SOME REMARKS ON SMALL SECRET LWE

Martin R. Albrecht @martinralbrecht

Auckland, December 4, 2015

Information Security Group, Royal Holloway, University of London

OUTLINE

Introduction

Warm Up

Modulus Switching

Coded-BKW

Swapping Error and Secret

A Different Embedding Approach

Exploiting Sparse Secrets

OUTLINE

Introduction

Warm Up

Modulus Switching
Coded-BKW

Swapping Error and Secret

A Different Embedding Approach

Exploiting Sparse Secrets

LEARNING WITH ERRORS

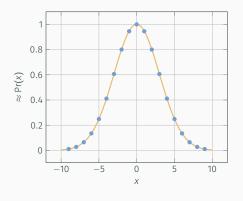
The Learning with Errors (LWE) problem was defined by Oded Regev¹. Given (\mathbf{A}, \mathbf{c}) with $\mathbf{c} \in \mathbb{Z}_a^m$, $\mathbf{A} \in \mathbb{Z}_a^{m \times n}$, $\mathbf{s} \in \mathbb{Z}_a^n$ and $\mathbf{e} \in \mathbb{Z}_a^m$ do we have

$$\left(\begin{array}{c} c \\ \end{array}\right) = \left(\begin{array}{ccc} \leftarrow & n & \rightarrow \\ & A \\ \end{array}\right) \times \left(\begin{array}{c} s \\ \end{array}\right) + \left(\begin{array}{c} e \\ \end{array}\right)$$

or $\mathbf{c} \leftarrow_{\$} \mathcal{U}(\mathbb{Z}_q^m)$.

¹Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In: *37th ACM STOC*. ed. by Harold N. Gabow and Ronald Fagin. ACM Press, May 2005, pp. 84–93.

PARAMETERS



- · Parameters are:
 - · dimension n,
 - modulus q,
 - noise size α ,
 - number of samples m.
- Elements of A, s, e, c are in \mathbb{Z}_q .
- **e** is sampled from a discrete Gaussian with width

$$\sigma = \frac{\alpha q}{\sqrt{2\pi}}.$$

LWE NORMAL FORM

Given samples

$$(\mathbf{a}, c) = (\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$$

with $\mathbf{a} \leftarrow \mathcal{U}(\mathbb{Z}_q^n)$, $e \leftarrow D_{\alpha q,0}$ and $\mathbf{s} \in \mathbb{Z}_q^n$, we can construct samples

$$(\mathbf{a}, c) = (\mathbf{a}, \langle \mathbf{a}, \mathbf{e} \rangle + e) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$$

with $\mathbf{a} \leftarrow \mathcal{U}(\mathbb{Z}_q^n)$, $e \leftarrow D_{\alpha q,0}$ and \mathbf{e} such that all components

$$e_i \leftarrow D_{\alpha q,0}$$

in polynomial time.²

²Benny Applebaum et al. Fast Cryptographic Primitives and Circular-Secure Encryption Based on Hard Learning Problems. In: *CRYPTO 2009*. Ed. by Shai Halevi. Vol. 5677. LNCS. Springer, Heidelberg, Aug. 2009, pp. 595–618.

SMALL SECRET LWE

- · Some applications use much smaller secrets.
- For example, $\mathbf{s}_i \leftarrow \{-1, 0, 1\}$ or $\mathbf{s}_i \leftarrow \{0, 1\}$.
- HElib³ chooses **s** such that h = 64 entries are ± 1 and all remaining entries are 0, regardless of dimension n.

Question

How much security does this cost?

³Shai Halevi and Victor Shoup. Algorithms in HElib. In: *CRYPTO 2014, Part I.* ed. by Juan A. Garay and Rosario Gennaro. Vol. 8616. LNCS. Springer, Heidelberg, Aug. 2014, pp. 554–571. DOI: 10.1007/978-3-662-44371-2_31.

HARDNESS: REDUCTIONS

"A major part of our reduction [...] is therefore dedicated to showing reduction from LWE (in dimension n) with arbitrary secret in \mathbb{Z}_q^n to LWE (in dimension $n \log_2 q$) with a secret chosen uniformly over $\{0,1\}$."

⁴Zvika Brakerski et al. Classical hardness of learning with errors. In: 45th ACM STOC. ed. by Dan Boneh, Tim Roughgarden, and Joan Feigenbaum. ACM Press, June 2013, pp. 575–584.

HARDNESS: ALGORITHMS

"[This work] suggests that this is overkill and that even n log log n may be more than sufficient."⁵

⁵Shi Bai and Steven D. Galbraith. Lattice Decoding Attacks on Binary LWE. . In: ACISP 14. Ed. by Willy Susilo and Yi Mu. Vol. 8544. LNCS. Springer, Heidelberg, July 2014, pp. 322–337. DOI: 10.1007/978-3-319-08344-5_21.

HARDNESS: CONSTRUCTIONS

"This brings up the question of whether one can get better attacks against LWE instances with a very sparse secret (much smaller than even the noise). [...] it seems that the very sparse secret should only add maybe one bit to the modulus/noise ratio." 6

⁶Craig Gentry, Shai Halevi, and Nigel P. Smart. Homomorphic Evaluation of the AES Circuit. Cryptology ePrint Archive, Report 2012/099. http://eprint.iacr.org/2012/099. 2012.

BINARY LWE SECRET DISTRIBUTIONS

- \mathcal{B}^+ each component is independently sampled uniformly from $\{0,1\}.$
- \mathcal{B}^- each component is independently sampled uniformly from $\{-1,0,1\}$.
- \mathcal{B}_{hw}^+ like \mathcal{B}^+ but with guarantee that hw components are non-zero.
- \mathcal{B}^-_{hw} like \mathcal{B}^- but with guarantee that hw components are non-zero.

LATTICE REDUCTION

In the guestimates below, we assumed

•
$$\delta_0 \approx \left(\frac{k}{2\pi e}(\pi k)^{\frac{1}{k}}\right)^{\frac{1}{2(k-1)}};$$

- the SVP oracle in BKZ is realise using sieving;
- sieving in blocksize k costs $t_k = 2^{0.3366 \, k + 12.31}$ clock cycles;
- BKZ-k costs $\frac{n^3}{k^2} \log(n) \cdot t_k$ clock cycles in dimension n.

https://github.com/dstehle/fplll
https://github.com/malb/fpylll

ROLLING EXAMPLE

We use the following LWE parameters as a rolling example throughout this talk.

- dimension n = 2048,
- modulus $q \approx 2^{63.4}$,
- noise parameter $\alpha \approx 2^{-60.4}$, i.e. standard deviation $\sigma \approx 3.2$,
- h=64 components of the secret are ± 1 , all other components are zero, $\sigma_{\rm S}\approx 0.44$: ${\cal B}_{64}^-$

This is inspired by parameters chosen in HElib.

OUTLINE

Introduction

Warm Up

Modulus Switching
Coded-BKW

Swapping Error and Secret

A Different Embedding Approach

Exploiting Sparse Secrets

EXHAUSTIVE SEARCH

- · Clearly, exhaustive search is an option for solving.
- This gives a complexity of about 2^n for \mathcal{B}^+ and 3^n for \mathcal{B}^- .
- For \mathcal{B}_{64}^- we get a complexity of about $2^{64} \cdot \binom{n}{64}$.

Meet in the Middle

We can about square-root these complexities using standard timememory trade-offs.

HElib

Plugging our example in gives expected costs of $\approx 2^{470}$ and $\approx 2^{235}$ operations, respectively.

STANDARD APPROACHES

Given
$$A, c$$
 with $c = A \times s + e$ or $c \leftarrow_s \mathbb{Z}_q^m$

 Solve the Short Integer Solutions problem (SIS) in the left kernel of A, i.e.

find a short w such that $\mathbf{w} \times \mathbf{A} = 0$

and check if $\langle w, c \rangle = w \times (A \times s + e) = \langle w, e \rangle$ is short.

STANDARD APPROACHES

Given
$$A$$
, c with $c = A \times s + e$ or $c \leftarrow_s \mathbb{Z}_q^m$

 Solve the Short Integer Solutions problem (SIS) in the left kernel of A, i.e.

find a short w such that $\mathbf{w} \times \mathbf{A} = 0$

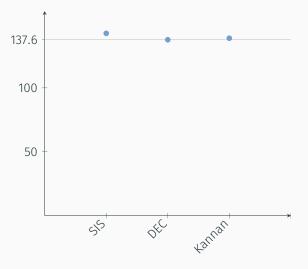
and check if $\langle w, c \rangle = w \times (A \times s + e) = \langle w, e \rangle$ is short.

· Solve the Bounded Distance Decoding problem (BDD), i.e.

find s' such that |w - c| with $w = A \times s'$ is minimised.

via Kannan's embedding or Babai's nearest planes.

STANDARD APPROACHES VS. ROLLING EXAMPLE



OUTLINE

Introduction

Warm Up

Modulus Switching

Coded-BKW

Swapping Error and Secret

A Different Embedding Approach

Exploiting Sparse Secrets

Let $(\mathbf{a},c)=(\mathbf{a},\langle \mathbf{a},\mathbf{s}\rangle+e)\in\mathbb{Z}_q^n\times\mathbb{Z}_q$ be an LWE sample and

$$p \approx \sqrt{\frac{2\pi n}{12}} \cdot \frac{\sigma_{\rm S}}{\alpha},$$

where σ_s is the standard deviation of components of the secret **s**. If p < q then

$$\left(\left\lfloor \frac{p}{q} \cdot \mathbf{a} \right
brace, \left\lfloor \frac{p}{q} \cdot c \right
brace \right)$$
 in $\mathbb{Z}_p^n \times \mathbb{Z}_p$

follows a distribution close to an LWE distribution with $n, \sqrt{2}\alpha, p$.

Zvika Brakerski and Vinod Vaikuntanathan. Efficient Fully Homomorphic Encryption from (Standard) LWE. . In: *52nd FOCS*. ed. by Rafail Ostrovsky. IEEE Computer Society Press, Oct. 2011, pp. 97–106

SLIGHTLY BETTER MODULUS SWITCHING FOR CRYPTANALYSIS I

- We usually simply assume that the rounding noise is also some Gaussian distribution.
- However, the rounding noise is not completely out of our control.
- · We know one component that goes into making it:

$$\frac{p}{q} \cdot \mathbf{a} - \left\lfloor \frac{p}{q} \cdot \mathbf{a} \right\rfloor$$

SLIGHTLY BETTER MODULUS SWITCHING FOR CRYPTANALYSIS II

Given known vectors $\mathbf{r}_i \longleftrightarrow \left(-\frac{1}{2},\frac{1}{2}\right]^n$ and an unknown fixed vector $\mathbf{s} \longleftrightarrow \mathcal{B}$, we call $\mathcal{Q}_{\mathbf{s}}(\mathbf{r}_i)$ the distribution obtained by outputting $\lfloor \langle \mathbf{r}_i, \mathbf{s} \rangle \rfloor$.

$$Q_{s}\left(\frac{p}{q}\cdot\mathbf{a}-\left\lfloor\frac{p}{q}\cdot\mathbf{a}\right\rfloor\right)=\left\langle\frac{p}{q}\cdot\mathbf{a}-\left\lfloor\frac{p}{q}\cdot\mathbf{a}\right\rfloor,s\right\rangle_{p}+e'$$

SLIGHTLY BETTER MODULUS SWITCHING FOR CRYPTANALYSIS III

Let $\mathbf{s} \leftarrow_{\mathbf{s}} \mathcal{B}^+$. Let $\mathbf{r}_i = \frac{p}{q} \cdot \mathbf{a}_i - \left\lfloor \frac{p}{q} \cdot \mathbf{a}_i \right\rfloor$. Let $L_{\mathbf{s},\chi}^{(n)}$ be a distribution which outputs those (\mathbf{a}_i',c_i') where $\sum \mathbf{r}_i' \leq c \cdot \sigma$ with σ the standard deviation of $\mathcal{Q}_{\mathbf{s}}(\mathbf{r}_i)$.

Then, $Q_s\left(\mathbf{r}_i'\right)$ for $\mathbf{r}_i'=\frac{p}{q}\cdot\mathbf{a}_i'-\left\lfloor\frac{p}{q}\cdot\mathbf{a}_i'\right
ceil$ satisfies:

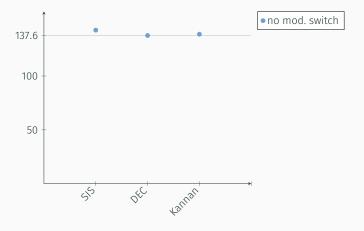
$$\Pr[\mathcal{Q}_{s}(\mathbf{r}'_{i}) > C \cdot \sigma] \leq \frac{\exp(-C^{2} + cC - c^{2}/2)}{2\pi \cdot (C^{2} - cC)}.$$

Compare:

$$\Pr[Q_{s}(\mathbf{r}_{i}) > C \cdot \sigma] \leq \frac{\exp(-C^{2}/2)}{C\sqrt{2\pi}}.$$

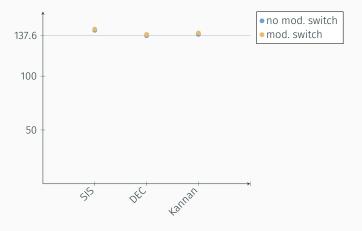
MODULUS SWITCHING + LATTICE REDUCTION |

Applied to our rolling example:



MODULUS SWITCHING + LATTICE REDUCTION |

Applied to our rolling example:



BKW ALGORITHM I

The BKW algorithm was first proposed for the Learning Parity with Noise (LPN) problem which can be viewed as a special case of LWE over \mathbb{Z}_2 .

BKW ALGORITHM II

We considering $a \approx \log n$ 'blocks' of b elements each.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & c_0 \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} & c_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} & c_m \end{pmatrix}$$

BKW ALGORITHM III

For each block we build a table of all q^b possible values indexed by \mathbb{Z}_q^b .

$$T^{0} = \begin{bmatrix} -\lfloor \frac{q}{2} \rfloor & -\lfloor \frac{q}{2} \rfloor & & & t_{1n} & C_{t,0} \\ -\lfloor \frac{q}{2} \rfloor & -\lfloor \frac{q}{2} \rfloor + 1 & & t_{2n} & C_{t,1} \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ \lfloor \frac{q}{2} \rfloor & \lfloor \frac{q}{2} \rfloor & t_{q^{2}3} & \cdots & t_{q^{2}n} & C_{t,q^{2}} \end{bmatrix}$$

For each $\mathbf{z} \in \mathbb{Z}_q^b$ we try to find a row in **A** such that it contains **z** as a subvector at the target indices.

BKW ALGORITHM IV

We use these tables to eliminate b entries in other rows. Assume $(\mathbf{a}_{21}, \mathbf{a}_{22}) = (\lfloor \frac{q}{2} \rfloor, \lfloor \frac{q}{2} \rfloor + 1)$, then:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & c_0 \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} & c_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} & c_m \end{pmatrix}$$

$$+ \begin{bmatrix} -\lfloor \frac{q}{2} \rfloor & -\lfloor \frac{q}{2} \rfloor & t_{13} & \cdots & t_{1n} & c_{t,0} \\ -\lfloor \frac{q}{2} \rfloor & -\lfloor \frac{q}{2} \rfloor + 1 & t_{23} & \cdots & t_{2n} & c_{t,1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \lfloor \frac{q}{2} \rfloor & \lfloor \frac{q}{2} \rfloor & t_{q^23} & \cdots & t_{q^2n} & c_{t,q^2} \end{bmatrix}$$

$$\Rightarrow \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & c_0 \\ 0 & 0 & \tilde{a}_{23} & \cdots & \tilde{a}_{2n} & \tilde{c}_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} & c_m \end{pmatrix}$$

LAZY MODULUS SWITCHING

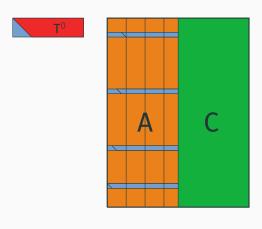
- When running the BKZ algorithm, only eliminate the most significant bits
- · This can be seen as a lazy variant of modulus switching.

Martin R. Albrecht et al. Lazy Modulus Switching for the BKW Algorithm on LWE. . In: *PKC 2014*. Ed. by Hugo Krawczyk. Vol. 8383. LNCS. Springer, Heidelberg, Mar. 2014, pp. 429–445. DOI: 10.1007/978-3-642-54631-0_25

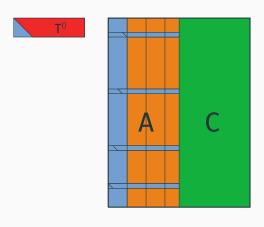
UNEVEN NOISE CONTRIBUTION |

When eliminating higher order bits in latter tables of BKW, this leads to an increase in the noise of the components covered by earlier tables.

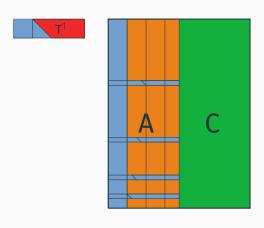
UNEVEN NOISE CONTRIBUTION II



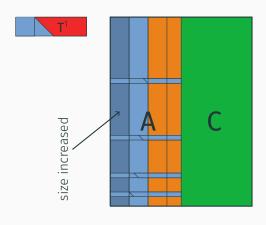
UNEVEN NOISE CONTRIBUTION III



UNEVEN NOISE CONTRIBUTION IV



UNEVEN NOISE CONTRIBUTION V



BALANCING NOISE

- We could pick decreasing moduli (increasing noise levels) for consecutive blocks to address this problem.
- This, however, would increase the complexity which would now be dominated by the size of the table T^0 .
- To compensate for this, we may choose increasing blocksizes b_i for each of the a blocks

Paul Kirchner and Pierre-Alain Fouque. An Improved BKW Algorithm for LWE with Applications to Cryptography and Lattices. In: CRYPTO 2015, Part I. ed. by Rosario Gennaro and Matthew J. B. Robshaw. Vol. 9215. LNCS. Springer, Heidelberg, Aug. 2015, pp. 43–62. DOI: 10.1007/978-3-662-47989-6_3

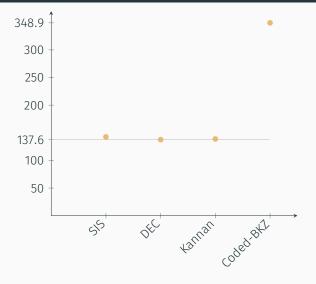
CODED-BKW I

This approach can be generalised

- Consider our modulus switching as a special form of quantisation (also done in [KF15])
- · Choose appropriate lattice code to find good quantisation
- Consider blocks of size b_i as messages which are thrown into buckets based on the codeword they correspond to.

Qian Guo, Thomas Johansson, and Paul Stankovski. Coded-BKW: Solving LWE Using Lattice Codes. In: *CRYPTO 2015, Part I.* ed. by Rosario Gennaro and Matthew J. B. Robshaw. Vol. 9215. LNCS. Springer, Heidelberg, Aug. 2015, pp. 23–42. DOI: 10.1007/978-3-662-47989-6_2

CODED-BKW II



OUTLINE

Introduction

Warm Up

Modulus Switching
Coded-BKW

Swapping Error and Secret

A Different Embedding Approach

Exploiting Sparse Secrets

SWAPPING ERROR AND SECRET |

"applying the reduction technique of Applebaum et al. to switch the key with part of the error vector, thus getting a smaller LWE error."

Craig Gentry, Shai Halevi, and Nigel P. Smart. Homomorphic Evaluation of the AES Circuit. Cryptology ePrint Archive, Report 2012/099. http://eprint.iacr.org/2012/099. 2012

Benny Applebaum et al. Fast Cryptographic Primitives and Circular-Secure Encryption Based on Hard Learning Problems. In: *CRYPTO 2009.* Ed. by Shai Halevi. Vol. 5677. LNCS. Springer, Heidelberg, Aug. 2009, pp. 595–618

SWAPPING ERROR AND SECRET II

 We are given a random m × n matrix A mod q, and also an m-vector

$$c = A \cdot s + e \mod q$$
.

- Let A_0 denotes the first n rows of A, A_1 the next n rows, etc.
- e_0, e_1, \ldots are the corresponding parts of the error vector and c_0, c_1, \ldots the corresponding parts of c.
- We have $\mathbf{c}_0 = \mathbf{A}_0 \cdot \mathbf{s} + \mathbf{e}_0$ or $\mathbf{A}_0^{-1} \cdot \mathbf{c}_0 = \mathbf{s} + \mathbf{A}_0^{-1} \mathbf{e}_0$.
- Also, for i > 0 we have $\mathbf{c}_i = \mathbf{A}_i \cdot \mathbf{s} + \mathbf{e}_i$, which together with the above gives

$$A_i A_0^{-1} c_0 - c_i = A_i A_0^{-1} e_0 - e_i.$$

SWAPPING ERROR AND SECRET III

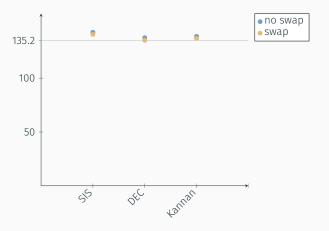
· Set $B=(A_0^{-1}\mid A_1\cdot A_0^{-1}\mid\dots)$ and $z=(A_0^{-1}c_0\mid A_1A_0^{-1}c_1\mid\dots)$, and also $f=(s|e_1|\dots)$ then we get the LWE instance

$$z = B \cdot e_0 + f$$

• For our rolling example, this reduces α from $2^{-60.4}$ to $\approx 2^{-60.8}$.

SWAPPING ERROR AND SECRET IV

Applied to our rolling example:



OUTLINE

Introduction

Warm Up

Modulus Switching
Coded-BKW

Swapping Error and Secret

A Different Embedding Approach

Exploiting Sparse Secrets

BAI-GAL ALGORITHM I

- Let m' = m + n.
- We may embed our LWE lattice into a different lattice with uSVP structure:

$$L = \{ \mathbf{v} \in \mathbb{Z}^{m'} | \mathbf{A}' \mathbf{v} \equiv 0 \bmod q \}$$

where

$$A' = (A|I_m).$$

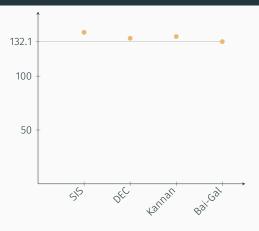
- The target short vector is now (s||e)
- When $|\mathbf{s}| \ll |\mathbf{e}|$, the vector $(\mathbf{s}||\mathbf{e})$ is uneven.
- We may want to rescale the first components to have same size as the last components.

BAI-GAL ALGORITHM II

- When $\mathbf{s} \leftarrow_{\mathbf{s}} \mathcal{B}^-$, after an appropriate rescaling, the volume of the lattice is increased by σ^n .
- When $\mathbf{s} \leftarrow_{\mathbf{s}} \mathcal{B}^+$ the volume is increased by $(2\sigma)^n$ because we can scale by 2σ and then rebalance.
- When $\mathbf{s} \leftarrow_{\mathbf{s}} \mathcal{B}^{\pm}_{hw}$ the volume increases further based on the hw.

Shi Bai and Steven D. Galbraith. Lattice Decoding Attacks on Binary LWE. . In: ACISP 14. Ed. by Willy Susilo and Yi Mu. Vol. 8544. LNCS. Springer, Heidelberg, July 2014, pp. 322–337. DOI: 10.1007/978-3-319-08344-5_21

BAI-GAL ALGORITHM III



Note: We don't know the performance of this algorithm in the low advantage regime.

OUTLINE

Introduction

Warm Up

Modulus Switching
Coded-BKW

Swapping Error and Secret

A Different Embedding Approach

Exploiting Sparse Secrets

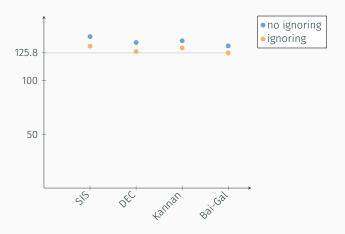
IGNORING COMPONENTS I

- All approaches so far tried to exploit small secrets. However, in our rolling example, the secret is sparse, i.e. most components are zero.
- In our example, the probability that a random coordinate is non-zero is $64/2048 = 1/32 \Rightarrow$ with probability 1 1/32 a coordinate is zero.
- Ignoring k random components will ignore only non-zero components with probability

$$P_k = \prod_{i=0}^{k-1} \left(1 - \frac{64}{n-i} \right)$$

• Solving $\approx 1/P_k$ instances in dimension n-k solves our instance at dimension n.

IGNORING COMPONENTS II



SUMMARY

To summarise the results for our rolling example, we get:

- $\cdot \approx 2^{137.6}$ operations when ignoring small, sparse secret
- $\boldsymbol{\cdot} \approx 2^{125.5}$ operations when exploiting small, sparse secret

THANK YOU



Questions?