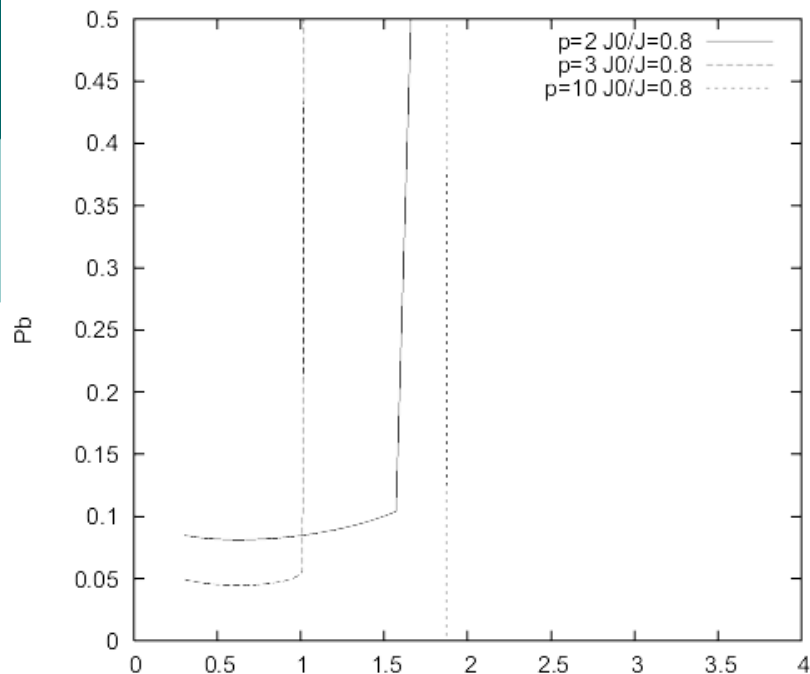
$$\frac{J_0}{J} = 0.6$$

シグナルノイズ比



誤り訂正不可能

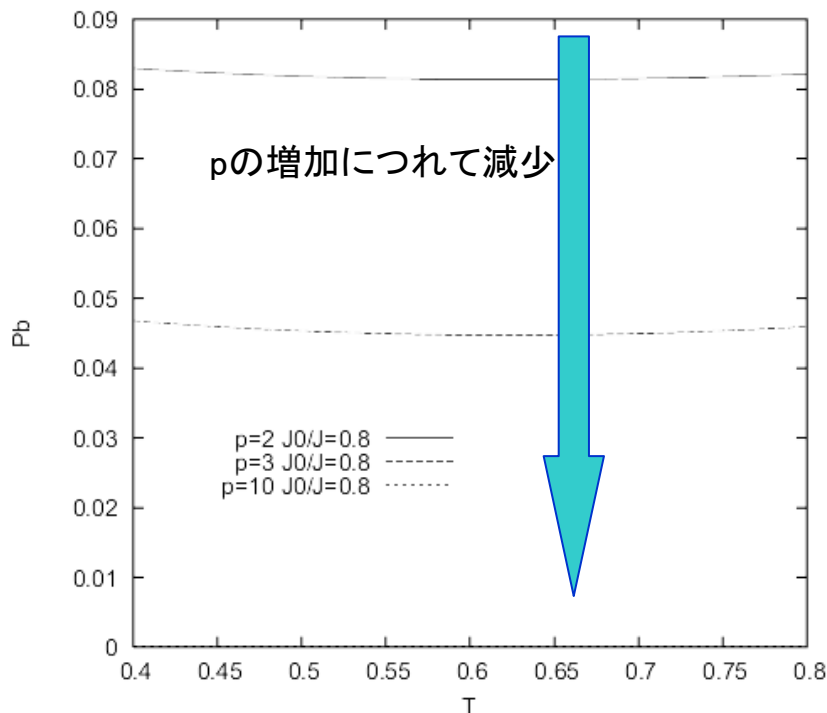
解析結果 #4



$p \rightarrow \infty$:

$$\left(\frac{J_0}{J} \right) \geq \sqrt{\log 2},$$

強磁性相の存在限界



p の増加につれて減少

$$R = C$$

Sourlas符号は漸近的にシャノン限界を達成する

次回はCDMAに移ります