Error-correcting pairs and arrays from algebraic geometry codes

Irene Márquez-Corbella and Ruud Pellikaan

Dept. of Algebra, Geometry and Topology, University of Valladolid
Facultad de Ciencias, 47005 Valladolid, Spain. E-mail: imarquez@agt.uva.es
Dept. of Mathematics and Computing Science, Eindhoven University of Techn.
P.O. Box 513, 5600 MB Eindhoven, The Netherlands. E-mail: g.r.pellikaan@tue.nl

Abstract

The security of the most popular number-theory public key crypto (PKC) systems will be devastatingly affected by the success of a large quantum computer. Code-based cryptography is one of the promising alternatives that are believed to resist classical and quantum computer attacks. Many families of codes have been proposed for these cryptosystems, one of the main requirements is having an efficient t-bounded decoding algorithm.

In [16, 17] it was shown that for the so called very strong algebraic geometry codes $\mathcal{C}$ which is a collection of codes $C = C_L(\mathcal{X}, \mathcal{P}, E)$, where $\mathcal{X}$ is an algebraic curve over $\mathbb{F}_q$, $\mathcal{P}$ is an n-tuple of mutually distinct $\mathbb{F}_q$-rational points of $\mathcal{X}$ and E is a divisor of $\mathcal{X}$ with disjoint support from $\mathcal{P}$, an equivalent representation can be found. Moreover in [19] an efficient computational approach is given to retrieve a triple that is isomorphic with the original representation, and, from this representation, an efficient decoding algorithm is obtained.

In this talk, we will show how an efficient decoding algorithm can be retrieved from an algebraic geometry code $\mathcal{C}$ by means of error-correcting pairs [20] and arrays directly, that is without the detour via the representation $(\mathcal{X}, \mathcal{P}, E)$ of the code $C = C_L(\mathcal{X}, \mathcal{P}, E)$.

As a consequence we will have that algebraic geometry codes with certain parameters are not secure for the code-based McEliece public key cryptosystem.

Keywords

Code based cryptography, McEliece public key cryptosystem,
algebraic geometry codes, error-correcting pairs and arrays.

1 Introduction

The security of code-based cryptosystems is founded on the (supposedly) hardness of decoding up to half the minimum distance. The minimum distance decoding problem was shown by Berlekamp-McEliece-Van Tilborg [1, 3] to be NP-hard. McEliece [21] proposed a PKC system using binary Goppa codes.

All known minimum distance decoding algorithms for general codes have exponential complexity in the length of the code. The complexity exponent of decoding general binary codes up to half the minimum distance has been lowered in a series of papers from above 1/3 for brute force decoding to below 1/20 by [2]. However there are several classes of codes such as the generalized Reed-Solomon (GRS), BCH, Goppa or algebraic geometry codes which have polynomial decoding algorithms that correct up to a certain bound which is at most half the minimum distance.

In 1986 [23] Niederreiter presented a dual version of McEliece cryptosystem which is equivalent in terms of security. This system differs from McEliece's system since it uses a parity check matrix instead of a generator matrix of the code. Several classes of codes are proposed for code-base PKC systems such as subcodes of GRS codes, alternant codes which contains the Goppa codes as subclass, and algebraic geometry codes [12].

It was shown in [6, 14, 24, 26, 28] that the known efficient bounded distance decoding algorithms of the before mentioned codes can be described by a basic algorithm using an error-correcting pair. That means that the proposed McEliece cryptosystem that use these classes of codes can be viewed as using the error-correcting pair as a secret key. Hence the security of these PKC systems is not only based on the inherent intractability of bounded distance decoding but also on the assumption that it is difficult to retrieve an error-correcting pair.

2 Error-correcting pairs and arrays

From now on the dimension of a linear code C will be denoted by $k(C)$ and its minimum distance by $d(C)$. Given two elements $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{F}_q^n$, the *star multiplication* is defined by coordinatewise multiplication, that is $\mathbf{a} * \mathbf{b} = (a_1b_1, \ldots, a_nb_n)$ while the *standard inner multiplication* is defined by $\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_ib_i$. In general, for two subsets A and B of $\mathbb{F}_q^n$ the set $A * B$ is given by $\{\mathbf{a} * \mathbf{b} \mid \mathbf{a} \in A \text{ and } \mathbf{b} \in B\}$. Furthermore $A \perp B$ if and only if $\mathbf{a} \cdot \mathbf{b} = 0$ for all $\mathbf{a} \in A$ and $\mathbf{b} \in B$.

Let C be a linear code in $\mathbb{F}_q^n$. The pair (A, B) of linear codes over $\mathbb{F}_{q^e}$ of length n is called a *t-error-correcting pair* (ECP) for C if the following properties hold:

E.1 $(A * B) \perp C$,

E.2 $k(A) > t$,

E.3 $d(B^{\perp}) > t$,

E.4 $d(A) + d(C) > n$.

The notion of an error-correcting pair for a linear code was introduced in 1988 by Pellikaan [24, 26] and independently by Kötter in [14, 15] in 1992. It is shown that a linear code in $\mathbb{F}_q^n$ with a t-error-correcting pair has a decoding algorithm which corrects up to t errors with complexity $\mathcal{O}\left((en)^3\right)$.

The existence of ECP's for GRS and algebraic geometry codes was shown in [24, 26]. For many cyclic codes Duursma and Kötter in [6, 14, 15] have found ECP's which correct beyond the designed BCH capacity.

An *error-correcting array* is defined in [13, 27] for a sequence of codes. From it follows the *Feng-Rao designed minimum distance* of the codes and the majority voting scheme of Feng-Rao [4, 5, 8] gives a decoding algorithm that decodes these codes up to half the Feng-Rao designed minimum distance with complexity $\mathcal{O}(n^3)$. An equivalent formulation is given in terms of *(weakly) well-behaving sequences* [9, 10, 11].

3 Algebraic geometry codes

Let $\mathcal{X}$ be an algebraic curve defined over $\mathbb{F}_q$ with genus g. Let $\mathcal{P}$ be an n-tuple of $\mathbb{F}_q$-rational points on $\mathcal{X}$ and let E be a divisor of $\mathcal{X}$ with disjoint support from $\mathcal{P}$ of degree m. Then the algebraic geometry code $C_L(\mathcal{X}, \mathcal{P}, E)$ is the image of the Riemann-Roch space $L(E)$ of rational functions with prescribed behavior of zeros and poles at E under the evaluation map $\mathrm{ev}_{\mathcal{P}}$. If $m < n$, then the dimension of the code $C_L(\mathcal{X}, \mathcal{P}, E)$ is at least $m + 1 - g$ and its minimum distance is at least $n - m$. If $m > 2g - 2$, then its dimension is $m + 1 - g$. The dual code $C_L(\mathcal{X}, \mathcal{P}, E)^{\perp}$ is again AG. If $m > 2g - 2$, then the dimension of the code $C_L(\mathcal{X}, \mathcal{P}, E)^{\perp}$ is at least $n - m - 1 + g$ and its minimum distance is at least $d^* = m - 2g + 2$, which is called the *designed minimum distance.* If $m < n$, then its dimension is $n - m - 1 + g$.

Algebraic geometry codes were proposed by Niederreiter [23] and Janwa-Moreno [12] for code-based PKC systems. This system was broken for genus zero [29], one and two [7, 22] and for arbitrary genus for so called VSAP codes [16, 17, 18, 19].

Let $r = l(E) - 1$ and $\{f_0, \ldots, f_r\}$ be a basis of $L(E)$. Consider the following map:

$$\varphi_E : \mathcal{X} \longrightarrow \mathbb{P}^r(\mathbb{F}_q)$$

defined by $\varphi_E(P) = (f_0(P) : \ldots : f_r(P))$. If $m > 2g$, then $r = m - g$. So φ_E defines an embedding of the curve $\mathcal{X}$ of degree m in $\mathbb{P}^r$. More precisely, let $\mathcal{Y} = \varphi_E(\mathcal{X})$, $Q_j = \varphi_E(P_j)$ and $\mathcal{Q} = (Q_1, \ldots, Q_n)$. Then $\mathcal{Y}$ is a curve in $\mathbb{P}^{m-g}$ of degree m and φ_E is an isomorphism from $\mathcal{X}$ to $\mathcal{Y}$. Now $\varphi_E(E) \equiv \mathcal{Y} \cdot H$ for every hyperplane H of $\mathbb{P}^{m-g}(\mathbb{F}_q)$. If moreover E is effective, then $\varphi_E(E) = \mathcal{Y} \cdot H$ for some hyperplane H of $\mathbb{P}^{m-g}(\mathbb{F}_q)$. Let $F = \varphi_E(E)$, then $(\mathcal{Y}, \mathcal{Q}, F)$ is a representation of $\mathcal{C}$ that is strict isomorphic with $(\mathcal{X}, \mathcal{P}, E)$.

If $m \geq 2g + 2$, then $I(\mathcal{Y})$ is generated by $I_2(\mathcal{Y})$. If moreover $n > 2m$, then $I_2(\mathcal{Q}) = I_2(\mathcal{Y})$. Now $C_L(\mathcal{X}, \mathcal{P}, E)$ is called a *very strong algebraic geometry* (VSAG) code if

$$2g + 2 \leq m < \frac{1}{2}n \quad \text{or} \quad \frac{1}{2}n + 2g - 2 < m \leq n - 4.$$

It was shown that the representation by the triple $(\mathcal{X}, \mathcal{P}, E)$ of a VSAG code $C_L(\mathcal{X}, \mathcal{P}, E)$ is unique up to isomorphisms [16, 17, 18] and that such a triple can be retrieved efficiently [19].

4 Error-correcting pairs and arrays from VSAG codes

Let $C = C_L(\mathcal{X}, \mathcal{P}, E)^\perp$ be an AG code on a curve of genus g with designed minimum distance d^* and $m = \deg(E) > 2g - 2$. Let $A = C_L(\mathcal{X}, \mathcal{P}, E - F)$, $B = C_L(\mathcal{X}, \mathcal{P}, F)$ and $C = C_L(\mathcal{X}, \mathcal{P}, E)^\perp$. Then $\langle A * B\rangle \subseteq C^\perp$. If moreover $t = \lfloor (d^* - 1 - g)/2 \rfloor$ and $\deg(F) = m - t - g$, then (A, B) is a t-ECP over $\mathbb{F}_q$ by [25, Theorem 1] and [26, Theorem 3.3]. So there are abundant ways to construct error-correcting pairs of an AG code.

This approach needs the efficient computation of the Riemann-Roch spaces $L(F)$ and $L(E-F)$ and such algorithms are available. If e is sufficiently large and $m > 4g - 3$, then there exists a $\lfloor (d^*-1)/2 \rfloor$-ECP over $\mathbb{F}_{q^e}$ by [28, Proposition 4.2], but no efficient way to obtain the pair is known.

In the following we construct ECP's directly using subspaces of $\mathbb{F}_q^n$ and circumventing the use of the Riemann–Roch spaces. If we take $F = (m-t-g)P_1$ where P_1 is the first rational point of $\mathcal{P}$, then $L(E-F)$ is a subspace of $L(E)$, and $A = C_L(\mathcal{X}, \mathcal{P}, E-F)$ is a subspace of $C^\perp = C_L(\mathcal{X}, \mathcal{P}, E)$.

In fact A is the space of those codewords in $C^\perp$ that are zero at the first position of multiplicity $m - t - g$ and this multiplicity can be controlled, since we have computed $I_2(\mathcal{Q})$ efficiently. Define $B_0 = \langle A * C\rangle^\perp$, then $B_0^\perp = \langle A * C\rangle \subseteq B^\perp$. So $d(B_0^\perp) \geq d(B^\perp) > t$. Hence (A, B_0) is a t-ECP for C. There is one technical detail, note that P_1 is in the support of $E - F$ and F, but there is a generalized way to define algebraic geometry codes, using a local parameter as explained in [19], where it is no longer necessary to assume that $\mathcal{P}$ is disjoint from the support of the divisor E in the definition of the code $C_L(\mathcal{X}, \mathcal{P}, E)$.

Similarly we can decode up to $\lfloor (d^* - 1)/2 \rfloor$ errors using arrays or well-behaving sequences and majority voting [4, 5, 9, 10, 11].

References

[1] A. Barg. Complexity issues in coding theory. In V.S. Pless and W.C. Huffman, editors, *Handbook of coding theory*, volume 1, pages 649–754. North-Holland, Amsterdam, 1998.

[2] 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*, volume 7237 of *Lecture Notes in Comput. Sci.*, pages 520–536. Springer, Heidelberg, 2012.

[3] E.R. Berlekamp, R.J. McEliece, and H.C.A. van Tilborg. On the inherent intractability of certain coding problems. *IEEE Transactions on Information*, 24:384–386, 1978.

[4] I.M. Duursma. *Decoding codes from curves and cyclic codes.* PhD thesis, Eindhoven University of Technology, 1993.

[5] I.M. Duursma. Majority coset decoding. *IEEE Transactions on Information Theory*, 39(3):1067–1070, 1993.

[6] I.M. Duursma and R. Kötter. Error-locating pairs for cyclic codes. *IEEE Trans. Inform. Theory*, 40:1108–1121, 1994.

[7] C. Faure and L. Minder. Cryptanalysis of the McEliece cryptosystem over hyperelliptic codes. In *Proceedings 11th Int. Workshop on Algebraic and Combinatorial Coding Theory, ACCT 2008*, pages 99–107, 2008.

[8] G.L. Feng and T.R.N. Rao. Decoding algebraic-geometric codes up to the designed minimum distance. *IEEE Transactions on Information Theory*, 39(1):37–45, 1993.

[9] G.L. Feng and T.R.N. Rao. A simple approach for construction of algebraic-geometric codes from affine plane curves. *IEEE Transactions on Information Theory*, 40(4):1003–1012, 1994.

[10] G.L. Feng and T.R.N. Rao. Improved geometric goppa codes. I. basic theory. *IEEE Transactions on Information Theory*, 41(6):1678–1693, 1995.

[11] O. Geil, R. Matsumoto, and D. Ruano. Feng-Rao decoding of primary codes. *Finite Fields and their Applications*, 23:35–52, 2013.

[12] H. Janwa and O. Moreno. McEliece public crypto system using algebraic-geometric codes. *Designs, Codes and Cryptography*, 8:293–307, 1996.

[13] C. Kirfel and R. Pellikaan. The minimum distance of codes in an array coming from telescopic semigroups. *IEEE Trans. Inform. Theory*, 41(6, part 1):1720–1732, 1995. Special issue on algebraic geometry codes.

[14] R. Kötter. A unified description of an error locating procedure for linear codes. In *Proceedings of Algebraic and Combinatorial Coding Theory*, pages 113–117. Voneshta Voda, 1992.

[15] R. Kötter. *On algebraic decoding of algebraic-geometric and cyclic codes.* PhD thesis, Linköping University of Technology, Linköping Studies in Science and Technology, Dissertation no. 419, 1996.

[16] I. Márquez-Corbella, E. Martínez-Moro, and R. Pellikaan. Cryptanalysis of public-key cryptosystems based on algebraic geometry codes. *Oberwolfach Preprints*, OWP 2012-01:1–17, 2012.

[17] I. Márquez-Corbella, E. Martínez-Moro, and R. Pellikaan. On the unique representation of very strong algebraic geometry codes. *Designs, Codes and Cryptography*, pages 1–16, 2013.

[18] I. Márquez-Corbella, E. Martínez-Moro, and R. Pellikaan. Evaluation of public-key cryptosystems based on algebraic geometry codes. In *Proceedings of the Third International Castle Meeting on Coding Theory and Applications*, pages 199–204, Cardona Castle, Barcelona, September 11-15, 2011.

[19] I. Márquez-Corbella, E. Martínez-Moro, R. Pellikaan, and D. Ruano. Computational aspects of retrieving a representation of an algebraic geometry code. *submitted to Designs, Codes and Cryptography*, 2013.

[20] I. Márquez-Corbella and R. Pellikaan. A characterization of MDS codes that have an error-correcting pair. Lyngby, Denmark, 2012.

[21] R.J. McEliece. A public-key cryptosystem based on algebraic coding theory. *DSN Progress Report*, 42–44:114–116, 1978.

[22] L. Minder. *Cryptography based on error correcting codes.* PhD thesis, 3846 EPFL, 2007.

[23] H. Niederreiter. Knapsack-type crypto systems and algebraic coding theory. *Problems of Control and Information Theory*, 15(2):159–166, 1986.

[24] R. Pellikaan. On decoding linear codes by error correcting pairs. Preprint Technical University Eindhoven, 1988.

[25] R. Pellikaan. On a decoding algorithm of codes on maximal curves. *IEEE Trans. Inform. Theory*, 35:1228–1232, 1989.

[26] R. Pellikaan. On decoding by error location and dependent sets of error positions. *Discrete Math.*, 106–107:369–381, 1992.

[27] R. Pellikaan. On the efficient decoding of algebraic-geometric codes. In *Eurocode '92 (Udine, 1992)*, volume 339 of *CISM Courses and Lectures*, pages 231–253. Springer, Vienna, 1993.

[28] R. Pellikaan. On the existence of error-correcting pairs. *Statistical Planning and Inference*, 51:229–242, 1996.

[29] V.M. Sidelnikov and S.O. Shestakov. On the insecurity of cryptosystems based on generalized Reed-Solomon codes. *Discrete Math. Appl.*, 2:439–444, 1992.