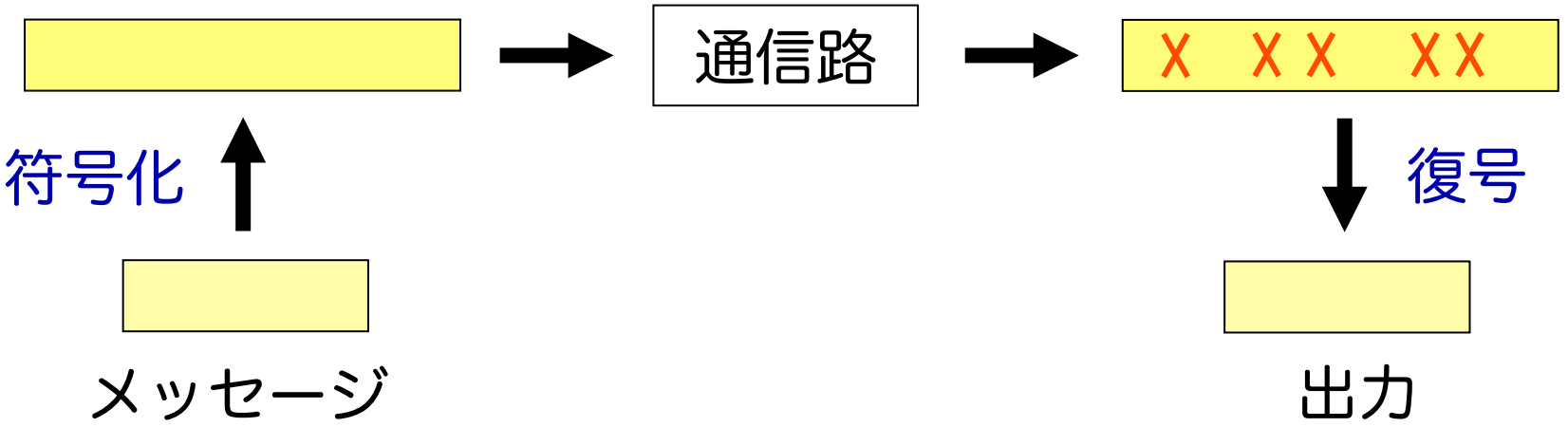
Error Correction in Computationally Bounded Channels

安永憲司

金沢大学

2013.9.12

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（二元対称通信路）
  - 各ビット毎に独立に一定確率 p で誤りが発生
  - 確率 $p < 1/2$ に対し
    符号化レート $1 - H(p)$ で訂正可能
    - レート $1 - H(p)$ は最適
  - 効率的な符号化・復号法が存在
    - 連接符号・Polar 符号

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 $p < 1/4$ に対し
    符号化レート $1 - H(2p)$ で訂正可能
    - 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
  - 誤り割合 $p \geq 1/4$ だと訂正不可能
    （符号化レートが 0 でない限り）

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

- 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

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

- 関連研究紹介
- Gurswami & Smith (2010) の符号化方式の紹介
- 標本可能な加法的誤りの訂正可能性

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
  - 符号語を擬似ランダムに置換することで、C^{comp} の誤り → ランダム誤りに
  - 一方向性関数の存在を仮定
  - レート 1 - H(p) の符号が存在

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
  - 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
  - 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 - γ, 正レートを達成
  - 多元符号で誤り割合 1 - R, レート R を達成

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

- 共有乱数・公開鍵は仮定しない
- 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

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

Gurswami & Smith (2010) の符号化方式

結果

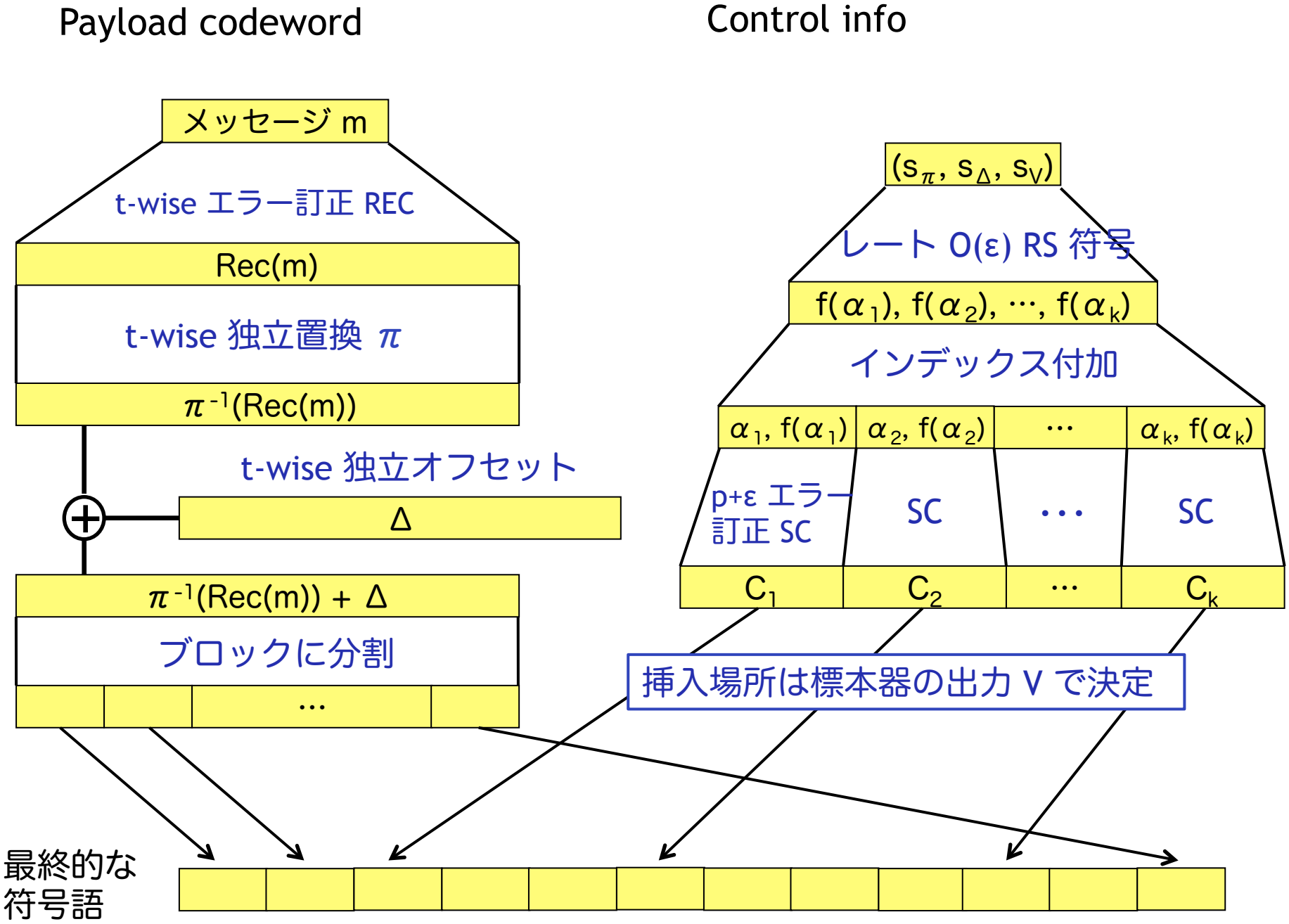
- 設定
  - 共有乱数・公開鍵は仮定しない
  - 誤り割合は高い確率で p 以下
- 最悪ケース加法的通信路に対する一意復号
  - 最適レート 1 - H(p) の符号の明示的構成法
- 多項式時間制限通信路に対するリスト復号
  - 最適レート 1 - H(p) の符号のモンテカルロ構成法
- リスト復号の必要性
  - p > 1/4 の一意復号は不可能
    - 通信路が符号方式を知っていて、簡単な計算ができるとき

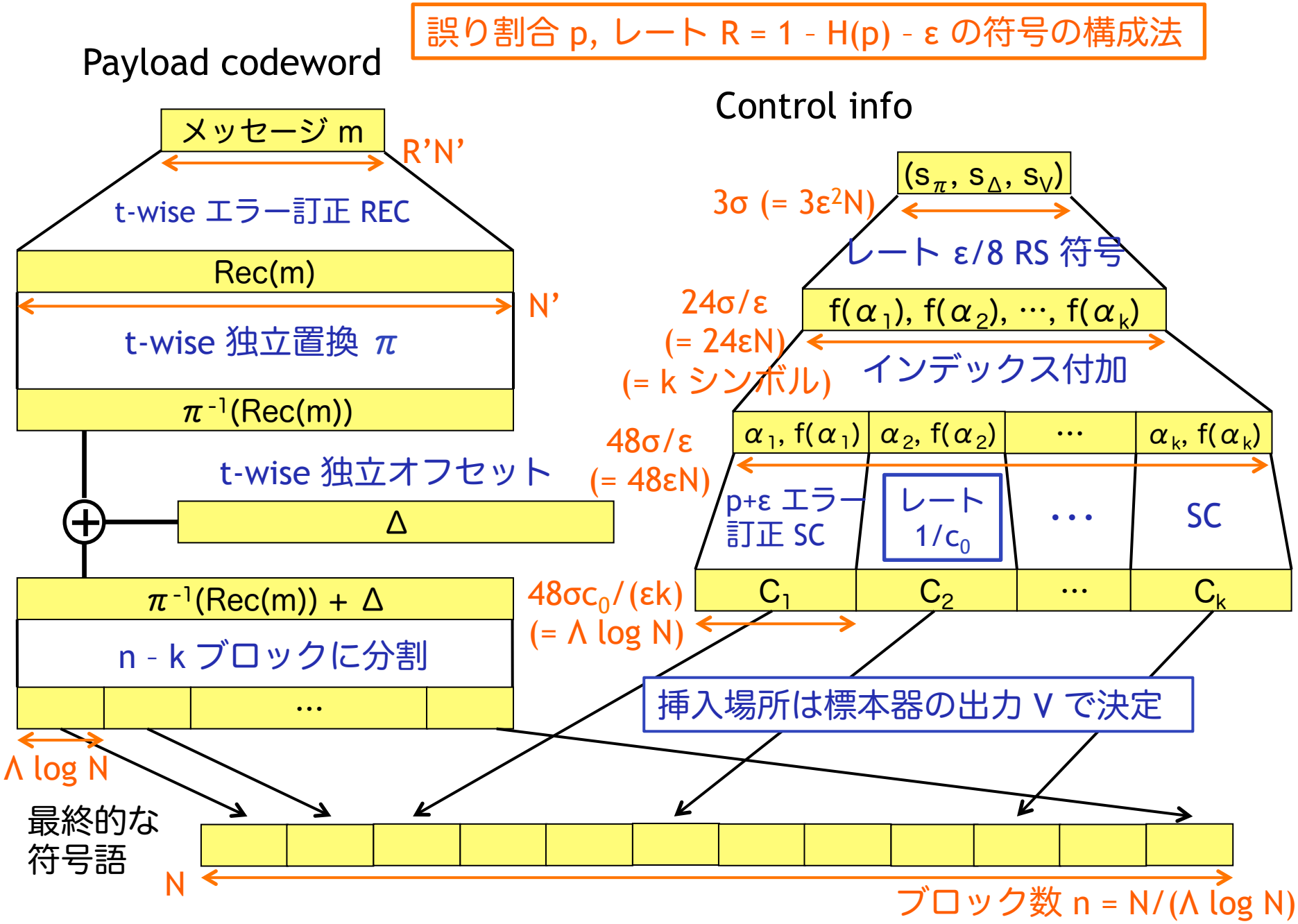
最悪ケース加法的通信路に対する一意復号

- アイディア：ランダムエラーに帰着
  - Lipton (1994) では擬似ランダム置換を共有
  - 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
Payload codeword
Control info
メッセージ m
R'N'
t-wise エラー訂正 REC
Rec(m)
N'
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
n - k ブロックに分割
…
Λ log N
(s_π, s_Δ, s_V)
3σ (= 3ε²N)
レート ε/8 RS 符号
24σ/ε
(= 24εN)
(= k シンボル)
f(α_1), f(α_2), …, f(α_k)
インデックス付加
48σ/ε
(= 48εN)
α_1, f(α_1)
α_2, f(α_2)
…
α_k, f(α_k)
p+ε エラー
訂正 SC
レート
1/c_0
…
SC
48σc_0/(εk)
(= Λ log N)
C_1
C_2
…
C_k
挿入場所は標本器の出力 V で決定
最終的な
符号語
N
ブロック数 n = N/(Λ log N)
R' = RN/N', N' = N - kΛ log N, σ = ε²N, k = 24εN/(log N), Λ = 2c_0,

構成要素（1/2）

- 定数レート符号 SC : $\{0,1\}^b \times \{0,1\}^b \to \{0,1\}^{c_0 b}$
  - $b = O(\log N)$ ($= 2 \log N$)
  - $p + O(\varepsilon)$ 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
  - 定数レートで十分なため、効率的な復号法が存在

- 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 : $\{0,1\}^\sigma \to [N]^k$
  - $\forall B \subseteq [N]$, $|B| \geq \mu N$,
    $|\mathrm{Samp}(s) \cap B| \approx \mu |\mathrm{Samp}(s)|$ 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' 要素置換集合
  - $\sigma = O(t \log N')$ [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 $REC : \{0,1\}^{R'N'} \to \{0,1\}^{N'}$
  - p 割合以下の t-wise 独立エラーを効率的に訂正

    $\Leftrightarrow$ 任意の m, 重み pn 以下の $e \in \{0,1\}^n$ に対し、
    REC-DEC(REC(m) + π(e)) = m
    w.p. $1 - \exp(-\Omega(\varepsilon^2 t))$ over $\pi \in$ range(KNR)
  - 連接符号で $\omega(\log N') < t < O(\varepsilon N')$ で可能 [Smith '07]

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

多項式時間制限通信路に対するリスト復号

- 加法的誤りと異なる部分

  (i) 正当な control block を簡単に挿入できる

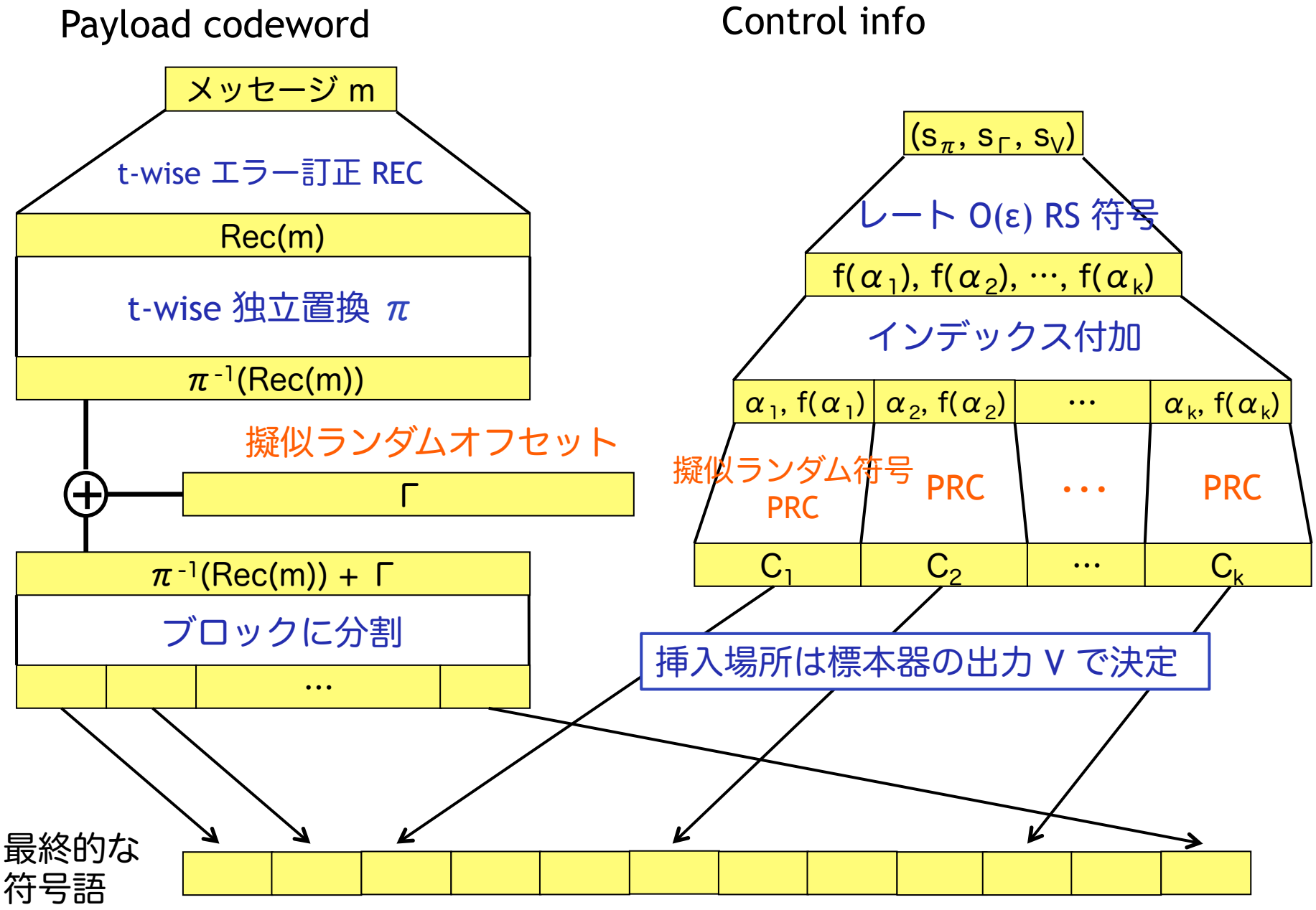
  → リスト復号にして、リストサイズを抑える

  (ii) 符号語を見てから誤りを入れるので、
  control block に誤りが集中する可能性
  (置換情報は隠す必要)

  → 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

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）
    - OWF の存在性を仮定すれば explicit に

- 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
  - 任意の $\delta \in (0,1/2)$ 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）

定理 7.5（多項式時間通信路のリスト復号）

任意の $p \in (0,1/2)$, $\varepsilon > 0$ with $0 < 2\varepsilon < 1/2 - p$,
pN-bounded T_0-time 通信路に対し、
レート $R = 1 - H(p) - \varepsilon$ の符号方式 (Enc, Dec) が存在し、
任意の $m \in \{0,1\}^{RN}$ に対して、高確率でサイズ $\mathrm{poly}(1/\varepsilon)$ の
リストを出力し、リストは確率 $1 - O(N/T_0)$ 以上で m を含む。
Dec の実行時間は $\mathrm{poly}(N, T_0)$

証明は補題 7.6, 7.7 より

補題 7.6 (Few control candidates)

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 $\mathrm{poly}(1/\varepsilon)$

補題 7.7 (Payload decoding succeeds)

正しい control info のもと、
確率 $1 - \beta_{\mathrm{payload}}$ 以上で正しくメッセージを復元
ただし、$\beta_{\mathrm{payload}} \le \beta_\pi + \beta_\Gamma(T_2) + N\,\beta_{\mathrm{PRC}}(T_2) \le (N+3)/T_0$
$T_2 = T_0 + O(N\,2^{\mathrm{poly}(1/\varepsilon)})$

control info が正しければ正しく復号

証明のスケッチ：

- oblivious error $\pi(e)$ に対して、復号失敗確率は $\beta_\pi \le 1/T_0$
- 擬似ランダムの時にも満たすことを示すには、
  正しい復号ができるかをテストできればよく、$O(N2^{\mathrm{poly}(1/\varepsilon)})$で可能

補題 7.8 (Hiding lemma)

任意の m, π, V に対して

$\mathrm{Enc}(m; \pi, V, \cdot) \approx^{\beta}_{\text{time-T}} U_N$

ただし、$\beta \le \beta_{\mathrm{Hide}}(T) = N\,\beta_{\mathrm{PRC}}(T') + \beta_{\Gamma}(T'),\ T' = T+N$

m, π, V を固定しても、time-T に対して符号語は擬似ランダム

証明はハイブリッド論法より
（符号語の擬似ランダムな部分を真正ランダムに順次置き換えていく）

系 7.10 (Oblivious error corollary)

任意の m, π, V, time-T 通信路 W に対して

$\mathrm{Err}_W(U_N) \approx^{\beta}_{\text{time-T'}} \mathrm{Err}_W(\mathrm{Enc}(m; \pi, V, \cdot))$

ただし、$\mathrm{Err}_W(x)$：通信路 W に x を入力したときの誤り

$\beta \le \beta_{\mathrm{Hide}}(T+T'+N)$

m, π, V を固定しても、
挿入される誤りは、ランダム入力の場合の誤りと識別できない

補題 7.6 の証明のために、補題 7.11, 7.12, 7.13 を示す

補題 7.11 (Good sampler; time-bounded version of Lemma 6.3)

任意の pN-bounded T_0-time 通信路 $T_0 > N, m, s_\pi$ に対して、
Samp の出力 V は good for e
w.p. $1 - (\beta_V + 2\,\beta_\Gamma(T_2) + 2N \cdot \beta_{PRC}(T_2)) \geq 1 - 2(N+3)/T_0$
ただし、$T_2 = T_0 + N \log N$
（確率は control info および通信路でとる）

標本器 Samp の出力は、高い確率で good for e

証明のスケッチ：
- 系 7.10 より、e と V はほぼ独立
- ランダムな場合、Samp の性質より、高い確率で V は good for e
- 擬似ランダムな場合にも good for e であることを示すには、
  good for e という性質が効率的にテストできることを示せばよい
  → e, V が与えられ、誤りの数を数えるだけなので、
  O(N log N) でテスト可能

補題 7.12 (Correct control blocks – list decoding version)

任意の e, V s.t. V が good for e に対して
PRC は、$\varepsilon n_{ctrl}/2 = \Theta(\varepsilon^2 N/\log(T_0))$ 個以上の control block で、
正しいものを含むサイズ L のシンボルを出力する

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

証明のスケッチ：
・符号は (δ,L)-list decodable であり、good for e なら誤り $p+\varepsilon < \delta$ だから

補題 7.13 (Bounding mistaken control blocks)

任意の m, e, $\omega = (s_\pi, s_\Gamma, s_V)$ に対して、
control info の候補数は $NL/b_{ctrl} = \Theta(NL/\log(T_0))$ 以下

control info の候補は (ブロック数)×(リストサイズ) 程度

証明のスケッチ：
・各 control block は $b_{ctrl} = \Theta(\log(T_0))$ ビット
→ ブロック数は $N/\log(T_0)$
→ 各々でサイズ L のリストなので、候補は $NL/\log(T_0)$

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、$L' \le 1/\varepsilon^{O(1)} = \mathrm{poly}(1/\varepsilon)$

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 C.1

任意の Enc, Dec でメッセージ空間サイズ |M| → ∞ 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 ≥ 1/2 – o(1)
2. ∀ 0 < v < 1/4, online space-(log(N)) 通信路が存在し、N(1/4 + v) 誤りで、復号誤り率 Ω(v)

証明のスケッチ：

- 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

- time-bounded 通信路のリスト復号 → 一意復号
  - p > 1/4 では簡単な計算で復号誤りを誘発
    - p < 1/4 なら不可能性を回避．他の方法は？
      - 通信路以上の Enc 計算？　送信者が秘密鍵保持？
      - PKI 設定では署名鍵を知らず、符号語を計算できないため回避
- time-bounded でモンテカルロ構成 → explicit
  - PRG は OWF 仮定で explicit にできる
  - PRC は可能か？
- space-bounded でモンテカルロ構成 → explicit
  - arXiv, ver.3 で構成（Nisan's PRG & log-space 帰着）
  - log-length PRC for log-space があれば explicit に
- よりシンプルな構成法は？

その他の不可能性

命題 [MPSW '05]

動的な符号方式において
(1) 送受信者が状態をもたないとき、
誤り割合 $p > 1/4$ で無視できない誤り率の通信路が存在
→ 状態が必要
(2) 送信者の入力すべてを通信路が知っているとき、
誤り割合 $p > 1/4$ で無視できない誤り率の通信路が存在
→ 送信者が秘密の入力をもつ必要

証明のスケッチ：
(1)ランダムメッセージ m, m' の符号語 x, x' を観察し、
$M(x,x') = x''$ s.t. $dist(x,x'') < n/4$, $dist(x',x'') < n/4$ とする
最初に x を、次に $dist(x,x') < n/2$ なら $M(x,x')$ を送れば w.p. 1/2 で間違い
(2)通信路は符号化をシミュレートできるため、2n+1個のメッセージのうち
$dist(x,x') < n/2$ である x, x' に対して $M(x,x')$ を出力する
最初の符号語が x でありかつ復号を失敗する確率は無視できない程度

標本可能な加法的誤りの訂正可能性

標本可能な加法的誤り

- 確率分布 Z が標本可能
  ⇔ 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

- 標本可能な分布 Z による
  加法的通信路 $C^Z : \{0,1\}^n \rightarrow \{0,1\}^n$
  - $C^Z(x) = x + z,\ z \sim Z$
  - 発生する誤りの数は制限しない
    - 誤り数がまばらだが規則性のある誤りを含む
  - 符号化方式は C^Z に依存して存在性を議論

  → どのような Z なら訂正可能か？

訂正可能性に関する考察

- H(Z) = 0 ならば簡単に訂正可能
  - 誤りの系列を知っているので
- H(Z) = n ならば訂正不可能
  - 受信系列は乱数
- H(Z) = n･H(p) のとき
  レート R > 1 - H(p) では訂正不可能
  - Z = BSC_p を計算できる場合

標本可能な Z の訂正可能性

- $H(Z) \leq n^{\varepsilon}$ で効率的に訂正できない Z が存在
  - 任意の $0 < \varepsilon < 1$
- 証明
  - 擬似乱数生成器 $G : \{0,1\}^m \rightarrow \{0,1\}^n$ に対し $Z = G(U_m)$ とする
  - $y = x + G(U_m)$ から x を効率的に復号できると、$G(U_m)$ が擬似ランダムであることに矛盾
  - 一方向性関数の存在を仮定した場合、任意の $0 < \varepsilon < 1$ について $m = n^{\varepsilon}$ とできる

シンドローム復号による訂正可能性

- $H(Z) = \omega(\log n)$ のとき
  レート $R > \Omega((\log n)/n)$ では
  シンドローム復号による効率的な訂正は不可能
  - あるオラクルへのアクセスを許すとき
- 証明
  - $H(Z) = \omega(\log n)$ で長さ $< n - \Omega(\log n)$ に効率的に圧縮できない標本可能分布が存在 (Wee '04)
    - あるオラクルへのアクスを許すとき
  - レート R で Z をシンドローム復号訂正可能
    $\Leftrightarrow$ Z を長さ $n(1 - R)$ に線形圧縮可能

訂正可能性のまとめ

- 標本可能な Z による加法的誤りの訂正可能性

H(Z)	訂正可能性
0	効率的に訂正可能
$\omega(\log n)$	レート $R > \Omega((\log n)/n)$ で シンドローム復号による 効率的な訂正は不可能
n^{ε} for $0 < \varepsilon < 1$	効率的に訂正不可能
$n \cdot H(p)$ for $0 < p < 1$	レート $R > 1 - H(p)$ では 訂正不可能
n	訂正不可能

今後の研究

- 訂正可能な Z の特徴付け
  - 訂正可能性に関する議論
- 無損失濃縮器との関係
  - 濃縮器: エントロピーを高くする関数
  - 平坦分布 Z に対する線形無損失濃縮器
    ⇔ 加法的誤り Z を線形関数で訂正可能
    - Cheraghchi (ISIT '09)
    - 復号の効率性は考えていない

→ 標本可能な Z に対する
無損失濃縮器の存在の可能性を探る

まとめ

- 計算能力制限通信路
- Gurswami & Smith (2010) の符号化方式
  - 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
  - p > 1/4 の一意復号の不可能
- 標本可能な加法的誤りの訂正可能性
  - 訂正限界の考察

オラクルアクセスについて

- $H(Z) = \omega(\log n)$ のとき
  レート $R > \Omega((\log n)/n)$ では
  シンドローム復号による効率的な訂正は不可能
  - あるオラクルへのアクセスを許すとき

- (a) から (b) のブラックボックス構成は存在しない

  (a) $H(Z) = \omega(\log n)$ の Z

  (b) Z をシンドローム復号で効率的に訂正する
  レート $R > \Omega((\log n)/n)$ の符号