Smoothing of codes, uniform distributions, and applications

Alexander Barg
(based on joint works with Madhura Pathegama)
University of Maryland

IITH, August 21, 2024

Outline

- Uniformly distributed point sets
- Smoothing binary codes: Asymptotic limits
- BSC wiretap channel, strong secrecy
- Linear hashing and randomness extraction
- Smoothing and hardness of Learning Parity with Noise (LPN)

References (Madhura Pathegama and A.B.):
arXiv:2308.11009, arXiv:2405.04406, arXiv:2408:03742

Uniformly distributed point sets

U.D. point sets approximate random subsets of the metric space

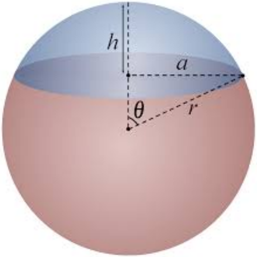
Classical theory of uniform distributions (H. WEYL, 1916) developed to measure errors in numerical (QMC) integration on $\mathcal{X} = [0,1)^d$

A set $Z_N = \{z_1, \dots, z_N\} \subset \mathcal{X}$ is used to approximate $\int_{\mathcal{X}} f(x)dx \approx \frac{1}{N}\sum_{i=1}^{N} f(z_i)$

Uniformly distributed sets of points are studied in multiple contexts. In information theory the most relevant are:

- "uniform" subsets in $\mathcal{H}_n := \{0,1\}^n$
- uniformly distributed points on the sphere $S^n(\mathbb{R})$

Uniform distributon on the sphere



Spherical cap $\mathrm{Cap}(x,t) = \{y \in S^d : (x,y) \geqslant t\};\ t = \cos\theta$

A *spherical code* is a finite set $Z_N \subset S^d$

A sequence of spherical codes $(Z_N)_N$ is called uniformly distributed if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{C(x,t)}(z_i) = \sigma(C(x,t)) \text{ for all } x \in S^d, t \in [-1,1]$$

BORODACHOV-HARDIN-SAFF, Discrete energy on rectifiable sets, Springer 2019

Uniform distribution in $\mathcal{H}_n := \{0,1\}^n$

- Code $\mathcal{C} \subset \mathcal{H}_n$
- View $\mathcal{C}$ as a distribution on the space, $P_{\mathcal{C}}(y) = \frac{1}{|\mathcal{C}|} \mathbb{1}_{\mathcal{C}}(y)$
- We could attempt a similar definition in the Hamming space: A subset (code) $\mathcal{C} \subset \mathcal{H}_n$ is (approximately) uniform if

$$d_{\mathrm{TV}}\left(\frac{\mathbb{1}_{\mathcal{C}}}{|\mathcal{C}|}, U_n\right) \leqslant \epsilon, \quad \text{where } U_n(z) = 1/2^n \text{ for all } z$$

- It is useful to think of “smoothed” code distributions. E.g., Bernoulli noise:

$$(T_{\beta_\delta}\mathcal{C})(x) := \frac{1}{|\mathcal{C}|} \sum_{y \in \mathcal{H}_n} \mathbb{1}_{\mathcal{C}}(x+y)\delta^{|y|}(1-\delta)^{n-|y|}$$

Uniformity and smoothing

Noise operator on $\mathcal{H}_n := \{0,1\}^n$

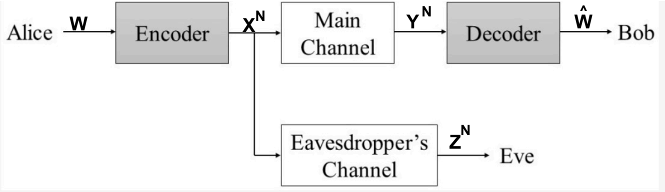
- Let $r : \mathcal{H}_n \to \mathbb{R}$ be a function
- $T_r f(x) = (r * f)(x) := \sum_{z \in \mathcal{H}_n} r(z) f(x-z)$
- Let $\mathcal{C} \subset \mathcal{H}_n$ be a code, $f_{\mathcal{C}} := \frac{\mathbb{1}_{\mathcal{C}}}{|\mathcal{C}|}$
- If r is a pmf, then $T_r f_{\mathcal{C}}$ is also a pmf
- Examples:
  - Bernoulli noise $(T_{\beta_\delta} \mathcal{C})(x) := \frac{1}{|\mathcal{C}|} \sum_{y \in \mathcal{H}_n} \mathbb{1}_{\mathcal{C}}(x+y) \delta^{|y|}(1-\delta)^{n-|y|}$
  - Ball noise $(T_{b_t} \mathcal{C})(x) := \frac{1}{V(t)} \sum_{y : |y| \leqslant t} \mathbb{1}_{\mathcal{C}}(x+y)$
- Clearly if $\mathcal{C}$ is of small size, $T_r f_{\mathcal{C}}$ cannot be close to uniform U_n, $U_n(x) = 2^{-n}$
- Let $|\mathcal{C}| = 2^{Rn}$. We are interested in conditions on $\mathcal{C}$ or R for $T_r f_{\mathcal{C}}$ to be close to U_n and applications of this property

A. Samorodnitsky 2016, O. Ordentlich–Y. Polyanskiy 2018, T. Debris-Alazard e.a., 2023

Related work in Information theory and Cryptography

- **Channel resolvability**
  Given $P^n_{Y|X} : \mathcal{X}^n \to \mathcal{Y}^n$, what is $\min |\mathcal{C}|$ such that $d_{\mathrm{TV}}(\frac{1}{|\mathcal{C}|}\mathbb{1}_{\mathcal{C}} \circ P^n_{Y|X}, Q^n_Y)$ is small?
  HAN-VERDÚ '94, HAYASHI '06, YU-TAN '18, PATHEGAMA-B. '23, YU '23
- **Strong coordination**
  BLOCH-LANEMAN '13, CHOU E.A. '18, COVER-PERMUTER '07, CUFF E.A., '10-'13
- **Entropy of noisy functions and BSC decoding**
  SAMORODNITSKY '16, ORDENTLICH-POLYANSKIY '18
  Decoding: HĄZŁA E.A. '21, SPRUMONT-RAO '23, PATHEGAMA-B. '23
- **Wiretap Channels**
  Discrete: HAYASHI '06, YU-TAN '19, PATHEGAMA-B. '23
  Gaussian: BELFIORE-OGGIER '10, LUZZI E.A. '23
- **Linear hashing**
  General results: IMPAGLIAZZO E.A. '89, HAYASHI-TAN '16-'18
  Linear hashing: PATHEGAMA-B. '24, YAN-LING '24
- **WDP-to-LPN (worst-to-average) case reductions**
  BRAKERSKY '19, DEBRIS-ALAZARD E.A. '22, DEBRIS-ALAZARD–RESCH '22, PATHEGAMA-B. '24

Application of uniformity: The Wiretap Channel



Goals:

(1) Main receiver B can recover W with high probability,

(2) $I(Z^N, W) \leqslant \epsilon$

- ▷ Strong secrecy $I(Z^N, W) \to 0$
- ▷ Weak secrecy $\frac{1}{N} I(Z^N, W) \to 0$

Capacity of the BSC wiretap channel $\mathscr{C} = H(\delta_e) - H(\delta_b)$

Smoothing of binary codes

Definition: A sequence of codes $(\mathcal{C}_n)_n$ is asymptotically perfectly D_p-smoothable with respect to kernels $(r_n)_n$ if

$$\lim_{n\to\infty} D_p(T_{r_n} f_{\mathcal{C}_n} \| U_n) = 0, \quad 0 \leqslant p \leqslant \infty.$$

Here $D_p(P\|Q)$ is the Rényi divergence of order p :

$$D_p(P\|Q) = \frac{1}{p-1} \log \left(\mathbb{E}_P \frac{dP}{dQ} \right)^{p-1} = \frac{1}{p-1} \log \sum_i P_i^p Q_i^{-(p-1)}$$

For $p = 1$ this is the KL divergence.

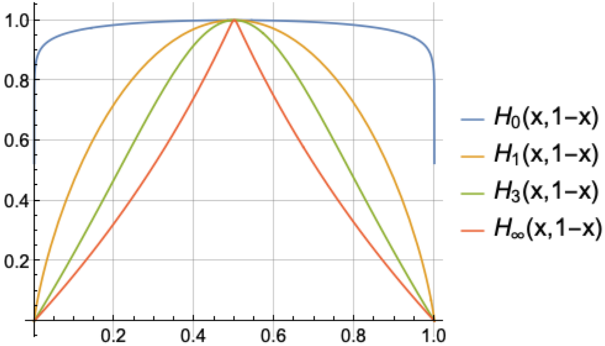
If Q is uniform, D_p is related to the Rényi entropy:

$$D_p(P\|U_n) = n - H_p(P)$$

$$H_p(P) = \frac{1}{1-p} \log \left(\sum_i P_i^p \right)$$

For $P \sim \text{Ber}(\delta)$, $H_p(P) = H_p(\delta, 1-\delta) = \frac{1}{1-p} \log(\delta^p + (1-\delta)^p)$

Rényi entropy



Our results

- Finding the threshold rates for binary codes to achieve asymptotically perfect smoothing under Bernoulli noise
- Threshold rates for ball noise
- Bounds on the achievable rate of Reed-Muller codes on the wiretap channel

Remark: Recent results established that Reed-Muller codes achieve capacity of the binary-input symmetric channels

- BEC: Kudekar, Kumar, Mondelli, Pfister, Şaşŏglu, Urbanke, ’17
- BMS: Abbe-Sandon, ’23

The property of achieving BEC capacity is the reason that RM codes figure in our examples

Threshold for perfect smoothing

Definition: Let $(r_n)_n$ be a sequence of noise kernels. Rate R is achievable for perfect D_p-smoothing if there exists a sequence of codes $(\mathcal{C}_n)_n$ such that $R(\mathcal{C}_n) \to R$ as $n \to \infty$ and $(\mathcal{C}_n)_n$ is perfectly D_p-smoothable.

Define the capacity of perfect smoothing S_p^r as inf(achievable rates) for D_p smoothing w.r.t. $(r_n)_n$.

This is a particular case of the general problem of channel resolvability (T.-S. Han and S. Verdú, 1983)

The current state of the art for Bernoulli kernels is given in the next theorem.

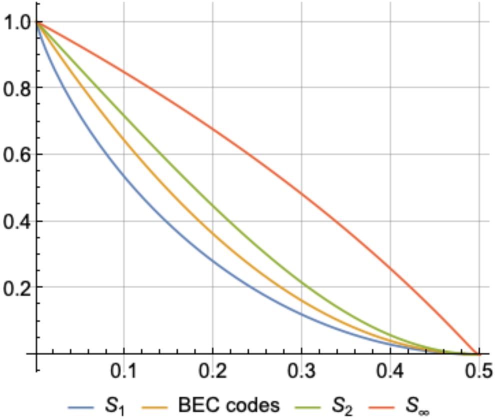
Theorem

$$S_p^{\beta_\delta} = \begin{cases} 0 & \text{if } p = 0 \\ 1 - H(\delta) & \text{if } p \in (0, 1] \\ 1 - H_p(\delta) & \text{if } p \in (1, \infty], \end{cases}$$

where $H_p(\delta) = \frac{1}{1-p} \log(\delta^p + (1-\delta)^p)$ *is the Rényi entropy of order* p.

Here the results for $p \in [0, 2] \cup \{\infty\}$ are due to L. Yu and V.Y.F. Tan, 2018, while the cases $2 < p < \infty$ are new.

Smoothing capacity plots



Threshold rates for perfect asymptotic smoothing (top to bottom)

- ▷ $S_\infty^{\beta_\delta}$
- ▷ $S_2^{\beta_\delta}$
- ▷ D_1 smoothing threshold for (duals of) codes achieving BEC capacity
- ▷ D_p smoothing capacity, $p \in (0, 1]$ = Shannon capacity of BSC(δ)

Remarks on the proof

To show that this rate is attainable, we use random coding.

For the lower bound on R we use the equality

$$D_p(f_{\mathcal{C}} \| U_n) = \frac{p}{p-1} \log \|2^n f_{\mathcal{C}}\|_p.$$

Then use induction to establish the bound for all rational p and a density argument to prove it for all real p.

Which (explicit) codes achieve perfect smoothing?

- Achieving BSC capacity does not imply perfect smoothing
- Polar codes achieve perfect smoothing at smoothing capacity
- RM codes (and other BEC capacity achieving codes) achieve a certain rate $R > S_1$

Good codes for the erasure channel and perfect smoothing

- A. SAMORODNITSKY 2016-'21 proved general bounds for the entropy of noisy functions on $\mathcal{H}_n$
- Using these results, HĄZŁA-SAMORODNITSKY-SBERLO '21 connected performance of codes on the BEC and on the BSC
- Extending these ideas, we derive smoothing properties from erasure correction properties of codes
- Key statement: Bernoulli smoothing is "bounded above" by BEC performance.
  Take a linear code $\mathcal{C}$, let $X_{\mathcal{C}^\perp}$ be a random codeword of $\mathcal{C}^\perp$. Let $Y_{X_{\mathcal{C}}^\perp,\lambda}$ be the output of BEC(λ) for the input $X_{\mathcal{C}^\perp}$. Then

$$D_p(T_\delta f_{\mathcal{C}} \| U_n) \leqslant H(X_{\mathcal{C}}^\perp | Y_{X_{\mathcal{C}}^\perp,\lambda}),$$

  where $\lambda = (1-2\delta)^2$ for $p=1$ and $\lambda = 1 - H_p(\delta)$ for $p \geqslant 2$

Good codes for the erasure channel and perfect smoothing

Using this lemma, we show that certain ***explicit*** code families attain perfect smoothing

Theorem

Let $(\mathcal{C}_n)_n$ be a sequence of linear codes with rate $R_n \to R$. Suppose that the dual sequence $(\mathcal{C}_n^\perp)_n$ achieves capacity of the BEC(λ) with $\lambda = R$. Assume that $d(\mathcal{C}_n^\perp) = \omega(\log(n))$ and $R > (1-2\delta)^2$, then

$$D(T_\delta f_{\mathcal{C}_n} \| U_n) \to 0 \quad \text{as} \quad n \to \infty.$$

Let $p \in \{2, 3, \ldots, \infty\}$. If $R > 1 - h_p(\delta)$, then

$$D_p(T_\delta f_{\mathcal{C}_n} \| U_n) \to 0 \quad \text{as} \quad n \to \infty$$

In particular, the sequence $\mathcal{C}_n$ achieves D_p-smoothing capacity $S_p^{\beta_\delta}$ for $p \in \{2, 3, \ldots, \infty\}$.

Example: Reed-Muller codes

Smoothing for the wiretap channel: Main ideas

A. Wyner ('75) suggested the following scheme for transmission over the wiretap channel:

Let $\mathcal{C}_0 \subset \mathcal{C}_1 \subset \mathcal{H}_n$ be linear codes. Encode messages into cosets $\mathcal{C}_1/\mathcal{C}_0$; transmit a random vector from the coset.

Reliability: $P(\hat{W} \neq W) \to 0$; Secrecy: $T_{\delta_e} f_{\mathcal{C}_0}$ approaches U_n

$$P_{Z|M=m}(z) = P_{X_{C_m+W}}(z) = P_{C_m} * P_W(z) = P_{C_0+c_m} * P_W(z) = P_{C_0} * P_W(z+c_m)$$

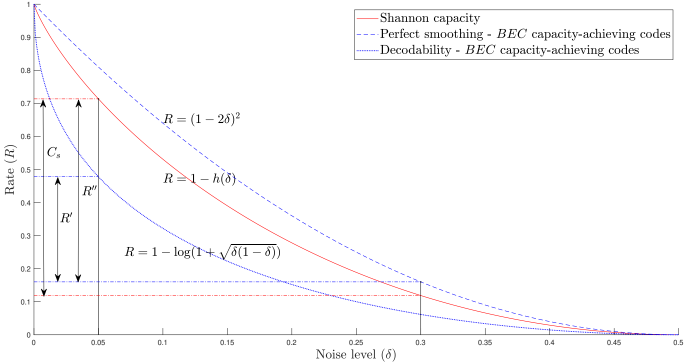
Lemma: Under Wyner's coding scheme, if

$$D(T_{\delta_e} f_{\mathcal{C}_0} \| U_n) < \epsilon, \text{ then } I(M;Z) < \epsilon.$$

If a code attains BEC capacity, then its dual is smoothed by Bernoulli noise

$\Rightarrow$ Duals of good codes for the BEC support strong security on the BSC wiretap channel

Transmission rates for the wiretap channel



Achievable rates in BSC wiretap channel with BEC capacity-achieving codes.

For instance, take $\delta_m = 0.05$; $\delta_e = 0.3$. Then

$$\mathcal{C}_s = H(0.3) - H(0.05) = 0.5949$$

For Reed-Muller codes,

$$R'' = 0.5536$$

Open problem about RM codes

Do nested sequences of RM codes attain secrecy capacity of the BSC wiretap channel under strong security? This does not follow directly from either

- ▷ Abbe-Sandon's proof of the BSC capacity result for RM codes
- ▷ The approach via classical-quantum duality (RENES '18; RENGASWAMY E.A. '21)

Remark: MAHDAVIFAR-VARDY ('11) showed that polar codes attain the BSC wiretap capacity with weak secrecy; later GÜLCÜ-B. ('16) showed strong secrecy

Universal hash families

- Given a source $Z \sim P_Z$ with unknown distribution P_Z on $\mathbb{F}_q^n$
- $\mathscr{F}_{n,m} := \{f : \mathbb{F}_2^n \to \mathbb{F}_2^m\}$ forms a universal hash family (UHF) if

$$\Pr_{f \sim \mathscr{F}}(f(u) = f(v)) \leqslant \frac{1}{2^m} \quad \forall u \neq v$$

LHL (classic): Let $f \sim \mathscr{F}_{n,m}$. If $m \leqslant H_\infty(Z) - 2\log(1/\epsilon)$, then

$$d_{\text{TV}}(P_{f(Z),f}, P_{U_m} \times P_f) \leqslant \epsilon/2$$

- Since $d_{\text{TV}}(P,Q) = \frac{q^n}{2}\|P - Q\|_1$, we can rewrite the LHL as:

$$\mathbb{E}_{f\sim\mathfrak{F}}\|q^m P_{f(Z)} - 1\|_1 \leqslant \epsilon$$

- (Strengthened LHL) (BENNET E.A., '95) If $m \leqslant H_2(Z) - \log(1/\epsilon)$, then

$$\mathbb{E}_{f\sim\mathscr{F}}[D(P_{f(Z)}\|P_{U_m})] \leqslant \frac{\epsilon}{\ln 2}$$

- We establish similar results using p-norms and linear hashing

Linear Hashing

- Given a source Z on $\mathbb{F}_q^n$, $Z \sim P_Z$
- The set of linear codes $\mathscr{C} := \{\mathcal{C}[n, n-m]_q\}$;
- Let H be a parity-check matrix of $\mathcal{C} \in \mathscr{C}$

$P_Z \rightsquigarrow P_{\mathrm{HZ}}$ – smoothing as hashing

Theorem

Let $\epsilon > 0$ and let $p \geqslant 2$ be an integer. If Z is a random vector from $\mathbb{F}_q^n$ with $Z \sim P_Z$ and $m \leqslant H_p(Z) - p - \log_q(1/\epsilon)$, then

$$\mathbb{E}_{\mathcal{C}\sim\mathscr{C}}[D_p(P_{\mathrm{HZ}} \| P_{U_m})] \leqslant \frac{p\epsilon}{(p-1)\ln q}$$

- Rewriting the conclusion:

$$\mathbb{E}_{\mathcal{C}\sim\mathscr{C}} \|q^m P_{\mathrm{HZ}} - 1\|_p \leqslant 2^{1-1/p}((1+\epsilon)^p - 1)^{1/p}$$

  This says that P_{HZ} is almost independent of the code. Measuring uniformity by l_p rather than d_{TV} is a stronger guarantee.
- Previous works (Hayashi-Tan, '16-'18) proved p-uniformity guarantees for memoryless sources Z

Remarks on the proof

The technical claim behind this theorem can be stated as follows.

Theorem

Let Z be a random vector in $\mathbb{F}_q^n$. Let $\mathscr{C}$ be the set of all $[n,k]_q$ linear codes. Then for all natural $p \geqslant 2$,

$$\mathbb{E}_{\mathcal{C}\sim\mathscr{C}}[\|q^n P_{X_{\mathcal{C}}+Z}\|_p^p] \leqslant \sum_{d=0}^{p} \binom{p}{d} q^{(p-d)(d+n-k-H_p(Z))},$$

where $X_{\mathcal{C}}$ is a uniform random codeword of $\mathcal{C}$.

Bernoulli sources and hashing with Reed-Muller matrices

Theorem

Let $R \in (0,1)$ and let $\mathcal{C}_n$ be a sequence of RM codes whose rate R_n approaches R. Let H_n be the parity check matrix of $\mathcal{C}_n$ and let Z_n be a binary vector formed of independent Bernoulli(δ) random bits. If $R > 1 - h_p(\delta)$, then

$$\lim_{n\to\infty} D_p(P_{\mathsf{H}_n Z_n} \| P_{U_{n(1-R_n)}}) = 0, \quad p \in \{2, \dots, \infty\}$$

If $p = 1$ and $R > (1-2\delta)^2$, then

$$\lim_{n\to\infty} D(P_{\mathsf{H}_n Z_n} \| P_{U_{n(1-R_n)}}) = 0.$$

The proof follows from the results on threshold rates for smoothing capacity

Learning Parity with Noise

LPN problem $\mathrm{LPN}(k, \delta, N, \alpha)$ with *noise rate* $\delta \in (0, 1/2)$, *sample complexity* N, and *success probability* α:
Given a collection of samples $(a_i, a_i^\intercal m + b_i)_{i=1}^N$, $a_i, m \in \mathbb{F}_2^k, b_i \in \mathbb{F}_2$, where

(1) $m \sim P_{U_k}$ is fixed across all samples

(2) $(a_i, b_i) \sim P_{U_k} P_{\mathrm{Ber}(\delta)}$ are chosen independently for each sample,

find $\hat{m}$ with $\Pr(\hat{m} = m) \geqslant \alpha$.

LPN underlies several cryptographic primitives:

- symmetric encryption HOPPER-BLUM '01, JUELS-WEIS '05
- public key cryptography ALEKHNOVICH '03
- collision-resistant hashing BRAKERSKI E.A. '19, YU E.A., '19

Is LPN computationally hard?

WDP-to-LPN reduction

The worst-case decoding problem $\mathrm{WDP}(n, k, w)$ is defined as follows: Given

(1) a matrix $\mathsf{G} \in \mathbb{F}_2^{k \times n}$

(2) a vector $y \in \mathbb{F}_2^n$ of the form $y = \mathsf{G}^\intercal m' + e'$ for some $m' \in \mathbb{F}_2^k$ and $e' \in \mathbb{F}_2^n$ with $|e'| = w$,

find m such that $y = G^\intercal m + e$ for some $e \in \mathbb{F}_2^n$ with $|e| = w$.

Finding an efficient solver for LPN would amount to constructing an efficient probabilistic decoder for linear codes.

Reduction

An LPN solver $\mathcal{A}$ can be converted into a decoder

G is $k \times n$ is a generator matrix of $\mathcal{C}$; noisy codeword $G^\intercal m + e$, where $|e| = w$

- Sample $m' \xleftarrow{U_k} \mathbb{F}_2^k$.
- Find P on $\mathbb{F}_2^n$ and $\varepsilon > 0$ such that

$$d_{\mathrm{TV}}(P_{\mathrm{G}Z, e^\intercal Z}, P_{U_k} P_{\mathrm{Ber}(\delta)}) \leqslant \varepsilon \quad (Z \sim P)$$

- $\{Z_i\}_{i=1}^N \leftarrow P$; $a_i = \mathrm{G}Z_i$ $b_i = e^\intercal Z_i$, $i = 1, \dots, N$

$$Z_i^\intercal(\mathrm{G}^\intercal m' + \mathrm{G}^\intercal m + e) = a_i^\intercal(m + m') + b_i.$$

- $\mathcal{A} \leftarrow (a_i, a_i^\intercal(m + m') + b_i)_{i=1}^N$
- If $N\varepsilon < \alpha$, Algorithm $\mathcal{A}$ outputs $m + m'$ with success probability at least $\alpha - N\varepsilon$ in time T.
- In conclusion, with probability $\alpha - N\varepsilon$ the message m is found in time $T \cdot \mathrm{poly}(n, k)$.

Thus, we need ***fast smoothing***: for large N, decoding error rate is large

Meaningful reductions

- For $\xi \sim \mathrm{Ber}(\delta)$ with $P_\xi(1) = \delta$, $\mathrm{bias}(\xi) := \frac{1}{2} - \delta$
- $\mathrm{bias}(\xi) = o(1/\operatorname{poly}(k))$ is too small; $\mathrm{bias}(\xi) = \Omega(1/\operatorname{poly}(k))$ supports symmetric cryptography
- $d_{\mathrm{TV}}(P_{\mathsf{GZ},e^\intercal Z}, P_{U_k} P_{\mathrm{Ber}(\delta)}) < \epsilon$ (fast smoothing)

In summary, a meaningful reduction should satisfy

$$d_{\mathrm{TV}}(P_{\mathsf{GZ},}, P_{U_k}) < d_{\mathrm{TV}}(P_{\mathsf{GZ},e^\intercal Z}, P_{U_k} P_{\mathrm{Ber}(\delta)}) < \alpha/N, \tag{1a}$$

$$\mathrm{bias}(e^\intercal Z) = \Omega\Big(\frac{1}{\operatorname{poly}(k)}\Big). \tag{1b}$$

Brakerski e.a. '19 showed that meaningful reductions are possible under the assumption of vanishing code rate k/n

Yu-Zhang '22 and Debriz-Alazard & Resch, '22 show similar results with additional assumptions on codes and smoothing distributions.

In particular, it was not clear whether meaningful reductions with constant rate were possible

Our results

We show that for constant-rate codes, the necessary conditions are violated, so an efficient reduction is generally not possible.

Theorem: Let $(\mathcal{C}_n, n = 1, 2, \dots)$ be a sequence of $[n, k]$ linear codes of increasing length and let $\frac{k}{n} \to R > 0$ and $\frac{d^\perp}{n} \to \delta^\perp > 0$.

For any sequence of random vectors Z defined on $\mathbb{F}_2^n$, there exists a sequence of vectors e with $|e|/n \to \omega$ such that the following holds true:

- If $d_{\text{TV}}(P_{\mathsf{G}_n Z}, P_{U_k}) = o(\frac{1}{\text{poly}(k)})$, then $\text{bias}(e^\intercal Z) = o(\frac{1}{\text{poly}(k)})$
- If $d_{\text{TV}}(P_{\mathsf{G}_n Z}, P_{U_k}) = 2^{-\Omega(k)}$, then $\text{bias}(e^\intercal Z) = 2^{-\Omega(k)}$,

where G_n is the generator matrix of $\mathcal{C}_n$.

Remark: Slow smoothing supports reduction

Previously we assumed $d_{\mathrm{TV}}(P_{\mathsf{G}Z, e^\intercal Z}, P_{U_k} P_{\mathrm{Ber}(\delta)}) = o(1/\operatorname{poly}(k))$. Relaxing this allows reduction, but degrades the decoder's performance.

Theorem:
Let $R \in (0,1)$, $\omega \in (0, 1/2)$, and $l \in \mathbb{N}$. Let $\mathcal{C}_n$ be a sequence of $[n,k]$ linear codes of increasing length n such that $k/n \to R$. Let G_n be a generator matrix of $\mathcal{C}_n$ and let $e \in \mathbb{F}_2^n$ be a vector satisfying $|e| = \lfloor \omega n \rfloor$. Then there exists a sequence of distributions $(P_Z)_n$ satisfying the following conditions:

(i) $d_{\mathrm{TV}}(P_{\mathsf{G}_n Z, e^\intercal Z}, P_{U_k} P_{e^\intercal Z}) = O(k^{-l})$

(ii) $\mathrm{bias}(e^\intercal Z) = \Omega(k^{-l})$.

Thank you!