An Improvement on Ajtai-GGH Hash Function

Giovanni Tognolini

University of Trento

April 2022

Index

The Framework

The notion of hardness

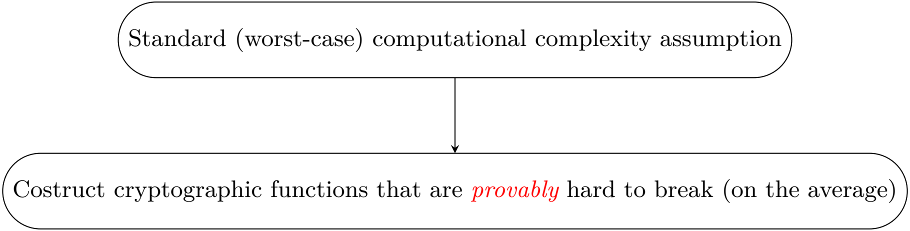
In computational complexity theory it is possible to distinguish two notions of *hardness*:

- *Problems hard in the worst-case.*
- *Problems hard in the average-case.*

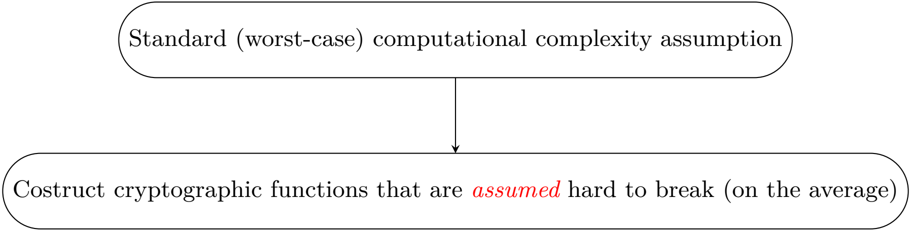
The notion of worst-case hardness is clearly not enough for cryptography

We want some reasonable guarantee that, if our key is chosen at random according to the prescribed key generation procedure, then with very high probability our key is hard to break.

What We would Like to Have in Cryptography



What we (almost) always have in cryptography



The Framework

Example

RSA requires the assumption that factoring is hard not only in the worst case, but also on the average, for a suitable distribution of n.

But with lattices we can do more!

In the following we construct a hash function as hard to break (on the average) as the worst case instance of solving certain lattice problem.

Lattices

Definition (Lattice)

Let $\mathbb{R}^m$ be the m-dimensional Euclidean space. Suppose we are given n linearly independent vectors $b_1, \ldots, b_n$. The *lattice* in $\mathbb{R}^m$ associated to $b_1, \ldots, b_n$ is the set

$$\mathcal{L}(b_1, \ldots, b_n) := \left\{ \sum_{i=1}^{n} x_i b_i : x_i \in \mathbb{Z} \right\}$$

- From a computational point of view, the basis vectors are assumed to be in $\mathbb{Q}^m$.
- We can multiply each element of the basis by an appropriate scaling factor and consider only integer lattices.

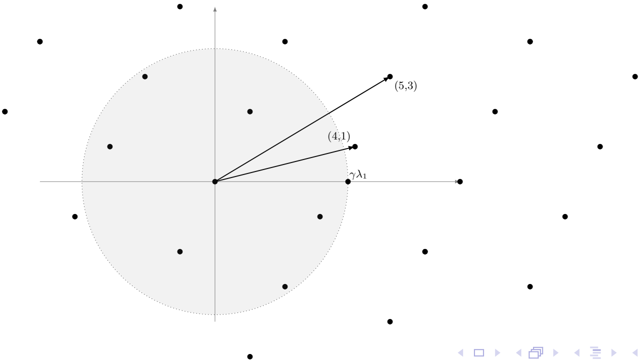
Problems on Lattices: SVP

Definition (Shortest Vector Problem, SVP)

Given a basis $B \in \mathbb{Z}^{m \times n}$, find a nonzero lattice vector Bx (with $x \in \mathbb{Z}^n \setminus \{0\}$) such that $\|Bx\| \leq \|By\|$ for any other $y \in \mathbb{Z}^n \setminus \{0\}$.

Definition (Approximate Shortest Vector Problem, SVP_γ)

Given a basis $B \in \mathbb{Z}^{m \times n}$, find a nonzero lattice vector Bx ($x \in \mathbb{Z}^n \setminus \{0\}$) such that $\|Bx\| \leq \gamma \cdot \|By\|$ for any other $y \in \mathbb{Z}^n \setminus \{0\}$.



Problems on Lattices: CVP

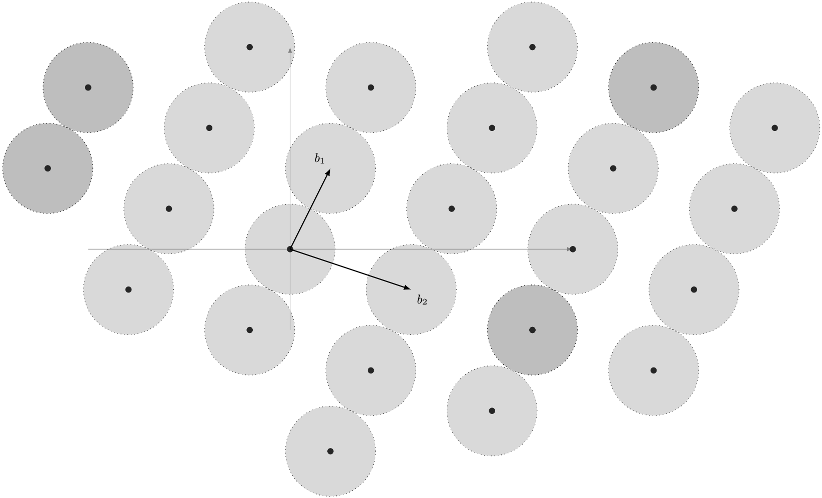
Definition (Closest Vector Problem, CVP)

Given a lattice basis $B \in \mathbb{Z}^{m \times n}$ and a target vector $t \in \mathbb{Z}^m$, find a lattice vector Bx closest to the target t, i.e., find an integer vector $x \in \mathbb{Z}^n$ such that $\|Bx - t\| \leq \|By - t\|$ for any other $y \in \mathbb{Z}^n$.

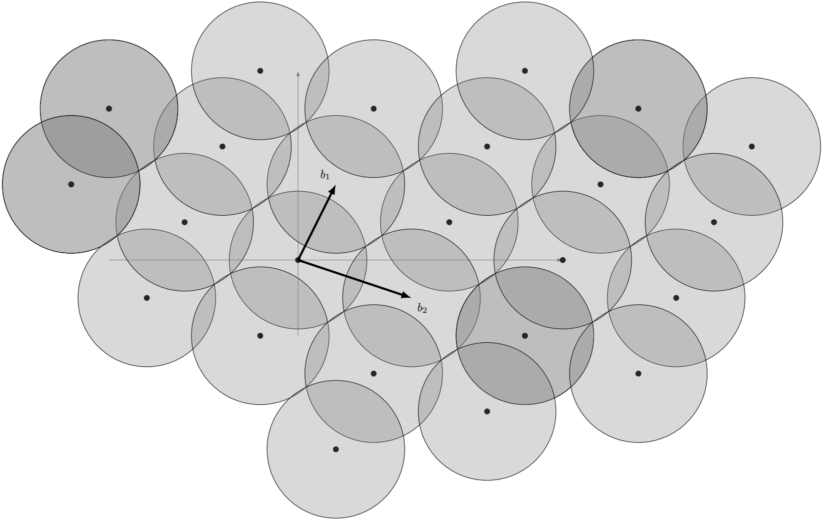
For the last problem we need to define two fundamental constant associated to any lattice:

- The packing radius.
- The covering radius.

Packing Radius ($\frac{\lambda_1}{2}$)



Covering Radius (ρ)

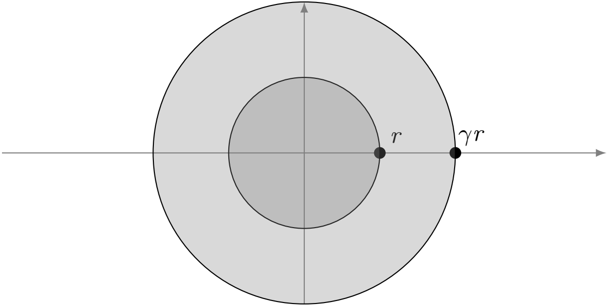


Problems on Lattices: GAPCRPγ

Definition (Approximate Covering Radius Problem, GAPCRPγ)

For any approximation factor γ, the (approximate) Covering Radius Problem (denoted GAPCRPγ) is the following promise problem. Instances are pairs (B, r). Moreover

- (B, r) is a YES instance if $\rho(\mathcal{L}(B)) \leq r$.
- (B, r) is a NO instance if $\rho(\mathcal{L}(B)) \geq \gamma \cdot r$.



Quotient Groups

Suppose we have

- L rank n lattice.
- M full rank sublattice of L (i.e. $M = LA$, for some nonsingular integer matrix $A \in \mathbb{Z}^{n \times n}$).

We can define a relation $\mathcal{R} \subseteq \mathcal{L}(L) \times \mathcal{L}(L)$ where

$$(x, y) \in \mathcal{R} \iff x - y \in \mathcal{L}(M)$$

Observation

- This is an equivalence relation.
- The quotient $\mathcal{L}(L)/\mathcal{L}(M)$ is an additive abelian group with $[x] + [y] = [x + y]$.

Quotient Groups

Question

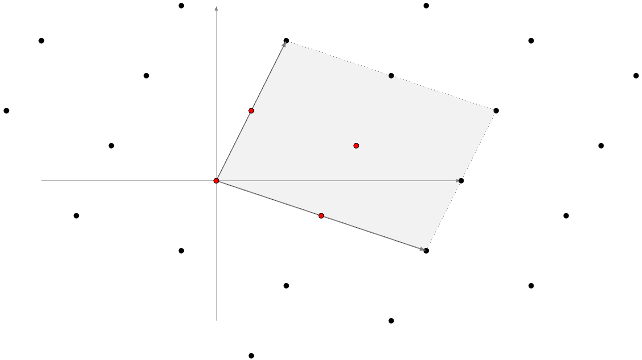
How to represent the elements of

$$G = \mathcal{L}(L)/\mathcal{L}(M)?$$

Quotient Groups

1. We could use the set of lattice points $\mathcal{L}(L) \cap P(M)$ in the half open parallelepipeded

$$P(M) := \{Mz : 0 \leq z_i < 1 \text{ for all } i\}$$



Observation

Computing representatives can be done efficiently (just by traslation).

Quotient Groups

2. Alternative way: use integer points inside $P(A^*)$, i.e. taking $z \in P(A^*) \cap \mathbb{Z}^n$.

In this case the geometric interpretation is lost, but it we can show that

$$\begin{aligned} P(A^*) \cap \mathbb{Z}^n &\longrightarrow \mathcal{L}(L)/\mathcal{L}(M) \\ z &\longmapsto Lz \end{aligned}$$

is a bijection

Observation

A possible way to compute the representative z' for Lz is to use a variant of the nearest plane algorithm:

- Given $\mathcal{L}(A)$ and a target z
- We find a vector $a \in \mathcal{L}(A)$ such that $z - a$ belongs to

$$P^{'}(A^*) := \left\{ A^* z : -\frac{1}{2} \leq z_i < \frac{1}{2} \text{ for all } i \right\} = P(A^*) - \frac{1}{2} \sum_i a_i^*$$

Quotient Groups

Observation

With these two methods we can represent the elements of G with strings of length polynomial in the size of the bases L and M

Question

Could we do better?

Quotient Groups

Observation

G only depends on $\mathcal{L}(L)$ and $\mathcal{L}(M)$

$\downarrow$

We can apply unimodular transformation to either basis without changing G

$$M = LA$$

$$MU = L(AU)$$

Remark

Any matrix A is column equivalent to a (unique) matrix in Hermite Normal Form.

Quotient Groups

Definition

A square nonsingular integer matrix $A \in \mathbb{Z}^{n \times n}$ is in Hermite Normal Form if

- A is upper triangular.
- All diagonal elements of A are strictly positive.
- All non diagonal elements are reduced modulo the corresponding diagonal element on the same row.

Observation

We can assume A is in HNF

Quotient Groups

Advantage

- The orthogonalized vectors a_i^* are simply given by $a_{i,i}e_i$.

$$A^* = \begin{bmatrix} a_{1,1} & 0 & \cdots & 0 \\ 0 & a_{2,2} & 0 & \cdots \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdot & a_{n,n} \end{bmatrix}$$

- $\mathbb{Z}^n \cap P(A^*) = \left\{ v \in \mathbb{Z}^n \text{ such that } 0 \leq v_i < a_{i,i} \right\}.$

Results:

- Each coordinate can be represented with $\log a_{i,i}$ bits
- The size of the group element representation is

$$\sum_{i=1}^{n} \log a_{i,i} = \log \prod_{i=1}^{n} a_{i,i} = \log \det(A) = \log |G|$$

Quotient Groups

In conclusion:

- Element of G are represented using $\log|G|$ bits.
- The group operation can be computed in *polynomial* time.

Observation

There is another way to represents group elements (using matrices in Smith Normal Form). In this way:

- We have a space efficient representation.
- We can perform group operation in *linear* time.

Ajtai's Intuition

$$\begin{aligned} f_A : \mathbb{Z}_q^m &\longrightarrow \mathbb{Z}_q^n \\ x &\longmapsto Ax \mod q \end{aligned}$$

- If A is chosen random
- If we consider a suitable restriction of the domain

(Worst case) lattice problem $\longrightarrow$ Inverting f_A (on the average)

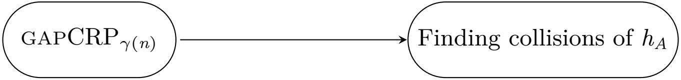
Ajtai-GGH Hash Function

Observation

Goldreich, Goldwasser and Halevi observed that a similar function

$$\begin{aligned} h_A : \{0,1\}^m &\longrightarrow \mathbb{Z}_q^n \\ x &\longmapsto Ax = \textstyle\sum_{i:x_i=1} a_i \end{aligned}$$

is collision resistant.



The Plan

In the folllowing...

- We construct a hash function family that generalizes and improves the Ajtai-GGH hash function.
- We prove its collision resistance.

The Construction

The Construction

1. We consider Λ a full-rank-n dimensional lattice such that
   a. The CVP in Λ can be solved easily.

Observation

Lattices where the CVP can be solved in polynomial time exist: consider for example $\Lambda := \mathbb{Z}^n$. Given $t \in \mathbb{Q}^n$, a lattice vector $x \in \Lambda$ closest to t can be easily found by rounding each coordinate of t to the closest integer.

   b. The packing-covering ratio $\tau = \frac{2\rho}{\lambda_1}$ is as small as possible ($\tau \in (1, \sqrt{n}]$).

Observation

τ is always greater that 1.

The Construction

Observation

Setting $\Lambda := \mathbb{Z}^n$ we obtain

$$\lambda_1 = 1 \qquad \rho = \sqrt{\left(\frac{1}{2}\right)^2 + \cdots + \left(\frac{1}{2}\right)^2} = \frac{1}{2}\sqrt{n}$$

and $\tau = \frac{2 \cdot \rho}{\lambda_1} = \sqrt{n}$.

Observation

For every n, there exists a lattice with $\tau < 4$.

The Construction

2. We construct an almost orthogonal sublattice $\mathcal{L}(M) \subseteq \Lambda$. The construction is as follows:
   - We define a scaling factor α.
   - For all $i \in \{1, ..., n\}$ let $m_i := CVP_\Lambda(\alpha\rho e_i)$

   In matrix notation:

$$\begin{cases} M := \alpha\rho I + R \\ \|r_i\| \leq \rho \end{cases}$$

3. We consider the abelian group $G := \Lambda/\mathcal{L}(M)$

Observation (Properties of G)

Using matrices in Smith Normal Form we can prove that:

1. Elements of G can be represented using $\log|G|$ bits.
2. Group operation can be computed in polynomial time.
3. There is an efficient homomorphism $\psi : \Lambda \longrightarrow G$ which maps each lattice vector to the corresponding group element ($\psi(x) = 0 \iff x \in \mathcal{L}(M)$).

The Construction

4. We define a family of G valued hash functions as follows:
   - We take an integer m.
   - We fix $a := (a_1, ..., a_m) \in G^m$.
   - We define the (hash) function

$$\begin{aligned} h_a \colon \{0,1\}^m &\longrightarrow G \\ x &\longmapsto \sum_{i=1}^m x_i a_i \end{aligned}$$

Collision Resistance

Collision Resistance

Our Aim

$a \in G^m$ random $\implies$ collisions are hard to find

Observation (How could we represent collisions?)

$$\begin{cases} x, y \in \{0,1\}^m \\ h_a(x) = h_a(y) \end{cases} \iff \begin{cases} z \in \{-1,0,1\}^m \setminus \{0\} \\ h_a(z) = 0 \end{cases}$$

We will refer to such vectors as h_a-collisions.

Collision Resistance

How to prove it?

$a \in G^m \longrightarrow$ Collision Finder Algorithm $\mathcal{F} \longrightarrow z$

Where z is an h_a-collision with nonnegligible probability δ.

$\textsc{gap}\mathrm{CRP}_{\gamma(n)} \longrightarrow$ Finding collisions of h_a

with $\gamma \in \omega(\tau n^2 \log n)$.

The Reduction

The Reduction

Steps

1. We consider a GAPCRPγ instance (B, r).
2. We try to find linearly indipendent vectors $s_1, ..., s_n \in \mathcal{L}(B)$ s.t. the length of the diagonal of the orthogonalized parallelepipeded $\sigma(S) := \sqrt{\sum_i \|s_i^*\|^2} < 2\gamma r$.

The Reduction

Why?

$$\begin{cases} \sigma(S) < 2\gamma r \\ 2\rho \leq \sigma(S) \qquad \text{(known result)} \end{cases} \implies \quad \rho < \gamma r$$

The Reduction

How?

We proceed iteratively as follows.

- Fix $S = B$. W.l.o.g. $\|s_1\| \leq \cdots \leq \|s_n\|$.
- If $\sigma(S) \geq 2\gamma\rho$ we can efficiently find (with nonnegligible probability) a new lattice vector $s \in \mathcal{L}(B)$ l.i. from $s_1, ..., s_{n-1}$ such that $\|s\| \leq \frac{1}{2}\|s_n\|$.
- Replace s_n with s and repeat.

Since the vectors cannot be reduced infinitely, at some point the iterative step must fail. If the iterative step repeatedly fails to find a short vector s, it must be the case (with very high prob.) that the assumption $\sigma(S) \geq 2\gamma\rho$ is false, i.e. $\sigma(S) < 2\gamma\rho$.

The Reduction

Result

We can build a randomized reduction that rejects all NO instances with probability 1, and accepts all YES instances with probability exponentially close to 1.

Extra

The Reduction

Main obstacle of the previous reduction

Suppose:

- We are under the same hypothesis of the previous construction.
- $B \in \mathbb{Z}^{n \times n}$ basis.
- $S := (s_1, ..., s_n)$ sequence of l.i. vectors in $\mathcal{L}(B)$ as before.

Then we have to efficiently find (with probability $\Omega(\delta)$) a lattice vector s such that:

- $s \in \mathcal{L}(B)$
- $s \notin \text{span}(s_1, ..., s_{n-1})$
- $\|s\| \leq \|s_n\|/2$

How Could We Do?

The Reduction

- Consider $\beta := \frac{\sqrt{n}\sigma(S)}{\alpha\rho(\Lambda)}$
- Consider the almost orthogonal matrix βM.
- We approximate each vector βm_i with a lattice point $c_i \in \mathcal{L}(S) \subseteq \mathcal{L}(B)$.

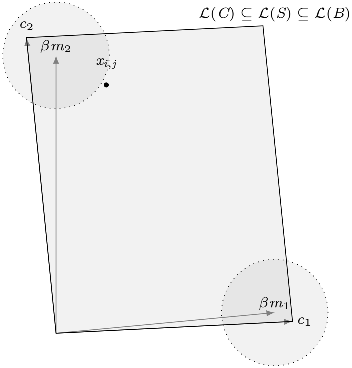
Observation

Using the nearest plane algorithm we can find c_i within distance $\sigma(S)/2$ from βm_i, i.e., in matrix notation:

$$\begin{cases} C = \beta M + Q \\ \|q_i\| \leq \sigma(S)/2 \end{cases}$$

The Reduction

- Define $k := 3\log n + \log(1/\delta)$.
- Consider the quotient $\mathcal{L}(B)/\mathcal{L}(C)$ and sample mk group elements $[x_{i,j}]_C \ \ (i \leq m, j \leq k)$.



The Reduction

At this point we would like to work in the lattice where the CVP is easy.

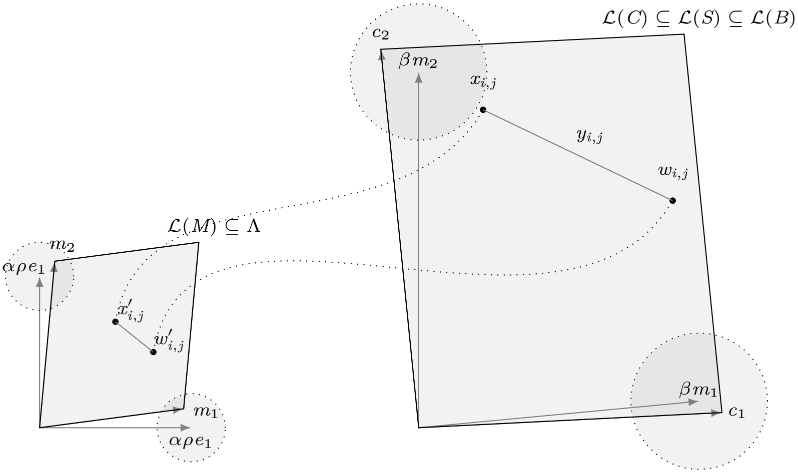
Observation

There is an invertible map

$$\begin{aligned} \mathcal{L}(B)/\mathcal{L}(C) &\longrightarrow \mathcal{L}(MC^{-1}B)/\mathcal{L}(M) \\ x &\longmapsto MC^{-1}x \end{aligned}$$

So we can choose $[x'_{i,j}]_M$ uniformly at random in $\mathcal{L}(B')/\mathcal{L}(M)$, where $B' := MC^{-1}B$, and then set $x_{i,j} = CM^{-1}x'_{i,j}$.

The Reduction



The Reduction

- Use CVP_Λ to find $w'_{i,j} \in \Lambda$ within distance $\rho(\Lambda)$ from $x'_{i,j}$.
- Fix $y_{i,j} := x_{i,j} - w_{i,j}$.
- Let $a_{i,j} := \psi(w'_{i,j})$ the group element corresponding to lattice point $w'_{i,j}$.
- Let $a_i := \sum_{j=1}^{k} a_{i,j}$, for every $i = 1, ..., m$.
- Call the collision finder algorithm $\mathcal{F}$ to obtain $z := \mathcal{F}(a)$.

The output of the reduction is the vector

$$s := \sum_{i=1}^{m} z_i \sum_{j=1}^{k} y_{i,j}$$

The Reduction

Observation

We have to prove that s satisfies:

- $s \in \mathcal{L}(B)$.
- $s \notin \text{span}(s_1, ..., s_{n-1})$.
- $\|s\| \leq \|s_n\|/2$.

with probability $\Omega(\delta)$.

The Reduction

Observation

$\mathcal{G} := \{g \in G^m \text{ s.t. } \mathcal{F}(g) \text{ is an } h_g\text{-collision}\} \Longrightarrow \mathbb{P}(u \in \mathcal{G}) = \delta$

We can prove the following:

1. If $a \in \mathcal{G}$ (i.e., if $\mathcal{F}(a)$ is an h_a-collision) then $s \in \mathcal{L}(B)$.
2. $\mathbb{P}(a \in \mathcal{G}) = \delta \cdot (1 - o(1))$.
3. $\mathbb{P}\left(s \notin \text{span}(s_1, ..., s_{n-1}) \middle| a \in \mathcal{G}\right) \geq 1/6$.
4. $\mathbb{P}(a \in \mathcal{G} \wedge \|s\| > \|s_n\|/2) < \delta \cdot o(1)$.

The Reduction

It follows that the success probability of the reduction is

$$
\begin{aligned}
\mathbb{P}(s \in \mathcal{L}(B), s \not\leq s_n^*, \|s\| \leq \|s_n\|/2) &\geq \mathbb{P}(a \in \mathcal{G}, s \not\leq s_n^*, 2\|s\| \leq \|s_n\|) \\
&= \mathbb{P}(a \in \mathcal{G}, s \not\leq s_n^*) \cdot \mathbb{P}\left(2\|s\| \leq \|s_n\| \middle| a \in \mathcal{G}, s \not\leq s_n^*\right) \\
&= \mathbb{P}(\cdots) \cdot \left(1 - \mathbb{P}\left(2\|s\| > \|s_n\| \middle| \cdots\right)\right) \\
&= \mathbb{P}(a \in \mathcal{G}, s \not\leq s_n^*) - \mathbb{P}(a \in \mathcal{G}, s \not\leq s_n^*, 2\|s\| > \|s_n\|) \\
&\geq \mathbb{P}(a \in \mathcal{G}, s \not\leq s_n^*) - \mathbb{P}(a \in \mathcal{G}, 2\|s\| > \|s_n\|) \\
&= \mathbb{P}(a \in \mathcal{G})\mathbb{P}\left(s \not\leq s_n^* \middle| a \in \mathcal{G}\right) - \mathbb{P}(a \in \mathcal{G}, 2\|s\| > \|s_n\|) \\
&\geq \delta(1 - o(1)) \cdot \frac{1}{6} - \delta \cdot o(1) \\
&= \Omega(\delta)
\end{aligned}
$$

Conclusions

Conclusions

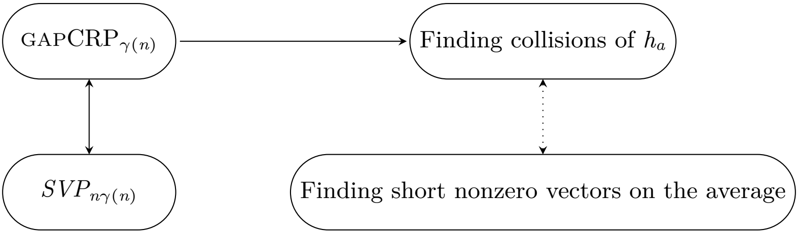
Observation (Collisions)

$$\begin{cases} x, y \in \{0,1\}^m \\ h_a(x) = h_a(y) \end{cases} \iff \begin{cases} z \in \{-1,0,1\}^m \setminus \{0\} \\ h_a(z) = 0 \end{cases}$$

Collisions "correspond" to short vectors z in the lattice

$$\Lambda_a := \{z : h_a(z) = 0\}$$

Conclusions



Thanks