# Correction of Samplable Additive Errors

Kenji Yasunaga

Kanazawa University, Japan

ISIT 2014 @ Honolulu, USA

# Two Well-Studied Channel Models

- Binary Symmetric Channels (BSC)
  - Each bit is flipped with constant probability p

- Worst-Case (or Adversarial) Channels
  - Only weight of error vectors is restricted

# Two Well-Studied Channel Models

- Binary Symmetric Channels (BSC)
  - Each bit is flipped with constant probability p

  → Introduce errors by simply flipping a coin
  ( Low-cost computation )

- Worst-Case (or Adversarial) Channels
  - Only weight of error vectors is restricted

  → Introduce errors based on the full knowledge
  ( High-cost computation )

# Computationally-Bounded Channels

- Introduced by Lipton (STACS 1994)
- Channels = Polynomial-time bounded algorithms (with error weight restriction)
  - Lie between BSC and Worst-Case Channels
- Related results
  - PKI setting [Micali et al. (TCC 2005)]
  - Without shared randomness or PKI [Guruswami, Smith (FOCS 2010)]

# This Work

- Introduce “Samplable Additive-Error Channels” as another intermediate model

- Investigate the possibilities and limitations for the reliable communication over this channel

# Samplable Distributions

- Distribution Z over $\{0,1\}^n$ is *samplable*

$\Leftrightarrow$ Exist probabilistic polynomial-time algorithm S s.t. $S(1^n)$ is distributed according to Z

- Related work on samplable distributions
  - Data compression [GS91, Wee04, TVZ05]
  - Randomness extraction [TV00, Vio11, DW12, DRV12, DPW14]

# Samplable Additive-Error Channels

- Additive channel $C^Z : \{0,1\}^n \rightarrow \{0,1\}^n$ is defined by a samplable distibution Z s.t.

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

  - Error vector z does not depend on x
  - Dist. Z does not depend on the code
    - Conversely, the code can depend on Z
  - Weight of z is not bounded

# Reasons for Introducing This Model

- Error distr. is same for every code/codeword
  → Error-correction problem is simple

- Error distribution is samplable
  → Computational constraints on error vectors
  can help error correction?

- Errors without weight restriction
  → High-weight errors are correctable
  if they have "nice structure" ?

What Z is correctable?

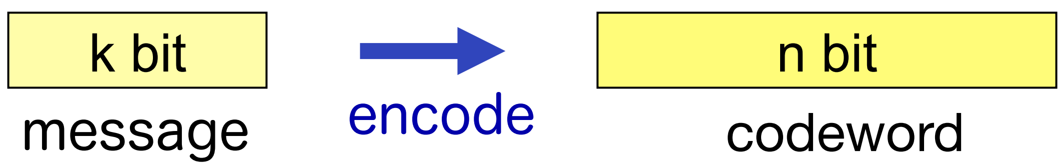
# Correctability of Samplable Additive Errors

- Use Shannon entropy H(Z) as criterion

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

  - $H(Z) \in [0, n]$ since Z is over $\{0,1\}^n$
  - $H(Z) = 0 \rightarrow$ easily correctable
  - $H(Z) = n \rightarrow$ uncorrectable

- What coding rate $R = k/n$ is achievable?



# Observation 1: Z is BSC

- Z can simulate BSC

**Theorem 1**

$\exists$ Z with $H(Z) = n \cdot h(p)$ ( $= \Omega(n)$ )
uncorrectable for $R > 1 - h(p)$

# Observation 2: Z is pseudo-random

- Z is the output of pseudo-random generator,
  → Not correctable by poly-time algorithms

**Theorem 2**

∃ Z with $H(Z) \leq n^{\varepsilon}$, $0 < \varepsilon < 1$,
uncorrectable by polynomial-time algorithms

- Assume OWF exists

# Observation 3: Z forms a linear subspace

**Theorem 3**

$\forall$ vector in linear space $Z \subseteq \{0,1\}^n$ of dim. m is correctable by a linear code with rate $R = 1 - m/n$

- $H(Z) = m$
- Decoding is efficient

- Proof sketch:
  - For basis $\{z_1, \ldots, z_m\}$,
    $\exists$ linear map $T : \{0,1\}^n \to \{0,1\}^m$ s.t.
    $T(z) = (a_1, \ldots, a_m)$ for $\forall$error vector $z = \Sigma_i a_i z_i$
  - The code = $\{ x : T(x) = 0 \}$

# Observation 4: Z is flat

**Theorem 4.1**

$\forall$ flat Z is correctable by linear code with rate $R < 1 - m/n - \Omega(\log(1/\varepsilon)/n)$ and error $\varepsilon > 0$

- $H(Z) = m$ ( $|\mathrm{supp}(Z)| = 2^m$ )
- Code is not explicitly given

- Proof sketch: Equivalence between linear code ensemble and linear lossless condenser [Cheraghchi (ISIT 2009)]

**Theorem 4.2**

$\forall$ flat Z is uncorrectable for rate $R \geq 1 - m/n + O(1/n)$ and error $\varepsilon < 1/2$ ( $|\mathrm{supp}(Z)| = 2^m$ )

- Proof sketch: Need to divide received word space $\{0,1\}^n$ into $2^{Rn}$ disjoint sets each of size $(1-\varepsilon)2^m$

# Observation 5: Uncorrectable Z with low entropy

**Theorem 5**

$\forall\ \omega(\log n) < m < n$, $\exists$ samplable Z with $H(Z) = m$ uncorrectable by "efficient syndrome decoding" for rate $R > \omega((\log n)/n)$

- Assume "oracle access" to some oracle

- Proof sketch:
  - $\exists$ samplable Z with $H(Z) = \omega(\log n)$ not efficiently compressible to length $< n - \omega(\log n)$ [Wee (CCC2004)] (Assuming "oracle access")
  - Z is compressible to length $n(1 - R)$ by linear function $\Leftrightarrow$ Z is correctable with rate R by syndrome decoding [Cair et al. 2004]

# Observation 6: Z is small-biased distribution

- Sample space $S \subseteq \{0,1\}^n$ is δ-biased
  $\Leftrightarrow \forall$ non-zero $a \in \{0,1\}^n$, $| E_{x\sim S}[ (-1)^{a \cdot x} ] | \leq \delta$

  Z is small-biased ⇔ Indistinguishable from uniform
  by linear functions

**Theorem 6**

∀ Z is uncorrectable for rate
$R > 1 - \Omega( \log(1/\delta) / n )$ with error $\varepsilon < 1/2$
if Z is uniform over δ-biased sample space S

- Proof sketch: Z can work as the key of the one-time pad if message has entropy [Dodis, Smith (TCC2005)]

**Corollary 6.1**

∃ Z with $H(Z) = m$ uncorrectable for
rate $R \geq 1 - m/n + O( (\log n)/n )$

# Correctability of Samplable Additive Errors (Summary)

| H(Z) | Correctability | Assump. | References |
|---|---|---|---|
| 0 | Efficiently correctable | | Trivial |
| $\omega(\log n)$ | Efficiently uncorrectable by syndrome decoding for $R > \omega( (\log n)/n )$ | Oracle access | Theorem 5 |
| $n^{\varepsilon}$ $(0 < \varepsilon < 1)$ | Efficiently uncorrectable | OWF | Theorem 2 |
| $n \cdot h(p)$ $(0 < p < 1)$ | Uncorrectable for $R > 1 - h(p)$ | | Theorem 1 |
| $0 \le m \le n$ | $\forall$ linear space Z of dim. m is correctable for $R \le 1 - m/n$ | | Theorem 3 |
| $0 \le m \le n$ | $\forall$ flat Z is correctable for $R \le 1 - m/n - \Omega(\log(1/\varepsilon)/n)$ | | Theorem 4.1 |
| $0 \le m \le n$ | $\forall$ flat Z is uncorrectable for $R > 1 - m/n + O(1/n)$ | | Theorem 4.2 |
| $0 \le m \le n$ | $\forall$ $\delta$-biased Z is uncorrectable for $R \ge 1 - m/n + O( (\log n)/n )$ | | Corollary 6.1 |
| n | Uncorrectable | | Trivial |

# Conclusions

## Our Results

- Introduce “Samplable Additive-Error Channels”
- Investigate the correctability

## Future Work

- Any practical situations captured by this model (with positive results)?
- More positive results (on more restricted Z?)
  - Log-space/constant-depth samplable Z
- Prove without assumptions (OWF, oracle access)
  - Or prove the assumption is necessary