空間結合符号

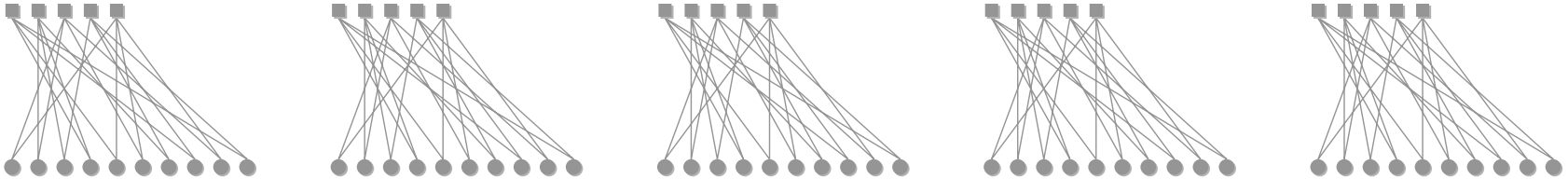
笠井 健太

東京工業大学

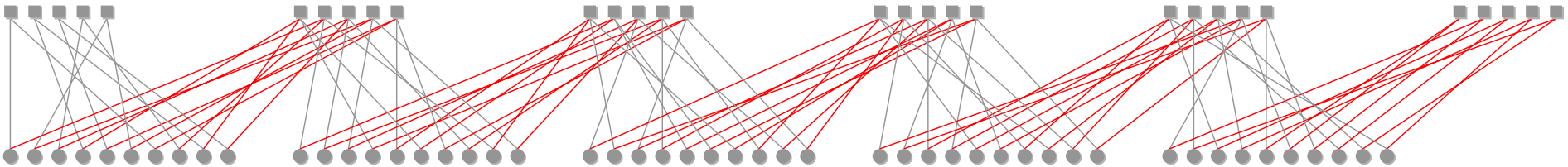
平成２６年１１月１８日
IBIS2014 @名古屋大学

最近の符号理論の研究で分かったこと

ルーピー BP によって解かれるシステムをいくつか用意して，チェイン状に結合させると大域的な MAP 解をルーピー BP で解くことができる.



システムを $L(=5)$ 個コピーしてもルーピー BP で訂正できるノイズレベルは個々のシステムのそれと同じだが,



隣と結合して大きな結合システムを作ると，ルーピー BP で訂正できるノイズレベルは個々のシステムが MAP で訂正できるノイズレベルまで向上する

符号理論が解決しようとしている問題

$\underline{x} = (x_1, \ldots, x_n) \in C(H)$ を送信して $\underline{y} = (y_1, \ldots, y_n)$ を受信する.

$$C(H) := \{x \in \{0,1\}^n | H\underline{x} = \underline{0}\} \quad (H \in \{0,1\}^{m \times n}, n \simeq 10^5)$$

送信語は一様分布することを仮定

$$\Pr(\hat{X} = \hat{x}) := \frac{\mathbb{1}[\underline{x} \in C]}{\#C}$$

ノイズの統計的性質 $\Pr(\underline{Y}|\underline{X})$ を知っていることを仮定する.
簡単のために独立性も仮定する.

$$\Pr(\underline{Y}|\underline{X}) = \prod_{j=1}^{n} \Pr(Y_j|X_j)$$

符号理論の目的

このとき，H をうまく設計して $\#C(H)$ を $2^{nI(X;Y)}$ にできるだけ近づけて，エラーレート $\Pr(x_j(Y) \neq X_j)$ をできるだけ小さくする推定 $\hat{x}_j(\underline{y})$ を $O(n)$ の計算量で実現したい.

符号理論が解決しようとしている問題の難しさ

エラーレートを最小にする推定は MAP 推定 $\hat{x}^{\mathrm{MAP}}(\underline{y})$ である.

$$
\begin{aligned}
\hat{x}_j^{\mathrm{MAP}}(\underline{y}) :&= \operatorname*{argmax}_{x_j \in \{0,1\}} p_{X_j|\underline{Y}}(x_j|\underline{y}) \\
&= \operatorname*{argmax}_{x_j \in \{0,1\}} \sum_{\sim x_j} p_{\underline{X}|\underline{Y}}(\underline{x}|\underline{y}) \qquad \sim x_j \text{は } x_j \text{ 以外の } \underline{x} \\
&= \operatorname*{argmax}_{x_j \in \{0,1\}} \sum_{\sim x_j} p_{\underline{Y}|\underline{X}}(\underline{y}|\underline{x}) p_{\underline{X}}(\underline{x}) / p_{\underline{Y}}(\underline{y}) \\
&= \operatorname*{argmax}_{x_j \in \{0,1\}} \sum_{\sim x_j} p_{\underline{Y}|\underline{X}}(\underline{y}|\underline{x}) \mathbb{1}[\underline{x} \in C(H)] / \#C(H)
\end{aligned}
$$

$\mathbb{1}[\underline{x} \in C(H)] \neq 0$ な $\underline{x} \in C$ の数は n の指数オーダなので，素朴な $\hat{x}_j^{\mathrm{MAP}}(\underline{y})$ の計算は困難である.

現代符号理論では，解くべき問題の拘束条件を利用者が設計できる

前のスライドの最後の式 MAP 推定 $\hat{x}^{\mathrm{MAP}}(\underline{y})$ をルーピー BP で解けるように行列 H を設計し，符号理論の目的を達成することに成功している．

$p_{\underline{Y}|\underline{X}}(\underline{y}|\underline{x})$ と $\mathbb{1}[\underline{x} \in C(H)]$ は次のように因子分解できる

$$p_{\underline{Y}|\underline{X}}(\underline{y}|\underline{x}) = \prod_{i=1}^{n} p_{Y_j|X_j}(y_j|x_j)$$

$$\mathbb{1}[\underline{x} \in C(H)] = \prod_{i=1}^{m} \mathbb{1}\Big[\sum_{j:H_{i,j}=1} x_j = 0\Big]$$

- ルーピー BP の信頼性
  グラフィカルモデルの局所木性により担保される
- ルーピー BP の低計算量
  スパースに因子分解できる (各因子の引数が少ない) ことで担保される

ルーピー BP に適した行列 H の設計方法

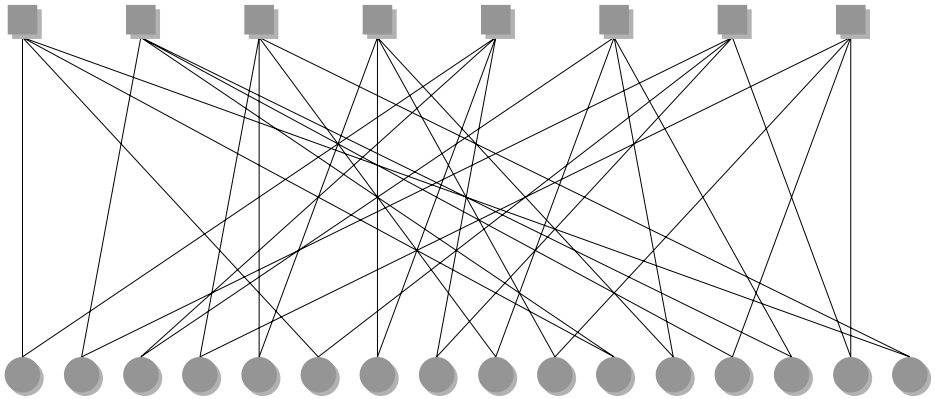
Low-Density Parity-Check 符号

列重み l，行重み r である列数 n の行列 H

例．$l = 2, r = 4, n = 20$

$$H = \begin{pmatrix}
\mathbf{1} & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & \mathbf{1} \\
0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & \mathbf{1} & \mathbf{1} & 0 & 0 \\
0 & 0 & 0 & \mathbf{1} & \mathbf{1} & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} \\
0 & 0 & 0 & 0 & \mathbf{1} & 0 & \mathbf{1} & 0 & 0 & \mathbf{1} & 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\
\mathbf{1} & 0 & \mathbf{1} & 0 & 0 & 0 & \mathbf{1} & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & \mathbf{1} & 0 & \mathbf{1} & 0 & 0 \\
0 & \mathbf{1} & 0 & 0 & 0 & \mathbf{1} & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\
0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & \mathbf{1} & 0 & \mathbf{1} & 0
\end{pmatrix}, C = \{\mathrm{x} \in \{0,1\}^n \mid H\boldsymbol{x} = \mathbf{0}\}$$

グラフィカルモデル



$$\mathbb{1}[\underline{x} \in C(H)]$$
$$= \prod_{i=1}^{m} \mathbb{1}\Big[\sum_{j:H_{i,j}=1} x_j = 0\Big]$$

消失通信モデルでルーピー BP の性能を測ろう

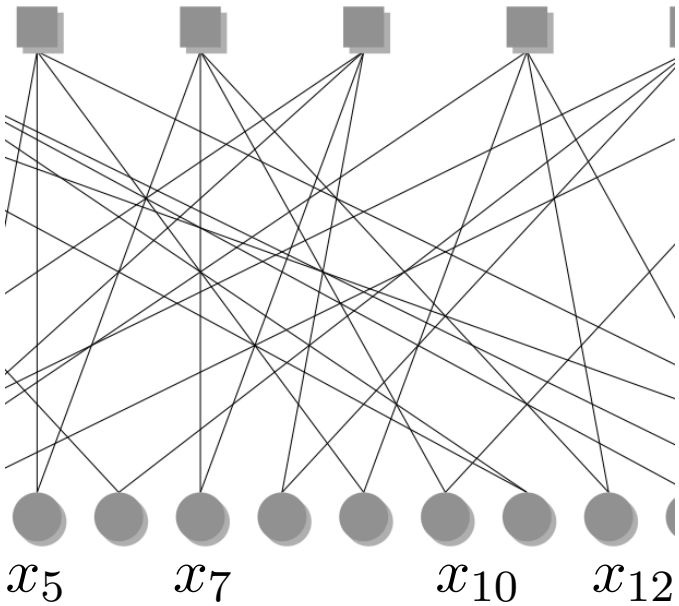
消失確率 ϵ の消失通信路 BEC(ϵ)

$$Y_j = \begin{cases} ? & \text{with probability } \epsilon \\ X_j & \text{with probability } 1-\epsilon \end{cases}$$

1. 1 次元のパラメータで表されるシンプルな方程式を解くことで，消失通信路における復号誤り率を評価できる.
2. 消失通信路の評価に用いた 1 次元のパラメータを確率分布に拡張することで，容易に他の通信路における復号誤り率を評価することができる.
3. 消失通信路でシャノン限界に達成する符号の構成法を開発すれば，一般の無記憶通信路でも同様な構成法でシャノン限界に接近する符号を構成できる.

ループ BP の動作のもよう

$$x_5 + x_7 + x_{10} + x_{12} = 0 \quad (*)$$



ループ BP の途中で $x_5 =?, x_7 = 1, x_{10} = 1, x_{12} = 0$ とすると，式 $(*)$ から，$x_5 = 1$ であることが分かる．

- ？がなくなるか，決定できる？がなくなるまで繰り返し更新する
- 辺の数に比例する計算量 $O(n)$ で，訂正を完了できる．

(l, r) 符号の BEC(ϵ) におけるルーピー BP の訂正性能

次数 r である拘束条件ノードのメッセージ消失確率

$$q^{(t)} = 1 - (1 - p^{(t-1)})^{r-1}$$

次数 l であるビットノードのメッセージ消失確率

$$p^{(t)} = \epsilon (q^{(t)})^{l-1}$$

まとめると,

$$\begin{aligned} p^{(t)} &= \epsilon (1 - (1 - p^{(t-1)})^{r-1})^{l-1} \\ &= f(g(p^{(t-1)}); \epsilon) \end{aligned}$$

BP 閾値：(l, r) 符号の性能を測る指標

$$\epsilon^{\mathrm{BP}}(l, r) = \sup\{\epsilon \mid \lim_{t \to \infty} p^{(t)} = 0\}$$

ルーピー BP で訂正可能な最大ノイズレベルを表している.

非空間結合システム (l, r) 符号の性能解析 (つづき)

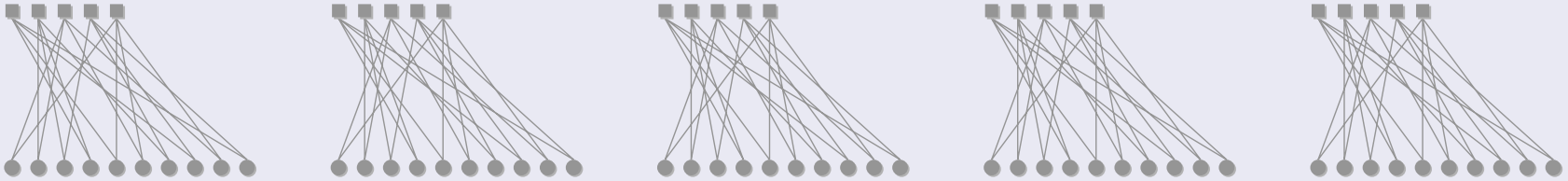
$(\mathtt{l}, \mathtt{r})$	$\epsilon^{\mathrm{BP}}(l, r)$	$\epsilon^{\mathrm{MAP}}(l, r)$
(2,4)	0.3333	0.33333
(3,6)	0.4294	0.48815
(4,8)	0.3834	0.49774
(5,10)	0.3416	0.49949
(6,12)	0.3075	0.49988
(7,14)	0.2797	0.49997

非結合システム (l, r) 符号の MAP 閾値 $\epsilon^{\mathrm{MAP}}(l, r)$

$$\epsilon^{\mathrm{MAP}}(l, r) = \sup\left\{\epsilon \mid \frac{1}{n}\sum_{i=1}^{n} \Pr(X_i \neq \hat{x}_i^{\mathrm{MAP}}(\boldsymbol{Y}(\boldsymbol{\epsilon}))) = 0\right\}$$

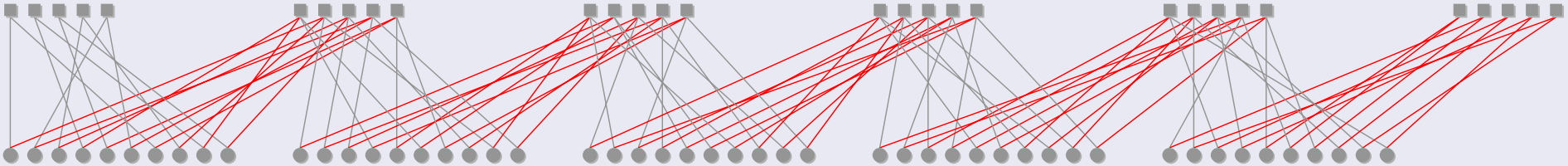
$\underline{x} = (x_1, \ldots, x_n)$: (l, r) 符号の符号語

非空間結合システム (l, r) 符号の性能解析



$p_1^{(t)} = f(g(p_1^{(t-1)}); \epsilon)$ $\cdots\cdots\cdots\cdots$ $p_5^{(t)} = f(g(p_5^{(t-1)}); \epsilon)$

空間結合システム (l, r, L, w) 符号の性能解析



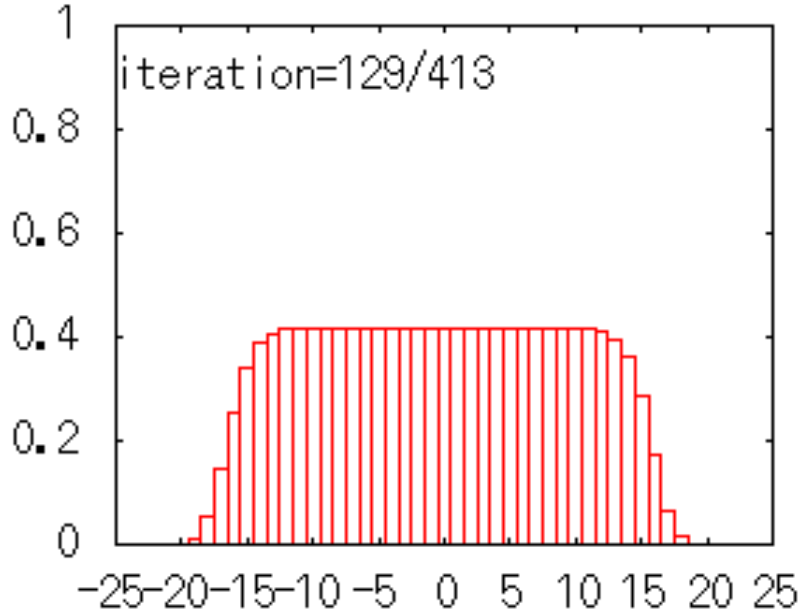
$$p_i^{(t+1)} = \frac{1}{w} \sum_{k=0}^{w-1} f\left(\frac{1}{w} \sum_{j=0}^{w-1} g(p_{i+j-k}^{(t)}); \epsilon_{i-k} \right)$$

$$i = 1, 2, \ldots, L$$

w は結合幅: w 離れたシステムまで結合する (例. $w = 2$)

空間結合符号システムの訂正過程デモのページ

`reg36_eps048_animation.gif` を開く



$l = 3, r = 6, L = 51, w = 2, \epsilon = 0.48$

- なぜ端から訂正されていくのかを説明する
- なぜ中から訂正されていかないのかを説明する
- スライディング復号法を説明する

結合システムの訂正性能はどこまで向上しているのか？

結合システム (l, r, L, w) 符号の BP 閾値 $\epsilon^{\mathrm{BP}}(l, r, L, w)$

$$\epsilon^{\mathrm{BP}}(l, r, L, w) = \sup\left\{\epsilon \mid \lim_{t\to\infty} p_i^{(t)} = 0, i = 1, \ldots, L\right\}$$

結合システムの BP 閾値は非結合システムの MAP 閾値と一致する

[2010 Kudekar et al., 2011 Takeuchi et al., 2012 Yedla et al.]

$$\lim_{w\to\infty} \lim_{L\to\infty} \epsilon^{\mathrm{BP}}(l, r, L, w) = \epsilon^{\mathrm{MAP}}(l, r)$$

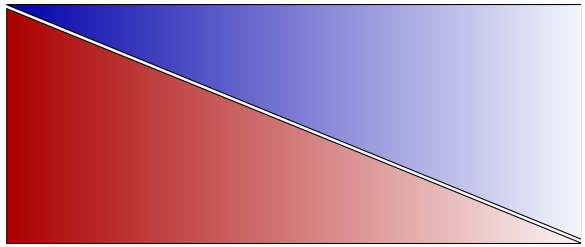
$(\mathtt{l}, \mathtt{r})$	$\epsilon^{\mathrm{BP}}(l, r)$	$\epsilon^{\mathrm{MAP}}(l, r)$
(2,4)	0.3333	0.33333
(3,6)	0.4294	0.48815
(4,8)	0.3834	0.49774
(5,10)	0.3416	0.49949
(6,12)	0.3075	0.49988
(7,14)	0.2797	0.49997

空間結合符号の定義 (範囲)

このように，BP 復号の連鎖反応が端から起こり結合前よりも強いレベルの誤りを訂正することができる事を空間結合効果と呼び，空間結合効果が起こることを狙って設計された符号を空間結合符号と呼ぶ.

Felström-Zigangirov の空間結合符号の構成法

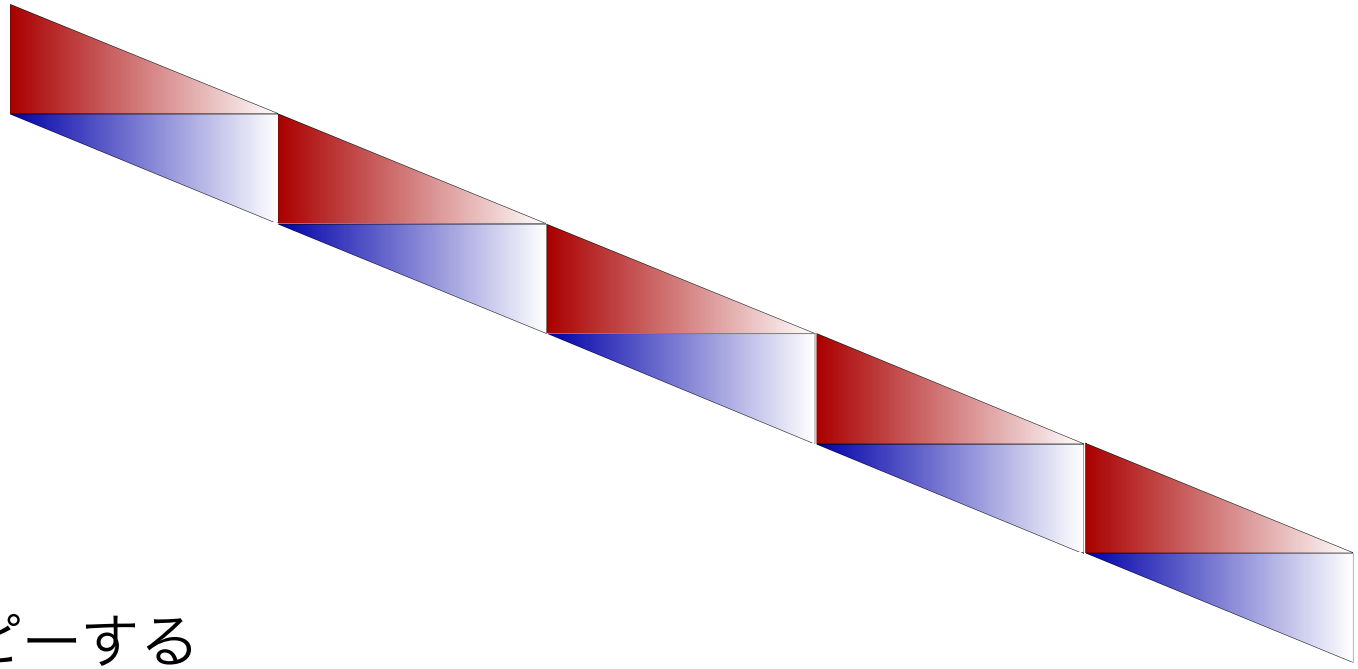
$H =$

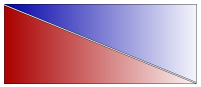


1. 行列 H を斜めに切る

Felström-Zigangirov の空間結合符号の構成法 (続き)

$H' =$



2.  を L 個コピーする

3. 対角状にコピー行列を並べる.

4. 疎な帯行列であるパリティ検査行列 H' が得られる

空間結合符号の万能性

- 空間結合符号より前に開発されていた誤り訂正符号は通信路に依存した設計が必要であった
- 空間結合符号は，任意の無記憶通信路だけでなく有記憶通信路に対してもその理論限界の訂正性能を達成することが知られている

$$y_j = \sum_{i \geq 0} a_i x_{j-i} + \xi_j$$

- さらに，任意の通信路容量を有する任意の通信路に対して，レートレス符号器と復号器の単一のペアを用いてその理論限界性能を達成することが実験から予想されている

$$(x_1, x_2, \ldots, x_{10^5})$$
$$(x_1, x_2, \ldots, x_{10^5}, \ldots, x_{2\times 10^5})$$
$$(x_1, x_2, \ldots, x_{10^5}, \ldots, x_{2\times 10^5}, \ldots, x_{3\times 10^5})$$

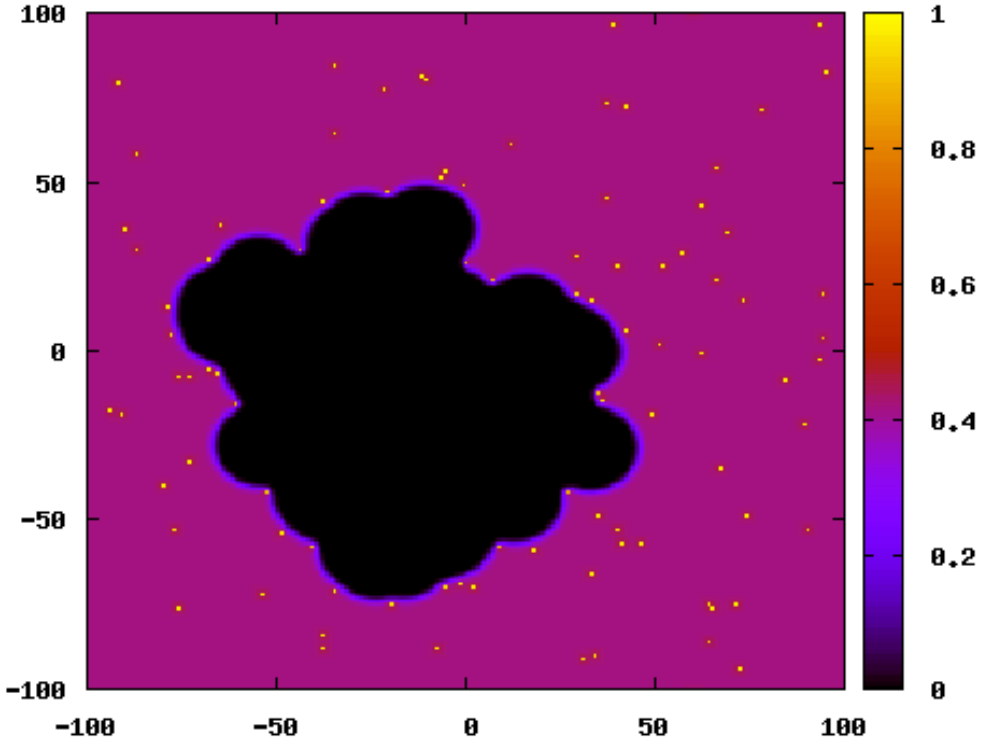
どこで通信をやめてもその系列の符号化率でシャノン限界を達成することができる.

バースト誤りに対する弱さ

- 空間結合符号の訂正過程はドミノ倒しのような連鎖反応である
- 一時的にノイズレベルが高くなり訂正の連鎖反応を止めてしまうことがある
- インターリーバを使うとスライディング窓法が使えなくなってしまう

- 連鎖反応の迂回路を作って連鎖反応が止まらないようにすれば良い

バースト誤り訂正のデモ



- なぜ中心から復号が始まるのかを説明する

有限計算量でシャノン限界達成へ

- シャノン限界を達成するために，MAP 閾値が良いものを探したい
- BP 閾値< MAP 閾値であるから，BP 閾値がよい符号を空間結合すればいいんじゃないか？
- (l, r) 符号の MAP 閾値をシャノン限界まで到達させるためには $l \to \infty, r \to \infty$ が必要だが，l, r は訂正に必要な情報ビットあたりの計算量に比例するので大きくしたくない
- MacKay-Neal 符号は 2 つの疎行列 H と G を用いて定義される符号

$$C_{\mathrm{MN}}(H, G) = \{y \in \{0,1\}^n | Gx = Hy, x \in \{0,1\}^k\}$$

  である．MacKay-Neal 符号はルーピー BP 復号では，反復復号の初期状態でアルゴリズムの不動点にいるので，結合しないと全く役に立たない．情報をドーピングして符号化率を落とすことで，動作するように修正して使っていた．
- MacKay-Neal 符号の MAP 閾値はシャノン限界と一致していて空間結合 MacKay-Neal 符号はルーピー BP により情報ビットあたり有限の計算量でシャノン限界を達成する事ができる．

まとめ

- 空間結合符号の定義，性能評価，万能性，レートレス符号への応用を紹介した．
- ループー BP がうまく動かないシステムに対する適用法を紹介した．
- 紹介しなかった応用例として，圧縮センシング，k-SAT，CDMA マルチユーザ検出，量子通信，ニューラル連想記憶などへの応用がなされている