A polynomial time key-recovery attack on the Sidon cryptosystem.

Pierre Briaud[1,2], Jean-Pierre Tillich[1], and Javier Verbel[3]

[1] Sorbonne Universités, UPMC Univ Paris 06, Paris, France
[2] Inria, Team COSMIQ, Paris, France
pierre.briaud@inria.fr
jean-pierre.tillich@inria.fr
[3] Cryptography Research Centre, Technology Innovation Institute, Abu Dhabi, UAE
javier.verbel@tii.ae

Abstract. The Sidon cryptosystem [21] is a new multivariate encryption scheme based on the theory of Sidon spaces which was presented at PKC 2021. As is usual for this kind of schemes, its security relies on the hardness of solving particular instances of the MQ problem and of the MinRank problem. A nice feature of the scheme is that it enjoys a homomorphic property due the bilinearity of its public polynomials. Unfortunately, we will show that the Sidon cryptosystem can be broken by a polynomial time key-recovery attack. This attack relies on the existence of solutions to the underlying MinRank instance which lie in a subfield and which are inherent to the structure of the secret Sidon space. We prove that such solutions can be found in polynomial time. Our attack consists in recovering an equivalent key for the cryptosystem by exploiting these particular solutions, and this task can be performed very efficiently.

Keywords: Multivariate cryptography · Encryption scheme · Algebraic attack · MinRank Problem.

1 Introduction.

In recent years, many Public Key Cryptographic (PKC) primitives have been proposed to fulfill the need of having quantum-secure Key Encapsulation Mechanisms (KEMs) and Digital Signatures Schemes (DSS).

In the field of DSS, Multivariate Public Key Cryptography (MPKC) has proven to be one of the most promising alternatives. In the ongoing NIST-PQC standardization process, the MPKC schemes Rainbow [11] and GeMSS [8] are in the final round, even though the last as an alternative candidate. But when it comes to KEMs, the situation is less promising. Throughout the last 25 years, many MPKC encryption schemes have been proposed such as HFE, ZHFE, Extension Field Cancellation (EFC), SRP, HFERP, EFLASH and the Simple Matrix Encryption Scheme, see [19,25,24,20,16,7], and all of them are either extremely inefficient or have been successfully cryptanalyzed [4,6,1,23,18,17]. A relative common structure among the MPKC schemes is that the public key is a system of multivariate polynomials which is hard to solve directly. On the contrary, the private key is a sequence of polynomials which is easy to solve and which is masked in certain way to produce the public key. The two main hard problems considered to build MPKC are the problem of solving a system of multivariate quadratic equations over a finite field (the MQ problem) and the following MinRank problem, which was originally defined and proven NP-complete in [5]. The version of the MinRank problem which is relevant in our case is given by

Problem 1 (MinRank problem).

Input: an integer $r \in \mathbb{N}$, n matrices $\boldsymbol{M}^{(1)}, \ldots, \boldsymbol{M}^{(n)} \in \mathbb{F}_q^{k \times k}$, $\mathbb{L}$ a finite extension of $\mathbb{F}_q$.

Output: field elements $x_1, x_2, \ldots, x_n \in \mathbb{L}$, not all zero, such that

$$\operatorname{Rank}\left(\sum_{i=1}^{n} x_i \boldsymbol{M}^{(i)}\right) \leq r.$$

Note that the standard formulation in the literature considers $\mathbb{L} = \mathbb{F}_q$ and possibly non-square matrices. In all cases, solving a particular instance of one of these two problems leads to either a key-recovery attack or a message-recovery attack on a given scheme. Thus, the security of the scheme is usually estimated via the hardness of solving particular instances of MQ or MinRank.

At PKC 2021, Raviv, Langton and Tamo proposed, for the first time, a MPKC encryption scheme based on the theory of Sidon spaces [21]. The concept of Sidon space was originally defined in [2]. A Sidon space is an $\mathbb{F}_q$-subspace of an extension field $\mathbb{F}_{q^n}$ in which the product of any two elements factors uniquely up to multiplicative constants in $\mathbb{F}_q$. The basic idea of the scheme is as follows: the plaintext is the equivalence class of pairs of two elements in a Sidon space $\mathcal{V}$ of dimension k, two pairs with the same product being equivalent, while the ciphertext is the product of these two elements. The private key is some information related to the structure of $\mathcal{V}$ that allows to factor efficiently any ciphertext, while the public key is a bilinear sequence $(p_1, \ldots, p_{2k})$ of $2k$ homogeneous equations in two blocks of variables $\boldsymbol{a}$ and $\boldsymbol{b}$ over $\mathbb{F}_q$ of equal size k. This makes the Sidon cryptosystem to have an additive homomorphic property since the sum of the encryptions of two messages like $\{\boldsymbol{a}_1, \boldsymbol{b}\}$, $\{\boldsymbol{a}_2, \boldsymbol{b}\}$ results in the encryption of $\{\boldsymbol{a}_1 + \boldsymbol{a}_2, \boldsymbol{b}\}$. To the best of our knowledge, this is first MQ- or MinRank- based scheme doted with any kind of homomorphic property.

The private key can be obtained from a solution to the MinRank problem given by the matrices $\boldsymbol{M}^{(i)} \in \mathbb{F}_q^{k \times k}$ defined by $p_i(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{a}\boldsymbol{M}^{(i)}\boldsymbol{b}^\mathsf{T}$ for $1 \leq i \leq n = 2k$, with $r = 1$ and $\mathbb{L} = \mathbb{F}_{q^{2k}}$. It turns out that this MinRank instance has many solutions, and the structure of the Sidon space can be fully extracted from at least one of these solutions. However, to perform a key-recovery attack, the authors argue that

(i) it is not clear how to solve this particular MinRank instance, since standard techniques are strongly based either on the fact that the base field is small (Linear

Algebra Search [15]) or the number of solutions is small (Minors modeling [12] + XL, and Support-Minors modeling [3]).
(ii) Even if one is able to find a solution to the MinRank problem, it is not clear how this solution can be used to develop a key recovery.

Therefore, the message-recovery attack is arguably the most threatening attack which may be used to design parameters, and this attack can be performed by inverting the public bilinear system. When q is large enough, the Gröbner basis approach outperforms the exhaustive search on $\boldsymbol{a}$ or $\boldsymbol{b}$ in $\mathcal{O}(k^3q^{k-1})$, and the authors claim a complexity of

$$\mathcal{O}\left(\binom{3k+1}{k+1}^{\omega}\right) \tag{1}$$

operations in $\mathbb{F}_q$ for this attack, where $2 \leq \omega \leq 3$ is the linear algebra constant. This cost is clearly exponential in k.

Contributions. The purpose of this paper is to give a polynomial-time attack breaking the Sidon cryptosystem. Our attack relies on a rigorous analysis of the solution set of the underlying MinRank problem. In particular, we show here that in addition to generic solutions over $\mathbb{F}_{q^{2k}}$, there exist solutions over the subfield $\mathbb{F}_{q^k}$. Moreover, all these solutions over $\mathbb{F}_{q^k}$ are inherent to the Sidon space used in the scheme. Our attack can be summarized as follows

- The first step of our attack consists in recovering these solutions over $\mathbb{F}_{q^k}$. To this end, we propose a dedicated modeling of the MinRank problem and prove that for this modeling the Gröbner basis computation on this algebraic system terminates at degree 3 independently from the value of k. This shows that this first step can be achieved in polynomial time.
- Second, it is possible to exploit these solutions in order to find an equivalent key which is another Sidon space. This second step can be performed by simple linear algebra operations followed by a sub-algorithm of the original key generation process. Therefore, its cost is also expected to be polynomial.

Along with this paper, we provide a sage implementation of our attack in [26]. This tool can also be used to verify experimentally all the theoretical claims made in the paper and to reproduce the experiments we performed.

Roadmap. The Sidon cryptosystem from [21] is presented in Section 2. In Section 3, we provide a detailed analysis of the underlying MinRank instance, and this material allows us to introduce our key-recovery attack in Section 4 and in Section 5.

Notation. Row vectors are denoted by bold lowercase letters $(\boldsymbol{u}, \boldsymbol{v}, \dots)$ and matrices are denoted by bold uppercase letters $(\boldsymbol{M}, \boldsymbol{N}, \dots)$. For a vector $\boldsymbol{v}$ we use the notation v_i for the i-th component of $\boldsymbol{v}$, and for a matrix $\boldsymbol{M}$ we use the notation $\boldsymbol{M}_{i,j}$ for the entry in row i and column j.

For $\boldsymbol{v}$ a vector of length k, we denote by $\boldsymbol{M}(\boldsymbol{v})$ the rank 1 symmetric matrix of size $k \times k$ which is equal to $\boldsymbol{v}^{\mathsf{T}}\boldsymbol{v}$. In the following, $q \geq 3$ is a prime power, and we will consider finite extensions of $\mathbb{F}_q$, namely $\mathbb{F}_{q^k}$ and $\mathbb{F}_{q^{2k}}$. For $j \in \mathbb{Z}_{\geq 0}$ and $\boldsymbol{v} = (v_1, \dots, v_k)$ a vector whose entries are elements or polynomials over a finite extension of $\mathbb{F}_q$, we define

$$\boldsymbol{v}^{[j]} := (v_1^{q^j}, \dots, v_k^{q^j}).$$

This corresponds to applying the Frobenius automorphism $x \mapsto x^q$ j times on each coordinate of $\boldsymbol{v}$. Note that this field automorphism is the identity on $\mathbb{F}_q$. We will adopt the same notation for matrices, namely the matrix $\boldsymbol{M}^{[j]}$ is the matrix obtained from $\boldsymbol{M}$ by raising all its entries to the power q^j.

We will also adopt in several places a coding theoretic point of view and view a subspace $\mathcal{C}$ of $\mathbb{F}_q^N$ as a linear code and use the term *parity-check* for it to denote a matrix $\boldsymbol{H}$ whose null-space is $\mathcal{C}$, that is:

$$\mathcal{C} = \{\boldsymbol{x} \in \mathbb{F}_q^N : \boldsymbol{H}\boldsymbol{x}^{\mathsf{T}} = 0\}.$$

Finally, it is convenient to consider for a vector space V, ordered bases for it, and we will use a vector notation for the basis.

Polynomial systems. In the following, the expression $\mathbb{K}[\boldsymbol{x}]$ denotes the polynomial ring over the field $\mathbb{K}$ in the coordinates of $\boldsymbol{x} = (x_1, \dots, x_k)$. We will use Gröbner basis techniques to solve polynomial systems, and we refer the reader to [9] for basic definitions and properties of monomial orderings and Gröbner bases.

2 The Sidon cryptosystem.

Several explicit constructions of Sidon spaces with relevant parameters and factoring properties were proposed in [22], and the one used in [21] to instantiate the Sidon cryptosystem is of this kind. In Section 2.1, we give some background on Sidon spaces in general and on this specific construction. The Sidon cryptosystem is presented in Section 2.2, and in Section 2.3 we discuss the important notion of equivalent keys for this scheme.

2.1 Sidon spaces.

For integers k and n and q a prime power, let $\mathcal{G}_q(n, k)$ be the set of all $\mathbb{F}_q$-subspaces of $\mathbb{F}_{q^n}$ of dimension k. The formal definition of a Sidon space is the following.

Definition 1. *A subspace $\mathcal{V} \in \mathcal{G}_q(n, k)$ is called a Sidon space if for all non-zero $a, b, c, d \in \mathcal{V}$, if $ab = cd$, then $\{a\mathbb{F}_q, b\mathbb{F}_q\} = \{c\mathbb{F}_q, d\mathbb{F}_q\}$.*

Then, a first natural question is whether if there exist Sidon spaces of arbitrary dimension. A constraint on k was given by [2, Thm. 18][22, Prop. 3], where it is proven that if $\mathcal{V} \in \mathcal{G}_q(n,k)$ is a Sidon space, then one has

$$\dim_{\mathbb{F}_q}(\mathcal{V}^2) \geq 2k$$

with $\mathcal{V}^2 = \mathrm{span}_{\mathbb{F}_q}\{uv \mid u, v \in \mathcal{V}\}$. Since $\mathcal{V}^2 \subset \mathbb{F}_{q^n}$, this implies that $k \leq n/2$, and Sidon spaces for which this bound is an equality are referred as *min-span.* Note that regardless of the existence of any factoring algorithm for $\mathcal{V}$, it is crucial for the security of the cryptosystem that the dimension of $\mathcal{V}$ satisfies $k = \Theta(n)$, as pointed out in [21, Rem. 2]. In particular, the construction considered to devise the scheme is a *min-span* Sidon space, *i.e.* $n = 2k$. To describe this construction, let $W_{q-1} = \{u^{q-1} | u \in \mathbb{F}_{q^k}\}$ and $\overline{W_{q-1}} = \mathbb{F}_{q^k} \setminus W_{q-1}$.

Construction 1 *[22, Const. 15] For $q \geq 3$ a prime poewer and k a positive integer, let $n = 2k$ and let $\gamma \in \mathbb{F}_{q^n}^*$ be a root of an irreducible polynomial x^2+bx+c over $\mathbb{F}_{q^k}$ such that $c \in \overline{W_{q-1}}$* [4]*. Then, the subspace $\mathcal{V} = \{u+u^q\gamma | u \in \mathbb{F}_{q^k}\} \subset \mathbb{F}_{q^n}$ is a Sidon space of dimension k.*

A Sidon space $\mathcal{V}$ given by Construction 1 admits the following efficient factoring algorithm. This algorithm fully uses the knowledge of the element γ such that $\mathcal{B} := \{1, \gamma\}$ is a basis of $\mathbb{F}_{q^{2k}}$ over $\mathbb{F}_{q^k}$, and for $x \in \mathbb{F}_{q^{2k}}$ we will denote by $[1](x)$ and $[\gamma](x)$ the components of x in this basis. Given a product $\pi = \pi_1\pi_2$ where π_1 and π_2 lie in $\mathcal{V}$, Algorithm 1 recovers π_1 and π_2 up to constant factors in $\mathbb{F}_q$.

```
Input: A product $\pi = \pi_1\pi_2$, where $\pi_1 = u + u^q\gamma$ and $\pi_2 = v + v^q\gamma \in \mathcal{V}$,
       the element $\gamma \in \mathbb{F}_{q^n}^*$ such that $\gamma^2 + b\gamma + c = 0$ from
       Construction 1.
Output: $\{\pi_1\mathbb{F}_q, \pi_2\mathbb{F}_q\}$.
Decompose $\pi$ in the basis $\{1, \gamma\}$:
  $q_0 \leftarrow [1](\pi)$                                  // $q_0 = uv - c(uv)^q$
  $q_1 \leftarrow [\gamma](\pi)$                             // $q_1 = uv^q + u^qv - b(uv)^q$
  $A \leftarrow T^{-1}(q_0)$            // where $T$ is map $x \mapsto x - cx^q$, $A = uv$
  $B \leftarrow q_1 + bA^q$                                  // $B = uv^q + u^qv$
Compute the roots $\alpha, \beta$ of $A + Bx + A^qx^2$
// $\alpha = -1/u^{q-1}$, $\beta = -1/v^{q-1}$
From $\alpha$ and $\beta$, recover $\{u\mathbb{F}_q, v\mathbb{F}_q\}$ uniquely and therefore $\{\pi_1\mathbb{F}_q, \pi_2\mathbb{F}_q\}$.
```

Algorithm 1: Factoring algorithm for Sidon space from Construction 1.

2.2 Description of the cryptosystem.

The Sidon cryptosystem relies on Construction 1, but it might be possible to consider another type of Sidon space such that $k = \Theta(n)$ for which an efficient factoring algorithm exists. In this section, we briefly describe the building blocks of the scheme, and we refer the reader to [21, §3] for further details.

[4] Such a polynomial is known to exist by [22, Corollary 14].

Keygen:

- Select a random element $\gamma \in \mathbb{F}_{q^n}$ satisfying the constraints given in Construction 1 in order to build the Sidon space $\mathcal{V} := \{u + u^q\gamma | u \in \mathbb{F}_{q^k}\}$.
- Select $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_k)$ a random basis of $\mathcal{V}$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$ a random basis of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$.
- Represent the matrix $\boldsymbol{M}(\boldsymbol{\nu}) = \boldsymbol{\nu}^\mathsf{T}\boldsymbol{\nu} \in \mathbb{F}_{q^n}^{k\times k}$ over the basis $\boldsymbol{\beta}$:

$$\boldsymbol{M}(\boldsymbol{\nu}) = \boldsymbol{\nu}^\mathsf{T}\boldsymbol{\nu} = \sum_{i=1}^{n} \beta_i \boldsymbol{M}^{(i)},$$

  where $\boldsymbol{M}^{(i)} \in \mathbb{F}_q^{k\times k}$ for $1 \leq i \leq n$.
- Output $\mathsf{sk} = (\boldsymbol{\beta}, \boldsymbol{\nu}, \gamma)$ as secret key and $\mathsf{pk} = (\boldsymbol{M}^{(1)}, \ldots, \boldsymbol{M}^{(n)})$ as public key.

As explained in the introduction, the message space correspond to the equivalence class of pairs of elements $\{a, b\}$ in the Sidon space $\mathcal{V}$, two pairs $\{a, b\}$ and $\{c, d\}$ being equivalent if their product is the same: $ab = cd$. If one views an element a of $\mathcal{V}$ as a vector $\boldsymbol{a} \in \mathbb{F}_q^k$, i.e. $a = \sum_{i=1}^{k} a_i\nu_i$, then the equivalence class associated to $\{\boldsymbol{a}, \boldsymbol{b}\}$ corresponds to all pairs $\{\boldsymbol{c}, \boldsymbol{d}\}$ such that either $\boldsymbol{a}^\mathsf{T}\boldsymbol{b} = \boldsymbol{c}^\mathsf{T}\boldsymbol{d}$ or $\boldsymbol{a}^\mathsf{T}\boldsymbol{b} = \boldsymbol{d}^\mathsf{T}\boldsymbol{c}$. This space is of size $\dfrac{(q^k-1)(q^k-q)}{2(q-1)} + q^k - 1$ as shown in [21, App A].

Encrypt($\{\boldsymbol{a}, \boldsymbol{b}\}$, $\mathsf{pk} = (\boldsymbol{M}^{(i)})_{i=1}^n$):

- The ciphertext associated to (the equivalence class of) $\{\boldsymbol{a}, \boldsymbol{b}\}$ is

$$\boldsymbol{c} = (c_i)_{i=1}^n = (\boldsymbol{a}\boldsymbol{M}^{(i)}\boldsymbol{b}^\mathsf{T})_{i=1}^n \in \mathbb{F}_q^n. \tag{2}$$

Note that this definition is compatible with the way the plaintext is defined: the ciphertext does not depend on the particular pair $\{\boldsymbol{a}, \boldsymbol{b}\}$ chosen in the equivalence class of the message. An interesting property of the Sidon cryptosystem is that it is homomorphic under the addition on half of the plaintext. That is, for two given plaintexts $\{\boldsymbol{a}_1, \boldsymbol{b}\}$ and $\{\boldsymbol{a}_2, \boldsymbol{b}\}$ we have

$$\textbf{Encrypt}(\{\boldsymbol{a}_1, \boldsymbol{b}\}, \mathsf{pk}) + \textbf{Encrypt}(\{\boldsymbol{a}_2, \boldsymbol{b}\}, \mathsf{pk}) = \textbf{Encrypt}(\{\boldsymbol{a}_1 + \boldsymbol{a}_2, \boldsymbol{b}\}, \mathsf{pk}).$$

To decrypt with the secret key, Bob views the ciphertext $\boldsymbol{c}$ as a product of elements in $\mathcal{V}$. Then, he is able to recover the factors using Algorithm 1 since he completely knows the structure of $\mathcal{V}$.

Decrypt($\boldsymbol{c}, \mathsf{sk} = (\boldsymbol{\beta}, \boldsymbol{\nu}, \gamma)$):

- Compute

$$\begin{aligned}\sum_{i=1}^{n} \beta_i c_i = \sum_{i=1}^{n} \beta_i \left(\boldsymbol{a}\boldsymbol{M}^{(i)}\boldsymbol{b}^\mathsf{T} \right) &= \boldsymbol{a}\boldsymbol{M}(\boldsymbol{\nu})\boldsymbol{b}^\mathsf{T} \\ &= \boldsymbol{a}\boldsymbol{\nu}^\mathsf{T}\boldsymbol{\nu}\boldsymbol{b}^\mathsf{T} = \left(\sum_{i=1}^{k} a_i \nu_i \right) \left(\sum_{i=1}^{k} b_i \nu_i \right) = ab, \end{aligned} \tag{3}$$

  and ab is a product of elements in $\mathcal{V}$.
- From the knowledge of γ, use Algorithm 1 to recover $\{a, b\}$ up to a multiplicative factor in $\mathbb{F}_q$.
- Finally, retrieve $\{\boldsymbol{a}, \boldsymbol{b}\}$ (up to a multiplicative factor) by representing $\{a, b\}$ over the basis $\boldsymbol{\nu}$. Such an $\{\boldsymbol{a}, \boldsymbol{b}\}$ defines the message in a unique way.

2.3 Equivalent keys for the Sidon cryptosystem.

An important notion for multivariate schemes in general is that of equivalent keys. Two secret keys are equivalent if they lead to the same public key. In the case of the Sidon cryptosystem, one can easily obtain the following result by using the definition of the decryption process given in Equation (3):

Fact 1 *Any Sidon space $\mathcal{V}'$ generated using Construction 1 with basis $\boldsymbol{\nu}'$ and such that the matrix $\boldsymbol{M}(\boldsymbol{\nu}')$ lies in the linear span of the $\boldsymbol{M}^{(i)}$'s can be used as an equivalent key.*

Equivalent keys are an important feature for our attack, since it will consist in recovering a Sidon space $\mathcal{V}' \neq \mathcal{V}$ which allows to decrypt any ciphertext.

3 Analysis of the underlying MinRank problem.

Given the public key $(\boldsymbol{M}^{(1)}, \dots, \boldsymbol{M}^{(n)})$ such that $\boldsymbol{M}^{(i)} \in \mathbb{F}_q^{k \times k}$ for $1 \leq i \leq n$, one has

$$\boldsymbol{M}(\boldsymbol{\nu}) = \boldsymbol{\nu}^\mathsf{T}\boldsymbol{\nu} = \sum_{i=1}^{n} \beta_i \boldsymbol{M}^{(i)},$$

where $(\beta_1, \dots, \beta_n)$ is the basis of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ and $\boldsymbol{\nu}$ is the basis of the Sidon space which are part of the private key. In other words, the matrix $\boldsymbol{M}(\boldsymbol{\nu})$ is a linear combination of the $\boldsymbol{M}^{(i)}$'s over $\mathbb{F}_{q^n}$ which has rank 1, and a key-recovery attack on the scheme requires to find specific solutions over $\mathbb{F}_{q^n}$ of the MinRank instance described by the $\boldsymbol{M}^{(i)}$'s. This MinRank problem is not standard in at least two ways. First, solutions are searched in the extension field $\mathbb{F}_{q^n}$, whereas the $\boldsymbol{M}^{(i)}$'s have their entries in $\mathbb{F}_q$. Second, this system has surprisingly many solutions and it is not clear at all whether the Sidon structure can be recovered from an arbitrary solution to it. Note that the authors of [21] only studied the hardness of finding any solution to this MinRank problem, and the task of determining the whole solution set was not addressed explicitly.

In this section, we examine in greater depth the properties of this solution set. A first remark is that the solutions correspond to rank 1 matrices in the space generated by the $\boldsymbol{M}^{(i)}$'s over $\mathbb{F}_{q^n}$:

$$\mathcal{C}_{mat} := \left\langle \boldsymbol{M}^{(1)}, \dots, \boldsymbol{M}^{(n)} \right\rangle_{\mathbb{F}_{q^n}}. \tag{4}$$

Our main result will be the existence of solutions over the subfield $\mathbb{F}_{q^k}$, *i.e.* rank 1 matrices in

$$\mathcal{D}_{mat} := \mathcal{C}_{mat}|_{\mathbb{F}_{q^k}} := \mathcal{C}_{mat} \cap \mathbb{F}_{q^k}^{k \times k}.$$

Fact 2 *The subspace $\mathcal{D}_{mat} = \mathcal{C}_{mat}|_{\mathbb{F}_{q^k}}$ contains elements of rank* 1.

Such elements are described in Section 3.3, and our experiments suggest that these are the only ones in $\mathcal{D}_{mat}$. A reader only interested in our key-recovery attack can directly go to Section 4 and Section 5.

3.1 Restricting the number of the solutions.

We start by reviewing some elementary properties of the solution set. Since the generators $\boldsymbol{M}^{(i)}$ are symmetric, all the elements in $\mathcal{C}_{mat}$ are symmetric as well. Therefore, rank 1 elements in $\mathcal{C}_{mat}$ will be of the form $\boldsymbol{x}^\mathsf{T}\boldsymbol{y} \in \mathbb{F}_{q^n}^{k \times k}$ for $\boldsymbol{x}$ collinear with $\boldsymbol{y}$. In particular, we will be interested in the following subset of solutions defined by

$$\mathcal{Z}_{\mathbb{F}_{q^n}} := \left\{ \boldsymbol{x} \in \mathbb{F}_{q^n}^k,\ \boldsymbol{x}^\mathsf{T}\boldsymbol{x} \in \mathcal{C}_{mat} \right\}. \tag{5}$$

This set is non-trivial since it contains $\boldsymbol{\nu}$ from the private key. Also, there is still one degree of freedom coming from the $\mathbb{F}_{q^n}$-linearity of $\mathcal{C}_{mat}$. For instance, since $\nu_1 \neq 0$ in $\boldsymbol{\nu}$, the set

$$\mathcal{Z}_{\mathbb{F}_{q^n},s} := \left\{ \boldsymbol{x} \in \mathcal{Z}_{\mathbb{F}_{q^n}}, x_1 = s \right\}$$

is also non-trivial for $s \in \mathbb{F}_{q^n}^*$.

3.2 Generic solutions over $\mathbb{F}_{q^n}$.

In this section, we describe a generic way to generate many solutions to the MinRank problem. Consider any $\boldsymbol{\omega} \in \mathbb{F}_{q^n}^k$ such that $\boldsymbol{\omega} \in \mathcal{Z}_{\mathbb{F}_{q^n}}$. By definition, there exists $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n) \in \mathbb{F}_{q^n}^n$ such that

$$\boldsymbol{M}(\boldsymbol{\omega}) = \sum_{\ell=1}^{n} \eta_\ell \boldsymbol{M}^{(\ell)}. \tag{6}$$

In particular, one has for $1 \leq i, j \leq k$:

$$\omega_i \omega_j = \sum_{\ell=1}^{n} \eta_\ell \boldsymbol{M}_{i,j}^{(\ell)}. \tag{7}$$

Then, by iterating the Frobenius map p times on this equation for $0 \leq p \leq n-1$, one obtains

$$\omega_i^{[p]}\omega_j^{[p]} = \sum_{\ell=1}^{n} \eta_\ell^{[p]} \boldsymbol{M}_{i,j}^{(\ell)} \tag{8}$$

since the matrices $\boldsymbol{M}^{(\ell)}$ have entries in $\mathbb{F}_q$. This implies that the matrix

$$\boldsymbol{M}(\boldsymbol{\omega}^{[p]}) = \sum_{\ell=1}^{n} \eta_\ell^{[p]} \boldsymbol{M}^{(\ell)} \tag{9}$$

belongs to $\mathcal{C}_{mat}$ for $0 \leq p \leq n-1$. Overall, this observation can be summarized in the following Lemma 1.

Lemma 1 ("Stability by Frobenius") *Let $\mathcal{C}_{mat}$ as defined in Equation (4) and let $\mathcal{Z}_{\mathbb{F}_{q^n}}$ as defined in Equation (4). If $\boldsymbol{\omega} \in \mathcal{Z}_{\mathbb{F}_{q^n}}$, then $\boldsymbol{\omega}^{[j]} \in \mathcal{Z}_{\mathbb{F}_{q^n}}$ for any $j \geq 0$, and more generally if $\boldsymbol{M} \in \mathcal{C}_{mat}$, then $\boldsymbol{M}^{[j]} \in \mathcal{C}_{mat}$ for any $j \geq 0$.*

The fact that $\mathcal{V}$ is a Sidon space is not used at all in this reasoning. In particular, the very same argument can be applied to a random subspace $\mathcal{W} \subset \mathbb{F}_{q^n}$ of dimension k along with a random secret basis $\boldsymbol{\omega}$ for $\mathcal{W}$. In this case, the $\boldsymbol{M}^{(i)}$'s are obtained from the decomposition of the matrix $\boldsymbol{M}(\boldsymbol{\omega})$ in an arbitrary basis of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$, and $\mathcal{C}_{mat}$ and $\mathcal{Z}_{\mathbb{F}_{q^n}}$ are defined in the same way as before. For such a random subspace $\mathcal{W}$, the only solutions to the MinRank instance that we observe in practice are given by Lemma 1:

Observation 1 (From experiments) *Let $\mathcal{W} \subset \mathbb{F}_{q^n}$ be a random $\mathbb{F}_q$-subspace of dimension k, let $\boldsymbol{\omega}$ be a random basis of $\mathcal{W}$ and let $s \in \mathbb{F}_{q^n}^*$. One has*

$$\left|\mathcal{Z}_{\mathbb{F}_{q^n},s}\right| = n.$$

Moreover, if $s \in \mathbb{F}_q^*$, then there exists $\boldsymbol{u} \in \mathbb{F}_{q^n}^k$ with $u_1 = s$ such that

$$\mathcal{Z}_{\mathbb{F}_{q^n},s} = \left\{\boldsymbol{u}, \boldsymbol{u}^{[1]}, \ldots, \boldsymbol{u}^{[n-1]}\right\}.$$

However, when $\mathcal{V}$ is a Sidon space generated using Construction 1, we observed that there were many more solutions to the MinRank instance than those just described. This behavior can be explained by the fact that there also exist rank 1 linear combinations of the $\boldsymbol{M}^{(i)}$'s over the subfield $\mathbb{F}_{q^k}$, *i.e.* the set $\mathcal{Z}_{\mathbb{F}_{q^k}}$ is non-trivial. More precisely, we obtained the following experimental result.

Observation 2 (From experiments) *Let $\mathcal{V}$ be a Sidon space generated using Construction 1. For $s \in \mathbb{F}_{q^k}^*$, we observed that*

$$\left|\mathcal{Z}_{\mathbb{F}_{q^k},s}\right| = k(q^k - 1).$$

Moreover, if $t \in \mathbb{F}_{q^k}^*$, $t \notin \langle s \rangle_{\mathbb{F}_q}$, we observed that

$$\left|\left\{\boldsymbol{x} \in \mathcal{Z}_{\mathbb{F}_{q^k},s},\ x_2 = t\right\}\right| = k.$$

3.3 Rank 1 codewords in $\mathcal{D}_{mat}$ from the Sidon structure.

This section is dedicated to the study of the set $\mathcal{Z}_{\mathbb{F}_{q^k}}$ when $\mathcal{V}$ is a Sidon space generated using Construction 1 with secret basis $\boldsymbol{\nu}$. For $1 \leq i \leq k$, there exists $u_i \in \mathbb{F}_{q^k}$ such that

$$\nu_i = u_i + u_i^q \gamma. \tag{10}$$

Note that $\boldsymbol{u} := (u_1, \dots, u_k)$ is necessarily a basis of $\mathbb{F}_{q^k}$ over $\mathbb{F}_q$. In the following Proposition 1, we are interested in the rank 1 matrix

$$\boldsymbol{M}(\boldsymbol{u}) = (u_1, \dots, u_k)^\top (u_1, \dots, u_k) \in \mathbb{F}_{q^k}^{k \times k}.$$

Proposition 1 *Let $\boldsymbol{M}^{(1)}, \dots, \boldsymbol{M}^{(n)}$ be the public matrices associated to the secret Sidon space $\mathcal{V}$ and let $\boldsymbol{u}$ be the basis of $\mathbb{F}_{q^k}$ over $\mathbb{F}_q$ associated to $\boldsymbol{\nu}$ by Equation* (10)*. Then, $\boldsymbol{M}(\boldsymbol{u})$ is a rank* 1 *matrix in $\mathcal{D}_{mat}$. Moreover, the same is true for $\boldsymbol{M}(\boldsymbol{u}^{[j]})$ for $0 \leq j \leq k-1$.*

Proof. We do the proof for $\boldsymbol{M}(\boldsymbol{u}) = \boldsymbol{M}(\boldsymbol{u}^{[0]})$ and the rest easily follows by using Lemma 1. First, one can write $\boldsymbol{M}(\boldsymbol{\nu})$ as a sum

$$\boldsymbol{M}(\boldsymbol{\nu}) = \boldsymbol{A} + \gamma \boldsymbol{B} \tag{11}$$

where the matrices $\boldsymbol{A},\ \boldsymbol{B} \in \mathbb{F}_{q^k}^{k \times k}$ are such that

$$\begin{cases} \boldsymbol{A} = \displaystyle\sum_{i=1}^{n} \delta_i \boldsymbol{M}^{(i)} \\ \boldsymbol{B} = \displaystyle\sum_{i=1}^{n} \eta_i \boldsymbol{M}^{(i)} \end{cases} \tag{12}$$

and where $\beta_i := \delta_i + \gamma \eta_i$ is expressed in the basis $\{1, \gamma\}$ for $1 \leq i \leq n$ with $\delta_i, \eta_i \in \mathbb{F}_{q^k}$. Also, recall that the primitive element γ is a root of the irreducible polynomial $x^2 + bx + c$ over $\mathbb{F}_{q^k}$, so that one obtains for $1 \leq i, j \leq k$:

$$\begin{aligned} \nu_i \nu_j &= (u_i + u_i^q \gamma)(u_j + u_j^q \gamma) \\ &= (u_i u_j - c(u_i u_j)^q) + \gamma(u_i u_j^q + u_i^q u_j - b(u_i u_j)^q). \end{aligned} \tag{13}$$

Therefore, Equation (13) shows that $\boldsymbol{A} = \boldsymbol{M}(\boldsymbol{u}) - c\boldsymbol{M}(\boldsymbol{u}^{[1]})$, and this matrix belongs to $\mathcal{D}_{mat}$ by (12). By Lemma 1, the same is true for the following matrices

$$\begin{aligned} \boldsymbol{A}^{[1]} &= \boldsymbol{M}(\boldsymbol{u}^{[1]}) - c^q \boldsymbol{M}(\boldsymbol{u}^{[2]}) \\ \boldsymbol{A}^{[2]} &= \boldsymbol{M}(\boldsymbol{u}^{[2]}) - c^{q^2} \boldsymbol{M}(\boldsymbol{u}^{[3]}) \\ &\vdots \\ \boldsymbol{A}^{[k-1]} &= \boldsymbol{M}(\boldsymbol{u}^{[k-1]}) - c^{q^{k-1}} \boldsymbol{M}(\boldsymbol{u}^{[k]}) = \boldsymbol{M}(\boldsymbol{u}^{[k-1]}) - c^{q^{k-1}} \boldsymbol{M}(\boldsymbol{u}). \end{aligned}$$

Then, by performing linear combinations over $\mathbb{F}_{q^k}$, one gets

$$\boldsymbol{A} + \sum_{i=1}^{k-1} c^{1+q+\cdots+q^{i-1}} \boldsymbol{A}^{[i]} = (1 - c^{1+q+\cdots+q^{k-1}})\boldsymbol{M}(\boldsymbol{u}) = (1 - c^{\frac{q^k-1}{q-1}})\boldsymbol{M}(\boldsymbol{u}).$$

Finally, one has $c^{\frac{q^k-1}{q-1}} \neq 1$ since $c \in \overline{\mathcal{W}}_{q-1}$, and therefore the matrix $\boldsymbol{M}(\boldsymbol{u})$ can be expressed as a linear combination of the $\boldsymbol{A}^{[i]}$'s over $\mathbb{F}_{q^k}$. This proves $\boldsymbol{M}(\boldsymbol{u}) \in \mathcal{D}_{mat}$. □

Also, note that it is easy to find other rank 1 matrices in $\mathcal{D}_{mat}$. Indeed, by using the second term in the right hand side of Equation (13), one notices that the matrix $\boldsymbol{B} \in \mathbb{F}_q^{k\times k}$ defined in the proof of Proposition 1 by (11) satisfies $\boldsymbol{B}_{i,j} = u_i u_j^q + u_i^q u_j - b(u_i u_j)^q$ for $1 \leq i, j \leq k$, and this matrix also belongs to $\mathcal{D}_{mat}$ by (12). Notice that this equality implies that

$$\begin{aligned}\boldsymbol{B} &= \boldsymbol{u}^\intercal \boldsymbol{u}^{[1]} + \left(\boldsymbol{u}^{[1]}\right)^\intercal \boldsymbol{u} - b\left(\boldsymbol{u}^{[1]}\right)^\intercal \boldsymbol{u}^{[1]} \\ &= \boldsymbol{u}^\intercal \boldsymbol{u}^{[1]} + \left(\boldsymbol{u}^{[1]}\right)^\intercal \boldsymbol{u} - b\boldsymbol{M}(\boldsymbol{u}^{[1]}). \qquad (14)\end{aligned}$$

Now, let $\lambda \in \mathbb{F}_{q^k}$ and consider

$$\begin{aligned}\boldsymbol{M}(\boldsymbol{u} + \lambda\boldsymbol{u}^{[1]}) &= \left(\boldsymbol{u} + \lambda\boldsymbol{u}^{[1]}\right)^\intercal \left(\boldsymbol{u} + \lambda\boldsymbol{u}^{[1]}\right) \\ &= \boldsymbol{u}^\intercal \boldsymbol{u} + \lambda^2 \left(\boldsymbol{u}^{[1]}\right)^\intercal \boldsymbol{u}^{[1]} + \lambda\left\{\boldsymbol{u}^\intercal \boldsymbol{u}^{[1]} + \left(\boldsymbol{u}^{[1]}\right)^\intercal \boldsymbol{u}\right\} \\ &= \boldsymbol{M}(\boldsymbol{u}) + \lambda^2 \boldsymbol{M}(\boldsymbol{u}^{[1]}) + \lambda\boldsymbol{B} + \lambda b\boldsymbol{M}(\boldsymbol{u}^{[1]}) \quad \text{(by (14))}.\end{aligned}$$

This implies that $\boldsymbol{M}(\boldsymbol{u} + \lambda\boldsymbol{u}^{[1]})$ belongs to $\mathcal{D}_{mat}$. Since $\boldsymbol{M}(\boldsymbol{u} + \lambda\boldsymbol{u}^{[1]})$ is of rank 1, we have therefore proved the following generalization of Proposition 1.

Proposition 2 *Let $\boldsymbol{M}^{(1)}, \ldots, \boldsymbol{M}^{(n)}$ be the public matrices associated to the secret Sidon space $\mathcal{V}$ and let $\boldsymbol{u}$ be the basis of $\mathbb{F}_{q^k}$ over $\mathbb{F}_q$ associated to $\boldsymbol{\nu}$ by Equation* (10)*. One has*

$$\left\{\lambda\boldsymbol{u}^{[j]} + \mu\boldsymbol{u}^{[j+1]} : (\lambda, \mu) \in \mathbb{F}_{q^k}^2,\ 0 \leq j \leq k-1\right\} \subset \mathcal{Z}_{\mathbb{F}_{q^k}}.$$

Finally, note that Observation 2 suggests that the inclusion given in Proposition 2 is an equality. Combined with the content of Section 3.1, this gives a complete understanding of the set of rank 1 codewords in $\mathcal{D}_{mat}$.

In the next sections, we are going to describe our key-recovery attack which builds upon the following facts:

- As pointed out by [21], it seems infeasible to recover the Sidon space $\mathcal{V}$ directly as a solution to the MinRank problem.

– However, from Fact 1, it is possible to decrypt by using a different Sidon space. Moreover, contrary to the approach described in [21, §5.1] which introduces an algebraic system with too many variables to be solved in practice, we will find such a space efficiently. We will then show how this can be exploited to find sufficiently many elements of $\mathcal{Z}_{\mathbb{F}_{q^k}}$ that will be used to recover an equivalent key.

4 Solving the MinRank instance over $\mathbb{F}_{q^k}$.

In this section, we show how to determine elements in the set

$$\mathcal{Z}_{\mathbb{F}_{q^k}} := \left\{ \boldsymbol{x} \in \mathbb{F}_{q^k}^k,\ \boldsymbol{x}^\mathsf{T}\boldsymbol{x} \in \mathcal{C}_{mat} \right\} = \left\{ \boldsymbol{x} \in \mathbb{F}_{q^k}^k,\ \boldsymbol{x}^\mathsf{T}\boldsymbol{x} \in \mathcal{D}_{mat} \right\},$$

which corresponds to particular solutions to the MinRank instance described in the previous section. They will be exploited in Section 5 in order to derive an equivalent key.

4.1 Parity-check modeling.

Rather than using the generic techniques described in [21, §4] to target elements in $\mathcal{Z}_{\mathbb{F}_{q^k}}$, we found experimentally that it was more favorable to consider the following algebraic modeling which is largely inspired by [10, §5.4]. A first remark is that a square matrix of size k over $\mathbb{F}_{q^k}$ can also be viewed as a vector of length k^2 over $\mathbb{F}_{q^k}$. To make this correspondence explicit, we will use the linear isomorphism

$$\begin{aligned} \text{vec} :\ \mathbb{F}_{q^k}^{k\times k} &\to \mathbb{F}_{q^k}^{k^2} \\ \boldsymbol{M} &\mapsto \boldsymbol{m} \end{aligned}$$

such that $\boldsymbol{m}_{(i-1)k+j} = \boldsymbol{M}_{i,j}$ for $1 \le i, j \le k$, and we consider the following subspace of $\mathbb{F}_{q^k}^{k^2}$

$$\text{vec}(\mathcal{D}_{mat}) := \{\text{vec}(\boldsymbol{M}),\ \boldsymbol{M} \in \mathcal{D}_{mat}\}.$$

It is a subspace of dimension n over $\mathbb{F}_{q^k}$. We can consider it as linear code and let $\boldsymbol{H} \in \mathbb{F}_{q^k}^{(k^2-n)\times k^2}$ be an arbitrary parity-check matrix for it, *i.e.* a matrix such that

$$\text{vec}(\mathcal{D}_{mat}) = \left\{ \boldsymbol{x} \in \mathbb{F}_{q^k}^{k^2},\ \boldsymbol{H}\boldsymbol{x}^\mathsf{T} = 0 \right\}.$$

Note that since the $\boldsymbol{M}^{(i)}$'s have entries in $\mathbb{F}_q$, it is possible to choose a parity-check matrix whose entries lie in $\mathbb{F}_q$ as well. Finally, let

$$\boldsymbol{X} := \boldsymbol{x}^\mathsf{T}\boldsymbol{x} = \begin{pmatrix} x_1^2 & x_1x_2 & \cdots & x_1x_k \\ x_2x_1 & x_2^2 & \cdots & x_2x_k \\ \vdots & \vdots & \ddots & \vdots \\ x_kx_1 & x_kx_2 & \cdots & x_k^2 \end{pmatrix} \tag{15}$$

be a matrix in the unknowns x_i corresponding to a solution $\boldsymbol{x} \in \mathcal{Z}_{\mathbb{F}_{q^k}}$. Since the vector $\mathrm{vec}(\boldsymbol{X})$ belongs to $\mathrm{vec}(\mathcal{D}_{mat})$, one obtains the following system of k^2-2k quadratic equations given by

$$\boldsymbol{H}\ \mathrm{vec}(\boldsymbol{X})^{\mathsf{T}} = 0. \tag{16}$$

Lemma 2 *The sequence given by Equation* (16) *contains at most* $k^2-2k-\binom{k}{2}$ *linearly independent quadratic polynomials over* $\mathbb{F}_{q^k}$.

Proof. Let $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{k^2})$ be the canonical basis of $\mathbb{F}_{q^k}^{k^2}$. Owing to the symmetry of the $\boldsymbol{M}^{(i)}$'s, one obtains that for $1 \leq i < j \leq k$, the vector

$$\sigma_{i,j} = \boldsymbol{e}_{(i-1)k+j} - \boldsymbol{e}_{(j-1)k+i}$$

belongs to the dual code $\mathrm{vec}(\mathcal{D}_{mat})^{\perp}$. Therefore, there exists a parity-check matrix for $\mathrm{vec}(\mathcal{D}_{mat})$ of the form

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{U} \\ \boldsymbol{H}_\sigma \end{pmatrix},$$

where the rows of $\boldsymbol{H}_\sigma \in \mathbb{F}_{q^k}^{\binom{k}{2}\times k^2}$ are the $\sigma_{i,j}$ and $\boldsymbol{U} \in \mathbb{F}_{q^k}^{(k^2-2k-\binom{k}{2})\times k^2}$. The equations coming from $\boldsymbol{H}_\sigma \mathrm{vec}(\boldsymbol{X})^{\mathsf{T}} = 0$ all give the zero polynomial. Therefore, the useful part of the system is given by

$$\boldsymbol{U}\mathrm{vec}(\boldsymbol{X})^{\mathsf{T}} = 0,$$

and it contains $k^2-2k-\binom{k}{2}$ equations. □

Modeling 1 (Parity-check modeling over $\mathbb{F}_{q^k}$) *Let* $\boldsymbol{X}$ *be the matrix of unknowns defined in Equation* (15), *let* $n = 2k$ *and let* $\boldsymbol{H} = \begin{pmatrix} \boldsymbol{U} \\ \boldsymbol{H}_\sigma \end{pmatrix} \in \mathbb{F}_{q^k}^{(k^2-n)\times k^2}$ *be a parity-check matrix for the code* $vec(\mathcal{D}_{mat})$ *as described in the proof of Lemma 2, where* $\mathcal{D}_{mat} = \mathcal{C}_{mat} \cap \mathbb{F}_{q^k}^{k\times k}$. *We consider the system* $\mathcal{F}$ *over* $\mathbb{F}_{q^k}$ *given by*

$$\boldsymbol{U}\mathit{vec}(\boldsymbol{X})^{\mathsf{T}} = 0. \tag{17}$$

This system contains $k^2-n-\binom{k}{2}$ *quadratic equations in the* x_i *variables.*

It is readily verified that the solutions to Modeling 1 are in one-to-one correspondence with the elements of $\mathcal{Z}_{\mathbb{F}_{q^k}}$. Experimentally, these solutions were also all of the form described in Proposition 2. If one wants to find an element in $\mathcal{Z}_{\mathbb{F}_{q^k}}$ in practice, two variables must be fixed in Modeling 1 to obtain a zero-dimensional ideal. The corresponding variety over $\mathbb{F}_{q^k}$ has size $\geq k$ still by using Proposition 2, and experimentally this was always an equality.

Modeling 2 (Recovering an element in $\mathcal{Z}_{\mathbb{F}_{q^k}}$) *Let $(s,t) \in \mathbb{F}_{q^k}^2$ such that $t \notin \langle s \rangle_{\mathbb{F}_q}$. We consider the system $\mathcal{F}_{spec}$ obtained by fixing $x_{k-1} = s$ and $x_k = t$ in the equations of the sequence $\mathcal{F}$ from Modeling 1.*

The approach that we use to solve this system $\mathcal{F}_{spec}$ is standard. We start by computing a Gröbner basis for a suitable ordering and then we perform a change of order step to deduce a basis for the lexicographic ordering to get the solutions. Using Proposition 2, we expect k distinct solutions to Modeling 2, and therefore the complexity of this second step is polynomial in k by using the so-called FGLM algorithm [13]. In the following, we will focus on the complexity of the first step.

4.2 Complexity of solving the system $\mathcal{F}_{spec}$.

In this section, we show that, under the following Assumption 1 and Assumption 2, the system $\mathcal{F}_{spec}$ can always be solved at degree 3 independently from the value of k.

First, note that one can permute the coordinates of the row-vector $\mathrm{vec}(\boldsymbol{X})$ and the columns of $\boldsymbol{U}$ accordingly so that the $\binom{k+1}{2}$ leftmost entries of $\mathrm{vec}(\boldsymbol{X})$ correspond to all the distinct monomials $x_i x_j$ for $1 \leq i \leq j \leq k$. This is equivalent to choosing a grevlex ordering on the x_i variables to label the columns of $\boldsymbol{U}$. Also, by adding rows of $\boldsymbol{H}_\sigma$ to rows of $\boldsymbol{U}$ in $\boldsymbol{H}$, it is always possible to assume that the last $\binom{k}{2}$ columns of the matrix $\boldsymbol{U}$ are identically zero.

Assumption 1 *We assume that $\boldsymbol{U}$ is full-rank, and moreover we assume that the submatrix $\boldsymbol{U}_{*,\{1..\binom{k-1}{2}\}}$ is also full-rank.*

The first part of Assumption 1 ensures that one can find $\binom{k-1}{2}$ distinct leading monomials of the form $x_i x_j$ for $1 \leq i \leq j \leq k$ in the $\mathbb{F}_q$-span of the polynomials $\mathcal{F}$, and a fortiori the equations from Modeling 1 are linearly independent. The second part is a bit stronger: it implies that these leading monomials will not involve x_{k-1} or x_k. Therefore, by doing linear combinations between the equations from the specialized system $\mathcal{F}_{spec}$, one can obtain a set $\mathcal{G}_{spec}$ of $m := k^2 - 2k - \binom{k}{2} = \frac{k^2-3k}{2}$ equations $g_1 = 0, \cdots, g_m = 0$ with distinct leading monomials $x_i x_j$ for $1 \leq i, j \leq k-2$. Since the total number of quadratic monomials of this form is equal to

$$\binom{k-1}{2} = \frac{(k-1)(k-2)}{2} = \frac{k^2-3k}{2} + 1,$$

this implies that all monomials of degree 2 appear as leading terms of the g_i's but one. With this assumption it can be proved that computing the Gröbner basis of the algebraic system $\mathcal{F}_{spec}$ is extremely efficient: essentially it amounts to compute the aforementioned echelonized set of quadratic polynomials $\mathcal{G}_{spec}$, and then the Gröbner basis is either already computed or close to be computed. This is easily verified by making the further assumption that

Assumption 2 *The algebraic system $\mathcal{F}_{spec}$ has exactly k distinct solutions which do not belong to a common hyperplane of $\mathbb{F}_{q^k}^{k-2}$.*

This assumption was satisfied in all our experiments and is natural when considering Proposition 2 together with Observation 2, which suggests that the inclusion given in this Proposition is an equality. Indeed, the form of the solutions we get from this Proposition then suggests that Assumption 2 should typically hold.

Buchberger's algorithm for computing a Gröbner basis from $\mathcal{G}_{spec} = \{g_1 = 0, \cdots, g_m = 0\}$ would start by computing the S-polynomials $S(g_i, g_j)$ and reduce them. There are two cases to consider.

Case 1. The missing leading monomial in the g_i's is of the form $x_i x_j$. Note that the only case where $S(g_i, g_j)$ were not reduced to 0 would be when the leading monomials of g_i and g_j have a common factor (see [9, Prop. 4, p.106]). In such a case, the polynomial $S(g_i, g_j)$ is of degree at most 3 and since in our situation
(i) all the monomials of degree 3 appear as multiples of leading monomials of the g_i's,
(ii) all monomials of degree 2 appear as leading monomials in the g_i's but $x_i x_j$, this implies that $S(g_i, g_j)$ is reduced to a polynomial of the form $g_{m+1} := \mu x_i x_j + L(\boldsymbol{x})$ where L is affine in the x_i's. It is impossible that $\mu = 0$ and $L \neq 0$ since this would imply that all the k solutions to $\mathcal{F}_{spec}$ lie in the affine hyperplane $L(\boldsymbol{x}) = 0$, which contradicts Assumption 2. If $\mu \neq 0$, then it is clear by performing the same reasoning that all S-polynomials $S(g_{m+1}, g_i)$ would reduce to 0 (since they would this time reduce to affine forms which are necessarily 0 by the previous reasoning). We are therefore left with a Gröbner basis.

Case 2. The missing leading monomial in the g_i's is of the form x_i^2. The difference with the previous case is that all degree 3 monomials appear as multiples of leading monomials of the g_i's with the exception of x_i^3. In such a case, $S(g_i, g_j)$ reduces to a polynomial of the form $g_{m+1} := \lambda x_i^3 + \mu x_i^2 + L(\boldsymbol{x})$ where L is again an affine form. It is readily seen that we can not have $\lambda = \mu = 0$ without that $L = 0$ itself (this would contradict in the same way as before Assumption 2). From this, it is readily seen that all S-polynomials $S(g_{m+1}, g_j)$ reduce to 0 and that we have a Gröbner basis again.

Remark 1 *Actually the first part of Assumption 2 is already enough to prove this kind of behavior for the Gröbner basis computation by using the fact that the number of solutions for $\mathcal{G}_{spec}$ is equal to the number of monomials that can not be leading monomials of an element of the ideal generated by the g_i's (this is essentially a corollary of [14, Cor. 5, p.83]). We have avoided to use this result to keep the proof as simple as possible. The constant monomial is an example of such a kind (because $\mathcal{G}_{spec}$ has solutions), there are at most $k-2$ monomials of degree 1, at most one monomial of degree 2 and at most one monomial of degree 3 of such kind. From this, it is for instance straightforward to rule out the possibility that $g_{m+1} \neq 0$ in Case 1.*

Overall, one needs to go up to degree 3 in the worst case to compute the Gröbner basis for $\mathcal{F}_{spec}$. The final complexity is then dominated by that of performing Gaussian elimination at degree 3 on a matrix of size $A \times B$ with $A \leq B := \binom{k-2+3}{3}$, say

$$\mathcal{O}\left(\binom{k-2+3}{3}^{\omega}\right) \tag{18}$$

operations in $\mathbb{F}_q$, where $2 \leq \omega \leq 3$ is the linear algebra constant. Therefore, the complexity of solving the system is in $\mathcal{O}\left(k^{3\omega}\right)$, which is clearly polynomial in the dimension k of the Sidon space.

5 Finding an equivalent Sidon space $\mathcal{V}$'.

Even if recovering elements in $\mathcal{Z}_{\mathbb{F}_{q^k}}$ can be performed in an efficient way, it remains to explain how this leads to a key-recovery attack. In particular, we have yet to show how we obtain from those elements an equivalent key. We will prove here that we obtain from a set of $k+1$ elements $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_{k+1}$ in $\mathcal{Z}_{\mathbb{F}_{q^k}}$ a Sidon space $\mathcal{V}'$ obtained by Construction 1 that meets the criterion of Fact 1, namely that there is an ordered basis $\boldsymbol{\nu}'$ for it such that $\boldsymbol{M}(\boldsymbol{\nu}')$ lies in the space spanned by the $\boldsymbol{M}^{(i)}$'s. This procedure consists in

1. From $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_{k+1}$ in $\mathcal{Z}_{\mathbb{F}_{q^k}}$ we recover $\boldsymbol{t} = \lambda \boldsymbol{u}^{[j]}$ for some λ in $\mathbb{F}_{q^k}$ and j in $\{0, \cdots, k-1\}$ where $\boldsymbol{u} = (u_1, \cdots, u_k)$ is defined from the secret basis $\boldsymbol{\nu}$ of the Sidon space of the scheme by (10).
2. From such a $\boldsymbol{t}$, we deduce the aforementioned Sidon space $\mathcal{V}'$ as

   $$\mathcal{V}' = \langle t_1 + \gamma' t_1^q, \cdots, t_k + \gamma' t_k^q \rangle_{\mathbb{F}_q},$$

   where $(t_1, \cdots, t_k) = \boldsymbol{t}$ and γ' is an element generated like γ in **Keygen**, namely as a root of an irreducible polynomial $x^2 + ex + f$ over $\mathbb{F}_{q^k}$ such that $f \in \overline{W_{q-1}}$.

5.1 Targeting an element of the form $\boldsymbol{\lambda u}^{[j]}$.

Assuming that the inclusion in Proposition 2 is an equality, one obtains that the set $\mathcal{Z}_{\mathbb{F}_{q^k}}$ is equal to the union of vector spaces

$$\mathcal{Z}_{\mathbb{F}_{q^k}} = \bigcup_{i=1}^{k} \mathcal{W}_i, \quad \text{where } \mathcal{W}_i := \left\langle \boldsymbol{u}^{[i-1]}, \boldsymbol{u}^{[i]} \right\rangle_{\mathbb{F}_{q^k}}.$$

Let us notice that these vector spaces $\mathcal{W}_i$ satisfy a peculiar property, namely that

$$\mathcal{W}_i \cap \mathcal{W}_i^{[1]} = \left\langle \boldsymbol{u}^{[i]} \right\rangle \tag{19}$$

where for a set S of vectors, $S^{[1]}$ stands for the set $\{\boldsymbol{x}^{[1]} : \boldsymbol{x} \in S\}$. (19) follows from the fact that $\mathcal{W}_i^{[1]}$ is the $\mathbb{F}_{q^k}$-vector space generated by $\boldsymbol{u}^{[i]}$ and $\boldsymbol{u}^{[i+1]}$. In other words, we are able to recover one of the $\boldsymbol{u}^{[i]}$'s up to multiplication by an element of $\mathbb{F}_{q^k}$ if we are able to produce one of those $\mathcal{W}_i$'s. This can be achieved by using the pigeonhole principle: two among the solutions $\boldsymbol{t}_i$ for $1 \leq i \leq k+1$ will fall into a same vector space $\mathcal{W}_{j_0}$. These considerations lead to the following procedure for recovering one of those $\boldsymbol{u}^{[i]}$'s (up to a multiplicative constant)

```
Input: A set of $k+1$ non-collinear vectors $\boldsymbol{t}_1, \dots, \boldsymbol{t}_{k+1}$ in $\mathcal{Z}_{\mathbb{F}_{q^k}}$.
Output: A set $\mathcal{S}$ of elements containing at least one element collinear
        with one of the $\boldsymbol{u}^{[i]}$'s.
for $i = 1$ to $k$ do
    for $j = i$ to $k+1$ do
        $V \leftarrow \langle \boldsymbol{t}_i, \boldsymbol{t}_j \rangle_{\mathbb{F}_{q^k}}$
        if $\dim V \cap V^{[1]} = 1$ then
            $\mathcal{S} \leftarrow \mathcal{S} \cup \{\boldsymbol{x}\}$        // where $\boldsymbol{x}$ generates $V \cap V^{[1]}$
        end
    end
end
```

This algorithm is of complexity $\mathcal{O}(k^2)$ and it remains now just to explain how from one of those elements of $\mathcal{S}$ which is collinear with a $\boldsymbol{u}^{[i]}$ we are able to produce an equivalent key for the Sidon cryptosystem. Notice that we do not even need to have $k+1$ non-collinear vectors $\boldsymbol{t}_1, \dots, \boldsymbol{t}_{k+1}$ in $\mathcal{Z}_{\mathbb{F}_{q^k}}$, $\Theta(\sqrt{k})$ vectors are indeed sufficient by using the birthday paradox to get an $\mathcal{S}$ containing an element $\boldsymbol{t}$ collinear with some $\boldsymbol{u}^{[i]}$ with probability $\Omega(1)$.

5.2 Deducing $\mathcal{V}'$ from $\boldsymbol{t}$

How a Sidon space $\mathcal{V}'$ with the right properties can be deduced from $\boldsymbol{t}$ collinear with some $\boldsymbol{u}^{[i]}$ is explained by the following proposition.

Proposition 3 *Let γ' be a root of an irreducible polynomial $x^2 + ex + f$ over $\mathbb{F}_{q^k}$ such that $f \in \overline{W_{q-1}}$. Then, the $\mathbb{F}_q$-linear space $\mathcal{V}'$ generated by the ordered basis $\boldsymbol{v}' := \boldsymbol{t} + \gamma' \boldsymbol{t}^{[1]}$ is a Sidon space $\mathcal{V}'$ such that $\boldsymbol{M}(\boldsymbol{\nu}')$ is spanned by the $\boldsymbol{M}_i$'s.*

Proof. We have

$$\begin{aligned}
\boldsymbol{M}(\boldsymbol{\nu}') &= \boldsymbol{M}(\boldsymbol{t} + \gamma' \boldsymbol{t}^{[1]}) \\
&= \boldsymbol{t}^\intercal \boldsymbol{t} + {\gamma'}^2 \left(\boldsymbol{t}^{[1]}\right)^\intercal \boldsymbol{t}^{[1]} + \gamma' \boldsymbol{t}^\intercal \boldsymbol{t}^{[1]} + \gamma' \left(\boldsymbol{t}^{[1]}\right)^\intercal \boldsymbol{t} \\
&= \lambda^2 \boldsymbol{u}^\intercal \boldsymbol{u} + \lambda^{2q} {\gamma'}^2 \left(\boldsymbol{u}^{[1]}\right)^\intercal \boldsymbol{u}^{[1]} + \lambda^{1+q} \gamma' \left\{ \boldsymbol{u}^\intercal \boldsymbol{u}^{[1]} + \left(\boldsymbol{u}^{[1]}\right)^\intercal \boldsymbol{u} \right\} \quad \text{(since } \boldsymbol{t} = \lambda \boldsymbol{u} \text{ for } \lambda \in \mathbb{F}_{q^k}) \\
&= \lambda^2 \boldsymbol{M}(\boldsymbol{u}) + \lambda^{2q} {\gamma'}^2 \boldsymbol{M}(\boldsymbol{u}^{[1]}) + \lambda^{1+q} \gamma' \left\{ (\boldsymbol{u} + \boldsymbol{u}^{[1]})^\intercal (\boldsymbol{u} + \boldsymbol{u}^{[1]}) - \boldsymbol{u}^\intercal \boldsymbol{u} - \left(\boldsymbol{u}^{[1]}\right)^\intercal \boldsymbol{u}^{[1]} \right\} \\
&= \lambda^2 \boldsymbol{M}(\boldsymbol{u}) + \lambda^{2q} {\gamma'}^2 \boldsymbol{M}(\boldsymbol{u}^{[1]}) + \lambda^{1+q} \gamma' \left\{ \boldsymbol{M}(\boldsymbol{u} + \boldsymbol{u}^{[1]}) - \boldsymbol{M}(\boldsymbol{u}) - \boldsymbol{M}(\boldsymbol{u}^{[1]}) \right\} \\
&\in \left\langle \boldsymbol{M}^{(1)}, \cdots, \boldsymbol{M}^{(n)} \right\rangle_{\mathbb{F}_{q^n}} \quad \text{(by Proposition 2).}
\end{aligned}$$

□

In other words for finding $\mathcal{V}'$, we just have to

1. find an element γ' satisfying the same constraints as γ, i.e. γ' is a root of an irreducible polynomial $x^2 + ex + f$ over $\mathbb{F}_{q^k}$ such that $f \in \overline{W_{q-1}}$;
2. $\mathcal{V}'$ is then generated by the basis

$$\boldsymbol{\nu}' = \{t_1 + \gamma' t_1^q, \dots, t_k + \gamma' t_k^q\}$$

and leads to an equivalent key by Fact 1.

Note that Step 1. for finding $\gamma' \in \mathbb{F}_{q^n}$ can be performed in the same way as in **Keygen**. This was done at random in [21], and the success probability can be estimated using [22, Lemma 13]. Heuristically, this works in constant expected time.

6 Conclusion.

The use of Sidon spaces for cryptography is an interesting new idea initially proposed in [21]. However, in this paper we show that this first attempt to build a public-key encryption scheme based on Sidon spaces is insecure. Here, we develop a key-recovery attack which is polynomial in the dimension of the underlying Sidon space. Besides of that, we consider worth to further study the possibility of using Sidon spaces to devise other cryptographic primitives.

Acknowledgements

The authors would like to thank Magali Bardet for her valuable comments to improve the quality of this paper.

References

1. Daniel Apon, Dustin Moody, Ray Perlner, Daniel Smith-Tone, and Javier Verbel. Combinatorial rank attacks against the rectangular simple matrix encryption scheme. In Jintai Ding and Jean-Pierre Tillich, editors, *Post-Quantum Cryptography*, pages 307–322, Cham, 2020. Springer International Publishing.
2. Christine Bachoc, Oriol Serra, and Gilles Zémor. An analogue of Vosper's theorem for extension fields. *Mathematical Proceedings of the Cambridge Philosophical Society*, 163(3):423–452, November 2017.
3. Magali Bardet, Maxime Bros, Daniel Cabarcas, Philippe Gaborit, Ray Perlner, Daniel Smith-Tone, Jean-Pierre Tillich, and Javier Verbel. Improvements of algebraic attacks for solving the rank decoding and minrank problems. In Shiho Moriai and Huaxiong Wang, editors, *Advances in Cryptology – ASIACRYPT 2020*, pages 507–536, Cham, 2020. Springer International Publishing.
4. Luk Bettale, Jean-Charles Faugère, and Ludovic Perret. Cryptanalysis of HFE, multi-HFE and variants for odd and even characteristic. *Designs, Codes and Cryptography*, 69(1):1–52, 2013.

5. Jonathan F. Buss, Gudmund S. Frandsen, and Jeffrey O. Shallit. The computational complexity of some problems of linear algebra. *J. Comput. System Sci.*, 58(3):572–596, June 1999.
6. Daniel Cabarcas, Daniel Smith-Tone, and Javier Verbel. Key recovery attack for zhfe. In Tanja Lange and Tsuyoshi Takagi, editors, *Post-Quantum Cryptography*, pages 289–308, Cham, 2017. Springer International Publishing.
7. Ryann Cartor and Daniel Smith-Tone. EFLASH: A new multivariate encryption scheme. In *Selected Areas in Cryptography - SAC 2018 - 25th International Conference, Calgary, AB, Canada, August 15-17, 2018, Revised Selected Papers*, pages 281–299, 2018.
8. Antoine Casanova, Jean-Charles Faugère, Gilles Macario-Rat, Jacques Patarin, Ludovic Perret, and Jocelyn Ryckeghem. GeMSS: A Great Multivariate Short Signature. Research report, UPMC - Paris 6 Sorbonne Universités ; INRIA Paris Research Centre, MAMBA Team, F-75012, Paris, France ; LIP6 - Laboratoire d'Informatique de Paris 6, December 2017.
9. David A. Cox, John Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 3/e (Undergraduate Texts in Mathematics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.
10. Thomas Debris-Alazard and Jean-Pierre Tillich. Two attacks on rank metric code-based schemes: Ranksign and an identity-based-encryption scheme. In *Advances in Cryptology - ASIACRYPT 2018*, LNCS, pages 62–92, Brisbane, Australia, December 2018. Springer.
11. Jintai Ding. Rainbow. Second round submission to the NIST post-quantum cryptography call, April 2019.
12. Jean-Charles Faugère, Mohab Safey El Din, and Pierre-Jean Spaenlehauer. Computing loci of rank defects of linear matrices using Gröbner bases and applications to cryptology. In *International Symposium on Symbolic and Algebraic Computation, ISSAC 2010, Munich, Germany, July 25-28, 2010*, pages 257–264, 2010.
13. J.C. Faugère, P. Gianni, D. Lazard, and T. Mora. Efficient computation of zero-dimensional gröbner bases by change of ordering. *Journal of Symbolic Computation*, 16(4):329 – 344, 1993.
14. Ralf Fröberg. *An introduction to Gröbner bases.* Pure and applied mathematics. Wiley, 1998.
15. Louis Goubin and Nicolas Courtois. Cryptanalysis of the TTM cryptosystem. In Tatsuaki Okamoto, editor, *Advances in Cryptology - ASIACRYPT 2000*, volume 1976 of *LNCS*, pages 44–57. Springer, 2000.
16. Yasuhiko Ikematsu, Ray A. Perlner, Daniel Smith-Tone, Tsuyoshi Takagi, and Jeremy Vates. HFERP - A new multivariate encryption scheme. In *Post-Quantum Cryptography - 9th International Conference, PQCrypto 2018, Fort Lauderdale, FL, USA, April 9-11, 2018, Proceedings*, pages 396–416, 2018.
17. Dustin Moody, Ray A. Perlner, and Daniel Smith-Tone. Key recovery attack on the cubic ABC simple matrix multivariate encryption scheme. In *Selected Areas in Cryptography - SAC 2016 - 23rd International Conference, St. John's, NL, Canada, August 10-12, 2016, Revised Selected Papers*, pages 543–558, 2016.
18. Morten Øygarden, Patrick Felke, Håvard Raddum, and Carlos Cid. Cryptanalysis of the multivariate encryption scheme eflash. In Stanislaw Jarecki, editor, *Topics in Cryptology – CT-RSA 2020*, pages 85–105, Cham, 2020. Springer International Publishing.

19. Jacques Patarin. Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): Two New Families of Asymmetric Algorithms. In *EUROCRYPT*, pages 33–48, 1996.
20. Jaiberth Porras, John Baena, and Jintai Ding. Zhfe, a new multivariate public key encryption scheme. In Michele Mosca, editor, *Post-Quantum Cryptography*, pages 229–245, Cham, 2014. Springer International Publishing.
21. Netanel Raviv, Ben Langton, and Itzhak Tamo. Multivariate public key cryptosystem from sidon spaces. In Juan A. Garay, editor, *Public-Key Cryptography – PKC 2021*, pages 242–265, Cham, 2021. Springer International Publishing.
22. Ron M. Roth, Netanel Raviv, and Itzhak Tamo. Construction of Sidon spaces with applications to coding. *IEEE Transactions on Information Theory*, 64(6):4412–4422, 2018.
23. Daniel Smith-Tone and Javier A. Verbel. A rank attack against extension field cancellation. In Jintai Ding and Jean-Pierre Tillich, editors, *Post-Quantum Cryptography - 11th International Conference, PQCrypto 2020, Paris, France, April 15-17, 2020, Proceedings*, volume 12100 of *Lecture Notes in Computer Science*, pages 381–401. Springer, 2020.
24. Alan Szepieniec, Jintai Ding, and Bart Preneel. Extension field cancellation: A new central trapdoor for multivariate quadratic systems. In Tsuyoshi Takagi, editor, *Post-Quantum Cryptography*, pages 182–196, Cham, 2016. Springer International Publishing.
25. Chengdong Tao, Adama Diene, Shaohua Tang, and Jintai Ding. Simple matrix scheme for encryption. In *PQCrypto*, pages 231–242, 2013.
26. xxxx xxxx and xxxx xxxx. Crypanalysis tool for the sidon cryptosystem, 2021. https://github.com/Javierverbel/cryptanalysis-sidon-cryptosystem.