Cryptanalysis of the Subfield Bilinear Collision Problem

Pierre Briaud[1,2] and Romaric Neveu[1]

[1] XLIM, UMR 7252, France
[2] CNRS, France

Abstract. The security of a recent MPC-in-the-Head signature scheme introduced at Crypto 2024 by Huth and Joux required the introduction of a new ad hoc hardness assumption called the Subfield Bilinear Collision (SBC) problem. By relying on the VOLE-in-the-Head framework, the authors further improved the performance of this scheme at Asiacrypt 2025, resulting in a very compact construction.
In this paper, we improve the original cryptanalysis of SBC in several ways. First, we describe a link between the SBC problem and the decoding problem in the rank metric for codes linear over an extension field $\mathbb{F}_{q^m}$, strengthening its theoretical hardness and expanding the range of attacks on the SBC problem. Second, we analyze Gröbner basis algorithms applied to the bilinear modeling of SBC proposed by Huth and Joux and formulate conjectures on the behavior of this system. Finally, we describe another algebraic modeling of SBC obtained by using the Plücker relations between the maximal minors of a matrix. While we do not threaten the parameters of the proposed signature schemes relying on SBC, our work opens the door to a more accurate analysis. In particular, we were not able to analyze the inclusion of the field equations of $\mathbb{F}_q$ in the Gröbner basis algorithm, which is especially relevant since $q = 2$ in these parameters. We leave this task for future work.

1 Introduction

The use of the MPC-in-the-Head (MPCitH) paradigm [21] to build signature schemes has allowed the introduction of a new class of very compact proposals whose security can be reduced to the study of one unique underlying hardness assumption. Both this compactness and the absence of a trapdoor make these schemes quite competitive, as can be seen in the call for additional signature proposals launched by the NIST [27], in which five MPCitH constructions qualified in the second round (Mirath, MQOM, PERK, RYDE, and SDitH). While the underlying hardness assumptions for these five schemes had already been studied previously to the advent of such MPCitH constructions, this paradigm also enables the use of more ad hoc assumptions to build efficient schemes, as long as a representation of these assumptions as a polynomial system is known (as the MPCitH framework handles the efficiency part).

At Crypto 2024 [19], Huth and Joux proposed a new MPCitH signature scheme whose security is based on the so-called *Subfield Bilinear Collision (SBC)*

problem. This scheme was later improved at Asiacrypt 2025 [20] by relying on the VOLE-in-the-Head paradigm [6]. As depicted in Table 4 of [20], this second scheme compares very favorably with the VOLE-in-the-Head versions of the aforementioned NIST submissions. A clear drawback is that SBC remains ad hoc compared to the problems underlying these NIST submissions. The SBC problem can be defined as follows.

Problem 1 (Subfield Bilinear Collision (SBC) problem) *Let $n,\ m \in \mathbb{N}$, let $\mathbb{F}_q$ be a finite field and let $\mathbb{F}_{q^m}$ be a degree-m extension of $\mathbb{F}_q$. Given two linearly independent vectors $\boldsymbol{u},\ \boldsymbol{v} \in \mathbb{F}_{q^m}^n$, the Subfield Bilinear Collision (SBC) problem asks to find two non-collinear vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{F}_q^n$ such that*

$$T(\boldsymbol{x}, \boldsymbol{y}) \stackrel{def}{=} \langle \boldsymbol{u}, \boldsymbol{x} \rangle \langle \boldsymbol{v}, \boldsymbol{y} \rangle - \langle \boldsymbol{u}, \boldsymbol{y} \rangle \langle \boldsymbol{v}, \boldsymbol{x} \rangle = 0, \tag{1}$$

where $\langle ., . \rangle$ denotes the inner product between vectors in $\mathbb{F}_{q^m}^n$.

Remark 1 *The original definition in [19] uses k whereas we use m. We make this choice to align with the standard notation in rank-metric code-based cryptography, where the degree of the extension field is typically denoted by m.*

The definition of Problem 1 was inspired by a particular class of Index Calculus algorithms [22,23] to solve the discrete logarithm problem in the multiplicative group $\mathbb{F}_{q^m}^*$. Indeed, one step of these algorithms corresponds to solving a specific SBC instance (see [19, Page 6]). However, there is no known reduction to the discrete logarithm problem, and in particular, SBC is a candidate to be quantum-resistant.

The SBC instances used in [19,20] have a planted solution of the form $\boldsymbol{x} = (\boldsymbol{x}', 1, 0)$ and $\boldsymbol{y} = (\boldsymbol{y}', 0, 1)$ for $\boldsymbol{x}' \in \mathbb{F}_q^{n-2}$ and $\boldsymbol{y}' \in \mathbb{F}_q^{n-2}$. Such instances are referred to as *Normalised SBC (NSBC) instances*. The authors consider only $q = 2$ and focus on the parameter regime $n \approx m/2$, as it seems to be the hardest regime for SBC. In particular, the NSBC instance they propose for 128-bit security is such that $q = 2,\ m = 257$, and $n = 130$. The most efficient attack on these parameters is the generic collision technique described in [19, Choice of parameters, Page 2], whose complexity $\mathcal{O}(q^{m+1-n})$ indeed simplifies to 2^{128} for these parameters if the hidden constant in the $\mathcal{O}(.)$ notation is neglected. The only other attack analyzed in [19] is an algebraic attack based on a bilinear modeling $\mathcal{B}$ of SBC, in which the two blocks of variables $\boldsymbol{x}$ and $\boldsymbol{y}$ represent an SBC solution. Such a modeling is derived by applying a standard descent process from $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$ to Equation (1), which yields m bilinear equations in $\mathbb{F}_q[\boldsymbol{x}, \boldsymbol{y}]$. To attack specifically NSBC, one can further fix $x_{n-1} = 1,\ x_n = 0,\ y_{n-1} = 0$ and $y_n = 1$ in $\mathcal{B}$ and solve the resulting system in $\mathbb{F}_q[\boldsymbol{x}', \boldsymbol{y}']$ that we will denote by $\mathcal{B}_*$. Among known solving algorithms for such a $\mathcal{B}_*$, the folklore approach that consists in guessing the value of $\boldsymbol{x}'$ or $\boldsymbol{y}'$ and checking the guess by solving a linear system in the other block seems much more efficient than Gröbner basis algorithms. Indeed, for $q = 2$, the complexity of the former is 2^{n-2} times a polynomial factor corresponding to the cost of checking each guess, while the

complexity of the latter was estimated in [19, Page 7] as

$$\mathcal{O}\left(\binom{n_{x'}-1+n_{y'}-1+\min\left(n_{x'},n_{y'}\right)}{\min\left(n_{x'},n_{y'}\right)}^{\omega}\right)=\mathcal{O}\left(\binom{2(n-3)+(n-2)}{n-2}^{\omega}\right)$$
$$=\mathcal{O}\left(2^{2.75n\omega}\right), \tag{2}$$

where $2 \leq \omega \leq 3$ is the linear algebra constant.

Contributions. In this work, we strengthen the original cryptanalysis of SBC in several directions. While we cannot break the parameters of [19,20], our results suggest that a break is either unlikely or would require including the field equations of $\mathbb{F}_2$ in algebraic approaches, and analyzing this inclusion. This analysis is not trivial, as we discuss below.

First, we show that SBC can equivalently be viewed as the problem of finding a word of weight $\leq r = 2$ with respect to the rank metric in an $\mathbb{F}_{q^m}$-linear code of length n and dimension $k = n-2$. For arbitrary values of (m, n, k, r), the latter problem is closely related to the Rank Decoding (RD) problem, which is the main hardness assumption used in rank-metric code-based cryptography. While RD is typically not considered for such values of the weight and dimension, we observe that the instance used in [19,20] lies at the Rank Gilbert–Varshamov bound, which corresponds to the hardest regime for this problem. This connection to the rank metric implies that any attack from this area can be applied to the SBC problem. In particular, we give another attack with complexity 2^{128} times a polynomial factor that relies on the recent Hybrid Subsupport Guessing technique from PQCrypto 2026 [7].

Second, we highlight a specific structure present in $\mathcal{B}$ that has to be taken into account in the analysis of algebraic approaches. Namely, we observe that the equations of $\mathcal{B}$ are linear combinations of the maximal minors of the $2 \times n$ matrix

$$\boldsymbol{Z} \stackrel{def}{=} \begin{pmatrix} x_1 & \dots & x_n \\ y_1 & \dots & y_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix}. \tag{3}$$

This allows us to use results on determinantal ideals to analyze algebraic techniques applied to $\mathcal{B}_*$. It is important to mention here that the algebraic cryptanalysis of generic *overdetermined* bilinear systems is not well understood in comparison to that of *underdetermined* versions. For this reason, despite the progress made in this paper, we believe that a complete analysis of algebraic techniques to solve our system $\mathcal{B}_*$, which is both overdetermined and structured, remains out of reach. Still, we believe, as in [19], that the folklore approach described above or a very aggressive hybrid Gröbner basis approach will be more efficient than a plain Gröbner basis algorithm to solve $\mathcal{B}_*$, since we can afford to fix a large number of variables over $\mathbb{F}_2$ without increasing the cost too much.

Field equations of $\mathbb{F}_2$. A clear direction to improve our work would be to take the field equations of $\mathbb{F}_2$ into account in the Gröbner basis computation, which is done neither here or in [19,20] (note indeed that Equation (2) assumes that

no field equations are added). However, deriving a complexity analysis seems challenging. First, there is no published analysis of the solving of generic Boolean bilinear systems, whereas the case of generic Boolean quadratic systems is well understood [3,5,4]. In our case, we would also specifically need to understand the behavior of the ideal generated by the maximal minors of the matrix $\boldsymbol{Z}$ together with the field equations of $\mathbb{F}_2$. Unfortunately, our experiments have not suggested any meaningful conjectures regarding this behavior.

2 Preliminaries

Vectors are denoted by lower case boldface letters and matrices by uppercase letters. The j-th coordinate of a vector $\boldsymbol{x}$ is denoted by x_j and the submatrix of a matrix $\boldsymbol{M}$ on the rows in I and columns in J is denoted by $\boldsymbol{M}_{I,J}$. When I (resp. J) represents all the rows (resp. columns), we use the notation $\boldsymbol{M}_{*,J}$ (resp. $\boldsymbol{M}_{I,*}$). Finally, $|\boldsymbol{M}|$ is the determinant of a matrix $\boldsymbol{M}$, $|\boldsymbol{M}|_{I,J}$ is the determinant of the submatrix $\boldsymbol{M}_{I,J}$ and $|\boldsymbol{M}|_{*,J}$ the one of $\boldsymbol{M}_{*,J}$ (as long as all these matrices are square). In the following, let $\mathbb{F}_q$ be a finite field and let $\mathbb{F}_{q^m}$ be an extension of degree m.

2.1 Rank metric over $\mathbb{F}_{q^m}^n$

We present the rank metric specifically on vectors in $\mathbb{F}_{q^m}^n$.

Definition 1 (Support, rank weight). *The* support *of a vector $\boldsymbol{c} \in \mathbb{F}_{q^m}^n$ for the rank metric is the $\mathbb{F}_q$-subspace of $\mathbb{F}_{q^m}$ defined by*

$$Supp(\boldsymbol{c}) \stackrel{def}{=} \langle c_1, \ldots, c_n \rangle_{\mathbb{F}_q}.$$

The rank weight of $\boldsymbol{c}$ is $|\boldsymbol{c}| \stackrel{def}{=} \dim_{\mathbb{F}_q}(Supp(\boldsymbol{c}))$.

The main hard problem in code-based cryptography relying on the rank metric is the generic decoding problem for $\mathbb{F}_{q^m}$-linear codes in $\mathbb{F}_{q^m}^n$ equipped with this metric. This problem is referred to as the *Rank Decoding (RD) problem*. A closely related problem is the problem of finding low-weight codewords.

Problem 2 (Low Rank Codeword (LRC) problem) *Let q, m, n, ℓ, $r \in \mathbb{N}$, let $\mathbb{F}_q$ be a finite field and let $\mathbb{F}_{q^m}$ be a degree-m extension of $\mathbb{F}_q$. Given an $\mathbb{F}_{q^m}$-linear code $\mathcal{C}$ of length n and dimension ℓ, the Low Rank Codeword (LRC) problem asks to find a non-zero codeword $\boldsymbol{c} \in \mathcal{C}$ such that $|\boldsymbol{c}| \leq r$.*

While LRC is not known to be equivalent to RD, the former problem is still widely believed to be hard. For instance, given an RD problem consisting of an $\mathbb{F}_{q^m}$-linear code $\mathcal{D}$, a noisy codeword $\boldsymbol{y} \in \mathbb{F}_{q^m}^n$ and a target weight r, the current best RD algorithms are LRC algorithms in the augmented code $\mathcal{C} = \mathcal{D} \oplus \langle \boldsymbol{y} \rangle_{\mathbb{F}_{q^m}}$ with the same target weight. In the reverse direction, there is a quantum reduction from finding low-weight codewords to decoding in several

metrics [16], but it is unfortunately trivial in the rank-metric setting, as the RD instance resulting from this reduction is easy to solve.

The MaxMinors approach was introduced in [2] to solve the RD problem. It is, in fact, an LRC algorithm in the augmented code $\mathcal{C} = \mathcal{D} \oplus \langle \boldsymbol{y} \rangle_{\mathbb{F}_{q^m}}$. We describe this algorithm here in the case of an arbitrary LRC instance. A vector $\boldsymbol{c} \in \mathbb{F}_{q^m}^n$ of weight $\leq r$ can be expressed as

$$\boldsymbol{c} = (s_1, \ldots, s_r)\boldsymbol{C},$$

where $s_1, \ldots, s_r \in \mathbb{F}_{q^m}$ generate $\mathrm{Supp}(\boldsymbol{c})$ and where $\boldsymbol{C} \in \mathbb{F}_q^{r \times n}$ is a full-rank matrix. Then, $\boldsymbol{c} \in \mathcal{C}$ if and only if

$$(s_1, \ldots, s_r)\boldsymbol{C}\boldsymbol{H}^\mathsf{T} = \boldsymbol{0},$$

where $\boldsymbol{H} \in \mathbb{F}_{q^m}^{(n-\ell) \times n}$ is a full-rank parity-check matrix of $\mathcal{C}$. This implies that the matrix $\boldsymbol{C}\boldsymbol{H}^\mathsf{T}$ is not full-rank. The so-called MaxMinors system is given by the vanishing of all maximal minors of this matrix, namely

$$\mathrm{Minors}_r(\boldsymbol{C}\boldsymbol{H}^\mathsf{T}) = \boldsymbol{0}.$$

As in Equation (1), these $\binom{n-\ell}{r}$ equations have coefficients in $\mathbb{F}_{q^m}$ but variables in $\mathbb{F}_q$. In particular, by relying on the same descent process as the one used to derive the system $\mathcal{B}$ previously mentioned, one can obtain $m\binom{n-\ell}{r}$ equations with both coefficients and variables in $\mathbb{F}_q$. A last ingredient is the Cauchy–Binet formula, which yields a more compact expression for these equations.

Lemma 1 (Cauchy–Binet formula). *Let $\boldsymbol{A}$ be an $s \times t$ matrix and let $\boldsymbol{B}$ be an $t \times s$ matrix. Then*

$$|\boldsymbol{A}\boldsymbol{B}| = \sum_{T \subset \{1..t\},\ \#T = s} |\boldsymbol{A}|_{*,T} |\boldsymbol{B}|_{T,*}.$$

More precisely, for any subset $J \subset \{1..(n-\ell)\}$ of size r, this formula applied to $\boldsymbol{A} = \boldsymbol{C}$ and $\boldsymbol{B} = \boldsymbol{H}^\mathsf{T}_{*,J}$ shows that

$$|\boldsymbol{C}\boldsymbol{H}^\mathsf{T}|_{*,J} = \sum_{T \subset \{1..r\},\ \#T = r} |\boldsymbol{C}|_{*,T} |\boldsymbol{H}^\mathsf{T}|_{T,J} = \sum_{T \subset \{1..r\},\ \#T = r} |\boldsymbol{C}|_{*,T} |\boldsymbol{H}|_{T,J}.$$

Using this expression, the MaxMinors system can be interpreted as a linear system in the $\binom{n}{r}$ unknowns given by the maximal minors $|\boldsymbol{C}|_{*,T}$ of $\boldsymbol{C}$, which we call the *minor variables*. All attacks on RD adopt this interpretation of MaxMinors as a linear system.

2.2 Hilbert function and Hilbert series

Let $R = \mathbb{F}_q[\boldsymbol{x}]$ be a multivariate polynomial ring. A polynomial $f \in R$ is homogeneous of degree $d \in \mathbb{N}$ if all its monomials have the same degree d. We denote by R_d the corresponding vector space. An ideal $I \subset R$ is homogeneous

if it admits a set of homogeneous generators. In this case, the set $I \cap R_d$ is a subspace of R_d, denoted by I_d. Provided they are known for the associated ideal, the Hilbert function and the Hilbert series can be used to study the complexity of solving an algebraic system.

Definition 2 (Hilbert function). *Let $I \subset R$ be a homogeneous ideal. The Hilbert function of the quotient ring R/I is the function defined by*

$$\begin{aligned} HF_{R/I} : \mathbb{N} &\longrightarrow \mathbb{N} \\ d &\longmapsto \dim_{\mathbb{F}_q}(R_d/I_d). \end{aligned}$$

Definition 3 (Hilbert series). *Let $I \subset R$ be a homogeneous ideal. The Hilbert series of the quotient ring R/I is the power series defined by*

$$HS_{R/I}(t) \stackrel{def}{=} \sum_{d=0}^{\infty} HF_{R/I}(d)t^d.$$

An ideal I is said to be zero-dimensional if the quotient ring R/I is a finite-dimensional vector space. If, in addition, I is homogeneous, the degree of regularity of I is defined as the smallest integer d such that $I_d = R_d$.

Similar definitions can be given for a bi-graded polynomial ring $S = \mathbb{F}_q[\boldsymbol{x}, \boldsymbol{y}]$ with n_x variables $\boldsymbol{x}$ and n_y variables $\boldsymbol{y}$. The *bi-degree* of a polynomial $f \in S = \mathbb{F}_q[\boldsymbol{x}, \boldsymbol{y}]$ is defined as

$$\boldsymbol{d} = (d_x, d_y) \stackrel{def}{=} \big(\deg_{\boldsymbol{x}}(f), \deg_{\boldsymbol{y}}(f)\big),$$

where $\deg_{\boldsymbol{x}}(f)$ (resp. $\deg_{\boldsymbol{y}}(f)$) denotes the partial degree of f with respect to the variables $\boldsymbol{x}$ (resp. $\boldsymbol{y}$). Note that a polynomial of bi-degree $\boldsymbol{d}$ does not necessarily contain monomials of bi-degree $\boldsymbol{d}$. In particular, an element $f \in S$ will be called *bihomogeneous of bi-degree $\boldsymbol{d}$* if all its monomials have bi-degree $\boldsymbol{d}$. We denote by $S_{\boldsymbol{d}}$ the corresponding vector space. An ideal $J \subset S$ is bihomogeneous if it admits a set of bihomogeneous generators. The following Definitions 4 and 5 are the bihomogeneous versions of Definitions 2 and 3, respectively.

Definition 4 (Hilbert bi-function). *Let $J \subset S$ be a bihomogeneous ideal. The Hilbert bi-function of the quotient ring S/I is the function defined by*

$$\begin{aligned} HF_{S/J} : \mathbb{N}^2 &\longrightarrow \mathbb{N} \\ \boldsymbol{d} &\longmapsto \dim_{\mathbb{F}_q}(S_{\boldsymbol{d}}/J_{\boldsymbol{d}}). \end{aligned}$$

Definition 5 (Hilbert bi-series). *Let $J \subset S$ be a bihomogeneous ideal. The Hilbert bi-series of the quotient ring S/J is defined by*

$$H_{S/J}(s,t) \stackrel{def}{=} \sum_{\boldsymbol{d}=(d_x,d_y)\in\mathbb{N}^2} HF_{S/J}(\boldsymbol{d})s^{d_x}t^{d_y}.$$

In practice, we often need to solve equations in S that are not bihomogeneous. To handle these, let us order bi-degrees componentwise, namely $\boldsymbol{d}' \leq \boldsymbol{d}$ if and only if $d'_x \leq d_x$ and $d'_y \leq d_y$. We define

$$S_{\leq \boldsymbol{d}} \stackrel{def}{=} \oplus_{\boldsymbol{d}' \leq \boldsymbol{d}} S_{\boldsymbol{d}'}.$$

An element $f \in S_{\leq (1,1)}$ will be called *affine bilinear*.

As we stated in the introduction, the complexity of solving overdetermined bilinear systems is less well understood than that of solving general quadratic systems. Without going into too much detail, this stems from the fact that the Hilbert series of the quotient ring R/I is known for ideals I generated by a generic system, both in the underdetermined and overdetermined cases, whereas in the bilinear setting, the Hilbert bi-series of S/J is known only in the underdetermined case $m \leq n_x + n_y$, still under the assumption that J is generic [17, Theorem 5.1]. In the next subsection, we explain how overdetermined bilinear systems from the cryptanalytic literature have been solved and analyzed, even though no such Hilbert bi-series was available.

2.3 XL techniques for overdetermined bilinear systems

Several works have considered adapting the XL algorithm [13] to overdetermined bilinear systems, namely [1,26] for generic overdetermined bilinear systems, and [29,28,10] for specific overdetermined bilinear systems relevant to cryptanalysis. These works rely on Macaulay matrices indexed by bi-degrees, as introduced below. These matrices were also used previously in [17].

Definition 6 (Homogeneous Macaulay matrix). *Let $\mathcal{F} \in S^m$ be a bihomogeneous system and let $\boldsymbol{d} \in \mathbb{N}^2$. The homogeneous Macaulay matrix of $\mathcal{F}$ in bi-degree $\boldsymbol{d}$, denoted by $\boldsymbol{M}_{\boldsymbol{d}}(\mathcal{F})$, is the matrix whose columns are indexed by all monomials $\nu \in S_{\boldsymbol{d}}$, rows are indexed by pairs (μ, f) for all monomials $\mu \in S$ and $f \in \mathcal{F}$ such that $\mu f \in S_{\boldsymbol{d}}$, and coefficients are given by coef$(\nu, \mu f)$.*

Definition 7 (Affine Macaulay matrix). *Let $\mathcal{F} \in S^m$ be a polynomial system and let $\boldsymbol{d} \in \mathbb{N}^2$. The affine Macaulay matrix of $\mathcal{F}$ in bi-degree $\boldsymbol{d}$, denoted by $\boldsymbol{M}_{\leq \boldsymbol{d}}(\mathcal{F})$, is the matrix whose columns are indexed by all monomials $\nu \in S_{\leq \boldsymbol{d}}$, rows are indexed by pairs (μ, f) for all monomials $\mu \in S$ and $f \in \mathcal{F}$ such that $\mu f \in S_{\leq \boldsymbol{d}}$, and coefficients are given by coef$(\nu, \mu f)$.*

Let $\mathcal{F}$ be an affine bilinear system that is overdetermined. For a given bi-degree $\boldsymbol{d}$, the XL algorithm aims at finding a solution to $\mathcal{F}$ or showing that no solution exists by solving the linear system in the unknowns $\boldsymbol{z}$,

$$\boldsymbol{M}_{\leq \boldsymbol{d}}(\mathcal{F}) \boldsymbol{z}^{\mathsf{T}} = \boldsymbol{0}.$$

This will succeed as long as the rank of the matrix $\boldsymbol{M}_{\leq \boldsymbol{d}}(\mathcal{F})$ is equal to its number of columns when $\mathcal{F}$ has no solution or this number of columns minus a small positive integer constant κ when a solution exists, typically $\kappa = 1$.

The first bilinear adaptation of XL that we are aware of is $\boldsymbol{y}$-XL [29,1], which focuses on Macaulay matrices in bi-degrees $(1, d_y)$ for $d_y \geq 1$. Macaulay matrices in arbitrary bi-degrees were considered in [28,9,10,26]. Finally, a generalization using Macaulay matrices indexed by sets of bi-degrees instead of single bi-degrees was proposed recently in [26]. Mutant XL (MXL) [12] is a standard technique to solve the system at a lower degree than XL by exploiting *degree-fall polynomials*. The rough idea is that the Macaulay matrix $\boldsymbol{M}_{\leq \boldsymbol{d}}(\mathcal{F})$ may be too small to solve the system at this bi-degree, but an interesting point is that its rowspace may contain "new" polynomials of bi-degree smaller than $\boldsymbol{d}$ that we would like to use in the computation. Adapting MXL to the bilinear case was investigated for bi-degrees $(1, d_y)$ in [1], where a $\boldsymbol{y}$-MXL variant of $\boldsymbol{y}$-XL is given, while the extension to arbitrary bi-degrees was suggested in [26].

Since we do not have access to a Hilbert bi-series, it is not straightforward to estimate the ranks of the Macaulay matrices associated to an overdetermined bilinear system $\mathcal{F}$. Still, it is clear that knowing these ranks is important for studying the complexity of XL techniques. Several works [28,9,10] have *conjectured* these ranks via the following series

$$\frac{(1-st)^m}{(1-s)^{n_x}(1-t)^{n_y}}. \tag{4}$$

More precisely, the co-rank of $\boldsymbol{M}_{\boldsymbol{d}}(\mathcal{F})$ was estimated to be equal to the coefficient of $s^{d_x}t^{d_y}$ in this series.

Finally, as in the standard case, there are two relevant bihomogeneous systems associated with an affine bilinear system $\mathcal{F}$, which are useful for estimating the complexity of solving $\mathcal{F}$ via XL techniques. The first one, which is relevant for studying XL, corresponds to the following homogenization of $\mathcal{F}$.

Definition 8 (Bihomogenization). *The bihomogenization of a polynomial $f \in S_{\leq(1,1)}$ is the polynomial $f^{(h)} \in \mathbb{F}_q[\boldsymbol{x}, \boldsymbol{y}, x_0, y_0]_{(1,1)}$ defined by*

$$f^{(h)} \stackrel{def}{=} f^{(1,1)} + y_0 f^{(1,0)} + x_0 f^{(0,1)} + x_0 y_0 f^{(0,0)},$$

where $f^{\boldsymbol{d}}$ denotes the bihomogeneous component of f of bi-degree $\boldsymbol{d} \leq (1,1)$. The bihomogenization of a system $\mathcal{F} = \{f_1, \ldots, f_m\} \in S^m_{\leq(1,1)}$ is the system $\mathcal{F}^{(h)} \in \mathbb{F}_q[\boldsymbol{x}, \boldsymbol{y}, x_0, y_0]^m_{(1,1)}$ defined by

$$\mathcal{F}^{(h)} \stackrel{def}{=} \{f_1^{(h)}, \ldots, f_m^{(h)}\}.$$

The second one, which we may use to study MXL, is the bihomogeneous system $\mathcal{F}^{\text{top}}$ consisting of the highest bi-degree components of the polynomials in $\mathcal{F}$. This is because the study of degree-fall polynomials of $\mathcal{F}$ is typically reduced to determining syzygies of $\mathcal{F}^{\text{top}}$. However, the fact that $\mathcal{F}^{\text{top}}$ does not control the bihomogeneous components of smaller bi-degree might be problematic, in the sense that it does not allow us to predict the usefulness of the degree-fall polynomials. For instance, the lower bi-degree parts may cancel as well, leading to

no degree fall, and more generally, the resulting polynomials could be redundant in the rest of computation. In any case, the main issue here is that analyzing the complexity of MXL after including non-redundant degree-fall polynomials remains an open problem, *even in the generic overdetermined setting.*

3 SBC as a Rank Metric Problem

The starting point of the paper was to observe that an SBC instance readily corresponds to an instance of Problem 2 in which the code $\mathcal{C}$ has dimension $\ell = n - 2$ and the rank weight is $r = 2$.

Proposition 1 *The SBC problem with parameters q, m and n is equivalent to the LRC problem with parameters q, m, n, $\ell = n - 2$, and $r = 2$.*

Proof. ($\Rightarrow$) Let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{F}_{q^m}^n$ be two linearly independent vectors that define an SBC instance, and recall that this problem asks to find two linearly independent vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}_q^n$ such that

$$\langle \boldsymbol{x}, \boldsymbol{u} \rangle \langle \boldsymbol{y}, \boldsymbol{v} \rangle - \langle \boldsymbol{x}, \boldsymbol{v} \rangle \langle \boldsymbol{y}, \boldsymbol{u} \rangle = 0.$$

We remark that the left-hand side of this equality is the determinant of the matrix

$$\begin{pmatrix} \langle \boldsymbol{x}, \boldsymbol{u} \rangle & \langle \boldsymbol{x}, \boldsymbol{v} \rangle \\ \langle \boldsymbol{y}, \boldsymbol{u} \rangle & \langle \boldsymbol{y}, \boldsymbol{v} \rangle \end{pmatrix} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \times \begin{pmatrix} \boldsymbol{u}^\mathsf{T} & \boldsymbol{v}^\mathsf{T} \end{pmatrix} = \boldsymbol{Z}\boldsymbol{H}^\mathsf{T} \in \mathbb{F}_{q^m}^{2\times 2},$$

where $\boldsymbol{Z} \in \mathbb{F}_q[\boldsymbol{x}, \boldsymbol{y}]^{2\times n}$ is as in Equation (3) and

$$\boldsymbol{H} \stackrel{def}{=} \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} \in \mathbb{F}_{q^m}^{2\times n}.$$

Since $\boldsymbol{u}$ and $\boldsymbol{v}$ are linearly independent, the matrix $\boldsymbol{H}$ can be seen as a parity-check matrix for an $\mathbb{F}_{q^m}$-linear code $\mathcal{C}$ of length n and dimension $n-2$. We show now that a solution to the LRC problem in $\mathcal{C}$ with $r = 2$ provides a solution to the above SBC problem. A codeword of $\mathcal{C}$ of weight ≤ 2 is a vector $(s_1, s_2)\boldsymbol{C} \in \mathbb{F}_{q^m}^n$ such that $s_1, s_2 \in \mathbb{F}_{q^m}$, $\boldsymbol{C} \in \mathbb{F}_q^{2\times n}$ is full-rank and

$$(s_1, s_2)\boldsymbol{C}\boldsymbol{H}^\mathsf{T} = 0.$$

This implies that the determinant of $\boldsymbol{C}\boldsymbol{H}^\mathsf{T}$ is zero, and thus the rows of $\boldsymbol{C}$ provide an SBC solution. The converse ($\Leftarrow$) is analogous.

Remark 2 *The proof of Proposition 1 shows that the bilinear system $\mathcal{B}$ corresponds to the MaxMinors system used to solve the LRC instance associated to the SBC one. Before applying the descent process from $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$, we have $\binom{n-(n-2)}{2} = 1$ equation, namely*

$$\left| \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix}^\mathsf{T} \right| = \mathit{Minors}_2\left(\boldsymbol{Z}\boldsymbol{H}^\mathsf{T}\right),$$

where $\boldsymbol{Z} \in \mathbb{F}_q[\boldsymbol{x}, \boldsymbol{y}]^{2\times n}$ *is as in Equation* (3) *and* $\boldsymbol{H} = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} \in \mathbb{F}_{q^m}^{2\times n}$. *Expressing this equation in a basis of* $\mathbb{F}_{q^m}$ *over* $\mathbb{F}_q$ *yields the bilinear system* $\mathcal{B}$. *Finally, recall from Section 1 that the MaxMinors system is linear in the maximal minors* $z_T \stackrel{def}{=} |\boldsymbol{Z}|_{*,T}$ *of* $\boldsymbol{Z}$. *This allows us to interpret* $\mathcal{B}$ *as a linear system* $\mathcal{L} \in \mathbb{F}_q[z_T]^m$, *which we will use in Subsection 4.2.*

Before moving on to the cryptanalysis of SBC, let us mention that one can use arguments on the Rank Gilbert–Varshamov distance to recover the condition on m and n for the hardness of the SBC problem that was justified in [22, §2.2]. The Rank Gilbert–Varshamov distance for such a code $\mathcal{C}$ is defined as the smallest integer r_{RGV} such that

$$\sum_{j=0}^{r_{\mathrm{RGV}}} \prod_{\ell=0}^{j-1} (q^n - q^\ell) \binom{m}{j}_q \geq q^{m(n-(n-2))} = q^{2m},$$

where $\binom{m}{j}_q$ refers to the Gaussian binomial coefficient. When $n > m+1$, we have $r_{\mathrm{RGV}} = 1$ and thus $r = 2$ is an easy zone for LRC since the number of rank-2 codewords is exponential. When $n \approx m/2$ as chosen in [22,23], we can check that asymptotically $r_{\mathrm{RGV}} = 2$, so that $r = 2$ is the hardest weight in this case.

4 Revisiting the Cryptanalysis of SBC

In this section, we exploit the connection to LRC and we improve the analysis of the algebraic attack considered in [19] based on solving the system $\mathcal{B}_*$, that is, the system $\mathcal{B}$ that stems from Equation (1) to which we fix $x_{n-1} = y_n = 1$ and $x_n = y_{n-1} = 0$.

4.1 Known LRC algorithms

Due to Proposition 1, it is clear that any LRC algorithm yields an SBC algorithm. By studying the complexity of known LRC algorithms in the parameter regime corresponding to SBC, we found that it was crucial to use the so-called Hybrid Subsupport Guessing approach [7] to obtain a competitive complexity. This hybrid approach can be applied to any LRC algorithm to potentially reduce its initial complexity.

To solve an LRC instance with a unique solution, the Hybrid Subsupport Guessing approach consists in randomly sampling a matrix $\Pi \in \mathbb{F}_q^{n\times(n-h)}$ for some fixed integer $0 \leq h \leq r-1$ and making the bet that $\ker(\Pi) \subset \mathrm{Supp}(\boldsymbol{c})$ for the codeword $\boldsymbol{c}$ of weight $\leq r$ that is targeted. If the bet is correct, then $\boldsymbol{c}\Pi$ is of weight $\leq r-h$ in the code $\mathcal{C}\Pi$. Concretely, this bet is checked by solving the LRC problem with weight $r-h$ in the code $\mathcal{C}\Pi$, since no nonzero codewords of this weight should exist if the bet is incorrect. The code $\mathcal{C}\Pi$ is proven to behave as a random code if $\mathcal{C}$ is itself a random code, while the success probability

is estimated as $q^{-h(n-r)}$. Overall, the complexity of the Hybrid Subsupport Guessing technique when the parameter $0 \leq h \leq r-1$ is fixed can be estimated as

$$\mathcal{O}\left(q^{h(n-r)}\mathcal{T}_{\mathrm{LRC}}(q,m,n-h,\ell,r-h)\right),$$

where $\mathcal{T}_{\mathrm{LRC}}(q,m,n,\ell,r)$ is the time complexity of the LRC solver to which we apply this technique, and the actual complexity is obtained by taking the minimum over all such values of h.

If one applies this approach to SBC, i.e., to LRC with $r=2$ and $\ell=n-2$, the only nontrivial choice is $h=1$. This gives a hybrid factor of q^{n-2} in the complexity, whereas it is easy to see that solving the resulting LRC instance of weight $r-h=1$ takes polynomial time. On the SBC parameters $q=2,\ m=257,\ n=130$, the overall complexity is 2^{128} times a polynomial factor.

4.2 Approach based on Plücker relations

We propose another algebraic approach to solve SBC that could more generally apply to RD or LRC, but which has not yet been studied in detail in this context. This approach is based on the linear system $\mathcal{L} \in \mathbb{F}_q[z_T]^m$ from Remark 2, which we augment at a higher degree by multiplying its equations by monomials. Since $\mathcal{L}$ has fewer equations than variables in our parameter regime, this is meaningful only if we add other equations (and/or fix variables). In our case, we add the quadratic equations $\mathcal{Q}$ corresponding to all the Plücker relations between the z_T variables, and we solve $\mathcal{L} \cup \mathcal{Q}$ using Gröbner bases.

Definition 9 (Plücker relations). *Let $\boldsymbol{Z}$ be a formal $2 \times n$ matrix. The maximal minors of $\boldsymbol{Z}$, denoted by $z_{\{a,b\}} \stackrel{def}{=} |\boldsymbol{Z}|_{*,\{a,b\}}$ for $1 \leq a < b \leq n$, satisfy the quadratic relations*

$$z_{\{i,j\}}z_{\{k,l\}} - z_{\{i,k\}}z_{\{j,l\}} + z_{\{i,l\}}z_{\{j,k\}} = 0,$$

for any $1 \leq i < j < k < l \leq n$. We call $\mathcal{Q}$ the system in the variables z_T for $T \subset \{1..n\},\ \#T = 2$ that contains these $\binom{n}{4}$ quadratic relations.

Let $R = \mathbb{F}_q[z_T]$. To estimate the complexity of this approach, we try to determine the Hilbert function of the quotient ring $R/\langle \mathcal{L} \cup \mathcal{Q} \rangle$. We remark that the system $\mathcal{Q}$ readily corresponds to a generating set of the Pfaffian ideal of rank 2 of a skew-symmetric matrix $\boldsymbol{M} = (m_{i,j})$ of size n. Indeed, this ideal can be defined by

$$\mathcal{P}_2(\boldsymbol{M}) \stackrel{def}{=} \langle m_{i,j}m_{k,l} - m_{i,k}m_{j,l} + m_{i,l}m_{j,k},\ 1 \leq i < j < k < l \leq n \rangle.$$

Using the Hilbert function from [18, Theorem 4] that is recalled in [14, Proposition 14], we obtain

$$\mathrm{HF}_{R/\langle \mathcal{Q} \rangle}(d) = \binom{n+d-2}{d}^2 - \binom{n+d-2}{d-1}\binom{n+d-2}{d+1}. \tag{5}$$

Then, to determine the Hilbert function of $R/\langle \mathcal{L} \cup \mathcal{Q} \rangle$, the natural strategy would be to treat the linear equations of $\mathcal{L}$ one by one and assume that each of them behaves generically in the quotient by $\mathcal{Q}$ and the previous ones. This strategy amounts to applying [14, Conjecture 1], which was formulated for a similar modeling that contains $\mathcal{Q}$ as well as linear equations. Unfortunately, [14, Conjecture 1] holds under the assumption that this modeling has only zero for solution, whereas ours has a nonzero solution.

We focus instead on the analogue of the specialization $\mathcal{B}_*$ for the system $\mathcal{L} \cup \mathcal{Q}$, which still allows us to use Hilbert series arguments. We start from the specialized matrix

$$\boldsymbol{Z} = \begin{pmatrix} \boldsymbol{x}' & 1 & 0 \\ \boldsymbol{y}' & 0 & 1 \end{pmatrix},$$

and we add the equation $z_{\{n-1,n\}} = 1$ to $\mathcal{L} \cup \mathcal{Q}$, turning it into an affine system. To study degree falls in such a system, we consider the system obtained by taking the highest-degree homogeneous components of the equations, i.e., $\mathcal{L} \cup \mathcal{Q}$ with $z_{\{n-1,n\}} = 0$ instead of $z_{\{n-1,n\}} = 1$. Note that $\mathcal{L} \cup \mathcal{Q}$ typically admits no nonzero solution satisfying $z_{\{n-1,n\}} = 0$, as (i) the planted NSBC solution satisfies $z_{\{n-1,n\}} = 1$ and (ii) there are no other nonzero solutions in the parameter regime we consider (see the end of Section 3 or [19, Choice of parameters, Pages 2-3]). We are now in a setting to apply [14, Conjecture 1], which yields the following conjecture on the Hilbert function we are interested in. Our experiments support this conjecture, which we summarize in Appendix B.1.

Conjecture 1 *Let $\boldsymbol{u}$, $\boldsymbol{v} \in \mathbb{F}_{q^m}^n$ be an NSBC instance which admits a unique nontrivial solution, let $\mathcal{L} \in R^m$ be the set of MaxMinors polynomials that model this instance, and let $\mathcal{Q} \in R^{\binom{n}{4}}$ be the set of Plücker relations. For any $d \geq 0$, let*

$$HF_d \stackrel{def}{=} HF_{R/\langle \mathcal{L} \cup \mathcal{Q} \cup \{z_{\{n-1,n\}}\} \rangle}(d).$$

If $HF_{d'} > 0$ for any $d' < d$, then

$$HF_d = \max\left(0, \sum_{i=0}^{d} (-1)^i \binom{m+1}{i} HF_{\mathbb{F}_q[z_T]/\langle \mathcal{Q} \rangle}(d-i)\right),$$

where $HF_{R/\langle \mathcal{Q} \rangle}(\cdot)$ is defined in Equation (5). Otherwise, $HF_d = 0$.

Conjecture 1 predicts the smallest degree d at which degree falls occur in the affine system $\mathcal{L} \cup \mathcal{Q} \cup \{z_{\{n-1,n\}} - 1\}$ when applying a Gröbner basis algorithm with respect to a graded order. Moreover, it is reasonable to expect that solving the system will require at least as much work as finding these degree falls. However, for the SBC parameters, the corresponding value of d is already so large that computing the degree falls at this degree is too expensive. More precisely, for $m = 257$ and $n = 130$, the smallest integer d such that $\mathrm{HF}_d = 0$ is $d = 64$. This results in more than 500 bits of complexity for computing degree-fall polynomials at this degree. More generally, an important difference with $\mathcal{B}$ is that the number of variables is $\binom{n}{2}$ instead of $2n$, so we are much more limited in the degree we can feasibly consider with this approach.

Remark 3 *Similarly, we believe that the same approach for solving RD in the parameter regime (m, n, k, r) relevant to rank-based schemes is not competitive due to the large number of monomials already in degree 2. What is missing to derive a precise complexity analysis is the Hilbert function for the Plücker relations in an $r \times n$ matrix instead of a $2 \times n$ matrix. Indeed, we can no longer use the ideal $\mathcal{P}_2(\boldsymbol{M})$, and we do not have access to the Hilbert function $HF_{R/\langle \mathcal{Q} \rangle}(\cdot)$, which prevents us from adapting our analysis to the general case.*

4.3 Revisiting the complexity of solving $\mathcal{B}_*$

We return to the analysis of algebraic approaches for solving the affine system $\mathcal{B}_*$ introduced in the original paper.

Determinantal structure Recall from Remark 2 that the system $\mathcal{B}$ is linear in the maximal minors of the matrix

$$\boldsymbol{Z} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \in \mathbb{F}_q[\boldsymbol{x}, \boldsymbol{y}]^{2 \times n}.$$

In the previous subsection, we have exploited this fact by replacing these maximal minors $x_i y_j - x_j y_i$ for $1 \leq i < j \leq n$ by new variables $z_{\{i,j\}} = x_i y_j - x_j y_i$. In this subsection, we keep $\boldsymbol{x}$ and $\boldsymbol{y}$ as variables following the approach of [19,20], but we use this determinantal structure to better analyze the complexity of the solving algorithm. Our analysis actually applies to m generic equations that are linear combinations of the same quantities $x_i y_j - x_j y_i$ for $1 \leq i < j \leq n$.

Note that m such generic equations can always be viewed as the descent of one linear combination over $\mathbb{F}_{q^m}$ in the same maximal minors $x_i y_j - x_j y_i$. The difference in our case is that we start from an equation T which is not a generic linear combination of these 2×2 determinants, since it is itself a 2×2 determinant. Still, motivated by our experiments, we analyze the complexity of solving $\mathcal{B}$ by using such a generic benchmark.

Assumption 1 *We assume that the behavior of $\mathcal{B}$ matches that of m generic equations that are linear combinations of $x_i y_j - x_j y_i$ for $1 \leq i < j \leq n$ to study algebraic algorithms.*

Degree falls in $\mathcal{B}_*$ for Mutant XL To determine at which bi-degrees we may find degree falls in $\mathcal{B}_*$, we study the system $\mathcal{B}_*^{\text{top}}$ that contains the bihomogeneous components of bi-degree $(1, 1)$ of the equations of $\mathcal{B}_*$.

As $\mathcal{B}_*$ is the system $\mathcal{B}$ to which we specialized the values $x_{n-1} = y_n = 1$ and $x_n = y_{n-1} = 0$, we observe that the equations of $\mathcal{B}_*^{\text{top}}$ are linear combinations between the 2×2 minors of the matrix

$$\boldsymbol{Z}_0 = \begin{pmatrix} x_1 & \dots & x_{n-2} \\ y_1 & \dots & y_{n-2} \end{pmatrix}.$$

This implies that

$$I \stackrel{def}{=} \langle \mathcal{B}_*^{\text{top}} \rangle \subset I_2(\boldsymbol{Z}_0) \subset S \stackrel{def}{=} \mathbb{F}_q[x_1, \ldots, x_{n-2}, y_1, \ldots, y_{n-2}],$$

where the ideal $I_2(\boldsymbol{Z}_0)$ is defined by

$$I_2(\boldsymbol{Z}_0) \stackrel{def}{=} \langle x_i y_j - x_j y_i, 1 \leq i < j \leq n-2 \rangle .$$

Under Assumption 1, $\mathcal{B}$ behaves generically. Therefore, we expect that $\mathcal{B}_*^{\text{top}}$ obtained from $\mathcal{B}$ via $x_{n-1} = y_{n-1} = x_n = y_n = 0$, also behaves generically. In particular, we make the following conjecture on the rank of the homogeneous Macaulay matrix $\boldsymbol{M}_{\boldsymbol{d}}(\mathcal{B}_*^{\text{top}})$.

Conjecture 2 *Let $\boldsymbol{d} = (d_x, d_y)$. The rank of the Macaulay matrix $\boldsymbol{M}_{\boldsymbol{d}}(\mathcal{B}_*^{top})$ is equal to $\min(a_{\boldsymbol{d}}, b_{\boldsymbol{d}})$, where*

$$a_{\boldsymbol{d}} \stackrel{def}{=} [s^{d_x} t^{d_y}] \frac{1 - (1-st)^m}{(1-s)^{n-2}(1-t)^{n-2}},$$

and

$$b_{\boldsymbol{d}} \stackrel{def}{=} [s^{d_x} t^{d_y}] \left(\frac{1}{(1-s)^{n-2}(1-t)^{n-2}} - \frac{s(1-t)^{n-2} - t(1-s)^{n-2}}{(s-t)(1-s)^{n-2}(1-t)^{n-2}} \right).$$

Experiments 1 (Conjecture 2) *Our experiments to test Conjecture 2 can be found in Appendix B.2, where we also give an explanation for the cases in which the conjecture fails. To facilitate the reading, we do not modify the conjecture to account for these few cases.*

Conjecture 2 is motivated as follows. On the one hand, the value $a_{\boldsymbol{d}}$ is equal to $\dim S_{\boldsymbol{d}}$ minus the coefficient of $s^{d_x} t^{d_x}$ in the series obtained from Equation (4) with $n_x = n_y = n-2$ variables in both blocks. While this seems to be the natural series to consider here, we still have no proof for such an expression. On the other hand, the definition of $b_{\boldsymbol{d}}$ comes from the Hilbert bi-series of $S/I_2(\boldsymbol{Z}_0)$. The formula for this Hilbert series is well known in the literature on determinantal ideals.

Lemma 2. *The Hilbert bi-series of $S/I_2(\boldsymbol{Z}_0)$ is equal to*

$$HS_{S/I_2(\boldsymbol{Z}_0)}(s,t) = \frac{s(1-t)^{n-2} - t(1-s)^{n-2}}{(s-t)(1-s)^{n-2}(1-t)^{n-2}}.$$

Proof. We give a proof in Appendix A.

Coming back to MXL, Conjecture 2 gives an estimate for the number of degree falls that we expect in bi-degree $\boldsymbol{d} = (d_x, d_y)$, namely $a_{\boldsymbol{d}} - b_{\boldsymbol{d}}$. As explained at the end of Subsection 2.3, this is a possibly non-tight upper bound, since there may be linear dependencies among the degree-fall polynomials, and some of them could also be redundant for the computation. More importantly, we do not know the behavior of MXL once these degree falls are included in the computation. Since this behavior is also unknown for generic bilinear equations, which may correspond to a simpler case, we believe that a more precise analysis of MXL in our case is beyond the scope of this paper.

Termination of XL As in [1], we consider the bihomogenization $\mathcal{B}_*^{(h)}$ of $\mathcal{B}_*$, introduced in Definition 8, in order to analyze the complexity of plain XL applied to $\mathcal{B}_*$.

In the same spirit as in the previous subsection, we remark that the equations of $\mathcal{B}_*^{(h)}$ are linear combinations between the maximal minors of the matrix

$$\boldsymbol{Z}_1 = \begin{pmatrix} \boldsymbol{Z}_0 & \begin{matrix} x_0 & 0 \\ 0 & y_0 \end{matrix} \end{pmatrix}.$$

Let $S' = \mathbb{F}_q[x_1, \ldots, x_{n-2}, y_1, \ldots, y_{n-2}, x_0, y_0]$ and let $I_2(\boldsymbol{Z}_1)$ be the ideal generated by the maximal minors of $\boldsymbol{Z}_1$. We have a result similar to Lemma 2 for the Hilbert series of the quotient $S'/I_2(\boldsymbol{Z}_1)$.

Lemma 3. *The Hilbert bi-series of $S'/I_2(\boldsymbol{Z}_1)$ is equal to*

$$HS_{S'/I_2(\boldsymbol{Z}_1)}(s,t) = \frac{s(1-t)^n - t(1-s)^n}{(s-t)(1-s)^{n-1}(1-t)^{n-1}}.$$

Proof. We give a proof in Appendix A.

Then, following the same method as the one used to derive Conjecture 2, we obtain the following conjecture.

Conjecture 3 *Let $\boldsymbol{d} = (d_x, d_y)$. The rank of the Macaulay matrix $\boldsymbol{M}_{\boldsymbol{d}}(\mathcal{B}_*^{(h)})$ is equal to $\min(a'_{\boldsymbol{d}}, b'_{\boldsymbol{d}} - 1)$, where*

$$a'_{\boldsymbol{d}} \stackrel{def}{=} [s^{d_x} t^{d_y}] \frac{1-(1-st)^m}{(1-s)^{n-1}(1-t)^{n-1}},$$

and

$$b'_{\boldsymbol{d}} \stackrel{def}{=} [s^{d_x} t^{d_y}] \left(\frac{1}{(1-s)^{n-1}(1-t)^{n-1}} - \frac{s(1-t)^n - t(1-s)^n}{(s-t)(1-s)^{n-1}(1-t)^{n-1}} \right).$$

Experiments 2 (Conjecture 3) *Our experiments to test Conjecture 3 can be found in Appendix B.2. We observed the same phenomenon as with Conjecture 2.*

Under Conjecture 3, the system $\mathcal{B}_*$ can be solved by the XL algorithm at any bi-degree $\boldsymbol{d}$ such that $a'_{\boldsymbol{d}} \geq b'_{\boldsymbol{d}}$. For the complexity analysis, we may assume that the number of columns in $\boldsymbol{M}_{\boldsymbol{d}}(\mathcal{B}_*^{(h)})$ can be taken as $b'_{\boldsymbol{d}}$ instead of the total number of monomials in bi-degree $\boldsymbol{d}$ by fixing a basis of $I_2(\boldsymbol{Z}_1)_{\boldsymbol{d}}$. As is standard, we may further assume that we can restrict to a square submatrix of $\boldsymbol{M}_{\boldsymbol{d}}(\mathcal{B}_*^{(h)})$ without decreasing the rank. The complexity of dense linear algebra on this matrix is then

$$\mathcal{O}\left((b'_{\boldsymbol{d}})^\omega\right),$$

where $2 \leq \omega \leq 3$ is the linear algebra constant. Since we do not control the sparsity of the Macaulay matrix after replacing $S'_{\boldsymbol{d}}$ by $I_2(\boldsymbol{Z}_1)_{\boldsymbol{d}}$, we prefer not to give a complexity estimate for sparse linear algebra techniques on this matrix.

Comparison to the original estimate Note that we would already obtain a gain over the original complexity estimate given in Equation (2) by relying on estimates derived for generic overdetermined bilinear systems instead of underdetermined versions, namely [1] for bi-degrees $(1, d_y)$, and heuristic estimates based on the negative coefficients in the conjectured series of Equation (4) for general bi-degrees.

In our case, the condition $a_{\boldsymbol{d}} \geq b_{\boldsymbol{d}}$ from Conjecture 2 is easier to satisfy than the condition imposed by the negativity of the coefficient of bi-degree $\boldsymbol{d}$ in Equation (4), namely $a_{\boldsymbol{d}} \geq \dim S_{\boldsymbol{d}}$. We have an analogous statement for Conjecture 3. This suggests that solving m linear combinations of the $x_i y_j - x_j y_i$'s for $1 \leq i < j \leq n$ will involve Macaulay matrices in lower bi-degrees than those needed to solve a generic overdetermined bilinear system with m equations. Thus, we always obtain a complexity gain.

However, this complexity gain coming from the determinantal structure should not be significant asymptotically. Indeed, we have

$$\begin{aligned} b_{\boldsymbol{d}} &= \binom{n-3+d_x}{d_x}\binom{n-3+d_y}{d_y} - \binom{n-3+d_x+d_y}{d_x+d_y} \\ &= \dim S_{\boldsymbol{d}} - \binom{n-3+d_x+d_y}{d_x+d_y}, \end{aligned}$$

and thus $b_{\boldsymbol{d}}$ behaves as $\dim S_{\boldsymbol{d}}$ when either d_x or d_y grows to infinity.

For completeness, we characterize this gain over the original complexity by applying Conjecture 2 to the original SBC parameters, as we did for Conjecture 1. We find that the first degree falls appear in bi-degree $(1, 122)$, i.e., $a_{(1,122)} > b_{(1,122)}$, which can be found by constructing Macaulay matrices of size roughly 2^{250}. As for the approach based on Plücker relations, the complexity of the linear algebra on such matrices, while much lower than the one given by Equation (2), is still too high for practical purposes.

5 Adding the Field Equations of $\mathbb{F}_2$

We finish by explaining why analyzing the solving of $\mathcal{B}_*$ if we include the field equations of $\mathbb{F}_2$ seems difficult. This difficulty also applies to the approach described in Subsection 4.2, namely, we have not been able to obtain an analogue of Conjecture 1 for the resulting system after adding the field equations.

To begin with, the case of a generic affine system $\mathcal{F}_{\mathbb{F}_2} = \mathcal{F} \cup \{x_1^2 - x_1, \ldots, x_n^2 - x_n\}$ in $\mathbb{F}_2[x_1, \ldots, x_n]$ that contains the field equations of $\mathbb{F}_2$ was investigated in [3,5,4], which all reduce to studying a homogeneous system. More precisely, [4, Corollary 1] considers the homogenization of $\mathcal{F}_{\mathbb{F}_2}$ in $\mathbb{F}_2[x_1, \ldots, x_n, h]$, while [3,5] study $\mathcal{F}_{\mathbb{F}_2}^{\text{top}}$. We focus on [3,5]. There, the Hilbert series is derived by studying syzygies of the image of $\mathcal{F}^{\text{top}}$ in the quotient

$$\mathbb{F}_2[x_1, \ldots, x_n]/\langle x_1^2, \ldots, x_n^2 \rangle^{3}.$$

[3] In [4], the relevant quotient, with a similar analysis, is

$$\mathbb{F}_2[x_1, \ldots, x_n, h]/\langle x_1^2 - x_1 h, \ldots, x_n^2 - x_n h \rangle.$$

To try to adapt this methodology to affine bilinear systems over $\mathbb{F}_2$ with the field equations added, it could make sense to consider the quotient of $\mathbb{F}_2[\boldsymbol{x}, \boldsymbol{y}]$ by the bihomogeneous ideal

$$\langle x_1^2, \ldots, x_n^2, y_1^2, \ldots, y_n^2 \rangle,$$

and perform a similar analysis of syzygies. However, as we discussed above, the case of overdetermined generic bilinear systems is already not fully understood, and there is no reason to expect this analysis to be easier, even when starting from fewer equations than variables (the addition of the field equations renders the system overdetermined).

Another difficulty in our case is how to treat the generic setting defined right above Assumption 1. Even though the Hilbert series of the quotient by $\langle x_1^2, \ldots, x_n^2, y_1^2, \ldots, y_n^2 \rangle$ is trivial to derive, this generic setting may require considering the quotient by $\langle x_1^2, \ldots, x_n^2, y_1^2, \ldots, y_n^2 \rangle + \boldsymbol{I}_2(\boldsymbol{Z})$ (recall that the matrix $\boldsymbol{Z}$ is defined in Equation (3) and that $\boldsymbol{I}_2(\boldsymbol{Z})$ is the ideal generated by its maximal minors). We have not been able to obtain this Hilbert series, even though we can easily observe many nontrivial relations. For instance, for any two indexes $1 \leq i < j \leq n$, the polynomial $f = x_i y_j + x_j y_i \in \boldsymbol{I}_2(\boldsymbol{Z})$ satisfies

$$\begin{cases} x_i x_j \times f + x_j y_j \times x_i^2 + x_i y_i \times x_j^2 = 0, \\ x_i y_i \times f + y_i y_j \times x_i^2 + x_i x_j \times y_i^2 = 0, \\ x_j y_j \times f + y_i y_j \times x_j^2 + x_i x_j \times y_j^2 = 0, \\ y_i y_j \times f + x_j y_j \times y_i^2 + x_i y_i \times y_j^2 = 0. \end{cases} \tag{6}$$

If one can determine this Hilbert series and, if one were to improve the understanding of the generic overdetermined bilinear case, we believe that analyzing the generic setting relevant to studying $\mathcal{B}$ should be feasible.

A final difficulty is that Assumption 1 no longer holds if we include the field equations of $\mathbb{F}_2$. This is because the equation T is itself a 2×2 minor. This structure implies that the relations of Equation (6) extend from "monomials" to the equations themselves. More precisely, we have in the quotient of $\mathbb{F}_{2^m}[\boldsymbol{x}, \boldsymbol{y}]$ by $\langle x_1^2, \ldots, x_n^2, y_1^2, \ldots, y_n^2 \rangle$,

$$\begin{cases} \langle \boldsymbol{u}, \boldsymbol{x} \rangle \langle \boldsymbol{v}, \boldsymbol{x} \rangle \times T(\boldsymbol{x}, \boldsymbol{y}) = 0 \\ \langle \boldsymbol{u}, \boldsymbol{y} \rangle \langle \boldsymbol{v}, \boldsymbol{y} \rangle \times T(\boldsymbol{x}, \boldsymbol{y}) = 0 \\ \langle \boldsymbol{u}, \boldsymbol{x} \rangle \langle \boldsymbol{u}, \boldsymbol{y} \rangle \times T(\boldsymbol{x}, \boldsymbol{y}) = 0 \\ \langle \boldsymbol{v}, \boldsymbol{x} \rangle \langle \boldsymbol{v}, \boldsymbol{y} \rangle \times T(\boldsymbol{x}, \boldsymbol{y}) = 0. \end{cases}$$

These relations induce m times more relations in $\mathcal{B}$, namely m relations in bi-degrees $\boldsymbol{d} = (1, 3)$ and $\boldsymbol{d} = (3, 1)$ respectively, and $2m$ relations in bi-degree $\boldsymbol{d} = (2, 2)$ (this was confirmed in our experiments). However, if these first ones are easy to take into account to adapt Conjectures 2 and 3, a complete adaptation is out of reach for now. Note already that relations at degree $\boldsymbol{d} = (2, 3)$ will

involve both bi-degrees $(1,3)$ and $(2,2)$, and so we were not able to properly count them (this already calls for understanding the overlap between these two sets of relations when we augment them in bi-degree $(2,3)$).

6 Conclusion

In this paper, we have shown that the ad hoc Subfield Bilinear Collision problem introduced by Huth and Joux corresponds to particular instances of the problem of finding low-weight codewords in an $\mathbb{F}_{q^m}$-linear code with respect to the rank metric. While this connection does not allow us to find attacks that break the parameters of [19,20], it can still be seen as strengthening the cryptanalysis of [19]. Independently of this, we have highlighted a determinantal structure in the bilinear system proposed in [19] that has to be taken into account in the analysis of algebraic attacks. This structure may have applications beyond the Subfield Bilinear Collision problem. For example, it also arises in a subsequent[4] work [11] on the cryptanalysis of the McEliece cryptosystem [25].

We believe that a hybrid algebraic attack that could threaten the parameters of [19,20] would be along the lines of a bilinear adaptation of BooleanSolve [4] or of the Crossbred algorithm [24]. Note that both of them take into account the field equations of $\mathbb{F}_2$, so we believe that studying their complexity would be difficult, as detailed in Section 5.

References

1. Baena, J.B., Cabarcas, D., Verbel, J.: On the complexity of solving generic overdetermined bilinear systems. Advances in Mathematics of Communications **17**(5), 1210–1234 (2023)
2. Bardet, M., Bros, M., Cabarcas, D., Gaborit, P., Perlner, R., Smith-Tone, D., Tillich, J.P., Verbel, J.: Improvements of Algebraic Attacks for Solving the Rank Decoding and MinRank Problems. In: Advances in Cryptology – ASIACRYPT 2020: 26th International Conference on the Theory and Application of Cryptology and Information Security, Daejeon, South Korea, December 7–11, 2020, Proceedings. p. 507–536. Springer (2020)
3. Bardet, M., Faugère, J.C., Salvy, B.: Complexity of Gröbner basis computation for Semi-regular Overdetermined sequences over F_2 with solutions in F_2. Research Report RR-5049, INRIA (2003), https://inria.hal.science/inria-00071534
4. Bardet, M., Faugère, J.C., Salvy, B., Spaenlehauer, P.J.: On the complexity of solving quadratic Boolean systems. Journal of Complexity **29**(1), 53–75 (2013)
5. Bardet, M., Faugère, J.C., Salvy, B., Yang, B.Y.: Asymptotic expansion of the degree of regularity for semi-regular systems of equations. The Effective Methods in Algebraic Geometry Conference p. 1–14 (01 2005)
6. Baum, C., Braun, L., de Saint Guilhem, C.D., Klooß, M., Orsini, E., Roy, L., Scholl, P.: Publicly Verifiable Zero-Knowledge and Post-Quantum Signatures from VOLE-in-the-Head. In: Advances in Cryptology - CRYPTO 2023 - 43rd Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2023, Proceedings. vol. 14085, p. 581–615. Springer (2023)

[4] The first version of our work was made public on May 12, 2026.

7. Beeloo-Sauerbier Couvée, H., Wachter-Zeh, A., Weger, V.: Hybrid Subsupport Guessing: A New Hybrid Technique for the Rank Decoding Problem. In: Post-Quantum Cryptography. p. 130–155. Springer (2026)
8. Bernstein, D., Zelevinsky, A.: Combinatorics of maximal minors. J. Algebraic Comb. **2**(2), 111–121 (Jun 1993). https://doi.org/10.1023/A:1022492222930, https://doi.org/10.1023/A:1022492222930
9. Beullens, W.: Improved Cryptanalysis of UOV and Rainbow. In: Advances in Cryptology – EUROCRYPT 2021: 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, October 17–21, 2021, Proceedings. p. 348–373. Springer (2021)
10. Bouillaguet, C., Feneuil, T., Maire, J., Rivain, M., Sauvage, J., Vergnaud, D.: Multivariate Commitments and Signatures with Efficient Protocols. Cryptology ePrint Archive, Paper 2025/2035 (2025), https://eprint.iacr.org/2025/2035
11. Briaud, P., Lemoine, A., Randriambololona, H., Tillich, J.P.: A Heuristic Subexponential Attack on the McEliece Cryptosystem. Cryptology ePrint Archive, Paper 2026/1232 (2026), https://eprint.iacr.org/2026/1232
12. Buchmann, J.A., Ding, J., Mohamed, M.S.E., Mohamed, W.S.A.E.: MutantXL: Solving Multivariate Polynomial Equations for Cryptanalysis. In: Symmetric Cryptography. Dagstuhl Seminar Proceedings (DagSemProc), vol. 9031, p. 1–7. Dagstuhl, Germany (2009)
13. Courtois, N., Klimov, A., Patarin, J., Shamir, A.: Efficient Algorithms for Solving Overdefined Systems of Multivariate Polynomial Equations. In: Advances in Cryptology — EUROCRYPT 2000. p. 392–407. Springer (2000)
14. Couvreur, A., Mora, R., Tillich, J.P.: A New Approach Based on Quadratic Forms to Attack the McEliece Cryptosystem. In: Advances in Cryptology – ASIACRYPT 2023. p. 3–38. Springer, Singapore (2023)
15. Cox, D.A., Little, J., O'Shea, D.: Ideals, Varieties, and Algorithms. Springer International Publishing (2015). https://doi.org/10.1007/978-3-319-16721-3, https://doi.org/10.1007/978-3-319-16721-3
16. Debris-Alazard, T., Remaud, M., Tillich, J.P.: Quantum Reduction of Finding Short Code Vectors to the Decoding Problem. IEEE Transactions on Information Theory **PP**, 1–1 (01 2023)
17. Faugère, J.C., Safey El Din, M., Spaenlehauer, P.J.: Gröbner bases of bihomogeneous ideals generated by polynomials of bidegree (1,1): Algorithms and complexity. Journal of Symbolic Computation **46**(4), 406–437 (2011)
18. Ghorpade, S.R.: The Hilbert Series of Pfaffian Rings. In: Algebra, Arithmetic and Geometry with Applications: Papers from Shreeram S. Abhyankar's 70th Birthday Conference. p. 337–356 (2004)
19. Huth, J., Joux, A.: MPC in the Head Using the Subfield Bilinear Collision Problem. In: Advances in Cryptology – CRYPTO 2024: 44th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18–22, 2024, Proceedings. p. 39–70. Springer (2024)
20. Huth, J., Joux, A.: VOLE-in-the-Head Signatures from Subfield Bilinear Collisions. In: Advances in Cryptology - ASIACRYPT 2025 - 31st International Conference on the Theory and Application of Cryptology and Information Security, Melbourne, VIC, Australia, December 8-12, 2025, Proceedings. vol. 16248, p. 3–34. Springer (2025)
21. Ishai, Y., Kushilevitz, E., Ostrovsky, R., Sahai, A.: Zero-Knowledge from Secure Multiparty Computation. In: Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing. p. 21–30. STOC '07, Association for Computing Machinery (2007)

22. Joux, A.: A New Index Calculus Algorithm with Complexity L(1/4+o(1)) in Small Characteristic. In: Selected Areas in Cryptography - SAC 2013 - 20th International Conference, Burnaby, BC, Canada, August 14-16, 2013. vol. 8282, p. 355–379. Springer (2013)
23. Joux, A., Pierrot, C.: Algorithmic aspects of elliptic bases in finite field discrete logarithm algorithms. Advances in Mathematics of Communications **18**(5), 1266–1302 (2024)
24. Joux, A., Vitse, V.: A Crossbred Algorithm for Solving Boolean Polynomial Systems. In: International Conference on Number-Theoretic Methods in Cryptology (2017), https://api.semanticscholar.org/CorpusID:3776552
25. McEliece, R.J.: A Public-Key System Based on Algebraic Coding Theory, p. 114–116. Jet Propulsion Lab (1978), dSN Progress Report 42-44
26. Ning, K.C., Ran, L., Samardjiska, S.: Multi-homogeneous XL. Cryptology ePrint Archive, Paper 2025/2060 (2025), https://eprint.iacr.org/2025/2060
27. NIST: Post-Quantum Cryptography: Digital Signature Schemes (2023), https://csrc.nist.gov/projects/pqc-dig-sig/standardization/call-for-proposals
28. Perlner, R., Smith-Tone, D.: Rainbow Band Separation is Better than we Thought. Cryptology ePrint Archive, Paper 2020/702 (2020), https://eprint.iacr.org/2020/702
29. Verbel, J., Baena, J., Cabarcas, D., Perlner, R., Smith-Tone, D.: On the Complexity of "Superdetermined" Minrank Instances. In: Post-Quantum Cryptography. p. 167–186. Springer (2019)

A Proofs of Lemmata 2 and 3

For Lemma 2, we consider the polynomial ring $S = \mathbb{F}_q[x_1, \dots, x_{n-2}, y_1, \dots, y_{n-2}]$, the matrix

$$\boldsymbol{Z}_0 = \begin{pmatrix} x_1 & \dots & x_{n-2} \\ y_1 & \dots & y_{n-2} \end{pmatrix},$$

and the ideal $I_2(\boldsymbol{Z}_0) = \langle x_i y_j - x_j y_i, 1 \leq i < j \leq n-2 \rangle$.

Lemma 4 (Lemma 2 restated). *The Hilbert bi-series of the quotient ring $S/I_2(\boldsymbol{Z}_0)$ is equal to*

$$HS_{S/I_2(\boldsymbol{Z}_0)}(s,t) = \frac{s(1-t)^{n-2} - t(1-s)^{n-2}}{(s-t)(1-s)^{n-2}(1-t)^{n-2}}.$$

Proof. Let $n' \stackrel{def}{=} n-2$. A very powerful result states that the set of maximal minors of a matrix of variables forms a universal Gröbner basis [8]. Our case is the particular instance of this result in which the matrix has size $2 \times n'$. Moreover, note that for any graded monomial order $\prec$ on S, we have

$$\mathrm{HS}_{S/I_2(\boldsymbol{Z}_0)} = \mathrm{HS}_{S/\mathrm{in}_\prec(I_2(\boldsymbol{Z}_0))},$$

where $\mathrm{in}_\prec(I)$ denotes the initial ideal of an ideal I with respect to $\prec$ [15, Chapter 9]. Therefore, we may choose such a graded order $\prec$ for which $\mathrm{in}_\prec(I_2(\boldsymbol{Z}_0))$ is

convenient to study. We can, for example, consider the graded lexicographic order $\prec$ with variable ordering

$$x_1 > x_2 > \cdots > x_{n'} > y_1 > y_2 > \cdots > y_{n'},$$

for which $\mathrm{in}_\prec(I_2(\boldsymbol{Z}_0))$ is generated by all monomials $x_i y_j$ such that $1 \leq i < j \leq n'$. We are left with computing the Hilbert series of the quotient by this monomial ideal. A general monomial $\mu \in S$ can be written as

$$\mu = x_1^{a_1} \dots x_{n'}^{a_{n'}} y_1^{b_1} \dots y_{n'}^{b_{n'}},$$

and we set $A = \{i \mid a_i > 0\}$ and $B = \{i \mid b_i > 0\}$. The condition that μ is not divisible by $x_i y_j$ for any $1 \leq i < j \leq n'$ is equivalent to the inequality $\max(B) \leq \min(A)$ (with the convention $\min(\emptyset) = +\infty$ and $\max(\emptyset) = 0$). We now count such monomials μ for which $r = \min(A)$ is fixed. These are precisely the monomials in which the variables $x_1, \dots, x_{r-1}$ do not occur, the variable x_r occurs with a positive exponent, and the variables $y_{r+1}, \dots, y_{n'}$ do not occur. The associated generating series is thus

$$\underbrace{\frac{s}{1-s}}_{x_r} \times \underbrace{\frac{1}{(1-s)^{n'-r}}}_{x_{r+1},\dots,x_{n'}} \times \underbrace{\frac{1}{(1-t)^r}}_{y_1,\dots,y_r} = \frac{s}{(1-s)^{n'-r+1}(1-t)^r}.$$

We finish by summing these series over $1 \leq r \leq n'$ and adding the contribution $\frac{1}{(1-t)^{n'}}$ of the monomials that do not involve any variable from the first row of $\boldsymbol{Z}_0$. We obtain

$$\begin{aligned}
\mathrm{HS}_{S/I_2(\boldsymbol{Z}_0)} &= \frac{1}{(1-t)^{n'}} + \sum_{r=1}^{n'} \frac{s}{(1-s)^{n'-r+1}(1-t)^r} \\
&= \frac{1}{(1-t)^{n'}} + \frac{s}{(1-s)^{n'+1}} \sum_{r=1}^{n'} \left(\frac{1-s}{1-t}\right)^r \\
&= \frac{1}{(1-t)^{n'}} + \frac{s}{(s-t)(1-s)^{n'}} \left(1 - \left(\frac{1-s}{1-t}\right)^{n'}\right) \\
&= \frac{(s-t)(1-s)^{n'} + s(1-t)^{n'} - s(1-s)^{n'}}{(s-t)(1-s)^{n'}(1-t)^{n'}} \\
&= \frac{s(1-t)^{n'} - t(1-s)^{n'}}{(s-t)(1-s)^{n'}(1-t)^{n'}},
\end{aligned}$$

and the last line is the statement of the lemma.

For Lemma 3, we consider $S' = \mathbb{F}_q[x_1, \dots, x_{n-2}, y_1, \dots, y_{n-2}, x_0, y_0]$, the matrix

$$\boldsymbol{Z}_1 = \begin{pmatrix} \boldsymbol{Z}_0 & \begin{matrix} x_0 & 0 \\ 0 & y_0 \end{matrix} \end{pmatrix},$$

and the ideal $I_2(\boldsymbol{Z}_1)$ generated by the maximal minors of this matrix.

Lemma 5 (Lemma 3 restated). *The Hilbert bi-series of* $S'/I_2(\boldsymbol{Z}_1)$ *is equal to*

$$HS_{S'/I_2(\boldsymbol{Z}_1)}(s,t) = \frac{s(1-t)^n - t(1-s)^n}{(s-t)(1-s)^{n-1}(1-t)^{n-1}}.$$

Proof. From the previous proof, we know that the set of maximal minors of $\boldsymbol{Z}_0$ is a universal Gröbner basis. The remaining maximal minors of $\boldsymbol{Z}_1$ are x_0y_i for $1 \leq i \leq n-2$, y_0x_i for $1 \leq i \leq n-2$, and x_0y_0. In particular, we can easily check the whole set of maximal minors of $\boldsymbol{Z}_1$ still forms a Gröbner basis by computing S-polynomials. This gives us the initial ideal, and we can then proceed as at the end of the previous proof to compute the Hilbert bi-series.

B Experiments

We used MAGMA and SageMath for our experiments.

B.1 Experiments for Conjecture 1

Starting from an NSBC instance, we tested Conjecture 1 by computing the Hilbert series of $R/\langle \mathcal{L} \cup \mathcal{Q} \cup \{z_{n-1,n}\}\rangle$. We have been able to run experiments for n up to 14, and for these values of n we have tested $2n \leq m \leq 3n$. Our experiments always produced the expected result for these parameters when the input ideal was zero-dimensional.

B.2 Experiments for Conjectures 2 and 3

Our experiments to test Conjectures 2 and 3 can be found in Tables 1 and 2, respectively. We constructed the Macaulay matrices by hand and computed their ranks by using built-in functions.

n	m	$\boldsymbol{d}$	$a_{\boldsymbol{d}}$	$b_{\boldsymbol{d}}$	$\operatorname{rk} \boldsymbol{M}_{\boldsymbol{d}}(\mathcal{B}_*^{\mathrm{top}})$
12	20	$(1,1)$	20	45	20
		$(1,2)$	200	330	200
		$(1,3)$	1100	1485	1100
		$(2,2)$	1810	2310	1810
	24	$(1,1)$	24	45	24
		$(1,2)$	240	330	240
		$(1,3)$	1320	1485	1320
		$(1,4)$	5280	5148	5148
		$(2,2)$	2124	2310	2124
		$(2,3)$	10440	10098	10098
	25	$(2,2)$	2200	2310	2200
	26	$(2,2)$	2275	2310	2255
	28	$(1,1)$	28	45	28
		$(1,2)$	280	330	280
		$(1,3)$	1540	1485	1485
		$(2,2)$	2422	2310	2310
14	28	$(1,3)$	2184	3003	2184
		$(1,4)$	10192	12012	10192
		$(2,2)$	3654	4719	3654
		$(2,3)$	21672	24024	21672
	36	$(1,3)$	2808	3003	2808
		$(1,4)$	13104	12012	12012
		$(2,2)$	4554	4719	4524
	37	$(2,2)$	4662	4719	4602
	38	$(2,2)$	4769	4719	4680
	39	$(2,2)$	4875	4719	4719
15	40	$(2,2)$	5980	6461	5980
	41	$(2,2)$	6109	6461	6097
	42	$(2,2)$	6237	6461	6188
	43	$(2,2)$	6364	6461	6279
	44	$(2,2)$	6490	6461	6370
	45	$(2,2)$	6615	6461	6461

Table 1. Rank of the Macaulay matrix $\boldsymbol{M}_{\boldsymbol{d}}(\mathcal{B}_*^{\mathrm{top}})$. The quantities $a_{\boldsymbol{d}}$ and $b_{\boldsymbol{d}}$ are defined in Conjecture 2.

n	m	$\boldsymbol{d}$	$a'_{\boldsymbol{d}}$	$b'_{\boldsymbol{d}}$	$\operatorname{rk} \boldsymbol{M}_{\boldsymbol{d}}(\mathcal{B}_*^{(h)})$
11	17	$(1,2)$	170	385	170
		$(1,3)$	935	1705	935
		$(1,4)$	3740	5863	3740
		$(2,2)$	1564	2530	1564
		$(2,3)$	7990	10813	7990
	22	$(1,2)$	220	385	220
		$(1,3)$	1210	1705	1210
		$(1,4)$	4840	5863	4840
		$(2,2)$	1969	2530	1969
		$(2,3)$	9790	10813	9790
	32	$(1,2)$	320	385	320
		$(1,3)$	1760	1705	1704
		$(1,4)$	7040	5863	5862
		$(2,2)$	2720	2530	2529
		$(2,3)$	12640	10813	10812
12	19	$(1,2)$	209	506	209
		$(1,3)$	1254	2431	1254
		$(1,4)$	5434	9009	5434
		$(2,2)$	2128	3641	2128
		$(2,3)$	11913	16874	11913
	24	$(1,2)$	264	506	264
		$(1,3)$	1584	2431	1584
		$(1,4)$	6864	9009	6864
		$(2,2)$	2628	3641	2628
		$(2,3)$	14388	16874	14388
	34	$(1,2)$	374	506	374
		$(1,3)$	2244	2431	2244
		$(1,4)$	9724	9009	9008
		$(2,2)$	3553	3641	3553
		$(2,3)$	18513	16874	16873
13	40	$(2,2)$	4980	5083	4980
	41	$(2,2)$	5084	5083	5071
	42	$(2,2)$	5187	5083	5082
14	23	$(1,2)$	299	819	299
		$(1,3)$	2093	4550	2093
		$(1,4)$	10465	19292	10465
		$(2,2)$	3634	6916	3634
		$(2,3)$	23920	37037	23920
	28	$(1,2)$	364	819	364
		$(1,3)$	2548	4550	2548
		$(1,4)$	12740	19292	12740
		$(2,2)$	4354	6916	4354
		$(2,3)$	28210	37037	28210
	38	$(1,2)$	494	819	494
		$(1,3)$	3458	4550	3458
		$(1,4)$	17290	19292	17290
		$(2,2)$	5719	6916	5719
		$(2,3)$	35815	37037	35815
	46	$(2,2)$	6739	6916	6739
	47	$(2,2)$	6862	6916	6850
	48	$(2,2)$	6984	6916	6915

Table 2. Rank of the Macaulay matrix $\boldsymbol{M}_{\boldsymbol{d}}(\mathcal{B}_*^{(h)})$. The quantities $a'_{\boldsymbol{d}}$ and $b'_{\boldsymbol{d}}$ are defined in Conjecture 3.

We give here an explanation for the cases where Conjectures 2 and 3 fail, as hinted just below these conjectures. These conjectures would fail in the same way for m arbitrary linear combinations of the maximal minors of the matrix $\boldsymbol{Z}$ defined in Equation (3). Let $\mathcal{B}$ denote the corresponding bilinear system, let $\mathcal{L}$ be the system $\mathcal{B}$ interpreted as a linear system in $R = \mathbb{F}_q[\boldsymbol{z}_T]$ and let $\mathcal{Q} \in R^{\binom{n}{4}}$ be the set of Plücker relations introduced in Definition 9. If for some degree d, it holds that

$$\langle \mathcal{L} \rangle_d + \langle \mathcal{Q} \rangle_d = \mathbb{F}_q[\boldsymbol{z}_T]_d, \tag{7}$$

then some elements of $\langle \mathcal{Q} \rangle_d$ will be found in $\langle \mathcal{L} \rangle_d$. When replacing $z_{\{i,j\}}$ with $x_i y_j - x_j y_i$, these elements yield redundant rows in the Macaulay matrix $\boldsymbol{M}_{(d,d)}(\mathcal{B})$. More precisely, we can find a basis of $\boldsymbol{I}_2(\boldsymbol{Z})_{(d,d)}$, such that, if we construct a smaller Macaulay matrix whose columns correspond to this basis instead of the set of all monomials of bi-degree (d, d), we would obtain all-zero rows. Degrees d such that Equation (7) holds can be determined from the Hilbert series of $R/\langle \mathcal{L} \cup \mathcal{Q} \rangle$. This Hilbert series has already been discussed in Subsection 4.2.

We use the above explanation to understand the experimental differences in Table 1. Let $\boldsymbol{Z}_0$ be the matrix whose expression was recalled at the beginning of Appendix A, let R^0 be the polynomial ring in $\binom{n-2}{2}$ variables containing its maximal minors and let $\mathcal{Q}_0 \in (R^0)^{\binom{n-2}{4}}$ denote the set of Plücker relations associated with this matrix. We consider the following generating series

$$(1-t)^m \mathrm{HS}_{R^0/\langle \mathcal{Q}_0 \rangle}, \tag{8}$$

where

$$\mathrm{HS}_{R^0/\langle \mathcal{Q} \rangle}(t) = \frac{\left(\sum_{d=0}^{n-5} \binom{n-4}{d}^2 - \binom{n-5}{d-1}\binom{n-3}{d+1} \right) t^d}{(1-t)^{2n-7}}.$$

This last equation is just a restatement of Equation (5) with $n-2$ columns instead of n. We see that the generating series from Equation (8) accurately predicts the deviations from Conjecture 2. Indeed, the underlined values in Table 1 correspond to the cases when the coefficient of t^d is negative in Equation (8) and $a_{(d,d)} < b_{(d,d)}$. Moreover, the observed rank defect matches the opposite of this negative coefficient.

For Table 2, we can proceed similarly with the Hilbert series of $R/\langle \mathcal{L} \cup \mathcal{Q} \rangle$. For instance, the two underlined values this table correspond to cases where $a'_{(d,d)} < b'_{(d,d)}$ and the coefficient of t^d in the associated generating series is negative. In these two cases, the observed rank defect appears to be equal to the absolute value of this negative coefficient *plus one*. This "plus one" is explained by the fact that, contrary to $\mathcal{B}_*^{\mathrm{top}}$, the system $\mathcal{B}_*^{(h)}$ admits a nontrivial solution.