Weak Keys for the Quasi-Cyclic MDPC Public Key Encryption Scheme

Magali Bardet(✉), Vlad Dragoi, Jean-Gabriel Luque, and Ayoub Otmani

Normandie Univ, France; UR, LITIS, 76821 Mont-saint-aignan, France
{magali.bardet,vlad.dragoi1,jean-gabriel.luque,
ayoub.otmani}@univ-rouen.fr

Abstract. We analyze a new key recovery attack against the Quasi-Cyclic MDPC McEliece scheme. Retrieving the secret key from the public data is usually tackled down using exponential time algorithms aiming to recover minimum weight codewords and thus constructing an equivalent code. We use here a different approach and give under certain hypothesis an algorithm that is able to solve a key equation relating the public key to the private key. We relate this equation to a well known problem the *Rational Reconstruction Problem* and therefore propose a natural solution based on the extended Euclidean algorithm. All private keys satisfying the hypothesis are declared weak keys. In the same time we give a precise number of weak keys and extend our analysis by considering all possible cyclic shifts on the private keys. This task is accomplished using combinatorial objects like Lyndon words. We improve our approach by using a generalization of the Frobenius action which enables to increase the proportion of weak keys. Lastly, we implement the attack and give the probability to draw a weak key for all the security parameters proposed by the designers of the scheme.

Keywords: Quasi-cyclic MDPC codes · McEliece scheme · Rational reconstruction problem · Extended euclidean algorithm

1 Introduction

Moderate Density Parity Check (MDPC) codes were introduced in [MTSB12] in order to propose a public-key encryption scheme following McEliece's general approach [McE78]. These codes can be viewed as Low Density Parity Check (LDPC) codes where the parity-check matrices defining them have higher density. LDPC codes are classically constructed from matrices with constant row weights whereas the codes chosen in [MTSB12] have row weights $O(\sqrt{n \log n})$ assuming n is the length. They can be decoded likewise with Gallager's bit-flipping decoding algorithm. Even if using MDPC codes comes at the cost of a degraded error-correction compared to standard LDPC codes, it is still possible to obtain a probability of decoding failure below an acceptable threshold. Furthermore, because of the presence of low-weight codewords, LDPC codes are vulnerable to key recovery attacks based on Information Set Decoding algorithms

D. Pointcheval et al. (Eds.): AFRICACRYPT 2016, LNCS 9646, pp. 346–367, 2016.
DOI: 10.1007/978-3-319-31517-1_18

(see for instance [HS13], while MDPC codes are purposely designed to resist to such attacks. MDPC codes tend to become a serious choice in cryptography because they display the interesting feature of being less structured than codes that are traditionally encountered in code-based cryptography.

In this work we consider the quasi-cyclic variant of MDPC (QC-MDPC) codes, and instead of searching for relatively small weight codewords, we try to solve an equation relating the public polynomial (public data) to secret polynomials (private key). This equation is related to a well-known problem called the *rational reconstruction problem*, which can be solved for instance by the extended Euclidean algorithm (EEA). Solving this equation, which would give a trapdoor to the corresponding scheme, is expected to be hard in general. Nevertheless, in some cases, the solutions are rather easy to compute. We will call this type of (secret) configurations *weak keys* because they can be recovered efficiently from public data.

The main advantage of our technique is the low complexity of the algorithms that are able to check whether a private key is weak. If the original extended Euclidean algorithm is used then the time complexity is quadratic $O(p^2)$ if p is the length of the input. The first optimizations were proposed by Lehmer [Leh38] in 1938 where the constant factor was improved but the complexity was still quadratic. The first sub-quadratic algorithm was proposed in 1970 by Knuth [Knu71] with complexity $O(p(\log p)^5 \log\log p)$ and shortly after revisited by Schönhage in 1971 [Sch71] who obtained a better complexity $O(p(\log p)^2 \log\log p)$. The Least-Significant-Bit version of the Knuth-Schönhage algorithm is due to Stehlé and Zimmermann in 2004 [SZ04]. Even though the time complexity of this algorithm is not improved the description and the proof of their algorithm is significantly simpler in this case. The average behaviour was studied in [LV06,LV08,CCD+09]. Throughout the paper we call weak keys all pairs of private keys that can be recovered using the EEA algorithm from public data. We extend the collection of weak keys thanks to a group action that preserves the key equation. This permits to consider rather a *weak orbit* whenever the orbit under the action of the group contains at least one weak key.

The main contribution of this paper is to provide a fine analysis of the probability of weak keys and weak orbits for the QC-MDPC scheme, under two different actions. Let p be a prime number and consider a random $(2p, p, \omega)$-QC-MDPC code over $\mathbb{F}_2$ (a precise definition will be given in Sect. 2). Such a code is given by two vectors from $\mathbb{F}_2^p$ with a total Hamming weight ω. A rough estimate shows that, when used in a McEliece scheme, the resulting key is vulnerable to the EEA if the non-zero coefficients are all located at the same block. The probability of getting this configuration is $\frac{\binom{p}{\omega}}{\binom{2p}{\omega}}$. In the article we compute the asymptotic equivalence for the suggested range of parameters in [MTSB12],

$$\omega = \sqrt{2cp\log p}(1 + O(1)) \text{ and } p \to \infty. \tag{1}$$

In this case the probability is equivalent to $p^{-c/2}2^{-\omega}$.

In Sect. 4 we compute the exact proportion of weak keys and show that it is asymptotically ω times the previous estimate for the conditions in (1):

$$\frac{\omega\binom{p+1}{\omega}}{\binom{2p}{\omega}-(-1)^{\omega/2}\binom{p}{\omega/2}} = \frac{\omega}{p^{c/2}2^{\omega}}\left(1+O\left(\sqrt{\log^3 p/p}\right)\right). \tag{2}$$

We remark that the cyclic structure of the code defines a natural group action of $(\mathbb{Z}_p,+)$ over the set of public keys. If the coset of a private key contains a weak key, then it is possible to recover the private key by applying EEA to the shifted public key.

To count explicitly the number of weak orbits, we link orbits to Lyndon words and show that counting weak orbits is equivalent to counting Lyndon words with a fixed longuest run value (see Sect. 5 for precise definitions). In [GR61], Gilbert and Riordan count Lyndon words of length p and weight ω. We extend their results and give in Theorem 1 a formula for the number of Lyndon words of length p, weight ω and longuest run less than or equal to k.

This technique permits to increase the quantity of weak keys by a multiplicative factor equal to ω^3 for the conditions in (1), that is to say

$$\omega p^2\frac{\binom{p-1}{\omega-2}}{\binom{2p}{\omega}+(-1)^{\omega/2+1}\binom{p}{\omega/2}} = \frac{\omega^3}{p^{c/2}2^{\omega}}\left(1+O\left(\sqrt{\log^3 p/p}\right)\right). \tag{3}$$

In Sect. 6 we define another action of $(\mathbb{Z}_p^*,\times)$ over the set of public keys, that is compatible with the action of $(\mathbb{Z}_p,+)$. We explain how to apply EEA to every element of an orbit under both actions, and show that the attack will succeed if there exists at least one weak key in the orbit of a public key.

We prove that the quantity of keys our algorithm is able to attack is increased using this technique, by a multiplicative factor that is linear in the block length for the conditions in (1), that is to say

$$\frac{\omega p^3\binom{p-1}{\omega-2}}{\binom{2p}{\omega}+(-1)^{\omega/2+1}\binom{p}{\omega/2}} = \frac{\omega^3 p}{p^{c/2}2^{\omega}}\left(1+O\left(\sqrt{\log^3 p/p}\right)\right). \tag{4}$$

In Sect. 7 we give numerical values for the proportion of weak keys for all the security parameters suggested by the designers of the cryptosystem. Finaly in Sect. 8 we give experimental timings for our attack.

Due to space constraints, many proofs are just sketched. The full version of this paper is available on the `arXiv.org` preprint server.

2 QC-MDPC Encryption Scheme

We present here the most relevant material for describing the public-key encryption scheme [MTSB12]. We focus on the quasi-cyclic variant of [MTSB12] which is defined through circulant matrices. Throughout the paper, the weight of a vector or a polynomial refers to the Hamming weight and is denoted by $\|\ \|$.

2.1 QC-MDPC Codes and the Algebra of Circulant Matrices

Definition 1 (Moderate Density Parity Check codes). *A (n, r, w)-code is a linear code defined by a $r \times n$ parity-check matrix ($r < n$) where each row has weight w. A* Moderate Density Parity-Check (MDPC) *code is a (n, r, w)-code with $w = O\left(\sqrt{n \log n}\right)$, when $n \to \infty$.*

Definition 2. *A* circulant *matrix $\boldsymbol{M}$ of order p is a $p \times p$ matrix obtained by cyclically right shifting its first row $\boldsymbol{m} = (m_0, m_1, \ldots, m_{p-1})$.*

Any circulant matrix is thus completely described by its first row. A circulant matrix is also obtained by cyclically down shifting its first column. It is well-known that the matrix operations of addition and multiplication preserve the circulant structure of matrices.

Proposition 1. *[Dav79] The algebra of $p \times p$ circulant matrices with entries in a field $\mathbb{K}$ denoted by $\left(\mathfrak{C}_p(\mathbb{K}), +, \times\right)$ is isomorphic to the polynomial algebra $\left(\mathbb{K}[x]/(x^p - 1), +, \cdot\right)$ through the mapping*

$$\begin{aligned} \mathfrak{C}_p(\mathbb{K}) &\longrightarrow \mathbb{K}[x]/(x^p - 1) \\ \boldsymbol{M} &\longmapsto m(x) = \sum_{i=0}^{p-1} m_i x^i \pmod{x^p - 1}. \end{aligned}$$

Corollary 1. *A $p \times p$ circulant matrix $\boldsymbol{M}$ defined by m is invertible if and only if $m(x)$ is coprime to $x^p - 1$. In particular, the weight of m is necessarily odd.*

The algebra of circulant matrices enables to define the algebra of block matrices where the blocks are circulant. Any such matrix can be viewed as a matrix with entries in $\mathbb{K}[x]/(x^p - 1)$. This will define quasi-cyclic codes which represent the unique focus of this article.

Definition 3. *A* Quasi-Cyclic MDPC (QC-MDPC) code *is a MDPC code defined by a block parity-check matrix where each block is a circulant matrix.*

We now have defined all objects that permit to fully describe the scheme [MTSB12]. We will focus however, exclusively on the key generation algorithm since it is the only compound of the scheme that is of interest in this paper.

2.2 QC-MDPC Public-Key Encryption Scheme

The private key is a parity check matrix $\boldsymbol{H}$ of an (n, r, w) QC-MDPC code where $n = n_0 p$ and $r = p$ for some non-negative integer n_0. There exist therefore $p \times p$ circulant matrices $\boldsymbol{H}_1, \ldots, \boldsymbol{H}_{n_0}$ such that

$$\boldsymbol{H} = \left(\boldsymbol{H}_1 \; \boldsymbol{H}_2 \cdots \boldsymbol{H}_{n_0}\right). \tag{5}$$

This private key is obtained by taking at random the first row of $\boldsymbol{H}$ until $\boldsymbol{H}_{n_0}$ is invertible. The public key is the block parity-check matrix $\boldsymbol{F} \stackrel{\text{def}}{=} \boldsymbol{H}_{n_0}^{-1}\boldsymbol{H}$, or

$$\boldsymbol{F} = \left(\boldsymbol{H}_{n_0}^{-1}\boldsymbol{H}_1 \cdots \boldsymbol{H}_{n_0}^{-1}\boldsymbol{H}_{n_0-1} \; \boldsymbol{I}_p\right) \stackrel{\text{def}}{=} \left(\boldsymbol{F}_1 \cdots \boldsymbol{F}_{n_0-1} \; \boldsymbol{I}_p\right). \tag{6}$$

Using the isomorphism defined in Proposition 1, the private and public keys are fully described by the sequences $h_1, \ldots, h_{n_0}$ and $f_1, \ldots, f_{n_0-1}$ of polynomials in $\mathbb{K}[x]/(x^p - 1)$ such that for all $i \in \{1, \ldots, n_0 - 1\}$,

$$f_i = \frac{h_i}{h_{n_0}} \pmod{x^p - 1}. \tag{7}$$

The secret polynomials are taken so that $\sum_{i=1}^{n_0} \|h_i\| = w$.

Hence, the key generation of QC-MDPC scheme can be summarised as follows

- **Private key.** Pick at random $h_1, \ldots, h_{n_0}$ from $\mathbb{K}[x]/(x^p - 1)$ such that $\sum_{i=1}^{n_0} \|h_i\| = w$ and h_{n_0} is prime with $x^p - 1$.
- **Public key.** $f_1, \ldots, f_{n_0-1}$ where $f_i = \dfrac{h_i}{h_{n_0}} \pmod{x^p - 1}$.

2.3 Discussion on the Choice of the Parameters

Since [MTSB12] considers solely *binary* matrices, we assume from now $\mathbb{K} = \mathbb{F}_2$. Furthermore, the weights $\|h_i\|$ are "smoothly" distributed and p is always a prime number for security reasons. During the Key Generation step one must randomly choose the polynomials h_i until at least one of them is invertible. So we might expect, for security reasons, that the designers selected those parameters for which the set of invertible polynomials in the polynomial algebra $\mathbb{K}[x]/(x^p-1)$ is the largest possible. Using a ring isomorphism we give the number of invertible polynomials and thus show which are the proper parameters to be selected.

Proposition 2. *Let p be a prime number and assume $(x-1)\prod_{i=1}^{d} g_i(x)$ is the decomposition of $x^p - 1$ into irreducible polynomials over $\mathbb{F}_2[x]$ for some $d \geqslant 1$ then $\deg g_i = \frac{p-1}{d}$ for all $i \in \{1, \ldots, d\}$. In particular, the number of invertible polynomials in $\mathbb{F}_2[x]/(x^p - 1)$ equals $\left(2^{(p-1)/d} - 1\right)^d$.*

For the choice of secure parameters, it is recommended to choose p so that the Folding attack [Gen01, Loi01, FOP+14] is inefficient. The most favorable situation is when d is as small as possible, for instance $d = O(1)$ when p tends to infinity. The designers of the scheme considered this option since all the parameters respect this condition. Hence, the number of invertible polynomials in $\mathbb{F}_2[x]/(x^p - 1)$ tends to be 2^{p-1} which is exactly the number of polynomials with an odd Hamming weight. So the probability of choosing an invertible polynomial from the set of polynomials with an odd Hamming weight is $\left(1 - 2^{-(p-1)/d}\right)^d$ which tends to 1 when $d = O(1)$. One very interesting case is when $d = 1$, since it seems to be the most secure choice for the cryptosystem. In Sect. 4 we investigate this particular case.

3 Rational Reconstruction Problem

We are interested in a *key-recovery under a chosen plaintext attack*. When applied on a (pn_0, p, w) QC-MDPC scheme whose public key is the sequence

of polynomials $(f_1, \ldots, f_{n_0-1})$, the attack can be reformulated as the problem of finding $(h_1, \ldots, h_{n_0})$ satisfying

$$f_i = \frac{h_i}{h_{n_0}} \pmod{x^p - 1} \text{ and } \sum_{i=1}^{n_0} \|h_i\| \leqslant w. \tag{8}$$

This problem can be tackled by applying classical techniques based on exponential algorithms seeking low-weight codewords. It can also be recast as the problem of solving the *rational reconstruction problem* that is described in full details in Sect. 3. The extended Euclidean algorithm solves (8) when there exists an integer $t > 0$ such that $\deg h_i < t \leqslant p$ and $\deg h_{n_0} \leqslant p - t$. Actually, (8) is a special case of a well-known problem called the *Rational Reconstruction problem*. It will be used in Sect. 4 as a general framework within which it is possible to perform a polynomial time key recovery attack.

Remark 1. Because of the bit-flipping decoding algorithm for MDPC codes, an attacker does not necessarily have to find the exact same secret polynomials for decrypting any ciphertext. Indeed, *any* sequence of polynomials satisfying the conditions (8) will lead to an efficient decoding of any ciphertext. It also means that there might exist several equivalent secret keys for a single QC-MDPC scheme.

Definition 4 (Rational reconstruction). *Let g and f be polynomials in $\mathbb{K}[x]$ where $\mathbb{K}$ is a field such that $0 < \deg f < \deg g$. For a given integer r satisfying $1 \leqslant r \leqslant \deg g$, the* rational reconstruction *of f modulo g consists in finding φ and ψ in $\mathbb{K}[x]$ such that $\gcd(\varphi, g) = 1$, $\deg \psi < r$ and $\deg \varphi \leqslant \deg g - r$ and satisfying*

$$\frac{\psi}{\varphi} = f \pmod{g}. \tag{RR}$$

Remark 2. When $g = x^p$ then we rather speak of Padé approximation.

Note that if (RR) has a solution (φ, ψ) then the quotient ψ/φ is unique. Furthermore if $(\varphi, \psi) \in \mathbb{K}[x]^2$ is a solution of the problem (RR), then it is also a solution to the following problem.

Definition 5. *Let $\mathbb{K}$ be a field, g be a polynomial in $\mathbb{K}[x]$ of degree $p > 0$ and f be in $\mathbb{K}[x]$ of degree $< p$. For a given r with $1 \leqslant r \leqslant p$, the* (SRR) *problem consists in finding ψ and φ in $\mathbb{K}[x]$ such that $(\varphi, \psi) \neq (0, 0)$ and*

$$\varphi f = \psi \pmod{g} \quad \textit{with } \deg \psi < r \textit{ and } \deg \varphi \leqslant p - r. \tag{SRR}$$

Clearly, any solution to (SRR) is solution to (RR) if and only if $\gcd(\varphi, g) = 1$. Moreover, (SRR) always has a non-trivial solution since recovering φ and ψ can be done by solving a linear system of p equations with $r + (p - r + 1) = p + 1$ unknowns representing the coefficients of φ and ψ.

A very efficient way to solve (RR) is to apply the Extended Euclidean Algorithm (EEA) to (f, g). Recall that if we denote by $(\varphi_i, \delta_i, \psi_i)$, with $i \geqslant 0$, the polynomials obtained at the i-th step of $\mathsf{EEA}(f, g)$ then we have $\psi_0 \stackrel{\text{def}}{=} g$, $\psi_1 \stackrel{\text{def}}{=} f$ and for all $i \geqslant 0$:

$$\begin{cases} \psi_i = Q_{i+1}\psi_{i+1} + \psi_{i+2} & \text{with } 0 \leqslant \deg \psi_{i+2} < \deg \psi_{i+1}, \\ \\ \psi_i = \varphi_i f + \delta_i g & \text{with } (\varphi_0, \varphi_1) \stackrel{\text{def}}{=} (0,1) \text{ and } (\delta_0, \delta_1) \stackrel{\text{def}}{=} (1,0). \end{cases}$$

We also have the relations $\varphi_{i+2} = -Q_{i+1}\varphi_{i+1} + \varphi_i$ and $\delta_{i+2} = -Q_{i+1}\delta_{i+1} + \delta_i$. We are now able to prove that this approach provides a non-trivial solution. We require the following proposition.

Proposition 3. *At each step $i \geqslant 0$ of $\mathsf{EEA}(f, g)$ it holds that*

$$\deg \varphi_{i+1} = p - \deg \psi_i. \tag{9}$$

The following proposition characterises a solution to (RR) when it exists.

Proposition 4. *Let j be the smallest integer such that $\deg(\psi_j) < r$ then (φ_j, ψ_j) is a non-trivial solution to (SRR). Furthermore, if (φ, ψ) is a solution to (RR) then there exists λ in $\mathbb{K}\backslash\{0\}$ such that $\varphi = \lambda\varphi_j$ and $\psi = \lambda\psi_j$.*

4 Weak Keys

This section is devoted to the identification of private keys $h_1, \ldots, h_{n_0}$ that can be recovered from public key $f_1, \ldots, f_{n_0-1}$ by means of the extended Euclidean algorithm. Since $f_i = \dfrac{h_i}{h_{n_0}} \pmod{x^p - 1}$, the idea of our attack is to start by finding a rational reconstruction of f_1 modulo $x^p - 1$. At each step t of $\mathsf{EEA}(f_1, x^p - 1)$, the attacker checks if the ongoing computed polynomials denoted by $(\psi_t^{(1)}, \varphi_t^{(1)})$ where $\psi_t^{(1)} = f_1\varphi_t^{(1)}$ satisfy the inequality

$$\|\varphi_t^{(1)}\| + \sum_{i=1}^{n_0-1} \|f_i\varphi_t^{(1)}\| \leqslant w. \tag{10}$$

If such a solution is found then by Proposition 4 we have found (equivalent) secret polynomials. Otherwise, the attacker performs the same attack to f_2 instead of f_1. If this fails again the attack goes on with the other polynomials $f_3, \ldots, f_{n_0-1}$. The main problem is to estimate precisely the number of keys that can be recovered with this technique.

We restrict the study to the case of two blocks $(2p, p, \omega)$ QC-MDPC scheme that is to say $n_0 = 2$. Nevertheless all our results can be extended to $n_0 > 2$. Let p be a prime number and ω an even integer with $1 < \omega < p$. Let $(\omega_1, \omega_2) \in \mathbb{N}^2$ be odd integers such that $\omega_1 + \omega_2 = \omega$. We define the set of *private pairs with fixed weights* by

$$\mathcal{P}_{\omega_1,\omega_2} = \left\{ (h_1, h_2) \in (\mathbb{K}[x]/(x^p - 1))^2 \mid \|h_i\| = \omega_i \text{ and } \omega_i \text{ odd} \right\},$$

and the set of all *private pairs* of a $(2p, p, \omega)$ QC-MDPC scheme by $\mathcal{P}_\omega = \bigcup_{\omega_1+\omega_2=\omega} \mathcal{P}_{\omega_1,\omega_2}$.

Private pairs that can be recovered using the extended Euclidiean algorithm are declared weak pairs.

Definition 6. *A pair* $(h_1, h_2) \in \mathcal{P}_\omega$ *is called a* weak pair *if*

$$\deg h_1 + \deg h_2 < p. \tag{11}$$

The set of weak pairs is denoted by $\mathcal{W}_\omega = \{(h_1, h_2) \in \mathcal{P}_\omega \mid \deg h_1 + \deg h_2 < p\}$. *Similarly,* $\mathcal{W}_{\omega_1,\omega_2}$ *is defined as* $\mathcal{W}_\omega \cap \mathcal{P}_{\omega_1,\omega_2}$.

Remark 3. It is important to notice that *true* collection of private keys of a general $(2p, p, \omega)$ QC-MDPC scheme is actually the set ${\mathcal{P}_\omega}^* = \bigcup_{\omega_1+\omega_2=\omega} \mathcal{P}^*_{\omega_1,\omega_2}$ where

$$\mathcal{P}^*_{\omega_1,\omega_2} = \left\{ (h_1, h_2) \in \mathcal{P}_{\omega_1,\omega_2} \mid \gcd(h_2, x^p - 1) = 1 \right\}.$$

But in order to simplify our analysis, we will only count weak pairs (h_1, h_2) and not weak keys for a $(2p, p, \omega)$ QC-MDPC scheme. This approximation is also justified by the fact we know from Sect. 2.3 that

$$\lim_{p\to\infty} \left(\sum_{\omega=2}^{2p} |{\mathcal{P}_\omega}^*| \right) \Big/ \left(\sum_{\omega=2}^{2p} |\mathcal{P}_\omega| \right) = 1.$$

Remark also that there is one case where the two sets are equal. Indeed if $x^p - 1 = (x-1)\prod_{i=1}^{d} g_i(x)$ is the factorization of $x^p - 1$ into irreducible factors (see Sect. 2.3 for more details) then when $d = 1$ we have $\mathcal{P}_{\omega_1,\omega_2} = \mathcal{P}^*_{\omega_1,\omega_2}$ and $\mathcal{P}_\omega = {\mathcal{P}_\omega}^*$. For several reasons we consider this case in the article. The first one is that this is the strongest possible case for the QC-MDPC scheme since it avoids folding-type attacks. The second reason is that the number of private keys reaches its maximum since all todd weight polynomials are invertible.

Proposition 5.

$$|\mathcal{W}_{\omega_1,\omega_2}| = \tbinom{p+1}{\omega_1+\omega_2} \quad \textit{and} \quad |\mathcal{W}_\omega| = \frac{\omega}{2}\binom{p+1}{\omega}. \tag{12}$$

$$|\mathcal{P}_{\omega_1,\omega_2}| = \binom{p}{\omega_1}\binom{p}{\omega_2} \quad \textit{and} \quad |\mathcal{P}_\omega| = \frac{1}{2}\left(\binom{2p}{\omega} - (-1)^{\frac{\omega}{2}}\binom{p}{\frac{\omega}{2}}\right). \tag{13}$$

The asymptotic expansion when $\frac{\omega_i^2}{2p} = c_i + O(\frac{1}{\sqrt{p}})$ *is*

$$\frac{|\mathcal{W}_{\omega_1,\omega_2}|}{|\mathcal{P}_{\omega_1,\omega_2}|} = \sqrt{2\pi\alpha(1-\alpha)}e^{-2\sqrt{c_1 c_2}}\omega^{\frac{1}{2}}2^{-\omega H(\alpha)}\left(1 + O(1/\sqrt{p})\right)$$

where $\alpha = 1/(1+\sqrt{c_2/c_1})$ *and* $H(\alpha) = -\alpha\log_2\alpha - (1-\alpha)\log_2(1-\alpha)$ *is the entropy function. The asymptotic expansion for* $\frac{\omega_i^2}{2p} = c_i \log p + O(\sqrt{\log p/p})$ *is*

$$\frac{|\mathcal{W}_{\omega_1,\omega_2}|}{|\mathcal{P}_{\omega_1,\omega_2}|} = \sqrt{2\pi\alpha(1-\alpha)}p^{-2\sqrt{c_1c_2}}\omega^{\frac{1}{2}}2^{-\omega H(\alpha)}\left(1+O(\sqrt{\log^3 p/p})\right).$$

$$\frac{|\mathcal{W}_\omega|}{|\mathcal{P}_\omega|} = \omega 2^{-\omega} \times \begin{cases} e^{-\frac{c}{2}}\left(1+O(\frac{1}{\sqrt{p}})\right) \; if \frac{\omega^2}{2p} = c + O(\frac{1}{\sqrt{p}}), \\ p^{-\frac{c}{2}}\left(1+O(\sqrt{\frac{\log^3 p}{p}})\right) \; if \frac{\omega^2}{2p} = c\log p + O(\sqrt{\frac{\log p}{p}}). \end{cases}$$

Proof. Let $(h_1,h_2) \in \mathcal{P}_{\omega_1,\omega_2}$. Then h_i has w_i non-zero coefficients, and a degree less than p, hence $|\mathcal{P}_{\omega_1,\omega_2}| = \binom{p}{\omega_1}\binom{p}{\omega_2}$. For $(h_1,h_2) \in \mathcal{W}_{\omega_1,\omega_2}$ we have $\deg(h_1) + \deg(h_2) < p$. If $k = \deg(h_1)$, then h_1 has a leading coefficient x^k and $\omega_1 - 1$ non-zero coefficients between x^0 and x^{k-1}. The number of such polynomials is $\binom{k}{\omega_1-1}$. Furthermore the number of polynomials h_2 with ω_2 non-zero coefficients and $\deg(h_2) < p-k$ equals $\binom{p-k}{\omega_2}$. Using the Gould's formulae [Gou72], we get

$$|\mathcal{W}_{\omega_1,\omega_2}| = \sum_{k=0}^{p-1}\binom{k}{\omega_1-1}\binom{p-k}{\omega-\omega_1} = \binom{p+1}{\omega},$$

$$|\mathcal{P}_\omega| = \sum_{\substack{\omega_1+\omega_2=\omega \\ \omega_i \text{ odd}}}\binom{p}{\omega_1}\binom{p}{\omega_2} = \frac{1}{2}\left[\binom{2p}{\omega} - (-1)^{\frac{\omega}{2}}\binom{p}{\frac{\omega}{2}}\right].$$

As for $\mathcal{W}_\omega$ we obtain:

$$|\mathcal{W}_\omega| = \sum_{\substack{\omega_1+\omega_2=\omega \\ \omega_i \text{ odd}}}\binom{p+1}{\omega} = \binom{p+1}{\omega}\sum_{\substack{\omega_1+\omega_2=\omega \\ \omega_i \text{ odd}}} 1 = \frac{\omega}{2}\binom{p+1}{\omega}.$$

For the asymptotic expansion use the Stirling formula and obtain the results.

Corollary 2. *In particular*

$$\frac{|\mathcal{W}_{\omega/2,\omega/2}|}{|\mathcal{P}_{\omega/2,\omega/2}|} = \frac{\binom{p+1}{\omega}}{\binom{p}{\omega/2}^2},$$

with asymptotic equivalence

$$\frac{|\mathcal{W}_{\omega/2,\omega/2}|}{|\mathcal{P}_{\omega/2,\omega/2}|} \sim \begin{cases} \sqrt{\pi}p^{\frac{1}{4}}e^{-2}2^{\frac{1}{4}-2\sqrt{2p}} \; if\, \omega = 2\sqrt{2p}, \\ \sqrt{\pi}p^{\frac{1}{4}-2}\log^{\frac{1}{4}} p 2^{\frac{1}{4}-2\sqrt{2p\log p}} \; if\, \omega = 2\sqrt{2p\log p}. \end{cases}$$

The number of weak pairs can be easily increased by considering all possible cyclic shifts on the polynomials (h_1,h_2). We formally define the cyclic shift of a polynomial in terms of group action and explain how we extend the weak pairs to weak orbits.

5 Weak Pairs Derived from the Action of $(\mathbb{Z}_p, +)$

Let $f \in \mathbb{F}_2[x]/(x^p-1)$ be a public key, and $(h_1, h_2) \in \mathbb{F}_2[x]/(x^p-1) \times \mathbb{F}_2[x]/(x^p-1)$ the corresponding private key. We have $f = \frac{h_1}{h_2} \mod (x^p-1)$. Now assume that there exists $\alpha_1, \alpha_2 \in \mathbb{Z}_p^2$ such that $(x^{\alpha_1}h_1, x^{\alpha_2}h_2)$ is a weak key, then the public key $x^{\alpha_1-\alpha_2}f = \frac{x^{\alpha_1}h_1}{x^{\alpha_2}h_2}$ can be attacked by EEA, which is equivalent to say that

$$\exists \alpha_1, \alpha_2 \in \mathbb{Z}_p^2 \text{ such that } \deg(x^{\alpha_1}h_1) + \deg(x^{\alpha_2}h_2) < p. \tag{14}$$

Using this idea if our attack does not work on f we repeat it on all p cyclic shifts of f, namely $xf, x^2f, \ldots, x^{p-1}f$. If there is a shift such that the outgoing polynomials satisfy the weight conditions in (10) then we have succesfully recovered (equivalent) secret polynomials by Proposition 4. As in the previous section we want to estimate precisely the number of keys that can be recovered with this technique.

Definition 7. *The additive group $(\mathbb{Z}_p, +)$ acts on the set of polynomials as:*

$$\begin{aligned} \mathbb{Z}_p \times \mathbb{F}_2[x]/(x^p-1) &\longrightarrow \mathbb{F}_2[x]/(x^p-1) \\ (\alpha, h) &\longmapsto x^{\alpha}h. \end{aligned}$$

The orbit of $h \in \mathbb{F}_2[x]/(x^p-1)$ under the action of $(\mathbb{Z}_p, +)$ is denoted by $\mathcal{O}_h$.

Definition 8 (Weak orbit). *The set $\mathcal{O}_{h_1} \times \mathcal{O}_{h_2}$ defined by a private key (h_1, h_2) in $\mathbb{F}_2[x]/(x^p-1)^2$ is called a* weak orbit *if it contains at least one weak key, i.e. satisfies (14).*

Potentially, we would get $p^2\,|\mathcal{W}_\omega|$ such keys. But this statement overestimates the real number of weak pairs since it counts several times the same private keys. Nevertheless it gives a first intuition on the quantity of weak pairs that can be recovered using the rational reconstruction.

Lemma 1. *Let $\overline{h_i} = \min \mathcal{O}_{h_i}$ be the minimum polynomial for the lexicographical order of $h_i \in \mathbb{F}_2[x]/(x^p-1)$. Then the set $\mathcal{O}_{h_1} \times \mathcal{O}_{h_2}$ is a weak orbit if and only if $\deg \overline{h_1} + \deg \overline{h_2} < p$.*

We define the *longest run of zeros of a polynomial* in $\mathbb{F}_2[x]/(x^p-1)$ by the longest sequence of consecutive zero coefficients. We remark that there is a relation connecting the degree of the minimum polynomial and the longest run of zeros. If k_i denotes the longest run of zeros of $h_i \in \mathbb{F}_2[x]/(x^p-1)$ we have that $\deg \overline{h_i} = p - k_i - 1$. Since we have the relation between the degree and the longest run of zeros for the minimal polynomial in the equivalence class we can redefine a weak orbit in terms of longest run:

Proposition 6 (Weak orbit). *The set $\mathcal{O}_{h_1} \times \mathcal{O}_{h_2}$ defined by a private key $(h_1, h_2) \in \mathbb{F}_2[x]/(x^p-1)^2$ is a* weak orbit *if and only if it satisfies the equation:*

$$k_1 + k_2 \geqslant p - 1. \tag{15}$$

At this point we have reduced our key recovery attack to a well-known problem. To count all pairs (h_1, h_2) with the restriction mentioned above, we have to solve another problem: *What is the distribution of the longest run of zeros for the equivalence class of all cyclic shifts of a $\mathbb{K}^p$ vector with fixed Hamming weight?*

Definition 9. *[Lot02] A Lyndon word l is a word satisfying the conditions:*

- *l is a primitive word (i.e. it cannot be written $l = uv$, where u and v commute and $u, v \neq 1$)*
- *l is the smallest element in its conjugacy class for the lexicographical order*

Example 1.

1. Let $\mathcal{O}_{00011} = \{00011, 00110, 01100, 11000, 10001\}$. The Lyndon word here is 00011 since it is the strictly smallest than all the cyclic shifts.
2. Let $\mathcal{O}_{0101} = \{0101, 1010, 0101, 1010\}$. There is no Lyndon word here, since there is no strictly smallest element in the orbit.

An important property is that when p is prime there is a one-to-one mapping between the Lyndon words and the orbits if the weight is different from zero or p. So each equivalence class has p different shifts and the strictly smallest (since it exists) is the Lyndon word.

Theorem 1. *Let p, k, ω be integers, such that $1 \leqslant \omega \leqslant p$ and $k \leqslant p - \omega$. The number of binary Lyndon words with length p, longest run less than or equal to k and weight equal to ω is:*

$$\left|L^{\leqslant k}(p,\omega)\right| = \frac{1}{\omega} \sum_{j \in \mathbb{N}^*,\, j \mid gcd(p,\omega)} \mu(j) \binom{\frac{\omega}{j}}{\frac{p}{j} - \frac{\omega}{j}}_k, \tag{16}$$

where μ is the Möbius function, defined by $\mu(j) = 0$ if j has a squared prime factor, $\mu(j) = 1$ if j is square-free with an even number of prime factors and $\mu(j) = -1$ otherwise. The standard multinomial coefficient $\binom{j}{i}_k$ is defined as the coefficient of x^i in $\left(1 + x + \cdots + x^k\right)^j$.

The full proof of Theorem 1 is given in Appendix A and it uses a bijection between the Lyndon words with some specific properties on two alphabets: the binary alphabet and an $(k+1)$-ary alphabet. Straightforward we obtain:

Corollary 3. *The number of Lyndon words of length p and Hamming weight equal to ω over the binary alphabet (result already found in [GR61] by Gilbert and Riordan) is:*

$$|L(p,\omega)| = \frac{1}{p} \sum_{j \mid \gcd(p,\omega)} \mu(j) \binom{\frac{p}{j}}{\frac{\omega}{j}}. \tag{17}$$

Corollary 4. *When p is prime we have*

$$\left|L^{\leqslant k}(p,\omega)\right| = \frac{1}{\omega} \binom{\omega}{p-\omega}_k \text{ and } |L(p,\omega)| = \frac{1}{p} \binom{p}{\omega}. \tag{18}$$

As we already stated we will consider only the case p prime. Since all the orbits have the same length (p) and each orbit is defined by the corresponding Lyndon word, there is a uniform distribution over the set of Lyndon words when p is prime. So we consider a discrete probability model where the probability space is the set of Lyndon words with length p and weight ω with cardinal $\frac{1}{p}\binom{p}{\omega}$ and the probability of choosing a Lyndon word equals $p/\binom{p}{\omega}$. Furthermore we put a condition on the longest run of each Lyndon word and obtain a different distribution over the same set. In other words we write $L(p,\omega) = \bigcup_{k=\lfloor \frac{p-1}{\omega} \rfloor}^{p-\omega} L^k(p,\omega)$ and denote by $X_{p,\omega}$ a discrete random variable that represents the longest run of zeros of Lyndon words with length p and weight ω. Using Corollary 4 we define:

Definition 10. *The cumulative distribution and mass function for $X_{p,\omega}$ are:*

$$F_{X_{p,\omega}}(k) = \frac{\left|L^{\leqslant k}(p,\omega)\right|}{|L(p,\omega)|} \text{ and } f_{X_{p,\omega}}(k) = \frac{\left|L^{k}(p,\omega)\right|}{|L(p,\omega)|}.$$

Let $Y_{p,\omega_1,\omega_2} = X_{p,\omega_1} + X_{p,\omega_2}$ a discrete random variable that represents the sum of two independent random variables X_{p,ω_1} and X_{p,ω_2}. So the probability of a weak orbit is:

$$P\left(Y_{p,\omega_1,\omega_2} \geqslant p-1\right) = \sum_{k_1+k_2 \geqslant p-1} f_{X_{p,\omega_1}}(k_1) f_{X_{p,\omega_2}}(k_2)$$

As p is prime, using Corollary 4 and Definition 10 we get the exact value:

$$P\left(Y_{p,\omega_1,\omega_2} \geqslant p-1\right) = \sum_{k_1+k_2 \geqslant p-1} \frac{\binom{\omega_1}{p-\omega_1}_{k_1} - \binom{\omega_1}{p-\omega_1}_{k_1-1}}{\binom{p-1}{\omega_1-1}} \frac{\binom{\omega_2}{p-\omega_2}_{k_2} - \binom{\omega_2}{p-\omega_2}_{k_2-1}}{\binom{p-1}{\omega_2-1}} \quad (19)$$

The first case that seems interesting is when each variable has a longest run greater than or equal to half of the wanted quantity $\frac{p-1}{2}$.

Proposition 7. *Let ω_1 and $\omega_2 \geqslant 2$, then we have:*

$$P\left(X_{p,\omega_1} \geqslant \frac{p-1}{2}\right) P\left(X_{p,\omega_2} \geqslant \frac{p-1}{2}\right) = \omega_1\omega_2 \times \frac{\binom{\frac{p-1}{2}}{\omega_1-1}\binom{\frac{p-1}{2}}{\omega_2-1}}{\binom{p-1}{\omega_1-1}\binom{p-1}{\omega_2-1}}, \quad (20)$$

with asymptotic equivalence

$$\omega_1\omega_2 2^{-\omega} \times \begin{cases} e^{-\frac{c_1+c_2}{2}} & if\ \omega_i^2 = c_i p + O(\sqrt{p}), \\ p^{-\frac{c_1+c_2}{2}} & if\ \omega_i^2 = c_i p \log p + O(\sqrt{p \log p}). \end{cases}$$

Proof. We apply the formula for the generalized Pascal-DeMoivre coefficient from [Lot02, BBK08]:

$$\binom{\omega}{p-\omega}_k = \sum_{j=0}^{\lfloor \frac{p-\omega}{k+1} \rfloor} (-1)^j \binom{\omega}{j} \binom{p-j(k+1)-1}{\omega-1}.$$

For asymptotic expansion as before use the Stirling approximation for factorials.

Remark 4. We observe that using the shifts increased the probability of a weak key with a multiplicative factor equal to $\omega^{\frac{3}{2}}$. From Sect. 4 when $\omega_1 = \omega_2 = \frac{\omega}{2}$ we have that

$$P\left(X_{p,\omega_1} \geqslant \frac{p-1}{2}\right)^2 \sim \omega^{\frac{3}{2}} \frac{|\mathcal{W}_{\omega_1,\omega_2}|}{|\mathcal{P}_{\omega_1,\omega_2}|}.$$

We step forward and analyze the probability for a weak orbit in the general case. We remark that if either ω_1 or ω_2 equals 1 then the probability of a weak orbit equals 1. But the interesting analysis is when ω_1 and ω_2 are relatively close and $\omega = O\left(\sqrt{p \log p}\right)$.

Proposition 8. *If $\omega_1 \geqslant \omega_2$ and $\omega_i^2 = 2c_i p \log p + O(\sqrt{p \log p})$ then we have*

$$P(Y_{p,\omega_1,\omega_2} \geqslant p-1) \sim \omega_1 \omega_2 \frac{\binom{p-1}{\omega-2}}{\binom{p-1}{\omega_1-1}\binom{p-1}{\omega_2-1}} \quad \text{when } p \to \infty, \tag{21}$$

with asymptotic equivalence

$$P(Y_{p,\omega_1,\omega_2} \geqslant p-1) \sim \omega^2 \sqrt{2\pi\alpha(1-\alpha)} p^{-2\sqrt{c_1 c_2}} \omega^{\frac{1}{2}} 2^{-\omega H(\alpha)}.$$

where $\alpha = 1/(1+\sqrt{c_2/c_1})$ and $H(\alpha) = -\alpha \log_2 \alpha - (1-\alpha)\log_2(1-\alpha)$

Proof. See Appendix A page 21.

We can easily check that for $\omega_i = \sqrt{c_i p \log p}$ and $c_1 > c_2$ the condition in Proposition 21 is satisfied. Experiments show that if we release the conditions on ω_i the approximation is still sharp. So a deeper investigation of the generalized Pascal-DeMoivre triangles might be used to prove this statement but this is no longer our purpose here.

Corollary 5. *We have the asymptotic equivalences*

$$P(Y_{p,\omega/2,\omega/2} \geqslant p-1) \sim \left(\frac{\frac{\omega}{2}}{\binom{p-1}{\frac{\omega}{2}-1}}\right)^2 \binom{p-1}{\omega-2} \quad \text{when } p \to \infty \text{ and } \omega = o(p),$$

$$P(Y_{p,\omega/2,\omega/2} \geqslant p-1) \sim \sqrt{\pi/2} p^{-\frac{1}{4}} \omega^{\frac{5}{2}} 2^{-\omega} \quad \text{if} \quad \omega_i^2 = \frac{p \log p}{4} + O(\sqrt{\log p/p}).$$

Remark 5. If we recall the results obtained with the first method in Proposition 5 and Corollary 2 we conclude that we gain a multiplicative factor equal to ω^2 using the shifts:

$$P(Y_{p,\omega_1,\omega_2} \geqslant p-1) \sim \omega^2 \times \frac{|\mathcal{W}_{\omega_1,\omega_2}|}{|\mathcal{P}_{\omega_1,\omega_2}|}.$$

Even though only "smooth" repartition is considered in the original article [MTSB12], we continue our analysis in the general case for all possible values $\omega_1 + \omega_2 = \omega$:

Proposition 9. *Let* $Y_{p,\omega} = \sum_{\omega_1+\omega_2=\omega} Y_{p,\omega_1,\omega_2}$ *and* $\omega^2 = p\log p + O(1/\sqrt{p})$. *Then*

$$P(Y_{p,\omega/2,\omega/2} \geqslant p-1) \leqslant P(Y_{p,\omega} \geqslant p-1) \leqslant \omega p^2 \frac{\binom{p-1}{\omega-2}}{\binom{2p}{\omega} + (-1)^{\frac{\omega}{2}+1}\binom{p}{\frac{\omega}{2}}}. \tag{22}$$

The upper bound is asymptotically equivalent to $p^{-\frac{1}{4}}\omega^3 2^{-\omega}$.

Proof. For the upper bound we use Eq. (35) from Appendix A and the formula

$$P(Y_{p,\omega} \geqslant p-1) = \sum_{\omega_1+\omega_2=\omega} P(Y_{p,\omega_1,\omega_2} \geqslant p-1)P(\omega_1,\omega_2)$$

Remark 6. If we recall the result in Sect. 4 we obtain a gain factor that is close to ω^2.

6 Improvements Under the Group Action of $(\mathbb{Z}_p^*, \times)$

In this section we define another group action that leaves the code invariant.

Definition 11. *We denote by "* $\cong''$ *the equivalence relation corresponding to the cyclic shifts equivalence class. The action of* $\mathbb{Z}_p^*$ *over* $\mathbb{F}_2[x]/(x^p-1)/\cong$ *can be defined as follow:*

$$\begin{array}{ccccc} \mathbb{Z}_p^* & \times & (\mathbb{F}_2[x]/(x^p-1)/\cong) & \longrightarrow & (\mathbb{F}_2[x]/(x^p-1)/\cong) \\ (\alpha & , & \mathcal{O}_h) & \longmapsto & \alpha\cdot\mathcal{O}_h, \end{array}$$

where $\alpha \cdot (\sum_{i=0}^{p-1} a_i x^i) = \sum_{i=0}^{p-1} a_i x^{\alpha i}$ *with* $\sum_{i=0}^{p-1} a_i x^i \in \mathcal{O}_h$.

So we start our attack by fixing $\alpha \in \mathbb{Z}_p^*$ and try to find a rational reconstruction of $\alpha \cdot f$ modulo $x^p - 1$. If the algorithm finds a solution (ψ_t, φ_t) where $\psi_t = \alpha \cdot f\varphi_t$ satisfy the inequality

$$\|\varphi_t\| + \|\psi_t\| \leqslant w. \tag{23}$$

then we have found as before (equivalent) secret polynomials.

Otherwise, the attacker performs the same attack to all shifts of f, namely $\alpha \cdot x^j f$. If the attack fails, another α is chosen and the procedure is repeated until the good combination of α and shifts are founded. As before, we want to estimate precisely the number of keys that can be recovered with this technique.

Lemma 2. *The group action previously defined is a ring morphism.*

Proof. We can easily check that $\alpha \cdot \left(x^a + x^b\right) = \alpha \cdot x^a + \alpha \cdot x^b$ and $\alpha \cdot \left(x^{a+b}\right) = \alpha \cdot x^a \times \alpha \cdot x^b$.

We give now the most relevant properties related to the group action defined above.

Proposition 10. *Let $\alpha \in \mathbb{Z}_p^*$ and $\mathcal{O}_h \in \mathbb{F}_2[x]/(x^p-1)/\cong$. The following equivalence holds:*

$$\alpha \cdot \mathcal{O}_h = \mathcal{O}_h \Leftrightarrow \exists\, h^* \in \mathcal{O}_h,\ \alpha \cdot h^* = h^*. \tag{24}$$

Proof. The ($\Leftarrow$) implication comes from the definition of the orbits. For the other implication, let h be an element of the $\mathcal{O}_h$ class so that $\alpha \cdot h \in \mathcal{O}_h$. This means that there exits $j < p$ so that $\alpha \cdot h = x^j h$. Then by setting $k = -j\alpha^{-1}(1-\alpha^{-1})^{-1}$ we have $\alpha \cdot (x^k h) = x^k h$.

Corollary 6. *Let $h \in \mathbb{F}_2[x]/(x^p-1)$ and $\overline{\mathcal{O}_h}$ be the orbit of $\mathcal{O}_h$ under the action of $(\mathbb{Z}_p^*, \times)$. Let Γ_h be the subgroup of $(\mathbb{Z}_p^*, \times)$ which stabilizes $\mathcal{O}_h$. Then the cardinality of the orbit $\overline{\mathcal{O}_h}$ is*

$$\left|\overline{\mathcal{O}_h}\right| = \frac{p-1}{|\Gamma_h|}. \tag{25}$$

Proposition 11. *Let $\alpha \in (\mathbb{Z}_p^*, \times)$ and $h \in \mathbb{F}_2[x]/(x^p-1)$ so that $\|h\| = \omega_1 < p$ and $\alpha \cdot \mathcal{O}_h = \mathcal{O}_h$. Then the order of α divides either ω_1 or $\omega_1 - 1$.*

So only group elements that respect the order property given above can fix elements in the set of polynomials with weight restrictions. Thus a natural consequence is that we can use the Burnside lemma for counting the number of orbits in this case, but this is no longer the purpose here.

As before we say that the set $\overline{\mathcal{O}_{h_1}} \times \overline{\mathcal{O}_{h_2}}$ is a weak orbit if and only if it contains at least one weak pair and denote by $P([Y_{p,\omega}] \geqslant p-1)$ the probability of a extended weak orbit. We also denote by Γ_{h_1,h_2} the subgroup that stabilize $\mathcal{O}_{h_1} \times \mathcal{O}_{h_2}$. We remark from Proposition 11 that for any pair of polynomials h_i with weight ω_i we have that any $\alpha \in (\mathbb{Z}_p^*, \times)$ that stabilizes the orbit $\mathcal{O}_{h_1} \times \mathcal{O}_{h_2}$ has to satisfy the condition

$$(\mathsf{ord}(\alpha) | \omega_1 \text{ or } \mathsf{ord}(\alpha) | \omega_1 - 1) \quad \text{and} \quad (\mathsf{ord}(\alpha) | \omega - \omega_1 \text{ or } \mathsf{ord}(\alpha) | \omega - \omega_1 - 1)\,.$$

In order to estimate the probability of such weak configurations, two main factors must be taken into consideration: the length of an orbit $\overline{\mathcal{O}_{h_1}} \times \overline{\mathcal{O}_{h_2}}$ and the intersection of two weak orbits.

Proposition 12. *If the intersection of any two weak orbits $\overline{\mathcal{O}_{h_1}} \times \overline{\mathcal{O}_{h_2}} \cap \overline{\mathcal{O}_{h_1^*}} \times \overline{\mathcal{O}_{h_2^*}} = \emptyset$ and $\Gamma_{h_1,h_2} = \{1,-1\}$ for any orbit then we have:*

$$\frac{p-1}{2}\left(\frac{\frac{\omega}{2}}{\binom{p-1}{\frac{\omega}{2}-1}}\right)^2\binom{p-1}{\omega-2} \leqslant P([Y_{p,\omega}] \geqslant p-1) \leqslant \frac{\omega p^3}{2}\frac{\binom{p-1}{\omega-2}}{\binom{2p}{\omega} + (-1)^{\frac{\omega}{2}+1}\binom{p}{\frac{\omega}{2}}}. \tag{26}$$

The asymptotic values for the upper and the lower bound can be computed as in Propositions 8 and 9.

Remark 7. We observe that with this extra group action we improved our probability by a multiplicative factor equal to $p-1$ in the best case. In the worst case the factor is still linear in the block length (see Proposition 11 and Corrolary 6).

7 Numerical Results

The parameters chosen for the experimental part are those suggested by the designers of the scheme [MTSB12]. The security levels correspond to the best known attacks given in [MTSB12]. The probabilities displayed in Figs. 1 and 2 are computed directly from the formulas given in Corollary 2, Proposition 7, Corollary 5 and Proposition 5.

In Fig. 1 we compute the exact values directly from Corollary 2 and Proposition 7 for the first and the second probability. In the last column we give the asymptotic value of the probability of a weak orbit from Corollary 5. The asymptotic value approaches very precisely the exact value, at least when the exact computation is possible. We used the following procedure to obtain our results:

- We generate the list $L := \left[\binom{\frac{\omega}{2}}{p-\frac{\omega}{2}}_k - \binom{\frac{\omega}{2}}{p-\frac{\omega}{2}}_{k-1} \right]_{k \in \{(p-1)/\frac{\omega}{2}, \ldots, p-\frac{\omega}{2}\}}$.
- We compute the convolution from Eq. 19

$$P\left(Y_{p,\omega_1,\omega_2} \geqslant p-1\right) = \sum_{\substack{k_1+k_2 \geqslant p-1 \\ k_1,k_2 \in \{(p-1)/\frac{\omega}{2}, \ldots, p-\frac{\omega}{2}\}}} L[k_1]L[k_2].$$

The results are amazingly faithful to the asymptotic value in the sense that for all the parameters the exponential factor is the same for the two probabilities up to the last digit. This result is quite amazing since the inequalities used in Appendix A page 21. for the asymptotic expansion are not very sharp. But one of the reasons why the two values are so close might come from the compensation phenomenon when computing the convolution in Eq. 19.

In Fig. 2, we display the probability values for all $\omega_1 + \omega_2 = \omega$. In the first column we compute the exact value of the probability from Proposition 5. Whereas in the next column we compute the asymptotic value of lower bound and the upper bound. In the last column we give only the asymptotic value for the upper

Security level	p	$\frac{\omega}{2}$	$\frac{\lvert\mathcal{W}_{\omega/2,\omega/2}\rvert}{\lvert\mathcal{P}_{\omega/2,\omega/2}\rvert}$ Corollary 2 exact value	$P(X_{p,\frac{\omega}{2}} \geqslant \frac{p-1}{2})^2$ Proposition 7 exact value	$P(Y_{p,\frac{\omega}{2},\frac{\omega}{2}} \geqslant p-1)$ Equation 19 exact value	Corollary 5 asympt. value
80	4801	45	2^{-87}	2^{-78}	$2^{-74.04}$	$2^{-74.04}$
	3593	51	2^{-99}	2^{-90}	$2^{-86.02}$	$2^{-86.02}$
	3079	55	2^{-108}	2^{-98}	$2^{-94.12}$	$2^{-94.12}$
128	9857	71	2^{-139}	2^{-128}	$2^{-124.52}$	$2^{-124.52}$
	7433	81	2^{-159}	2^{-149}	$2^{-145.58}$	$2^{-144.58}$
	6803	85	2^{-167}	2^{-157}	$2^{-153.67}$	$2^{-152.67}$
256	32771	132	2^{-260}	2^{-249}		$2^{-244.3}$
	22531	155	2^{-307}	2^{-295}		$2^{-290.5}$
	20483	161	2^{-319}	2^{-307}		$2^{-302.7}$

Fig. 1. Probability of a weak key (orbit) for the QC-MDPC when $\omega_1 = \omega_2 = \frac{\omega}{2}$.

Security level	p	$\frac{\omega}{2}$	$\frac{\lvert\mathcal{W}_\omega\rvert}{\lvert\mathcal{P}_\omega\rvert}$ Proposition 5 exact value	$P(Y_{p,\omega} \geqslant p-1)$ Proposition 9 bounds Eq. (22)	$P([Y_{p,\omega}] \geqslant p-1)$ Proposition 12 upper bound
80	4801	45	2^{-84}	$[2^{-74}, 2^{-71}]$	2^{-60}
	3593	51	2^{-96}	$[2^{-86}, 2^{-83}]$	2^{-72}
	3079	55	2^{-105}	$[2^{-94}, 2^{-91}]$	2^{-80}
128	9857	71	2^{-136}	$[2^{-125}, 2^{-121}]$	2^{-109}
	7433	81	2^{-156}	$[2^{-145}, 2^{-141}]$	2^{-129}
	6803	85	2^{-164}	$[2^{-153}, 2^{-149}]$	2^{-137}
256	32771	132	2^{-257}	$[2^{-244}, 2^{-241}]$	2^{-227}
	22531	155	2^{-303}	$[2^{-291}, 2^{-287}]$	2^{-273}
	20483	161	2^{-315}	$[2^{-303}, 2^{-299}]$	2^{-285}

Fig. 2. Probability of a weak key, extended weak pairs and improvements on extended weak pairs for the QC-MDPC for all $\omega_1 + \omega_2 = \omega$.

bound. One might think that the upper bound is not very tight and that the exact value of the probability is way lower than the value of the upper bound. Even though we share this concern we want to insist on the following fact. In order to obtain real sharp bounds many unanswered questions concerning the generalized Pascal-DeMoivre triangles are to deal with and this is clearly not the purpose here. Nevertheless the experiments show that the probability is quite close to the upper bound. As p goes to infinity and $\omega = O\left(\sqrt{p \log p}\right)$ the difference between the two values tends to zero. We compute the probabilities for the first cryptographic parameters $p = 4801$ and $\omega = 90$. The exact value for the probability equals $2^{-71.26}$ whereas the upper bound equals $2^{-71.12}$.

8 Complexity and Experimental Timings

The cost of the attack on public key using the two group actions previously defined, is in theory $p - 1$ action of $(\mathbb{Z}_p^*, \times)$ times p action of $(\mathbb{Z}_p, +)$ times the cost of the EEA. This is the worst case scenario and also the case where our attack in applied on a random key (potentially which is not weak).

The first set of parameters that we used were not in the scale of the cryptographic values. More precisely we considered $p = 101$ and $\omega_1 = \omega_2 = 9$. The purpose was to confront the theoretical values for the probabilities of a weak keys and the experimental results. In this sense using MAGMA's random generator we computed 10^5 pair of polynomials for the QC-MDPC scheme and executed the attack on the shifted keys. In theory the probability of finding a weak orbit equals 0.0032. Meanwhile in practice we obtained 317 weak orbits and the time needed to test all the orbits was approximately 6000 s.

In the second part we used the first parameters for the 2^{80} security level which are $p = 4801$ and $\omega = 90$ and consider the most frequent case $\max_{i \in \{1,2\}} \omega_i = 47$. In the first case we applied the EEA on a weak key. In the second part we

generated a weak key that we shifted. Therefore we randomly choose an integer $i \in (\mathbb{Z}_p, +)$ and applied the EEA on the i^{th} shift. We repeated the procedure until a weak key was found. In the worst case we had to compute all the p shifts, whereas in average we only needed a small number of trials until the weak key was discovered. The last column corresponds to the following experience. We generated a weak key, then we applied the action of $(\mathbb{Z}_p^*, \times)$ and the we shifted. In this case the procedure is the same: we randomly pick an element of the group $(\mathbb{Z}_p^*, \times)$ and consider the key under the action of this element. Then we apply the Shifted(EEA) until the proper pair of shift and extension in founded. In the worst case we compute all the possible combinations of shifts and extensions.

On a 4-core Intel(R) Xeon(R) CPU ES-2690 @ 2.90 GHz, using MAGMA V2.19-9 we applied two variants of the EEA : the recursive original variant with complexity $O(p^2)$ and the MAGMA implementation using the Knuth–Schönhage version with complexity $O(p \log p^2 \log\log p)$.

	EEA	Shifted(EEA)		Extended(Shifted(EEA))	
	Best	Average	Worst	Average	Worst
Recursive version	0.12 s	4.5 min	9.5 min	5.3 days	1 month
MAGMA version	0.86 ms	2 s	4.1 s	1 h	5 h 30 min

9 Conclusion

The rational reconstruction attack turns out to be a very efficient solution for the key recovery attack on the QC-MDPC scheme. The main advantages of the algorithm is its low complexity, that is sub-quadratic in the code length, and the fact that it can be computed in parallel for several instances of the public key.

We proposed a first technique to estimate the number of private keys that can be recovered with the extended Euclidean algorithm. Furthermore in order to increase the success probability, equivalence classes of the public key have been considered. Formally this operation was defined in terms of two group actions ($(\mathbb{Z}_p, +)$ and $(\mathbb{Z}_p^*, \times)$) over the set of polynomials in $\mathbb{F}_2[x]/(x^p - 1)$. Counting equivalence classes turned out to be a combinatorial problem based the theory of Lyndon words. This technique increased the quantity of weak keys by a multiplicative factor equal to ω^2. The second group action $(\mathbb{Z}_p^*, \times)$ increased the number by a multiplicative factor p.

In order to avoid such type of attacks one can easily check if the longest run of the private keys satisfy the conditions given in (15). The designer has to check if the group action previously defined increase or not the longest run in order to insure the security of the key.

We stress out the importance of our counting technique since it can be applied to other cryptographic schemes, for instance the NTRU cryptosystem.

Acknowledgement. We would like to thank the anonymous referees for their careful reading and helpful comments.

A Appendix

Proof of Theorem 1 First of all we define the variables involved in the theorem. Let p, ω, k be integers, such that $1 \leqslant \omega \leqslant p$ and $k \leqslant p - \omega$. A finite word w is a Lyndon word if w is strictly smaller for the lexicographical order than all of its cyclic shifts. We denote by $\mathcal{L}(\mathcal{A})$ the set of Lyndon words over an alphabet $\mathcal{A}$. Let $\mathcal{B}$ be a binary alphabet, and $\mathcal{L}^{\leqslant k}(\mathcal{B}, p, \omega)$ the set of all Lyndon words with length p, number of ones equal to ω and the longest run of zeros less or equal to k over $\mathcal{B}$. Let $\mathcal{A}_k = \{a_0, a_1, \ldots, a_k\}$ be an alphabet. Monoids $\mathcal{A}_k^*$ and $\mathcal{B}^*$ are endowed with the lexicographic orders satisfying $0 < 1$ and $a_k < \cdots < a_0$. The morphism

$$\begin{aligned} \varphi : \mathcal{A}_k^* &\to (0^*1)^* \subset \mathcal{B}^* \\ a_i &\to 0^i 1 \end{aligned}$$

is clearly an order preserving isomorphism. We deduce that $w \in \mathcal{A}_k^*$ is a Lyndon word if and only if $\varphi(w)$ is a Lyndon word (see [Ric03] for details). Setting $\psi(a_{l_0} \ldots a_{l_{j-1}}) = j + \sum_{m=0}^{j-1} l_m$ we obtain $\psi(w) = |\varphi(w)|$.

If we set $\mathcal{L}_\psi(\mathcal{A}_k, \omega, p) = \left\{ l \in \mathcal{L}(\mathcal{A}_k) \quad \middle| \quad |l| = \omega \text{ and } \psi(l) = p \right\}$ then

$$\varphi\left(\mathcal{L}_\psi(\mathcal{A}_k, \omega, p)\right) = \mathcal{L}^{\leqslant k}(\mathcal{B}, p, \omega).$$

Hence, it suffices to compute $\left|\mathcal{L}_\psi(\mathcal{A}_k, \omega, p)\right|$. We use the fact that the alphabet $\mathcal{A}_k$ is the generating basis for all words in the free monoid $\mathcal{A}_k^*$. In terms of formal series this means

$$\sum_{w \in \mathcal{A}_k^\star} w = \frac{1}{1 - \sum_{i=0}^{k} a_i}. \tag{27}$$

Then we use the Chen-Fox-Lyndon theorem that states that each word can be uniquely expressed as a decreasing product of Lyndon words [KTC58, Lot02]

$$\sum_{w \in \mathcal{A}_k^\star} w = \prod_{l \in \mathcal{L}(\mathcal{A}_k)}^{\swarrow} \frac{1}{1 - l}. \tag{28}$$

Sending each letter a_{l_m} to $z x^{l_m + 1}$ one obtains

$$\frac{1}{1 - z \sum_{i=1}^{k+1} x^i} = \prod_{1 \leqslant j \leqslant i}^{\infty} \left(\frac{1}{1 - x^i z^j} \right)^{\left|\mathcal{L}_\psi(\mathcal{A}_k, j, i)\right|}. \tag{29}$$

We apply the logarithm in each side of the equality above and develop using the Taylor expansion. In the resulting formula we compare the coefficient of $z^\omega x^p$ in the left hand side and the right hand side and obtain

$$\sum_{\substack{j|\omega \\ \frac{\omega}{j}|p}} j \left| \mathcal{L}_\psi(\mathcal{A}_k, j, \frac{p}{\omega} j) \right| = \binom{\omega}{p-\omega}_k, \tag{30}$$

where $\binom{\omega}{p}_k$ denotes the coefficient of x^p in $(1 + x + x^2 + \cdots + x^k)^\omega$.

We rewrite the last equation as

$$\sum_{j|\gcd(\omega,p)} \frac{\omega}{j} \left| \mathcal{L}_\psi(\mathcal{A}_k, \frac{\omega}{j}, \frac{p}{j}) \right| = \binom{\omega}{p-\omega}_k, \tag{31}$$

and apply the Möbius Inversion [Mob32, Lan09]to find the wanted result.

Proof of Proposition 8 By definition we have:

$$P\left(Y_{p,\omega_1,\omega_2} \geqslant p-1\right) = \sum_{\omega_2 - 1 \leqslant k \leqslant p-\omega_1} f_{X_{p,\omega_1}}(k) \left(1 - F_{X_{p,\omega_2}}(p-k-1-1)\right).$$

Lemma 3. *Let $\omega \geqslant 2$ and p prime. Then for $k > \lfloor \frac{p-\omega}{2} \rfloor$ we have*

$$f_{X_{p,\omega}}(k) = \frac{\omega\binom{p-k-2}{\omega-2}}{\binom{p-1}{\omega-1}}, \quad F_{X_{p,\omega}}(k-1) = 1 - \frac{\omega\binom{p-k-1}{\omega-1}}{\binom{p-1}{\omega-1}}. \tag{32}$$

For $k \leqslant \lfloor \frac{p-\omega}{2} \rfloor$ the bounds are

$$\frac{\omega\binom{p-k-2}{\omega-2} - \binom{\omega}{2}\left[\binom{p-2k-1}{\omega-1} - \binom{p-2k-3}{\omega-1}\right]}{\binom{p-1}{\omega-1}} \leqslant f_{X_{p,\omega}}(k) \leqslant \frac{\omega\binom{p-k-2}{\omega-2}}{\binom{p-1}{\omega-1}}, \tag{33}$$

$$\frac{\omega\binom{p-k-1}{\omega-1} - \binom{\omega}{2}\binom{p-2k-1}{\omega-1}}{\binom{p-1}{\omega-1}} \leqslant 1 - F_{X_{p,\omega}}(k-1) \leqslant \frac{\omega\binom{p-k-1}{\omega-1}}{\binom{p-1}{\omega-1}}. \tag{34}$$

For the upper bound, this gives

$$P\left(Y_{p,\omega_1,\omega_2} \geqslant p-1\right) \leqslant \sum_{k=\omega_2-1}^{p-\omega_1} \omega_1 \frac{\binom{p-k-2}{\omega_1-2}}{\binom{p-1}{\omega_1-1}} \omega_2 \frac{\binom{k}{\omega_2-1}}{\binom{p-1}{\omega_2-1}} = \frac{\omega_1\omega_2\binom{p-1}{\omega_1+\omega_2-2}}{\binom{p-1}{\omega_1-1}\binom{p-1}{\omega_2-1}}. \tag{35}$$

For the lower bound, we separate our sum into three different sums, for $k \leqslant \lfloor \frac{p-\omega_1}{2} \rfloor$, $\lfloor \frac{p-\omega_1}{2} \rfloor < k < p-1-\lfloor \frac{p-\omega_2}{2} \rfloor = \lceil \frac{p+\omega_2}{2} \rceil - 1$ and $\lceil \frac{p+\omega_2}{2} \rceil - 1 \leqslant k \leqslant p-\omega_1$

and use relations (32), (33) and (34):

$$P\left(Y_{p,\omega_1,\omega_2} \geqslant p-1\right) \geqslant \sum_{k=\omega_2-1}^{p-\omega_1} \omega_1 \frac{\binom{p-k-2}{\omega_1-2}}{\binom{p-1}{\omega_1-1}} \omega_2 \frac{\binom{k}{\omega_2-1}}{\binom{p-1}{\omega_2-1}} - \sum_{k=\omega_2-1}^{\lfloor \frac{p-\omega_1}{2} \rfloor} \binom{\omega_1}{2} \frac{\binom{p-2k-1}{\omega_1-1} - \binom{p-2k-3}{\omega_1-1}}{\binom{p-1}{\omega_1-1}} \omega_2 \frac{\binom{k}{\omega_2-1}}{\binom{p-1}{\omega_2-1}} - \sum_{k=\lceil \frac{p+\omega_2}{2} \rceil-1}^{p-\omega_1} \binom{\omega_2}{2} \frac{\binom{p-k-2}{\omega_1-2}}{\binom{p-1}{\omega_1-1}} \omega_1 \frac{\binom{2k-p+1}{\omega_2-1}}{\binom{p-1}{\omega_2-1}}$$

We use the relations $\binom{p-2k-1}{\omega_1-1} - \binom{p-2k-3}{\omega_1-1} = \binom{p-2k-2}{\omega_1-2} + \binom{p-2k-3}{\omega_1-2} \leqslant 2\binom{p-2k-2}{\omega_1-2}$ (as $\omega_1 \geqslant 2$), $\frac{\omega_1\omega_2}{\binom{p-1}{\omega_1-1}\binom{p-1}{\omega_2-1}} = \frac{p^2}{\binom{p}{\omega_1}\binom{p}{\omega_2}}$ and a change of variable $k \to p-k-2$ in the last sum to get

$$\frac{\binom{p}{\omega_1}\binom{p}{\omega_2}}{p^2} P\left(Y_{p,\omega_1,\omega_2} \geqslant p-1\right) \geqslant \binom{p-1}{\omega-2} - \omega_1 \sum_{k=\omega_2-1}^{\lfloor \frac{p-\omega_1}{2} \rfloor} \binom{p-2k-2}{\omega_1-2}\binom{k}{\omega_2-1} - \frac{1}{2}\omega_2 \sum_{k=\omega_1-2}^{\lfloor \frac{p-\omega_2}{2} \rfloor-1} \binom{p-2k-3}{\omega_2-1}\binom{k}{\omega_1-2}$$

Now we use the bound $\binom{p-2k-2}{\omega_1-2}\binom{k}{\omega_2-1} \leqslant \binom{p-k-2}{\omega-3}$ and the relation from [Gou72] $\sum_{k=r}^{s}\binom{a-k}{b} = \binom{a-r+1}{b+1} - \binom{a-s}{b+1} \leqslant \binom{a-r+1}{b+1}$ to get

$$\frac{\binom{p}{\omega_1}\binom{p}{\omega_2}}{p^2} P\left(Y_{p,\omega_1,\omega_2} \geqslant p-1\right) \geqslant \binom{p-1}{\omega-2} - \omega_1\binom{p-\omega_2}{\omega-2} - \frac{1}{2}\omega_2\binom{p-\omega_1}{\omega-2} \geqslant \binom{p-1}{\omega-2} - \frac{3}{2}\omega_1\binom{p-\omega_2}{\omega-2}.$$

if $\omega_1 = \max(\omega_1, \omega_2)$. We finally get the bounds

$$1 - \frac{3\omega_1}{2}\frac{\binom{p-\omega_2}{\omega-2}}{\binom{p-1}{\omega-2}} \leqslant \frac{P\left(Y_{p,\omega_1,\omega_2} \geqslant p-1\right)}{\frac{p^2\binom{p-1}{\omega-2}}{\binom{p}{\omega_1}\binom{p}{\omega_2}}} \leqslant 1. \tag{36}$$

We check that the lower bound tends to 1 when $w_i = O(\sqrt{p \log p})$.

References

[BBK08] Belbachir, H., Bouroubi, S., Khelladi, A.: Connection between ordinary multinomials, fibonacci numbers, bell polynomials and discrete uniform distribution. Ann. Math. Inform. **35**, 21–30 (2008)

[CCD+09] Cesaratto, E., Clément, J., Daireaux, B., Lhote, L., Maume-Deschamps, V., Vallée, B.: Regularity of the euclid algorithm, application to the analysis of fast GCD algorithms. J. Symbolic Comput. **44**(7), 726 (2009)

[Dav79] Davis, P.J.: Circulant Matrices. Pure and applied mathematics. Wiley, New York (1979)

[FOP+14] Faugère, J.C., Otmani, A., Perret, L., de Portzamparc, F., Tillich, J.P.: Folding alternant and goppa codes with non-trivial automorphism groups, submitted, [cs.IT] (2014). arxiv:1405.5101

[Gen01] Gentry, C.: Key recovery and message attacks on NTRU-composite. In: Pfitzmann, B. (ed.) EUROCRYPT 2001. LNCS, vol. 2045, pp. 182–194. Springer, Heidelberg (2001)

[Gou72] Gould, H.W.: Combinatorial identities: a standardized set of tables listing 500 binomial coefficient summations. Morgantown, W Va (1972)

[GR61] Gilbert, E.N., Riordan, J.: Symmetry types of periodic sequences. Illinois J. Math. **5**, 657–665 (1961)

[HS13] Hamdaoui, Y., Sendrier, N.: A non asymptotic analysis of information set decoding. In: Cryptology ePrint Archive, Report /162 (2013)

[Knu71] Knuth, D.E.: The analysis of algorithms. Actes Congr. Internat. Math. **3**, 269–274 (1971). http://cr.yp.to/bib/entries.html#1971/knuth-gcd

[KTC58] Lyndon, R.C., Chen, K.T., Fox, R.H.: Free differential calculus, iv. the quotient groups of the lower central series. Ann. Math. **68**(1), 81–95 (1958)

[Lan09] Landau, E.: Handbuch der Lehre von der Verteilung der Primzahlen. Teubner(1909)

[Leh38] Lehmer, D.H.: Euclid's algorithm for large numbers. Am. Math. Monthly **45**(4), 227–233 (1938)

[Loi01] Loidreau, P.: Codes derived from binary goppa codes. Probl. Inf. Transm. **37**(2), 91–99 (2001)

[Lot02] Lothaire, M.: Algebraic Combinatorics on Words. Encyclopedia of mathematics and its applications. Cambridge University Press, New York (2002)

[LV06] Lhote, L., Vallée, B.: Sharp estimates for the main parameters of the euclid algorithm. In: Correa, J.R., Hevia, A., Kiwi, M. (eds.) LATIN 2006. LNCS, vol. 3887, pp. 689–702. Springer, Heidelberg (2006)

[LV08] Lhote, L., Vallée, B.: Gaussian laws for the main parameters of the euclid algorithms. Algorithmica **50**(4), 497–554 (2008)

[McE78] McEliece, R.J.: A Public-Key System Based on Algebraic Coding Theory, pp. 114–116. Jet Propulsion Lab, DSN Progress Report, 44 (1978)

[Mob32] Möbius, A.F.: Über eine besondere art von umkehrung der reihen. Journal für die reine und angewandte Mathematik **9**, 105–123 (1832)

[MTSB12] Misoczki, R., Tillich, J.-P., Sendrier, N., Barreto, P.S.L.M.: MDPC-McEliece: New McEliece variants from moderate density parity-check codes. IACR Cryptology ePrint Archive, 409 (2012)

[Ric03] Richomme, G.: Lyndon morphisms. Bull. Belg. Math. Soc. Simon Stevin **10**(5), 761–785 (2003)

[Sch71] Schönhage, A.: Schnelle Berechnung von Kettenbruchentwicklungen. (German) [Fast calculation of expansions of continued fractions]. ACTA-INFO, 1, 139–144 (1971)

[SZ04] Stehlé, D., Zimmermann, P.: A binary recursive GCD algorithm. In: Buell, D.A. (ed.) ANTS 2004. LNCS, vol. 3076, pp. 411–425. Springer, Heidelberg (2004)