Boolean Modeling and Analysis of Learning With Rounding

Jules Baudrin, Rachelle Heim Boissier, François-Xavier Standaert

Université Catholique de Louvain

Jan. 2025

Outline

1 Introduction: motivation and setting

2 A symmetric point of view

3 Results

4 Some intuition

5 A taste of the algorithms

Hard learning problems

Learning With Error (LWE), Learning With Rounding (LWR), Learning Parity with Noise (LPN)

and their ring/module variants.

Central importance in post-quantum cryptography

- Encryption, Key encapsulation mechanisms: CRYSTALS-Kyber, Saber
- Signatures: CRYSTALS-Dilithium, BLISS

and in symmetric cryptography:

- Essentially to build (key homomorphic) PRFs for a variety of applications.
- E.g. distributed PRFs, proxy re-encryption, updatable encryption (Boneh et al., 2013)

Learning With Errors

In a nutshell: solving a noisy linear system over a ring.

Search Learning With Errors (Regev 05)

Parameters: $q \in \mathbb{N}$, $n \in \mathbb{N}^*$, small (Gaussian) distribution χ over $\mathbb{Z}_q$, secret $\boldsymbol{x} \overset{\$}{\leftarrow} \mathbb{Z}_q^n$

Given samples from the distribution

$$\mathscr{D}^{\mathsf{LWE}} = \{ \ (\boldsymbol{a}, \langle \boldsymbol{a}, \boldsymbol{x} \rangle + e), \ \ \boldsymbol{a} \overset{\$}{\leftarrow} \mathbb{Z}_q^n, \ \ e \leftarrow \chi \ \}$$

Find $\boldsymbol{x}$.

Learning With Errors

In a nutshell: solving a noisy linear system over a ring.

Search Learning With Errors (Regev 05)

Parameters: $q \in \mathbb{N}$, $n \in \mathbb{N}^*$, small (Gaussian) distribution χ over $\mathbb{Z}_q$, secret $\boldsymbol{x} \overset{\$}{\leftarrow} \mathbb{Z}_q^n$

Given samples from the distribution

$$\mathscr{D}^{\mathsf{LWE}} = \{\ (\boldsymbol{a}, \langle \boldsymbol{a}, \boldsymbol{x} \rangle + e),\ \ \boldsymbol{a} \overset{\$}{\leftarrow} \mathbb{Z}_q^n,\ \ e \leftarrow \chi\ \}$$

Find $\boldsymbol{x}$.

Decision LWE: distinguish from $\mathscr{D}_0 = \{(\boldsymbol{a}, r) \mid \boldsymbol{a} \overset{\$}{\leftarrow} \mathbb{Z}_q^n, r \overset{\$}{\leftarrow} \mathbb{Z}_q\}$

Learning With Errors

In a nutshell: solving a noisy linear system over a ring.

Search Learning With Errors (Regev 05)

Parameters: $q \in \mathbb{N}$, $n \in \mathbb{N}^*$, small (Gaussian) distribution χ over $\mathbb{Z}_q$, secret $\boldsymbol{x} \overset{\$}{\leftarrow} \mathbb{Z}_q^n$

Given samples from the distribution

$$\mathscr{D}^{\mathsf{LWE}} = \{\ (\boldsymbol{a}, \langle \boldsymbol{a}, \boldsymbol{x} \rangle + e),\ \ \boldsymbol{a} \overset{\$}{\leftarrow} \mathbb{Z}_q^n,\ \ e \leftarrow \chi\ \}$$

Find $\boldsymbol{x}$.

Decision LWE: distinguish from $\mathscr{D}_0 = \{(\boldsymbol{a}, r) \mid \boldsymbol{a} \overset{\$}{\leftarrow} \mathbb{Z}_q^n, r \overset{\$}{\leftarrow} \mathbb{Z}_q\}$

- Security level is determined by n, q, and standard deviation σ of χ.
- Drawback: LWE cannot be used to build deterministic primitives such as PRFs.

Learning with Rounding

'A way of partially 'derandomizing' the LWE problem, i.e. generating errors efficiently and deterministically'.

Banerjee, Peikert, Rosen, EC' 2012.

Search Learning With Rounding

Parameters: $q \in \mathbb{N}$, $p, n \in \mathbb{N}^*$, $p < q$, rounding function $\lfloor \cdot \rceil_p : \mathbb{Z}_q \to \mathbb{Z}_p$, secret $\boldsymbol{x} \overset{\$}{\leftarrow} \mathbb{Z}_q^n$

Given samples from the distribution

$$\mathscr{D}^{\mathrm{LWR}} = \{ \ (\boldsymbol{a}, \lfloor \langle \boldsymbol{a}, \boldsymbol{x} \rangle \rceil_p), \ \boldsymbol{a} \overset{\$}{\leftarrow} \mathbb{Z}_q^n \ \}$$

Find $\boldsymbol{x}$.

Today: PQ cryptography (e.g. Saber [B+18]), symmetric cryptography (PRFs indeed).

Power-of-two moduli

Today: Many use-cases rely on LWR with a power-of-two modulus.

Search Learning With Rounding

Parameters: $q \in \mathbb{N}$, $p, n \in \mathbb{N}^*$, $p < q$, rounding function $\lfloor \cdot \rceil_p : \mathbb{Z}_q \to \mathbb{Z}_p$, secret $\boldsymbol{x} \xleftarrow{\$} \mathbb{Z}_q^n$

Given samples from the distribution

$$\mathscr{D}^{\mathrm{LWR}} = \{\ (\boldsymbol{a}, s_{\boldsymbol{a}} = \lfloor \langle \boldsymbol{a}, \boldsymbol{x} \rangle \rceil_p),\ \boldsymbol{a} \xleftarrow{\$} \mathbb{Z}_q^n\ \}$$

Find $\boldsymbol{x}$.

Power-of-two moduli

Today: Many use-cases rely on LWR with a power-of-two modulus.

Search Learning With Rounding

Parameters: $q \in \mathbb{N}$, $p, n \in \mathbb{N}^*$, $p < q$, rounding function $\lfloor \cdot \rfloor_{2^p} : \mathbb{Z}_{2^q} \to \mathbb{Z}_{2^p}$, secret $\boldsymbol{x} \xleftarrow{\$} \mathbb{Z}_{2^q}^n$

Given samples from the distribution

$$\mathscr{D}^{\mathrm{LWR}} = \{ \ (\boldsymbol{a}, s_{\boldsymbol{a}} = \lfloor \langle \boldsymbol{a}, \boldsymbol{x} \rangle \rfloor_p), \ \boldsymbol{a} \xleftarrow{\$} \mathbb{Z}_{2^q}^n \ \}$$

Find $\boldsymbol{x}$.

Power-of-two moduli

Today: Many use-cases rely on LWR with a power-of-two modulus.

Search Learning With Rounding

Parameters: $q \in \mathbb{N}$, $p, n \in \mathbb{N}^*$, $p < q$, rounding function $\lfloor \cdot \rceil_{2^p} : \mathbb{Z}_{2^q} \to \mathbb{Z}_{2^p}$, secret $\boldsymbol{x} \xleftarrow{\$} \mathbb{Z}_{2^q}^n$

Given samples from the distribution

$$\mathscr{D}^{\mathrm{LWR}} = \{ \ (\boldsymbol{a}, s_{\boldsymbol{a}} = \lfloor \langle \boldsymbol{a}, \boldsymbol{x} \rangle \rceil_p), \ \boldsymbol{a} \xleftarrow{\$} \mathbb{Z}_{2^q}^n \ \}$$

Find $\boldsymbol{x}$.

In this case: the rounding function $\lfloor \cdot \rceil_{2^p}$ simply removes the $q - p$ least significant bits.

- Security level is determined by n, q and $q - p$: noise $\sim \mathit{Uniform}[-2^{q-p}, 0)$

e.g. LightSaber: $n = 512, q - p = 3$, dPRF LaKey $n = 256, q - p = 4$.

Power-of-two moduli

Today: Many use-cases rely on LWR with a power-of-two modulus.

Search Learning With Rounding

Parameters: $q \in \mathbb{N}$, $p, n \in \mathbb{N}^*$, $p < q$, rounding function $\lfloor \cdot \rfloor_{2^p} : \mathbb{Z}_{2^q} \to \mathbb{Z}_{2^p}$, secret $\boldsymbol{x} \xleftarrow{\$} \mathbb{Z}_{2^q}^n$

Given samples from the distribution

$$\mathscr{D}^{\mathrm{LWR}} = \{\ (\boldsymbol{a}, s_{\boldsymbol{a}} = \lfloor \langle \boldsymbol{a}, \boldsymbol{x} \rangle \rfloor_p),\ \ \boldsymbol{a} \xleftarrow{\$} \mathbb{Z}_{2^q}^n\ \}$$

Find $\boldsymbol{x}$.

In this case: the rounding function $\lfloor \cdot \rfloor_{2^p}$ simply removes the $q - p$ least significant bits.

- Security level is determined by n, q and $q - p$: noise $\sim \mathit{Uniform}[-2^{q-p}, 0)$

e.g. LightSaber: $n = 512, q - p = 3$, dPRF LaKey $n = 256, q - p = 4$.

Hardness

Theory

- LWE: Solid theoretical foundations (e.g. Brakerski et al. 13).
- LWR is as hard as LWE (asymptotic reduction, underlying assumptions).

Practice

- Parameter selection driven by best known attacks (Lattice estimator, Albrecht et al.)

> *'The hardness of (ring or module) LWR can be analyzed as an LWE problem, since there is no known attacks that make use of the additional structure offered by these variants'.*
>
> SABER specifications

Open question: what does a deterministic error do to (practical) security?

Linearisation attack by Arora & Ge (2011)

- Low noise, large number of samples (not SABER).

Parameters: $n \in \mathbb{N}^*$, Noise in set E.

Any sample $(\boldsymbol{a}, s_{\boldsymbol{a}} = \langle \boldsymbol{a}, \boldsymbol{x} \rangle + e_0)$, yields the following equation in the secret:

$$\prod_{e \in E} (s_{\boldsymbol{a}} - \langle \boldsymbol{a}, \boldsymbol{x} \rangle - e) = 0. \qquad \text{(degree } |E| \text{ polynomial)}$$

Nr of monomials: $\binom{n+|E|}{|E|}$

Linearisation: $\binom{n+|E|}{|E|}$ in data, $\binom{n+|E|}{|E|}^{\omega}$ in time, ω linear algebra constant.

And Gröbner? Some asymptotic results for prime or odd moduli (not 2^q).

- **LWE**: Gaussian distribution: bounded noise for a well-chosen number of samples.
- **LWR**: $|E| = 2^{q-p}$.

Linearisation attack by Arora & Ge (2011)

- Low noise, large number of samples (not SABER).

Parameters: $n \in \mathbb{N}^*$, Noise in set E.

Any sample $(\boldsymbol{a}, s_{\boldsymbol{a}} = \langle \boldsymbol{a}, \boldsymbol{x} \rangle + e_0)$, yields the following equation in the secret:

$$\prod_{e \in E} (s_{\boldsymbol{a}} - \langle \boldsymbol{a}, \boldsymbol{x} \rangle - e) = 0. \qquad \text{(degree } |E| \text{ polynomial)}$$

Nr of monomials: $\binom{n+|E|}{|E|}$

Linearisation: $\binom{n+|E|}{|E|}$ in data, $\binom{n+|E|}{|E|}^{\omega}$ in time, ω linear algebra constant.

And Gröbner? Some asymptotic results for prime or odd moduli (not 2^q).

- **LWE**: Gaussian distribution: bounded noise for a well-chosen number of samples.
- **LWR**: $|E| = 2^{q-p}$.

Our main result: in the case of LWR, one can do better.

Outline

1 Introduction: motivation and setting

2 A symmetric point of view

3 Results

4 Some intuition

5 A taste of the algorithms

A symmetric point of view

Def: A Boolean function is a function $f : \mathbb{F}_2^s \longrightarrow \mathbb{F}_2$.

Ingredients to generate an LWR sample: $(\boldsymbol{a},\ s_{\boldsymbol{a}} = \lfloor \langle \boldsymbol{a}, \boldsymbol{x} \rangle \rceil_p)$.

- Inner product $\langle \cdot, \cdot \rangle : \mathbb{Z}_{2^q}^n \times \mathbb{Z}_{2^q}^n \to \mathbb{Z}_{2^q}$
- Rounding function $\lfloor \cdot \rceil_{2^p} : \mathbb{Z}_{2^q} \to \mathbb{Z}_{2^p}$.

A symmetric point of view

Def: A Boolean function is a function $f : \mathbb{F}_2^s \longrightarrow \mathbb{F}_2$.

Ingredients to generate an LWR sample: $(\boldsymbol{a},\ s_{\boldsymbol{a}} = \lfloor \langle \boldsymbol{a}, \boldsymbol{x} \rangle \rfloor_p)$.

- Inner product $\langle \cdot, \cdot \rangle : \mathbb{Z}_{2^q}^n \times \mathbb{Z}_{2^q}^n \to \mathbb{Z}_{2^q}$
- Rounding function $\lfloor \cdot \rfloor_{2^p} : \mathbb{Z}_{2^q} \to \mathbb{Z}_{2^p}$.

Identify $\mathbb{Z}_{2^q}$ with $\mathbb{F}_2^q$.

A symmetric point of view

Def: A Boolean function is a function $f : \mathbb{F}_2^s \longrightarrow \mathbb{F}_2$.

Ingredients to generate an LWR sample: $(\boldsymbol{a},\ s_{\boldsymbol{a}} = \lfloor \langle \boldsymbol{a}, \boldsymbol{x} \rangle \rfloor_p)$.

- Inner product $\langle \cdot, \cdot \rangle : (\mathbb{F}_2^q)^n \times (\mathbb{F}_2^q)^n \to \mathbb{F}_2^q$
- Rounding function $\lfloor \cdot \rfloor_{2^p} : \mathbb{F}_2^q \to \mathbb{F}_2^p$.

Identify $\mathbb{Z}_{2^q}$ with $\mathbb{F}_2^q$.

A symmetric point of view

Def: A Boolean function is a function $f : \mathbb{F}_2^s \longrightarrow \mathbb{F}_2$.

Ingredients to generate an LWR sample: $(\boldsymbol{a},\ s_{\boldsymbol{a}} = \lfloor \langle \boldsymbol{a}, \boldsymbol{x} \rangle \rfloor_p)$.

- Inner product $\langle \cdot, \cdot \rangle : (\mathbb{F}_2^q)^n \times (\mathbb{F}_2^q)^n \to \mathbb{F}_2^q$
- Rounding function $\lfloor \cdot \rfloor_{2^p} : \mathbb{F}_2^q \to \mathbb{F}_2^p$.

Identify $\mathbb{Z}_{2^q}$ with $\mathbb{F}_2^q$.

The LWR function $(\boldsymbol{a}, \boldsymbol{x}) \mapsto \lfloor \langle \boldsymbol{a}, \boldsymbol{x} \rangle \rfloor_{2^p}$ is a vectorial Boolean function.

Symmetric crypto ♡ Boolean functions

The **LWR problem** looks like other symmetric-key problems ($\approx$ Weak-PRF).

Boolean monomials

- $x[i]$ is the bit of index i of x.
- R is the multivariate polynomial ring $\mathbb{F}_2[x[0], \ldots, x[s-1]]$ quotiented by the field equations $x[i]^2 + x[i]$.

Boolean monomial. Let $x, m \in \mathbb{F}_2^s$.

$$x^m := \prod_{i,\ m[i]=1} x[i] \in R\,.$$

Boolean monomials

- $x[i]$ is the bit of index i of x.
- R is the multivariate polynomial ring $\mathbb{F}_2[x[0], \ldots, x[s-1]]$ quotiented by the field equations $x[i]^2 + x[i]$.

Boolean monomial. Let $x, m \in \mathbb{F}_2^s$.

$$x^m := \prod_{i,\ m[i]=1} x[i] \in R\,.$$

Since $3 = (11)_2$, x^3 is the product of the two least significant bits of x, $x[0]x[1]$.

Boolean monomials

- $x[i]$ is the bit of index i of x.
- R is the multivariate polynomial ring $\mathbb{F}_2[x[0], \dots, x[s-1]]$ quotiented by the field equations $x[i]^2 + x[i]$.

Boolean monomial. Let $x, m \in \mathbb{F}_2^s$.

$$x^m := \prod_{i,\ m[i]=1} x[i] \in R\,.$$

Since $3 = (11)_2$, x^3 is the product of the two least significant bits of x, $x[0]x[1]$.

Since $4 = 2^2 = (100)_2$, x^4 is the third bit of x, $x[2]$.

Boolean monomials

- $x[i]$ is the bit of index i of x.
- R is the multivariate polynomial ring $\mathbb{F}_2[x[0], \ldots, x[s-1]]$ quotiented by the field equations $x[i]^2 + x[i]$.

Boolean monomial. Let $x, m \in \mathbb{F}_2^s$.

$$x^m := \prod_{i,\ m[i]=1} x[i] \in R\,.$$

Since $3 = (11)_2$, x^3 is the product of the two least significant bits of x, $x[0]x[1]$.

Since $4 = 2^2 = (100)_2$, x^4 is the third bit of x, $x[2]$.

Since $2^i = (10\ldots0)_2$, x^{2^i} is the $i+1$th bit of x, $x[i]$.

Symmetric point of view on LWR

The LWR function $(\boldsymbol{a}, \boldsymbol{x}) \mapsto \lfloor \langle \boldsymbol{a}, \boldsymbol{x} \rangle \rfloor_{2^p}$ is a vectorial Boolean function.

It has domain $\mathbb{Z}_{2^q}^n \times \mathbb{Z}_{2^q}^n$ (or $(\mathbb{F}_2^q)^n \times (\mathbb{F}_2^q)^n$) and co-domain $\mathbb{Z}_{2^p}$ (or $\mathbb{F}_2^p$).

We study the composition of the inner product with a (product of) bit:

$$F^{m,n} : (\boldsymbol{a}, \boldsymbol{x}) \mapsto (\langle \boldsymbol{a}, \boldsymbol{x} \rangle)^m \qquad F_{\boldsymbol{a}}^{m,n} : \boldsymbol{x} \mapsto (\langle \boldsymbol{a}, \boldsymbol{x} \rangle)^m \ .$$

Symmetric point of view on LWR

The LWR function $(\boldsymbol{a}, \boldsymbol{x}) \mapsto \lfloor \langle \boldsymbol{a}, \boldsymbol{x} \rangle \rceil_{2^p}$ is a vectorial Boolean function.

It has domain $\mathbb{Z}_{2^q}^n \times \mathbb{Z}_{2^q}^n$ (or $(\mathbb{F}_2^q)^n \times (\mathbb{F}_2^q)^n$) and co-domain $\mathbb{Z}_{2^p}$ (or $\mathbb{F}_2^p$).

We study the composition of the inner product with a (product of) bit:

$$F^{m,n} : (\boldsymbol{a}, \boldsymbol{x}) \mapsto (\langle \boldsymbol{a}, \boldsymbol{x} \rangle)^m \qquad F_{\boldsymbol{a}}^{m,n} : \boldsymbol{x} \mapsto (\langle \boldsymbol{a}, \boldsymbol{x} \rangle)^m .$$

Most important example:

- $m = 2^{q-p}$. $F^{2^{q-p},n}$ returns the $q - p + 1$th bit of $\langle \boldsymbol{a}, \boldsymbol{x} \rangle$.
  This corresponds to the LSB of the LWR sample.

Algebraic Normal Form

The Algebraic Normal Form (ANF) of $f : \mathbb{F}_2^s \to \mathbb{F}_2$ is the unique multivariate polynomial

$$\sum_{u \in \mathbb{F}_2^s} \alpha_u(f) x^u \in R \qquad \text{s.t.} \qquad \forall x \in \mathbb{F}_2^s, \quad f(x) = \sum_{u \in \mathbb{F}_2^m} \alpha_u(f) x^u \, .$$

Algebraic degree.

$$\deg(f) := \max_{u \in \mathbb{F}_2^m | \alpha_u \neq 0} hw(u).$$

Algebraic Normal Form

The Algebraic Normal Form (ANF) of $f : \mathbb{F}_2^s \to \mathbb{F}_2$ is the unique multivariate polynomial

$$\sum_{u \in \mathbb{F}_2^s} \alpha_u(f) x^u \in R \qquad \text{s.t.} \qquad \forall x \in \mathbb{F}_2^s, \quad f(x) = \sum_{u \in \mathbb{F}_2^m} \alpha_u(f) x^u \, .$$

Algebraic degree.

$$\deg(f) := \max_{u \in \mathbb{F}_2^m | \alpha_u \neq 0} hw(u).$$

Example. $f : \mathbb{F}_2^2 \to \mathbb{F}_2$ defined by $f(00) = 0, f(01) = 1, f(10) = 0, f(11) = 0.$

$$f = x[1]x[0] + x[0] = x^3 + x^1 \qquad\qquad \text{degree } 2.$$

Algebraic Normal Form

The Algebraic Normal Form (ANF) of $f : \mathbb{F}_2^s \to \mathbb{F}_2$ is the unique multivariate polynomial

$$\sum_{u \in \mathbb{F}_2^s} \alpha_u(f) x^u \in R \qquad \text{s.t.} \qquad \forall x \in \mathbb{F}_2^s, \quad f(x) = \sum_{u \in \mathbb{F}_2^m} \alpha_u(f) x^u \,.$$

Algebraic degree.

$$\deg(f) := \max_{u \in \mathbb{F}_2^m \mid \alpha_u \neq 0} hw(u).$$

Example. $f : \mathbb{F}_2^2 \to \mathbb{F}_2$ defined by $f(00) = 0, f(01) = 1, f(10) = 0, f(11) = 0$.

$$f = x[1]x[0] + x[0] = x^3 + x^1 \qquad\qquad \text{degree } 2.$$

Computing the ANF of a function is a necessary step for algebraic attacks. But it is hard:

- costs $s2^{s-1}$, requires to store the full LUT of size 2^s.
- **LWR:** $s = 2nq > 256$.

Algebraic Normal form of LWR

$$\begin{aligned} F^{m,n} : (\mathbb{Z}_{2^q})^n \times (\mathbb{Z}_{2^q})^n &\longrightarrow \quad \mathbb{F}_2 \\ (a, x) \quad &\longmapsto (\langle a, x \rangle)^m \ . \end{aligned}$$

e.g. $m = 2^{q-p}$ is the LSB of the sample.

Algebraic Normal form of LWR

$$\begin{aligned} F^{m,n} : (\mathbb{Z}_{2^q})^n \times (\mathbb{Z}_{2^q})^n &\longrightarrow \quad \mathbb{F}_2 \\ (\boldsymbol{a}, \boldsymbol{x}) \quad &\longmapsto (\langle \boldsymbol{a}, \boldsymbol{x} \rangle)^m \ . \end{aligned}$$

e.g. $m = 2^{q-p}$ is the LSB of the sample.

Notation. Let $\boldsymbol{u}, \boldsymbol{v} \in (\mathbb{Z}_{2^q})^n$.

$$\boldsymbol{a}^{\boldsymbol{u}} \boldsymbol{x}^{\boldsymbol{v}} := \prod_{i \in \llbracket 0, n \rrbracket} a_i^{u_i} x_i^{v_i} \ .$$

$\alpha_{\boldsymbol{u},\boldsymbol{v}}(F^{m,n}) \in \mathbb{F}_2$ coefficient of $\boldsymbol{a}^{\boldsymbol{u}} \boldsymbol{x}^{\boldsymbol{v}}$.

Example.

$$\begin{aligned} F^{2,2}((a_0, a_1), (x_0, x_1)) &= a_0^1 x_0^1 + a_1^1 x_1^1 + a_0^1 a_1^1 x_0^1 x_1^1 + a_0^1 x_0^2 + a_0^2 x_0^1 + a_1^1 x_1^2 + a_1^2 x_1^1 \\ &= \boldsymbol{a}^{(1,0)} \boldsymbol{x}^{(1,0)} + \boldsymbol{a}^{(0,1)} \boldsymbol{x}^{(0,1)} + \boldsymbol{a}^{(1,1)} \boldsymbol{x}^{(1,1)} + \ldots \end{aligned}$$

$\alpha_{(1,0),(1,0)} = 1$, $\alpha_{(1,0),(0,1)} = 0$.

Algebraic Normal form of LWR

$$\begin{aligned} F^{m,n} : (\mathbb{Z}_{2^q})^n \times (\mathbb{Z}_{2^q})^n &\longrightarrow \quad \mathbb{F}_2 \\ (\boldsymbol{a}, \boldsymbol{x}) \quad &\longmapsto (\langle \boldsymbol{a}, \boldsymbol{x} \rangle)^m \ . \end{aligned}$$

e.g. $m = 2^{q-p}$ is the LSB of the sample.

Notation. Let $\boldsymbol{u}, \boldsymbol{v} \in (\mathbb{Z}_{2^q})^n$.

$$\boldsymbol{a}^{\boldsymbol{u}} \boldsymbol{x}^{\boldsymbol{v}} := \prod_{i \in [\![0,n]\!]} a_i^{u_i} x_i^{v_i} \ .$$

$\alpha_{\boldsymbol{u},\boldsymbol{v}}(F^{m,n}) \in \mathbb{F}_2$ coefficient of $\boldsymbol{a}^{\boldsymbol{u}} \boldsymbol{x}^{\boldsymbol{v}}$.

Example.

$$\begin{aligned} F^{2,2}((a_0, a_1), (x_0, x_1)) &= a_0^1 x_0^1 + a_1^1 x_1^1 + a_0^1 a_1^1 x_0^1 x_1^1 + a_0^1 x_0^2 + a_0^2 x_0^1 + a_1^1 x_1^2 + a_1^2 x_1^1 \\ &= \boldsymbol{a}^{(1,0)} \boldsymbol{x}^{(1,0)} + \boldsymbol{a}^{(0,1)} \boldsymbol{x}^{(0,1)} + \boldsymbol{a}^{(1,1)} \boldsymbol{x}^{(1,1)} + \dots \end{aligned}$$

$\alpha_{(1,0),(1,0)} = 1$, $\alpha_{(1,0),(0,1)} = 0$.

- If $\boldsymbol{a} = (2, 1)$ we get the equation

$$x_1^1 + x_0^1 + x_1^2 = \boldsymbol{x}^{(0,1)} + \boldsymbol{x}^{(1,0)} + \boldsymbol{x}^{(0,2)} = \text{ some bit of } s_{\boldsymbol{a}} \ .$$

Algebraic Normal form of LWR

$$\begin{aligned} F^{m,n} : (\mathbb{Z}_{2^q})^n \times (\mathbb{Z}_{2^q})^n &\longrightarrow \quad \mathbb{F}_2 \\ (\boldsymbol{a}, \boldsymbol{x}) \quad &\longmapsto (\langle \boldsymbol{a}, \boldsymbol{x} \rangle)^m \ . \end{aligned}$$

e.g. $m = 2^{q-p}$ is the LSB of the sample.

Notation. Let $\boldsymbol{u}, \boldsymbol{v} \in (\mathbb{Z}_{2^q})^n$.

$$\boldsymbol{a}^{\boldsymbol{u}} \boldsymbol{x}^{\boldsymbol{v}} := \prod_{i \in [\![0,n]\!]} a_i^{u_i} x_i^{v_i} \ .$$

$\alpha_{\boldsymbol{u},\boldsymbol{v}}(F^{m,n}) \in \mathbb{F}_2$ coefficient of $\boldsymbol{a}^{\boldsymbol{u}} \boldsymbol{x}^{\boldsymbol{v}}$.

Example.

$$\begin{aligned} F^{2,2}((a_0, a_1), (x_0, x_1)) &= a_0^1 x_0^1 + a_1^1 x_1^1 + a_0^1 a_1^1 x_0^1 x_1^1 + a_0^1 x_0^2 + a_0^2 x_0^1 + a_1^1 x_1^2 + a_1^2 x_1^1 \\ &= \boldsymbol{a}^{(1,0)} \boldsymbol{x}^{(1,0)} + \boldsymbol{a}^{(0,1)} \boldsymbol{x}^{(0,1)} + \boldsymbol{a}^{(1,1)} \boldsymbol{x}^{(1,1)} + \dots \end{aligned}$$

$\alpha_{(1,0),(1,0)} = 1$, $\alpha_{(1,0),(0,1)} = 0$.

- If $\boldsymbol{a} = (2, 1)$ we get the equation

$$x_1^1 + x_0^1 + x_1^2 = \boldsymbol{x}^{(0,1)} + \boldsymbol{x}^{(1,0)} + \boldsymbol{x}^{(0,2)} = \text{ some bit of } s_{\boldsymbol{a}} \ .$$

Complexity of solving: Nr of monomials$^{\omega}$ $= |\mathrm{Exp}_{\boldsymbol{x}}(F^{m,n})|^{\omega}$ with $\mathrm{Exp}_{\boldsymbol{x}}(F^{m,n}) = \{\boldsymbol{v} \mid \boldsymbol{x}^{\boldsymbol{v}} \text{ appears in } F^{m,n}\}$.

Outline

1 Introduction: motivation and setting

2 A symmetric point of view

3 Results

4 Some intuition

5 A taste of the algorithms

A ‘free’ improvement

- **Degree.** $\deg(F^{m,n}) \leq m$. If $m \leq n$, we can also show $\deg_{x}(F^{m,n}) = m$.
- **Number of monomials.** $|\mathrm{Exp}_{x}(F^{m,n})| \leq \binom{n+m}{m}$. Conjectured equality when m is a power of two.

A ‘free’ improvement

- **Degree.** $\deg(F^{m,n}) \leq m$. If $m \leq n$, we can also show $\deg_{x}(F^{m,n}) = m$.
- **Number of monomials.** $|\mathrm{Exp}_{x}(F^{m,n})| \leq \binom{n+m}{m}$. Conjectured equality when m is a power of two.

- Recall over $\mathbb{Z}_{2^q}$: $\binom{n+2^{q-p}}{2^{q-p}}$ monomials (Arora & Ge).

A ‘free’ improvement

- **Degree.** $\deg(F^{m,n}) \leq m$. If $m \leq n$, we can also show $\deg_{x}(F^{m,n}) = m$.
- **Number of monomials.** $|\mathrm{Exp}_{x}(F^{m,n})| \leq \binom{n+m}{m}$. Conjectured equality when m is a power of two.

- Recall over $\mathbb{Z}_{2^q}$: $\binom{n+2^{q-p}}{2^{q-p}}$ monomials (Arora & Ge).
- If we use the LSB of the sample ($m = 2^{q-p}$), same number of monomials (Surprising!).

A ‘free’ improvement

- **Degree.** $\deg(F^{m,n}) \leq m$. If $m \leq n$, we can also show $\deg_x(F^{m,n}) = m$.
- **Number of monomials.** $|\mathrm{Exp}_x(F^{m,n})| \leq \binom{n+m}{m}$. Conjectured equality when m is a power of two.

- Recall over $\mathbb{Z}_{2^q}$: $\binom{n+2^{q-p}}{2^{q-p}}$ monomials (Arora & Ge).
- If we use the LSB of the sample ($m = 2^{q-p}$), same number of monomials (Surprising!).
- Free benefit: we work over $\mathbb{F}_2$ rather than $\mathbb{Z}_{2^q}$:
  1. operations are cheaper (we gain at least q).
  2. it’s a field.

A ‘free’ improvement

- **Degree.** $\deg(F^{m,n}) \leq m$. If $m \leq n$, we can also show $\deg_{x}(F^{m,n}) = m$.
- **Number of monomials.** $|\mathrm{Exp}_{x}(F^{m,n})| \leq \binom{n+m}{m}$. Conjectured equality when m is a power of two.

- Recall over $\mathbb{Z}_{2^q}$: $\binom{n+2^{q-p}}{2^{q-p}}$ monomials (Arora & Ge).
- If we use the LSB of the sample ($m = 2^{q-p}$), same number of monomials (Surprising!).
- Free benefit: we work over $\mathbb{F}_2$ rather than $\mathbb{Z}_{2^q}$:
  1. operations are cheaper (we gain at least q).
  2. it’s a field.

NB: Once the $q - p$ LSBs are recovered, the cost of recovering the full secret is negligible.

Set of exponents

Let S_k^n be the set of ordered integer partitions of k of length n . $|S_k^n| = \binom{n+k-1}{k}$.

Combining results from Braeken and Semaev (FSE, 2005) on the ANF of addition and multiplication:

Theorem (Exponents).

$$\mathrm{Exp}_{\boldsymbol{x}}(F^{m,n}) \subset \bigcup_{i=1}^{m} S_i^n \, .$$

Set of exponents

Let S_k^n be the set of ordered integer partitions of k of length n . $|S_k^n| = \binom{n+k-1}{k}$.

Combining results from Braeken and Semaev (FSE, 2005) on the ANF of addition and multiplication:

Theorem (Exponents).

$$\mathrm{Exp}_{\boldsymbol{x}}(F^{m,n}) \subset \bigcup_{i=1}^{m} S_i^n .$$

$\mathrm{Exp}_{\boldsymbol{x}}(F^{2,4}) = S_2^4 \cup S_1^4$.

$$S_2^4 = \{(2,0,0,0),(0,2,0,0),(0,0,2,0),(0,0,0,2),(1,1,0,0),$$
$$(1,0,1,0),(1,0,0,1),(0,1,1,0),(0,1,0,1),(0,0,1,1)\}$$
$$S_1^4 = \{(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\} .$$

Thus possible $\boldsymbol{x}^{\boldsymbol{v}}$: x_0^2, x_1^2 , . . . , $x_0^1 x_1^1$ . . . , x_3^1

Set of exponents

Let S_k^n be the set of ordered integer partitions of k of length n . $|S_k^n| = \binom{n+k-1}{k}$.

Combining results from Braeken and Semaev (FSE, 2005) on the ANF of addition and multiplication:

Theorem (Exponents).

$$\mathrm{Exp}_{\boldsymbol{x}}(F^{m,n}) \subset \bigcup_{i=1}^{m} S_i^n \,.$$

$\mathrm{Exp}_{\boldsymbol{x}}(F^{2,4}) = S_2^4 \cup S_1^4$.

$$S_2^4 = \{(2,0,0,0),(0,2,0,0),(0,0,2,0),(0,0,0,2),(1,1,0,0),$$
$$(1,0,1,0),(1,0,0,1),(0,1,1,0),(0,1,0,1),(0,0,1,1)\}$$
$$S_1^4 = \{(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\}\,.$$

Thus possible $\boldsymbol{x}^{\boldsymbol{v}}$: $x_0^2,\ x_1^2\ ,\ldots,\ x_0^1x_1^1\ \ldots,x_3^1$

Work in progress: equality when m a power of two.

Effective computation of the ANF

We improve Arora & Ge generically... .

Effective computation of the ANF

We improve Arora & Ge generically... as long as we can compute the ANF.

Effective computation of the ANF

We improve Arora & Ge generically... as long as we can compute the ANF.

Computing the ANF is the bottleneck:

- **Möbius transform** costs $s2^{s-1}$ in time and 2^s in memory where s is the number of variables.
- **Number of variables.** $F^{m,n}$ depends on at most $s = n \cdot (\lfloor \log_2(m) \rfloor + 1)$ variables of $\boldsymbol{x}$.

Our result

The ANF (and some additional properties) can be 'computed' for arbitrary large n and for m up to 16.

In practice, we:

- stored the ANF whenever possible.
- computed relevant parameters for algebraic attacks (number of monomials and more) for 'arbitrary' large n.

Additional improvement of the attack

Default approach: linearisation basis is the set of monomials $\boldsymbol{x}^{\boldsymbol{v}}$ that appear in the ANF of $F^{m,n}$.

Example.

$$\begin{aligned} F((a_0, a_1), (x_0, x_1)) &= a_0^1 x_0^1 + a_0^1 x_1^1 + a_0^1 a_1^1 x_0^1 x_1^1 + a_0^1 x_0^2 + a_0^2 x_0^1 + a_1^1 x_1^2 + a_0^2 x_1^1 \\ &= (a_0^1 + a_0^2) x_0^1 + (a_0^1 + a_0^2) x_1^1 + \ldots \end{aligned}$$

- Monomials x_0^1 and x_1^1 have the same coefficient $\alpha = a_0^1 + a_0^2$: $\forall \boldsymbol{a}$, they appear/vanish together.
- It makes sense to use $y = \boldsymbol{x}^{(1,0)} + \boldsymbol{x}^{(0,1)}$ as linear variable.

Additional improvement of the attack

Default approach: linearisation basis is the set of monomials $\boldsymbol{x}^{\boldsymbol{v}}$ that appear in the ANF of $F^{m,n}$.

Example.

$$\begin{aligned} F((a_0, a_1),(x_0, x_1)) &= a_0^1 x_0^1 + a_0^1 x_1^1 + a_0^1 a_1^1 x_0^1 x_1^1 + a_0^1 x_0^2 + a_0^2 x_0^1 + a_1^1 x_1^2 + a_0^2 x_1^1 \\ &= (a_0^1 + a_0^2) x_0^1 + (a_0^1 + a_0^2) x_1^1 + \ldots \end{aligned}$$

- Monomials x_0^1 and x_1^1 have the same coefficient $\alpha = a_0^1 + a_0^2$: $\forall \boldsymbol{a}$, they appear/vanish together.
- It makes sense to use $y = \boldsymbol{x}^{(1,0)} + \boldsymbol{x}^{(0,1)}$ as linear variable.

- Linear system in less variables: improved attack (data and time).
- Ideal situation: being able to compute the rank and an associated basis (work in progress).

Additional improvement of the attack

Default approach: linearisation basis is the set of monomials $\boldsymbol{x}^{\boldsymbol{v}}$ that appear in the ANF of $F^{m,n}$.

Example.

$$F((a_0, a_1), (x_0, x_1)) = a_0^1 x_0^1 + a_0^1 x_1^1 + a_0^1 a_1^1 x_0^1 x_1^1 + a_0^1 x_0^2 + a_0^2 x_0^1 + a_1^1 x_1^2 + a_0^2 x_1^1$$
$$= (a_0^1 + a_0^2) x_0^1 + (a_0^1 + a_0^2) x_1^1 + \dots$$

- Monomials x_0^1 and x_1^1 have the same coefficient $\alpha = a_0^1 + a_0^2$: $\forall \boldsymbol{a}$, they appear/vanish together.
- It makes sense to use $y = \boldsymbol{x}^{(1,0)} + \boldsymbol{x}^{(0,1)}$ as linear variable.

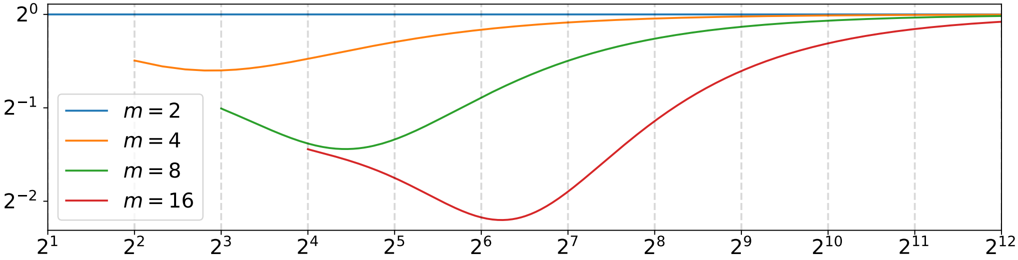
- Linear system in less variables: improved attack (data and time).
- Ideal situation: being able to compute the rank and an associated basis (work in progress).

Additional parameters (fixed m, arbitrary large n)

- Let $\alpha_{\boldsymbol{v}}$ denote the coeff of $\boldsymbol{x}^{\boldsymbol{v}}$, we compute the number of distinct $\alpha_{\boldsymbol{v}}$'s and associated generating family $Q^{m,n}$ (new y type variables). $m \leq 16$.
- The average sparsity in both generating families (monomials, $Q^{m,n}$). $m \leq 8$.

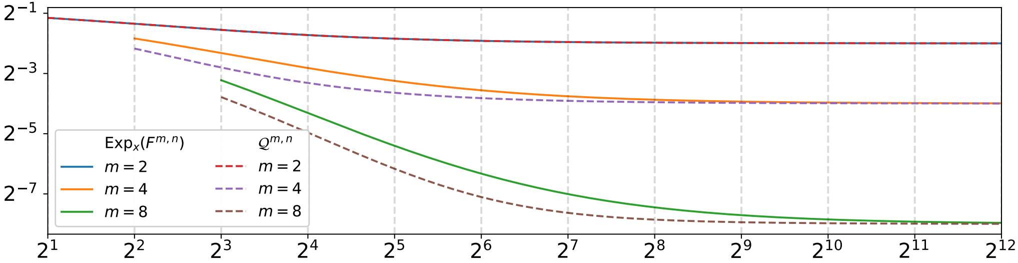
Our results: Upper bound on the rank

Ratio $\left|Q^{m,n}\right| / \binom{n+m}{m} - 1$ **as a function of** n**, for** $n \in [\![m, 4096]\!]$.



Our results: Sparsity

Average fraction of terms in a random equation $F_{\boldsymbol{a}}^{m,n}(\boldsymbol{x})$ as a function of n.



Improvements over Arora-Ge

Comparison to the linearisation attack by Arora & Ge using $\omega = 3$.

and cost of modular addition/multiplication $\approx q - p$.

$m = 2^{q-p}$	n	Arora-Ge	Our work ($\leq$ rank only)	Our work ($\leq$ rank and sparsity)
8	64	$2^{106.4}$	$2^{97.8}$	$2^{87.2}$
	128	$2^{129.3}$	$2^{121.8}$	$2^{110.8}$
	256	$2^{152.7}$	$2^{145.9}$	$2^{134.7}$
16	64	$2^{170.7}$	$2^{157.2}$	Non-available
	128	$2^{214.6}$	$2^{202.00}$	Non-available
	256	$2^{260.5}$	$2^{250.1}$	Non-available

Outline

1 Introduction: motivation and setting

2 A symmetric point of view

3 Results

4 Some intuition

5 A taste of the algorithms

Overview

Our result

The ANF (and some additional properties) can be computed for arbitrary large n and for m up to 16.

NB: used in practice! $m = 8$ for SABER, $m = 16$ for LaKEY.

Some intuition. Commutativity of modular addition creates strong symmetries in the ANF.

$F^{m,n}(\boldsymbol{a}, \boldsymbol{x}) = (\langle \boldsymbol{a}, \boldsymbol{x} \rangle)^m = (\sum_{0 \leq i < n} a_i \times x_i)^m$.

We define $(0\ 1) \cdot \boldsymbol{a} := (a_1, a_0, a_2, \ldots, a_{n-1}), (0\ 1) \cdot \boldsymbol{x} := (x_1, x_0, x_2, \ldots, x_{n-1})$.

Then

$$(\boldsymbol{a}, \boldsymbol{x}) \mapsto F^{m,n}(\boldsymbol{a}, \boldsymbol{x}) \qquad \text{and} \qquad (\boldsymbol{a}, \boldsymbol{x}) \mapsto F^{m,n}((0\ 1) \cdot \boldsymbol{a}, (0\ 1) \cdot \boldsymbol{x})$$

are the same functions since $a_0 \times x_0 + a_1 \times x_1 = a_1 \times x_1 + a_0 \times x_0$.

- Generalizes to any arbitrary permutation σ in $\mathfrak{S}_n$: notion of $\mathfrak{S}_n$-invariant function (not symmetric!).

Reduction to a system of representatives

Example. The function

$$F : (\boldsymbol{a}, \boldsymbol{x}) \mapsto \boldsymbol{a}^{(1,0)}\boldsymbol{x}^{(3,0)} + \boldsymbol{a}^{(0,1)}\boldsymbol{x}^{(0,3)} = a_0^1 x_0^3 + a_1^1 x_1^3$$

is $\mathfrak{S}_2$-invariant (not symmetric!) since $\mathfrak{S}_2 = \{\mathrm{id}, (0\ 1)\}$

$$(0\ 1) \cdot F = F((0\ 1) \cdot \boldsymbol{a}, (0\ 1) \cdot \boldsymbol{x}) = F(\boldsymbol{a}, \boldsymbol{x})\,.$$

Reduction to a system of representatives

Example. The function

$$F : (\boldsymbol{a}, \boldsymbol{x}) \mapsto \boldsymbol{a}^{(1,0)}\boldsymbol{x}^{(3,0)} + \boldsymbol{a}^{(0,1)}\boldsymbol{x}^{(0,3)} = a_0^1 x_0^3 + a_1^1 x_1^3$$

is $\mathfrak{S}_2$-invariant (not symmetric!) since $\mathfrak{S}_2 = \{\mathrm{id}, (0\ 1)\}$

$$(0\ 1) \cdot F = F((0\ 1) \cdot \boldsymbol{a}, (0\ 1) \cdot \boldsymbol{x}) = F(\boldsymbol{a}, \boldsymbol{x})\,.$$

Remarks: $\alpha_{\boldsymbol{u},\boldsymbol{v}} \in \mathbb{F}_2$ coefficient of $\boldsymbol{a}^{\boldsymbol{u}}\boldsymbol{x}^{\boldsymbol{v}}$, $\alpha_{\boldsymbol{v}} \in \mathbb{F}_2[\boldsymbol{a}]$ (mod field equations) coeff of $\boldsymbol{x}^{\boldsymbol{v}}$.

- $\alpha_{(1,0),(3,0)} = \alpha_{(0,1),(0,3)}$.
- $\alpha_{(3,0)} = \boldsymbol{a}^{(1,0)}$, $\alpha_{(0,3)} = \boldsymbol{a}^{(0,1)}$.

Reduction to a system of representatives

Example. The function

$$F : (\boldsymbol{a}, \boldsymbol{x}) \mapsto \boldsymbol{a}^{(1,0)}\boldsymbol{x}^{(3,0)} + \boldsymbol{a}^{(0,1)}\boldsymbol{x}^{(0,3)} = a_0^1 x_0^3 + a_1^1 x_1^3$$

is $\mathfrak{S}_2$-invariant (not symmetric!) since $\mathfrak{S}_2 = \{\mathrm{id}, (0\ 1)\}$

$$(0\ 1) \cdot F = F((0\ 1) \cdot \boldsymbol{a}, (0\ 1) \cdot \boldsymbol{x}) = F(\boldsymbol{a}, \boldsymbol{x})\,.$$

Remarks: $\alpha_{\boldsymbol{u},\boldsymbol{v}} \in \mathbb{F}_2$ coefficient of $\boldsymbol{a}^{\boldsymbol{u}}\boldsymbol{x}^{\boldsymbol{v}}$, $\alpha_{\boldsymbol{v}} \in \mathbb{F}_2[\boldsymbol{a}]$ (mod field equations) coeff of $\boldsymbol{x}^{\boldsymbol{v}}$.

- $\alpha_{(1,0),(3,0)} = \alpha_{(0,1),(0,3)}$.
- $\alpha_{(3,0)} = \boldsymbol{a}^{(1,0)}$, $\alpha_{(0,3)} = \boldsymbol{a}^{(0,1)}$.

Notation. Action of $\mathfrak{S}_n$ on vectors of length n: $\sigma \cdot \boldsymbol{u} := (u_{\sigma^{-1}(0)}, \ldots, u_{\sigma^{-1}(n-1)})$.

The following are equivalent:

- F is $\mathfrak{S}_n$-invariant
- $\alpha_{(\sigma\cdot\boldsymbol{u},\sigma\cdot\boldsymbol{v})}(F) = \alpha_{(\boldsymbol{u},\boldsymbol{v})}(F)$ for any $\boldsymbol{u}, \boldsymbol{v}$, $\sigma \in \mathfrak{S}_n$.
- $\sigma \cdot \alpha_{\boldsymbol{v}}(F) = \alpha_{\sigma^{-1}\cdot\boldsymbol{v}}(F)$ for any $\boldsymbol{v}$, $\sigma \in \mathfrak{S}_n$.

Allows to represent the ANF in a compact way.

Scaling

How do we get results for arbitrary n?

Example. Let's look at the $\mathfrak{S}_3$-invariant function

$$F^{(3)}((a_0, a_1, a_2), (x_0, x_1, x_2)) \mapsto \boldsymbol{a}^{(1,0,0)}\boldsymbol{x}^{(3,0,0)} + \boldsymbol{a}^{(0,1,0)}\boldsymbol{x}^{(0,3,0)} + \boldsymbol{a}^{(0,0,1)}\boldsymbol{x}^{(0,0,1)} .$$

Then

$$F^{(3)}((a_0, a_1, 0), (x_0, x_1, 0)) = \boldsymbol{a}^{(1,0)}\boldsymbol{x}^{(3,0)} + \boldsymbol{a}^{(0,1)}\boldsymbol{x}^{(0,3)} .$$

is $\mathfrak{S}_2$-invariant.

Scaling

How do we get results for arbitrary n?

Example. Let's look at the $\mathfrak{S}_3$-invariant function

$$F^{(3)}((a_0, a_1, a_2), (x_0, x_1, x_2)) \mapsto \boldsymbol{a}^{(1,0,0)}\boldsymbol{x}^{(3,0,0)} + \boldsymbol{a}^{(0,1,0)}\boldsymbol{x}^{(0,3,0)} + \boldsymbol{a}^{(0,0,1)}\boldsymbol{x}^{(0,0,1)} .$$

Then

$$F^{(3)}((a_0, a_1, 0), (x_0, x_1, 0)) = \boldsymbol{a}^{(1,0)}\boldsymbol{x}^{(3,0)} + \boldsymbol{a}^{(0,1)}\boldsymbol{x}^{(0,3)} .$$

is $\mathfrak{S}_2$-invariant.

- Conversely, all terms of support of cardinal at most m in the ANF of $F^{(n)}$ can be understood from the ANF of $F^{(m)}$.

Scaling

How do we get results for arbitrary n?

Example. Let's look at the $\mathfrak{S}_3$-invariant function

$$F^{(3)}((a_0, a_1, a_2), (x_0, x_1, x_2)) \mapsto \boldsymbol{a}^{(1,0,0)} \boldsymbol{x}^{(3,0,0)} + \boldsymbol{a}^{(0,1,0)} \boldsymbol{x}^{(0,3,0)} + \boldsymbol{a}^{(0,0,1)} \boldsymbol{x}^{(0,0,1)} .$$

Then

$$F^{(3)}((a_0, a_1, 0), (x_0, x_1, 0)) = \boldsymbol{a}^{(1,0)} \boldsymbol{x}^{(3,0)} + \boldsymbol{a}^{(0,1)} \boldsymbol{x}^{(0,3)} .$$

is $\mathfrak{S}_2$-invariant.

- Conversely, all terms of support of cardinal at most m in the ANF of $F^{(n)}$ can be understood from the ANF of $F^{(m)}$.

- **Using** m you can bound the cardinal of the support.

(Properties of) $F^{m,n}$ can be derived from $F^{m,m}$.

Outline

1 Introduction: motivation and setting

2 A symmetric point of view

3 Results

4 Some intuition

5 A taste of the algorithms

Group actions

Let $(\mathbb{G}, \circ)$ a group, E a set, $(\sigma, x) \longmapsto \sigma \cdot x$ a group action.

- **Orbit.** For any $x \in E$, $\mathrm{Orb}(x) = \{\sigma \cdot x,\ \sigma \in \mathbb{G}\} \subset E$.
- **Stabilizer.** For any $x \in E$, $\mathrm{Stab}(x) = \{\sigma,\ \sigma \cdot x = x\}$ subgroup of $\mathbb{G}$.

Example. Let $\boldsymbol{v} = (3, 1, 1)$, $\mathfrak{S}_3 = \{\mathrm{id}, (0\ 1), (0\ 2), (2\ 1), (0\ 1\ 2), (0\ 2\ 1)\}$.

- $\mathrm{Orb}(\boldsymbol{v}) = \{(3, 1, 1), (1, 3, 1), (1, 1, 3)\}$.
- $\mathrm{Stab}(\boldsymbol{v}) = \{\mathrm{id}, (1\ 2)\}$.

Important properties

- The set of orbits $\{\mathrm{Orb}(x),\ x \in E\}$ is a partition of E.
- It thus induces an equivalence relation $\sim$ defined as $x \sim x'$ if and only if $x' \in \mathrm{Orb}(x)$.
- For any $x \in E$,

$$|\mathrm{Orb}(x)| \cdot |\mathrm{Stab}(x)| = |\mathbb{G}| \ .$$

Effective computation of the ANF

Recall: S_m^n ordered partitions of length n of m. Let $\mathscr{C}_m^n$ be a system of representatives (unordered partitions).

$$F^{m,n}(\boldsymbol{a},\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} \sum_{\boldsymbol{c}'\in\mathrm{Orb}(\boldsymbol{c})} \prod_{i=0}^{n-1} (a_i \times x_i)^{c_i'} = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} \sum_{\boldsymbol{c}'\in\mathrm{Orb}(\boldsymbol{c})} H_{\boldsymbol{c}'}(\boldsymbol{a},\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} G_{\boldsymbol{c}}(\boldsymbol{a},\boldsymbol{x})\,.$$

Effective computation of the ANF

Recall: S_m^n ordered partitions of length n of m. Let $\mathscr{C}_m^n$ be a system of representatives (unordered partitions).

$$F^{m,n}(\boldsymbol{a},\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} \sum_{\boldsymbol{c}'\in\mathrm{Orb}(\boldsymbol{c})} \prod_{i=0}^{n-1}(a_i \times x_i)^{c_i'} = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} \sum_{\boldsymbol{c}'\in\mathrm{Orb}(\boldsymbol{c})} H_{\boldsymbol{c}'}(\boldsymbol{a},\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} G_{\boldsymbol{c}}(\boldsymbol{a},\boldsymbol{x})\,.$$

Effective computation of the ANF

Recall: S_m^n ordered partitions of length n of m. Let $\mathscr{C}_m^n$ be a system of representatives (unordered partitions).

$$F^{m,n}(\boldsymbol{a},\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} \sum_{\boldsymbol{c}'\in\mathrm{Orb}(\boldsymbol{c})} \prod_{i=0}^{n-1} (a_i \times x_i)^{c_i'} = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} \sum_{\boldsymbol{c}'\in\mathrm{Orb}(\boldsymbol{c})} H_{\boldsymbol{c}'}(\boldsymbol{a},\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} G_{\boldsymbol{c}}(\boldsymbol{a},\boldsymbol{x})\,.$$

Theorem. Let $\boldsymbol{c} \in \mathbb{Z}_{2^q}^n$, $(\boldsymbol{u},\boldsymbol{v}) \in \mathbb{Z}_{2^q}^n \times \mathbb{Z}_{2^q}^n$. Define $E_{\boldsymbol{c}}(\boldsymbol{u},\boldsymbol{v}) := \mathrm{Exp}(H_{\boldsymbol{c}}) \cap \mathrm{Orb}(\boldsymbol{u},\boldsymbol{v})$. For any $\sigma \in \mathfrak{S}_n$,

$$\alpha_{\sigma\cdot(\boldsymbol{u},\boldsymbol{v})}(G_{\boldsymbol{c}}) \quad = \quad \alpha_{(\boldsymbol{u},\boldsymbol{v})}(G_{\boldsymbol{c}}) \quad = \quad \frac{|E_{\boldsymbol{c}}(\boldsymbol{u},\boldsymbol{v})|\; n!}{|\mathrm{Stab}(\boldsymbol{c})|\,|\mathrm{Orb}(\boldsymbol{u},\boldsymbol{v})|} \mod 2\,.$$

Effective computation of the ANF

Recall: S_m^n ordered partitions of length n of m. Let $\mathscr{C}_m^n$ be a system of representatives (unordered partitions).

$$F^{m,n}(\boldsymbol{a},\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} \sum_{\boldsymbol{c}'\in\mathrm{Orb}(\boldsymbol{c})} \prod_{i=0}^{n-1} (a_i \times x_i)^{c_i'} = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} \sum_{\boldsymbol{c}'\in\mathrm{Orb}(\boldsymbol{c})} H_{\boldsymbol{c}'}(\boldsymbol{a},\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} G_{\boldsymbol{c}}(\boldsymbol{a},\boldsymbol{x})\,.$$

Theorem. Let $\boldsymbol{c} \in \mathbb{Z}_{2^q}^n$, $(\boldsymbol{u},\boldsymbol{v}) \in \mathbb{Z}_{2^q}^n \times \mathbb{Z}_{2^q}^n$. Define $E_{\boldsymbol{c}}(\boldsymbol{u},\boldsymbol{v}) := \mathrm{Exp}(H_{\boldsymbol{c}}) \cap \mathrm{Orb}(\boldsymbol{u},\boldsymbol{v})$. For any $\sigma \in \mathfrak{S}_n$,

$$\alpha_{\sigma\cdot(\boldsymbol{u},\boldsymbol{v})}(G_{\boldsymbol{c}}) \quad = \quad \alpha_{(\boldsymbol{u},\boldsymbol{v})}(G_{\boldsymbol{c}}) \quad = \quad \frac{|E_{\boldsymbol{c}}(\boldsymbol{u},\boldsymbol{v})|\; n!}{|\mathrm{Stab}(\boldsymbol{c})|\;|\mathrm{Orb}(\boldsymbol{u},\boldsymbol{v})|} \mod 2\,.$$

(i) For all canonical $\boldsymbol{c}$, compute $\{(\boldsymbol{u},\boldsymbol{v})^\star, \alpha_{(\boldsymbol{u},\boldsymbol{v})^\star}(G_{\boldsymbol{c}}) = 1\}$ from $H_{\boldsymbol{c}'}$ for some $\boldsymbol{c}' \in \mathrm{Orb}(\boldsymbol{c})$. (Theorem)

Effective computation of the ANF

Recall: S_m^n ordered partitions of length n of m. Let $\mathscr{C}_m^n$ be a system of representatives (unordered partitions).

$$F^{m,n}(\boldsymbol{a},\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} \sum_{\boldsymbol{c}'\in\mathrm{Orb}(\boldsymbol{c})} \prod_{i=0}^{n-1} (a_i \times x_i)^{c_i'} = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} \sum_{\boldsymbol{c}'\in\mathrm{Orb}(\boldsymbol{c})} H_{\boldsymbol{c}'}(\boldsymbol{a},\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} G_{\boldsymbol{c}}(\boldsymbol{a},\boldsymbol{x}) \,.$$

Theorem. Let $\boldsymbol{c} \in \mathbb{Z}_{2^q}^n$, $(\boldsymbol{u},\boldsymbol{v}) \in \mathbb{Z}_{2^q}^n \times \mathbb{Z}_{2^q}^n$. Define $E_{\boldsymbol{c}}(\boldsymbol{u},\boldsymbol{v}) := \mathrm{Exp}(H_{\boldsymbol{c}}) \cap \mathrm{Orb}(\boldsymbol{u},\boldsymbol{v})$. For any $\sigma \in \mathfrak{S}_n$,

$$\alpha_{\sigma\cdot(\boldsymbol{u},\boldsymbol{v})}(G_{\boldsymbol{c}}) \quad = \quad \alpha_{(\boldsymbol{u},\boldsymbol{v})}(G_{\boldsymbol{c}}) \quad = \quad \frac{|E_{\boldsymbol{c}}(\boldsymbol{u},\boldsymbol{v})|\; n!}{|\mathrm{Stab}(\boldsymbol{c})|\,|\mathrm{Orb}(\boldsymbol{u},\boldsymbol{v})|} \mod 2\,.$$

(i) For all canonical $\boldsymbol{c}$, compute $\{(\boldsymbol{u},\boldsymbol{v})^\star, \alpha_{(\boldsymbol{u},\boldsymbol{v})^\star}(G_{\boldsymbol{c}}) = 1\}$ from $H_{\boldsymbol{c}'}$ for some $\boldsymbol{c}' \in \mathrm{Orb}(\boldsymbol{c})$. (Theorem)

(ii) Compute a SOR of the ANF of $F^{m,n}$ from the $G_{\boldsymbol{c}}$'s i.e. $\{(\boldsymbol{u},\boldsymbol{v})^\star, \alpha_{(\boldsymbol{u},\boldsymbol{v})^\star}(F^{m,n}) = 1\}$.

Effective computation of the ANF

Recall: S_m^n ordered partitions of length n of m. Let $\mathscr{C}_m^n$ be a system of representatives (unordered partitions).

$$F^{m,n}(\boldsymbol{a},\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} \sum_{\boldsymbol{c}'\in\mathrm{Orb}(\boldsymbol{c})} \prod_{i=0}^{n-1}(a_i \times x_i)^{c_i'} = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} \sum_{\boldsymbol{c}'\in\mathrm{Orb}(\boldsymbol{c})} H_{\boldsymbol{c}'}(\boldsymbol{a},\boldsymbol{x}) = \sum_{\boldsymbol{c}\in\mathscr{C}_m^n} G_{\boldsymbol{c}}(\boldsymbol{a},\boldsymbol{x})\,.$$

Theorem. Let $\boldsymbol{c} \in \mathbb{Z}_{2^q}^n$, $(\boldsymbol{u},\boldsymbol{v}) \in \mathbb{Z}_{2^q}^n \times \mathbb{Z}_{2^q}^n$. Define $E_{\boldsymbol{c}}(\boldsymbol{u},\boldsymbol{v}) := \mathrm{Exp}(H_{\boldsymbol{c}}) \cap \mathrm{Orb}(\boldsymbol{u},\boldsymbol{v})$. For any $\sigma \in \mathfrak{S}_n$,

$$\alpha_{\sigma\cdot(\boldsymbol{u},\boldsymbol{v})}(G_{\boldsymbol{c}}) \quad = \quad \alpha_{(\boldsymbol{u},\boldsymbol{v})}(G_{\boldsymbol{c}}) \quad = \quad \frac{|E_{\boldsymbol{c}}(\boldsymbol{u},\boldsymbol{v})|\; n!}{|\mathrm{Stab}(\boldsymbol{c})|\,|\mathrm{Orb}(\boldsymbol{u},\boldsymbol{v})|} \mod 2\,.$$

(i) For all canonical $\boldsymbol{c}$, compute $\{(\boldsymbol{u},\boldsymbol{v})^\star, \alpha_{(\boldsymbol{u},\boldsymbol{v})^\star}(G_{\boldsymbol{c}}) = 1\}$ from $H_{\boldsymbol{c}'}$ for some $\boldsymbol{c}' \in \mathrm{Orb}(\boldsymbol{c})$. (Theorem)

(ii) Compute a SOR of the ANF of $F^{m,n}$ from the $G_{\boldsymbol{c}}$'s i.e. $\{(\boldsymbol{u},\boldsymbol{v})^\star, \alpha_{(\boldsymbol{u},\boldsymbol{v})^\star}(F^{m,n}) = 1\}$.

(iii) Compute the multiset $\{\{\alpha_{\boldsymbol{v}^\star}(F^{m,n}) \neq 0\}\}$.

NB: $\{\alpha_{\boldsymbol{v}^\star}(F^{m,n}) \neq 0\}$ is not a SOR of $\{\alpha_{\boldsymbol{v}}(F^{m,n}) \neq 0\}/\sim$.

Effective computation of the number of monomials

Recall. For each random $\boldsymbol{a}$, the function in $\{F_{\boldsymbol{a}}^{m,n} := F^{m,n}(\boldsymbol{a},\cdot)\}$ is evaluated on the secret $\boldsymbol{x}$ to yield $s_{\boldsymbol{a}}$.

If we can compute the ANF of $F_{\boldsymbol{a}}^{m,n}$, we obtain an equation in the secret $\boldsymbol{x}$:

$$\sum_{\boldsymbol{v}\in(\mathbb{F}_{\mathbf{2}}^{q})^{n}} \alpha_{\boldsymbol{v}}(F_{\boldsymbol{a}}^{m,n})\ \boldsymbol{x}^{\boldsymbol{v}} = \sum_{\boldsymbol{v}\in(\mathbb{F}_{\mathbf{2}}^{q})^{n}} \alpha_{\boldsymbol{v}}(F^{m,n})(\boldsymbol{a})\ \boldsymbol{x}^{\boldsymbol{v}} = (s_{\boldsymbol{a}})^{m/2^{q-p}}\ .$$

Effective computation of the number of monomials

Recall. For each random $\boldsymbol{a}$, the function in $\{F_{\boldsymbol{a}}^{m,n} := F^{m,n}(\boldsymbol{a}, \cdot)\}$ is evaluated on the secret $\boldsymbol{x}$ to yield $s_{\boldsymbol{a}}$.

If we can compute the ANF of $F_{\boldsymbol{a}}^{m,n}$, we obtain an equation in the secret $\boldsymbol{x}$:

$$\sum_{\boldsymbol{v} \in (\mathbb{F}_2^q)^n} \alpha_{\boldsymbol{v}}(F_{\boldsymbol{a}}^{m,n})\ \boldsymbol{x}^{\boldsymbol{v}} = \sum_{\boldsymbol{v} \in (\mathbb{F}_2^q)^n} \alpha_{\boldsymbol{v}}(F^{m,n})(\boldsymbol{a})\ \boldsymbol{x}^{\boldsymbol{v}} = (s_{\boldsymbol{a}})^{m/2^{q-p}}\ .$$

NB: Number of monomials in the system:

$$\left| \bigcup_{\boldsymbol{a} \in \mathbb{Z}_{2^q}^n} \{\boldsymbol{v}, \alpha_{\boldsymbol{v}}(F_{\boldsymbol{a}}^{m,n}) \neq 0\} \right| = \left| \{\boldsymbol{v}, \alpha_{\boldsymbol{v}}(F^{m,n}) \neq 0\} \right|\ .$$

Effective computation of the number of monomials

Recall. For each random $\boldsymbol{a}$, the function in $\{F_{\boldsymbol{a}}^{m,n} := F^{m,n}(\boldsymbol{a}, \cdot)\}$ is evaluated on the secret $\boldsymbol{x}$ to yield $s_{\boldsymbol{a}}$. If we can compute the ANF of $F_{\boldsymbol{a}}^{m,n}$, we obtain an equation in the secret $\boldsymbol{x}$:

$$\sum_{\boldsymbol{v} \in (\mathbb{F}_2^q)^n} \alpha_{\boldsymbol{v}}(F_{\boldsymbol{a}}^{m,n})\ \boldsymbol{x}^{\boldsymbol{v}} = \sum_{\boldsymbol{v} \in (\mathbb{F}_2^q)^n} \alpha_{\boldsymbol{v}}(F^{m,n})(\boldsymbol{a})\ \boldsymbol{x}^{\boldsymbol{v}} = (s_{\boldsymbol{a}})^{m/2^{q-p}} .$$

NB: Number of monomials in the system:

$$\left| \bigcup_{\boldsymbol{a} \in \mathbb{Z}_{2^q}^n} \{\boldsymbol{v}, \alpha_{\boldsymbol{v}}(F_{\boldsymbol{a}}^{m,n}) \neq 0\} \right| = \left|\{\boldsymbol{v}, \alpha_{\boldsymbol{v}}(F^{m,n}) \neq 0\}\right| .$$

- Our algorithm returned the multiset $\{\{\alpha_{\boldsymbol{v}^\star}(F^{m,n}) \neq 0\}\}$, which has cardinal equal to $\left|\{\boldsymbol{v}^\star, \alpha_{\boldsymbol{v}^\star}\} \neq 0\right|$.

  Nr of monomials $= |\{\boldsymbol{v}, \alpha_{\boldsymbol{v}} \neq 0\}| = \sum_{\boldsymbol{v}^\star, \alpha_{\boldsymbol{v}^\star} \neq 0} \mathrm{Orb}(\alpha_{\boldsymbol{v}^\star}) \binom{n}{|\mathrm{Supp}(\boldsymbol{v}^\star)|}$

Effective computation of the number of monomials

Recall. For each random $\boldsymbol{a}$, the function in $\{F_{\boldsymbol{a}}^{m,n} := F^{m,n}(\boldsymbol{a},\cdot)\}$ is evaluated on the secret $\boldsymbol{x}$ to yield $s_{\boldsymbol{a}}$.

If we can compute the ANF of $F_{\boldsymbol{a}}^{m,n}$, we obtain an equation in the secret $\boldsymbol{x}$:

$$\sum_{\boldsymbol{v}\in(\mathbb{F}_{\mathbf{2}}^{q})^{n}} \alpha_{\boldsymbol{v}}(F_{\boldsymbol{a}}^{m,n})\ \boldsymbol{x}^{\boldsymbol{v}} = \sum_{\boldsymbol{v}\in(\mathbb{F}_{\mathbf{2}}^{q})^{n}} \alpha_{\boldsymbol{v}}(F^{m,n})(\boldsymbol{a})\ \boldsymbol{x}^{\boldsymbol{v}} = (s_{\boldsymbol{a}})^{m/2^{q-p}}\ .$$

NB: Number of monomials in the system:

$$\left|\bigcup_{\boldsymbol{a}\in\mathbb{Z}_{\mathbf{2}^q}^{n}} \{\boldsymbol{v}, \alpha_{\boldsymbol{v}}(F_{\boldsymbol{a}}^{m,n}) \neq 0\}\right| = |\{\boldsymbol{v}, \alpha_{\boldsymbol{v}}(F^{m,n}) \neq 0\}|\ .$$

- Our algorithm returned the multiset $\{\{\alpha_{\boldsymbol{v}^\star}(F^{m,n}) \neq 0\}\}$, which has cardinal equal to $\left|\{\boldsymbol{v}^\star, \alpha_{\boldsymbol{v}^\star}\} \neq 0\right|$.

$$\text{Nr of monomials } = |\{\boldsymbol{v}, \alpha_{\boldsymbol{v}} \neq 0\}| = \textstyle\sum_{\boldsymbol{v}^\star, \alpha_{\boldsymbol{v}^\star}\neq 0} \mathrm{Orb}(\alpha_{\boldsymbol{v}^\star})\binom{n}{|\mathrm{Supp}(\boldsymbol{v}^\star)|}$$

In practice, when m is a power of two, this is always equal to the upper bound $\binom{n+m}{m} - 1$.

Conclusion

Results

- Deterministic noise impacts practical complexity.
- Generic improvement over Arora & Ge:
  - same number of monomials + working over $\mathbb{F}_2$.
  - computation of additional parameters.
- Other stuff e.g.: Time-data trade-offs work better in mod/ring LWR.

Many, many open questions

- Understanding the ANF even better (work in progress).
- Help us solve this system (with less data)!
  - E.g. super structured ANF guides guess-and-solve strategies.
- More applications of group actions.