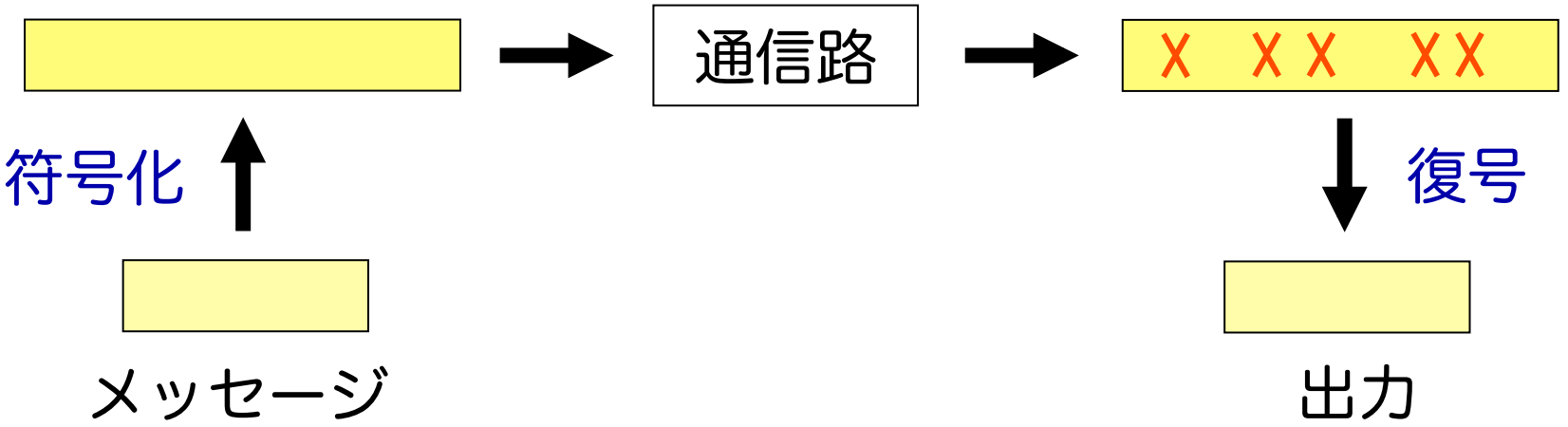
計算量制限通信路における誤り訂正

安永憲司

金沢大学

第３回 誤り訂正符号のワークショップ @千葉県 鳩山荘 松庵
2014.9.17-19

誤り訂正符号

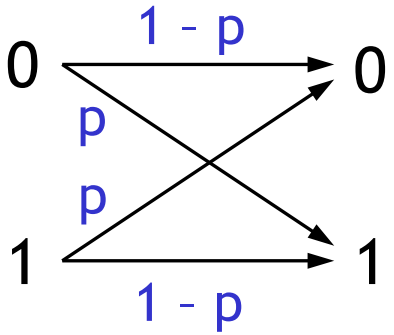


- 多くの誤りを訂正したい
- 多くのメッセージを送りたい（高い符号化レート）

→ その限界は通信路モデルに依存

通信路モデル

- 確率的通信路（２元対称通信路）
  - 各ビット独立に一定確率 p で誤りが発生 (p < 1/2)



$h(p) = -p \log p - (1-p) \log (1-p)$

  - 符号化レート R < 1 – h(p) で訂正可能 [Shannon ’48]
    - レート 1 – h(p) は最適 [Shannon ’48]
  - 効率的な符号化・復号法が存在
    - 連接符号 [Forney ’66]・ポーラ符号 [Arikan ’09]

通信路モデル

- 最悪ケース通信路
  - 符号語に挿入される誤りの数だけを制限
  - 誤り割合 $p \geq 1/4$ だと訂正不可能 [Plotkin '60]
    （符号化レート $R \to 0$ でない限り）
  - 誤り割合 $p < 1/4$ に対し、符号化レート $R \geq 1 - h(2p)$ で訂正可能 [Gilbet '52, Varshamov '57]
    - レート $1 - h(2p)$ が最適化かどうかは未解決
    - 明示的な構成法・効率的な復号法の存在も未解決

通信路モデルのギャップ

- 確率的通信路では、単純な方法で誤りが発生

  → 低コスト計算を行う通信路

- 最悪ケース通信路では、
  符号に関する十分な知識・考察から誤りが発生

  → 高コスト計算を行う通信路

計算量制限通信路

- Lipton (STACS ’94) が導入

- 通信路の計算量は、符号長の多項式時間
  - 確率的/最悪ケース通信路の中間モデル

以降の発表

- 計算量制限通信路に対する既存研究・方式の紹介
  - Lipton (STACS ’94)
  - Micali, Peikert, Sudan, Wilson (TCC ’05)
  - Gurswami, Smith (FOCS ’10)

- サンプル可能な加法的誤りの訂正
  - ISIT 2014 で発表した内容

Lipton (STACS '94)

- 計算量制限通信路 $C^{comp} : \{0,1\}^n \rightarrow \{0,1\}^n$
  - C^{comp} は多項式時間計算アルゴリズム
  - 誤り割合 p 以下

- BSC に対する符号 → C^{comp} に対する符号
  - C^{comp} に秘密の共有乱数を仮定
  - 一方向性関数の存在を仮定
    - 共有乱数長を短くするため
  - 最適レート 1 – h(p) の符号が存在
  - 計算量を制限しない場合 → Langberg (FOCS '04)

Lipton (1994) 方式

- (E, D) : BSC に対する符号化方式
- (E', D') : C^{comp} に対する符号を以下のように定める
  - $E'(x, s) = \pi(E(x)) + r$
  - $D'(y, s) = D(\pi^{-1}(y + r))$
  - ただし、π : ランダム置換, r : 一様ランダム系列
    これらは共有乱数 s を擬似乱数生成器で伸長して得られたもの
- C^{comp} でうまくいく直観的理由
  - r でマスクし、C^{comp} に対し符号語情報を隠す
  - π で置換し、C^{comp} に対しビット毎に独立

→ 復号器 D にとっては、重み制限ありのランダムエラー

Micali, Peikert, Sudan, Wilson (TCC '05, IEEE IT '10)

- 計算量制限通信路 C^{comp}
- 公開鍵基盤を仮定（送信者の公開鍵は既知）
  - 共有乱数は仮定しない

- リスト復号可能符号 → C^{comp} に対する符号
  - 「メッセージ＋カウンター＋署名」を符号化
    - カウンターは「動的符号化モデル」への対処のため
  - 一方向性関数（電子署名）の存在を仮定
  - 2元符号で誤り割合 $1/2 - \gamma$, レート $\Omega(\gamma^3)$ を達成
  - 多元符号で誤り割合 $1 - R$, レート R を達成

アルファベットが大きい場合に最適レートを達成

Guruswami, Smith (FOCS '10, arXiv '13)


- 共有乱数・公開鍵基盤ともに仮定しない
- 誤り割合 p 以下

- 符号化方式の構成
  - 最悪ケース加法的通信路に対する一意復号
    - 最適レート 1 – h(p) の符号の明示的構成法
  - 多項式時間制限通信路に対するリスト復号
    - 最適レート 1 – h(p) の符号のモンテカルロ構成法
- リスト復号の必要性
  - p > 1/4 の一意復号は不可能
    - 通信路が簡単な計算（符号語保持とカウント）可能なとき

Dey, Jaggi, Langberg, Sarwate (IEEE IT ’13)

- オンライン通信路
  - 符号語を1ビットずつ見て反転するかを決める
  - 誤り割合は制限
  - 共有乱数は仮定しない
  - 通信路の計算能力は制限しない

結果の詳細（最悪ケース加法的通信路）

定理 1（最悪ケース加法的通信路の一意復号）

任意の $p \in (0,1/2)$, $\varepsilon > 0$ に対して、
レート $R = 1 - h(p) - \varepsilon$ の符号 (Enc, Dec) が存在し、
任意の $m \in \{0,1\}^{RN}$, 重み pN の誤り $e \in \{0,1\}^N$ に対して

$$\Pr_\omega[\, \mathrm{Dec}(\, \mathrm{Enc}(m;\omega) + e \,) = m \,] \geq 1 - \exp(\, - \Omega(\varepsilon^2 N / \log^2 N) \,)$$

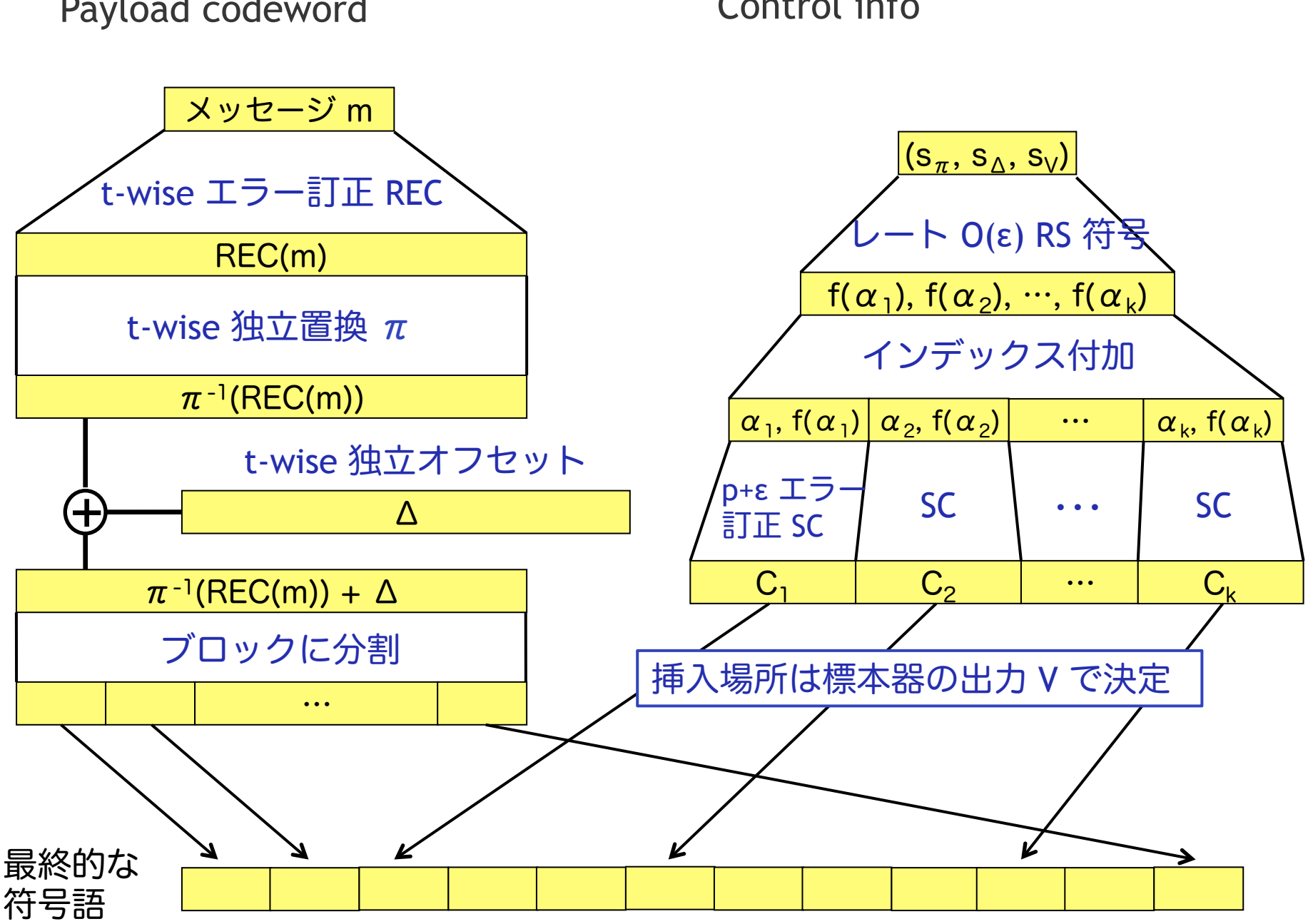
- (Enc, Dec) は poly(N) 時間計算可能
- 誤りは、符号化方式 (Enc, Dec) やメッセージ m に依存してよい（符号語はダメ）
- Enc が確率的であることが本質
  - 確定的である場合、通常の最悪ケース誤り

最悪ケース加法的通信路に対する一意復号

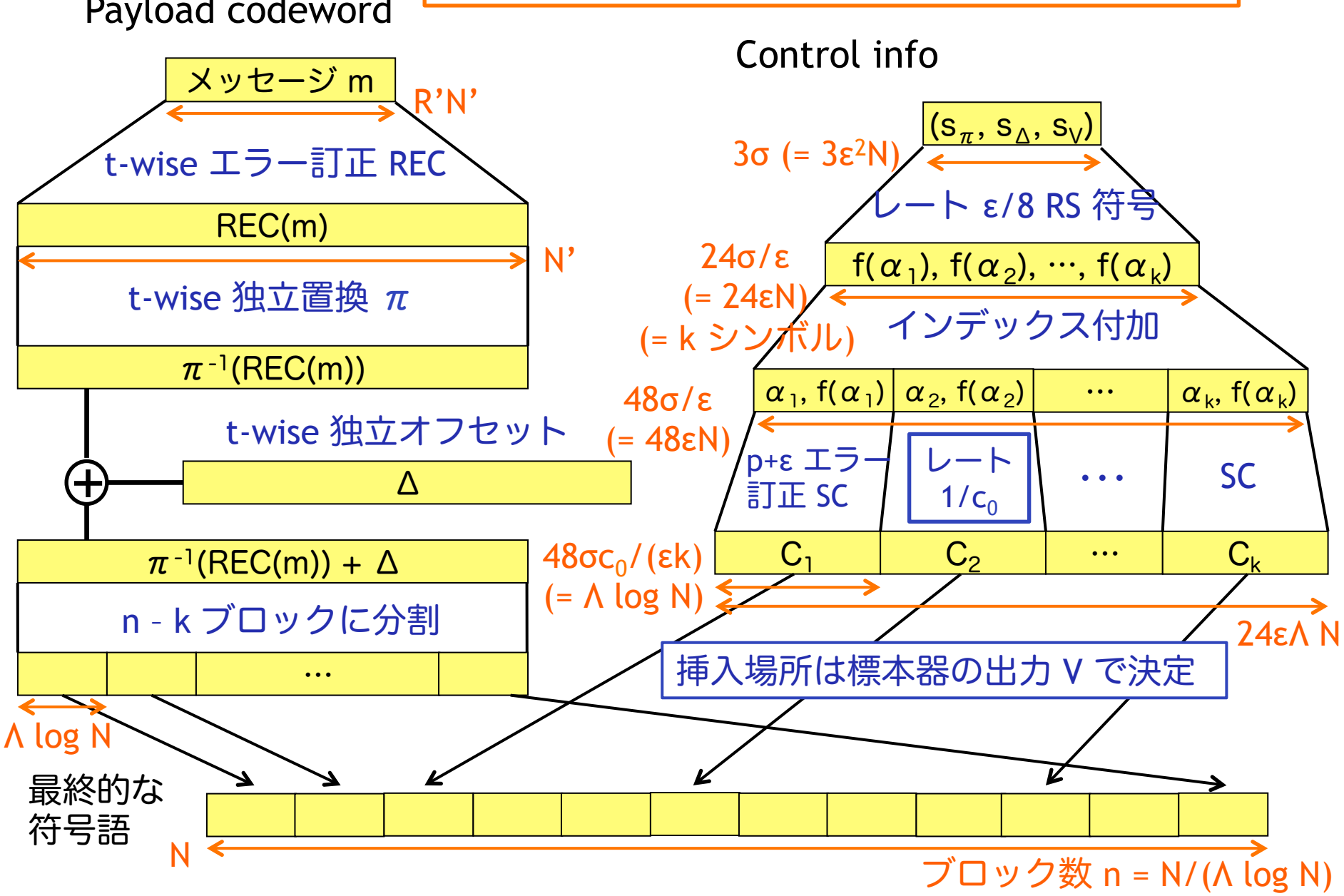
- アイディア：ランダムエラーに帰着
  - Lipton (1994) では擬似乱数生成器のシードを共有
  - 共有はできないため、シード情報 (control info) だけは訂正能力が高くなるように符号化
  - control block にエラーが集中すると困る

    → 標本器で block を分散し、エラーの入り方を平均化
  - 受信側で control / payload block を識別する必要

    → payload block に擬似ランダムオフセットを施し、control block の復号で誤り検出させる
  - control info を Reed-Solomon 符号化しておけば、一定の control block が集まれば control info を復元可


Payload codeword
Control info
メッセージ m
t-wise エラー訂正 REC
REC(m)
t-wise 独立置換 π
π⁻¹(REC(m))
t-wise 独立オフセット
Δ
π⁻¹(REC(m)) + Δ
ブロックに分割
…
(sπ, sΔ, sV)
レート O(ε) RS 符号
f(α1), f(α2), …, f(αk)
インデックス付加
α1, f(α1)
α2, f(α2)
…
αk, f(αk)
p+ε エラー訂正 SC
SC
・・・
SC
C1
C2
…
Ck
挿入場所は標本器の出力 V で決定
最終的な符号語

誤り割合 p, レート R = 1 - h(p) - ε の符号の構成法



R' = RN/N', N' = N - kΛ log N, σ = ε²N, k = 24εN/(log N), Λ = 2c0, t = Ω(ε²N/log N)

構成要素（1/2）

- 定数レート符号 SC : $\{0,1\}^b \times \{0,1\}^b \to \{0,1\}^{c_0 b}$
  - $b = O(\log N)$ ($= 2 \log N$)
  - $p + O(\varepsilon)$ 割合の加法的誤りを w.p. $1 - O(1/N)$ で訂正
  - ランダム系列の入力を w.p $1 - O(1/N)$ で検出
  - poly(N) 時間復号の方式が存在（リスト復号+AMD符号）

- レート $\varepsilon/8$ の RS 符号 : $\{0,1\}^{3\sigma}$ ($= F_q^{\varepsilon k/8}$) $\to F_q^k$, $q \approx N$
  - 正しいものが $\varepsilon k/4$ 以上、間違いが $\varepsilon k/12$ 以下のシンボル集合からメッセージを復元
    - $\varepsilon k/4 - \varepsilon k/12 \geq \varepsilon k/8$ なので、通常の RS 符号で復元可能

- 標本器 Samp : $\{0,1\}^\sigma \to [N]^k$ （ $[N] = \{1, 2, \ldots, N\}$ ）
  - $\forall B \subseteq [N]$, $|B| \geq \mu N$, $\forall\ 0 < \theta < \mu$,
    $|\mathrm{Samp}(s) \cap B| \geq (\mu - \theta)\,|\mathrm{Samp}(s)|$ w.h.p. over $s \in \{0,1\}^\sigma$
  - $\sigma \leq O(\log N + k \log(1/\theta))$ [Vadhan '04]

構成要素（2/2）

- almost t-wise 独立置換生成器 KNR: $\{0,1\}^{\sigma} \to S_{N'}$
  - $S_{N'}$: N' 要素置換の全体集合
  - $\forall\, i_1, \ldots, i_t, \{\pi(i_1), \ldots, \pi(i_t)\}_{\pi \leftarrow KNR}$ が一様分布と距離 2^{-t} 以下
  - $\sigma = O(t \log N')$ [Kaplan, Naor, Reingold '06]

- t-wise 独立分布生成器 $POLY_t : \{0,1\}^{\sigma} \to \{0,1\}^{N'}$
  - $\sigma = O(t \log N')$; 標数 2 の t 次多項式の評価で構成可能

- レート $R = 1 - H(p) - O(\varepsilon)$ 符号 $REC : \{0,1\}^{R'N'} \to \{0,1\}^{N'}$
  - 誤り割合 p 以下の t-wise 独立誤りを効率的に訂正

    $\Leftrightarrow \forall$ m, 重み pn 以下の $e \in \{0,1\}^n$ に対し、
    $Pr_{\pi \in KNR}[\, D_{REC}(REC(m) + \pi(e)) = m \,] \geq 1 - \exp(-\Omega(\varepsilon^2 t))$
  - 連接符号で $\omega(\log N') < t < O(\varepsilon N')$ で可能 [Smith '07]

証明の概要（control info の復元）

補題 1.1: ∀ 誤り e に対し、標本器の出力 V は、高い確率で good for e

- V が good for e ⇔ ε/2 割合以上の control block で誤り割合 p+ε 以下
- 証明概要: 標本器のシード s_V は e と独立 → 標本器の性質より、誤りの少ない block が一定割合存在

s_V と e の独立性

補題 1.2: V が good for e のとき、高い確率で、 (i) 正しく復号される control block は ε/4 以上、(ii) 誤復号される control block は ε/24 以下

- 証明概要: SC の Enc の乱数は e と独立 → SC の性質より、一定割合は正しく復元

SC の Enc 乱数と e の独立性

補題 1.3: ∀ m, e, s_V, s_π に対し、高い確率で、payload block が control block に間違われるのは ε/24 割合以下

- 証明概要: payload には t-wise 独立オフセット Δ
  → 各 block は長さ Λ log N ≥ t = Ω(ε²N/log N)
  → ランダム系列に対し、SC は 高い確率で誤り検出

s_Δ と m, e, s_V, s_π の独立性

補題 1.4: 任意の m, e に対し、高い確率で、control info は正しく復元

- 証明概要: 補題 1.1, 1.2, 1.3 より

証明の概要（payload の復元）

- control info (s_π, s_Δ, s_V) は正しく復元されたと仮定
- m, e, s_V を固定したとき、payload 部分の誤り e_{pay} の相対ハミング重みは $p\,(1 + 25\Lambda\varepsilon)$ 以下
- s_π は V と独立 → π は e_{pay} と独立 s_π と m, e, V の独立性
- 復号器 D_{REC} への入力は $\pi(y + \Delta) = x + \pi(e_{pay})$
  - 受信語 $y = \pi^{-1}(x) + \Delta + e_{pay}$
  - π は t-wise 独立置換であるため、
    訂正すべき誤り $\pi(e_{pay})$ は t-wise 独立
    → REC の性質より訂正可能

多項式時間制限通信路に対するリスト復号

- 最悪ケース加法的誤りと異なる部分

  (i) 正当な control block を広範囲に挿入できる

  → リスト復号にして、リストサイズを抑える

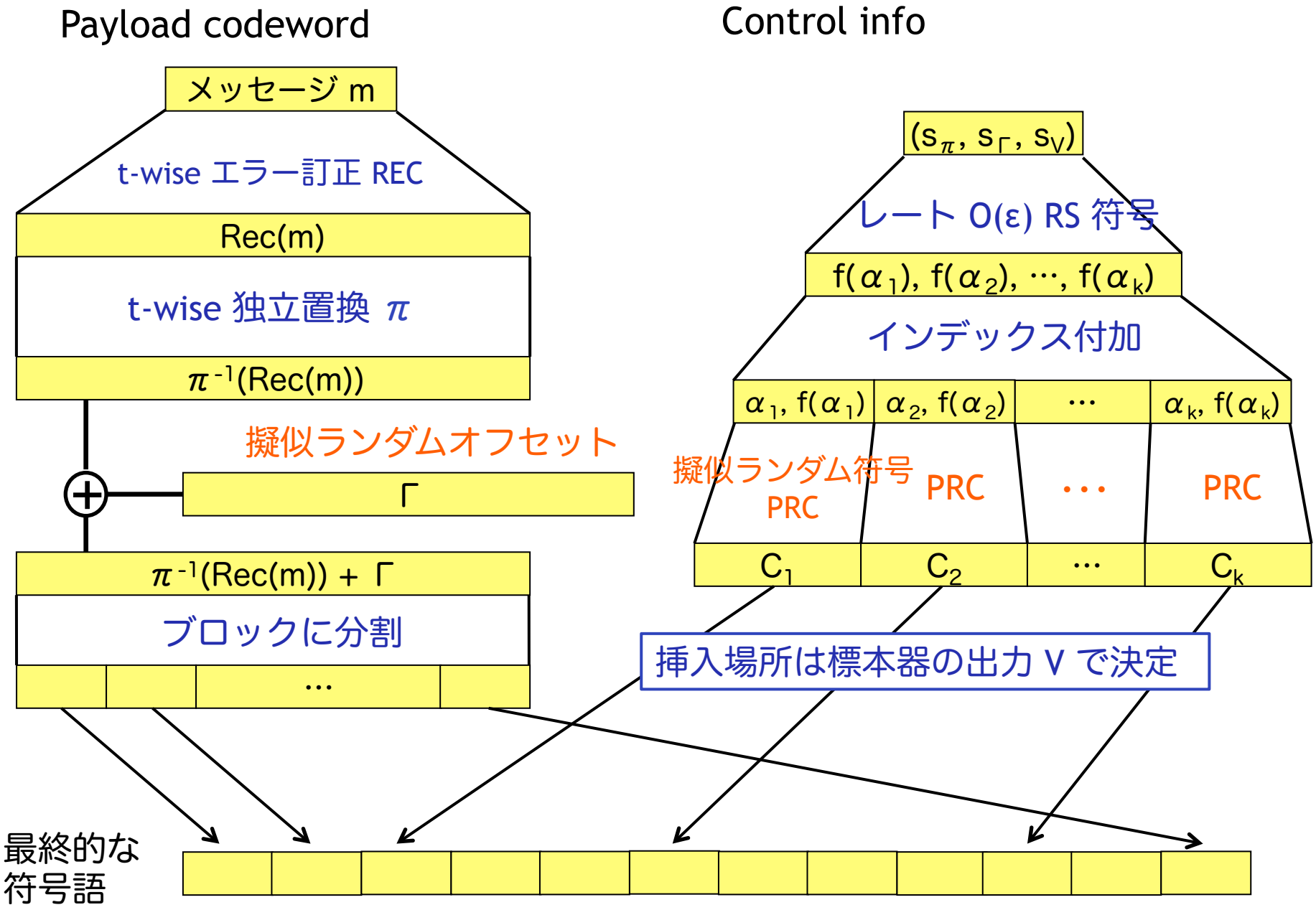
  (ii) control block に誤りが集中する可能性があり、また、置換情報は通信路から隠す必要がある

  → 符号語を擬似ランダム系列にする

  → 通信路が挿入した誤りは(擬似)ランダム

  → 誤りが control info と(擬似)独立なので、最悪ケース加法的誤りの議論に帰着

T_0-time 通信路に対する，誤り割合 p，レート R = 1 - H(p) - ε の符号の構成法



新しい構成要素

- 擬似乱数生成器 PRG: $\{0,1\}^{\zeta N} \to \{0,1\}^N$
  - (T, 1/T)-pseudorandom

    ⇔ 任意のサイズ T の回路 A に対して
    $|\Pr[A(PRG(U_{\zeta N})) = 1] - \Pr[A(U_N) = 1]| \le 1/T$
  - poly-size PRG のモンテカルロ構成法（ランダム関数）

- 擬似ランダム符号 PRC : $\{0,1\}^{Rb} \times \{0,1\}^s \to \{0,1\}^b$
  - (δ, L)-list decodable & (T, 1/T)-pseudorandom
  - ∀ $\delta \in (0,1/2)$ に対し、レート $R \ge (1/2 - \delta)^{\Omega(1)}$,
    $L \le 1/(1/2 - \delta)^{O(1)}$, $b = \Lambda_0 \log(T)$ のモンテカルロ構成法
    - Enc(m, r) = C(m) + BPPRG(r), $s = 10 \log(T)$, $b = \Lambda_0 \log(T)$
    - C: linear list-decodable code [Guruswami, Sudan '00]
    - BPPRG: 各出力を一様ランダムに選択（poly(T)-time 構成可）

結果の詳細（多項式時間制限通信路）

定理 2（多項式時間通信路のリスト復号）

誤り数 pN 以下の T_0 時間制限通信路 W ($p \in (0,1/2)$)
および $\varepsilon > 0$ に対して、レート $R = 1 - h(p) - \varepsilon$ の
符号 (Enc, Dec) の効率的なモンテカルロ構成法が存在し、
任意の $m \in \{0,1\}^{RN}$ に対して、Dec(W(Enc(m))) は
サイズ poly(1/ε) 以下のリストを出力し、
確率 $1 - O(N/T_0)$ 以上でリストに m が含まれる

- (Enc, Dec) は poly(N) 時間計算可能
- (Enc, Dec) は W の計算時間 T_0 に依存して構成
  - すべての多項式時間通信路に対して動作するのではない
- CRS (common reference string) を仮定すれば実現可能
  - CRS の乱数系列をモンテカルロ構成の乱数とする

証明の概要（control info のリスト復号）

補題 2.1: $\forall$ m, π, V に対して, $\{Enc(m; \pi, V, \cdot)\} \approx_T \{U_{|Enc|}\}$

- 証明概要: ランダムオフセット Γ と PRC の擬似ランダム性より（PRC は任意のメッセージを擬似ランダムにできる）

系 2.1: $\forall$ m, π, V, 通信路 W に対し、
$\{e_W(Enc(m;\pi,V,\cdot)\} \approx_{T'} \{e_W(U_N)\}$

e_W は m, π, V と(擬似)独立

- $e_W(x)$: 入力 x に対して通信路 W が加える誤り

補題 2.2: $\forall$ T_0 時間制限通信路 W, m, s_π に対し、標本器の出力 V は、高い確率で good for e_W

- 証明概要: 系 2.1 より, $\forall$ m, π, V に対し, $\{e_W\} \approx_{T'} \{e_W(U_N)\}$
  → e_W は V と計算量的独立 & V が good かは効率的検査可
  → あとは標本器の性質より

補題 2.3: V が good for e のとき, PRC は ε/2 割合以上の control block において正しいシンボル α_iを復元

補題 2.4: $\forall$m, e, ω, 各シンボル α_i の候補数は Θ(NL/log T) 以下

リスト復号の必要性
（p > 1/4 の bit-fixing 通信路の不可能性）

定理 3

任意の (Enc, Dec) でメッセージ空間サイズ $|M| \to \infty$ を送るとき、通信路にて一様ランダムなメッセージを送ると

1. 平均 n/4 ビット誤りの記憶のない通信路が存在し、復号誤り率 $\geq 1/2 - o(1)$
2. $\forall\ 0 < \nu < 1/4$, online space-(log(N)) 通信路が存在し、$N(1/4 + \nu)$ 誤りで、復号誤り率 $\Omega(\nu)$

証明概要：

- Swapping 通信路 $W_s(c)$ と $W_c(s)$は同一
  → s が codeword なら復号器はこれらを区別できず w.p. 1/2 で間違う
- 誤り数は dist(c,s) の半分なので、dist(c,s) がおよそ n/2 であれば、n/4 程度
- online space-(log(N)) で近いことができる（誤りの数を数えておくだけ）

Swapping 通信路 $W_s(c)$:

入力 $c = (c_1, \ldots, c_N)$ に対して

$$W_s(c)_i = \begin{cases} c_i & \text{if } c_i = s_i \\ U \in_R \{0,1\} & \text{o.w.} \end{cases}$$

$$= \begin{cases} c_i & \text{w.p. } 1/2 \\ s_i & \text{w.p. } 1/2 \end{cases}$$

Open Questions

- 時間制限通信路のリスト復号 → 一意復号
  - p > 1/4 では簡単な計算可能な通信路に対して不可能（正当な符号語の保持と誤り数のカウント程度）
  - 不可能性の回避方法は？
    - p < 1/4. 一意復号で達成可能なレートは？GV 超えは可能？
    - 送信符号語の部分情報しか通信路に渡さない
    - 共有乱数・PKI では事前に正当な符号語を保持できず回避
- 時間制限通信路のモンテカルロ構成 → explicit
  - PRG は通常の OWF 仮定で explicit にできる
  - PRC は？ 指数時間安全 OWF にすればよいが…
  - 明示的構成法に OWF 仮定は必要か？
    - 上限なしの多項式時間通信路では？

サンプル可能な加法的誤りの訂正

本研究の内容

- 計算量制限通信路の１つとして
  「サンプル可能な加法的通信路」を導入

- 誤り訂正の可能性・限界について考察

サンプル可能な分布

- $\{0,1\}^n$ 上の確率分布 Z がサンプル可能

  ⇔ 確率的多項式時間アルゴリズム S が存在し、
  $S(1^n)$ が Z に従って分布

- サンプル可能な分布に関する関連研究
  - データ圧縮 [GS91, Wee04, TVZ05]
  - 乱数抽出 [TV00, Vio11, DW12, DRV12, DPW14]

サンプル可能な加法的通信路

- サンプル可能な分布 Z による
  加法的通信路 $W^Z : \{0,1\}^n \rightarrow \{0,1\}^n$

$$W^Z(x) = x + z,\ z \sim Z$$

  - 誤りベクトル z は x に依存しない
  - 分布 Z は符号化方式に依存しない
    - 逆に、符号化方式は W^Z の知識を使って構成可
  - 発生する誤りの数（z の重み）は制限しない
    - 既存研究のほとんどは、誤りの数を制限

サンプル可能な加法的通信路を導入する理由

- シンプルな通信路モデル
  - 誤り分布は、どの符号・符号語に対しても同一
  - ２元対称通信路の計算能力強化版
    - ただし、記憶ありの通信路モデル

- 「誤りが多い → 訂正できない」を反証したい
  - 誤りの数が多くても、
    そこに構造があれば訂正可能（かもしれない）

→ どのような Z なら訂正可能か？

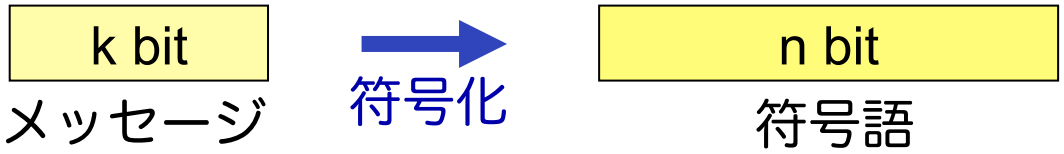
訂正可能性を調べるための基準

- 分布 Z のエントロピー H(Z) を基準とする

$$H(Z) = \mathop{\mathrm{E}}_{z}\left[\log \frac{1}{p_Z(z)}\right] = \sum_{z \in \mathrm{supp}(Z)} p_Z(z) \log \frac{1}{p_Z(z)}$$

  - Z が $\{0,1\}^n$ 上のとき、$H(Z) \in [0,n]$
  - H(Z) = 0 ならば簡単に訂正可能
  - H(Z) = n ならば訂正不可能

- 達成可能な符号化レート R = k/n は？



考察１：Ｚが２元対称通信路

- Ｚは２元対称通信路をシミュレートできる

定理１

H(Z) = n･h(p) (= Ω(n)) のとき、
R > 1 - h(p) では訂正不可能な Z が存在

考察２：Z が擬似ランダム分布

- Z が擬似乱数生成器の出力のとき
  効率的に（多項式時間で）訂正はできない

定理２

$H(Z) \leq n^{\varepsilon}$ で効率的に訂正不可能な Z が存在

- 任意の $0 < \varepsilon < 1$
- 一方向性関数の存在を仮定

考察３：Z が線形空間

定理３

次元 m の線形空間 $Z \subseteq \{0,1\}^n$ に対して、
任意の $z \in Z$ を訂正可能なレート $R = 1 - m/n$
の線形符号が存在（復号も効率的）

- 証明の概要：
  - Z の基底 $\{z_1, \ldots, z_m\}$ と $\forall z = \Sigma_i a_i z_i$ に対し、
    $T(z) = (a_1, \ldots, a_m)$ を満たす
    線形変換 $T : \{0,1\}^n \rightarrow \{0,1\}^m$ が存在
  - $T(x) = (0, \ldots, 0)$ を満たす x を符号語とする

考察４：Z が平坦分布

定理４.１

任意の $\varepsilon > 0$, 平坦分布 Z に対して、
Z を誤り率 ε, レート $R \geq 1 - m/n - O(\log(1/\varepsilon)/n)$ で
訂正可能な線形符号が存在（ $|\mathrm{supp}(Z)| = 2^m$ ）
- 構成は明示的でなく、復号法も効率的でない

証明の概要：線形符号集合と線形無損失濃縮器の等価性
および線形無損失濃縮器の存在性から [Cheraghchi (ISIT'09)]

定理４.２

任意の平坦分布 Z は、レート $R \geq 1 - m/n + O(1/n)$,
誤り率 $\varepsilon < 1/2$ では訂正不可能（$|\mathrm{supp}(Z)| = 2^m$）

証明の概要：受信語空間 $\{0,1\}^n$ を各サイズ $(1-\varepsilon)2^m$ の
2^{Rn} 個の素集合に分割する必要があるため

考察５：Z が小バイアスサンプル空間

- サンプル空間 $S \subseteq \{0,1\}^n$ のバイアスが δ
  $\Leftrightarrow \forall$ 非零 $a \in \{0,1\}^n$ に対して、$|E_{x\sim S}[(-1)^{a\cdot x}]| \leq \delta$

Z が小バイアス $\Leftrightarrow$ 線形関数では一様分布と識別不能

定理５

バイアス δ のサンプル空間上の一様分布 Z は、
レート $R > 1 - \Omega(\log(1/\delta)/n)$, 誤り率 $\varepsilon < 1/2$
では訂正不可能

証明の概要：最小エントロピーが保証される分布に対して、
Z は使い捨て鍵暗号の鍵になる [DS05] → 訂正できない

系５

$H(Z) = m$ のサンプル可能な分布 Z が存在し、
レート $R \geq 1 - m/n + O((\log n)/n)$ では訂正不可能

考察６：低エントロピーの訂正不可能分布

定理６.１

$\forall\, \omega(\log n) < m < n$ に対して、
あるオラクル O にアクセス可能なとき、
$H(Z) = m$ であるサンプル可能な Z が存在し、
Z はレート $R > \omega((\log n)/n)$ のとき
「効率的なシンドローム復号」では訂正不可能

証明の概要：

- $H(Z) = \omega(\log n)$ で長さ $< n - \omega(\log n)$ に効率的に圧縮できないサンプル可能分布が存在 [Wee04]
  （オラクル O にアクセス可能なとき）
- レート R で Z をシンドローム復号訂正可能
  ⇔ Z を長さ $n(1 - R)$ に線形圧縮可能 [Caire et al. ’04]

考察６：低エントロピーの訂正不可能分布

定理６.２

$\forall\, \omega(\log n) < m < n$ に対して、
あるオラクル O にアクセス可能なとき、
$H(Z) = m$ であるサンプル可能な Z が存在し、
Z はレート $R > \omega((\log n)/n)$ のとき
「効率的な復号」では訂正不可能

証明の概要：
[Wee04] のテクニック（再構成補題/圧縮補題）を利用
- O はランダム関数 f にアクセスするオラクル
- $Z = f(U_m)$ とし、
「Z を多項式サイズ回路で訂正可」→「f を簡潔に記述可」
という事実から、多項式サイズ回路で訂正可能な f は
それほど多くないため、ランダム関数 f は訂正できない

サプル可能な Z の訂正可能性（まとめ）

H(Z)	訂正可能性	仮定	参照
0	効率的に訂正可能		自明
$\omega(\log n)$	レート $R > \omega((\log n)/n)$ で効率的に訂正不可能	オラクルアクセス	定理 6.2
n^{ε} $(0 < \varepsilon < 1)$	効率的に訂正不可能	OWF	定理 2
$n \cdot h(p)$ $(0 < p < 1)$	$R > 1 - H(p)$ で訂正不可能		定理 1
$0 \le m \le n$	∀ 次元 m の線形空間 Z は $R \le 1 - m/n$ で訂正可能		定理 3
$0 \le m \le n$	∀ 平坦 Z は $R \le 1 - m/n - \Omega(\log(1/\varepsilon)/n)$ で訂正可能		定理 4.1
$0 \le m \le n$	∀ 平坦 Z は $R > 1 - m/n + O(1/n)$ で訂正不可能		定理 4.2
$0 \le m \le n$	∀バイアス δ 分布は $R > 1 - m/n + O((\log n)/n)$ で訂正不可能		系 5
n	訂正不可能		自明

今後の研究

- 提案モデルの妥当性
  - 現実に、対応する状況は存在するのか？
- （より限定的な Z に対する）訂正可能性の結果
  - 定数段回路でサンプル可能な Z
  - 対数領域サンプル可能な Z
- オラクルアクセス・OWF の仮定なしでの証明
  - または、仮定が不可避であることの証明

まとめ

- 計算量制限通信路
  - 確率的通信路と最悪ケース通信路の中間モデル
- 既存研究
  - Lipton (1994)：共有乱数設定
  - Micali, Peikert, Sudan, Wilson (2005): PKI 設定
  - Guruswami, Smith (2010): 共有乱数・PKI 仮定なし
    - 最悪ケース加法的誤り通信路に対する一意復号
    - 多項式時間制限通信路に対するリスト復号
- サンプル可能な加法的誤り通信路
  - 中間モデルの１つとして導入
  - 訂正可能性・限界についての考察

オラクルアクセスについて

定理６．２

$\forall\, \omega(\log n) < m < n$ に対して、
あるオラクル O にアクセス可能なとき、
(a) $H(Z) = m$ であるサンプル可能な Z が存在し
(b) Z はレート $R > \omega((\log n)/n)$ で効率的に復号できない

$H(Z) = \omega(\log n)$ の任意のサンプル可能な分布 Z を
効率的に訂正する符号のブラックボックス構成は存在しない

理由: ブラックボックス構成が存在するとき、任意の
オラクルにアクセスしても構成可能 → 定理に矛盾

また、符号構成が Z に依存して存在する場合は、
非ブラックボックス構成であり、排除されてない

定理 6.1（最悪ケース通信路の一意復号）

任意の $p \in (0,1/2)$, $\varepsilon > 0$ に対して、
レート $R = 1 - H(p) - \varepsilon$ の符号 (Enc, Dec) が存在し、
任意の $m \in \{0,1\}^{RN}$, 重み pN の誤り $e \in \{0,1\}^N$ に対して

$$\Pr_{\omega}[\ \mathrm{Dec}(\ \mathrm{Enc}(m;\omega) + e\ ) = m\ ] \geq 1 - \exp(\ -\Omega(\varepsilon^2 N/\log^2 N)\ )$$

定義 6.2 (Good sampler seeds)

標本集合 V が good for error vector e
⇔ 誤り割合が p + ε 以下の control block の割合が ε/2 以上

V is good for e ⇔ SC-Dec で正しく復号されるものが ε/2 割合以上

補題 6.3 (Good sampler lemma)

相対重み p 以下の任意の error vector e に対し、
Samp の出力 V は good for e w.p. $1 - \exp(-\Omega(\varepsilon^3 N/\log N))$
（確率は Samp のシードでとる）

標本器 Samp の出力は、高い確率で good for e

証明のスケッチ：
- $B \subseteq [n]$：誤り割合 $\leq p + \varepsilon$ のブロック集合
- ブロック全体における B の割合は ε 以上
- Samp のシードは error vector e とは独立に選ばれるため、
  control block における B の割合も ε 程度

補題 6.4 (Control block lemma)

任意の e, V s.t. V が good for e に対して
確率 $1 - \exp(-\Omega(\varepsilon^3 N/\log N))$ で以下が起きる
（確率は k 個の SC-Enc の乱数）
(i) SC-Dec で正しく復号される control block は $\varepsilon k/4$ 個以上
(ii) SC-Dec で間違って復号される control block は $\varepsilon k/24$ 個未満
（出力が正しくなく、⊥ でもない場合）

V が good for e のとき、SC-Dec の結果は、RS-Dec につなぐことができる

証明のスケッチ：
- control block C_j に誤り e_i が加わるとき
- e_i は C_j を生成する SC-Enc の乱数とは独立（つまり加法的誤り）
  → SC の性質より
  (i) e_i の誤り割合 $\leq p + \varepsilon$ のとき、高確率で正しく復号
  (ii) e_i の誤り割合 $> p + \varepsilon$ のとき、高確率で ⊥ 出力
- (i) でも (ii) でもない C_j の数が $\varepsilon k/24$ を越える確率は
  Chernoff bound より、非常に小さい

補題 6.5 (Payload block lemma)

任意の m, e, s_V, s_π に対して
確率 $1 - \exp(-\Omega(\varepsilon^2 N/\log^2 N))$ で、
（確率はオフセットのシード s_Δ でとる）
payload block のうち SC-DEC で control block と
間違われて復号される数は $\varepsilon k/24$ 以下

payload block のうち control block に間違われるのは $\varepsilon k/24$ 以下

証明のスケッチ：

・オフセット Δ は t-wise 独立（ $t = \Omega(\varepsilon^2 N/\log N)$ ）であり、
各ブロックは $\Lambda \log N$ ビットなので、
各 payload block には一様ランダムな Δ_i が加わっている
→ SC-Dec は高確率で $\perp$ を出力

・各ブロックも $(t/\Lambda \log N)$-wise 独立であるため、
control block に間違われる payload block が $\varepsilon k/24$ 個以上の確率は小さい
（t-wise independence の tail bound [Bellare, Rompel ‘94]）

補題 6.6 (Control information lemma)

任意の m, e に対して
確率 $1 - \exp(-\Omega(\varepsilon^2 N/\log^2 N))$ で、
（確率は control info および SC-Enc の乱数）
control info は正しく復元される

高い確率で control block は正しく復元

証明のスケッチ：

・補題 6,4 & 6.5 より、
$\varepsilon k/4$ 個以上の control block が正しく復元され、
$\varepsilon k/24$ 個以下の control block に誤りが含まれ、
$\varepsilon k/24$ 個以下の payload block が control block に間違われている

・$\varepsilon k/4 - 2(\varepsilon k/24) = \varepsilon k/6 > \varepsilon k/8$ であるため、
RS-Dec で正しく control info を復元できる

定理 6.1（最悪ケース通信路の一意復号）

任意の $p \in (0,1/2)$, $\varepsilon > 0$ に対して、
レート $R = 1 - H(p) - \varepsilon$ の符号 (Enc, Dec) が存在し、
任意の $m \in \{0,1\}^{RN}$, 重み pN の誤り $e \in \{0,1\}^N$ に対して

$$\Pr_\omega[\, \mathrm{Dec}(\, \mathrm{Enc}(m;\omega) + e \,) = m \,] \geq 1 - \exp(\, - \Omega(\varepsilon^2 N / \log^2 N) \,)$$

証明のスケッチ：

・補題 6.6 から、任意の m, e に対して
高確率で control info は正しく復元される

control info が正しいとき

・m, e, s_V を固定したとき payload 側の誤り割合は $p(1 + 25\Lambda\varepsilon)$ 以下

・s_π は V とは独立に選ばれるため、REC-Dec への入力は
重み$p(1 + 25\Lambda\varepsilon)$ 以下の t-wise 独立誤りが挿入された系列
→ REC-Dec で payload block は高確率で正しく復号

補題 7.6 (Few control candidates) [再掲]

control info のリストはサイズ $L' \le \mathrm{poly}(1/\varepsilon)$ であり、
リストは確率 $1 - \beta_{\mathrm{control}}$ 以上で正しいものを含む
ただし、$\beta_{\mathrm{control}} \le \beta_V + \beta_\Gamma(T_2) + N\,\beta_{\mathrm{PRC}}(T_2) \le (N+3)/T_0$
$T_2 = T_0 + O(N \log N)$

control info の候補リストのサイズは poly(1/ε)

証明のスケッチ：
- 補題 7.11, 7.12, 7.13 より、
  あとは RS 符号のレートを定め、リストサイズを保証するだけ
- データ数 $n \le NL/b_{\mathrm{ctrl}}$ のうち一致数 $t \ge \Theta(\varepsilon^2 N/\log(T_0))$ とするには
  レート $O(\varepsilon^4/L)$ の RS 符号でリストサイズ $O(L/\varepsilon^2)$ のリスト復号が可能
  （Guruswami-Sudan リスト復号）
- リスト復号半径 $\delta = p+\varepsilon < 1/2 - \varepsilon$ なので、$L' \le 1/\varepsilon^{O(1)} = \mathrm{poly}(1/\varepsilon)$