

Bilinear systems arising from code-based cryptography

Algebraic Methods in Post-Quantum Cryptography, August 13

- Summarize what is (un)known about Gröbner basis techniques for bilinear systems.
- Give examples of bilinear systems derived from code-based applications.

Preliminaries

Bilinear polynomials

We consider $R := \mathbb{F}_q[\mathbf{x}, \mathbf{y}]$ with bi-grading

$$R = \bigoplus_{(d_x, d_y) \in \mathbb{N}^2} R_{(d_x, d_y)},$$

where the bi-graded component of bi-degree (d_x, d_y) is defined as

$$R_{(d_x, d_y)} := \text{Span}_{\mathbb{F}_q} \left\{ \text{monomials } \mu \in R, \deg_{\mathbf{x}}(\mu) = d_x, \deg_{\mathbf{y}}(\mu) = d_y \right\}.$$

Bilinear polynomials

We consider $R := \mathbb{F}_q[\mathbf{x}, \mathbf{y}]$ with bi-grading

$$R = \bigoplus_{(d_x, d_y) \in \mathbb{N}^2} R_{(d_x, d_y)},$$

where the bi-graded component of bi-degree (d_x, d_y) is defined as

$$R_{(d_x, d_y)} := \text{Span}_{\mathbb{F}_q} \left\{ \text{monomials } \mu \in R, \deg_{\mathbf{x}}(\mu) = d_x, \deg_{\mathbf{y}}(\mu) = d_y \right\}.$$

A bihomogeneous polynomial belongs to such a bi-graded component, and a bilinear polynomial $f \in R$ corresponds to the special case $f \in R_{(1,1)}$.

Affine bilinear polynomials

To study the affine case, it is convenient to consider the partial order

$$(d'_x, d'_y) \preceq (d_x, d_y) \Leftrightarrow d'_x \leq d_x \text{ and } d'_y \leq d_y,$$

and to set

$$R_{\preceq d} := \bigoplus_{d' \preceq d} R_{d'}.$$

Then an affine bilinear polynomial $f \in R$ is an element $f \in R_{\preceq(1,1)}$.

Affine bilinear polynomials

To study the affine case, it is convenient to consider the partial order

$$(d'_x, d'_y) \preceq (d_x, d_y) \Leftrightarrow d'_x \leq d_x \text{ and } d'_y \leq d_y,$$

and to set

$$R_{\preceq d} := \bigoplus_{d' \preceq d} R_{d'}.$$

Then an affine bilinear polynomial $f \in R$ is an element $f \in R_{\preceq(1,1)}$.

This is also the deshomogenization $\delta(F)$ of some $F \in S_{(1,1)}$, where

$$S := \mathbb{F}_q[x_0, \mathbf{x}, y_0, \mathbf{y}],$$

and

$$\delta : S \longrightarrow R,$$

$$F(x_0, \mathbf{x}, y_0, \mathbf{y}) \longmapsto F(1, \mathbf{x}, 1, \mathbf{y}).$$

A system $\mathcal{F} \in R_{\succeq(1,1)}^m$ is just a particular MQ system, so we can apply a general-purpose Gröbner basis algorithm to solve it.

A system $\mathcal{F} \in R_{\succeq(1,1)}^m$ is just a particular MQ system, so we can apply a general-purpose Gröbner basis algorithm to solve it.

Instead, we may want to exploit the bi-graded structure when constructing Macaulay matrices: **columns indexed by bi-degrees rather than by total degrees.**

A system $\mathcal{F} \in R_{\preceq(1,1)}^m$ is just a particular MQ system, so we can apply a general-purpose Gröbner basis algorithm to solve it.

Instead, we may want to exploit the bi-graded structure when constructing Macaulay matrices: columns indexed by bi-degrees rather than by total degrees.

In either case, we must understand syzygies for the analysis. There are non-trivial ones.

Analysis of syzygies and limitations

Jacobian matrix of $\mathcal{F} \in R^m$ with respect to the $\mathbf{x}$ variables

$$\text{Jac}_{\mathbf{x}}(\mathcal{F}) := \left(\frac{\partial f_i}{\partial x_j} \right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n_x}}.$$

For $\mathcal{F} \in R_{(1,1)}^m$, the entries are linear forms in $\mathbf{y}$.

[FSS11] Faugère, Safey El Din, and Spaenlehauer, “Gröbner bases of bihomogeneous ideals generated by polynomials of bidegree (1,1): Algorithms and complexity”. *J. Symbolic Comput.*

Jacobian matrix of $\mathcal{F} \in R^m$ with respect to the $\mathbf{x}$ variables

$$\text{Jac}_{\mathbf{x}}(\mathcal{F}) := \left(\frac{\partial f_i}{\partial x_j} \right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n_x}}.$$

For $\mathcal{F} \in R_{(1,1)}^m$, the entries are linear forms in $\mathbf{y}$.

Euler's formula gives

$$(h_1, \dots, h_m) \text{Jac}_{\mathbf{x}}(\mathcal{F}) \mathbf{x}^T = \sum_{i=1}^m h_i f_i.$$

This implies that

$$\text{Syz}(\mathcal{F}) \cap \mathbb{F}_q[\mathbf{y}]^m = \text{Ker}(\text{Jac}_{\mathbf{x}}(\mathcal{F})).$$

[FSS11] Faugère, Safey El Din, and Spaenlehauer, “Gröbner bases of bihomogeneous ideals generated by polynomials of bidegree (1,1): Algorithms and complexity”. *J. Symbolic Comput.*

Defined for sequences in $S_{(1,1)}^m$, $m \leq n_x + n_y$. For such bi-regular sequences:

1. Syzygies are the result of trivial syzygies or Jacobian syzygies.
2. Kernels of Jacobians are well understood: generated by vectors of maximal minors.

Defined for sequences in $S_{(1,1)}^m$, $m \leq n_x + n_y$. For such bi-regular sequences:

1. Syzygies are the result of trivial syzygies or Jacobian syzygies.
2. Kernels of Jacobians are well understood: generated by vectors of maximal minors.

Complete analysis in this case

- We know the Hilbert bi-series of bi-regular sequences.
- We can tweak F5 to avoid reductions to zero arising from Jacobian syzygies.

Defined for sequences in $S_{(1,1)}^m$, $m \leq n_x + n_y$. For such bi-regular sequences:

1. Syzygies are the result of trivial syzygies or Jacobian syzygies.
2. Kernels of Jacobians are well understood: generated by vectors of maximal minors.

Complete analysis in this case

- We know the Hilbert bi-series of bi-regular sequences.
- We can tweak F5 to avoid reductions to zero arising from Jacobian syzygies.

It is also shown in [FSS11] that:

- Condition 1. is generic (the proof uses the Jacobian Criterion).
- Bi-regularity is generic for $m < n_x + n_y$.

For a generic affine bilinear system $\mathcal{F} \in R_{(1,1)}^m$ with $m = n_x + n_y$, we have upper bounds on the maximal degree reached by the solving algorithm:

$\min(n_x + 1, n_y + 1)$ for F5 in [FSS11];

$\min(n_x + 2, n_y + 2)$ for F4 in [Spa12].

The first bound is a bound on the degree of regularity. The second bound is tied more directly to the F4 algorithm.

[Spa12] Spaenlehauer, “Solving multi-homogeneous and determinantal systems: algorithms, complexity, applications”.

Overdetermined case: algorithms

Cryptanalysis has motivated several bilinear XL variants. They can be viewed as linearization over a specific subset of monomials.

[BCV23] Baena, Cabarcas, and Verbel, “On the complexity of solving generic overdetermined bilinear systems”. *Advances in Mathematics of Communications*.

[NRS25] Ning, Ran, and Samardjiska, *Multi-homogeneous XL*.

Overdetermined case: algorithms

Cryptanalysis has motivated several bilinear XL variants. They can be viewed as linearization over a specific subset of monomials.

$$\preceq (1, d_y)$$

[BCV23]

$$\preceq (d_x, d_y)$$

Several works

$$\preceq \text{a set of bi-degrees}$$

[NRS25]

Unfortunately, this type of XL approach is not well understood for $m > n_x + n_y$.

[BCV23] Baena, Cabarcas, and Verbel, “On the complexity of solving generic overdetermined bilinear systems”. *Advances in Mathematics of Communications*.

[NRS25] Ning, Ran, and Samardjiska, *Multi-homogeneous XL*.

In the overdetermined case, there is no analogue of [FSS11] to find the Hilbert bi-series.

In the overdetermined case, there is no analogue of [FSS11] to find the Hilbert bi-series.

To **conjecture** Macaulay matrix ranks, several works use the series

$$\frac{(1 - st)^m}{(1 - s)^{n_x}(1 - t)^{n_y}}. \quad (1)$$

Caveats:

- Does (1) accurately capture the random behavior?
- Is there really such a thing as random behavior, like “semi-bi-regularity” ?

As they target bi-degrees only $(1, d_y)$, [BCV23] focus on Jacobian syzygies to control the rank of Macaulay matrices.

As they target bi-degrees only $(1, d_y)$, [BCV23] focus on Jacobian syzygies to control the rank of Macaulay matrices.

For $\mathcal{B} \in R_{(1,1)}^m$ and $(h_1, \dots, h_m) \in R_{(0,d-2)}^m$, consider

$$(h_1, \dots, h_m) \text{Jac}_x(\mathcal{B}) = \mathbf{0}_{n_x}.$$

By treating the coefficients of the h_i 's as unknowns, this yields a linear system with

$$m \binom{n_y + d - 3}{d - 2} \text{ unknowns} \quad \text{and} \quad n_x \binom{n_y + d - 2}{d - 1} \text{ equations.}$$

Assuming generic linear forms in $\text{Jac}_x(\mathcal{B})$, [BCV23] expect syzygies in bi-degree $(1, d - 1)$ whenever

$$m \binom{n_y + d - 3}{d - 2} > n_x \binom{n_y + d - 2}{d - 1}.$$

- The authors assume that the Macaulay matrix has full row rank when d is below this threshold. Sequences that satisfy this assumption are called $\mathbf{y}$ -semiregular.
- They analyze their XL strategy for $\mathbf{y}$ -semiregular sequences.

- The authors assume that the Macaulay matrix has full row rank when d is below this threshold. Sequences that satisfy this assumption are called $\mathbf{y}$ -semiregular.
- They analyze their XL strategy for $\mathbf{y}$ -semiregular sequences.

Caveats:

This amounts to assuming that the st^d coefficients of the series (1) are correct. So this is weaker than conjecturing the entire series, but we are still playing the same game.

We have no idea whether $\mathbf{y}$ -semiregularity is generic; [BCV23] only give experiments.

Partial conclusion

So far, we have looked at affine bilinear systems with no extra structure (and we already struggle to analyze the overdetermined case).

But often, there is some extra structure in cryptanalysis. Let's see two examples.

**A first example: Subfield Bilinear
Collision problem
([eprint/2026/916](#))**

Subfield Bilinear Collision problem

This problem was introduced by Huth and Joux to construct MPCitH signatures [HJ24].

Given linearly independent $\mathbf{u}, \mathbf{v} \in \mathbb{F}_{2^m}^n$, find nonzero $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$ such that

$$T(\mathbf{x}, \mathbf{y}) := \langle \mathbf{x}, \mathbf{u} \rangle \langle \mathbf{y}, \mathbf{v} \rangle - \langle \mathbf{x}, \mathbf{v} \rangle \langle \mathbf{y}, \mathbf{u} \rangle = 0. \quad (2)$$

[HJ24] Huth and Joux, “MPC in the Head Using the Subfield Bilinear Collision Problem”. *CRYPTO 2024*.

Subfield Bilinear Collision problem

This problem was introduced by Huth and Joux to construct MPCitH signatures [HJ24].

Given linearly independent $\mathbf{u}, \mathbf{v} \in \mathbb{F}_{2^m}^n$, find nonzero $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$ such that

$$T(\mathbf{x}, \mathbf{y}) := \langle \mathbf{x}, \mathbf{u} \rangle \langle \mathbf{y}, \mathbf{v} \rangle - \langle \mathbf{x}, \mathbf{v} \rangle \langle \mathbf{y}, \mathbf{u} \rangle = 0. \quad (2)$$

Initial algebraic attack

Rewrite (2) as a bilinear system $\mathcal{F} \in \mathbb{F}_2[\mathbf{x}, \mathbf{y}]_{(1,1)}^m$ using a basis of $\mathbb{F}_{2^m}/\mathbb{F}_2$, and solve this system using Gröbner bases.

This attack is not the main threat here, but we wanted to improve its analysis further.

[HJ24] Huth and Joux, “MPC in the Head Using the Subfield Bilinear Collision Problem”. *CRYPTO 2024*.

Rank metric perspective

We have

$$\begin{aligned} T(\mathbf{x}, \mathbf{y}) &= \langle \mathbf{x}, \mathbf{u} \rangle \langle \mathbf{y}, \mathbf{v} \rangle - \langle \mathbf{x}, \mathbf{v} \rangle \langle \mathbf{y}, \mathbf{u} \rangle \\ &= \det \left(\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}^T \right) := \det(\mathbf{Z} \mathbf{H}^T). \end{aligned}$$

Thus, $T(\mathbf{x}, \mathbf{y}) = 0$ if and only if there exists a nonzero vector $\mathbf{mZ} \in \mathbb{F}_{q^m}^n$ in $\text{Ker}(\mathbf{H}^T)$.

By definition of the rank metric, this vector has rank weight ≤ 2 .

We have

$$\begin{aligned} T(\mathbf{x}, \mathbf{y}) &= \langle \mathbf{x}, \mathbf{u} \rangle \langle \mathbf{y}, \mathbf{v} \rangle - \langle \mathbf{x}, \mathbf{v} \rangle \langle \mathbf{y}, \mathbf{u} \rangle \\ &= \det \left(\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}^T \right) := \det(\mathbf{Z}\mathbf{H}^T). \end{aligned}$$

Thus, $T(\mathbf{x}, \mathbf{y}) = 0$ if and only if there exists a nonzero vector $\mathbf{mZ} \in \mathbb{F}_{q^m}^n$ in $\text{Ker}(\mathbf{H}^T)$.

By definition of the rank metric, this vector has rank weight ≤ 2 .

So, we are just looking for a codeword of weight ≤ 2 in a code of codimension 2 in $\mathbb{F}_{q^m}^n$. We can apply the usual attacks from rank-metric cryptanalysis.

Another algebraic attack

By the Cauchy–Binet formula, $\det(\mathbf{ZH}^T)$ is a linear combination of the 2×2 minors z_T of $\mathbf{Z}$. Descending this linear polynomial over $\mathbb{F}_2$ yields $\mathcal{L} \in \mathbb{F}_2[z_T]^m$.

Unfortunately, $\mathcal{L}$ is underdetermined since $m < \binom{n}{2}$. In rank metric, one uses various hybrid techniques to deal with this.

Another algebraic attack

By the Cauchy–Binet formula, $\det(\mathbf{Z}\mathbf{H}^T)$ is a linear combination of the 2×2 minors z_T of $\mathbf{Z}$. Descending this linear polynomial over $\mathbb{F}_2$ yields $\mathcal{L} \in \mathbb{F}_2[z_T]^m$.

Unfortunately, $\mathcal{L}$ is underdetermined since $m < \binom{n}{2}$. In rank metric, one uses various hybrid techniques to deal with this.

Adding quadratic Plücker relations

Another possibility is to add a quadratic system $\mathcal{Q}$ given by the Plücker relations between the z_T 's, and solve $\mathcal{Q} \cup \mathcal{L}$ using Gröbner bases in $\mathbb{F}_2[z_T]$.

We know how to analyze this second algebraic attack, even though it is not very efficient. But this helps us analyze the initial algebraic attack.

Coming back to the initial algebraic attack.

We saw that the initial system $\mathcal{F}$ was linear in the maximal minors $(x_i y_j - x_j y_i)$ for $1 \leq i < j \leq n$ in addition to being bilinear. We have to take this into account.

Coming back to the initial algebraic attack.

We saw that the initial system $\mathcal{F}$ was linear in the maximal minors $(x_i y_j - x_j y_i)$ for $1 \leq i < j \leq n$ in addition to being bilinear. We have to take this into account.

Let $I \subset \mathbb{F}_2[\mathbf{x}, \mathbf{y}]$ be the ideal generated by these maximal minors.

- We cannot use the series (1) if the predicted rank is $> \dim I_{(d_x, d_y)}$, rather than $\dim \mathbb{F}_2[\mathbf{x}, \mathbf{y}]_{(d_x, d_y)}$.
- Otherwise, we conjecture that we can still use (1).
- This is valid as long as we do not reach the degree of regularity d of $\mathcal{Q} \cup \mathcal{L}$.
Beyond this point, additional syzygies appear in bi-degree (d, d) in the system $\mathcal{F}$.

**A second related example: McEliece
cryptanalysis
([eprint/2026/1232](#))**

Reduction to MinRank

Some recent key-recovery attacks on McEliece reduce to a particular MinRank instance [CMT23b]; [Mor25]; [BLT26]. More precisely, in even characteristic, we are given a public space

$$\langle \mathbf{M}_1, \dots, \mathbf{M}_K \rangle_{\mathbb{F}_q}$$

of skew-symmetric matrices, and look for a matrix $\mathbf{M}$ of rank 2 in this space.

Why we care

The public space can be constructed from either a McEliece public key or a random code. In a suitable regime, the MinRank instance has a solution in the public-key case but not in the random-code case. → Solving MinRank distinguishes the two cases and, in the public-key case, the solution can be used to decode.

We look for a skew-symmetric matrix $\mathbf{M}$ of size N of rank 2 in characteristic 2 inside a public space $\langle \mathbf{M}_1, \dots, \mathbf{M}_K \rangle$.

- For a skew-symmetric matrix, rank ≤ 2 is equivalent to the vanishing of all 4×4 Pfaffians. For every $1 \leq i < j < k < \ell \leq N$, we have

$$\text{Pf}_{i,j,k,\ell}(\mathbf{M}) = m_{ij}m_{k\ell} + m_{ik}m_{j\ell} + m_{i\ell}m_{jk} = 0.$$

This gives a quadratic system $\mathcal{Q}$ of $\binom{N}{4}$ equations in the $m_{i,j}$'s.

- On the other hand, the condition $\mathbf{M} \in \langle \mathbf{M}_1, \dots, \mathbf{M}_K \rangle$ can be expressed by linear equations $\mathcal{L}$ in the same unknowns. These correspond to the parity-check equations of the matrix code.

Pfaffian approach

Solve the combined system $\mathcal{Q} \cup \mathcal{L}$ using Gröbner bases.

We can replace the Pfaffian equations above by the parametrization

$$M = \mathbf{x}^T \mathbf{y} + \mathbf{y}^T \mathbf{x}$$

of a rank-2 skew-symmetric matrix. Expressing the parity-check equations in terms of these new variables gives a bilinear system $\mathcal{F}$ in $\mathbb{F}_q[\mathbf{x}, \mathbf{y}]$.

Our approach

Solve $\mathcal{F}$ using Gröbner bases.

The apparent advantage is that the number of variables drops from $\binom{N}{2}$ to $2N$, but of course we have to analyze the complexity of the resulting solving algorithm.

Analyzing the solving algorithm

The analysis is similar to the one we had for Subfield Bilinear Collision.

- To estimate the rank of the Macaulay matrices, we use:
 - the fact that the equations of $\mathcal{F}$ are linear combinations of the maximal minors

$$x_i y_j + y_i x_j, \quad 1 \leq i < j \leq N,$$

- the conjectured generating series (1).

Analyzing the solving algorithm

The analysis is similar to the one we had for Subfield Bilinear Collision.

- To estimate the rank of the Macaulay matrices, we use:
 - the fact that the equations of $\mathcal{F}$ are linear combinations of the maximal minors

$$x_i y_j + y_i x_j, \quad 1 \leq i < j \leq N,$$

- the conjectured generating series (1).
- To Reviewer A: this analysis is valid whenever we do not reach the degree of regularity d of $\mathcal{Q} \cup \mathcal{L}$. In this case, there is no gain in considering the bi-degree (d, d) compared to [CMT23b]; [BLT26].

Thank you !
