An MQ/Code Cryptosystem Proposal

Leonard J. Schulman*

March 6, 2013

Abstract

We describe a new trap-door (and PKC) proposal. The proposal is "multivariate quadratic" (relies on the hardness of solving systems of quadratic equations); it is also code-based, and uses the code-scrambling technique of McEliece (1978). However, in the new proposal, the error-correcting code is not revealed in the public key, which protects against the leading attacks on McEliece's method.

Keywords: Multivariate quadratic cryptosystem, MinRank, tensor decomposition, post-quantum cryptography, code-based cryptography.

1 Introduction

We describe a new trap-door proposal based on elementary methods. The proposal belongs to the MQ family of cryptosystems, i.e., its security rests on the hardness of solving systems of multivariate quadratic equations.

The proposal is also code-based, and like McEliece's well-known method, it relies on a "scrambled" efficiently-decodable error correcting code. Apart from the adoption of the scrambling technique, the method is distinct from all previous methods. (In particular contrary to possible casual impression, it does not include McEliece's method as some kind of special case.)

An advantage of the new method is that the scrambled generator matrix of the code is not public, thereby reducing the reliance of the cryptosystem on the hardness of decoding random linear codes [12, 74]. A disadvantage of the new method is that its efficiency is limited by the current state of the art in coding theory. In the current state, the proposal is considerably less efficient than competing "post-quantum"-grade cryptosystems: lattice cryptosystems [2, 57, 72, 55, 71, 58, 56, 59] and the closely related subset-sum method in [52], McEliece [54, 63], or other MQ cryptosystems [70, 24, 29, 48]. It is therefore not the likely replacement for RSA/ECC if quantum computers materialize without further attacks on these competing systems. This could change depending on improvements in the error-correcting codes needed for the present proposal, and, of course, depending on the state of attacks on the various systems, which use trap-doors that are completely different from the one we introduce.

In Sec. 2 we describe an idealized form of the proposal, using hypothetical codes of a quality that is combinatorially feasible, but for which no efficiently-decodable code is yet known. In Sec. 4 we

*Engineering and Applied Science Division, California Institute of Technology. Email: schulman@caltech.edu. Supported in part by the NSF.

describe the actual state of the art in coding theory. In Sec. 5 we then show how the cryptosystem can be implemented (less efficiently than in the ideal proposal) using these state-of-the-art codes. In Sec. 6 we discuss hypothetical improved implementations.

It will be simplest to directly describe the idea as a PKC encrypting n bits; the underlying trap-door claim will then be easy to state.

2 The Ideal Proposal

Let $X_0 \cong (GF2)^n$; this is the plaintext space of the cryptosystem. The construction relies in an essential way on a linear error-correcting code over $GF2$; however, unlike in the McEliece cryptosystem, X_0 is *not* the message space of the error correcting code. That message space is $U \cong (GF2)^\kappa$; $W \cong (GF2)^\ell$ is the codeword space, and we let $G \in \mathrm{Hom}(U, W)$ be the generator matrix for the error correcting code. (G encodes κ-bit row vectors by ℓ-bit row vectors.) The code G is required to have the following properties for δ, d as large as possible.

1. G has an efficient decoding algorithm that can correct *almost* all patterns of $\delta\ell$ errors for some $0.25 < \delta < 0.5$, say $\delta = 0.3$.
2. Every d columns of G are linearly independent; or what is the same, $G^\perp$ (the code spanned by vectors in the right nullspace of G) has minimum (Hamming) distance $d+1$. No decoding algorithm is required for the code $G^\perp$.

In addition to κ, ℓ, δ, d the construction also uses two less-critical parameters n_1, ρ. All six parameters can be adjusted to optimize the security-efficiency tradeoff, subject to the following constraints (whose reasons will be apparent later): $n_1 \le n$; $\kappa c_1 < \ell$ for a $c_1 > 1$ which depends on the quality of error-correcting-code constructions; $0 \le \rho \le n$ (ρ is a semantic-security parameter); $c_2(n+\rho) < \ell$ for some $c_2 > 4$; $2(n+\rho) < d$. As a guideline one should consider κ, ℓ, d linear in n (although little will change if they are slightly larger, e.g., quasilinear), while n_1, ρ are each proportional to n^c for some $0 < c \le 1$ (not necessarily the same c).

Let $R_0 \cong R_1 \cong (GF2)^\rho$; these are auxiliary random-bit spaces used for semantic security.

Let $X = X_0 \oplus R_0$.

Let $X_1 \cong (GF2)^{n_1}$; let D be a full-rank matrix in $\mathrm{Hom}(X_0, X_1)$. For simplicity one may let $n_1 = n$ and take D to be any fixed full-rank matrix. (But generally D may be chosen randomly as part of the private key.) Let $V = X_1 \oplus R_1$.

Select, uniformly at random, a tensor $M \in X^\dagger \otimes V^\dagger \otimes U$. (So, its dimensions are $(n+\rho)\times(n_1+\rho)\times\kappa$.)

Let P be a uniformly random $\ell \times \ell$ permutation matrix.

The public key of the cryptosystem is a tensor in $X^\dagger \otimes V^\dagger \otimes W$ that is the sum of two terms. The first term, T_{mask}, is defined by

$$T_{\mathrm{mask}} = M \cdot G \cdot P. \tag{1}$$

Note that G acts on the U component of M. Here $\cdot$ denotes tensor contraction, which of course specializes to matrix product or inner product. In Einstein notation Eqn. 1 is

$$(T_{\mathrm{mask}})_{ijk} = M_{ij\ell} G^\ell_m P^m_k.$$

Observe that every "row" of T_{mask}, by which we mean, every vector of the form $(x \otimes v) \cdot T_{\text{mask}}$ for $x \in X$, $v \in V$ (equivalently $x^i v^j (T_{\text{mask}})_{ijk}$), belongs to the "scrambled" error-correcting code spanned by the rows of $G \cdot P$.

The second term, T_{pure}, is constructed as follows. Select, uniformly and independently, vectors $a_1, \ldots, a_\ell \in X^\dagger$ and $b_1, \ldots, b_\ell \in V^\dagger$. Form a tensor T_{pure} whose k'th "slice" transverse to W is $a_k \otimes b_k$. (Throughout this paper "slice" will mean a smaller-order tensor, in this case a matrix, obtained by restricting the full tensor to a single coordinate k in the output space W.) More formally, with w_k being the standard basis for W, let

$$T_{\text{pure}} = \sum_{k=1}^{\ell} a_k \otimes b_k \otimes w_k \tag{2}$$

or more explicitly

$$(T_{\text{pure}})_{ijk} = (a_k)_i (b_k)_j.$$

Finally, set the public key to be

$$T = T_{\text{mask}} + T_{\text{pure}}. \tag{3}$$

Summary. The public key is the tensor T. The private key is $M, P, D, a_1, \ldots, a_\ell, b_1, \ldots, b_\ell$. The code G is fixed in the specification of the cryptosystem.

2.1 Encryption

In order to encrypt $x \in X_0$, pick uniformly at random $r_0 \in R_0$ and $r_1 \in R_1$. The encryption is the following vector in W:

$$z = ((x + r_0) \otimes (x \cdot D + r_1)) \cdot T$$

or in Einstein notation,

$$z_k = (x + r_0)^i (x \cdot D + r_1)^j T_{ijk}$$

2.2 Honest-Party Decryption

Given the encryption $z = ((x + r_0) \otimes (x \cdot D + r_1)) \cdot T$, apply the decoding algorithm for G to the vector $z \cdot P^{-1}$ to decompose z into $z = z_{\text{mask}} + z_{\text{pure}}$,

$$z_{\text{mask}} = (x + r_0) \otimes (x \cdot D + r_1) \cdot T_{\text{mask}} = (x + r_0) \otimes (x \cdot D + r_1) \cdot M \cdot G \cdot P$$
$$z_{\text{pure}} = (x + r_0) \otimes (x \cdot D + r_1) \cdot T_{\text{pure}}$$

The decoding algorithm works because the 1's in z_{pure} are a random set of density w.h.p. $\sim 1/4$, which is $< \delta$.

Next we obtain x from z_{pure}. Generically this would be a hard problem because it requires solving a system of quadratic equations. However, the essence of the trap-door is that these equations are structured to be easy to solve. We obtain x (and r_0 which we do not need) from z_{pure} as follows: find any set K of $(n+\rho+\ell/4)/2$ coordinates k for which $(z_{\text{pure}})_k = 1$. (That this is almost always possible uses the fact that the set of 1's in z_{pure} is a random set of size w.h.p. $\sim 1/4$, and also that $\ell > 4(n + \rho)$.) Observe that if $(z_{\text{pure}})_k = 1$ then necessarily both $(x + r_0) \cdot a_k = 1$ and $(x \cdot D + r_1) \cdot b_k = 1$.

Then solve the (w.h.p. full-rank) system of $(n + \rho + \ell/4)/2$ linear equations in $n + \rho$ variables:

$$(x + r_0) \cdot a_k = 1 \qquad \text{for } k \in K. \tag{4}$$

2.3 Trap-Door Interpretation

Conjecturally, even fixing $\rho = 0$ and using a fixed D, the family of encryption functions is trap-door, one-way, and w.h.p. injective.

3 Security of the Ideal Proposal: Known Attacks

3.1 Message Attacks

A successful adversary must almost certainly use the special structure of the system, because without it, even if D is public, decrypting a message requires solving a system of $\ell \in O(n)$ random quadratic equations in $n + 2\rho$ variables. This system is over-determined, but only slightly. Gröbner (relinearization) algorithms can typically solve in polynomial time highly over-determined quadratic systems. (This is simple when the number of equations exceeds the number of pairs of variables, but more interesting below this regime.) The best known approaches are based on Faugère's F5 algorithm. However the performance of all known algorithms degrades rapidly as the number of equations decreases. A full theoretical understanding of the rate of degradation does not yet exist, but all existing studies [49, 25, 40, 27, 26, 95, 93, 5, 9, 60, 39, 13, 35, 87, 3, 8, 88, 86, 89] are apparently consistent with (and some actually suggest [25, 26]) the following runtime scaling rule for ℓ randomly chosen equations in n variables over the field GFq:

$$q^{\Theta(n/\sqrt{\ell})}\ \mathrm{poly}(\ell, n) \tag{5}$$

In our context $q = 2$ and $\ell \in O(n)$ which gives an estimate of $\Theta(\sqrt{n})$ for the security parameter for message attacks.

No quantum approaches to solving systems of quadratic equations are known, apart from the generic Grover search running in time $q^{n/2}$ (as compared with classical search in time q^n).

We conclude that any method that breaks the cryptosystem is likely to have to uncover significant information about the key.

3.2 Key Attacks

How does the public key T differ from a uniformly random tensor? We note one quantitative difference (Sec. 3.2.1) and two structural differences (Secs. 3.2.2 and 3.2.3), and discuss to what extent they may enable key attacks. Finally in Sec. 3.2.4 we discuss an important structural sense in which T does *not* share a vulnerability that has been exploited in attacks on the McEliece cryptosystem.

3.2.1 Quantitative Property: Tensor Rank

The first quantity to look at is *tensor rank*. (The rank of T is the smallest r such that for some vectors f_i, g_i, h_i, $T = \sum_{i=1}^{r} f_i \otimes g_i \otimes h_i$.) However, for sufficient values of the parameters k and ℓ, this quantity apparently does not distinguish T from a uniformly random tensor. Actually tensor rank is poorly understood; e.g., it is known that the maximum rank of any $n \times n \times n$ tensor is between $(1/3)n^2$ and $(3/4)n^2$ [47] (a better upper bound of $n(n+1)/2$ holds for algebraically closed

fields [64]), but the leading constant is not known; also, the rank of a random tensor is w.h.p. at least $(1/3)n^2$, but it is not even known whether it is w.h.p. close to the maximum possible rank.

Nonetheless, for κ sufficiently larger than n, it is almost unavoidable that the rank of a random $n \times n \times \kappa$ tensor (our M) is w.h.p. close to that of a random $n \times n \times \ell$ tensor (and both approach the limiting value of $(n+\rho)(n_1+\rho)$); yet remarkably, these statements do not appear to be established yet as theorems. The questions are perhaps not easy due to the challenges in obtaining any bounds on tensor rank, but, as experience with MAX-SAT has shown [19], computational intractability of the underlying random variable need preclude progress on bounds for the ensemble. In any case, because of the difficulties in estimating tensor rank, even assuming that the preceding assertions are correct, it is not clear if κ "a large constant times n" is, for instance, sufficient to make these two ranks almost the same; the author's guess is that the answer is yes. If this is indeed the case, then the rank of T does not distinguish it from that of a random tensor. If on the other hand this fails to be the case, then there is a slight possibility of attacks based on the *tensor decomposition problem:* find an optimal or even nearly-optimal decomposition of T (in terms of the number of rank 1 tensors used in the sum). This could reveal the terms $a_k \otimes b_k$ and crack the cryptosystem. However it is very unlikely that there is an efficient algorithm for this problem. Håstad [46] showed that the problem "Is rank $T \leq r$" is NP-complete over finite fields. There is also empirical evidence for the hardness of the tensor decomposition problem from work on efficiently computing bilinear problems, including integer multiplication and, most notably, matrix multiplication. The essential target of that work has been, indeed, to obtain low rank decompositions of some specific families of tensors. And (just as for this cracking this cryptosystem), it is not necessary toward that goal to find decompositions that achieve the exact rank of the tensor; decompositions which are not too much larger are also useful. It is therefore notable that upper bounds for the matrix multiplication exponent improved only incrementally despite efforts spanning two decades [83, 65, 11, 66, 67, 76, 68, 75, 22, 84, 85, 23], then stalled entirely for another two decades, until two small improvements recently with considerable computational investment [82, 92]. (Perhaps another sign of the hardness of the decomposition problem is that, after the first few papers, efforts ceased to upper bound the rank directly; progress on the exponent came through improved upper bounds on the border rank, a number that can be smaller but still gives an equivalent-exponent algorithm, with some additional constant-factor overhead.)

Structurally, however, T has two properties that distinguish it from a random tensor, which we now discuss.

3.2.2 Structural Property: Scrambled Code Structure

There is a permutation of the (known) code G, such that all the rows of T are within relative Hamming distance $\sim 1/4$ of a codeword.

We refer to this as the "scrambled code" structure of the public key. We now discuss the exposure of the cryptosystem to attacks based on this liability.

It appears to be hard to detect this structure. Sendrier [77, 16, 78] made a breakthrough on the *code equivalence problem*, in which a code G and its scrambled version $G \cdot P$ are provided, and the problem is to find P; however, there is no evident way to use his ideas here, simply because $G \cdot P$ is not provided in the clear. What is provided are only noisy versions of $(n+\rho)^2$ unknown, randomly chosen codewords in $G \cdot P$. (Each codeword has added to it a noise vector of fractional weight $\sim 1/4$.) Conceptually one may think of this as "McEliece noise" added permanently in the public key (as compared with the McEliece cryptosystem where the noise is added per-message).

3.2.3 Structural Property: Min-Rank Structure

If y is a linear dependence of the columns of the permuted code, i.e., if $y \in G^{\perp} \cdot P$, then $T_{mask} \cdot y = 0$ and therefore $T \cdot y = T_{pure} \cdot y$.

We refer to this as the "Min-Rank" structure of the public key. We now discuss the exposure of the cryptosystem to attacks based on this liability.

The linear dependencies among the "slices" of T_{mask} are w.h.p. identical to those among the columns of $G^{\perp} \cdot P$ (the only exception being the exponentially-rare event that M does not have linearly independent slices). Gaining any information about these dependencies is very useful to the cryptanalyst; the information can be exploited in two ways:

1. If the cryptanalyst learns a complete set of dependencies, i.e., a basis for the code $G^{\perp} \cdot P$, then he also gains a basis for the code $G \cdot P$ (although not that same matrix), so that he is in precisely the situation of an attacker of the McEliece cryptosystem with known error-correcting code, i.e., in the position of knowing both G and $S \cdot G \cdot P$ for some nonsingular S. He can then employ Sendrier's attack on the code equivalence problem and obtain T_{mask} and T_{pure}, thus cracking the cryptosystem.

   (We note in passing that known quantum algorithms are ill-suited to the code equivalence problem [30, 31].)

2. If the cryptanalyst learns any particular dependence y then he obtains $T_{\text{pure}} \cdot y = \sum_{k:y_k=1} a_k \otimes b_k$. If y is a low-weight vector then this is a low-rank matrix and therefore discloses information about the a_k's and b_k's in the support of y.

   Of course, the cryptanalyst is not provided with dependences y's. However, he does have a means of gaining access to such y's precisely in the case that they are most useful—namely in the case that they are of low weight. This is because if y is a dependence of weight r, then the matrix $T \cdot y = T_{\text{pure}} \cdot y$ is of rank $\leq r$. The *Min-Rank* problem with input (T, r) is the problem of finding linear combinations of the slices of T that form matrices of rank at most r. (Or as a decision problem, whether such linear combinations exist.) Although this problem is hard for general r, as well as for the particular setting $r = n - 1$, and although the least r is even NP-hard to approximate to within factor 527/520 [14], the problem is nonetheless easy for small r. Indeed, Min-Rank attacks have been successful against previous MQ cryptosystems [49]. There are two known forms of this attack. One, the "kernel attack" [20, 21, 49, 40] finds all rank-r linear combinations in time $2^{\ell r/n}\text{poly}(\ell n)$. (Actually n in this expression is more properly $n+\rho$ but this does not affect the asymptotics because $\rho \leq n$.) The other is the "small minors" attack which solves the system of equations consisting of setting the determinants of all $(r + 1) \times (r + 1)$ minors (of a general linear combination of the slices) equal to 0. By a Gröbner/relinearization approach (see, e.g., [36]) this can usually be solved in time $\text{poly}(\binom{n}{r+1})$.

The second of these items is the more imminent threat to the cryptosystem since, first, it can obtain some information about the private key without necessarily having to crack the key entirely; and second, because it imposes a clear design constraint, namely, that low-weight codewords in $G^{\perp}$ leak information. That is why we have designed the "ideal" cryptosystem with the requirement that $G^{\perp}$ have no codewords of weight $\leq d$ for $d > 2(n + \rho)$. This is a conservative design: it is not clear that having a small number of low-weight vectors is a significant security threat.

We note that if the small-minors attack is indeed the best attack on this system then d is, up to a constant factor, the security parameter (log of the attack time) for attacks on the key. (As discussed in Sec. 3.1, the security parameter for message attacks appears to be lower, approximately $\sqrt{n}$.)

The setting of d at $d > 2(n+\rho)$ is not made only in order to set this as the key security parameter. This threshold has another attractive property, which we will show in Sec. 7: the sum of $2(n+\rho)$ random rank 1 matrices of dimensions $(n+\rho) \times (n+\rho)$ over $GF2$ (i.e., the matrices $a_k \otimes b_k$ in a linear dependence) has a distribution that is exponentially close to the uniform distribution. There is therefore exponentially little information released about the private key from any such sum, even if the linear dependence is known.

3.2.4 An Absent Structural Property: Basis for the Code

The leading attacks on the McEliece cryptosystem are not based on uncovering the scrambling permutation as in [77, 78], but on decoding noisy codewords in a random linear code. This attack is performed by reducing to the problem of finding a low-weight vector in the linear space that is spanned by the scrambled code (which is public) and the ciphertext. For more information see [81, 15, 16, 10]. There is no obvious way to conduct this attack in our system for the simple reason that the generator matrix is not public.

4 Known Codes

Our requirements in Sec. 2 for the code G demand that both the code and its dual be "good" codes, and there is tension between these two requirements. Nonetheless, it is not hard to see that for sufficiently large constant c, we can satisfy the requirements by letting $\ell = c\kappa = c^2 n$ and choosing G to be a random linear code.

In spite of this existence proof, no construction is known of a code G which has these properties and is also efficiently decodable. The difficulty lies in the small (i.e., binary) alphabet. Requiring δ to lie in the range $\delta > 1/4$ precludes (because of the Plotkin bound) G from being a minimum-distance code, which is what one obtains from all algebraic constructions. Even below the Plotkin threshold of $1/4$ the problem remains challenging.

The best current construction is due to Guruswami (see the appendix of [80]) which produces a family of codes G satisfying our requirements with the parameters

$$\text{rate of G:} \qquad \kappa/\ell = 1/2, \tag{6}$$

$$\text{fractional minimum distance of } G\text{:} \qquad \delta = 1/60, \tag{7}$$

$$\text{fractional minimum distance of } G^{\perp}\text{:} \qquad d/\ell = 1/18. \tag{8}$$

The implicit trade-off in these parameters is not ideal for us. The efficiency of the cryptosystem degrades only moderately as κ/ℓ and d/ℓ decrease; by contrast, the public key has to increase in size drastically in to accommodate primal distance $1/60$, as we describe in the next section.

We hope developments in coding theory will produce different parameter trade-offs, and thereby improve the efficiency of our design. There is no in-principle obstacle to obtaining efficiently-decodable codes with the parameters we have specified; and even restricting to algebraic constructions which produce minimal-distance guarantees, one may hope for $\delta > 1/8$.

5 Best Current Implementation: Higher Order Tensors

The only way that we know to realize our proposal with currently-available codes is by applying Guruswami's construction. Of course this will not work with the ideal proposal as described in Sec. 2, which requires $\delta > 1/4$. However, a very simple modification suffices. Rather than using "order-3 tensors" in the construction we use "order-7 tensors". Unfortunately this requires that the public key of size $O(n^3)$ of the ideal proposal, be replaced with a public key of size $O(n^7)$. But it is an in-principle realization with current technology of the cryptosystem/trap-door.

To be specific, let X_0 be as in Sec. 2, but now define:

Auxiliary randomization spaces $R_0 \cong \ldots \cong R_5 \cong (GF2)^\rho$.

"Input image" spaces $X_1 \cong \ldots \cong X_5 \cong (GF2)^{n_1}$.

$X = X_0 \oplus R_0$ and $V_i = X_i \oplus R_i$ for $1 \le i \le 5$.

Select M uniformly in $X^\dagger \otimes V_1^\dagger \otimes \ldots \otimes V_5^\dagger \otimes W$.

As before, P is a random $\ell \times \ell$ permutation matrix, and $T_{\rm mask} = M \cdot G \cdot P$.

$T_{\rm pure}$ is constructed by selecting, uniformly, $a_1, \ldots, a_\ell \in X^\dagger$, and $b_{i,1}, \ldots, b_{i,\ell} \in V_i^\dagger$ for $1 \le i \le 5$. Then $T_{\rm pure} = \sum_{k=1}^{\ell} a_k \otimes b_{1,k} \otimes \ldots \otimes b_{5,k}$.

D_i $(1 \le i \le 5)$ are chosen in respectively $\mathrm{Hom}(X_0, X_i)$.

As before the encryption of $x \in X_0$ is (with randomly chosen $r_i \in R_i$), is the following vector in W:

$$z = ((x + r_0) \otimes (x \cdot D_1 + r_1) \otimes \ldots \otimes (x \cdot D_5 + r_5)) \cdot T$$

or in Einstein notation,

$$z_k = (x + r_0)^{i_0} (x \cdot D_1 + r_1)^{i_1} \cdots (x \cdot D_5 + r_5)^{i_5} T_{i_1 \ldots i_5 k}$$

z is a codeword plus a noise vector whose weight is concentrated around $\ell/64$. Guruswami's code is powerful enough to efficiently identify the noise vector, i.e., $z_{\rm pure}$. Now, provided $\ell > 64(n + \rho)$, we can w.h.p. find in $z_{\rm pure}$ a set K of $(n + \rho + \ell/64)/2$ coordinates k for which $(z_{\rm pure})_k = 1$, and then solve the linear system

$$(x + r_0) \cdot a_k = 1 \qquad \text{for } k \in K$$

to obtain x.

As noted above, the public key in this implementation is large enough that it is not at present competitive with other post-quantum cryptosystems. Nevertheless it is still gives honest users "exponential advantage" over attackers, by which we mean that the following ratio is bounded away from 0:

$$\text{"Advantage" of cryptosystem} = \liminf_n \frac{A(n)}{H(n)} \tag{9}$$

where, on inputs of length n,

$$H(n) = \log[\text{honest player's work}] \tag{10}$$

$$A(n) = \log\log[\text{adversary's work with best known (incl. quantum) methods}] \tag{11}$$

The question remains whether this "advantage" measure of asymptotic quality of the cryptosystem (which we have adapted from a discussion in [51]) can be further improved to make the system competitive. The most obvious attack is to revisit Guruswami's construction. Two more speculative suggestions will be discussed in the next section.

6 Potential Variations of the Proposal

6.1 Larger Alphabet

It is not entirely essential that the construction be performed over $GF2$; one may substitute larger fields of characteristic 2, provided the "pure" vectors $a_k \otimes b_k$ now become $\alpha_k a_k \otimes b_k$ for α_k's which may be chosen randomly in the extension field, and a_k, b_k's which are vectors of 0's and 1's in the extension field. Also the input x remains a vector of 0's and 1's in the extension field.

This approach allows considerable additional latitude in the code design. It remains unknown whether this helps to obtain the dual code properties (which must still hold over the base field $GF2$).

6.2 Dual Code with Relaxed Properties

The "advantage" of the cryptosystem (as defined in Sec. 5) will remain positive even if we reduce d, the minimum weight of the dual code, to $d = n^c$ for some $c < 1$. This weakening of the security parameter may make it possible to obtain better error correcting codes. However, it does eliminate the statistical-security guarantee coming from Sec. 7, and now, a Min-Rank attack identifying linear combinations of slices having matrix rank $\leq d$, will also be able (at least information-theoretically although perhaps not efficiently) to obtain some information about the terms $a_k \otimes b_k$ contributing to the linear combination.

Alternatively—and more hypothetically than the previous suggestion—it may be possible to relax the requirement that the dual code has minimum weight d, in favor of the requirement that it have very few codewords of weight less than d. This makes many more codes possible. There will now be a relatively small number of small sets K for which we learn the sum $\sum_{k \in K} a_k \otimes b_k$. *If* this is the total extent of compromise to the cryptosystem, then message security may still be easy to ensure merely by a preprocessing hash (e.g., padding x with a small number of random bits and then applying a random, but fixed and public, linear transformation).

7 Convergence of Sums of Random Rank 1 Matrices to Uniform

In this section we justify our statement at the end of Sec. 3.2.3.

Proposition 1. *Form an $n \times n$ $GF2$ matrix by summing m matrices $a_k \otimes b_k$, each chosen by selecting a_k, b_k uniformly at random. For $m = cn$, $c > 2$, the distribution of the resulting matrix is within variation distance $2^{1-(c-2)n}$ of the uniform distribution on $n \times n$ matrices.*

Proof. Consider the group of $n \times n$ $GF2$-matrices under addition; this is isomorphic to $(\mathbb{Z}/2)^{n^2}$. The characters of the group are in bijective correspondence with matrices M,

$$\chi_M(N) = (-1)^{\sum_{1 \leq i,j \leq n} M_{ij} N_{ij}}.$$

We are implementing a random walk on this group: a single step is an addition of $u \otimes v$ for u, v uniformly distributed. Letting P be this single-step distribution on the group, we are to show that P convolved with itself m times is close to uniform. Write

$$\chi_M(P) = 2^{-2n} \sum_{u,v} (-1)^{\sum_{1 \leq i,j \leq n} M_{ij} u_i v_j}.$$

For nonzero M, fix any nonzero row i of M. For any v such that $\sum M_{ij}v_j = 1 \bmod 2$ (which is true of half the vectors v), the involution $u \to u + e_i$ (flipping the i'th bit) is a bijection between pairs (u, v) with $\sum_{1\le i,j\le n} M_{ij}u_iv_j = 0 \bmod 2$ and pairs (u, v) with $\sum_{1\le i,j\le n} M_{ij}u_iv_j = 1 \bmod 2$. Consequently,

$$-1/2 \le \chi_M(P) \le 1/2.$$

A more careful version of this argument is to consider v according to whether or not its dot product with every row of M is 0. If this is the case, then for every u, $\sum_{1\le i,j\le n} M_{ij}u_iv_j = 0 \bmod 2$. If this is not the case, then the same involution argument may be applied to the first row of M for which the dot product is 1. Therefore,

$$\chi_M(P) = 2^{-\operatorname{rank} M}.$$

Taking m steps of this walk gives a distribution P^{*m} whose Fourier coefficients are

$$\chi_M(P^{*m}) = 2^{-m \operatorname{rank} M}.$$

The Fourier coefficient of the uniform distribution is also 1 at $M = 0$, and is 0 elsewhere, so the variation distance of our distribution from uniform is equal to

$$2^{-n^2} \sum_N \left| \sum_{M\ne 0} 2^{-m \operatorname{rank} M} (-1)^{\sum_{1\le i,j\le n} M_{ij}N_{ij}} \right|.$$

The number of matrices M of rank r is bounded by 2^{2nr} (this is a wasteful estimate for large r but no matter), so the variation distance is

$$\begin{aligned}
&\le 2^{-n^2} \sum_N \left| \sum_{r=1}^{n} 2^{2nr} 2^{-mr} \right| \\
&= \sum_{r=1}^{n} 2^{(2-c)nr} \\
&\le 2^{1-(c-2)n}
\end{aligned}$$

(Incidentally, observe that almost all of the contribution comes from the rank-one Fourier coefficients.) □

8 Discussion

Like many other cryptosystems, we have no proofs of security. This is an obvious but challenging goal for future work. Indeed, we have provided no proofs of *anything* apart from an information-theoretic bound in Sec. 7.

However, we feel the proposal is worthy of serious consideration on the basis of its resistance to known or envisaged algorithmic developments (especially focusing on post-quantum security):

1. First, the new proposal is not apparently vulnerable to defeat of the McEliece cryptosystem. While that system appears so far secure when used with randomly-chosen Goppa codes, the originally-proposed parameters have been shown to be insufficient (see [10], based on an attack in [81]); also, a recent attack [37] has been effective, in a certain parameter range, against the

problem of distinguishing a random Goppa code generator matrix from a random matrix; the hardness of this problem had been regarded as one of the bases for security of the McEliece cryptosystem. So it seems quite desirable to have a code-based cryptosystem in which the generator matrix of the scrambled code is not provided as part of the public key.

2. Second, the post-quantum security of lattice cryptosystems is not as convincing as their classical security. Lattice cryptosystems are particularly attractive due to the fact they have worst-case rather than average-case security reductions [1]. However their quantum security rests on the hardness of the hidden subgroup problem (HSP) in dihedral groups [72]. This hardness assumption is in question because two positive results are known about dihedral HSP:

   (a) Single-register coset measurements are information-theoretically sufficient to solve the problem [33]. By contrast this is known to be false for more general HSP problems (e.g., in S_n) [45, 42, 61, 44]. Proposed multi-register measurements [34, 6, 7] do not seem to lead toward an efficient algorithm.

   (b) An elegant and nontrivial algorithm is known for the problem [50] (and see [73]), running in time $\sim 2^{\sqrt{n}}$ on groups of size 2^n. By contrast it is known that this algorithm does not work in the symmetric group S_n [62].

3. Third, the trap-door is completely different from that offered in Patarin's MQ cryptosystem HFEv$^-$, and this offers at least diversity in the sources of hardness within MQ cryptography. HFEv$^-$ is in all likelihood secure, being the survivor of a series of attacks, but at the same time it is of course possible that that history is not at its end. To briefly recap the relevant sequence of MQ proposals (for a survey through 2005 see [94]): [53] was cracked by Patarin [69], who then replaced that method by a more flexible generalization [70] called HFE; the most basic form of HFE has been cracked [49, 28, 38, 41], as have some variants [32], but the variant HFEv$^-$ (used in Quartz) appears to remain viable with appropriate parameters [24, 29], although perhaps not as efficient as one might wish [90].

   As discussed earlier in the paper, all MQ cryptosystems are subject to attack by Gröbner basis algorithms for solving systems of polynomial equations, the leading method currently being Faugère's F5 algorithm; the question is whether the systems of polynomials generated by the cryptosystem have features that enable such an algorithm to run in sub-exponential time.

   Quantum algorithms have not been shown to possess any advantage over classical algorithms in solving systems of polynomial equations (other than the slight Grover search advantage [43]), despite a few successful attacks (see [17, 18] for some representatives) on problems that lack obvious linear, abelian, or normal group structure. For this reason, MQ cryptosystems are particularly promising for post-quantum cryptography.

Finally, we hope our proposal refocuses attention on some long-standing questions:

1. What is the hardness of approximating tensor rank? Although Håstad [46] showed that exact rank computation is hard, it is not even known, say, that the rank is hard to approximate within $\pm \log n$. Such additive hardness might be attempted without necessarily having to explicitly produce proofs of high rank for tensors, which is the key barrier for attempts to prove that multiplicative factors of approximation are hard [4]. (This is an imprecise discussion; it is possible to envision hardness results operating within the regime of rank

below $3n$, where explicit tensor constructions are already known [79, 4, 91]. An additive hardness of approximation could then be translated into a, perhaps small, multiplicative hardness of approximation.)

2. What is the distribution of the rank of random tensors?

Acknowledgments

Thanks to the organizers of post-quantum cryptography workshops at Dagstuhl and the Lorentz Center for creating stimulating environments in which some of these ideas were developed. I am grateful to the participants of those meetings, in particular Enrico Thomae, for comments. Thanks also to Alex Vardy, Madhu Sudan, Venkat Guruswami and Zvika Brakerski for helpful discussions.

References

[1] M. Ajtai. Generating hard instances of lattice problems (extended abstract). In *Proc. 28th Annual ACM STOC*, pages 99–108, 1996.

[2] M. Ajtai and C. Dwork. A public-key cryptosystem with worst-case/average-case equivalence. In *Proc. 29th Annual ACM STOC*, pages 284–293, 1997.

[3] M. Albrecht, C. Cid, J.-C. Faugère, and L. Perret. On the relation between the MXL family of algorithms and Gröbner basis algorithms. Cryptology ePrint Archive, Report 2011/164, 2011.

[4] B. Alexeev, M. Forbes, and J. Tsimerman. Tensor rank: some lower and upper bounds. (Preprint arXiv:1102.0072v1), 2011.

[5] G. Ars, J.-C. Faugère, H. Imai, M. Kawazoe, and M. Sugita. Comparison Between XL and Gröbner Basis Algorithms. In *ASIACRYPT*, pages 338–353, 2004.

[6] D. Bacon, A. Childs, and W. van Dam. From optimal measurement to efficient quantum algorithms for the hidden subgroup problem over semidirect product groups. In *Proc. 46th Annual IEEE FOCS*, pages 469–478, 2005. (Preprint quant-ph/0504083).

[7] D. Bacon, A. Childs, and W. van Dam. Optimal measurements for the dihedral hidden subgroup problem. *Chicago J. Theoret. Comput. Sci.*, 2:1–25, 2006. (Preprint quant-ph/0501044).

[8] M. Bardet, J.-C. Faugère, B. Salvy, and P.-J. Spaenlehauer. On the complexity of solving quadratic boolean systems. arXiv:1112.6263, 2011.

[9] M. Bardet, J.-C. Faugère, B. Salvy, and B.-Y. Yang. Asymptotic behaviour of the degree of regularity of semi-regular polynomial systems. In *Eighth International Symposium on Effective Methods in Algebraic Geometry (MEGA)*, 2005.

[10] D. J. Bernstein, T. Lange, and C. Peters. Attacking and Defending the McEliece Cryptosystem. In *Proceedings of the 2nd International Workshop on Post-Quantum Cryptography*, PQCrypto '08, pages 31–46, Berlin, Heidelberg, 2008. Springer-Verlag. Cryptology ePrint Archive, Report 2008/318.

[11] D. Bini, M. Capovani, F. Romani, and G. Lotti. O($n^{2.7799}$) complexity for n*n approximate matrix multiplication. *Inf. Process. Lett.*, 8(5):234–235, 1979.

[12] A. Blum, A. Kalai, and H. Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. *J. ACM*, 50(4):506–519, 2003.

[13] J. Buchmann, J. Ding, M. S. E. Mohamed, and W. S. A. E. Mohamed. MutantXL: solving multivariate polynomial equations for cryptanalysis. Dagstuhl seminar proceedings 09031, http://drops.dagstuhl.de/opus/volltexte/2009/1945, 2009.

[14] J. F. Buss, G. S. Frandsen, and J. Shallit. The computational complexity of some problems of linear algebra. *J. Comput. Syst. Sci.*, 58(3):572–596, 1999.

[15] A. Canteaut and F. Chabaud. A new algorithm for finding minimum-weight words in a linear code: application to McEliece's cryptosystem and to narrow-sense BCH codes of length 511. *IEEE Transactions on Information Theory*, 44(1):367–378, January 1998.

[16] A. Canteaut and N. Sendrier. Cryptoanalysis of the Original McEliece Cryptosystem. In *ASIACRYPT*, pages 187–199, 1998.

[17] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. A. Spielman. Exponential algorithmic speedup by a quantum walk. In *Proc. 35th Annual ACM STOC*, pages 59–68, New York, NY, USA, 2003. ACM.

[18] A. M. Childs, L. J. Schulman, and U. V. Vazirani. Quantum algorithms for hidden nonlinear structures. In *Proc. 48th Annual IEEE FOCS*, pages 395–404, 2007.

[19] D. Coppersmith, D. P. Gamarnik, M. Hajiaghayi, and G. B. Sorkin. Random MAX SAT, Random MAX CUT, and Their Phase Transitions. *Random Structures and Algorithms*, 24(4):502–545, 2004.

[20] D. Coppersmith, J. Stern, and S. Vaudenay. Attacks on the birational permutation signature schemes. In *CRYPTO*, pages 435–443, 1993.

[21] D. Coppersmith, J. Stern, and S. Vaudenay. The security of the birational permutation signature schemes. *J. Cryptology*, 10(3):207–221, 1997.

[22] D. Coppersmith and S. Winograd. On the asymptotic complexity of matrix multiplication. *SIAM J. Comput.*, 11(3):472–492, 1982.

[23] D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. *Journal of Symbolic Computation*, 9(3):251–280, 1990.

[24] N. Courtois, M. Daum, and P. Felke. On the Security of HFE, HFEv- and Quartz. In *Proceedings of the 6th International Workshop on Theory and Practice in Public Key Cryptography*, PKC '03, pages 337–350, London, UK, 2003. Springer-Verlag.

[25] N. Courtois, A. Klimov, J. Patarin, and A. Shamir. Efficient algorithms for solving overdefined systems of multivariate polynomial equations. In *EUROCRYPT*, pages 392–407, 2000.

[26] N. Courtois and J. Patarin. About the XL Algorithm over GF(2). In *CT-RSA*, pages 141–157, 2003.

[27] N. Courtois and J. Pieprzyk. Cryptanalysis of block ciphers with overdefined systems of equations. In *ASIACRYPT*, pages 267–287, 2002.

[28] N. T. Courtois. The security of Hidden Field Equations (HFE). In D. Naccache, editor, *Cryptographer's Track at RSA Conference 2001, volume 2020 of Lecture Notes in Computer Science*, pages 266–281. Springer, 2001.

[29] N. T. Courtois. Short Signatures, Provable Security, Generic Attacks and Computational Security of Multivariate Polynomial Schemes such as HFE, Quartz and Sflash. Cryptology ePrint Archive, Report 2004/143, 2004.

[30] H. Dihn, C. Moore, and A. Russell. McEliece and Niederreiter cryptosystems that resist quantum Fourier sampling attacks. In *Proceedings of the 31st annual conference on Advances in cryptology*, CRYPTO'11, pages 761–779, Berlin, Heidelberg, 2011. Springer-Verlag.

[31] H. Dihn, C. Moore, and A. Russell. Quantum Fourier sampling, Code Equivalence, and the quantum security of the McEliece and Sidelnikov cryptosystems. arXiv:1111.4382, 2012.

[32] V. Dubois, L. Granboulan, and J. Stern. Cryptanalysis of HFE with Internal Perturbation. In *Public Key Cryptography*, pages 249–265, 2007.

[33] M. Ettinger and P. Høyer. On quantum algorithms for noncommutative hidden subgroups. *Advances in Applied Mathematics*, 25(3):239–251, 2000. (LANL preprint quant-ph/9807029, 1998).

[34] M. Ettinger, P. Høyer, and E. Knill. Hidden subgroup states are almost orthogonal. (LANL preprint quant-ph/9901034), 1999.

[35] J.-C. Faugère, M. Safey El Din, and P.-J. Spaenlehauer. Gröbner bases of bihomogeneous ideals generated by polynomials of bidegree (1,1): Algorithms and complexity. arXiv:1001.4004, 2010.

[36] J.-C. Faugère, M. Safey El Din, and P.-J. Spaenlehauer. On the complexity of the generalized minrank problem. *CoRR*, abs/1112.4411, 2011.

[37] J.-C. Faugère, V. Gauthier-Umaña, A. Otmani, L. Perret, and J.-P. Tillich. A distinguisher for high rate McEliece cryptosystems. In *Proc. Information Theory Workshop (ITW)*, pages 282–286. IEEE, Oct. 2011.

[38] J.-C. Faugère and A. Joux. Algebraic Cryptanalysis of Hidden Field Equation (HFE) Cryptosystems Using Gröbner Bases. In *CRYPTO*, pages 44–60, 2003.

[39] J.-C. Faugère, F. Levy-Dit-Vehel, and L. Perret. Cryptanalysis of MinRank. In *Proceedings of the 28th Annual conference on Cryptology: Advances in Cryptology*, CRYPTO 2008, pages 280–296, Berlin, Heidelberg, 2008. Springer-Verlag.

[40] L. Goubin and N. Courtois. Cryptanalysis of the TTM Cryptosystem. In *ASIACRYPT*, pages 44–57, 2000.

[41] L. Granboulan, A. Joux, and J. Stern. Inverting HFE Is Quasipolynomial. In *CRYPTO*, pages 345–356, 2006.

[42] M. Grigni, L. J. Schulman, M. Vazirani, and U. Vazirani. Quantum mechanical algorithms for the nonabelian hidden subgroup problem. *Combinatorica*, 24(1):137–154, 2004. (STOC '01).

[43] L. Grover. A fast quantum mechanical algorithm for database search. In *Proc. 28th Annual ACM STOC*, pages 212–219, 1996.

[44] S. Hallgren, C. Moore, M. Rötteler, A. Russell, and P. Sen. Limitations of quantum coset states for graph isomorphism. *J. ACM*, 57:34:1–34:33, November 2010.

[45] S. Hallgren, A. Russell, and A. Ta-Shma. Normal subgroup reconstruction and quantum computation using group representations. In *Proc. 32nd Annual ACM STOC*, pages 627–635, 2000.

[46] J. Håstad. Tensor rank is NP-complete. *J. Algorithms*, 11:644–654, December 1990.

[47] T. D. Howell. Global properties of tensor rank. *Linear Algebra and its Applications*, 22:9 – 23, 1978.

[48] Yun-Ju Huang, Feng-Hao Liu, and Bo-Yin Yang. Public-key cryptography from new multivariate quadratic assumptions. In *Public Key Cryptography - PKC 2012*, volume 7293 of *Lecture Notes in Computer Science*, pages 190–205. Springer, 2012.

[49] A. Kipnis and A. Shamir. Cryptanalysis of the HFE public key cryptosystem. In M. Wiener, editor, *Advances in Cryptology — CRYPTO 1999, volume 1666 of Lecture Notes in Computer Science*, pages 19–30. Springer, 1999.

[50] G. Kuperberg. A subexponential-time quantum algorithm for the dihedral hidden subgroup. *SIAM J. Comput.*, 35(1):170–188, 2005.

[51] V. Lyubashevsky and D. Micciancio. Asymptotically efficient lattice-based digital signatures. In *Proc. Theory of Cryptography Conference (TCC)*, pages 37–54, 2008.

[52] Vadim Lyubashevsky, Adriana Palacio, and Gil Segev. Public-key cryptographic primitives provably as secure as subset sum. In *Theory of Cryptography, 7th Theory of Cryptography Conference, TCC 2010, Zurich, Switzerland, February 9-11, 2010. Proceedings*, volume 5978 of *Lecture Notes in Computer Science*, pages 382–400. Springer, 2010.

[53] T. Matsumoto and H. Imai. Public quadratic polynomial-tuples for efficient signature-verification and message-encryption. In *EUROCRYPT*, pages 419–453, 1988.

[54] R. J. McEliece. A public-key cryptosystem based on algebraic coding theory. Technical Report 42-44, Jet Propulsion Lab, Pasadena CA, 1978.

[55] D. Micciancio. Almost perfect lattices, the covering radius problem, and applications to Ajtai's connection factor. *SIAM J. Comput.*, 34:118–169, 2004. (STOC '02 and CCC '02).

[56] D. Micciancio. Generalized compact knapsacks, cyclic lattices, and efficient one-way functions. *Computational Complexity*, 16(4):365–411, 2007. (FOCS '02).

[57] D. Micciancio and S. Goldwasser. *Complexity of Lattice Problems*. Kluwer, 2002.

[58] D. Micciancio and O. Regev. Worst-case to average-case reductions based on Gaussian measures. *SIAM J. Comput.*, 37(1):267–302, 2007. (FOCS '04).

[59] D. Micciancio and O. Regev. Lattice-based cryptography. In D. J. Bernstein, J. Buchmann, and E. Dahmen, editors, *Post-Quantum Cryptography*, pages 147–191. Springer, 2009.

[60] M. S. E. Mohamed, W. S. A. E. Mohamed, J. Ding, and J. Buchmann. MXL2: Solving Polynomial Equations over GF(2) Using an Improved Mutant Strategy. In *PQCrypto*, pages 203–215, 2008.

[61] C. Moore, A. Russell, and L. J. Schulman. The symmetric group defies strong Fourier sampling. *SIAM J. Comput.*, 37(6):1842–1864, 2008. (FOCS '05).

[62] C. Moore, A. Russell, and P. Sniady. On the impossibility of a quantum sieve algorithm for graph isomorphism. *SIAM J. Comput.*, 39(6):2377–2396, 2010. (STOC '07).

[63] H. Niederreiter. Knapsack-type cryptosystems and algebraic coding theory. *Problems of Control and Information Theory (Problemy Upravlenija i Teorii Informacii)*, 15(2):159–166, 1986.

[64] P. Pamfilos. On the maximum rank of a tensor product. *Acta Math. Hung.*, 45(1-2):95–97, 1985.

[65] V. Pan. Strassen's algorithm is not optimal. In *Proc. 19th Annual IEEE FOCS*, pages 166–176, 1978.

[66] V. Pan. Field extension and trilinear aggregating, uniting and canceling for the acceleration of matrix multiplications. In *Proc. 20th Annual IEEE FOCS*, pages 28–38, 1979.

[67] V. Pan. New fast algorithms for matrix operations. *SIAM J. Comput.*, 9:321–342, 1980.

[68] V. Pan. New combinations of methods for the acceleration of matrix multiplication. *Comput. Math. with Appl.*, 7:73–125, 1981.

[69] J. Patarin. Cryptoanalysis of the Matsumoto and Imai Public Key Scheme of Eurocrypt'88. In *CRYPTO*, pages 248–261, 1995.

[70] J. Patarin. Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): Two New Families of Asymmetric Algorithms. In *EUROCRYPT*, pages 33–48, 1996.

[71] O. Regev. New lattice-based cryptographic constructions. *J. ACM*, 51(6):899–942, November 2004. (STOC '03).

[72] O. Regev. Quantum computation and lattice problems. *SIAM J. Comput.*, 33(3):738–760, June 2004. (FOCS '02).

[73] O. Regev. A subexponential time algorithm for the dihedral hidden subgroup problem with polynomial space. (Preprint arXiv:quant-ph/0406151v1), 2004.

[74] O. Regev. On lattices, learning with errors, random linear codes, and cryptography. In *Proc. thirty-seventh STOC*, pages 84–93, New York, NY, USA, 2005. ACM.

[75] F. Romani. Some properties of disjoint sums of tensors related to matrix multiplication. *SIAM J. Comput.*, 11:263–267, 1982.

[76] A. Schönhage. Partial and total matrix multiplication. *SIAM J. Comput.*, 10(3):434–455, 1981.

[77] N. Sendrier. An algorithm for finding the permutation between two equivalent binary codes. Technical Report RR-2853, INRIA, April 1996.

[78] N. Sendrier. Finding the permutation between equivalent linear codes: The support splitting algorithm. *IEEE Transactions on Information Theory*, 46(4):1193–1203, 2000.

[79] A. Shpilka. Lower bounds for matrix product. *SIAM J. Comput.*, 32(5):1185–1200, 2003.

[80] A. Shpilka. Constructions of low-degree and error-correcting ϵ-biased generators. *Comput. Complex.*, 18:495–525, 2009.

[81] J. Stern. A method for finding codewords of small weight. In *Proceedings of the 3rd International Colloquium on Coding Theory and Applications*, pages 106–113, London, UK, 1989. Springer-Verlag.

[82] A. J. Stothers. *On the complexity of matrix multiplication.* PhD thesis, U. Edinburgh, 2010.

[83] V. Strassen. Gaussian elimination is not optimal. *Numer. Math.*, 13:354356, 1969.

[84] V. Strassen. The asymptotic spectrum of tensors and the exponent of matrix multiplication. In *Proc. 27th Annual IEEE FOCS*, pages 49–54, 1986.

[85] V. Strassen. Relative bilinear complexity and matrix multiplication. *J. Reine Angew. Mathe.*, 375:406–443, 1987.

[86] E. Thomae. Quo Vadis Quaternion? Cryptanalysis of Rainbow over Non-Commutative Rings. *IACR Cryptology ePrint Archive*, 2012/270, 2012.

[87] E. Thomae and C. Wolf. Solving systems of multivariate quadratic equations or: from relinearization to MutantXL. Cryptology ePrint Archive 2010/596.

[88] E. Thomae and C. Wolf. Roots of square: Cryptanalysis of double-layer square and square+. In *PQCrypto*, pages 83–97, 2011.

[89] E. Thomae and C. Wolf. Cryptanalysis of Enhanced TTS, STS and All Its Variants, or: Why Cross-Terms Are Important. In *AFRICACRYPT*, pages 188–202, 2012.

[90] S. Tsujii, M. Gotaishi, K. Tadaki, and R. Fujita. Proposal of a Signature Scheme Based on STS Trapdoor. In *PQCrypto*, pages 201–217, 2010.

[91] B. Weitz. An improvement on rank of explicit tensors. (Preprint arXiv:1102.0580v2), 2011.

[92] V. V. Williams. Breaking the Coppersmith-Winograd barrier. Manuscript, 2011.

[93] C. Wolf, A. Braeken, and B. Preneel. Efficient Cryptanalysis of RSE(2)PKC and RSSE(2)PKC. In *4th International Conference on Security in Communication Networks (SCN)*, pages 294–309, 2004.

[94] C. Wolf and B. Preneel. Taxonomy of public key schemes based on the problem of multivariate quadratic equations. Cryptology ePrint Archive, Report 2005/077, 2005.

[95] B.-Y. Yang and J.-M. Chen. Theoretical Analysis of XL over Small Fields. In *ACISP*, pages 277–288, 2004.