expander 符号に関する文献調査

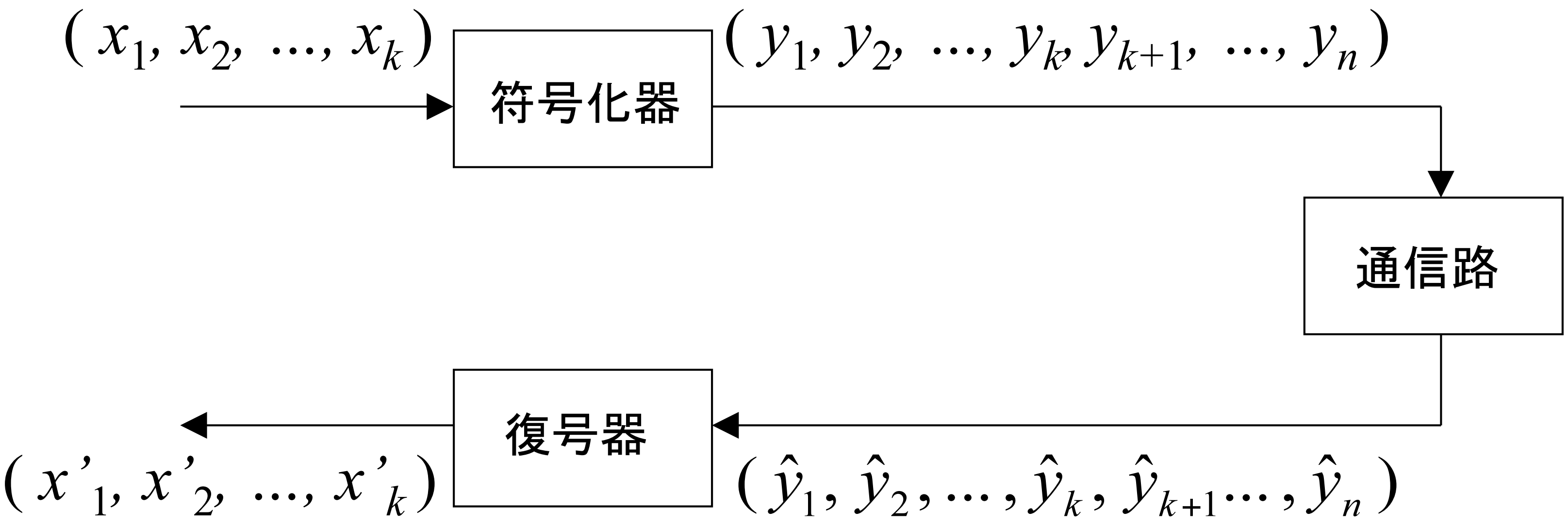
安永憲司

大阪大学 大学院情報科学研究科

2005年12月7日　大阪市立大学文化交流センター

誤り訂正符号

- 送信したい情報系列に冗長性を持たせて通信路での誤りを訂正する技術



2元 (n, k) 線形符号 C

- n : 符号長, k : 情報長, $R = k/n$: レート

- 例. (6, 2) 符号

情報系列		符号語
(0,0)	$\rightarrow$	(0,0,0,0,0,0)
(0,1)	$\rightarrow$	(0,0,0,1,1,1)
(1,0)	$\rightarrow$	(1,1,1,0,0,0)
(1,1)	$\rightarrow$	(1,1,1,1,1,1)

Shannonの通信路符号化定理

- Shannon (1948)
  - 通信路に対し、通信路容量 $C_{channel}$ を定義
  - すべての $R < C_{channel}$ に対し、任意に小さい復号誤り率を達成する符号が存在することを証明
  - $R < C_{channel}$ のとき、ランダム符号に最尤復号を用いた場合、復号誤り率は符号長に対し指数的に減少

$$P_{error} = 2^{-nE(R)}$$

誤り指数

$$E(R) > 0$$

符号の漸近性

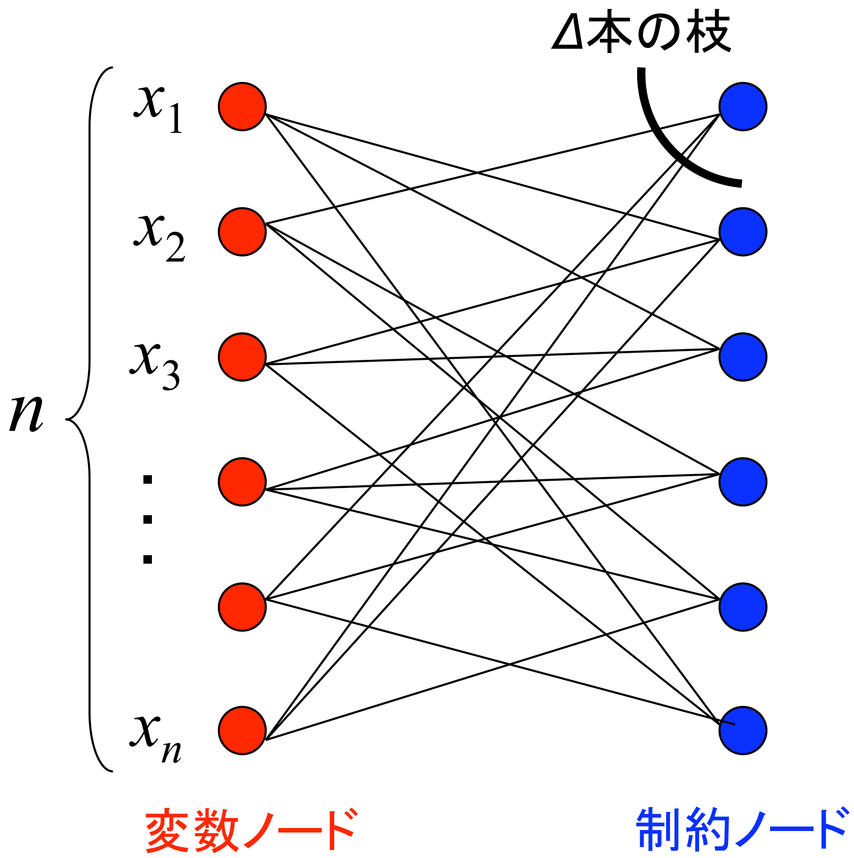
- (n, k) 符号の性能評価のパラメータ
  - R: レート（$= k/n$ ）
  - d: 最小距離（ $=$ 符号語間の距離の最小値）
    $\delta = d/n$: 相対最小距離
- 漸近的に良い符号のクラス
  - $n \to \infty$ としたとき $R, \delta \to 0$ とならない符号のクラス

符号理論の問題

- すべての $R < C_{channel}$ に対して、誤り指数が正の符号化・復号法
- 漸近的に良い符号
- 計算量が小さい符号化・復号法

上記3つを満たす符号として expander 符号が注目されている

2部グラフに基づく符号, Tanner (1981)



- グラフ G:
  n 個の変数ノードと次数Δの制約ノードをもつ2部グラフ
- 構成符号 C_0:
  長さΔの符号
- Tanner 符号 $C(G, C_0)$:
  $\{0,1\}^n$ を変数ノード列 $(x_1, x_2, \ldots x_n)$ に割り当てたとき、各制約ノード v_i に隣接する変数ノード列 $(x_{a(i,1)}, x_{a(i,2)}, \ldots, x_{a(i,\Delta)})$ が C_0 の符号語であるような $(x_1, x_2, \ldots, x_n)$ の集合からなる符号

C_0 が線形符号ならば C も線形符号

expander 符号

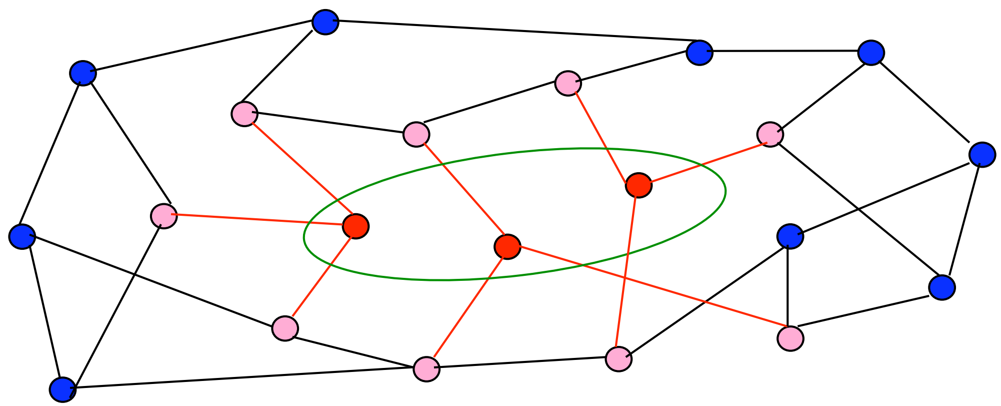
- Sipser & Spielman (1996) が expander グラフに基づく符号として提案
- Zémor (2001), Barg & Zémor (2002) 以降、グラフの expansion 性を性能分析に利用する符号の総称

expander グラフ

- 離散数学や計算機科学などで、1970年代頃から研究
  - ネットワーク設計、暗号、擬似乱数などの幅広い応用

- 直観的には、小さな頂点部分集合が大きな隣接頂点集合を持つグラフ

- グラフ (V, E) の expansion 性:
  m 個以下の頂点集合はすべて、βの係数で拡大
  ⇒ すべての部分集合 $S \subset V$ に対して,
  $$|S| \leq m \;\Rightarrow\; |\{ y: (x, y) \in E \text{ for } x \in S \}| > \beta |S|$$



Sipser & Spielman, *IEEE Tans. IT* (1996)

- expander 符号 C の最小距離 D の下界式を導出

  $D \geq N\delta^2 (1 - \varepsilon)$,　N: C の符号長, δ: C_0 の相対距離

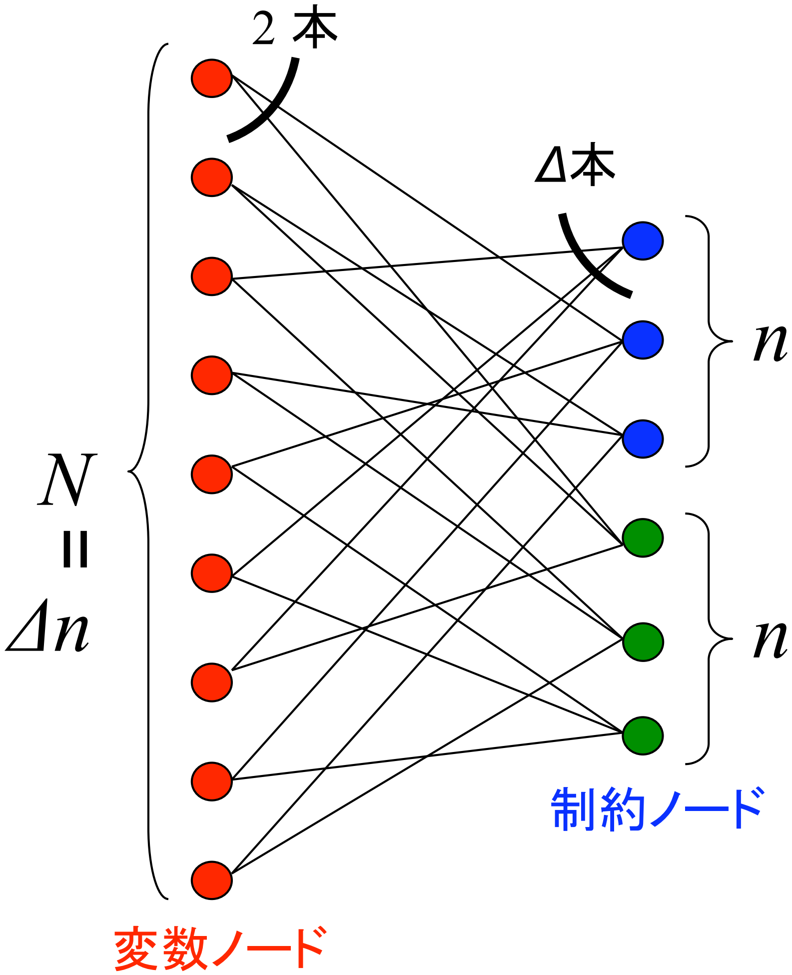
  - ε は C_0 とグラフの構造（expansion 性）に依存
  - グラフの expansion 性が高いと D は大きくなる

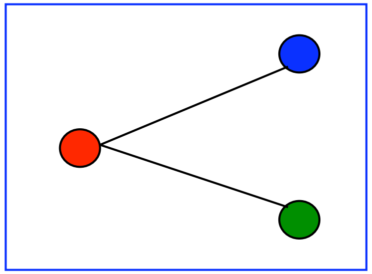
- 符号長の線形時間で復号可能な漸近的に良い符号
  - ビットフリッピングと呼ばれるシンプルな繰り返し復号
  - これまでに知られている多項式時間復号可能な漸近的に良い符号は Forney の連接符号だけ

- 繰り返し復号の性能を expander グラフと関連付け
  - $d/48$ までの誤りを訂正可能（d は最小設計距離）

Zémor, *IEEE Trans. IT* (2001)

- Sipser & Spielman 符号の特別な場合

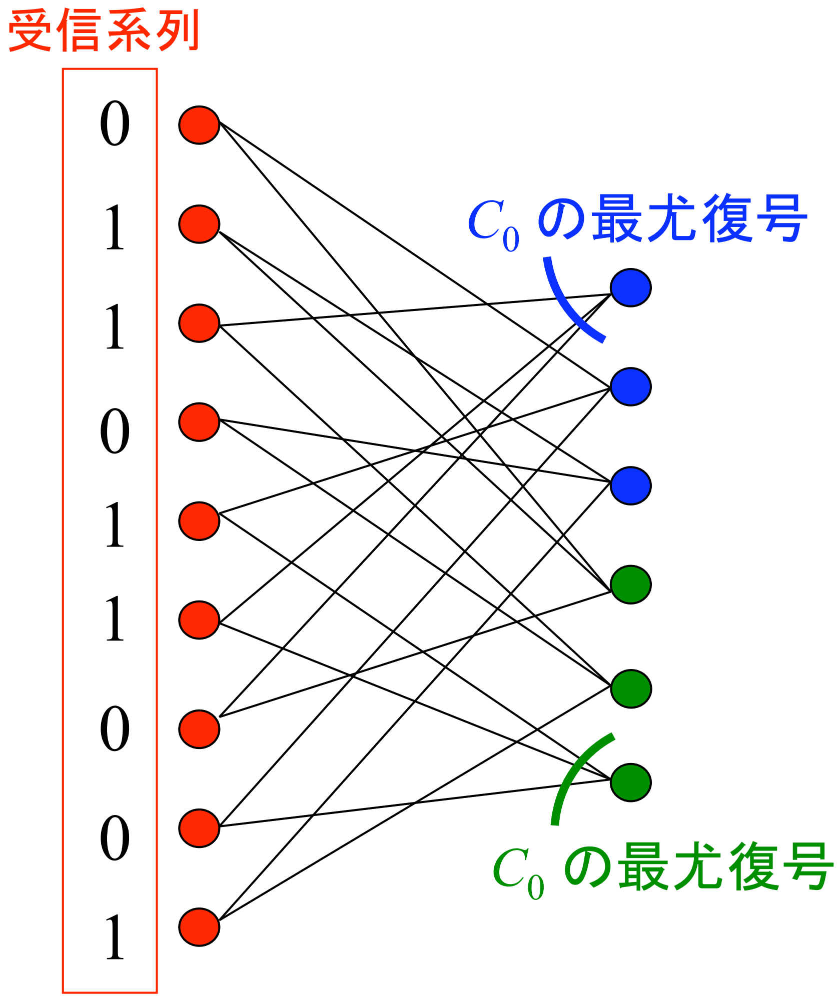




この形の枝だけ

Zémor, *IEEE Trans. IT* (2001)

- 復号は以下を繰り返す
  1. 各制約ノード ● について、隣接する変数ノード列に対して最尤復号
  2. ● について上と同様
- 各最尤復号はΔに対して指数だが、符号長 N に対しては定数
- 繰り返し回数は $O(\log N)$



Zémor, *IEEE Trans. IT* (2001)

- 復号計算量は符号長の線形
  - Sipser & Spielman (1996) と変わらない
- $d/4$ までの誤りを訂正可能（d は最小設計距離）
  - Sipser & Spielman (1996) は $d/48$ まで

- Skachek & Roth, *Proc. ITW* (2003)
  - Zémor (2001) の復号方法を修正することで、$d/2$ までの誤りを訂正可能

- Barg & Zémor, *IEEE Trans. IT* (2002)
  - Zémor (2001) を一般化した expander 符号は、すべての $R < C_{channel}$ に対して誤り指数が正
    - 誤り指数の大きさは、Forney の連接符号の方が大きい

- Barg & Zémor, *SIAM Journal on Disc. Math.* (2004)
  - R が $C_{channel}$ に近い領域での誤り指数は、expander 符号が連接符号を上回る

- Barg & Zémor, *IEEE Trans. IT* (2005)
  - 誤り指数が連接符号と同じ大きさの expander 符号が構成可能

Guruswami & Indyk, *Proc. STOC* (2002)

- 以下を満たす expander 符号の構成法を示した
  1. nearly-MDS（レートと最小距離の比が最適に近い）:
     すべてのレート R と十分小さい ε に対して、
     相対最小距離が $1-R-\varepsilon$ 以上
  2. 符号長の線形時間で符号化・復号可能
  3. 本質的に限界距離復号:
     $1-R-\varepsilon/2$ の割合の誤りを訂正可能
  - 欠点: アルファベットサイズが $1/\varepsilon$ に対して指数的に増加

- Roth & Skachek, *Proc. ISIT* (2004)
  - Guruswami & Indyk (2002) のアルファベットサイズを改良

Ashikhmin & Skachek, *Proc. ISIT* (2005)

- Roth & Skachek (2004) の expander 符号と LDPC 符号を連接符号化
- $R = (1 - \varepsilon)\, C_{channel}$ としたとき、復号計算量は符号長の線形、1 /εの多項式
  - Barg & Zémor (2002, 2005) の符号は復号計算量が符号長の線形、$1/\varepsilon^2$ の指数

まとめと考察

- expander 符号は、すべての $R < C_{channel}$ に対して誤り指数が正であり、かつ復号計算量が符号長の線形である唯一の符号

- 既存の符号化・復号テクニックと組み合わせることによって様々な性能を持つ expander 符号が提案されている

- expander グラフを導入することで様々な性能解析が可能
  - Tanner 符号における最小距離の下界
  - 繰り返し復号における訂正可能な誤り数

- 2元符号で線形時間符号化可能なものはほとんどない

参考文献

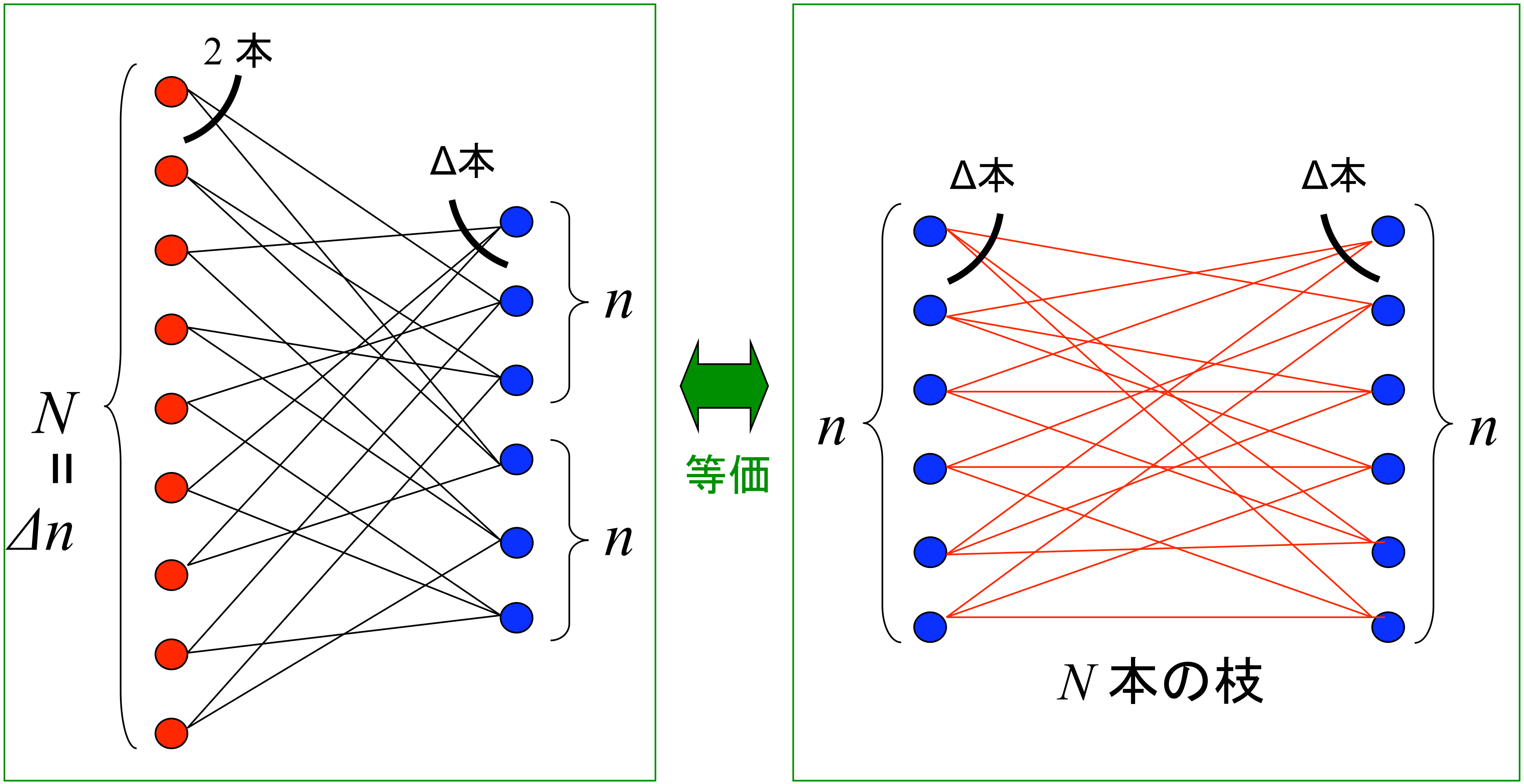
- C.E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, 1948.
- R.M. Tanner, “A recursive approach to low-complexity codes,” *IEEE Trans. IT*, 1981.
- M. Sipser and D.A. Spielman, “Expander codes,” *IEEE Trans. IT*, 1996.
- G. Zémor, “On expander codes,” *IEEE Trans. IT*, 2001.
- V. Skachek and R. Roth, “Generalized minimum distance decoding of expander codes,” *in Proc. ITW*, 2003
- A. Barg and G. Zémor, “Error exponents of expander codes,” *IEEE Trans. IT*, 2002.
- A. Barg and G. Zémor, “Error exponents of expander codes under linear-complexity decoding,” *SIAM Journal on Disc. Math*. 2004.
- A. Barg and G. Zémor, “Concatenated codes: serial and parallel,” *IEEE Trans. IT*, 2005.
- V. Guruswami and P. Indyk, “Near-optimal linear-time codes for unique decoding and new list-decodable codes over smaller alphabets,” *in Proc. STOC*, 2002.
- R. Roth and V. Skachek, “On nearly-MDS expander codes,” *in Proc. ISIT*, 2004.
- A. Ashikhmin and V. Skachek, “Decoding of expander codes at rate close to capacity,” *in Proc. ISIT*, 2005.

Feldman & Stein (2005)

- LP (Linear Programming) 復号
  - 線形計画法を利用した復号法
  - expander 符号に対して、$R < C_{channel}$ であるすべての R に対して誤り指数が正
  - ML certificate をもつ
    - ML certificate: 符号語を出力した場合は、最尤復号での符号語と同じ
    - ML certificate を持ち、通信路容量達成である、初めての多項式時間復号

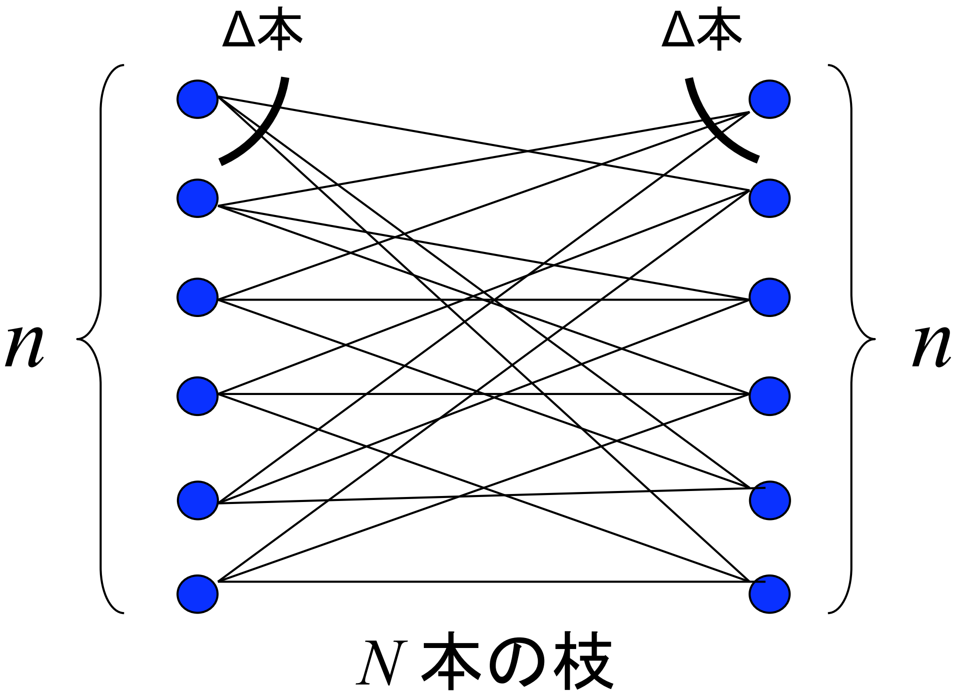
Zémor (2001)

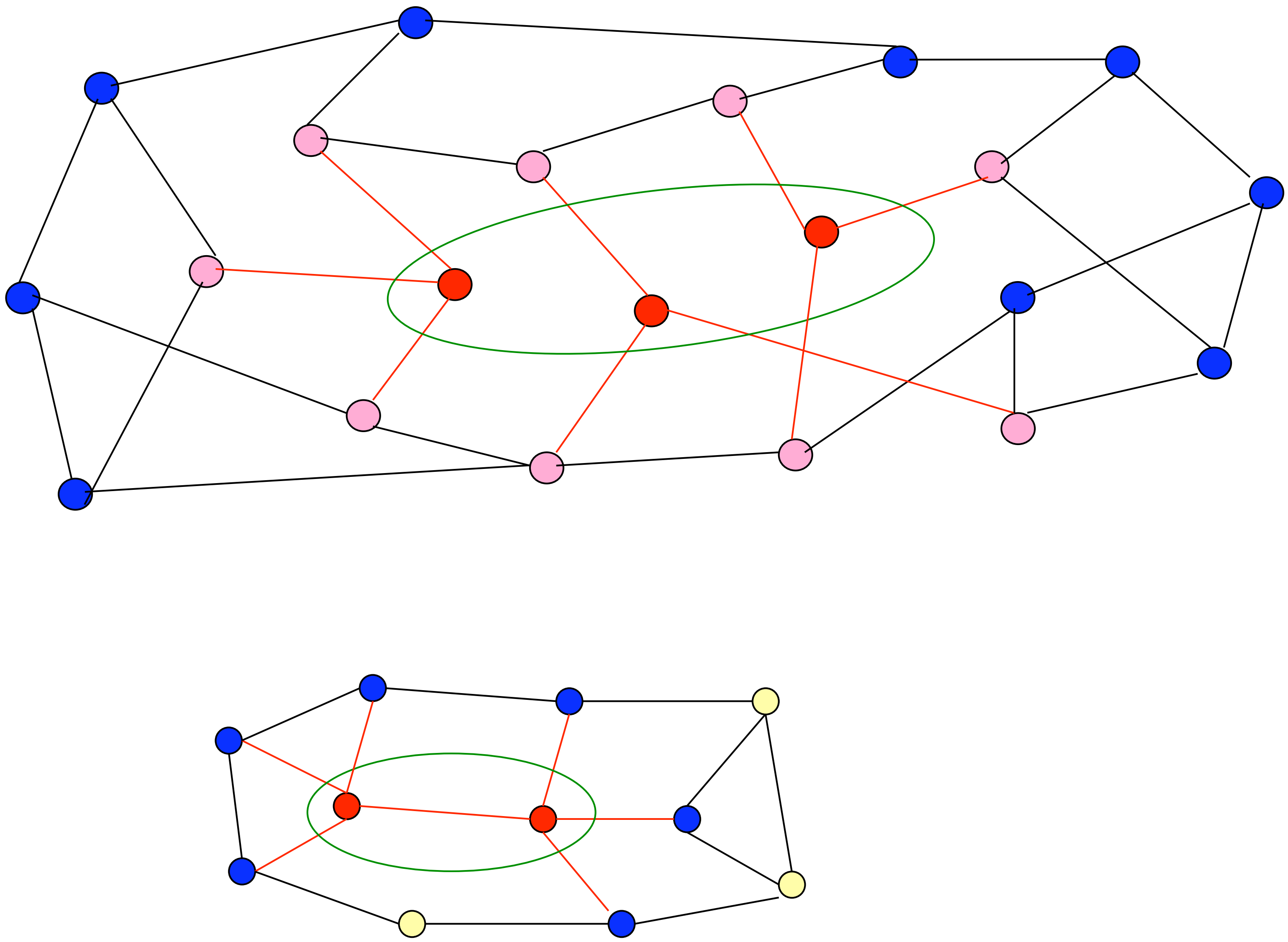
- 符号語は枝に割り当てられると考える



Zémor (2001)

- 枝に符号語を割り当て





Sipser & Spielman (1996)

- ビットフリッピング復号
  各変数ノードにおいて、反転させたほうが satisfy な制約ノードが増えるならば、そのノードを反転