Survey on Cryptanalysis of Code-Based Cryptography: from Theoretical to Physical Attacks

Vlad Drăgoi
Faculty of Exact Sciences, "Aurel Vlaicu" University of Arad
Arad, Romania
vlad.dragoi@uav.ro

Tania Richmond
TAMIS team, Inria Rennes - Bretagne Atlantique
Rennes, France
tania.richmond@inria.fr

Dominic Bucerzan
Faculty of Exact Sciences, "Aurel Vlaicu" University of Arad
Arad, Romania
dominic.bucerzan@uav.ro

Axel Legay
TAMIS team, Inria Rennes - Bretagne Atlantique
Rennes, France
axel.legay@inria.fr

***Abstract*—Nowadays public-key cryptography is based on number theory problems, such as computing the discrete logarithm on an elliptic curve or factoring big integers. Even though these problems are considered difficult to solve with the help of a classical computer, they can be solved in polynomial time on a quantum computer. Which is why the research community proposed alternative solutions that are quantum-resistant. The process of finding adequate post-quantum cryptographic schemes has moved to the next level, right after NIST's announcement for post-quantum standardization.**

One of the oldest quantum-resistant proposition goes back to McEliece in 1978, who proposed a public-key cryptosystem based on coding theory. It benefits of really efficient algorithms as well as a strong mathematical background. Nonetheless, its security has been challenged many times and several variants were cryptanalyzed. However, some versions remain unbroken.

In this paper, we propose to give some background on coding theory in order to present some of the main flawless in the protocols. We analyze the existing side-channel attacks and give some recommendations on how to securely implement the most suitable variants. We also detail some structural attacks and potential drawbacks for new variants.

***Index Terms*—Post-quantum cryptography, code-based cryptography, McEliece scheme, coding theory, side-channel analysis**

I. Introduction

The evolution of the Internet and its related security problems created a fertile ground for public-key cryptography (PKC). It is probably one of today's most spread solution to secure communications. Three of the main technologies used for security purposes, namely TLS, PGP, and SSH, all contain elliptic-curve cryptography. A major advantage of PKC compared to secret-key cryptography (SKC) is that today's requirements are all achievable by PKC, namely integrity, confidentiality, authentication, identification, and non-repudiation.

Current PKC bases its strength on mathematical problems from number theory, such as the integer factorization and discrete logarithm problems. In the past, these two problems were considered hard enough for a cryptographic purpose. Nowadays, the security of cryptosystems based on number theory is rather uncertain. This fact is mainly due to the discovery of polynomial time quantum algorithms for solving the aforementioned problems [1]. Even though a real quantum computer able to factor large numbers does not yet exist, the cryptographic community has already started to get ready for this event.

One of the institutes that prepares and elaborates standards for security solutions is the National Institute of Standards and Technology (NIST). It launched a vast program on post-quantum cryptography (PQC) standardization. The purpose of this process is to solicit, evaluate, and standardize one or more quantum-resistant public-key cryptographic algorithms. For that, they will organize an international conference.collocated with PQCrypto 2018. At this moment, the submission phase is finished and the list of candidates is public. Among the possible solutions, code-based cryptography has an important number of candidates. Roughly speaking, 3/8 proposals are code-based protocols.

McEliece introduced in 1978 the first code-based cryptosystem [2]. The scheme is not based on number theory primitives but rather difficult problems coming from coding theory. Its security relies on two problems: the hardness of the Syndrome Decoding Problem [3], and the difficulty to distinguish between a binary Goppa code and a random linear code [4]. When compared to other PKC, McEliece's scheme disposes of various advantages: the complexity of encryption and decryption algorithms are equivalent to those of symmetric schemes, i.e. are very efficient [5]. Also, the best attacks for solving the syndrome decoding problem are exponential in the code length, i.e. McEliece scheme presents a high potential [6].

Our contribution: In this article, we make a state-of-the-art of code-based cryptography, essentially for encryption and signature schemes. We provide the main ideas for theoretical and physical cryptanalysis. Note that in the literature, other surveys exists, such as [7], [8] or the well-known book of Pellikaan at al. [9]. Here, we choose to present this topic under

978-1-5386-1934-6/18/$31.00

today's requirements and process initiated by the NIST, and by that we offer a different point of view on the evolution of code-based PKC.

Organization of the paper: We start by giving necessary background information in cryptography and coding theory (Section II). In the same section, we detail some famous families of codes with a big impact in cryptography. Further, we describe public-key encryption schemes in Section III and signature schemes in Section IV. Section V presents some NIST proposals for standardization and Section VI exhibits the main cryptanalysis techniques for code-based PKC.

II. Preliminaries

A. Cryptography

Definition 1 (Encryption scheme): An encryption scheme is a function mapping a plaintext to a cryptogram, also called ciphertext, using a key. If this key is the same to encrypt and decrypt, i.e. shared between sender and receiver, then we call it a symmetric scheme as well as a secret-key cryptosystem. If a key pair is used, i.e. the receiver's public-key to encrypt and the receiver's private-key to decrypt, then we call it an asymmetric scheme as well as a public-key cryptosystem (PKC).

PKCs are quite inefficient to send long messages. The usual trick is to use a KEM-DEM hybrid encryption paradigm. It allows to combine the advantage of a PKC by avoiding the key distribution problem of a symmetric scheme with the advantage of large message space and lower cost for communications of a symmetric scheme.

Definition 2 (KEM-DEM): A key encapsulation mechanism (KEM) is an asymmetric scheme, more precisely a probabilistic algorithm producing a random symmetric key encrypted (encapsulated) by this scheme. That key is then decrypted (decapsulated) and can be used with a symmetric scheme. Each scheme has its own security condition in order to get an IND-CCA secure encryption scheme. A data encapsulation mechanism (DEM) is a symmetric scheme, more precisely a deterministic algorithm to encrypt a message of arbitrary length using the key given by a KEM (after decapsulation). To summarize, the initial message is encrypted/decrypted by the DEM.

Definition 3 (Signature scheme): A signature scheme is a function that allows to authenticate the sender, ensure the message integrity and the non-repudiation from the sender. A key pair is needed for a digital signature. The sender uses his private-key to sign a message then transmits the message and the signature. The receiver uses the sender's public-key to verify the signature.

Definition 4 (Key exchange): A key exchange (or key establishment) is a protocol defining the key agreement followed by its transport. Such a protocol is designed to share a secret between two (or more) parties.

B. Coding theory

Here we will give only a brief summary of the tools needed to understand how coding theory is used to provide efficient solutions for PKC. For details we address the reader to well-known books [10], [11], [12].

In the following we denote by $\mathbb{F}_q$ the finite field with q elements and $\mathcal{M}_{k,n}(\mathbb{F}_q)$ the set of $k \times n$ matrices with entries in $\mathbb{F}_q$. The *Hamming weight* of a given vector $\boldsymbol{x} \in \mathbb{F}_q^n$, denoted $|\boldsymbol{x}|$, is the number of non-zero coordinates of $\boldsymbol{x}$.

Definition 5 (Linear codes): A *linear code* $\mathscr{C}$ of length n and dimension k is a k-dimensional linear subspace of $\mathbb{F}_q^n$. A *generator matrix* of $\mathscr{C}$ is any $k \times n$ matrix $\boldsymbol{G} \in \mathcal{M}_{k,n}(\mathbb{F}_q)$ with rows that generate $\mathscr{C}$. The dual $\mathscr{C}^\perp$ of $\mathscr{C}$ is the $(n-k)$-dimensional linear subspace defined by

$$\mathscr{C}^\perp = \Big\{ \boldsymbol{z} \in \mathbb{F}_q^n \ : \ \forall \boldsymbol{c} \in \mathscr{C}, \ \sum_{i=1}^{n} c_i z_i = 0 \Big\}.$$

A parity-check matrix of $\mathscr{C}$ is a generator of $\mathscr{C}^\perp$. Another fundamental notion is the minimum distance of a linear code $\mathscr{C}$. It is defined as the minimum over the set of all possible weights $|\boldsymbol{c}|$, over all codewords $\boldsymbol{c} \in \mathscr{C}$. The minimum distance of $\mathscr{C}$ is strongly related to the error detection capability and error correction capacity of $\mathscr{C}$. Typically, a code with a large minimum distance has a big error correction capacity. In general, a random linear code has a minimum distance smaller than or equal to the well-known Gilbert-Varshamov bound, which is linear in the code length [10]. We also know (a non-constructive proof) that there exist binary linear codes that meet the Gilbert-Varshamov bound. Hence, a possible method to generate a code with large minimum distance, could be to randomly pick a linear code and then compute the minimum distance. Unfortunately, this turns out to be a costly solution because computing the minimum distance of a linear code is a difficult problem [13].

Since the initial purpose of linear codes was to detect and correct errors, to any code $\mathscr{C}$ we associate a decoding algorithm. Suppose that we send a codeword $\boldsymbol{c} \in \mathscr{C}$ through a noisy channel. The receiver gets a different vector. Let's denote the received vector by $\boldsymbol{x}$, and $\mathcal{D}ecode(\boldsymbol{x}, \mathscr{C})$ the decoding algorithm. This algorithm computes the most likely codeword $\boldsymbol{c}^\star \in \mathscr{C}$, given $\boldsymbol{x}$ and $\mathscr{C}$. Any solution to the following problem (Syndrome Decoding Problem) can be used to perform the aforementioned task.

Definition 6 (Syndrome Decoding Problem (SDP)):

Instance: A matrix $\boldsymbol{H} \in \mathcal{M}_{n-k,n}(\mathbb{F}_{q^m})$, a vector $\boldsymbol{s} \in \mathbb{F}_{q^m}^{n-k}$ and an integer $\omega > 0$.

Question: Is there a vector $\boldsymbol{x} \in \mathbb{F}_{q^m}^n$ of weight $\leq \omega$, s.t. $\boldsymbol{H}\boldsymbol{x}^T = \boldsymbol{s}$?

When the parity-check matrix $\boldsymbol{H}$ has no particular structure, i.e., is randomly chosen, SDP becomes difficult [3]. So, a challenging research question is to find families of codes with a sufficient large minimum distance and an efficient decoding algorithm.

C. Different families of codes

1) Algebraic codes:

Definition 7 (Generalized Reed-Solomon (GRS) codes): Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ where the α_i are distinct elements of $\mathbb{F}_{q^m}$, and let $\boldsymbol{v} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ where the v_i are nonzero (but not necessarily distinct) elements of $\mathbb{F}_{q^m}$. Then the GRS code, denoted by $GRS_k(\alpha, \boldsymbol{v})$, consists of all vectors $(\boldsymbol{v}_1 f(\alpha_1), \boldsymbol{v}_2 f(\alpha_2), \ldots, \boldsymbol{v}_n f(\alpha_n))$ where f(z) ranges over all polynomials of degree $< k$ with coefficients from $\mathbb{F}_{q^m}$.

It is known that $GRS_k(\alpha, \boldsymbol{v})$ is a $[n, k, n-k+1]$-code (see [10]).

Definition 8 (Alternant codes): The alternant code $\mathscr{A}(\alpha, \boldsymbol{y})$ consists of all codewords of $GRS_{k_0}(\alpha, \boldsymbol{v})$ which have components from $\mathbb{F}_q$, i.e. $\mathscr{A}(\alpha, \boldsymbol{y})$ is the restriction of $GRS_{k_0}(\alpha, \boldsymbol{v})$ to $\mathbb{F}_q$. Thus $\mathscr{A}(\alpha, \boldsymbol{y})$ consists of all vectors $\boldsymbol{a}$ over $\mathbb{F}_q$ such that $\boldsymbol{H}\boldsymbol{a}^T = 0$, where $\boldsymbol{H}$ is given by:

$$\boldsymbol{H} = \begin{pmatrix} y_1 & y_2 & \ldots & y_n \\ \alpha_1 y_1 & \alpha_2 y_2 & \ldots & \alpha_n y_n \\ \vdots & \vdots & \ldots & \vdots \\ \alpha_1^{r-1} y_1 & \alpha_2^{r-1} y_2 & \ldots & \alpha_n^{r-1} y_n \end{pmatrix}.$$

With any vector $\boldsymbol{a} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ over $\mathbb{F}_q$ we associate the rational function

$$R_{\boldsymbol{a}}(z) = \sum_{i=1}^{n} \frac{\boldsymbol{a}_i}{z - \alpha_i}.$$

Definition 9 (Goppa codes): The Goppa code $\Gamma(L, g)$ (or Γ) consists of all vectors $\boldsymbol{a}$ such that $R_{\boldsymbol{a}}(z) \equiv 0 \mod g(z)$, or equivalently such that $R_{\boldsymbol{a}}(z) = 0$ in the polynomial ring $\mathbb{F}_{q^m}[z]/g(z)$. If g(z) is irreducible then Γ is called an irreducible Goppa code.

There are several efficient decoding techniques for GRS, alternant and Goppa codes, such as the Berlekamp-Massey algorithm, the Extended Euclidean algorithm and many others (see [10], [11], [14]). They are mainly algebraic algorithms and exploit the polynomial structure of these codes.

Definition 10 (Generalized Srivastava (GS) codes): In the parity-check matrix for the alternant code $\mathscr{A}(\alpha, \boldsymbol{y})$, suppose $r = st$ and let $\alpha_1, \ldots, \alpha_n$, $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_s$ be $n + s$ distinct elements of $\mathbb{F}_{q^m}$, $z_1, \ldots, z_n$ be nonzero elements of $\mathbb{F}_{q^m}$. The GS code of order st, length $n < q^m - s$, dimension $k \geqslant n - mst$ and minimum distance $d \geqslant st + 1$ has a parity-check matrix given by

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{H}_1 \\ \boldsymbol{H}_2 \\ \vdots \\ \boldsymbol{H}_s \end{pmatrix}$$

where

$$\boldsymbol{H}_i = \begin{pmatrix} \frac{\boldsymbol{z}_1}{\alpha_1 - \boldsymbol{w}_i} & \frac{\boldsymbol{z}_2}{\alpha_2 - \boldsymbol{w}_i} & \ldots & \frac{\boldsymbol{z}_n}{\alpha_n - \boldsymbol{w}_i} \\ \frac{\boldsymbol{z}_1}{(\alpha_1 - \boldsymbol{w}_i)^2} & \frac{\boldsymbol{z}_2}{(\alpha_2 - \boldsymbol{w}_i)^2} & \ldots & \frac{\boldsymbol{z}_n}{(\alpha_n - \boldsymbol{w}_i)^2} \\ \vdots & \vdots & \ldots & \vdots \\ \frac{\boldsymbol{z}_1}{(\alpha_1 - \boldsymbol{w}_i)^t} & \frac{\boldsymbol{z}_2}{(\alpha_2 - \boldsymbol{w}_i)^t} & \ldots & \frac{\boldsymbol{z}_n}{(\alpha_n - \boldsymbol{w}_i)^t} \end{pmatrix}$$

for $i = 1, \ldots, s$.

2) Probabilistic codes:

Definition 11 (Low/Moderate Density Parity-Check (LDPC/MDPC) codes): A (n, r, ω)-code is a linear code defined by a $r \times n$ parity-check matrix ($r < n$) where each row has weight ω. A *LDPC* code is a (n, r, ω)-code with $\omega = O(1)$, when $n \to \infty$ [15]. A *MDPC* code is a (n, r, ω)-code with $\omega = O(\sqrt{n})$, when $n \to \infty$ [16].

Decoding LDPC and MDPC codes can be done using a probabilistic algorithm such as the Bit-Flipping algorithm, due to Gallager [15]. There are several variants, from the classical proposal to more involved ones [15], [17].

Definition 12 (Low Rank Parity-Check (LRPC) codes): A LRPC code of rank d, length n and dimension k over $\mathbb{F}_{q^m}$ is a code with parity-check matrix a $(n - k) \times n$ matrix $\boldsymbol{H}(h_{ij})$ such that the vector subspace of $\mathbb{F}_{q^m}$ generated by its coefficients h_{ij} has dimension at most d. We call this dimension the weight of $\boldsymbol{H}$ (in rank metric). Denoting F the vector subspace of $\mathbb{F}_{q^m}$ generated by the coefficients h_{ij} of $\boldsymbol{H}$, we denote by $\{F_1, F_2, \ldots, F_d\}$ one of its basis.
In practice it means that for any $1 \leqslant i \leqslant n - k$, $1 \leqslant j \leqslant n$, there exist $h_{ijl} \in \mathbb{F}_q$ such that $h_{ij} = \sum_{l=1}^{d} h_{ijl} F_l$.

3) Adding additional structure: This method was considerably employed by the scientific community in order to decrease the size of the public-key in the McEliece PKC.

Definition 13 (Quasi-Cyclic (QC) codes): A (n, r)-linear code over $\mathbb{F}_q$ is called a quasi-cyclic code of index l if $n = ml$ and for every codeword $c \in \mathscr{C}$, there exists a number l such that the codeword obtained by l cyclic shifts is also a codeword in $\mathscr{C}$.

Definition 14 (Quasi-Dyadic (QD) codes): Let $r = 2^k$ for some $k \in \mathbb{N}$. Given a ring $\mathcal{R}$ (in our case the finite field $\mathbb{F}_{q^m}$) and a vector $\boldsymbol{h} = (\boldsymbol{h}_0, \ldots, \boldsymbol{h}_{n-1}) \in \mathcal{R}^n$, the dyadic matrix $\Delta(\boldsymbol{h}) \in \mathcal{R}^{n \times n}$ is the symmetric matrix with components $\Delta_{ij} = h_{i \oplus j}$, where $\oplus$ stands for bitwise exclusive-or on the binary representations of the indices. The sequence $\boldsymbol{h}$ is called its signature. Moreover, $\Delta(t, \boldsymbol{h})$ denotes the matrix $\Delta(\boldsymbol{h})$ truncated to its first t rows. Finally, we call a matrix quasi-dyadic if it is a block matrix whose component blocks are $t \times t$ dyadic submatrices. If n is a power of 2, then every $2^k \times 2^k$ dyadic matrix can be described recursively as:

$$M = \begin{pmatrix} A & B \\ B & A \end{pmatrix}$$

where each block is a $2^{k-1} \times 2^{k-1}$ dyadic matrix (and where any 1×1 matrix is dyadic).

III. Public-key encryption schemes

A. McEliece encryption scheme

The McEliece PKC [2] is composed of three algorithms: *key generation* (KeyGen), *encryption* (Encrypt) and *decryption* (Decrypt). The key generation algorithm takes as input the integers n, m, k, t, q such that $k < n$ and $t < n$ and outputs the public-key/private-key pair $(\mathsf{pk}, \mathsf{sk})$.

a) $\mathsf{KeyGen}(n, m, k, t, q) = (\mathsf{pk}, \mathsf{sk})$*:*

1) Pick a generator matrix $\boldsymbol{G}$ of a $[n, k]$-code $\mathscr{C}$ that can corrects t errors.
2) Pick a random $\boldsymbol{S} \in \mathsf{GL}_k(\mathbb{F}_{q^m})$ and a $n \times n$ permutation matrix $\boldsymbol{P}$.
3) Compute $\boldsymbol{G}_{\text{pub}} \stackrel{\text{def}}{=} \boldsymbol{SGP}$.
4) Output

$$\mathsf{pk} = (\boldsymbol{G}_{\text{pub}}, t) \text{ and } \mathsf{sk} = (\boldsymbol{S}, \boldsymbol{G}, \boldsymbol{P}).$$

The encryption of a message $\boldsymbol{m} \in \mathbb{F}_{q^m}^k$ works as following.

b) $\mathsf{Encrypt}(\boldsymbol{m}, \mathsf{pk}) = \boldsymbol{z}$*:*

1) Generate at random $\boldsymbol{e} \in \mathbb{F}_{q^m}^n$ with $|\boldsymbol{e}| \leq t$.
2) Return $\boldsymbol{z} = \boldsymbol{mG}_{\text{pub}} \oplus \boldsymbol{e}$.

In order to decrypt the ciphertext $\boldsymbol{z}$ using sk, one uses the following function:

c) $\mathsf{Decrypt}(\boldsymbol{z}, \mathsf{sk}) = \boldsymbol{m}$*:*

1) Compute $\boldsymbol{z}^* = \boldsymbol{zP}^{-1}$ and $\boldsymbol{m}^* = \mathcal{D}ecode(\boldsymbol{z}^*, \boldsymbol{H})$.
2) Output $\boldsymbol{m}^*\boldsymbol{S}^{-1}$.

We denote by $\mathcal{D}ecode(.,.)$ an efficient decoding algorithm for $\mathscr{C}$. Notice that multiplying the error vector by a permutation does not change the weight of the vector. One can easily verify the correctness of the scheme by checking

$$\mathsf{Decrypt}(\mathsf{Encrypt}(\boldsymbol{m}, \mathsf{pk}), \mathsf{sk}) = \boldsymbol{m}.$$

B. Niederreiter encryption scheme

Often called the McEliece dual, Niederreiter proposed a new PKC [18], using a parity-check matrix instead of a generator matrix to encrypt. The securities of the McEliece and Niederreiter schemes are closely related, proved to be equivalent in [19]. In the classical Niederreiter scheme, the author proposed to employ the GRS code. This choice turned out to be very unwise since the structure of a GRS code turned out to be much easier to leak information than that of a binary Goppa code.

C. Hybrid McEliece Encription Scheme (HyMES)

An idea to improve the McEliece PKC appeared in [20], a decade after the first proposal: use the error to transfer more information. It seems that this paper was missed during several years by the community. One citation was made by Sendrier in 2002 [21], then another one for the implementation done by Bhaskar and Sendrier in 2008 called Hybrid McEliece Encryption Scheme (HyMES) [22]. This improvement allows to increase the information rate of the scheme. Goppa codes were chosen as well as in the original scheme.

D. McBits and QcBits

A KEM/DEM variant of the Niederreiter PKC, called McBits, was proposed by Bernstein, Chou and Schwabe in [23]. This proposition was done in order to close any timing attack. It is a Niederreiter implementation in fast and constant-time. The main steps are the same as in the Niederreiter scheme but different algorithms are used for critical parts. The additive fast Fourier transform over finite fields allows to recover the roots of the so-called error-locator polynomial and its transpose to compute the syndrome. Bernstein et al. also proposed sorting networks to avoid cache attacks on the permutation. Recently Chou published an improvement of the implementation, called McBits revisited [24]. He proposed to use Beneš network for a constant-time permutation and the same techniques as for decryption to generate the keys and encrypt also in constant-time. The main difference between both versions is that McBits decrypts many ciphertexts at the same time whereas McBits revisited exploits internal parallelism to get better performance for a higher security level. Moreover, both McBits versions use binary Goppa codes, and another variant called QcBits uses QC-MDPC codes [25]. They all share the idea of bit-slicing operations to achieve constant-time executions.

E. Other variants

One of the directions in code-based cryptography was to propose and analyze the security of the McEliece scheme using other types of codes. This process started with the GRS codes and the Neiderreiter scheme [18], then continued with the Reed-Muller codes and the Sidelnikov scheme [26], and many other codes. A quasi-complete survey on the evolution of the McEliece scheme is proposed in [8] and [7].

IV. Signature schemes

A. CFS scheme

The CFS [27] signature scheme exploits the decoding capacity of binary Goppa codes. The scheme works as follows. It takes the data that needs to be signed and hash it (denote the result $\boldsymbol{x}$), using any standard hash function. Then it searches for an integer such that when appended to $\boldsymbol{x}$ and hashed into a vector $\boldsymbol{s}$, $\boldsymbol{s}$ becomes a valid syndrome for a codeword in $\mathscr{C}$.

Notice that a code that has the capacity to decode many vectors from the ambient space is a good candidate for signature schemes. However, one expects that such a code has an important structure and thus might be vulnerable to cryptanalysis. Unfortunately, for binary Goppa codes the average number of trials is approximately $t!$, which makes the scheme not very efficient for practical purposes.

B. With other codes

RankSign is actually a NIST submission. The original proposal was made in [28] using LRPC codes (see Def 12). At the same time, low-density generator matrix (LDGM) codes were proposed in a signature scheme in [29], but the proposal was attacked in [30]. SURF is a recent signature scheme based on the "$\boldsymbol{u}|\boldsymbol{u}+\boldsymbol{v}$" construction proposed in [31].

V. Recent protocols

We do not pretend to be exhaustive here because the NIST standardization for post-quantum cryptography has just started.

A. Ouroboros

It is a key exchange protocol proposed by Deneuville, Gaborit and Zémor [32]. The protocol is inspired by two other schemes: Alekhnovitch proposal [33] and the QC-MDPC PKC [34]. The main idea of this approach is to derive a system with a security reduction to the problem of decoding random linear codes. The public code in this case is a random one and the secret key generates an MDPC code. It is also one of the candidates to the NIST standardization process.

B. Loidreau's cryptosystem

Loidreau's improvement for McEliece in rank metric with Gabidulin codes [35]. Several rank metric schemes based on Gabidulin codes were proposed but unfortunately they were cryptanalyzed, mainly due to their structure (see [8] for a detailed survey of the evolution of this scheme). However, Loidreau's scheme propose a different masking technique that avoids all the structural attacks on Gabidulin codes.

C. BIKE

BIKE is a KEM/DEM from Barreto, Gueron, Güneysu, Misoczki, Persichetti, Sendrier and Tillich [36]. It has three variants BIKE1, BIKE2 and BIKE3, and exploits the decoding algorithms used for LDPC and MDPC codes. If BIKE2 has a lot of similarities with the QC-MDPC scheme, BIKE3 is highly inspired by Ouroboros. This proposal is also a candidate to NIST standardization.

D. DAGS

DAGS is a KEM/DEM candidate to the NIST standardization from Cayrel, Persichetti, Gueye, N'diaye, Klamti, Dione and Boidje using QD-GS codes [37]. GS codes are a subclass of alternant codes. It would be interesting to check if attacks against alternant codes can not be applied on GS codes. We known that Goppa codes are still resilient so GS codes are maybe also resilient against existing attacks.

E. RLCE

Random Linear Code-based public-key Encryption Scheme (RLCE) from Wang [38] is based on the juxtaposition of a GRS code with a random linear code. The idea of this PKC is to use a distortion matrix that not only mix the random columns with the structured ones, but also adds them. In such a manner, previous attacks based on the square code are no longer possible. RLCE is also one of the candidates to NIST standardization.

VI. Main threats

A. Distinguisher for the McEliece scheme

The first problem that an adversary has to solve for the McEliece scheme is that of finding the nature of the code. It is also known in the literature as the Goppa Code Distinguishing problem [27], [21]: given a linear code $\mathscr{C}$, decide whether $\mathscr{C}$ is a Goppa code or a random code. The problem can be formulated for any other family of codes, for example Reed-Muller codes, Polar codes, GRS codes etc. Further, we explain some efficient solutions to solve this problem.

1) The Square code technique: The first solution is to determine the dimension of the square code and check out whether it satisfies the condition given in [4]. The work factor of this method is dominated by the computation of the square of a code ($n\binom{k}{2}$ bit operations, where k and n are the parameters of the code).

Definition 15: Let $\boldsymbol{x}$ and $\boldsymbol{y} \in \mathbb{F}_2^n$, then the component-wise product of $\boldsymbol{y}$ and $\boldsymbol{x}$ is

$$\boldsymbol{x} \star \boldsymbol{y} \stackrel{\text{def}}{=} (x_1 y_1, \ldots, x_n y_n) \ \in \ \mathbb{F}_2^n.$$

Definition 16 (Star product code): For $i \in \{1, 2\}$ let $\mathscr{C}_i$ be a $[n, k_i, d_i]$ binary linear code. Then the star product code of $\mathscr{C}_1$ and $\mathscr{C}_2$ is the binary linear code denoted $\mathscr{C}_1 \star \mathscr{C}_2$ defined as

$$\mathscr{C}_1 \star \mathscr{C}_2 \stackrel{\text{def}}{=} \mathrm{Span}_{\mathbb{F}_2} \left\{ \boldsymbol{c}_1 \star \boldsymbol{c}_2 \mid \boldsymbol{c}_1 \in \mathscr{C}_1 \ and \ \boldsymbol{c}_2 \in \mathscr{C}_2 \right\}.$$

Proposition 17: Let $\mathscr{C}_1$ be a $[n, k_1, d_1]$ binary linear code and $\mathscr{C}_2$ be a $[n, k_2, d_2]$ binary linear code. Then we have

$$\mathsf{d}_{\min} \left(\mathscr{C}_1 \star \mathscr{C}_2 \right) \leq \min\{d_1, d_2\}.$$

$$\dim \left(\mathscr{C}_1 \star \mathscr{C}_2 \right) \leq \min \left\{ n, k_1 k_2 - \binom{\dim \left(\mathscr{C}_1 \cap \mathscr{C}_2 \right)}{2} \right\}.$$

Using Proposition 17 one can efficiently compute given the public code $\mathscr{C}_p$, the dimension of the square code $\mathscr{C}_p \star \mathscr{C}_p$. If the dimension of the code $\mathscr{C}_p$ is such that $\binom{k+1}{2} < n$ then one can determine using the square code whether $\mathscr{C}_p$ is a random like linear code. If the dimension of $\mathscr{C}_p \star \mathscr{C}_p$ is far for being equal to $\binom{k+1}{2}$, then with high probability the code $\mathscr{C}_p$ is not a random one.

The square code technique was widely used in code-based cryptography. For the first time in [4] high rate Goppa codes were proved to be distinguishable from random codes. Other families of codes followed, such as Reed-Muller ([39], [40]), Generalized Reed-Solomon code ([41]). Another technique due to Sendrier [42], [43] is to analyze the structure of the Hull.

2) Properties of the hull: Sendrier pointed out that the nature of the code is related to the dimension of its hull. On one hand, for structured codes, such as Reed-Muller, Polar, GRS etc., the Hull is far from being trivial. On the other hand, for random linear codes the following holds.

Proposition 18 ([42], [31]): The expected dimension of the hull of a random linear code is $O(1)$. It is smaller than t with probability $\geqslant 1 - O(2^{-t})$.

The technique is used in [31] to prove that the $(\boldsymbol{u}|\boldsymbol{u}+\boldsymbol{v})$ code used for the SURF signature does not satisfy the random like hypothesis in the security proof. It can also be used in the case of the McEliece variant using Polar codes [44]. In this case it is highly probable that the Polar code used in the scheme will be equal to its hull (see [45], [46], [47]).

B. Message Recovery attacks

Message recovery attacks are an important issue in code-based cryptography. The problem of retrieving the private message from a ciphertext is directly related to the hardness of generic decoding for linear codes.

Decoding linear codes is an old studied problem in coding theory that started with Prange's algorithm [48]. All the improvements of this algorithm [49], [50], [51], [52], [53], [54] still have an exponential complexity in the number of errors that have to be corrected. In the literature they are known as variants of the Information Set Decoding (ISD).

The ISD algorithm searches for an information set such that the error positions are all out of the information set. Since almost all McEliece variants base their security on the assumption that the public code is indistinguishable from a random linear code, the security level against the MRA attacks is given by the complexity of the ISD, namely $O(e^{-\omega \log(1-R)(1+o(1))})$ when $\omega = o(n)$ [6].

A different approach is the statistical decoding proposed for the first time by Al Jabri in [55]. The algorithm was improved a little bit by Overbeck in [56] and later on by Fossorier, Kobara and Imai in [57]. The latest variant of the statistical decoding is by Debris-Alazard and Tillich [58].

C. Key Recovery Attacks

Being able to compute the private key given the public key is often reduced to solve the Code Equivalence Problem:

Definition 19: Let $\boldsymbol{G}$ and $\boldsymbol{G}^*$ be the generating matrices for two $[n, k]$ binary linear codes. Given $\boldsymbol{G}$ and $\boldsymbol{G}^*$ does there exist a $k \times k$ binary invertible matrix $\boldsymbol{S}$ and $n \times n$ permutation matrix $\boldsymbol{P}$ such that $\boldsymbol{G}^* = \boldsymbol{SGP}$?

The problem was first studied by Petrank and Roth over the binary field [59]. But the most common algorithm used to solve this problem is the Support Splitting Algorithm (SSA) [60]. SSA is very efficient in the random case, but cannot be used in the case of codes with large hulls or codes with large permutation group such as Goppa codes, Reed-Muller codes, etc.

When the SSA is infeasible, other efficient techniques can be employed such as the *Minimum Weight Codewords* approach. The main idea is to use the subcode spanned by the set of minimum weight codewords and solve the code equivalence problem for the latter code. Computing the set of minimum weight codewords of a given code is usually performed by means of any of the ISD variants.

The idea behind this technique is that in the case of several known families of codes, the code spanned by the set of minimum weight codewords is almost the entire code. This happens for Reed-Muller codes, Polar codes and more generally for any Decreasing Monomial codes (see [46], [47]). The technique was used to solve the code equivalence problem for Reed-Muller codes [61], Polar codes [45], [47], QC-LDPC codes [62], a variant of the McEliece scheme using QC-LDPC and QC-MDPC codes [63].

Distinguisher-based attacks are also very efficient for solving this problem (see [41] for Reed-Solomon variant, [39] for Sidelnikov's variant, [64] for algebraic codes). Notice that when we deal with the code equivalence problem, the various techniques that we might employ do not necessary preserve the permutation group of the initial code. By that we refer, for example, to the square code. In this case the permutation group of the code is included in the permutation group of the square code (see Appendix A in [47]).

Remark 20: A common way of avoiding typical key recovery attacks on encryption schemes is to take shortened or punctured codes, or possibly to combine these two techniques. These are really efficient ways of protecting the schemes since the structure of the private code is someway shattered.

Hence, few code families are still secure in a McEliece type scheme [8], like the binary Goppa codes or the QC-MDPC codes. If the key recovery attacks, for both binary Goppa and QC-MDPC codes are exponential in the error weight (see [34]), there are particular private keys that can be efficiently recovered. These type of keys are known in the literature as *weak keys* and have to be avoided in the key generation step. The existence of weak keys for binary Goppa codes was discovered by Loidreau and Sendrier in [65]. In the case of QC-MDPC codes, weak keys were discovered by Bardet et al. in [66] and secure key generation algorithm for the corresponding McEliece variant was proposed in [47].

D. Side-channel attacks

Side-channel attacks (SCAs) exploit a physical phenomenon of an implementation, e.g. running time in software or power consumption in hardware. The first SCA against the McEliece PKC was proposed in 2008 [67]. Main results in physical cryptanalysis were made during four PhD theses [68], [69], [70], [14]. We give in this subsection the main ideas for this kind of attacks. A table of all SCA (to the best of our knowledge) can find in Appendix.

1) SCA against Goppa codes: The only SCA against the McEliece key generation was proposed in [67]. The first idea is to detect through the power consumption the structure of the hidden parity-check matrix for the private Goppa code. Depending on the chosen description for the parity-check matrix, this attack is possible for the generalized BCH definition but not for the alternant definition of the code. The second idea is to use some precomputation of the Goppa polynomial over an assumed known support. A straightforward way to avoid this attack is to apply the Horner evaluation. More recently, the multi-point evaluation method was proposed in [23] via additive fast Fourier Transform and permutation.

We now focus on the McEliece decryption using Goppa codes. Most of those attacks target the Patterson algorithm [71], a decoding algorithm for binary Goppa codes. Starting by the ciphertext permutation, a timing attack on cache memory was proposed in [67]. Strenzke et al. gave a countermeasure based on address masking, but this was then attacked in [72].

Then a profiling was done in [73] for ciphertext permutation and syndrome computation. These profiles describe four ways to combine both consecutive steps. One method to permute the

ciphertext is to multiply it by the inverse of the permutation matrix. Otherwise, even the support is permuted before computation or the permutation matrix is multiplied by the parity-check matrix before application to the ciphertext. One method to compute the syndrome is to multiply the ciphertext by the parity-check matrix, otherwise some polynomial operations are made. Heyse et al. also provide in this paper the first power analysis attack, against the ciphertext permutation, syndrome computation and syndrome inversion. Strenzke proposed a timing attack against the syndrome inversion three years later [74]. The idea is to first locate the zero element into the support and try the decryption with an error vector of small degree by measuring the execution time of the extended Euclidean algorithm (EEA).

A similar attack was done against the determination of the error-locator polynomial in [75]. A timing difference is observed on the EEA to determine if flipping a bit added or erased a bit in the error vector. The associated countermeasure was to make the same amount of steps in the EEA (not depending on the error weight anymore). A second attack with randomly chosen ciphertext and error vector of small weight was proposed in [76] to get a system of linear equations whose unknowns are the coefficients of the error-locator polynomial. A similar countermeasure idea was provided: checking the degree of the result to be sure it is equal to the stop condition in the EEA.

The evaluation of the error-locator polynomial was firstly attacked in [67]. The idea is that the execution time depends on the polynomial degree. By flipping one bit in a ciphertext, an attacker can determine whether an error was corrected or not. A countermeasure is to increase the degree of the evaluated polynomial. This idea was then improved in [77] by choosing the evaluated polynomial as a product between the actual error-locator polynomial and another one with roots in a bigger field extension.

2) SCA against MDPC codes: Less attacks were done against MDPC codes. The first differential power analysis in code-based cryptography was proposed in [78]. This attack aimed the syndrome computation of chosen single-one ciphertext for QC-MDPC codes on a hardware implementation. Later Chen et al. proposed a countermeasure in [79] using Boolean masking of the same size than the parity-check matrix. Both works were extended in [80] including countermeasures with masking applied on the key and the syndrome. Chen et al. avoid first-order side channel attacks by using a threshold during the syndrome computation and decoding part. The syndrome computation in the QcBits implementation [25] was recently attacked by a differential power analysis in [81]. The syndrome computation is not done as a vector-matrix product but the parity-check matrix is described like an exclusive or between rotations. The associated countermeasure is to mask the ciphertext with a random codeword before the syndrome computation, previously proposed in [72].

VII. Conclusions and perspectives

Code-based cryptography became one of the most promising post-quantum security solutions. It is a dynamic field, especially because of the NIST's standardization. However, the community needs to be aware of several weaknesses, that are theoretical as well as physical.

The mathematical problems in code-based cryptography are well known. Depending on the chosen code, the first issue is to distinguish it from a random code. On one hand, message recovery attacks are mainly based on the information decoding problem. On the other hand, key recovery attacks are much difficult to classify, by their various methods, but much more efficient from the attacker point of view.

Side-channel analysis must be performed on digital signature schemes and key-establishment algorithms. These schemes are the most deployed in real-world. Side-channel analysis should also be improved on public-key encryption.

References

[1] P. W. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in *FOCS*, S. Goldwasser, Ed., 1994, pp. 124–134.

[2] R. J. McEliece, *A Public-Key System Based on Algebraic Coding Theory*. Jet Propulsion Lab, 1978, pp. 114–116, dSN Progress Report 44.

[3] E. Berlekamp, R. McEliece, and H. van Tilborg, "On the inherent intractability of certain coding problems," *IEEE Trans. Inform. Theory*, vol. 24, no. 3, pp. 384–386, May 1978.

[4] J.-C. Faugère, V. Gauthier, A. Otmani, L. Perret, and J.-P. Tillich, "A distinguisher for high rate McEliece cryptosystems," *IEEE Trans. Inform. Theory*, vol. 59, no. 10, pp. 6830–6844, Oct. 2013.

[5] R. Overbeck and N. Sendrier, "Code-based cryptography," in *Post-quantum cryptography*, D. J. Bernstein, J. Buchmann, and E. Dahmen, Eds. Springer, 2009, pp. 95–145.

[6] R. Canto-Torres and N. Sendrier, "Analysis of information set decoding for a sub-linear error weight," in *Post-Quantum Cryptography 2016*, ser. Lecture Notes in Comput. Sci., Fukuoka, Japan, Feb. 2016, pp. 144–161.

[7] N. Sendrier, "Code-based cryptography: State of the art and perspectives," *IEEE Security Privacy*, vol. 15, no. 4, pp. 44–50, 2017.

[8] D. Bucerzan, V. Dragoi, and H. T. Kalachi, *Evolution of the McEliece Public Key Encryption Scheme*. Cham: Springer International Publishing, 2017, pp. 129–149. [Online]. Available: https://doi.org/10.1007/978-3-319-69284-5_10

[9] R. Pellikaan, X.-W. Wu, S. Bulygin, and R. Jurrius, "Error-correcting codes and cryptology," 2012.

[10] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, 5th ed. Amsterdam: North–Holland, 1986.

[11] R. M. Roth, *Introduction to Coding Theory*. New York, NY, USA: Cambridge University Press, 2006.

[12] W. C. Huffman and V. Pless, *Fundamentals of error-correcting codes*. Cambridge University Press, Cambridge, 2003. [Online]. Available: http://dx.doi.org/10.1017/CBO9780511807077

[13] A. Vardy, "The intractability of computing the minimum distance of a code," *IEEE Trans. Inform. Theory*, vol. 43, no. 6, pp. 1757–1766, Nov. 1997.

[14] T. Richmond, "Implantation sécurisée de protocoles cryptographiques basés sur les codes correcteurs d'erreurs," Ph.D. dissertation, Université Jean Monnet, Saint-Étienne (France), October 2016.

[15] R. G. Gallager, *Low Density Parity Check Codes*. Cambridge, Massachusetts: M.I.T. Press, 1963.

[16] S. Ouzan and Y. Be'ery, "Moderate-density parity-check codes," *arXiv preprint arXiv:0911.3262*, 2009.

[17] J. Chaulet and N. Sendrier, "Worst case qc-mdpc decoder for mceliece cryptosystem," in *2016 IEEE International Symposium on Information Theory (ISIT)*, July 2016, pp. 1366–1370.

[18] H. Niederreiter, "Knapsack-type cryptosystems and algebraic coding theory," *Problems of Control and Information Theory*, vol. 15, no. 2, pp. 159–166, 1986.

[19] Y. X. Li, R. H. Deng, and X. M. Wang, "On the equivalence of McEliece's and Niederreiter's public-key cryptosystems," *IEEE Trans. Inform. Theory*, vol. 40, no. 1, pp. 271–273, 1994.

[20] C.-S. Park, "Improving code rate of McEliece's public-key cryptosystem," *Electronics Letters*, vol. 25, no. 21, pp. 1466–1467, October 1989. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=46216&queryText%3Dimproving+code+rate+of+McEliece%27s+public-key+cryptosystem

[21] N. Sendrier, "Cryptosystèmes à clé publique basés sur les codes correcteurs d'erreurs," Mémoire d'habilitation à diriger des recherches, Université Paris 6, Paris, France, Mars 2002.

[22] B. Biswas and N. Sendrier, "McEliece cryptosystem implementation: theory and practice," in *Post-Quantum Cryptography, Second International Workshop, PQCrypto 2008, Cincinnati, OH, USA, October 17-19, 2008, Proceedings*, ser. Lecture Notes in Comput. Sci., J. Buchmann and J. Ding, Eds., vol. 5299. Springer, 2008, pp. 47–62.

[23] D. J. Bernstein, T. Chou, and P. Schwabe, "Mcbits: Fast constant-time code-based cryptography," in *Cryptographic Hardware and Embedded Systems - CHES 2013*, ser. Lecture Notes in Comput. Sci., G. Bertoni and J. Coron, Eds., vol. 8086. Springer, 2013, pp. 250–272.

[24] T. Chou, *McBits Revisited*. Cham: Springer International Publishing, August 2017, pp. 213–231. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-66787-4_11

[25] ——, *QcBits: Constant-Time Small-Key Code-Based Cryptography*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 280–300. [Online]. Available: http://dx.doi.org/10.1007/978-3-662-53140-2_14

[26] V. M. Sidelnikov, "A public-key cryptosytem based on Reed-Muller codes," *Discrete Math. Appl.*, vol. 4, no. 3, pp. 191–207, 1994.

[27] N. Courtois, M. Finiasz, and N. Sendrier, "How to achieve a McEliece-based digital signature scheme," in *Advances in Cryptology - ASIACRYPT 2001*, ser. Lecture Notes in Comput. Sci., vol. 2248. Gold Coast, Australia: Springer, 2001, pp. 157–174.

[28] P. Gaborit, O. Ruatta, J. Schrek, and G. Zémor, "New results for rank-based cryptography," in *Progress in Cryptology - AFRICACRYPT 2014*, ser. Lecture Notes in Comput. Sci., vol. 8469, 2014, pp. 1–12.

[29] M. Baldi, M. Bianchi, F. Chiaraluce, J. Rosenthal, and D. Schipani, "Using LDGM codes and sparse syndromes to achieve digital signatures," in *Post-Quantum Cryptography 2013*, ser. Lecture Notes in Comput. Sci., vol. 7932. Springer, 2013, pp. 1–15.

[30] A. Phesso and J. Tillich, "An efficient attack on a code-based signature scheme," in *Post-Quantum Cryptography 2016*, ser. Lecture Notes in Comput. Sci., vol. 9606, Fukuoka, Japan, Feb. 2016, pp. 86–103. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-29360-8_7

[31] T. Debris-Alazard, N. Sendrier, and J. Tillich, "A new signature scheme based on $(u|u+v)$ codes," *CoRR*, vol. abs/1706.08065, 2017. [Online]. Available: http://arxiv.org/abs/1706.08065

[32] J.-C. Deneuville, P. Gaborit, and G. Zémor, "Ouroboros: A simple, secure and efficient key exchange protocol based on coding theory," in *Post-Quantum Cryptography*, T. Lange and T. Takagi, Eds. Cham: Springer International Publishing, 2017, pp. 18–34.

[33] M. Alekhnovich, "More on average case vs approximation complexity," *Computational Complexity*, vol. 20, no. 4, pp. 755–786, 2011. [Online]. Available: http://dx.doi.org/10.1007/s00037-011-0029-x

[34] R. Misoczki, J.-P. Tillich, N. Sendrier, and P. S. L. M. Barreto, "MDPC-McEliece: New McEliece variants from moderate density parity-check codes," in *Proc. IEEE Int. Symposium Inf. Theory - ISIT*, 2013, pp. 2069–2073.

[35] P. Loidreau, "A new rank metric codes based encryption scheme," in *Post-Quantum Cryptography*, T. Lange and T. Takagi, Eds. Cham: Springer International Publishing, 2017, pp. 3–17.

[36] N. Aragon, P. S. L. M. Barreto, S. Bettaieb, L. Bidoux, O. Blazy, J.-C. Deneuville, P. Gaborit, S. Gueron, T. Guneysu, C. Aguilar Melchor, R. Misoczki, E. Persichetti, N. Sendrier, J.-P. Tillich, and G. Zémor, "BIKE: Bit Flipping Key Encapsulation," Dec. 2017, submission to the NIST post quantum standardization process. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01671903

[37] G. Banegas, P. S. L. M. Barreto, B. O. Boidje, P.-L. Cayrel, G. N. Dione, K. Gaj, C. T. Gueye, R. Haeussler, J. B. Klamti, O. N'diaye, D. T. Nguyen, E. Persichetti, and J. E. Ricardini, "DAGS: Key encapsulation using dyadic GS codes," Cryptology ePrint Archive, Report 2017/1037, 2017, https://eprint.iacr.org/2017/1037.

[38] Y. Wang, "Quantum resistant random linear code based public key encryption scheme rlce," in *Information Theory (ISIT), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 2519–2523.

[39] I. V. Chizhov and M. A. Borodin, "The failure of McEliece PKC based on Reed-Muller codes." IACR Cryptology ePrint Archive, Report 2013/287, 2013, http://eprint.iacr.org/.

[40] A. Otmani and H. Talé-Kalachi, "Square code attack on a modified Sidelnikov cryptosystem," in *Codes, Cryptology, and Information Security - First International Conference, C2SI 2015, Rabat, Morocco, May 26-28, 2015, Proceedings - In Honor of Thierry Berger*, ser. Lecture Notes in Computer Science, S. E. Hajji, A. Nitaj, C. Carlet, and E. M. Souidi, Eds., vol. 9084. Springer, 2015, pp. 173–183.

[41] A. Couvreur, P. Gaborit, V. Gauthier-Umaña, A. Otmani, and J.-P. Tillich, "Distinguisher-based attacks on public-key cryptosystems using Reed-Solomon codes," *Des. Codes Cryptogr.*, vol. 73, no. 2, pp. 641–666, 2014. [Online]. Available: http://dx.doi.org/10.1007/s10623-014-9967-z

[42] N. Sendrier, "On the dimension of the hull," in *SIAM J. Discrete Math.*, vol. 10, no. 2, 1997, pp. 282–293.

[43] ——, "On the security of the McEliece public-key cryptosystem," in *Information, coding and mathematics*, ser. Kluwer International Series Engineering and Computer Science, H. C. A. v. T. Mario Blaum, Patrick G. Farrell, Ed., vol. 687, 2002, pp. 141–163.

[44] S. R. Shrestha and Y.-S. Kim, "New McEliece cryptosystem based on polar codes as a candidate for post-quantum cryptography," in *2014 14th International Symposium on Communications and Information Technologies (ISCIT)*. IEEE, 2014, pp. 368–372.

[45] M. Bardet, J. Chaulet, V. Dragoi, A. Otmani, and J.-P. Tillich, "Cryptanalysis of the McEliece public key cryptosystem based on polar codes," in *Post-Quantum Cryptography2016*, ser. Lecture Notes in Comput. Sci., Fukuoka, Japan, Feb. 2016, pp. 118–143. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-29360-8_9

[46] M. Bardet, V. Dragoi, A. Otmani, and J. Tillich, "Algebraic properties of polar codes from a new polynomial formalism," in *IEEE International Symposium on Information Theory, ISIT 2016, Barcelona, Spain, July 10-15, 2016*, 2016, pp. 230–234. [Online]. Available: http://dx.doi.org/10.1109/ISIT.2016.7541295

[47] V. Dragoi, "Algebraic approach for the study of algorithmic problems coming from cryptography and the theory of error correcting codes," Theses, Université de Rouen, France, Jul. 2017. [Online]. Available: https://hal.archives-ouvertes.fr/tel-01627324

[48] E. Prange, "The use of information sets in decoding cyclic codes," *IRE Transactions on Information Theory*, vol. 8, no. 5, pp. 5–9, 1962. [Online]. Available: http://dx.doi.org/10.1109/TIT.1962.1057777

[49] J. Stern, "A method for finding codewords of small weight," in *Coding Theory and Applications*, ser. Lecture Notes in Comput. Sci., G. D. Cohen and J. Wolfmann, Eds., vol. 388. Springer, 1988, pp. 106–113.

[50] I. Dumer, "On minimum distance decoding of linear codes," in *Proc. 5th Joint Soviet-Swedish Int. Workshop Inform. Theory*, Moscow, 1991, pp. 50–52.

[51] A. Barg, "Minimum distance decoding algorithms for linear codes," in *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, 12th International Symposium, AAECC-12, Toulouse, France, June 23-27, 1997, Proceedings*, ser. Lecture Notes in Comput. Sci., vol. 1255. Springer, 1997, pp. 1–14.

[52] A. May, A. Meurer, and E. Thomae, "Decoding random linear codes in $O(2^{0.054n})$," in *Advances in Cryptology - ASIACRYPT 2011*, ser. Lecture Notes in Comput. Sci., D. H. Lee and X. Wang, Eds., vol. 7073. Springer, 2011, pp. 107–124.

[53] A. Becker, A. Joux, A. May, and A. Meurer, "Decoding random binary linear codes in $2^{n/20}$: How $1 + 1 = 0$ improves information set decoding," in *Advances in Cryptology - EUROCRYPT 2012*, ser. Lecture Notes in Comput. Sci. Springer, 2012.

[54] A. May and I. Ozerov, "On computing nearest neighbors with applications to decoding of binary linear codes," in *Advances in Cryptology - EUROCRYPT 2015*, ser. Lecture Notes in Comput. Sci., E. Oswald and M. Fischlin, Eds., vol. 9056. Springer, 2015, pp. 203–228.

[55] A. A. Jabri, "A statistical decoding algorithm for general linear block codes," in *Cryptography and coding. Proceedings of the 8th IMA International Conference*, ser. Lecture Notes in Comput. Sci., B. Honary, Ed., vol. 2260. Cirencester, UK: Springer, Dec. 2001, pp. 1–8.

[56] R. Overbeck, "Statistical decoding revisited," in *Information security and privacy : 11th Australasian conference, ACISP 2006*, ser. Lecture Notes in Comput. Sci., R. S.-N. Lynn Batten, Ed., vol. 4058. Springer, 2006, pp. 283–294.

[57] M. P. C. Fossorier, K. Kobara, and H. Imai, "Modeling bit flipping decoding based on nonorthogonal check sums with application to iter-

ative decoding attack of McEliece cryptosystem," *IEEE Trans. Inform. Theory*, vol. 53, no. 1, pp. 402–411, 2007.

[58] T. Debris-Alazard and J. P. Tillich, "Statistical decoding," in *2017 IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 1798–1802.

[59] E. Petrank and R. Roth, "Is code equivalence easy to decide?" *IEEE Trans. Inform. Theory*, vol. 43, no. 5, pp. 1602–1604, 1997.

[60] N. Sendrier, "Finding the permutation between equivalent linear codes: The support splitting algorithm," *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1193–1203, 2000.

[61] L. Minder and A. Shokrollahi, "Cryptanalysis of the Sidelnikov cryptosystem," in *Advances in Cryptology - EUROCRYPT 2007*, ser. Lecture Notes in Comput. Sci., vol. 4515, Barcelona, Spain, 2007, pp. 347–360.

[62] A. Otmani, J.-P. Tillich, and L. Dallot, "Cryptanalysis of McEliece cryptosystem based on quasi-cyclic LDPC codes," in *Proceedings of First International Conference on Symbolic Computation and Cryptography*. Beijing, China: LMIB Beihang University, Apr. 28-30 2008, pp. 69–81.

[63] V. Dragoi and H. T. Kalachi, "Cryptanalysis of a public key encryption scheme based on qc-ldpc and qc-mdpc codes," *IEEE Communications Letters*, vol. PP, no. 99, pp. 1–1, 2017.

[64] A. Couvreur, I. Márquez-Corbella, and R. Pellikaan, "A polynomial time attack against algebraic geometry code based public key cryptosystems," in *Proc. IEEE Int. Symposium Inf. Theory - ISIT 2014*, Jun. 2014, pp. 1446–1450.

[65] P. Loidreau and N. Sendrier, "Weak keys in the McEliece public-key cryptosystem," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 1207–1211, 2001.

[66] M. Bardet, V. Dragoi, J. Luque, and A. Otmani, "Weak keys for the quasi-cyclic MDPC public key encryption scheme," in *Progress in Cryptology - AFRICACRYPT 2016 - 8th International Conference on Cryptology in Africa, Fes, Morocco, April 13-15, 2016, Proceedings*, 2016, pp. 346–367. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-31517-1_18

[67] F. Strenzke, E. Tews, H. G. Molter, R. Overbeck, and A. Shoufan, "Side channels in the McEliece PKC," in *Post-Quantum Cryptography 2008*, ser. Lecture Notes in Comput. Sci., J. Buchmann and J. Ding, Eds., vol. 5299. Springer, 2008, pp. 216–229.

[68] S. Heyse, "Post-quantum cryptography: Implementing alternative public key schemes on embedded devices," Ph.D. dissertation, Ruhr-Universitt Bochum, October 2013.

[69] F. Strenzke, "Efficiency and implementation security of code-based cryptosystems," Ph.D. dissertation, Technische Universität, Darmstadt, 2013.

[70] I. von Maurich, "Efficient implementation of code- and hash-based cryptography," Ph.D. dissertation, Ruhr-Universitt Bochum, October 2016. [Online]. Available: https://d-nb.info/1137379871/34

[71] N. Patterson, "The algebraic decoding of Goppa codes," *IEEE Trans. Inform. Theory*, vol. 21, no. 2, pp. 203–207, 1975.

[72] M. Petrvalský, T. Richmond, M. Drutarovský, P.-L. Cayrel, and V. Fischer, "Differential power analysis attack on the secure bit permutation in the McEliece cryptosystem," *RadioElektronika 2016*, pp. 132–137, April 2016. [Online]. Available: http://ieeexplore.ieee.org/xpl/abstractKeywords.jsp?arnumber=7477382&abstractAccess=no&userType=inst

[73] S. Heyse, A. Moradi, and C. Paar, "Practical power analysis attacks on software implementations of McEliece," in *Post-Quantum Cryptography 2010*, ser. Lecture Notes in Comput. Sci., N. Sendrier, Ed., vol. 6061. Springer, 2010, pp. 108–125.

[74] F. Strenzke, "Timing attacks against the syndrome inversion in code-based cryptosystems," in *Post-Quantum Cryptography 2013*, ser. Lecture Notes in Comput. Sci., vol. 7932. Limoges, France: Springer, Jun. 2013, pp. 217–230. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-38616-9_15

[75] A. Shoufan, F. Strenzke, H. G. Molter, and M. Stöttinger, "A timing attack against patterson algorithm in the McEliece PKC," in *Information, Security and Cryptology - ICISC 2009, 12th International Conference*, ser. Lecture Notes in Comput. Sci., vol. 5984. Seoul, Korea: Springer, Dec. 2010, pp. 161–175. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-14423-3_12

[76] F. Strenzke, "A Timing Attack against the Secret Permutation in the McEliece PKC," in *Post-Quantum Cryptography 2010*, ser. Lecture Notes in Comput. Sci., N. Sendrier, Ed., vol. 6061. Springer, 2010, pp. 95–107.

[77] R. Avanzi, S. Hoerder, D. Page, and M. Tunstall, "Side-channel attacks on the McEliece and Niederreiter public-key cryptosystems," *J. Cryptographic Engineering*, vol. 1, no. 4, pp. 271–281, 2011. [Online]. Available: http://dx.doi.org/10.1007/s13389-011-0024-9

[78] C. Chen, T. Eisenbarth, I. von Maurich, and R. Steinwandt, "Differential power analysis of a McEliece cryptosystem," in *Applied Cryptography and Network Security (ACNS)*, ser. Lecture Notes in Computer Science (LNCS), T. Malkin, V. Kolesnikov, A. B. Lewko, and M. Polychronakis, Eds. Springer International Publishing, 2015, vol. 9092, pp. 538–556. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-28166-7_26

[79] ——, "Masking large keys in hardware: A masked implementation of McEliece," *Selected Areas in Cryptography (SAC 2015)*, vol. 9566, pp. 293–309, September 2016. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-31301-6_18

[80] ——, "Horizontal and vertical side channel analysis of a McEliece cryptosystem," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 6, pp. 1093–1105, June 2016. [Online]. Available: http://ieeexplore.ieee.org/abstract/document/7360160/

[81] M. Rossi, M. Hamburg, M. Hutter, and M. E. Marson, "A side-channel assisted cryptanalytic attack against qcbits," in *Cryptographic Hardware and Embedded Systems – CHES 2017*, W. Fischer and N. Homma, Eds. Cham: Springer International Publishing, 2017, pp. 3–23.

[82] P.-L. Cayrel and P. Dusart, "McEliece/Niederreiter PKC: Sensitivity to fault injection," in *5th International Conference on Future Information Technology (FutureTech 2010)*, May 2010, pp. 1–6.

[83] H. G. Molter, M. Stöttinger, A. Shoufan, and F. Strenzke, "A simple power analysis attack on a McEliece cryptoprocessor," *Journal Cryptographic Engineering*, vol. 1, no. 1, pp. 29–36, 2011.

[84] F. Strenzke, "Message-aimed side channel and fault attacks against public key cryptosystems with homomorphic properties," *Journal of Cryptographic Engineering*, vol. 1, no. 4, pp. 283–292, 2011. [Online]. Available: http://dx.doi.org/10.1007/s13389-011-0020-0

[85] I. von Maurich and T. Güneysu, "Towards side-channel resistant implementations of QC-MDPC McEliece encryption on constrained devices," in *Post-Quantum Cryptography*, ser. Lecture Notes in Computer Science (LNCS), M. Mosca, Ed. Springer International Publishing, October 2014, vol. 8772, pp. 266–282. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-11659-4_16

[86] M. Petrvalský, T. Richmond, M. Drutarovský, P.-L. Cayrel, and V. Fischer, "Countermeasure against the SPA attack on an embedded McEliece cryptosystem," in *Radioelektronika (RADIOELEKTRONIKA), 2015 25th International Conference*. IEEE, April 2015, pp. 462–466. [Online]. Available: http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=7129055&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D7129055

[87] D. Bucerzan, P.-L. Cayrel, V. Dragoi, and T. Richmond, "Improved timing attacks against the secret permutation in the McEliece PKC," *International Journal of Computers Communications & Control*, vol. 12, no. 1, pp. 7–25, 2016.

APPENDIX

We summarize to the best of our knowledge in Table I all side-channel attacks in code-based cryptography.

TABLE I
STATE-OF-THE-ART IN SCAS

Title	Author(s)	Year	Ref.	Type of attacks
Side channels in the McEliece PKC	Strenzke Tews Molter Overbeck Shoufan	2008	[67]	TA [1]
A Timing Attack against Patterson Algorithm in the McEliece PKC	Shoufan Strenzke Molter Stöttinger	2009	[75]	TA error weight in EEA
A Timing Attack against the Secret Permutation in the McEliece PKC	Strenzke	2010	[**?**]	TA permutation matrix in EEA
Practical Power Analysis Attacks on Software Implementations of McEliece	Heyse Moradi Paar	2010	[73]	PA [2] private key
McEliece/Niederreiter PKC: sensitivity to fault injection	Cayrel Dusart	2010	[82]	FI [3]
Side-Channel Attacks on the McEliece and Niederreiter Public-Key Cryptosystems	Avanzi Hoerder Page Tunstall	2011	[77]	TA improvement of [67]
A simple power analysis attack on a McEliece cryptoprocessor	Molter Stöttinger Shoufan Strenzke	2011	[83]	SPA [4] error weight in EEA
Message-aimed side channel and fault attacks against public-key cryptosystems with homomorphic properties	Strenzke	2011	[84]	FI, TA message
Timing Attacks against the Syndrome Inversion in Code-based Cryptosystems	Strenzke	2013	[74]	TA Goppa polynomial
Towards Side-Channel Resistant Implementations of QC-MDPC McEliece Encryption on Constrained Devices	von Maurich Güneysu	2014	[85]	(TA,) (S)PA message (in Enc) private key (in Dec)
Differential Power Analysis of a McEliece Cryptosystem	Chen Eisenbarth von Maurich Steinwandt	2015	[78]	DPA [5] private key
Countermeasure against the SPA Attack on an Embedded McEliece Cryptosystem	Petrvalský Richmond Drutarovský Cayrel Fischer	2015	[86]	SPA private key attack of the [67] countermeasure
Differential Power Analysis of a McEliece Cryptosystem	Petrvalský Richmond Drutarovský Cayrel Fischer	2016	[72]	DPA private key improvement of [86]
Improved Timing Attacks against the Secret Permutation in the McEliece PKC	Bucerzan Cayrel Dragoi Richmond	2017	[87]	TA improvement of [76], [74]
A Side-Channel Assisted Cryptanalytic Attack Against QcBits	Rossi Hamburg Hutter Marson	2017	[81]	DPA private key

[1]TA : Timing Attack
[2]PA : Power Attack
[3]FI : Fault Injection
[4]SPA : Simple Power Analysis
[5]DPA : Differential Power Analysis