

混沌系工学特論 #4

情報科学研究科 井上純一

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

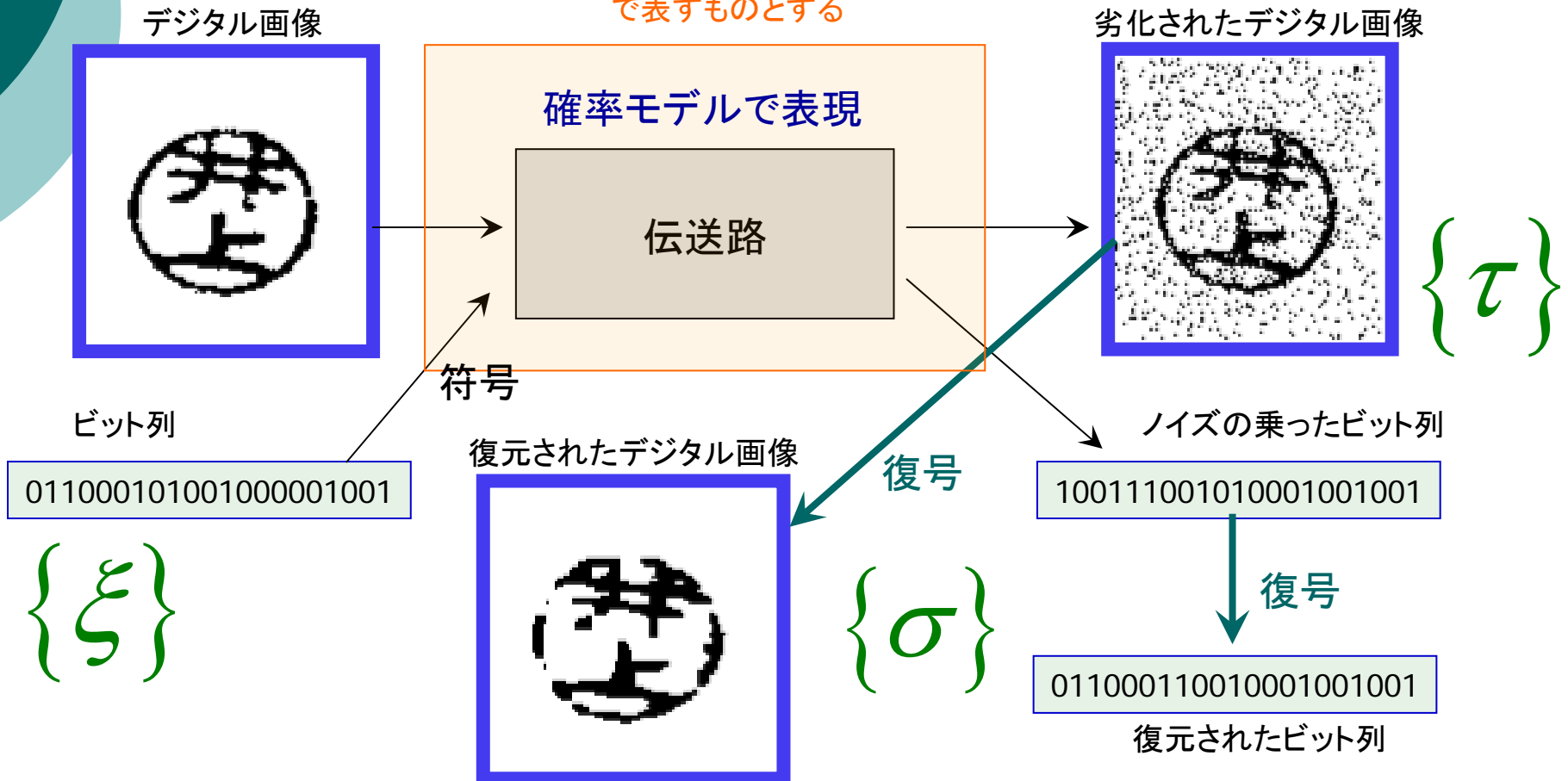
Mirror : http://www005.upp.so-net.ne.jp/j_inoue/index.html

平成17年11月7日 第4回講義

デジタルデータの転送と復元

前回までの復習

※ 古文書/文献の復元等では
時間経過による劣化を伝送路
で表すものとする



ベイズの公式と事後確率

前回までの復習

事後確率分布

尤度(ゆうど)

事前確率分布

$$P(\{\sigma\} | \{\tau\}) = \frac{P(\{\tau\} | \{\sigma\}) P(\{\sigma\})}{\sum_{\{\sigma\}} P(\{\tau\} | \{\sigma\}) P(\{\sigma\})}$$

原画像のピクセル列の推定値

ベイズの公式

尤度は劣化過程の確率モデル

2元対称通信路であることが既知ならば

$$P(\{\tau\} | \{\sigma\}) = \prod_i \frac{\exp(h \tau_i \sigma_i)}{2 \cosh(h)}$$

未知な変数を含む

事前分布の選び方は全く任意である



画像データの特徴を生かして適切に選択する必要がある

最大事後確率推定

前回までの復習

事後確率

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

エネルギー関数

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

事後確率最大



エネルギー関数最小
⇒ シミュレーテッド・アニーリングで最適化

画像復元が最適化問題として定式化される

このパラメータ T を温度としたアニーリング

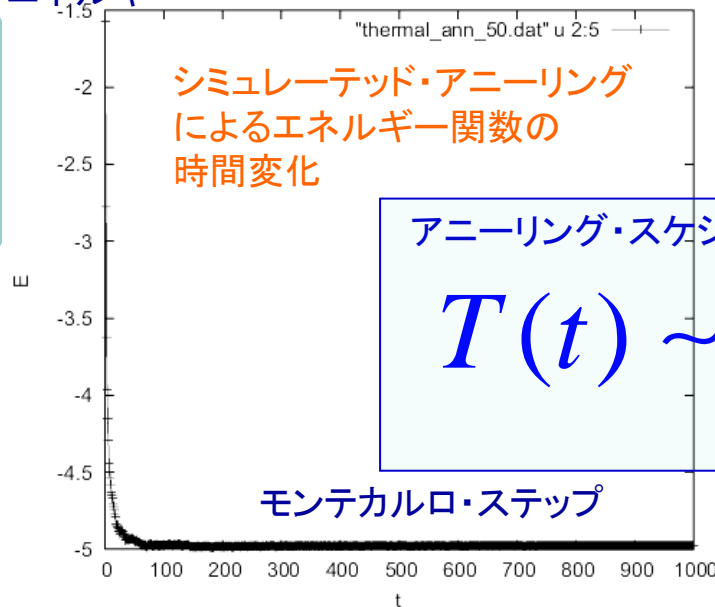
劣化画像データは与えられている

$$P(\{\sigma\}) = \frac{\exp\left(T^{-1} \left\{ J \sum_{\langle ij \rangle} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i \right\}\right)}{\sum_{\sigma} \exp\left(T^{-1} \left\{ J \sum_{\langle ij \rangle} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i \right\}\right)}$$

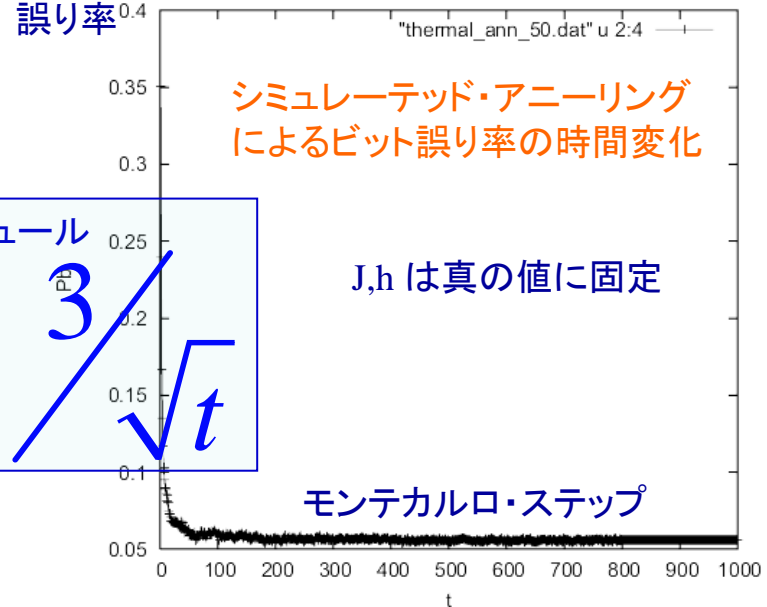
アニーリング法による画像復元の実例

前回までの復習

エネルギー



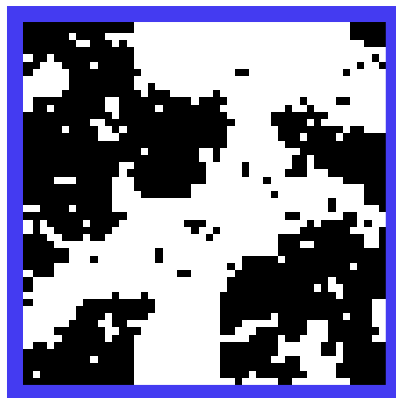
誤り率



アニーリング・スケジュール

$$T(t) \sim \frac{3}{\sqrt{t}}$$

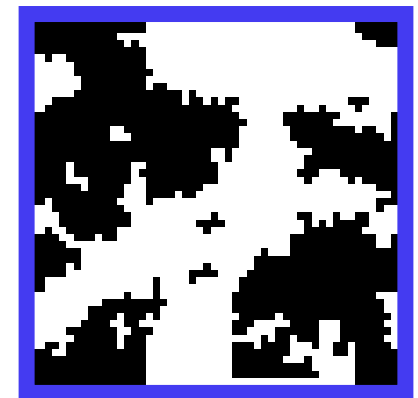
J, h は真の値に固定



原画像



劣化画像

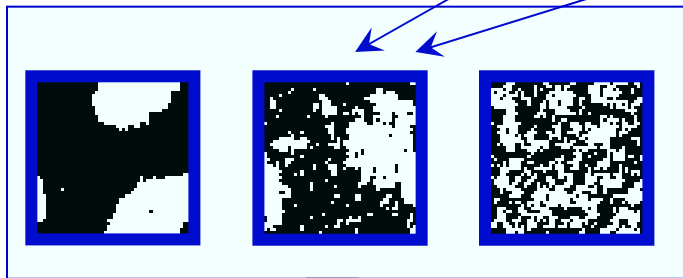


復元画像

確率モデルのパラメータ推定

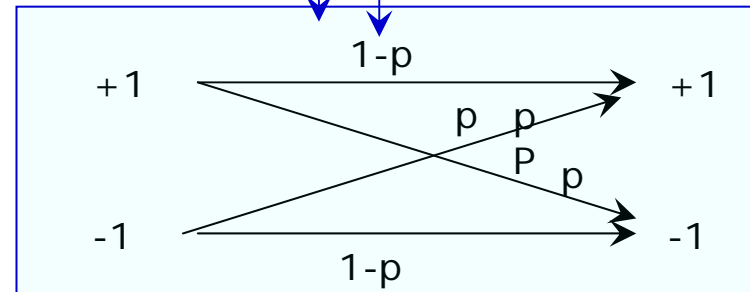
事後確率は次のように書けるが...

$$P(\{\sigma\}) = \frac{\exp\left(T^{-1} \left\{ J \sum_{\langle ij \rangle} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i \right\}\right)}{\sum_{\sigma} \exp\left(T^{-1} \left\{ J \sum_{\langle ij \rangle} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i \right\}\right)}$$



どの J で指定されるスナップショットの張り合わせで画像を表現するのか

未知



周辺尤度最大化の基準で決める

劣化率 h がどのくらいなのか

未知

用いることのできるのは劣化画像についての情報のみ

周辺尤度最大化

前回の復習

周辺尤度

$$L(J, h; \{\tau\}) = \log \text{tr}_{\{\sigma\}} P_h(\{\tau\} | \{\sigma\}) P_J(\{\sigma\})$$

観測データ

$$= \log \text{tr}_{\{\sigma\}} e^{J \sum_{\langle ij \rangle} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i}$$

$$+ \log \text{tr}_{\{\sigma\}} e^{J \sum_{\langle ij \rangle} \sigma_i \sigma_j}$$

$$- \log 2 \cosh(h)$$

一般的にKL情報量の非負性より

$$\left[L(J_*, h_*; \{\tau\}) \right]_{\{\tau\}} \geq \left[L(J, h; \{\tau\}) \right]_{\{\tau\}}$$

真のパラメータ値で最大

周辺尤度をコスト関数とみなしてパラメータに関して最大化すればよい

EM法による周辺尤度の最大化

前回の復習

周辺尤度を最大化する代わりに、次の尤度関数の事後確率での期待値を最大化

$$Q(J, h | J_t, h_t) = \underbrace{\text{tr}_{\{\sigma\}} P_{J_t, h_t}(\{\sigma\} | \{\tau\})}_{\text{事後確率}} \log \underbrace{P_h(\{\tau\} | \{\sigma\}) P_J(\{\sigma\})}_{\text{尤度関数}}$$

EMアルゴリズム

$$J_{t+1} = \arg \max_J Q(J, h | J_t, h_t)$$

$$h_{t+1} = \arg \max_h Q(J, h | J_t, h_t)$$

→ Eステップ(期待値計算)
とMステップ(最大化)
の繰り返しで構成される

一方の変数hが既知の場合

$$u_m(J_{t+1}) = \frac{\text{tr}_{\{\sigma\}} \varepsilon_B e^{J_{t+1} \sum_{\langle ij \rangle} \sigma_i \sigma_j}}{\text{tr}_{\{\sigma\}} e^{J_{t+1} \sum_{\langle ij \rangle} \sigma_i \sigma_j}} = \frac{\text{tr}_{\{\sigma\}} \varepsilon_B e^{J_t \sum_{\langle ij \rangle} \sigma_i \sigma_j + h_* \sum_i \tau_i \sigma_i}}{\text{tr}_{\{\sigma\}} e^{J_t \sum_{\langle ij \rangle} \sigma_i \sigma_j + h_* \sum_i \tau_i \sigma_i}} = u_p(J_t, h_*)$$

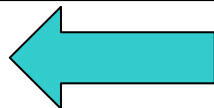
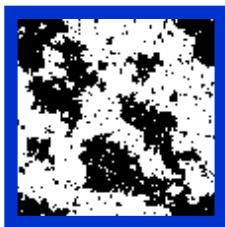
厳密解あり

厳密解ないのでMCMC法で計算

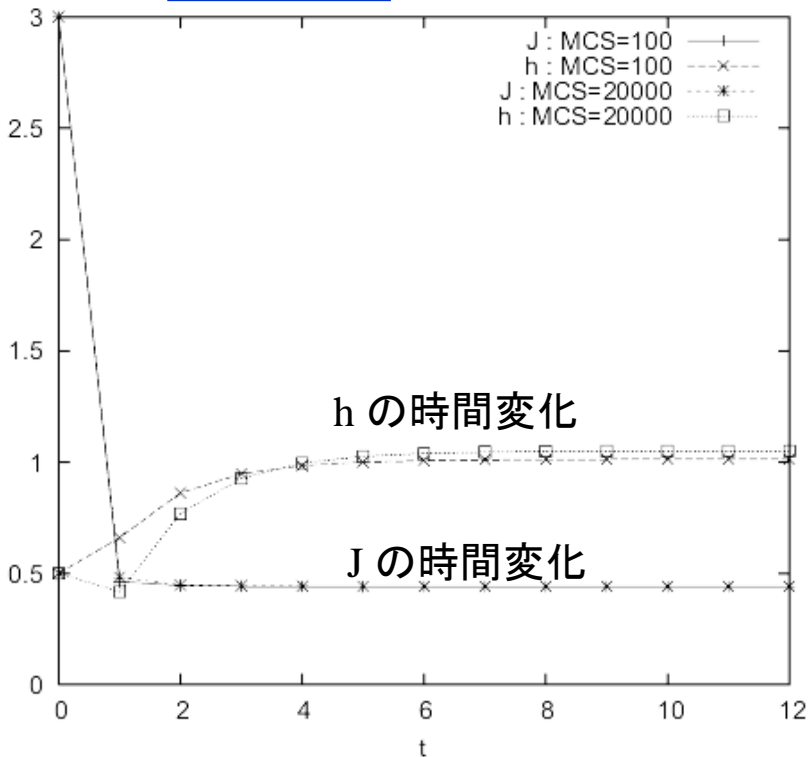
$$\varepsilon_B = -(1/2N) \sum_{\langle ij \rangle} \sigma_i \sigma_j$$

実行結果

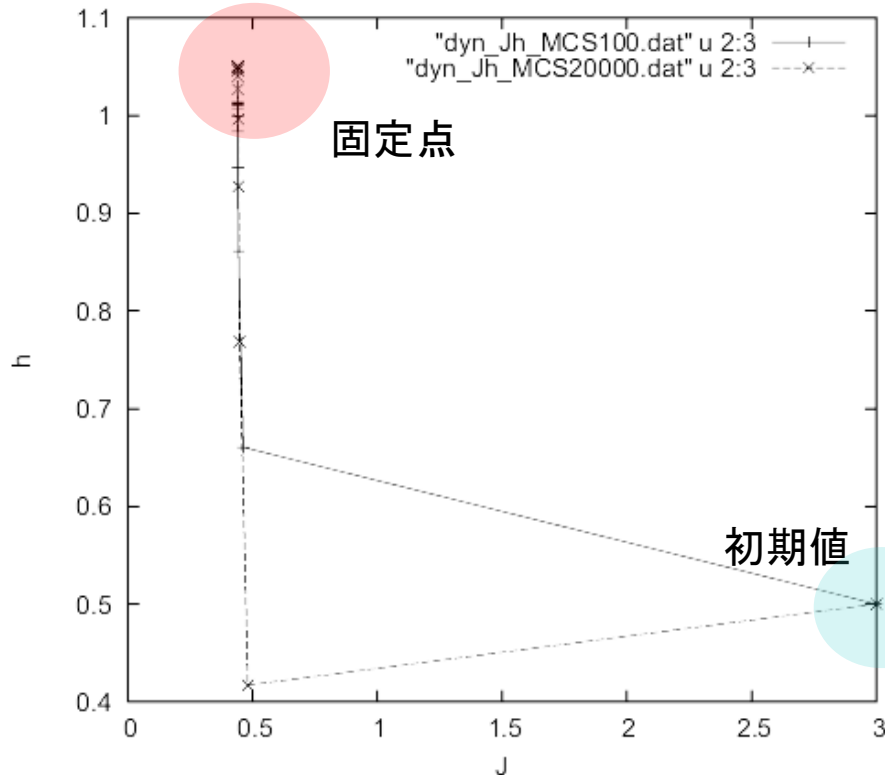
前回の復習 (原画像: イジングモデル)



原画像



ハイパーパラメータの時間発展



相空間での軌道

復元結果はデータに依存する

原画像依存性

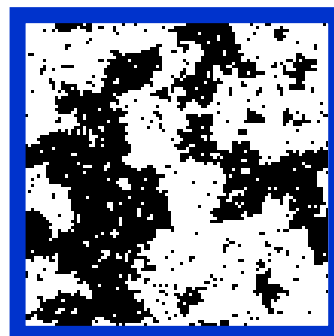
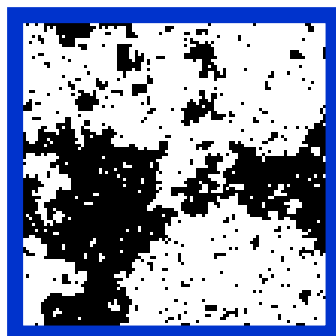
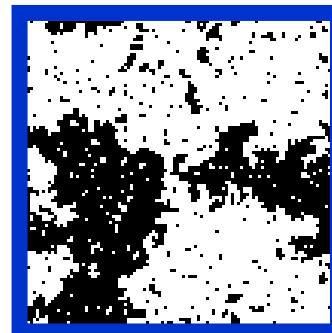
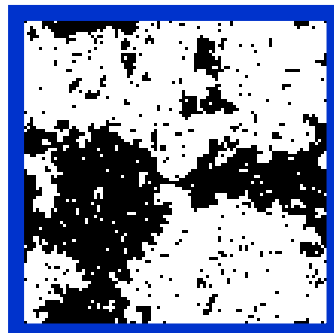
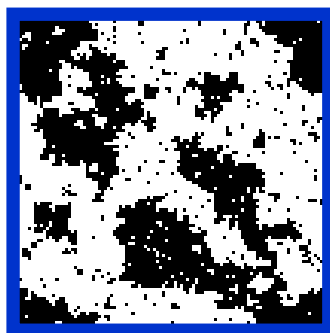
原画像をスナップショットに選ぶ (今は $T=1/J=2.15$ とする)

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

同じマクロな量

$$m = \frac{1}{N} \sum_{i=1}^N \sigma_i$$

であっても
ミクロにみれば
異なる

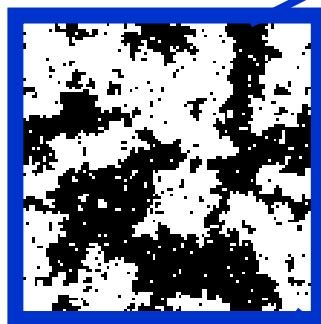


ビット誤り率について原画像の
生成分布に関する平均操作
を行わなければならない

復元結果はデータに依存する

劣化過程依存性

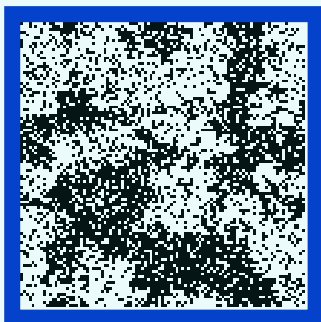
マクロな劣化率 p は同じ
でもミクロにみれば
劣化のされ方は異なる



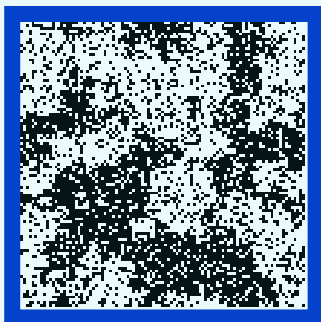
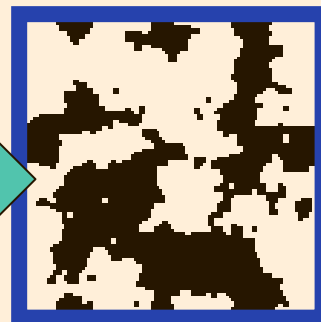
原画像

100 × 100

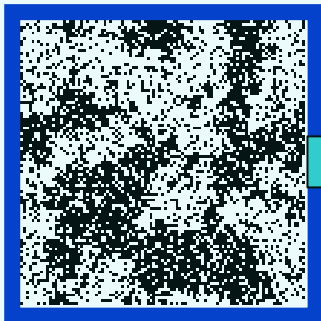
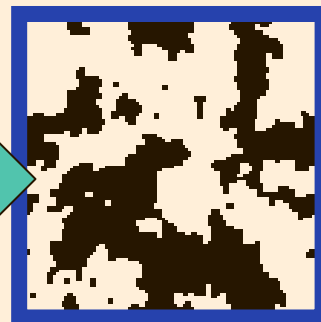
劣化率
 $P=0.2$ のBSC



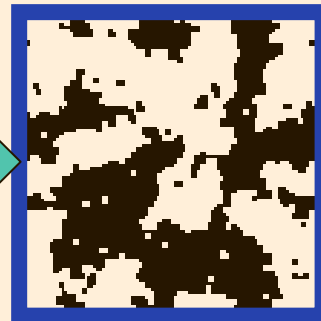
MAP
復元



MAP
復元



MAP
復元



復元結果
も劣化の
され方に依存
する

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

熱揺らぎを用いた画像の復元

周辺分布 $P(\sigma_i | \tau)$ による各画素ごとの期待値

画素 j の推定値

$$\hat{\xi}_j = \text{sgn}(\langle \sigma_j \rangle) = \text{sgn} \left[\sum_{\sigma_j = \pm 1} \sigma_j P(\sigma_j | \tau) \right]$$

MPM推定値

$$P(+1 | \tau) - P(-1 | \tau)$$

一つの画素に注目して
アンサンブルでの多数決をとる

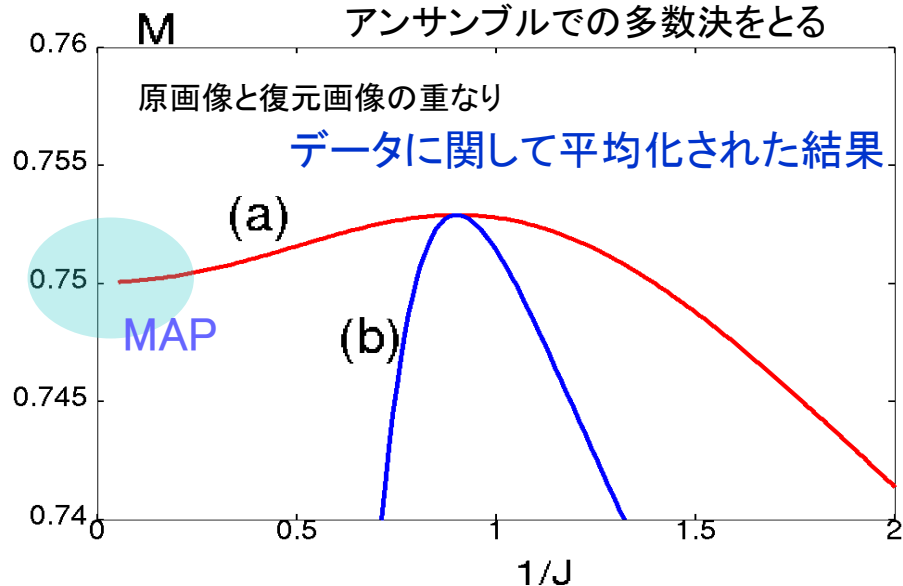
エネルギー関数 H で記述される磁性体の局所磁化

$$\langle \sigma_i \rangle = \frac{\sum_{\sigma} \sigma_i \exp[-H(\sigma | \tau)]}{Z} = m_i$$

各画素の期待値

どのようにして
平均的な性能を
評価するのか

今週の内容



原画像と復元画像の重なり
データに関して平均化された結果

MAP

平衡状態でのビット誤り率

平均場モデルの解析

エネルギー関数を書き直す(平均場近似)

$$H = -\frac{J}{N} \sum_{ij} \sigma_i \sigma_j - h \sum_i \tau_i \sigma_i = -Jm \sum_i \sigma_i - h \sum_i \tau_i \sigma_i, \quad m = \frac{1}{N} \sum_i \sigma_i$$

磁化に関して自己無撞着な方程式を作る

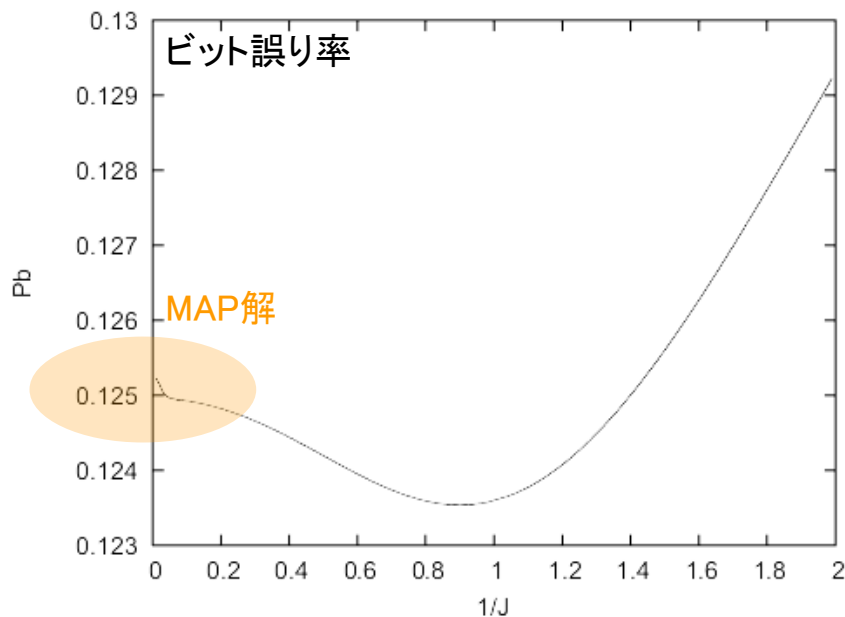
$$m = \left\langle \frac{1}{N} \sum_i \sigma_i \right\rangle = \frac{\text{tr}_{\{\sigma\}} \left(\left(\frac{1}{N} \sum_i \sigma_i \right) e^{Jm \sum_i \sigma_i + h \sum_i \tau_i \sigma_i} \right)}{\text{tr}_{\{\sigma\}} e^{Jm \sum_i \sigma_i + h \sum_i \tau_i \sigma_i}} = \frac{1}{N} \sum_i \tanh(Jm + h\tau_i) = \left[\tanh(Jm + h\tau) \right]$$

十分大きな1枚の画像に対する左辺の値はその平均値(右辺)に一致する

ビット誤り率について

$$\begin{aligned} \langle \sigma_i \rangle &= \tanh(Jm + h\tau_i) \\ P_b &= \frac{1}{2} \left(1 - \frac{1}{N} \sum_i \xi_i \text{sgn}(\langle \sigma_i \rangle) \right) \\ &= \frac{1}{2} \left(1 - \left[\xi \text{sgn}(Jm + h\tau) \right] \right) \end{aligned}$$

[...]はデータに関する平均



勾配法による周辺尤度の最大化

周辺尤度

$$L(J, h; \{\tau\}) = \log \text{tr}_{\{\sigma\}} P_h(\{\tau\} | \{\sigma\}) P_J(\{\sigma\})$$

$$= \log \text{tr}_{\{\sigma\}} e^{J \sum_{\langle ij \rangle} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i} + \log \text{tr}_{\{\sigma\}} e^{J \sum_{\langle ij \rangle} \sigma_i \sigma_j} + \frac{Na_0^2}{2a^2} - \frac{Na^2 h^2}{2}$$

に対し、次の勾配法を考える

データを含む平均値

$$C_J \frac{dJ}{dt} = -\frac{\partial L}{\partial J} = \frac{\text{tr}_{\{\sigma\}} \left(\sum_{ij} \sigma_i \sigma_j \right) e^{J \sum_{ij} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i}}{\text{tr}_{\{\sigma\}} e^{J \sum_{ij} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i}} - \frac{\text{tr}_{\{\sigma\}} \left(\sum_{ij} \sigma_i \sigma_j \right) e^{J \sum_{ij} \sigma_i \sigma_j}}{\text{tr}_{\{\sigma\}} e^{J \sum_{ij} \sigma_i \sigma_j}}$$

データを含まない平均値

$$C_h \frac{dh}{dt} = -\frac{\partial L}{\partial h} = \frac{\text{tr}_{\{\sigma\}} \left(\sum_i \tau_i \sigma_i \right) e^{J \sum_{ij} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i}}{\text{tr}_{\{\sigma\}} e^{J \sum_{ij} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i}} - Na^2 h$$

データを含む平均値

目標

期待値計算とデータ平均
を共に解析的に行う

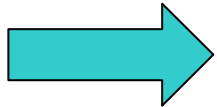
勾配法の動的性質

平均場モデルでの解析

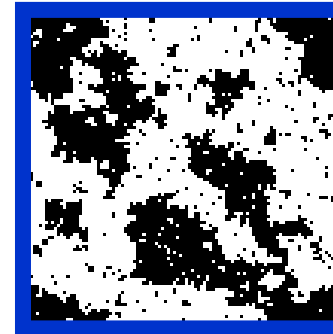
$$C_J \frac{dJ}{dt} = -\frac{\partial L}{\partial J} = \left[\frac{\text{tr}_{\{\sigma\}} \left(\sum_{ij} \sigma_i \sigma_j \right) e^{J \sum_{ij} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i}}{\text{tr}_{\{\sigma\}} e^{J \sum_{ij} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i}} \right] - \frac{\text{tr}_{\{\sigma\}} \left(\sum_{ij} \sigma_i \sigma_j \right) e^{J \sum_{ij} \sigma_i \sigma_j}}{\text{tr}_{\{\sigma\}} e^{J \sum_{ij} \sigma_i \sigma_j}}$$

$$C_h \frac{dh}{dt} = -\frac{\partial L}{\partial h} = \left[\frac{\text{tr}_{\{\sigma\}} \left(\sum_i \tau_i \sigma_i \right) e^{J \sum_{ij} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i}}{\text{tr}_{\{\sigma\}} e^{J \sum_{ij} \sigma_i \sigma_j + h \sum_i \tau_i \sigma_i}} \right] - Na^2 h$$

平均場モデルの解析では



詳細は講義ノート



原画像は
イジングモデル
のスナップショット

$$C_J \frac{dJ}{dt} = -\frac{m^2}{2} + m \frac{\text{tr}_{\xi} e^{\beta_s m_0 \xi}}{2 \cosh(\beta_s m_0)} \int_{-\infty}^{\infty} Dx \tanh(Jm + hax + ha_0 \xi) + \frac{m_1^2}{2} - m_1 \tanh(m_1 J)$$

$$C_h \frac{dh}{dt} = \frac{\text{tr}_{\xi} e^{\beta_s m_0 \xi}}{2 \cosh(\beta_s m_0)} \int_{-\infty}^{\infty} Dx (ax + a_0 \xi) \tanh(Jm + hax + ha_0 \xi) - a^2 h$$

$$C_m \frac{dm}{dt} = -m + \frac{\text{tr}_{\xi} e^{\beta_s m_0 \xi}}{2 \cosh(\beta_s m_0)} \int_{-\infty}^{\infty} Dx \tanh(Jm + hax + ha_0 \xi)$$

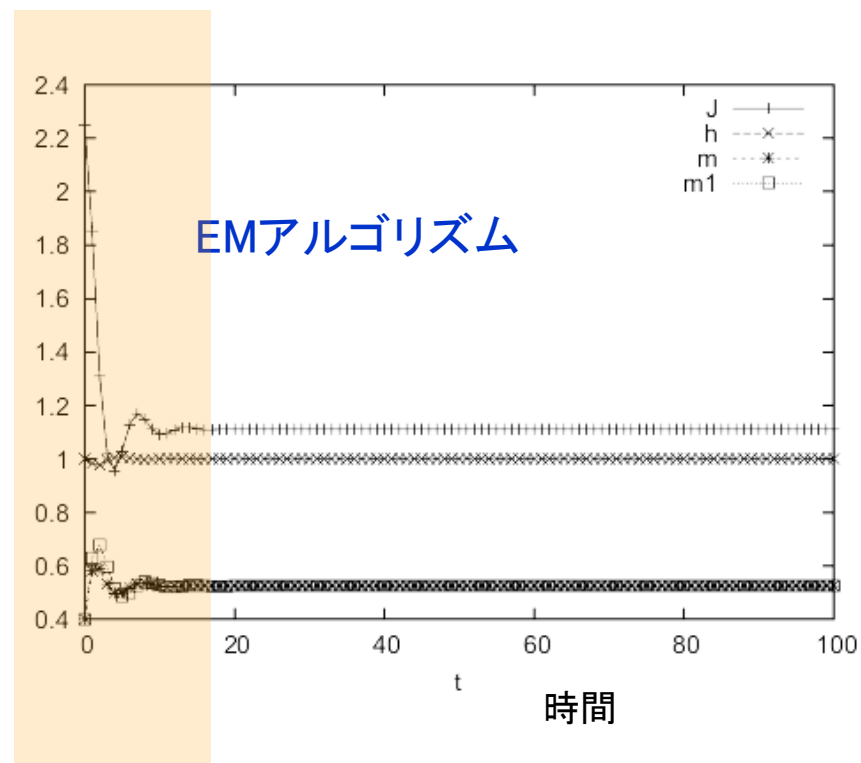
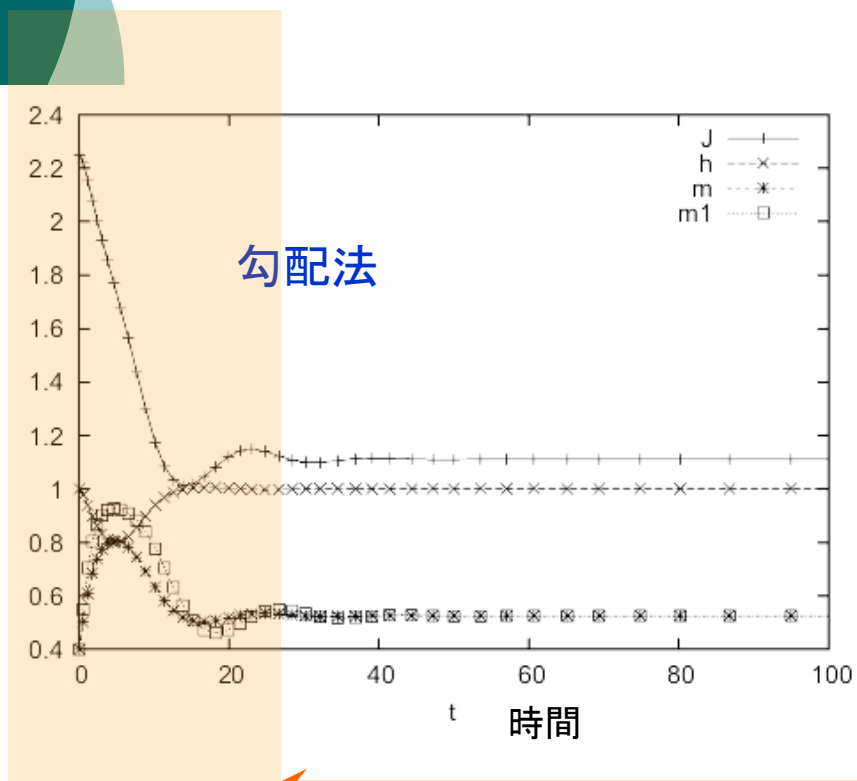
事後確率で与えられる系の磁化

$$C_{m_1} \frac{dm_1}{dt} = -m_1 + \tanh(Jm_1)$$

事前確率で与えられる系の磁化

解析結果

平均評価に基づくEMアルゴリズムとの比較

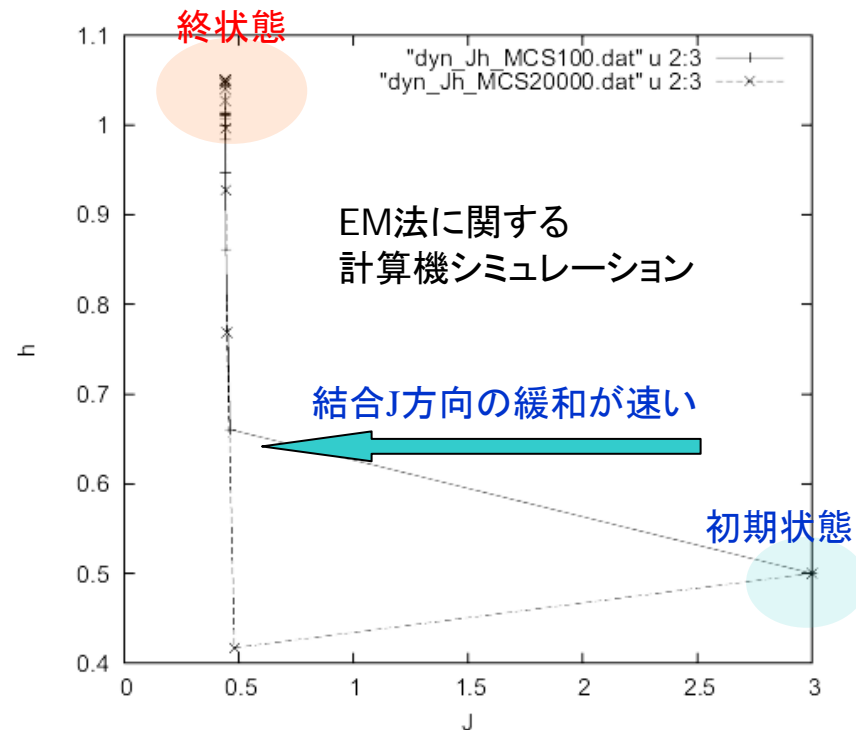
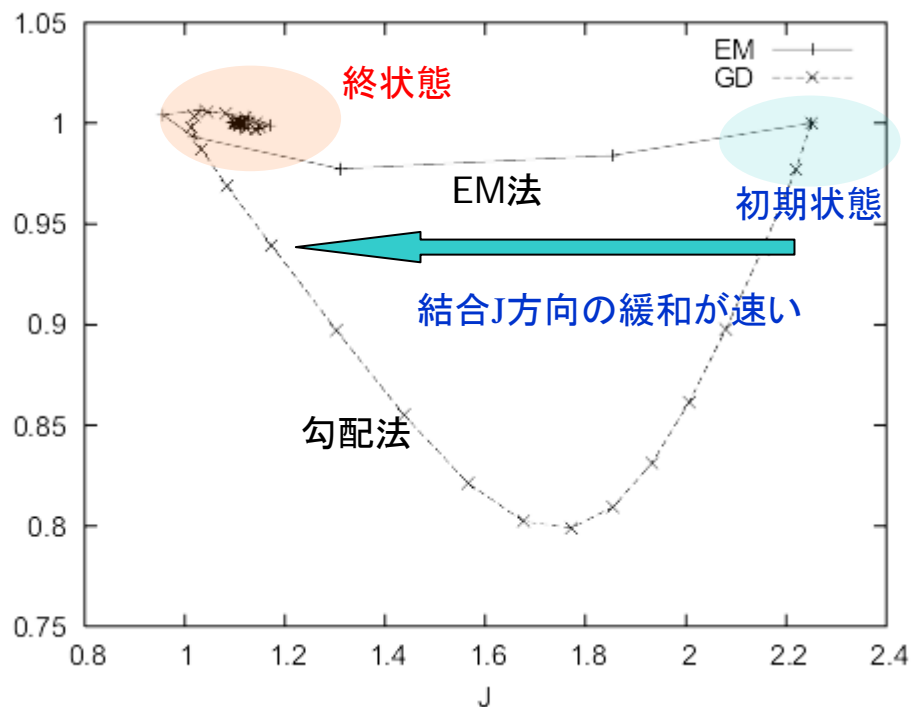


EM法がより速い初期収束

解析結果 #2

平均評価に基づくEMアルゴリズムとの比較

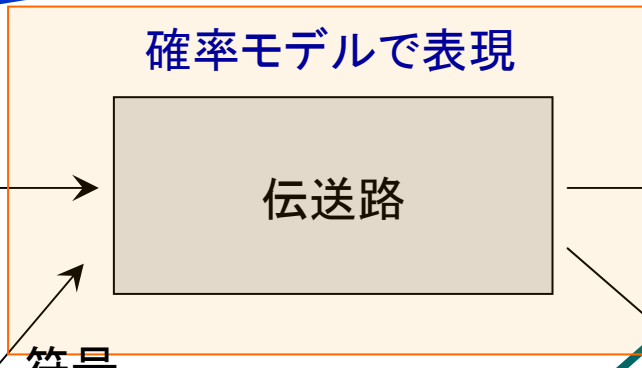
ハイパーパラメータ空間内での軌道



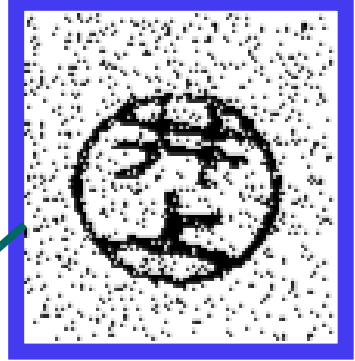
デジタルデータの転送と復元再考

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

画像の復元では画像の特質を
事前分布で表現できる



劣化されたデジタル画像



{ τ }

ビット列

011000101001000001001

{ ξ }

ビット列の場合は？

復元されたデジタル画像



{ σ }

ノイズの乗ったビット列

100111001010001001001

復号

011000110010001001001

復元されたビット列

復号

雑音のある状況でのデータ送信

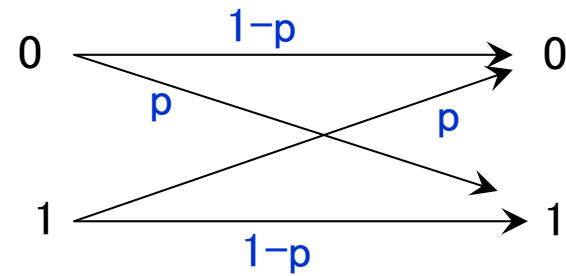
誤り訂正符号

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

$$f_e^{(5)}(p) = {}_5C_3 p^3 (1-p)^2 + {}_5C_4 p^4 (1-p) + p^5$$
$$= 10 p^3 (1-p)^2 + 5 p^4 (1-p) + p^5$$

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

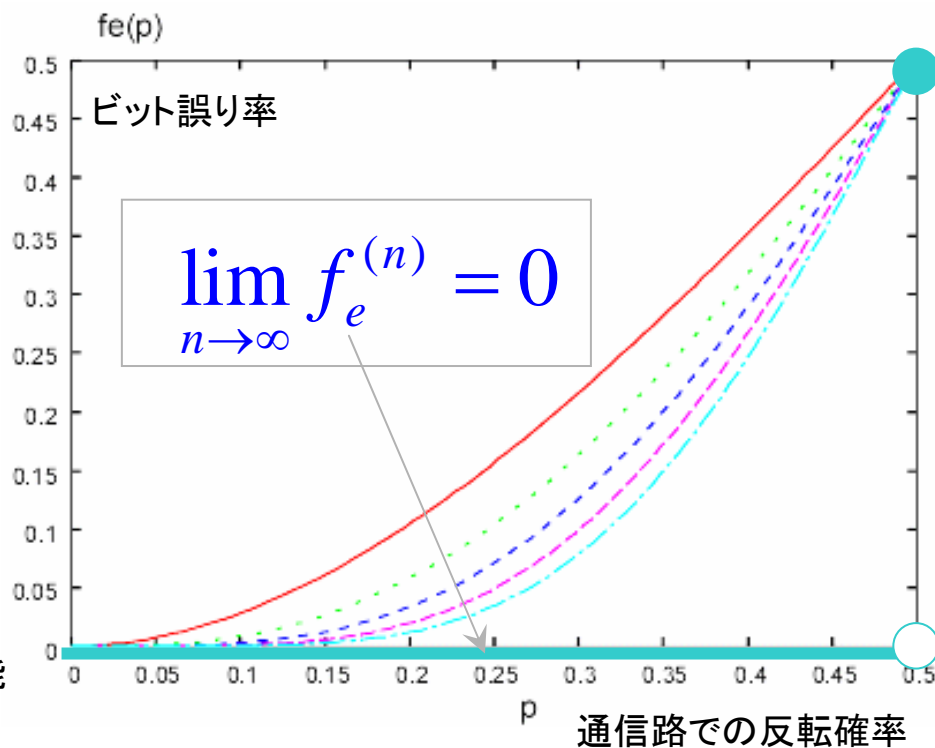
繰り返し回数を限りなく多くとると、
誤り確率は限りなくゼロに近づくが、
同時に伝送速度もゼロになってしまう

⇒ 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

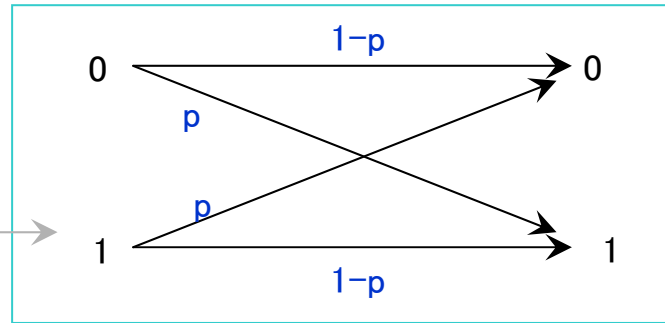
↓ 入力の確率分布に関して、相互情報量を最大化したものが通信路容量

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

$$P_{Y|X}(0|0) = P_{Y|X}(1|1) = 1-p$$
$$P_{Y|X}(0|1) = P_{Y|X}(1|0) = p$$



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

$S_1 \leftarrow 0010 \dots 000$

$S_2 \leftarrow 0100 \dots 001$

$S_3 \leftarrow 0010 \dots 100$

.....

$S_M \leftarrow 1000 \dots 010$

符号長 n

各ビットに $\frac{1}{2}$ の確率で 0,1 を割り振る

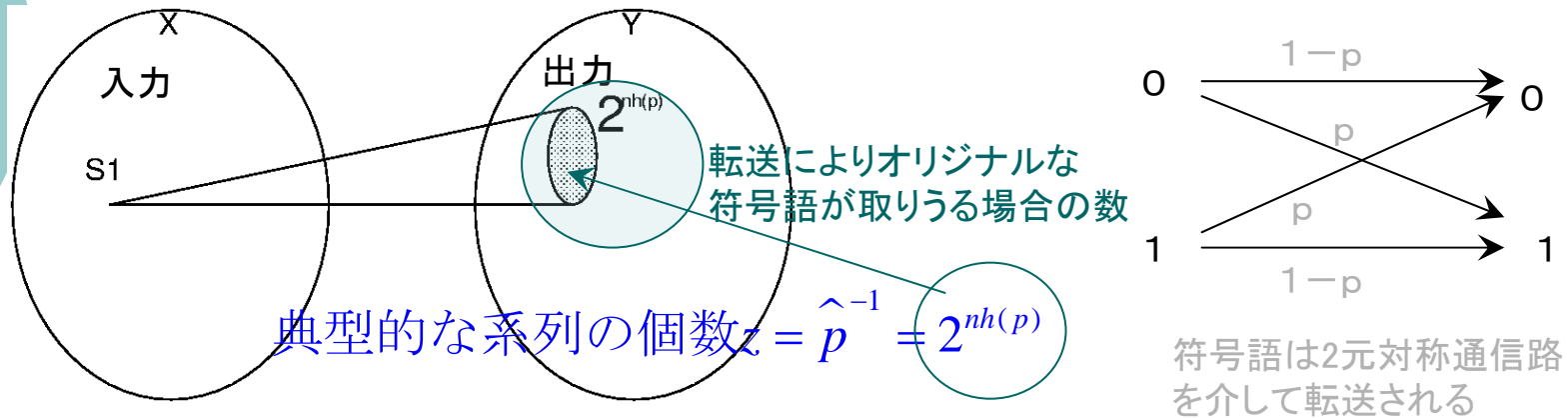
$$M = 2^{nR} < 2^n$$

$R \leq 1$ 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



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よそ np と見積もられる

$(0100 \dots 011)$	$(0111 \dots 100)$	\dots	$(0101 \dots 111)$
np ビットの誤り	np ビットの誤り		np ビットの誤り

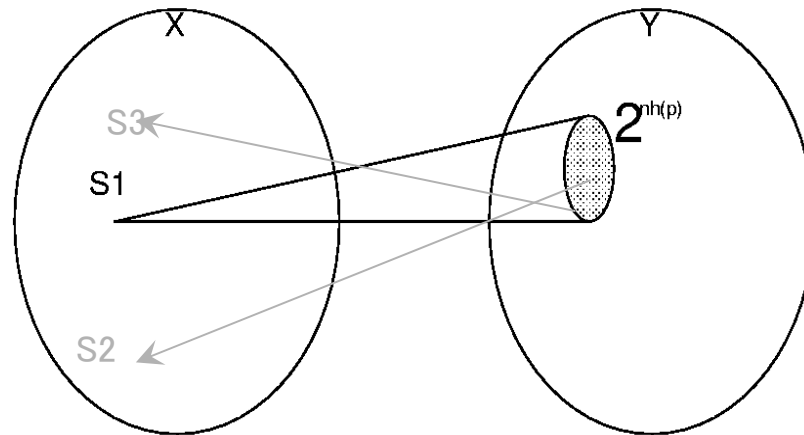
通信路出力の典型的な系列

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) の証明終わり

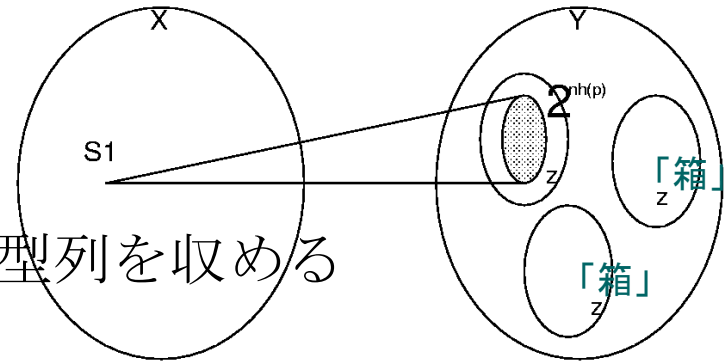
通信路符号化定理の証明#3

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

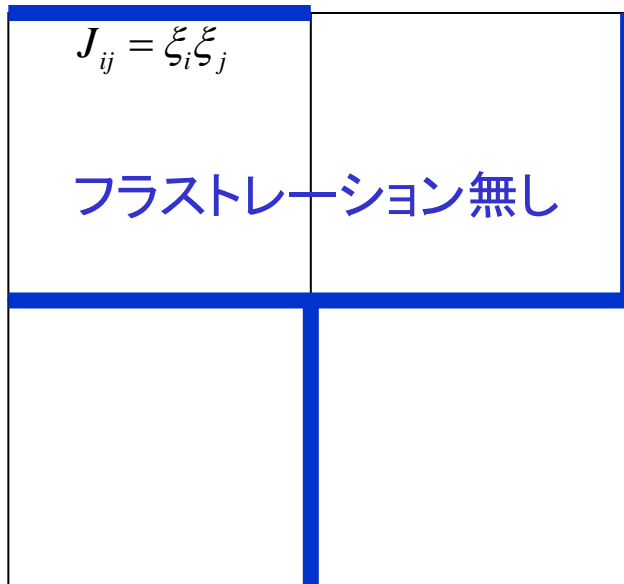
ii)の証明終わり

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

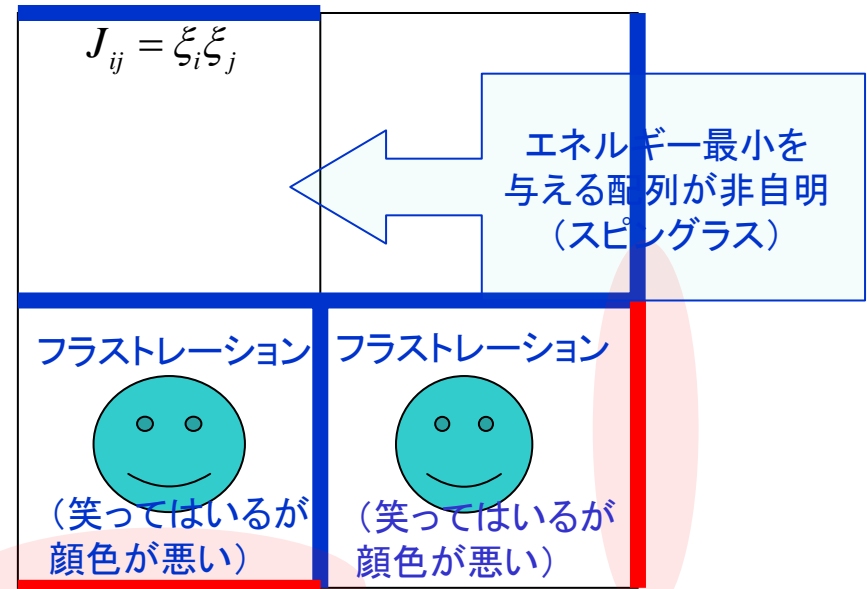
エネルギー関数 $H = -\sum_{ij} \xi_i \xi_j \sigma_i \sigma_j$ エネルギー関数の最小状態を選ぶことで復号する

ノイズが無い場合



詳細は次回

ノイズのある場合



ノイズにより反転したボンド