エクスパンダーグラフと
誤り訂正符号

安永憲司（大阪大学）

研究集会「実験計画法と符号および関連する組合せ構造」2019 @湯田温泉
2019.11.14

エクスパンダーグラフ (Expander Graphs)

- 疎なグラフで高い連結性をもつ
  - ある程度の頂点を切り離すには,
ある程度の辺を取り除くことが必要
- ランダムグラフは高い確率でエクスパンダー
  - 擬似ランダム性をもつグラフ
- 理論計算機科学において様々な応用
  - アルゴリズム，計算複雑さ理論，符号理論

グラフの連結性 (connectivity)

- $n^{1/2} \times n^{1/2}$ のグリッドグラフ，n 頂点
  - $O(1)$ 本の辺を切ると，$O(1)$ 個の頂点が切り離される
  - $n^{1/2}$ 本の辺を切ると，半分の頂点が切り離される
    - $O(1/n^{1/2})$ 割合だけ切れば，半分が残り半分と切断される

- k 次元超立方体 (hypercube)，$n = 2^k$ 頂点
  - $1/k = 1/\log_2 n$ 割合の辺を切れば，半分の頂点が切断

- 完全グラフ
  - p 割合の頂点の切り離しに，$p(1 - p)$ 割合の辺を切る必要

このぐらい欲しい

誤り訂正符号への応用

- 定数レート符号で線形時間符号化・復号

- 符号の最小距離の増幅

- リスト復号可能符号との等価性

誤り訂正符号の用語

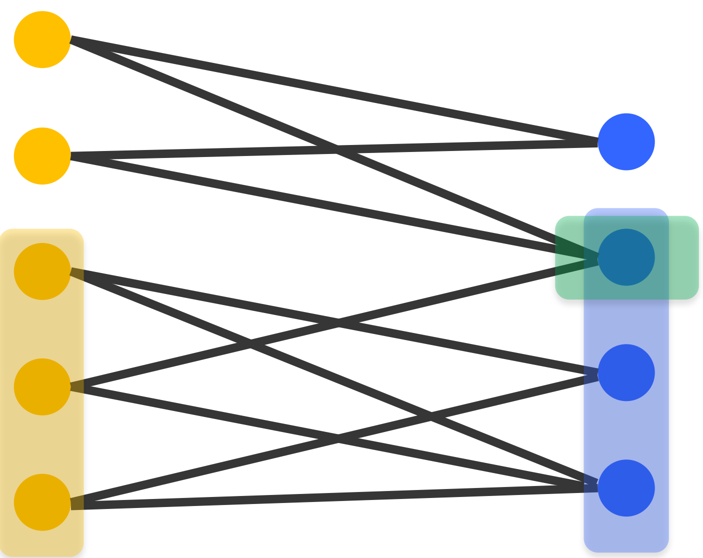
- 符号 $C \subseteq \Sigma^n$
  - $|\Sigma| = q$, $|C| = q^k$ のとき，$(n, k)_q$ 符号という
  - 符号化レート：$k/n \in [0,1]$
  - 最小（ハミング）距離： $d = \min_{u \neq v \in C} d_H(u,v)$
    - 相対最小距離 ：$d/n \in [0,1]$

- 線形符号 $C \subseteq \Sigma^n$
  - C がベクトル空間 Σ^n の部分空間のとき
  - 最小距離 $d = \min_{u \neq 0 \in C} w_H(u)$ が成立
  - 生成行列 $G \in \Sigma^{k \times n}$：$C = \{ xG \mid x \in \Sigma^k \}$ （次元 k ）
  - パリティ検査行列 $H \in \Sigma^{(n-k) \times n}$: $C = \{ z \in \Sigma^n \mid Hz = 0 \}$

グラフの用語

- グラフ $G = (V, E)$: 頂点集合 V, 辺集合 $E \in V \times V$
  - グラフが d-正則 : 各頂点の次数が d

- 二部グラフ $G = (L, R, E)$:
  左頂点集合 L, 右頂点集合 R,
  辺は L-R 間だけ存在
  - 二部グラフが d-左正則
    → 各左頂点の次数が d



- 頂点の部分集合 $S \subseteq V$ の近傍 (neighbor) $N(S) \subseteq V$
  とは，S に接続している頂点 $q \in V \setminus S$ の集合
  - S と接続している点がただ一つのとき，唯一近傍 (unique neighbor) といい，その集合は $U(S) \subseteq N(S)$

エクスパンダー符号

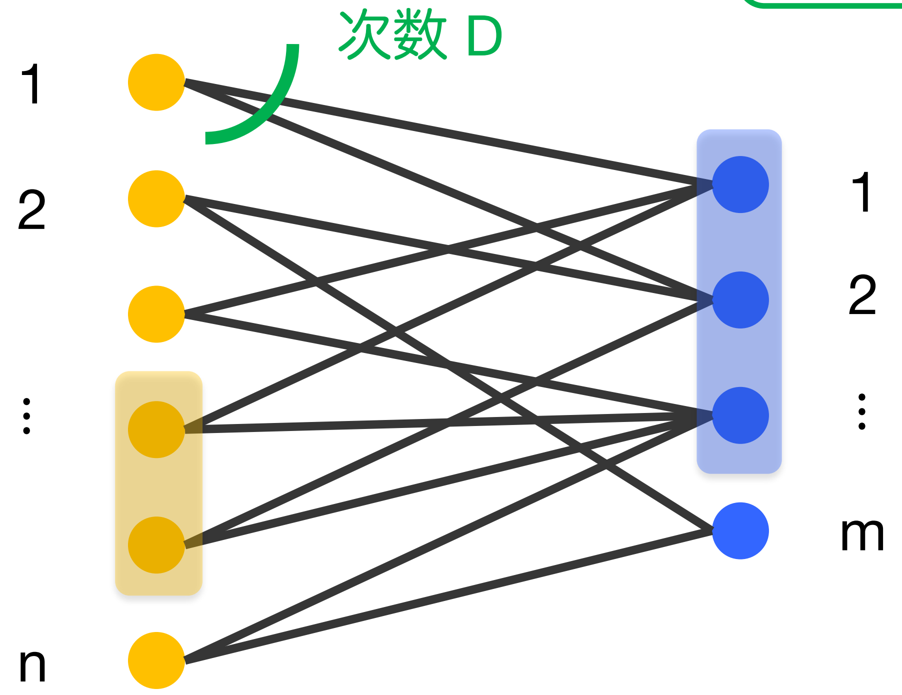
エクスパンダーグラフ

- 二部グラフ G = (L, R, E) が (n, m, D, γ, α)-エクスパンダー

  ⇔

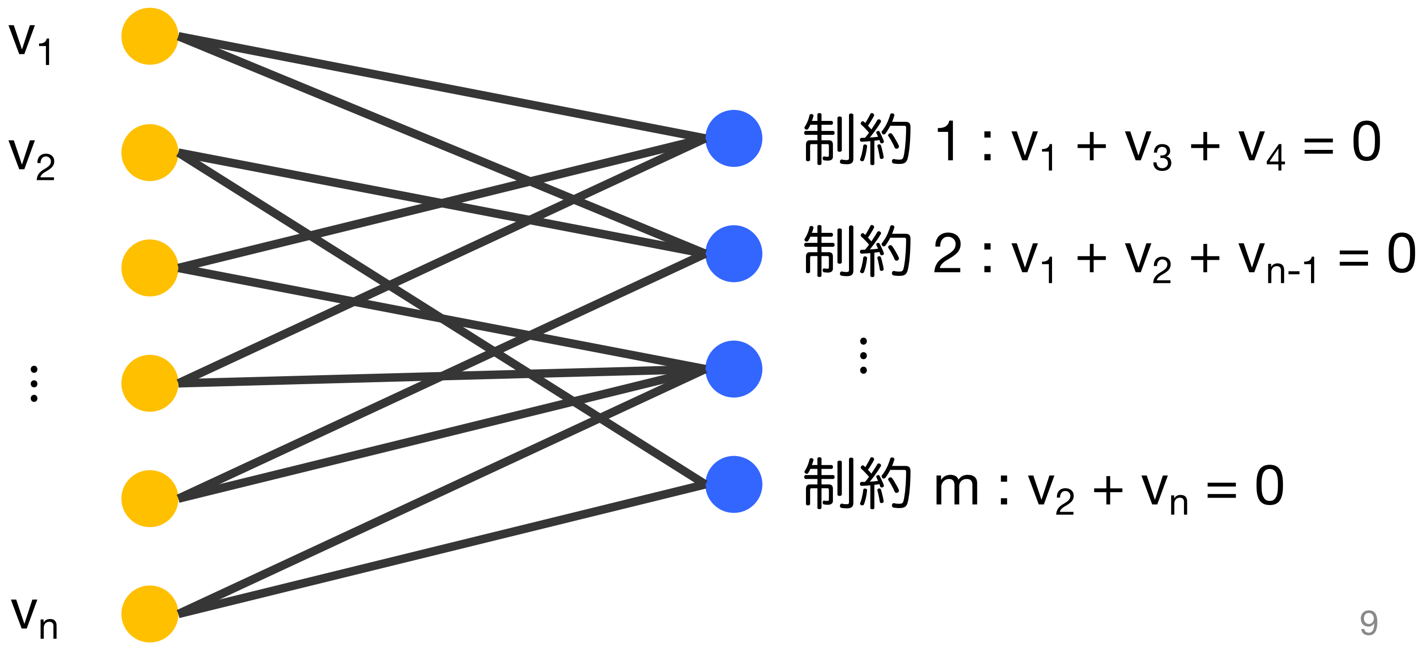
  G は D-左正則で，　$|L| = n$, $|R| = m$ であり，
  $\forall S \subseteq L$ with $|S| \leq \gamma n$, $|N(S)| \geq \alpha\, |S|$

拡大係数 α は D に近づけたい



基本構成：二部グラフによるパリティ検査行列

- 二元符号のパリティ検査行列の制約式を右頂点で表現
  - 制約式： $v_1 + v_3 + v_4 = 0$ 等
  - レート $\geq 1 - m/n$
  - グラフが疎 → Low Density Parity Check 符号
  - $\forall S \subseteq L$ with $|S| < d$, $|U(S)| > 0$ ⇒ 最小距離 $\geq d$



基本構成の性質

補題１．二部グラフ $G = (L, R, E)$ が，　$\varepsilon < 1/2$ に対し $(n, m, D, \gamma, D(1-\varepsilon))$-エクスパンダーのとき，
任意の $S \subseteq L$, $|S| \leq \gamma n$ に対し，$|U(S)| \geq D(1-2\varepsilon)\,|S|$

証明：

- S からは $D|S|$ 本の辺が出ている
- $|N(S)| \geq D(1-\varepsilon)|S|$ より，$\varepsilon D|S|$ 本以下の辺を除けばすべて唯一近傍.
- その $\varepsilon D|S|$ 本の辺により唯一近傍でなくなる頂点が $\varepsilon D|S|$ 個あるので，$U(S) \geq D(1-2\varepsilon)|S|$

（証明終）

これより，最小距離 γn 以上

基本構成の最小距離

定理２． 二部グラフ G が，$\varepsilon < 1/2$ に対し，
$(n, m, D, \gamma, D(1-\varepsilon))$-エクスパンダーのとき，Gで定義される符号 $C \subseteq \{0,1\}^n$ の最小距離は $2\gamma(1-\varepsilon)n$ 以上

証明：

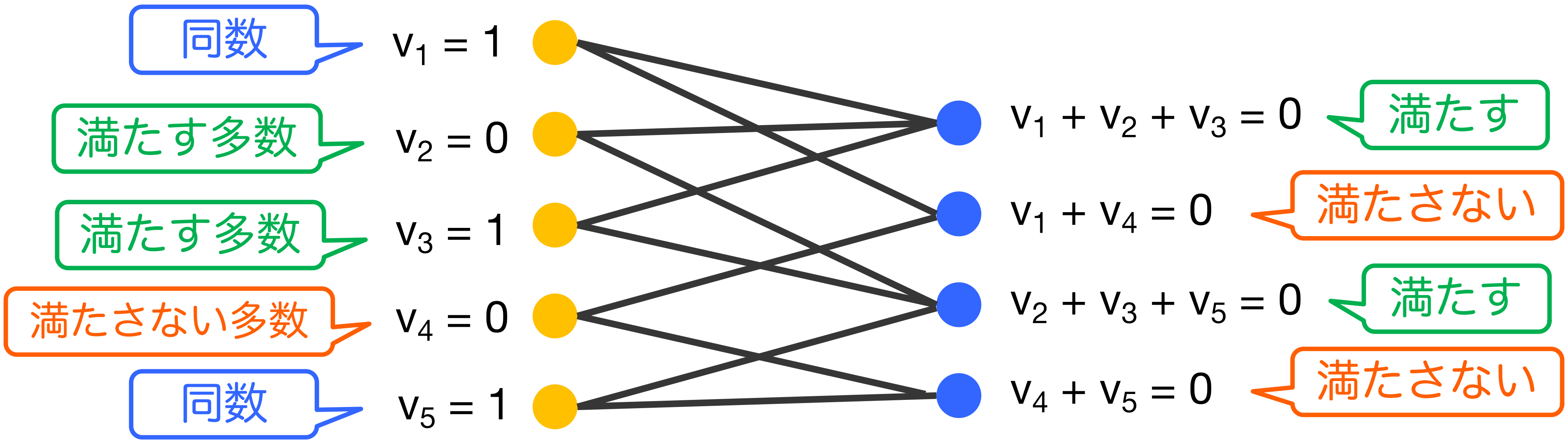
- 矛盾のため，重み $2\gamma(1-\varepsilon)n$ 未満の $v \in C$ を仮定
- v で値 1 をとる左頂点集合 $S \subseteq L$ に対し，
  $U(S)$ が空でないことを示せばよい
- $|S| \le \gamma n$ であれば補題 1 で証明終わり
- $|S| > \gamma n$ のとき，$Q \subseteq S$ として $|Q| = \gamma n$ を選ぶと，
  $|S| < 2\gamma(1-\varepsilon)n$ より，$|S \setminus Q| < \gamma(1-2\varepsilon)n$
- $S \setminus Q$ から出る辺の数は $D\gamma(1-2\varepsilon)n$ より少ない
- エクスパンダーの性質より，$|U(Q)| \ge D(1-2\varepsilon)\gamma n$ であり，
  $U(Q)$ を $N(S \setminus Q)$ でカバーできず，$U(S)$ は空でない

（証明終）

基本構成の復号法：ビット反転法

- 復号法として，$\varepsilon < 1/4$ のときに，
  重み $\gamma(1-2\varepsilon)n$ 未満の誤りの訂正方法を示す
- 受信語 $v \in \{0,1\}^n$ に対し，左頂点 $q \in L$ の値を v_q
  - 右頂点は制約を，満たす or 満たさない

アルゴリズム：左頂点 $q \in L$ において，$N(q)$ に満たさない頂点が多数である限り，v_q の値を反転



ビット反転法

補題３． 誤りの数が 1 以上 γn 以下のとき，
ある左頂点 $q \in L$ が存在し，$N(q)$ は満たさない頂点
を D/2 より多く含む（$\varepsilon < 1/4$ を仮定）

証明：

- 誤りの位置の頂点集合を T とする．
- 補題１より，$|T| \le \gamma n$ であるため，
  $|U(T)| \ge D(1 - 2\varepsilon)\,|T| > (D/2)|T|$
- $U(T)$ の各頂点は明らかに満たさない．
- 左頂点は D 正則であるため，T の頂点で近傍に
  満たさない頂点を D/2 より多く含むものが存在

（証明終）

ビット反転法

補題４．誤り数 $\gamma(1-2\varepsilon)n$ 未満の受信語に対してアルゴリズムを実行したとき，途中で誤り数が γn 以上になることはない

証明：

- 最初に，満たさない右頂点は $D(1-2\varepsilon)\gamma n$ 未満
- 左頂点の反転一回で，満たさない右頂点は１以上減る
- 途中で誤り数が γn になったとき，エクスパンダー性より，$D(1-2\varepsilon)\gamma n$ 以上の唯一近傍が存在.
- 唯一近傍の右頂点はすべて満たさないため矛盾.

（証明終）

ビット反転法

定理５． ビット反転法により，$\varepsilon < 1/4$ のとき，
誤り数 $\gamma(1 - 2\varepsilon)n$ 未満の受信語を正しく訂正できる

証明：

- 補題３より，誤り数が１以上 γn 以下のとき，
アルゴリズムは動き続ける
- 一回の反転で満たさない右頂点は１以上減るので，
m 回繰り返せば停止
- 補題４より，誤り数 $\gamma(1 - 2\varepsilon)n$ 未満の受信語に対して，途中で誤り数が γn 以上になることはない
- 符号の最小距離は $2\gamma(1 - 2\varepsilon)n > \gamma n$ であるため，
もとの符号語が出力される

（証明終）

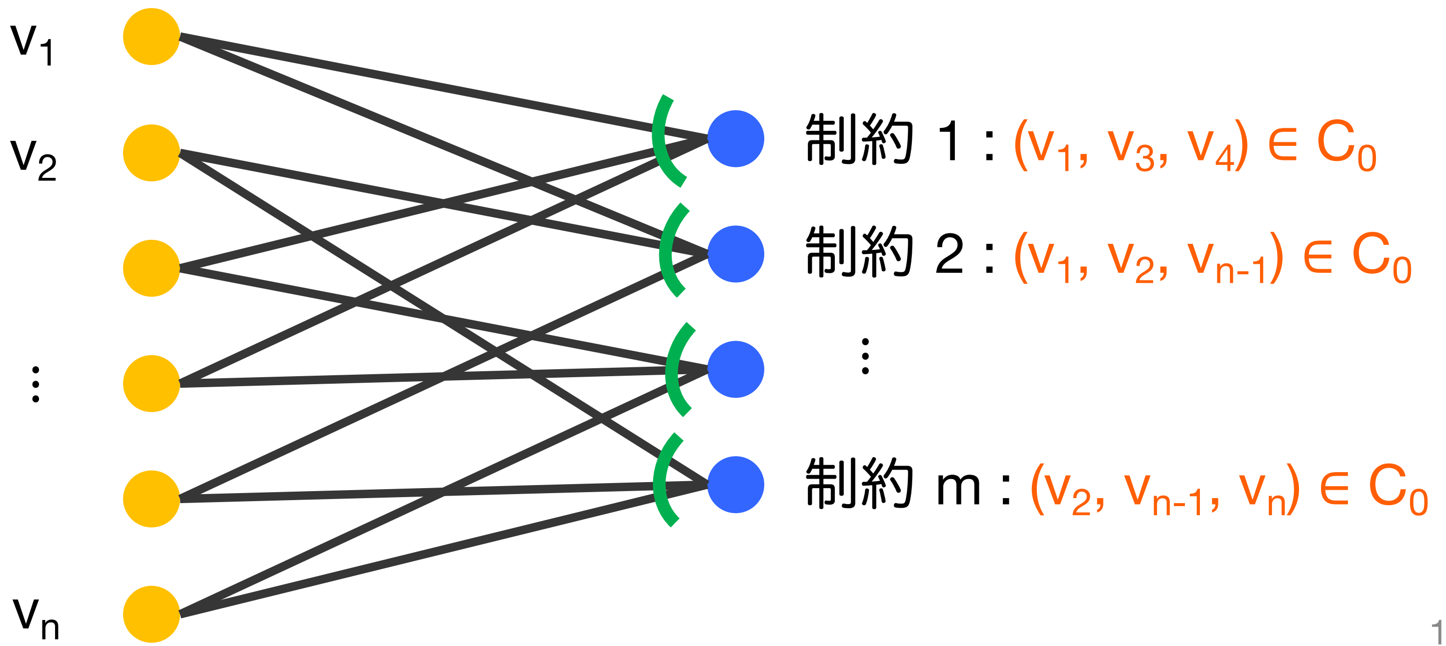
ビット反転法の実行時間

- 右頂点の最大次数を d とする
- 右頂点の「満たす or 満たさない」の計算に，O(md) 時間
- 左頂点の「満たす多数 or 満たさない多数 or 同数」の計算に，O(Dn) 時間
  - 満たさない多数の左頂点リスト Q を保持
- 各繰り返しでは,
  - Q の要素を 1 つ削除
  - 右頂点の「満たす or 満たさない」を O(D) 時間で更新
  - 左頂点のリスト Q を O(Dd) 時間で更新
- 繰り返しは m 回以下

以上より，実行時間は O(Ddm)．D, d が定数なら O(n)

一般化構成：タナーグラフによるパリティ検査

- 基本構成の制約を，二元線形符号 $C_0 \subseteq \{0,1\}^d$ の符号語であることに変更 → タナー符号
  - C_0 が線形であるため，タナー符号も線形
  - C_0 が $(d, d-1)_2$-パリティ検査符号のとき基本構成
  - タナー符号の次元は，$n - m(d - \dim(C_0))$ 以上

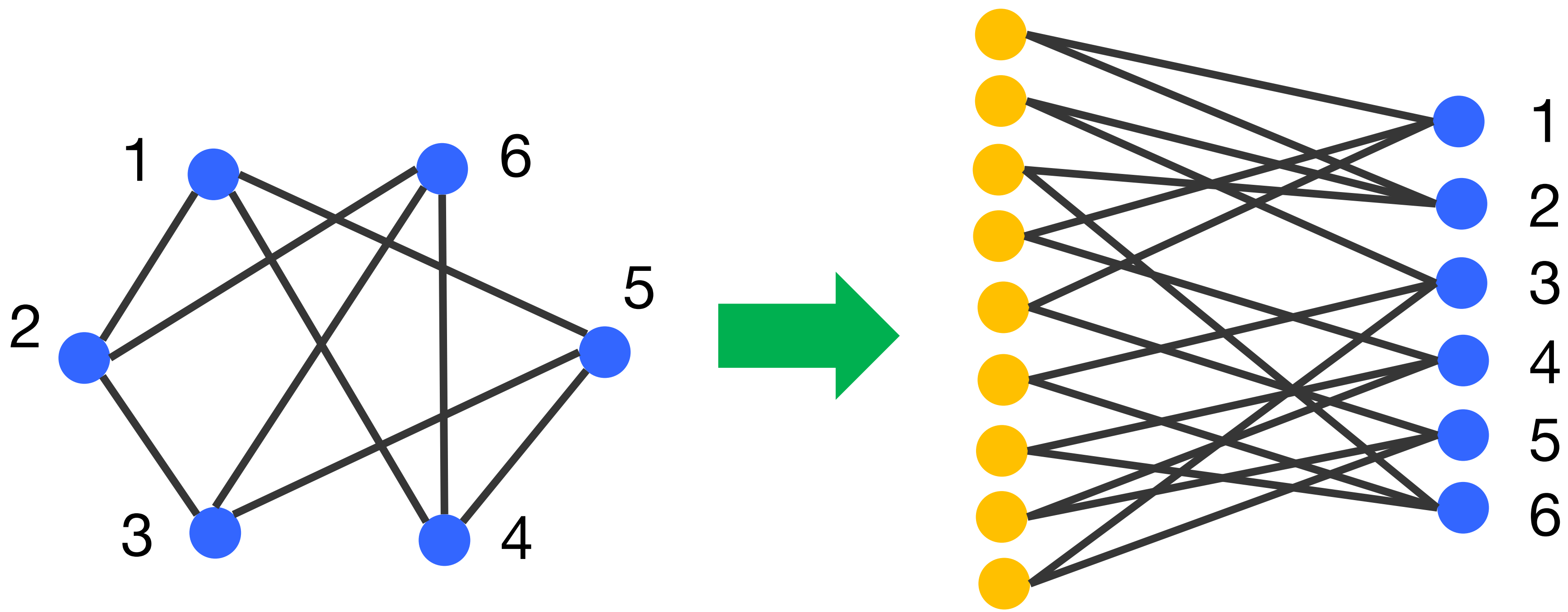


一般化構成：タナーグラフによるパリティ検査

- タナー符号のポイント
  - グラフのエクスパンダー性の要求が小さい
    - C_0 がパリティ検査符号 → 最小距離 2
      タナー符号の最小距離 ≥ γn を示すのに，$|S| \le \gamma n$
      の左頂点集合 S に対し，$|N(S)| \ge (D/2)|S|$ が必要
    - C_0 の最小距離 d_0 のとき，$|N(S)| \ge (D/d_0)|S|$ で十分

高レート符号の構成のために

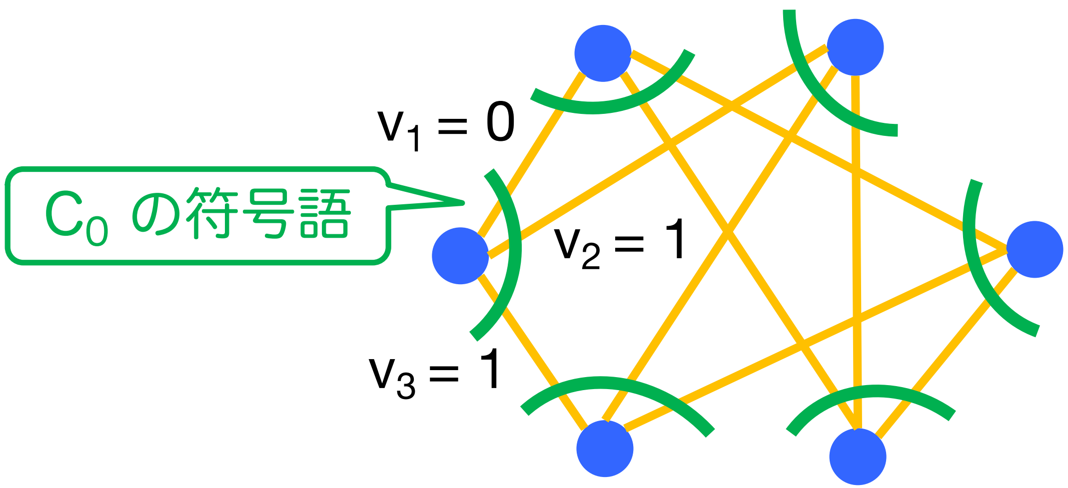
- グラフ G = (V, E) の，辺頂点接続グラフ (Edge Vertex Incidence Graph) とは，
以下の二部グラフ H_0 = (L, R, E')
  - L は E に対応，R は V に対応
  - E' は e ∈ E の端点が v ∈ V であれば辺が存在



高レート符号の構成のために

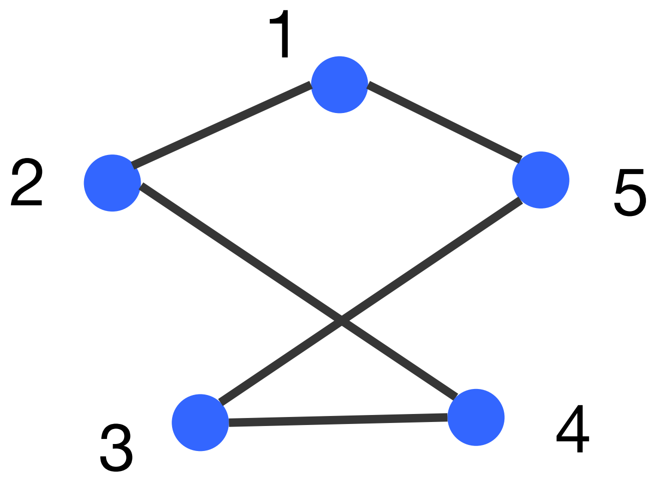
- $G = (V, E)$ を N 頂点の d-正則グラフとする
- G の辺頂点接続グラフ H_0 と線形符号 C_0 によるタナー符号を $T(G, C_0)$ と表す

- G から直接定義すると，$n = Nd/2$ 個ある各辺に値 $(v_1, \ldots, v_n) \in \{0,1\}^n$ を割り当て，各頂点の接続辺が C_0 の符号語という制約をもつ符号



グラフのスペクトル

- グラフ G = (V, E) の隣接行列 $A \in \{0,1\}^{n\times n}$, $|V| = n$

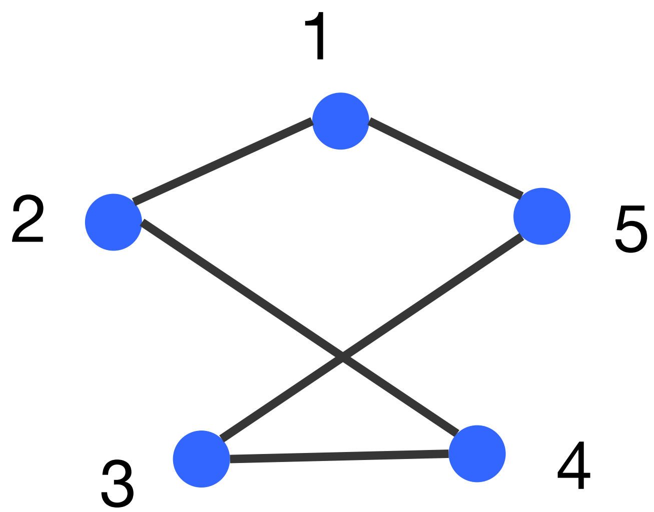


$$\begin{matrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{matrix}$$

- A の 固有値 $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ はスペクトルと呼ばれる
  - G が d-正則のとき $\lambda_1 = d$
  - d-正則 G が連結 $\Leftrightarrow$ $\lambda_2 < d$
  - d-正則 G が二部グラフ $\Leftrightarrow$ $\lambda_n = -d$

スペクトル版エクスパンダーグラフ

- グラフ G = (V, E) が (n, d, λ)-エクスパンダー
  ⇔
  G は d-正則 n 頂点グラフで，隣接行列の固有値が $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ のとき，$\lambda = \max\{ |\lambda_2|, |\lambda_n| \}$



$$\begin{matrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{matrix}$$

$\lambda_1 = 2, \lambda_2 = \lambda_3 = (5^{1/2} - 1)/2,$
$\lambda_4 = \lambda_5 = (- 5^{1/2} - 1)/2$

スペクトル版エクスパンダーグラフ

補題 6（Expander Mixing Lemma）

グラフ $G = (V, E)$ が (n, d, λ)-エクスパンダーのとき，$\forall S, T \subseteq V$ に対し

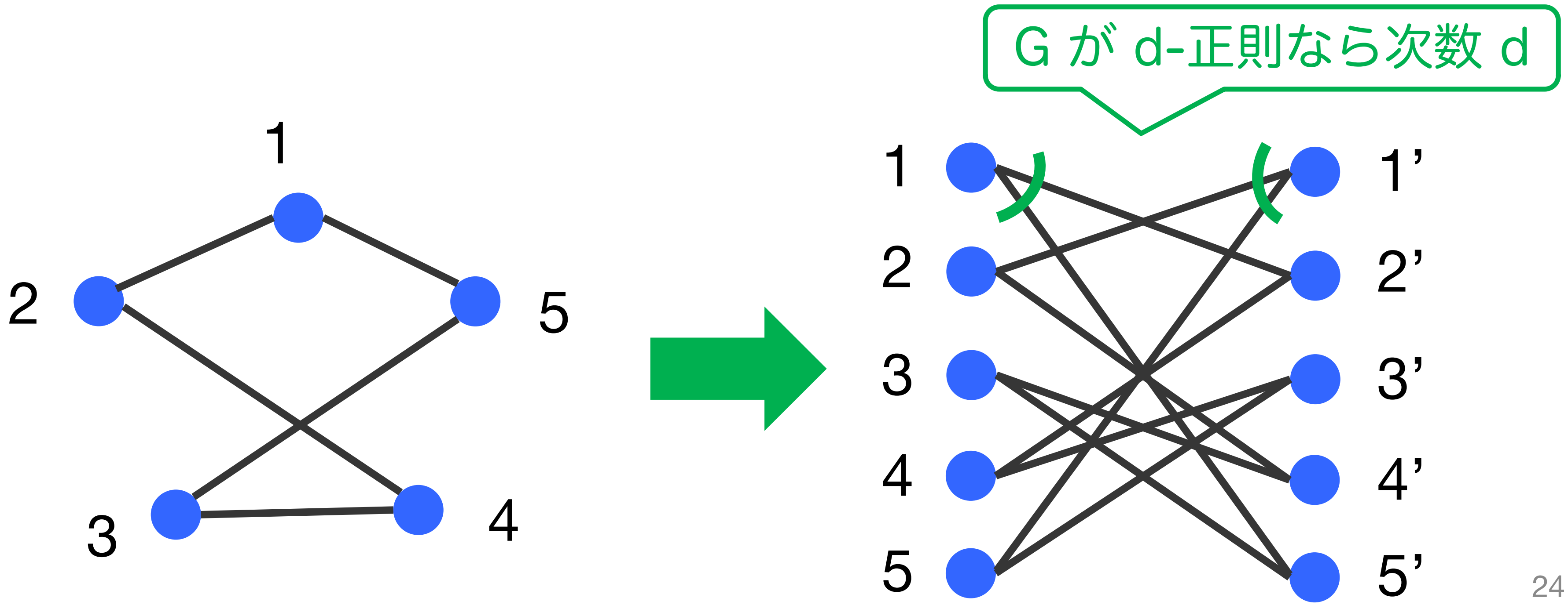
$$\left| |E(S, T)| - (d \cdot |S| \cdot |T|)/n \right| \leq \lambda (|S| \cdot |T|)^{1/2}$$

が成り立つ．ここで，$|E(S, T)|$ は S-T 間の辺の数であり，$S \cap T$ 間は二重に数える

- S からは $d \cdot |S|$ 本辺が出ており，ランダムグラフだとそれが T に入る確率は $|T|/n$
- 補題は，$|E(S, T)|$ からのズレが $\lambda (|S| \cdot |T|)^{1/2}$ で抑えられることを示している

二重被覆

- グラフ G = (V, E) の二重被覆 (double cover) とは, 以下の二部グラフ $H = (L, R, E_H)$
  - L = R = V
  - $(u, v) \in E$ のとき, $(u, v) \in E_H$, $(v, u) \in E_H$



二部グラフのスペクトル

- 二部グラフ H = (L, R, E) が，　L, R ともに d-正則で $|L| = |R| = n$ (biregular graph) のとき，　$n \times n$ 行列として，　列が L の頂点, 行が R の頂点の隣接行列のスペクトルを考える

- 二部グラフ H がグラフ G = (V, E) の二重被覆のとき，　その行列は G の隣接行列 と同じ！

補題７（二部グラフ版 Expander Mixing Lemma）

二部グラフ $H = (L, R, E_H)$ が，　(n, d, λ)-エクスパンダーグラフ G = (V, E) の二重被覆のとき，
$\forall S \subseteq L, T \subseteq R$ に対し

$$\left| |E(S, T)| - (d \cdot |S| \cdot |T|)/n \right| \leq \lambda \, (|S| \cdot |T|)^{1/2}.$$

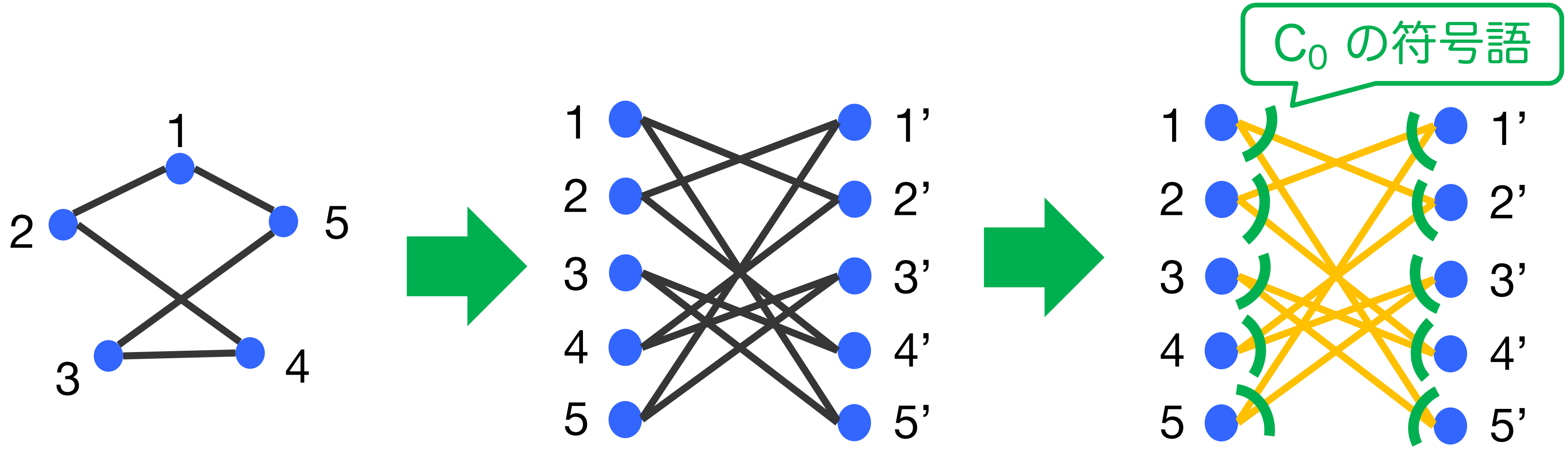
補題６と同じパラメータ

一般化構成の最小距離

定理８.

- (n, d, λ)-エクスパンダー $G = (V, E)$ の二重被覆 $H = (L, R, E_H)$
- 最小距離 $\delta_0 d$ の符号 $C_0 \subseteq \{0,1\}^d$

によるタナー符号 $T(H,C_0) \subseteq \{0,1\}^{nd}$ の最小距離は $\delta_0(\delta_0 - \lambda/d)nd$



エクスパンダーグラフ $G = (V, E)$

G の二重被覆 $H = (L, R, E_H)$

H の辺に値を割り当て，頂点の接続辺を C_0 で制約

- 証明：
  - 任意の符号語について，重みは 0 でなければ，$\delta_0(\delta_0 - \lambda/d)nd$ 以上であることを示す
  - 符号語 $v \in T(H, C_0) \subseteq \{0,1\}^{nd}$ の値を辺に割り当てたとき，値 1 の辺を $F \subseteq E_H$ とする
  - F に接続する左頂点を $S \subseteq L$, 右頂点を $T \subseteq R$
  - C_0 の最小距離が $\delta_0 d \Rightarrow |F| \geq \delta_0 d\ |S|$, $|F| \geq \delta_0 d\ |T|$

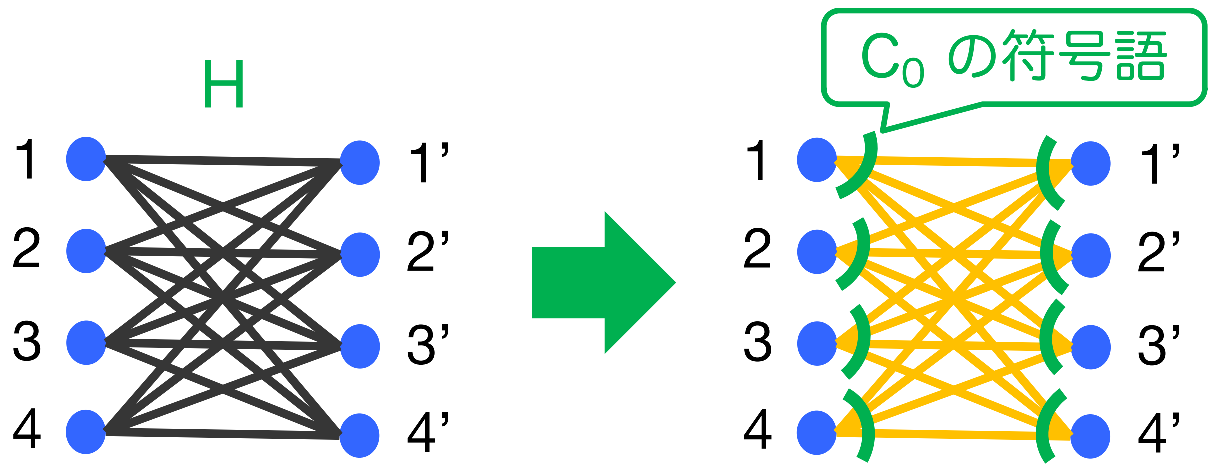
    $$\Rightarrow \quad |F| \geq \delta_0 d\ (|S|\cdot|T|)^{1/2}$$
  - 二部グラフ版 Expander Mixing Lemma より，

    $$|F| \leq |E(S, T)| \leq (d\cdot|S|\cdot|T|)/n + \lambda\ (|S|\cdot|T|)^{1/2}$$
  - 以上より，$\delta_0 d\ (|S|\cdot|T|)^{1/2} \leq (d\cdot|S|\cdot|T|)/n + \lambda\ (|S|\cdot|T|)^{1/2}$ であり，これを解くと，$(|S|\cdot|T|)^{1/2} \geq (\delta_0 - \lambda/d)n$
  - したがって，$|F| \geq \delta_0 d\ (|S|\cdot|T|)^{1/2} \geq \delta_0(\delta_0 - \lambda/d)nd$

（証明終）

補足

- H が完全二部グラフのとき，T(H, C_0) は C_0 自身とのテンソル積であり，相対最小距離は δ_0^2
  - $C_0 \subseteq \{0,1\}^d$ であり，符号長は d^2
- 定理８では，エクスパンダーグラフで $\lambda = o(d)$ のとき，相対最小距離 $\approx \delta_0^2$
  - 符号長は nd であり，大きい値をとれる



レートと距離の関係

- 符号 C_0 がレート R のとき，T(H, C_0) のレート R_T は

$$R_T = (nd - 2nd(1 - R))/(nd) = 2R - 1$$

- T(H, C_0) の相対最小距離 $\delta \approx \delta_0^2$ とし，
  C_0 を Gilbert-Varshamov 限界上の符号とし，
  $R \geq 1 - h(\delta_0)$ とすると，

$$R_T \geq 2(1 - h(\delta_0)) - 1 \approx 1 - 2h(\delta^{1/2})$$

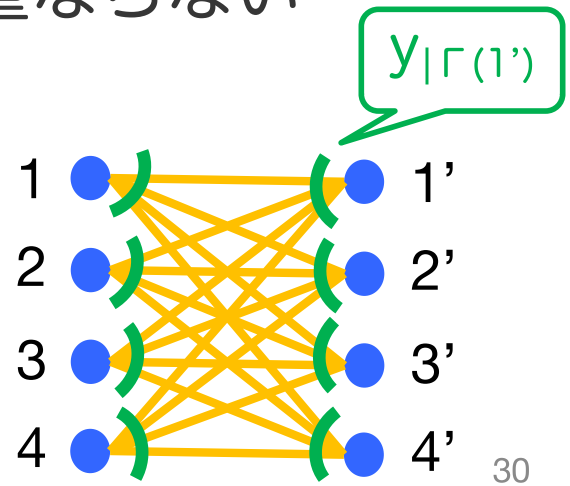
- $R_T > 0$ は，$\delta < 0.0121$ のときに限られる．

一般化構成の復号法：左右交互復号

- 重み $(1 - \varepsilon)(\delta_0/2)(\delta_0/2 - \lambda/d)nd$ 以下の誤りの訂正方法を示す
- 受信語 $y \in \{0,1\}^{|E|}$ に対し，$y_{|\Gamma(v)} \in \{0,1\}^d$ を，頂点 v 周りの辺に対応する部分系列とする
- 左復号：以下をすべての左頂点に対して並列に実行

$\forall v \in L$, $y_{|\Gamma(v)}$ を $d_H(c, y_{|\Gamma(v)})$ が最小の $c \in C_0$ に置き換え

  - C_0 は小さい符号（符号長 d）なので全数探索で実行
  - 二部グラフのため，置き換える辺は重ならない
  - 右復号も同様に定める



アルゴリズム：左復号と右復号を交互に A log n 回繰り返す

- A は十分大きな定数

左右交互復号

定理９． $\lambda/d < \delta_0/3$ のとき，左右交互復号により，誤り数 $(1 - \varepsilon)(\delta_0/2)(\delta_0/2 - \lambda/d)nd$ 以下の受信語を $A(\varepsilon)\log n$ 回の繰り返しで正しく訂正できる

証明：

- S_1 = { 最初の左復号後に，誤りの辺を含む左頂点 }
  - $v \in S_1 \Rightarrow v$ に誤り $\delta_0 d/2$ 以上
- 最初の誤り数 $E_0 \geq (\delta_0 d/2)\cdot|S_1|$ が成立．E_0 の上界より

$$|S_1| \leq (1 - \varepsilon)(\delta_0/2 - \lambda/d)n \quad \cdots \text{(a)}$$

- T_1 = { 最初の右復号後に，誤りの辺を含む右頂点 }
- 最初の右復号前の誤り数は，$|E(S_1, T_1)|$ 以下であるため，同様に，$|E(S_1, T_1)| \geq (\delta_0 d/2)\cdot|T_1|$ が成立．

（続く）

左右交互復号

（証明の続き）

- Expander Mixing Lemma より，

  $|E(S_1, T_1)| \leq (d \cdot |S_1| \cdot |T_1|)/n + \lambda \cdot (|S_1| \cdot |T_1|)^{1/2}$

- AM-GM 不等式 $(|S_1|+|T_1|)/2 \geq (|S_1| \cdot |T_1|)^{1/2}$ と (a) より，

  $|E(S_1, T_1)| \leq (1 - \varepsilon)(d\delta_0/2 - \lambda) \cdot |T_1| + \lambda \cdot (|S_1|+|T_1|)/2$

- 前ページの不等式 $|E(S_1, T_1)| \geq (\delta_0 d/2) \cdot |T_1|$ より，

  $(\delta_0 d/2) \cdot |T_1| \leq (1 - \varepsilon)(d\delta_0/2 - \lambda) \cdot |T_1| + \lambda \cdot (|S_1|+|T_1|)/2$

  であり，整理すると，

  $|T_1| \leq \lambda \cdot (\varepsilon\delta_0 d + (1 - 2\varepsilon)\lambda)^{-1} \cdot |S_1| \leq |S_1|/(1+\varepsilon).$

  ここで，$\lambda/d < \delta_0/3$ を利用.

- 左右一回の繰り返しで，誤り数が $1/(1+\varepsilon)$ 倍に減るので，$O(\log n)$ 回繰り返せば，誤りがなくなる.

（証明終）

左右交互復号の実行時間

- 左右復号は，d = O(1) であれば O(n) 時間であり，O(log n) 回繰り返すため，合計 O(n log n) 時間
- 復号を O(n) 時間にすることも可能
  - 各左右復号で対象とする頂点は，直前の復号で反転した辺と接続するものだけでよい
    - 反転していない辺は C_0 の符号語になっている
  - 復号対象の頂点集合を追っていけば，定理５の議論と同様，毎回一定数減るため，合計 O(n) 個だけ見ればよい

エクスパンダーグラフの構成

定理 10. [Alon, Roichman '94]（符号 → エクスパンダー）

非零符号語の重みが $(1/2 - \varepsilon)n$ から $(1/2 + \varepsilon)n$ の間にある二元線形符号の生成行列 $M \in \{0,1\}^{k\times n}$ を考える.
グラフ $G = (V, E)$ を，$V = \{0,1\}^k$,

$$(x, y) \in E \iff x - y \text{ が } M \text{ の列ベクトル}$$

と定めると，G は $(2^k, n, 2\varepsilon)$-エクスパンダー

- M の列による Cayley グラフ
- 頂点数 $N = 2^k$ より，次数 $O(\log k)$ のグラフを構成

定理 11. [Capalbo, Reingold, Vadhan, Wigderson '02]

$\forall \varepsilon > 0, n, m \leq n$ に対し，$(n, m, D, \gamma, D(1 - \varepsilon))$-エクスパンダーグラフが存在し，poly(n) 時間で構成可能.
ただし，$D = \Theta(\log(n/m)/\varepsilon)$, $\gamma n = \Theta(\varepsilon m/D)$

エクスパンダー符号の文献情報

- エクスパンダー符号は [Sipser & Spielman '96] の提案だが，当時はエクスパンダー性の高いグラフは知られておらず，タナー符号を利用
  - [SS96] の訂正可能な誤り割合は δ/48
- 定理 9 の誤り割合 δ/4 は [Zémor '01] による
- その後様々な改良
- [Spielman '96] は生成行列も工夫して，線形時間の符号化 & 復号を達成

エクスパンダーグラフによる符号の距離増幅

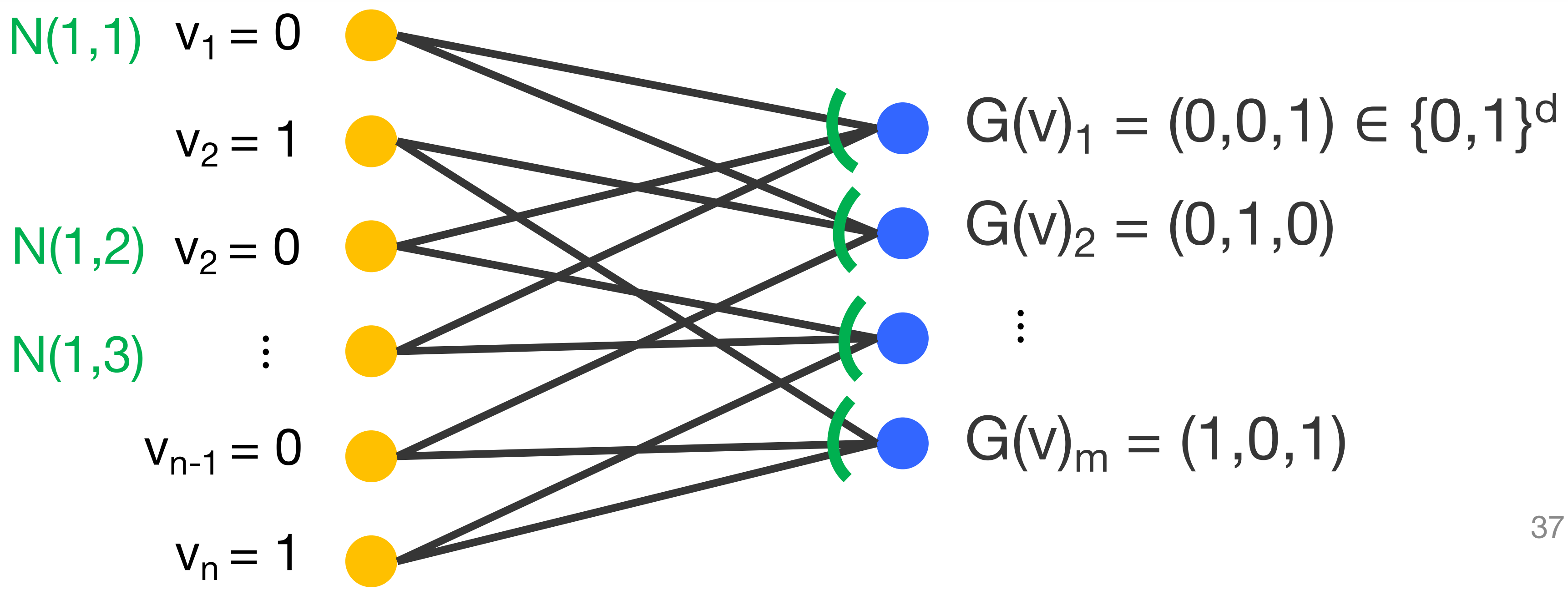
グラフによる距離増幅

定義． G = (L, R, E) を |L| = n, |R| = m の D-左正則, d-右正則グラフ， $C \subseteq \{0,1\}^n$ を二元線形符号とする．
$v \in C$ に対し， $G(v) \in (\{0,1\}^d)^m$ を， $j \in \{1, \ldots, m\}$ に,

$$G(v)_j = (v_{N(j,1)}, v_{N(j,2)}, \ldots, v_{N(j,1)})$$

と定め， $G(C) = \{ G(v) : v \in C \}$ とする．

- N(j, i) は右頂点 j と接続する i 番目の左頂点．



グラフによる距離増幅

- 二部グラフ $G = (L, R, E)$ が (γ, β)-分散器 (disperser)
  $\Leftrightarrow \forall S \subseteq L$ with $|S| \geq \gamma n$, $|N(S)| \geq \beta m$

定理 12. G が (γ, β)-分散器，C がレート R，最小距離 γn のとき，G(C) はレート R/D，最小距離 βm

証明：

- グラフより $nD = md$ でありレートは $Rn/md = R/D$
- $v, v' \in C$ に対し，$G(v)$ と $G(v')$ は左頂点で γn 個値が異なるため，(γ, β)-分散器の性質より，
  右頂点の βm シンボル以上は値が異なる　（証明終）

分散器の存在性

補題 13. $\forall\gamma, \varepsilon \in (0,1)$ に対し，多項式時間構成可能な $(\gamma, 1-\varepsilon)$-分散器，$D = d = \Theta(1/(\gamma\varepsilon))$, $|L| = |R|$ が存在

証明：

- $G = (L, R, E)$ を Ramanujan グラフの二重被覆とする
  - Ramanujan グラフは $(n, d, \lambda \le 2d^{1/2})$-エクスパンダー
- $\forall S \subseteq L$, $|S| = \gamma n$ に対し $|N(S)| \ge (1-\varepsilon)n$ を示す
- S を固定し，$T = R - N(S)$ とする．$|T| \le \varepsilon n$ を示す
- Expander Mixing Lemma より，

  $0 = |E(S, T)| \ge (d\cdot|S|\cdot|T|)/n - \lambda\,(|S|\cdot|T|)^{1/2}$

  $\Rightarrow\quad d^2\cdot|S|\cdot|T| \le \lambda^2 n^2 \quad\Rightarrow\quad d^2\,|T| \le \lambda^2 n/\gamma$

  $\Rightarrow\quad |T| \le (\lambda/d)^2 n/\gamma \quad\Rightarrow\quad |T| \le (4/\gamma d)n$
- $d \ge 4/(\gamma\varepsilon)$ であれば $|T| \le \varepsilon n$　（証明終）

最小距離の大きな符号

系 14. 相対距離 $1 - \varepsilon$, レート $\Omega(\varepsilon)$, アルファベットサイズ $2^{O(1/\varepsilon)}$ の多項式時間構成可能な符号が存在

証明：定数レート，定数相対距離を達成する多項式時間構成可能な符号を C とすれば，定理 12 と補題 13 より示される． (証明終)

- Singleton 限界（レート R に対し，相対距離 $\leq 1 - R$）と定数倍の差
- アルファベットサイズ $O(1/\varepsilon^2)$ の代数幾何符号と比べると，構成がシンプル
- この符号を外符号として連接符号化すれば，相対距離 $1/(2\varepsilon)$, レート $\Omega(\varepsilon^3)$ の二元符号に

グラフによる距離増幅の文献情報

- ここで紹介した構成法は [Alon, Bruck, Naor, Naor, Roth '92]
- C のアルファベットを $\{0,1\}^a$ にし,
次元 a の線形符号 $C_0 \subseteq \{0,1\}^D$ を使い,
左頂点から出る辺の値を C_0 の符号語とした
一般化構成法は [Alon, Edmonds, Luby '95]
    - もとの構成は $a = 1$, C_0 が D 回繰り返し符号
- [Guruswami & Indyk '02] は，レート R，距離 $1 - R - \varepsilon$，定数アルファベットサイズの符号に対し，$(1 - R - \varepsilon)/2$ 割合の誤りの線形時間復号

エクスパンダーグラフと符号に関するその他の文献

- [Guruswami & Indyk '03] 線形時間のリスト復号
- [Hemenway & Wootters '15] 線形時間のリスト復元 (recovery) をレート 1 – ε 符号に
    - [GI03] は低レート
- [Vadhan '10] 様々な擬似ランダムオブジェクトの等価性（リスト復号可能符号，エクスパンダーグラフ，擬似乱数生成器，標本器）
- [Dinur, Harsha, Kaufman, Navon '18] 二重標本器による効率的なリスト復号が可能な符号の一般的構成
    - 高次元エクスパンダーグラフによる構成

まとめ

- エクスパンダー符号
  - 基本構成：ビット反転法により線形時間復号
  - タナー符号による一般化構成：左右交互復号
- エクスパンダーグラフによる距離増幅
  - 定数アルファベットサイズで Singleton 限界に
  - 線形時間復号，線形時間リスト復号
- 特徴
  - 代数的構造の代わりにグラフの性質を利用
  - 復号法は，比較的シンプルで，線形時間復号が可能
  - 漸近的な性能を示すのに有効
- 今回の資料はおもに，Guruswami, “Introduction to Coding Theory” Spring 2010 の講義資料を参考にした