Variants of the Syndrome Decoding Problem and algebraic cryptanalysis

Pierre Briaud[1], joint work with Morten Øygarden[2]

Crypto Reading Club meeting, September 6, 2023

[1]Inria Paris & Sorbonne Université
[2]Simula UiB

Syndrome Decoding Problem (SDP)

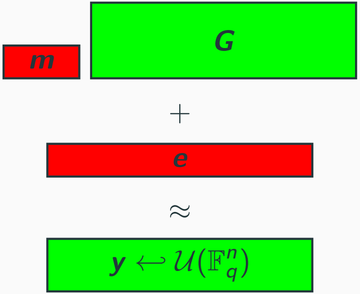
Code-based crypto — McEliece, BIKE, HQC, etc.

Secure Computation ("LPN") — Indistinguishability obfuscation

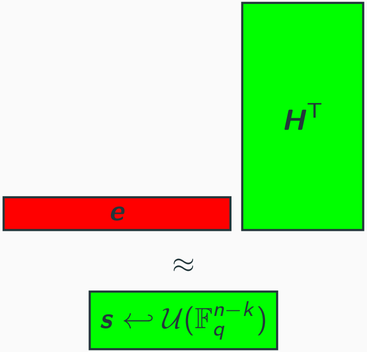
---

$\boldsymbol{G} \hookleftarrow \mathcal{U}(\mathbb{F}_q^{k\times n})$ full-rank, $\boldsymbol{m} \hookleftarrow \mathcal{U}(\mathbb{F}_q^{k})$

Error $\boldsymbol{e}$, $t \overset{def}{=} \mathrm{HW}(\boldsymbol{e})$ small



Parity-check $\boldsymbol{H} \hookleftarrow \mathcal{U}(\mathbb{F}_q^{(n-k)\times n})$ full-rank



SDP variants

What to change ?

- Public code: sparse, quasi-cyclic, ...
- Error distribution
- Metric: ~~HW~~ $\rightarrow$ rank metric, Lee metric

Goal

Efficiency gain ! (at least)

Known attacks

On plain version

Information Set Decoding (ISD), Statistical Decoding $\rightarrow$ combinatorial

More structure here !

- Improve generic solvers ?
- Other attack types ?

This talk

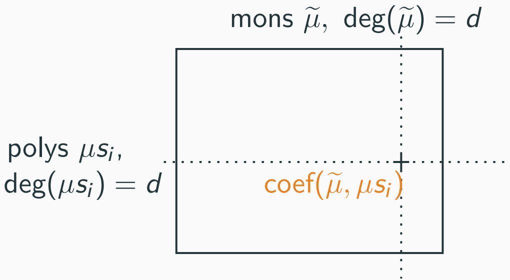
Algebraic cryptanalysis

- Reduction to polynomial system solving
- Applies to some variants

Regular Syndrome Decoding [BØ23] + Ongoing work

[BØ23] Briaud and Øygarden. "A New Algebraic Approach to the Regular Syndrome Decoding Problem and Implications for PCG Constructions". *Advances in Cryptology – EUROCRYPT 2023*.

1) $\times$ monomials: (Homogeneous) Macaulay matrix $\boldsymbol{M}_d$



2) Linear algebra: $\mathrm{RowEchelon}(\boldsymbol{M}_d)$ for $d \leq D$, solving degree

Regular Syndrome Decoding

Error distribution

Regular noise [AFS05]

Assume $n = N \times t$ for some $N \in \mathbb{N}$

- For $1 \leq i \leq t$, random $\boldsymbol{e}_i \in \mathbb{F}_q^N$, $\mathrm{HW}(\boldsymbol{e}_i) = 1$
- Error is $\boldsymbol{e} \stackrel{def}{=} (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_t) \in \mathbb{F}_q^n$

Secure Computation

Pseudorandom Correlation Generators (PCGs) [Boy+19]

[AFS05] Augot, Finiasz, and Sendrier. "A Family of Fast Syndrome Based Cryptographic Hash Functions". *MYCRYPT 2005*.

[Boy+19] Boyle et al. *Compressing Vector OLE*.

PCG for Vector OLE [Boy+19]

Want shares of long pseudorandom $\boldsymbol{u}$

1. Function Secret Sharing (FSS) $\rightarrow$ t-sparse vector $\boldsymbol{e}$
2. Decoding Problem $\rightarrow$ final $\boldsymbol{u}$

2 ways !

Code rate $R \stackrel{def}{=} k/n$

Primal	Dual
$\boldsymbol{u} = \boldsymbol{mG} + \boldsymbol{e}$	$\boldsymbol{u} = \boldsymbol{eH}^{\mathsf{T}}$
Very low R	Constant R

Regular $\boldsymbol{e}$ $\rightarrow$ reduce FSS cost

Contribution

Algebraic attack on Reg-SDP

- Tailored to noise distribution
- Can beat ISDs for low code rates (Primal)

Algebraic system for Reg-SDP

Modeling regular structure ($q = 2$)

Polynomial ring $R \stackrel{def}{=} \mathbb{F}_2[(e_{i,j})_{i,j}]$, n variables, block $\boldsymbol{e}_i \stackrel{def}{=} (e_{i,1}, \ldots, e_{i,N}) \in \mathbb{F}_2^N$

Coordinates $\in \mathbb{F}_2$ (field equations)

$$\forall i,\ \forall j,\ e_{i,j}^2 - e_{i,j} = 0 \tag{1}$$

One $\neq 0$ coordinate per block

$$\forall i,\ \forall j_1 \neq j_2,\ e_{i,j_1} e_{i,j_2} = 0 \tag{2}$$

Over $\mathbb{F}_2$, this coordinate is 1

$$\forall i,\ \sum_{j=1}^{N} e_{i,j} = 1 \tag{3}$$

We consider $\mathcal{Q} \stackrel{def}{=} (1) \cup (2) \cup (3)$

Parity-checks $eH^\top = s$

Linear equations ($\boldsymbol{h}_i$ i-th row in $\boldsymbol{H}$)

Parity-checks

$$\mathcal{P} \stackrel{def}{=} \{\forall i \in \{1..n-k\},\ \langle \boldsymbol{h}_i, \boldsymbol{e} \rangle - s_i\}$$

More when R small:

$$\boxed{\#\mathcal{P} = n - k = n(1 - R)}$$

Final system $\mathcal{S} \stackrel{def}{=} \mathcal{P} \cup \mathcal{Q}$

Estimating solving degree

Hilbert series (HS)

Homogeneous ideal I, $R_d \overset{def}{=} \text{span}\{\mu, \deg(\mu) = d\}$, $I_d \overset{def}{=} I \cap R_d$

$$\boxed{\mathcal{H}_{R/I}(z) \overset{def}{=} \sum_{d \in \mathbb{N}} \dim\left(R_d / I_d\right) z^d = \sum_{d \in \mathbb{N}} \dim\left(R_d\right) z^d - \sum_{d \in \mathbb{N}} \text{Rank}(\boldsymbol{M}_d) z^d}$$

Typical case in crypto: I zero-dimensional

Consequence

HS polynomial of degree $D - 1$

- Recover solving degree from HS !
- HS unknown in general :($\rightarrow$ need to estimate it

Structural part $\mathcal{Q}$

Easy to handle

$$\mathcal{Q}^{(h)} = \underbrace{\{\forall i \in \{1..t\}, \forall j \in \{1..N\},\ e_{i,j}^2\}}_{(1)} \cup \underbrace{\{\forall i, \forall j_1 \neq j_2,\ e_{i,j_1} e_{i,j_2}\}}_{(2)} \cup \underbrace{\{\forall i,\ \sum_{j=1}^{N} e_{i,j}\}}_{(3)}$$

HS 1

Combinatorics:

$$\boxed{\dim(R_d/\langle \mathcal{Q}^{(h)} \rangle_d) = \binom{t}{d}(N-1)^d}$$

$$\boxed{\mathcal{H}_{R/\langle \mathcal{Q}^{(h)} \rangle}(z) = (1 + (N-1)z)^t}$$

Parity-checks $\mathcal{P}$

Require assumption. Hope: HS known for random systems

Assumption ($\approx$ semi-regularity)

$\mathcal{P}^{(h)}$ behaves randomly in quotient $R/\langle \mathcal{Q}^{(h)} \rangle$

We have $\langle \mathcal{S}^{(h)} \rangle = \langle \mathcal{P}^{(h)} \rangle + \langle \mathcal{Q}^{(h)} \rangle$. Under Assumption, we get

$$\mathcal{H}_{R/\langle \mathcal{S}^{(h)} \rangle}(z) = \left[\frac{\mathcal{H}_{R/\langle \mathcal{Q}^{(h)} \rangle}(z)}{(1+z)^{n-k}} \right]_+ ,$$

$[.]_+$: truncation after first < 0 coef

HS 2 (under Assumption + using HS 1)

$$\boxed{\mathcal{H}_{R/\langle \mathcal{S}^{(h)} \rangle}(z) = \left[\frac{(1+(N-1)z)^t}{(1+z)^{n-k}} \right]_+}$$

Cost

Solving degree D

We had $D = \deg(\mathcal{H}_{R/\langle \mathcal{S}^{(h)} \rangle}) + 1$

$$\rightarrow \text{First} < 0 \text{ coef in } \frac{(1+(N-1)z)^t}{(1+z)^{n-k}}$$

- Linear algebra on Macaulay matrix $\boldsymbol{M}_D$, $2 \leq \omega < 3$

$$\boxed{T_{\text{solve}}(\mathcal{S}) = \mathcal{O}(\#\text{cols}(\boldsymbol{M}_D)^\omega) = \mathcal{O}\left(\binom{t}{D}^\omega (N-1)^{\omega D}\right)}$$

Improvements

- **Hybrid approach**
  $\rightarrow$ fix variables (here, in a structured way)

- **XL-Wiedemann**
  $\rightarrow$ exploit sparse Macaulay matrix

Cost with improvements

Parameters from Boyle *et al.* [Boy+19], updated analysis by Liu *et al.* [Liu+22]

n	k	t	$\mathbb{F}_2$ [Liu+22]	This work $\mathbb{F}_2$	$\mathbb{F}_{2^{128}}$ [Liu+22]	This work $\mathbb{F}_{2^{128}}$
2^{22}	64770	4788	147	**104**	156	**111**
2^{20}	32771	2467	143	**126**	155	**131**
2^{18}	15336	1312	139	**123**	153	**133**
2^{16}	7391	667	135	141	151	151
2^{14}	3482	338	132	140	150	152
2^{12}	1589	172	131	136	155	**152**
2^{10}	652	106	176	**146**	194	**180**

[Liu+22] Liu et al. *The Hardness of LPN over Any Integer Ring and Field for PCG Applications*.

Other SDP variants

Restricted Syndrome Decoding Problem (R-SDP)

CROSS signature scheme [Bal+23] $\rightarrow$ new NIST call

Constrained coefs

- Full Hamming weight

$$\boxed{\mathrm{HW}(\boldsymbol{e}) = n}$$

- Coefs $e_i \in \mathbb{F}_q^\times$ restricted to subgroup

$$\boxed{E = \langle g \rangle, \; g \in \mathbb{F}_q^\times \text{ of order } z}$$

Level 1 parameters $q = 127, \; n = 127, \; k = 76, \; z = 7$

[Bal+23] Baldi et al. *Zero Knowledge Protocols and Signatures from the Restricted Syndrome Decoding Problem*.

Permuted Kernel Problem (PKP)

Introduced by Shamir in 1989

- PKP-DSS [Beu+18] $\rightarrow$ Chinese PQC competition
- PERK $\rightarrow$ new NIST call

Formulation

Parity-check $\boldsymbol{H} \in \mathbb{F}_q^{(n-k)\times n}$, public vector $\boldsymbol{y} \in \mathbb{F}_q^n$.

Find secret $\sigma \in \mathfrak{S}_n$ s.t.

$$\boxed{\boldsymbol{y}_\sigma \boldsymbol{H}^\top = 0, \text{ where } \boldsymbol{y}_\sigma = (y_{\sigma(1)}, \ldots, y_{\sigma(n)})}$$

Level 1 PKP-DSS $q = 251, \; n = 69, \; n - k = 41$ $(n! \approx q^{n-k})$

[Beu+18] Beullens et al. *PKP-Based Signature Scheme*.

Algebraic systems ?

Parity-checks

$\rightarrow n - k$ linear eqs

Extra structure

$\rightarrow$ higher degree eqs

- R-SDP:

$$\forall i \in \{1..n\},\ {e_i}^z - 1 = 0$$

- PKP:

Model permutation matrix $\boldsymbol{P}_\sigma = (p_{i,j})$

Ongoing work

Same approach for Hilbert series ?

- Seems fine for R-SDP
- Much more complicated for PKP