WORST AND AVERAGE CASE HARDNESS OF DECODING VIA SMOOTHING BOUNDS

PKC '25

Thomas Debris-Alazard and Nicolas Resch

May 13, 2025

Inria, École Polytechnique

Code-Based Cryptography:

Hardness of decoding a random code

$\longrightarrow$ Average-case problem!

SELF-REDUCIBILITY

Self-reducibility:

Is decoding in average $\big($random code$\big)$ as hard as decoding **all** codes?

$\longrightarrow$ **Worst-to-average case reduction**

- [BLVW19]: *Worst-case hardness for LPN and cryptographic hashing via code smoothing*, Eurocrypt '19
- [YZ21]: *Smoothing out binary linear codes and worst-case sub-exponential hardness for LPN*, CRYPTO '21

$\longrightarrow$ Both papers rely on **smoothing bounds**

OUR CONTRIBUTIONS

- We developed general smoothing bounds to offer a greater degree of freedom when compared to [BLVW19,YZ21]
- We showed some inherent limitation of the worst-to-average case reduction [BLVW19,YZ21]
- We failed to improved parameters of [BLVW19,YZ21] by relying on stronger upper bounds

DECODING A RANDOM CODE

$\big($Binary linear$\big)$ code:

A $[n, k]$-code $\mathcal{C}$ is a subspace of $\mathbb{F}_2^n$

n : length k : dimension

- Basis/Generator matrix rep.: rows of $\mathbf{A} \in \mathbb{F}_2^{k \times n}$ form a basis,

$$\mathcal{C} = \left\{ \mathbf{s}\mathbf{A} : \ \mathbf{s} \in \mathbb{F}_2^k \right\}$$

- Knapsack/Parity-check rep.: $\mathcal{C}$ as null space of a full-rank $\mathbf{H} \in \mathbb{F}_2^{(n-k) \times n}$,

$$\mathcal{C} = \left\{ \mathbf{c} \in \mathbb{F}_2^n : \ \mathbf{H}\mathbf{c}^\top = \mathbf{0} \right\}$$

Hamming weight:

$$\forall \mathbf{x} \in \mathbb{F}_2^n, \quad |\mathbf{x}| \stackrel{\text{def}}{=} \{ i \in [1, n] : \ x_i \neq 0 \}$$

SOME NOTATION

- $X \leftarrow \mathcal{S}$: X picked uniformly at random in $\mathcal{S}$
- $\mathbf{e} \leftarrow \mathrm{Ber}(p)^{\otimes n}$: the e_i's are **independent** and $\mathbb{P}(e_i = x) = \begin{cases} 1-p & \text{if } x = 0 \\ p & \text{if } x = 1 \end{cases}$

$\mathrm{Ber}(p)^{\otimes n}$ concentrates over words of Hamming weight $\approx np$

$\mathrm{DP}(n,k,t)$ primal rep.

- **Input:** $(\mathbf{A}, \mathbf{y} \stackrel{\text{def}}{=} \mathbf{sA}+\mathbf{e})$ where $\mathbf{A} \leftarrow \mathbb{F}_2^{k\times n}$, $\mathbf{s} \leftarrow \mathbb{F}_2^k$ and $\mathbf{e} \leftarrow \mathrm{Ber}(t/n)^{\otimes n}$
- **Aim:** recover $\mathbf{s} \in \mathbb{F}_2^k$

Equivalent formulation:

- Syndrome decoding problem: given $\mathbf{H}$ and $\mathbf{He}^\top$ recover $\mathbf{e}\ldots$

Learning Parity with Noise (LPN): easier than $\mathrm{DP}(n,k,t)$

- n $\left(\text{number of samples}\right)$ can be chosen as large as we want

WORST TO AVERAGE CASE REDUCTION

We are given a fixed instance

$(\mathbf{G}, \mathbf{x}\mathbf{G} + \mathbf{r})$ where the Hamming weight of $\mathbf{r}$ is w

and we want to recover $\mathbf{r}$.

But, we have an algorithm $\mathcal{A}$ solving DP with probability ε

$\Big($it does not solve for any $\mathbf{A}, \mathbf{s}, \mathbf{e}\Big)$

$$\mathbb{P}_{\mathbf{A},\mathbf{s},\mathbf{e}}\left(\mathcal{A}(\mathbf{A}, \mathbf{s}\mathbf{A} + \mathbf{e}) = \mathbf{e}\right) = \varepsilon$$

$$\langle \mathbf{a}, \mathbf{b} \rangle \stackrel{\text{def}}{=} \sum a_i b_i$$

Key-idea:

From $(\mathbf{G}, \mathbf{y} \stackrel{\text{def}}{=} \mathbf{xG} + \mathbf{r})$, build a "uniform" instance that will being fed to $\mathcal{A}$

1. $\mathbf{e} \leftarrow \mathcal{D}$ (distribution which can be chosen as we want)
2. Compute,

$$\langle \mathbf{y}, \mathbf{e} \rangle = \langle \mathbf{xG}, \mathbf{e} \rangle + \langle \mathbf{r}, \mathbf{e} \rangle = \langle \underbrace{\mathbf{x}}_{\text{secret}}, \mathbf{eG}^\top \rangle + \underbrace{\langle \mathbf{r}, \mathbf{e} \rangle}_{\text{noise}}$$

To build a truly Decoding instance:

- We would like $\mathbf{eG}^\top$ "very close to uniform"
- Need to analyze noise distribution $e = \langle \mathbf{r}, \mathbf{e} \rangle$ (the easy part)

$$\langle \mathbf{y}, \mathbf{e} \rangle = \langle \mathbf{x}\mathbf{G}, \mathbf{e} \rangle + \langle \mathbf{r}, \mathbf{e} \rangle = \langle \underbrace{\mathbf{x}}_{\text{secret}}, \mathbf{e}\mathbf{G}^\top \rangle + \underbrace{\langle \mathbf{r}, \mathbf{e} \rangle}_{\text{noise}}$$

$\longrightarrow$ We want $\mathbf{e}\mathbf{G}^\top$ “very close to uniform”

A first approach:

Choose each bit of $\mathbf{e}$ with probability 1/2, then $\mathbf{e}\mathbf{G}^\top$ is uniform

But, doing this is useless: $\langle \mathbf{r}, \mathbf{e} \rangle$ will be a uniform noise...

Therefore, impossible to solve $\left(\mathbf{e}\mathbf{G}^\top, \langle \mathbf{x}, \mathbf{e}\mathbf{G}^\top \rangle + \underbrace{\langle \mathbf{r}, \mathbf{e} \rangle}_{\text{noise}} \right)$

$\longrightarrow$ We need to carefully choose the noise!

CLOSENESS: STATISTICAL DISTANCE

Statistical distance:

Given two random variables X, Y with distributions f, g,

$$\Delta(\mathsf{X}, \mathsf{Y}) = \Delta(f, g) = \frac{1}{2} \sum_{x} |f(x) - g(x)|$$

If an algorithm succeeds with inputs X *and probability* ε*, then it succeeds given* Y *with probability* $\geq \varepsilon - \Delta(\mathsf{X}, \mathsf{Y})$

True Random decoding sample

1. We want the following to be small:

$$\alpha \stackrel{\text{def}}{=} \Delta\Big((\mathbf{e}\mathbf{G}^\top, \langle \mathbf{x}, \mathbf{e}\mathbf{G}^\top\rangle + \langle \mathbf{r}, \mathbf{e}\rangle), (\underbrace{\mathbf{a}}_{\text{uniform}}, \langle \mathbf{x}, \mathbf{a}\rangle + \underbrace{e}_{\text{same distrib as } \langle \mathbf{r}, \mathbf{e}\rangle})\Big)$$

2. Then we feed $(\mathbf{e}\mathbf{G}^\top, \langle \mathbf{x}, \mathbf{e}\mathbf{G}^\top\rangle + \langle \mathbf{r}, \mathbf{e}\rangle)$ to the Decoding-solver $\mathcal{A}$ with prob. ε

3. If we give n samples to $\mathcal{A}$, it will recover $\mathbf{x}$ with prob. $\geq \varepsilon - n\alpha$

A simplification for the talk:

We will target $\Delta\left(\mathbf{e}\mathbf{G}^\top, \underbrace{\mathbf{a}}_{\text{uniform}}\right)$ to be small when $\mathbf{G}$ is fixed but $\mathbf{e}$ random variable

A GEOMETRICAL INTERPRETATION: PRIMAL REPRESENTATION

$$\text{Aim: } \Delta\left(\mathbf{e}\mathbf{G}^{\top}, \underbrace{\mathbf{a}}_{\text{uniform}}\right) \text{ small}$$

Which object is $\mathbf{e}\mathbf{G}^{\top}$*?*

$\longrightarrow$ Let us take the code $\mathcal{C} \subseteq \mathbb{F}_2^n$ point of view

$$\mathcal{C} = \left\{\mathbf{c} : \ \mathbf{c}\mathbf{G}^{\top} = \mathbf{0}\right\}$$

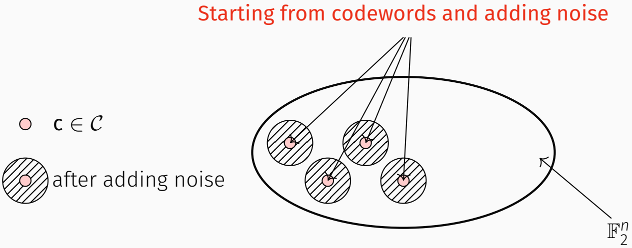
$\mathbf{e}\mathbf{G}^{\top}$ *defines a coset of* $\mathcal{C}$

Primal representation:

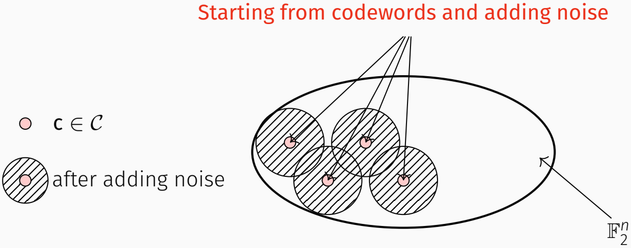
$$\mathbf{e}\mathbf{G}^{\top} \text{ uniform} \iff \text{uniform in } \mathbb{F}_2^n/\mathcal{C}, \textit{ i.e.} \text{ uniform modulo } \mathcal{C}$$

$$\mathbf{e}\mathbf{G}^{\top} \text{ uniform for } \mathbf{e} \leftarrow \mathcal{D} \iff \mathbf{c}+\mathbf{e} \text{ uniform in } \mathbb{F}_2^n \text{ where } \mathbf{c} \leftarrow \mathcal{C} \text{ and } \mathbf{e} \leftarrow \mathcal{D}$$

$\mathbf{c} + \mathbf{e}$ uniform in $\mathbb{F}_2^n$ where $\mathbf{c} \leftarrow \mathcal{C}$ and $\mathbf{e} \leftarrow \mathcal{D}$



$\mathbf{c} + \mathbf{e}$ uniform in $\mathbb{F}_2^n$ where $\mathbf{c} \leftarrow \mathcal{C}$ and $\mathbf{e} \leftarrow \mathcal{D}$



Optimal ball radius: Gilbert-Varshamov radius

$$t_{\mathrm{GV}}: \quad \text{smallest } t \text{ such that } \binom{n}{t} \cdot \sharp\mathcal{C} \geq 2^n = \sharp\mathbb{F}_2^n$$

SMOOTHING PARAMETER

OUR RESULT

Notation:

- unif: uniform distribution of $\mathbb{F}_2^n$
- $1_{\mathcal{C}}$: indicator function of $\mathcal{C} = \left\{\mathbf{c} : \mathbf{c}\mathbf{G}^\top = \mathbf{0}\right\}$
- Convolution, $f \star g(\mathbf{x}) \stackrel{\text{def}}{=} \sum_{\mathbf{y}\in\mathbb{F}_2^n} f(\mathbf{y})g(\mathbf{x}-\mathbf{y})$

If $\mathbf{X} \leftarrow f$ and $\mathbf{Y} \leftarrow g$ are independent, then $\mathbf{X}+\mathbf{Y} \leftarrow f \star g$

Our upper bound:

$$\Delta\left(\frac{1_{\mathcal{C}}}{\sharp\mathcal{C}} \star f, \text{unif}\right) \leq \sqrt{2^n \sum_{u>0} N_u(\mathcal{C}^\perp)|\widehat{f}(u)|^2}$$

where $\mathcal{C}^\perp$ dual code and $N_u(\mathcal{C}^\perp) = \sharp\left\{\mathbf{c}^\perp \in \mathcal{C}^\perp : |\mathbf{c}^\perp| = u\right\}$

$\Big($via Cauchy-Schwartz + Parseval$\Big)$

Our dream:

If f concentrates over words of Hamming weight t, then

$$\sqrt{2^n \sum_{u>0} N_u(\mathcal{C}^\perp)|\widehat{f}(u)|^2} \text{ is negligible as soon as } t \geq t_{\mathsf{GV}}$$

$\longrightarrow$ Optimal smoothing noise!

Upper bound in average:

$$\mathbb{E}_{\mathcal{C}^\perp}\left(\sqrt{2^n \sum_{u>0} N_u(\mathcal{C}^\perp)|\widehat{f}(u)|^2}\right) \leq \sqrt{2^n \sum_{u>0} \frac{\binom{n}{u}}{2^k} |\widehat{f}(u)|^2}$$

$\longrightarrow$ This “average” upper bound is only function of f

CHOICES OF DISTRIBUTION

- Choose f as $\mathrm{Ber}(t/n)^{\otimes n}$, then our bound is negligible when

$$t \geq \frac{n}{2}\left(1 - \sqrt{2^{k/n} - 1}\right) \gg t_{\mathrm{GV}}$$

- Choose f as uniform distribution over sphere with radius t, then our bound is negligible when

$$t \geq t_{\mathrm{GV}}$$

Conclusion:

Our bound enables optimal smoothing noise

We need to obtain a bound for a fixed $\mathcal{C}$, but how to upper-bound $N_u(\mathcal{C}^\perp)$?

$\longrightarrow$ We used the best known upper-bound $\Big($second linear programming bound$\Big)$

Failed attempt:

We obtained exactly the same constraint than already known smoothing bound using implicitly the trivial bound $N_u(\mathcal{C}^\perp) \leq \sharp\mathcal{C}^\perp$

CONCLUSION

CONCLUSION

- The worst-to-average case reduction can now be instantiated with any distribution for smoothing and our bound enables optimal parameter choice
- The Bernoulli distribution is not a good choice $\Big($unless to use a truncated argument$\Big)$
- The reduction has inherent limitation du to the constraint coming from the Gilbert-Varshamov radius