Towards provable security of the Unbalanced Oil and Vinegar signature scheme under direct attacks

Stanislav Bulygin[1], Albrecht Petzoldt[2], and Johannes Buchmann[1,2]

[1] Center for Advanced Security Research Darmstadt - CASED
Mornewegstraße 32, 64293 Darmstadt, Germany
{johannes.buchmann,Stanislav.Bulygin}@cased.de

[2] Technische Universität Darmstadt, Department of Computer Science
Hochschulstraße 10, 64289 Darmstadt, Germany
{apetzoldt,buchmann}@cdc.informatik.tu-darmstadt.de

Abstract. In this paper we show that solving systems coming from the public key of the Unbalanced Oil and Vinegar (UOV) signature scheme is on average at least as hard as solving a certain quadratic system with completely random quadratic part. In providing lower bounds on direct attack complexity we rely on the empirical fact that complexity of solving a non-linear polynomial system is determined by the homogeneous part of this system of the highest degree. Our reasoning explains, in particular, the results on solving the UOV systems presented by J.-C. Faugere and L. Perret at the SCC conference in 2008.

Keywords: Multivariate Cryptography, UOV Signature Scheme, provable security, security reduction, semi-regular sequence

1 Introduction

Multivariate public key cryptography is one of the alternatives for the post-quantum era, i.e. when a large enough quantum computer is built and the public key cryptosystems used today (RSA, ECC, El Gamal) are broken. Other than resistance to quantum computer attacks, multivariate public key cryptosystems (MPKCs) enjoy other useful properties. In particular, they are quite fast compared to conventional schemes and require only very moderate resources. This makes MPKCs excellent candidates for use in resource constraint devices, like RFIDs and smart cards. Still there are two issues that pose obstacles on the way of using MPKCs. The first one is the issue of key sizes. The second problem is that MPKC proposals are being broken on a regular basis, which weakens believe in a possibility of constructing both secure and efficient MPKC.

Quite a few attempts have been undertaken in order to tackle the first problem. Mainly, the researchers concentrated on reducing the secret key size. In the recent paper [23] the authors undertook an attempt to reduce the public key size, based on yet unbroken (under proper parameter choice) UOV scheme [15].

There has been no lack in proposals of MPKCs, see [7,10] for an overview. On the down side for the designers, the cryptanalytic progress has also been substantial. New proposals aim mainly at fixing problems exposed by the cryptanalysis, but then it often happens that "fixed" proposals get broken again (observe for example the sequence Matsumoto-Imai scheme [17] → its cryptanalysis [20] → HFE [21] → cryptanalysis of HFE challenge 1 [13] or a less known sequence MFE cryptosystem [26] → its SOLE cryptanalysis [8] → improved MFE [25] → cryptanalysis of the improved versions [5]; there are many more such "sequences"). There is a need in theoretical backing of design principles used in constructions of MPKCS. Note that for the classical cryptosystems one has some empirical certainty in the security of these systems. Namely, it is believed that breaking RSA in the classical computational model is as hard as factoring. For the ECC there is a believe that there exists no sub-exponential algorithm for solving the discrete logarithm problem in a group of points of an elliptic curve. Even better arguments are provided in the lattice-base and hash-based cryptography. For example, rigorous security reductions are provided for the cryptosystems based on random lattices as well as more compactly representable ideal lattices. Some attempts on providing "provability" or "reducibility" for MPKCs were undertaken by N. Courtois in his note [6]. There for providing security proofs he assumed strong properties of certain multivariate constructions themselves. It would be desirable instead to anchor security to some known problem(s) on which MPKCs are built. In this paper this will be a weaker version of the MQ-problem. Surprisingly enough, the methods used by the authors in [23] are also applicable to tackle the problem of "provable security" in the case of MPKCs as we will show.

The object of this paper is the Unbalanced Oil and Vinegar (UOV) scheme proposed in [15]. Note that for suitably chosen parameters (in particular, $v > o$, e.g. $v = 2o$, see Section 2) the progress in cryptanalysis of this scheme is connected mainly with the progress in solving generic quadratic systems over a finite field [4,14], which is one of the underlying hard problems the UOV is based on. Despite some progress, the above problem is still considered to be hard on average. Considering lack of structural attacks on the UOV for carefully chosen parameters, the system remains unbroken for more than ten years now. In this paper we show that breaking a UOV system directly is on average at least as hard as solving a quadratic system with a random quadratic part. We would like to be careful here on what we mean. What is meant is that using only direct (or general) attacks on the UOV, i.e. attacks based on Gröbner bases/XL-like, it is not possible to break the UOV if the parameters are large enough (and it is in principle possible to compute these). Still, our approach says nothing about structural attacks on the UOV. In particular, our approach says that the balanced Oil and Vinegar is secure against direct attacks, but it is a matter of the polynomial time algorithm to find an equivalent secret key [16]. So our claim is related to the direct solving attacks only. One, of course, should also be careful with "provably secure" in this context. If it is possible to solve a random quadratic system with certain parameters, this implies a jeopardy for a UOV scheme which anchors to such a system. In particular, in [14,4] it was shown that it is possible to forge signatures of the UOV with certain parameters. Still, moderate increase of parameters would render such an attack inefficient, due to high complexity of the anchoring problem.

The paper is organized as follows. In Section 2 we review the UOV scheme. Then in Section 3 we present the idea of [23] and show how it can be used for inserting a random (rather than partially cyclic) matrix in a UOV public key. Our reduction arguments follow in Section 4 where we present our main result in Theorem 1. Section 5 provides some lower bounds on direct attacks using results of the previous section. We conclude in Section 6.

2 The UOV signature scheme

The idea of the Oil and Vinegar trapdoor was first proposed by J. Patarin in [22] and stems from his cryptanalysis of the Matsumoto-Imai scheme [20].

Let K be a finite field. Let o and v be two integers and set $n = o+v$. Patarin suggested to choose $o = v$. The original scheme was broken by Kipnis and Shamir in [16], and it was recommended in [15] to choose $v > o$ (Unbalanced Oil and Vinegar (UOV)). Next we describe the idea of the UOV scheme.

The UOV scheme is a single field construction, so we work solely in the polynomial ring $K[X]$, where $X = \{x_1, \ldots, x_n\}$. We divide the variable set X into two sets: vinegar variables $(x_i)_{i\in V}, V = \{1, \ldots, v\}$ and oil variables $(x_i)_{i\in O}, O = \{v+1, \ldots, n\}$. Here $|V| = v, |O| = o$ and $v + o = n$. We define o quadratic polynomials $q_k(X) = q_k(x_1, \ldots, x_n)$ by

$$q_k(X) = \sum_{i\in V,\ j\in O} \alpha_{ij}^{(k)} x_i x_j + \sum_{i,j\in V,\ i\leq j} \beta_{ij}^{(k)} x_i x_j + \sum_{i\in V\cup O} \gamma_i^{(k)} x_i + \eta^{(k)}, k = 1, \ldots, o \qquad (1)$$

Note that oil and vinegar variables are not fully mixed, just like oil and vinegar in a salad dressing.

The map $\mathcal{Q} = (q_1(X), \ldots, q_o(X))$ can be easily inverted. First, we choose the values of the v vinegar variables $x_1, \ldots, x_v$ at random. Therewith we get a system of o linear equations in the o

variables $x_{v+1}, \ldots, x_n$ which can be solved by Gaussian elimination. If the system does not have a solution, choose other values of $x_1, \ldots, x_v$ and try again.

The public key $\mathcal{P}$ of the UOV scheme consists of o quadratic polynomials in n variables.

$$\begin{aligned} P &= (p^{(1)}, \ldots, p^{(o)}) \\ &= \left(\sum_{i=1}^{n} \sum_{j=i}^{n} p_{ij}^{(1)} x_i x_j + \sum_{i=1}^{n} p_i^{(1)} x_i + p_0^{(1)}, \ldots, \sum_{i=1}^{n} \sum_{j=i}^{n} p_{ij}^{(o)} x_i x_j + \sum_{i=1}^{n} p_i^{(o)} x_i + p_0^{(o)} \right) \end{aligned} \tag{2}$$

After having chosen an ordering on monomials, we can write down the public coefficients into an $o \times \frac{(n+1)\cdot(n+2)}{2}$-matrix M_P.

$$M_P = \begin{pmatrix} p_{11}^{(1)} & p_{12}^{(1)} & \ldots & p_{nn}^{(1)} & p_1^{(1)} & \ldots & p_n^{(1)} & p_0^{(1)} \\ \vdots & & & & & & & \vdots \\ p_{11}^{(o)} & p_{12}^{(o)} & \ldots & p_{nn}^{(o)} & p_1^{(o)} & \ldots & p_n^{(o)} & p_0^{(o)} \end{pmatrix}.$$

In the case of UOV, the public key is given as

$$\mathcal{P} = \mathcal{Q} \circ \mathcal{T}, \tag{3}$$

with an affine invertible map $\mathcal{T}$ and the central map $\mathcal{Q}$ as defined in (1).

Remark 1. In contrast to other multivariate schemes the second affine map $\mathcal{S}$ is not needed for the security of UOV and therefore is left out. So we indeed use $\mathcal{P} = \mathcal{Q} \circ \mathcal{T}$ and not $\mathcal{P} = \mathcal{S} \circ \mathcal{Q} \circ \mathcal{T}$.

Other than the attack on the balanced version [16] , there exists a number of attacks on different parameter choices, see e.g. [4,2]. Essentially, the UOV scheme remains unbroken. For example the parameter choice: $v = 2o, o = 26$ over the field $GF(2^8)$ is considered to be secure [2,4,7].

3 Inserting a random matrix in the UOV public key

Let $q_{ij}^{(k)}$ be the coefficients of quadratic terms of the central map polynomials from (1). Due to equations (2) and (3), we get the following equations for the coefficients of the quadratic terms of the public key:

$$p_{ij}^{(r)} = \sum_{k=1}^{n} \sum_{l=k}^{n} \alpha_{kl}^{ij} \cdot q_{kl}^{(r)} = \sum_{k=1}^{v} \sum_{l=k}^{n} \alpha_{kl}^{ij} \cdot q_{kl}^{(r)} \ (1 \le i \le j \le n, \ r = 1, \ldots, o) \tag{4}$$

with

$$\alpha_{kl}^{ij} = \begin{cases} t_{ki} \cdot t_{li} & (i = j) \\ t_{ki} \cdot t_{lj} + t_{kj} \cdot t_{li} & (i \ne j) \end{cases} \tag{5}$$

Note that the right hand side of equation (4) only contains coefficients of the quadratic terms of $\mathcal{Q}$ and coefficients of $\mathcal{T}$ and is linear in the former ones. The second "=" in equation (4) is due to the fact that all the q_{ij} $(i, j \in O)$ are zero.

Denote $D := \frac{v\cdot(v+1)}{2} + o \cdot v$. Let the monomials $x_i x_j, 1 \le i, j \le n$ be ordered w.r.t the given degree monomial ordering [1]. The given monomial ordering $<_{ord}$ also induces an ordering on the set of pairs $Pr = \{(i, j) | 1 \le i \le j \le n\}$, namely $(i', j') > (i'', j'')$ iff $x_{i'} x_{j'} >_{ord} x_{i''} x_{j''}$.

We define Q to be the $o \times D$ matrix containing the non-zero coefficients of the central polynomials with respect to the monomial ordering defined above. Additionally, we define a $D \times D$ matrix A containing the coefficients of the equations (4):

$$A = \left(\alpha_{kl}^{ij} \right),$$

[1] In fact we do not need a monomial ordering as is used in computer algebra; we just need some ordering of monomials. Still we prefer to work with monomial orderings in this paper.

where indices (k, l) are taken according to the monomial ordering as above and $1 \leq i \leq v,\ i \leq j \leq n$. Thus equation (4) yields

$$M' = Q \cdot A, \tag{6}$$

where M' is a submatrix of M_P composed of the first D columns.

In order to obtain a UOV scheme, we assign random values from K to the coefficients of $\mathcal{T}$. Then the entries of the matrix A can be computed by equation (5). Equation (6) yields a linear relation between the coefficients of $\mathcal{P}$ and $\mathcal{Q}$, given $\mathcal{T}$. To use this relation properly, we need the matrix A to be invertible. Practically for large enough fields (e.g. $K = GF(2^8)$) this property is satisfied in an overwhelming number of cases. Assuming A is invertible, we can prove the following proposition:

Proposition 1. *Given an $o \times D$ matrix B and an affine invertible map $\mathcal{T}$ such that the corresponding matrix A is invertible, it is possible to construct a UOV scheme with the secret key $(\mathcal{Q}, \mathcal{T})$ and the public key $\mathcal{P}$ with $M_P = (B|C)$, where C is a $(o \times ((n+1)(n+2)/2 - D))$-matrix.*

Proof. Under the assumption of A being invertible, equation (6) yields a bijection between the entries of M' and the quadratic coefficients of $\mathcal{Q}$. Therefore, if we assign the entries of M' the values of the matrix B, we get a uniquely determined quadratic part of the central map $\mathcal{Q}$. Since the linear part and constant terms of $\mathcal{Q}$ do not have any influence on the quadratic part of the public key, they can be chosen arbitrarily.

4 Security reduction

Lemma 1. *Let $\mathcal{T}$ be an invertible affine map which leads to an invertible transformation matrix A. Then, every UOV-scheme that has the affine map $\mathcal{T}$ as a part of the secret key can be obtained via the construction of Proposition 1.*

Proof. Let $(\mathcal{Q}, \mathcal{T})$ be a UOV scheme, such that $\mathcal{T}$ leads, via equation (5), to an invertible transformation matrix A. Let Q be the $o \times D$ matrix containing the non-zero quadratic coefficients of $\mathcal{Q}$ and L be the $o \times (n+1)$-matrix containing its linear coefficients and constant terms. Let B be defined as $B = Q \cdot A$. Since A was assumed to be invertible, this is an 1:1 relation between the matrices B and Q. We start with $(B, \mathcal{T})$ and follow the construction described in the Proposition. If the linear and constant part is chosen to be L, we will end up with the UOV scheme $(\mathcal{Q}, \mathcal{T})$.

In the classical UOV key generation we start with a random UOV central map $\mathcal{Q}$, random invertible affine $\mathcal{T}$ and then obtain a public key $\mathcal{P}$, which can be written with a matrix $M_P = (B|C)$. According to Lemma 1 we may equivalently start with B, the same $\mathcal{T}$ and end up with the same $\mathcal{Q}$ up to linear terms, which may be assigned arbitrarily. In this sense both constructions are equivalent. In order to provide a security reduction we will need the latter construction.

Let $(\mathcal{P}, \mathcal{Q}, \mathcal{T})$ be a UOV scheme obtained via Proposition 1. We impose the following monomial ordering. Let $Y, Z \subset X = \{1, \ldots, n\}$ be two disjoint subsets of X (note that X is now a set on indexes not to be confused with the variables set in Section 2) such that $X = Y \cup Z$. The sets Y and Z have cardinalities v and o resp., so that $o + v = n$. The monomial ordering with $x_1 > \cdots > x_n$ is then an ordering chosen in such a way that the following holds:

$$x_i x_j > x_k x_l > x_m x_p > x_u\ \forall i, j, k \in Y\ \forall l, m, p \in Z\ \forall u \in X. \tag{7}$$

One example of such an ordering is a weighted-degree ordering where each variable of Y has weight 3 and each variable of Z has weight 2. With this ordering quadratic monomials composed of Y-variables will be the largest, then follow "mixed" with variables from Y and Z, and finally those composed of Z-variables. We will need this ordering later in the proof of Proposition 2, in the follow-up procedure of the key generation, and in Theorem 1.

Now let $P(x)$ be polynomials of the public key $\mathcal{P}$. Let h be a hash value of the given document. The task of an attacker that wants to attack UOV directly is to find a solution of $P(x) = h$. Any solution to $P(x) = h$ provides a valid signature, therefore enables signature forgery.

Proposition 2. *Let $P(x)$ be polynomials of the public key $\mathcal{P}$ of a UOV scheme that is constructed by choosing completely random B and $\mathcal{T}$, following Proposition 1, and using an ordering satisfying (7). If it is possible to get a solution $x' = (x'_1, \ldots, x'_n)$ of $P(x) = h$, then it is possible to get a solution of a quadratic system of o equations and v variables with a completely random quadratic part.*

Proof. The public key $\mathcal{P}$ is represented by a matrix $M_P = (B|C)$, where the columns are ordered according to the chosen ordering that satisfies (7). Due to this ordering the public key polynomials may be written as

$$p^{(k)} = \sum_{i,j \in Y} a_{ij}^{(k)} x_i x_j + \sum_{i \in Y, j \in Z} b_{ij}^{(k)} x_i x_j + \sum_{i,j \in Z} c_{ij}^{(k)} x_i x_j + \sum_{i \in X} d_i^{(k)} x_i + e^{(k)}, \tag{8}$$

for $k = 1, \ldots, o$. Again according to the monomial ordering we have chosen, coefficients $a_{ij}^{(k)}$ and $b_{ij}^{(k)}$ are elements of the matrix B and therefore are chosen completely at random. Now a solution $x' = (x'_1, \ldots, x'_n)$ may be seen as $x' = ((x'_i)_{i \in Y}, (x'_j)_{j \in Z})$. Plug in values $(x'_j)_{j \in Z}$ for variables $(x_j)_{j \in Z}$ in (8). Therewith one obtains a quadratic system with o equations and v variables $(x_i)_{i \in Y}$ of the form

$$\tilde{p}^{(k)} = \sum_{i,j \in Y} a_{ij}^{(k)} x_i x_j + \sum_{i \in Y} \tilde{d_{ij}}^{(k)} x_i + \tilde{e}^{(k)}, k = 1, \ldots, o.$$

Note that coefficients $a_{ij}^{(k)}$ are completely random and are taken from the initial construction in (8). The system we need to solve is therefore

$$\tilde{P}((x_i)_{i \in Y}) = h. \tag{9}$$

So finding a solution x' of $P(x) = h$ provides a solution to (9), where $\tilde{P}$ has completely random quadratic part.

Proposition 2 seemingly provides a reduction for the problem of direct solving of $P(x) = h$ to the problem of solving a quadratic system with a completely random quadratic part. The way we presented the public key $\mathcal{P}$ this is not really true. Indeed, if the attacker knows the variable sets Y and Z he may simply fix the variables from Y ending up with a "non-random" system with variables from Z. In fact, if we suppose that the coefficients $c_{ij}^{(k)}$ are zero, then by fixing Y-variables the attacker ends up with a linear system as is the case for UOV maps. So the confusion here is that the attacker is not forced to solve (9) in order to solve $P(x) = h$.

Note that when the attacker uses Gröbner methods for solving, he/she would usually fix v variables first in order to end up with an $o \times o$ system. This is due to the fact that a random quadratic system with o equations and v variables over $GF(q)$ is expected to have q^{v-o} solutions. In order to be able to compute a solution it is preferable to "cut down" the solution space. By assigning values to some $v - o$ variables, the system still has o equations, but o variables, and is expected to have a unique solution, which is the found with Gröbner basis techniques. The idea of our reduction is to disguise the monomial ordering that was used and, in particular, the sets Y and Z. Below we show that after the process of fixing the attacker, at least on average, is intrinsically faced with solving a quadratic system which is at least as hard as a certain "random" one. For the reduction we need the following key generation procedure:

Key generation procedure:

1. Choose a $o \times D$ matrix B completely (and uniformly) at random.
2. Choose an affine map $\mathcal{T}$ at random. If it is not invertible, choose again.
3. Choose $Y \subset X, |Y| = v$ at random. Set $Z := X \setminus Y$.
4. Use the construction of Proposition 1 with a monomial ordering satisfying (7), obtain the central map $\mathcal{Q}$ and the matrix C. The secret key is $(\mathcal{Q}, \mathcal{T})$.

5. Let $M = (B|C)$ be a matrix with the columns indexed by monomials with degree up to 2 ordered with the ordering chosen in the previous step. Let M' be the matrix M whose columns are permuted according to the graded lexicographic ordering. The public key is the set of polynomials with the matrix $M_P = M'$.

Now the attacker observing the matrix M_P does not know which monomial ordering was used and what are the sets Y and Z. Therefore he is not able to figure out where the random part $\sum_{i,j\in Y} a_{ij}^{(k)} x_i x_j$ is (see proof of Proposition 2). Note also that except of a minor modification of the key generation procedure, the scheme stays essentially the same. In particular, one does not need to store the set Y: after obtaining a key pair, signature generation and verification are independent on a monomial ordering used.

Remark 2. A legitimate question here is whether an attacker observing the matrix M_P is able to figure out monomials $x_i x_j, i,j \in Z$ and therefore the sets Y and Z. In principle, he must be able to do so, since the coefficients of monomials $x_i x_j, i,j \in Z$ are not completely random, but obtained via the "reverse" computation, after computing $\mathcal{Q}$ from B and $\mathcal{T}$. It may be shown that the coefficients of the matrix C in the construction satisfy certain quadratic relations. So, in principle, by choosing a subset $S \subset X$ of cardinality o and checking if coefficients $x_i x_j, i,j \in S$ satisfy these quadratic relations, it may be possible to distinguish Z from other subsets of X. In order to do so one has to go through all o-subsets of n and this has complexity dominated by $\mathcal{O}(\binom{n}{o})$. It can be shown to be worse than one can do with system solving, as we proceed below.
Moreover, in the monomial ordering as in the above construction, we may also choose an arbitrary order of variables, unknown to the attacker. Then it is not really clear for the attacker, how to apply the quadratic relations that exist for the matrix C.
All in all, it seems that it is computationally impossible for the attacker to figure out the partition of X into Y and Z. It is a future research point to confirm this statement more rigorously.

The next theorem shows that the attacker applying the fixing+solving technique has to face some random system, at least on average.

Theorem 1. *Let $v = \alpha o, \alpha \geq 1$. Let*

$$P(x) = h \tag{10}$$

be an $o \times (o+v)$ system of public equations for a UOV scheme obtained with the procedure above. Suppose that the system (10) is solved by first fixing v variables (variables are chosen at random, as well as the values fixed) and then solving the $o \times o$ system, which is obtained after plugging in the fixed values in (10). Then solving (10) is on average at least as hard as solving an $o \times \frac{\alpha}{\alpha+1} o$ quadratic system with a completely random quadratic part.

Proof. Let Y and Z be disjoint variable sets as in the construction. So the public key is given by equations (8), where Y and Z are unknown to the attacker. The attacker fixes v variables to concrete values. Since $v = \alpha o$, we expect on average $\frac{\alpha}{\alpha+1} v$ variables to be fixed in Y and $\frac{1}{\alpha+1} v$ in Z. So there remains a set $Y_F \subset Y$ of variables in Y that are not fixed, $|Y_F| = v - \frac{\alpha}{\alpha+1} v = \frac{1}{\alpha+1} v = \frac{\alpha}{\alpha+1} o$. Denote the non-fixed variables in Z by Z_F. After plugging in the fixed values the attacker obtains a system

$$\tilde{P}(x_i | i \in Y_F \cup Z_F) = h. \tag{11}$$

Again note that the sets Y_F and Z_F are not known to the attacker. Let $x' = (x'_i | i \in Y_F \cup Z_F)$ be a solution of (11). Suppose the attacker is given the values of $x'_i, i \in Z_F$. After plugging in these values in (11) he/she obtains a system

$$\tilde{\tilde{P}}(x_i | i \in Y_F) = h \tag{12}$$

with a completely random quadratic part (all quadratic terms in (8) are "killed" except the ones with $i,j \in Y_F$). The system (12) has o equations and $|Y_F| = \frac{\alpha}{\alpha+1} o$ variables. Since the values $x'_i, i \in Z_F$ were given to the attacker, the actual solving of (11) is at least as hard as solving (12).

The corollary below specifies the above theorem to the choice of α that is used in the UOV to avoid the structural attack of [15].

Corollary 1. *If $v = 2o$, then the quadratic system with a completely random quadratic part from Theorem 1 has o equations and $\frac{2}{3}o$ variables. In other words, the ratio $\#eqs/\#vars = 3/2$ in this case.*

Let us discuss the above results. It may seem surprising that we have such a reduction to a hard problem, considering that the legitimate signer is able to get the solution x'. The catch here is that the system (12) is not completely random, only its quadratic part is. So it is not surprising that the signer, knowing the decomposition $\mathcal{P} = \mathcal{Q} \circ \mathcal{T}$ is able to get a solution which also yields a solution to (12). Similarly, the attacker, who via some structural attack is able to get the decomposition $\mathcal{P} = \mathcal{Q} \circ \mathcal{T}$ or an equivalent one, is also able to solve (12). If we consider only direct attacks, though, the situation is different. There is a quite strong experimental evidence that the complexity of solving a non-linear system is determined by its homogeneous part of the highest degree, e.g. [14,4]. In particular, in [3] an affine sequence of polynomials is defined to be semi-regular (practically speaking random) if its homogeneous part of the highest degree is semi-regular. Complexity estimates in [3] rely on the domination of the homogeneous highest degree part. If we take this domination assumption, we may state that from the point of view of Gröbner basis algorithms (in particular F5 [12]) complexity of solving (12) and a completely random quadratic system of o equations in $\frac{\alpha}{\alpha+1}o$ variables is roughly the same. Figuratively speaking, Gröbner basis algorithms are not able to see peculiarities that are hidden in the linear part of (12). To summarize the above discussion we state that the security reduction of Theorem 1 is based on the following empirical assumption.

Assumption: *Solving a random quadratic system with m equations and n variables is as hard as solving a quadratic system with m equations and n variables with a random quadratic part.*

In [14] J.-C. Faugère and L. Perret discuss security of the UOV scheme under direct attacks. In particular, they apply their implementation of the F5 algorithm to solve UOV systems with $o = 16$ and $v = 16, 32$ over $GF(2^4)$. Based on their experimental data, they conclude in particular: "These experiments suggest that the systems obtained when mounting a specify[2]+solve signature forgery attack against UOV behave like semi-regular systems". It is now clear why they came to such a conclusion: they intrinsically face solving a "random" system (12). In fact as we have seen in the proof of Theorem 1 a system the attacker faces is a $o \times o$ system (11), which is (much) harder than (12). We made a reduction to (12) to use a "random" quadratic system as a "provably secure" anchor. Also the following known observation is noticeable. We see from Theorem 1 that the larger α is, the more variables we expect in an "underlying" random system (and the number of equations stays the same). Therefore, as α (and thus v) increases, we expect this system to be "more random" and thus harder to solve. This is confirmed in [2], where the authors say "From experiments, we could conclude that the time complexity increases exponentially with increasing v. This fact can be understood intuitively by the observation that for increasing v, the scheme becomes more random, which makes it more difficult to solve.". Theorem 1 provides a theoretical explanation of this intuition. Note that although we are mentioning some previously known work here, clearly the key generation construction from Section 4 was not used there. Still, due to Lemma 1 the usual key generation and the modified one yield essentially the same result. Therefore, the reduction results naturally explain the older experimental results.

As a result, we have a theoretical argument for security of the UOV schemes under direct attacks. Note that such security is not that common for MPKCs. Many proposals may actually be broken already by direct attacks: Matsumoto-Imai, some instances of HFE [13], MQQ scheme [18], and many others. Some others although not broken by direct methods, show their distinction from random systems. In particular, solving succeeds at degrees lower than one would expect from a random system.

[2] We called it "fix".

5 Expected lower bounds on direct attacks complexity

In this section we present lower bounds of attacker's complexity, when using direct solving methods. These lower bounds are based on average hardness the attacker has to face as is described in Theorem 1 and Corollary 1. In order to provide concrete formulas for lower bounds we use complexity estimates for the F5 algorithm that exist for semi-regular sequences [3]. These complexity estimates are also in accordance with the assumption we made in the previous section. Similar complexity estimates exist for the XL algorithm, see [27].

In order to formalize the notion of a random system, the notion of a (semi-)regular system was introduced in [3]. The definition of a semi-regular system is as follows.

Definition 1 ([3]). *Let $f_1, \ldots, f_m \in K[X]$ be a sequence of homogeneous polynomials. This sequence of polynomials is* semi-regular *if*

- $\langle f_1, \ldots, f_m \rangle \neq K[X]$,
- *for all* $1 \leq i \leq m$ *and* $g \in K[X] : \deg(g \cdot f_i) < d_{reg}$ *and* $g \cdot f_i \in \langle f_1, \ldots, f_{i-1} \rangle \Rightarrow g \in \langle f_1, \ldots, f_{i-1} \rangle$.

Here d_{reg} is the degree of regularity *defined in [3] and it determines the degree at which a Gröbner basis algorithm like F5 terminates.*

The definition above is for homogeneous systems only. A sequence $f_1, \ldots, f_m$ of affine polynomials is called *semi-regular* if the sequence $f_1^h, \ldots, f_m^h$ is semi-regular, where f_i^h is the homogeneous part of f_i of the highest degree.

Further, there is a result saying what is the asymptotic complexity of solving a semi-regular system.

Proposition 3 ([3]). *Let $f_1, \ldots, f_m$ be an affine semi-regular sequence. Then the total number of arithmetic operations in K performed by the F5 algorithm is bounded by*

$$\mathcal{O}\Big(m \cdot d_{reg}^h \binom{n + d_{reg}^h - 1}{d_{reg}^h}\Big)^{\omega}.$$

Here d_{reg}^h is the degree of regularity of the corresponding homogeneous semi-regular sequence, and ω is the exponent of linear algebra elimination procedure, $2 < \omega \leq 3$.

For our results we also need explicit formulas for the degree of regularity that plays a role in the proposition above. Namely, the following result from [3] is of interest.

Theorem 2 ([3]). *With the notation as above, let $m = kn$, k is a constant $k > 1$. Then the degree of regularity of a homogeneous quadratic semi-regular sequence in m polynomials and n variables behaves asymptotically like:*

$$d_{reg} = (k - \frac{1}{2} - \sqrt{k(k-1)})n + \mathcal{O}(n^{1/3}), n \to \infty.$$

Note that the above results are asymptotic. Therefore, one has to be careful when applying these to concrete instances. Now having all this machinery we may state the main result of this section.

Theorem 3. *We use the same notation as in previous sections. Let $v = \alpha o, \alpha \geq 1$ and o is large enough. Let (10) be an $o \times (o + v)$ system of public equations for a UOV scheme obtained with the procedure of Section 3. Suppose that the system (10) is solved by first fixing v variables and then solving the $o \times o$ system with the F5 algorithm, which is obtained after plugging in the fixed values in (10). The complexity of this approach is lower bounded by*

$$\mathcal{O}\Big(o^2 \cdot DR(\alpha) \cdot \binom{\frac{\alpha}{\alpha+1} o + DR(\alpha) o - 1}{DR(\alpha) o}^{\omega}\Big), \tag{13}$$

where

$$DR(\alpha) = \Big(1 - \frac{\alpha}{2(\alpha+1)} - \frac{1}{\sqrt{\alpha+1}}\Big).$$

Proof. Due to Theorem 1 complexity of solving (10) is bounded from below by complexity of solving a $o \times \frac{\alpha}{\alpha+1} o$ affine quadratic semi-regular system. Now the result is obtained by setting $n = \frac{\alpha}{\alpha+1} o, m = o, k = \frac{\alpha+1}{\alpha}$ in Theorem 2.

The above lower bound is dominated by the binomial coefficient. The following result gives a simplified lower bound on the logarithm of complexity necessary for the direct attack

Proposition 4. *Using notation as above, if we denote by Compl the lower bound on complexity as in Theorem 3, then for large enough o we have*

$$\log Compl \geq \omega \cdot \frac{3\alpha + 2 - 2\sqrt{\alpha+1}}{2(\alpha+1)} \cdot H\Big(\frac{\alpha + 2 - 2\sqrt{\alpha+1}}{3\alpha + 2 - 2\sqrt{\alpha+1}}\Big) \cdot o,$$

where log *is the binary logarithm and* $H(x) = -x \log x - (1-x)\log(1-x)$ *is the binary entropy function.*

Proof. Let us rewrite the binomial coefficient in (13) in terms of o and α:

$$\binom{\frac{\alpha}{\alpha+1} o + DR(\alpha) o - 1}{DR(\alpha) o}^{\omega} = \binom{\frac{3\alpha+2-2\sqrt{\alpha+1}}{2(\alpha+1)} o}{\frac{\alpha+2-2\sqrt{\alpha+1}}{2(\alpha+1)} o}^{\omega}.$$

Using Stirling's approximation for large n and $0 < \lambda < 1$: $\log \binom{n}{\lambda n} \approx nH(\lambda)$ and the fact that (13) is dominated by the binomial coefficient, we have the result by a direct computation.

The following corollary shows the above results specified for the case $\alpha = 2$.

Corollary 2. *Using the notation above and assuming $\alpha = 2$ we have that complexity of direct F5-approach is lower bounded by*

$$\mathcal{O}\Big(o^2 \cdot \frac{2-\sqrt{3}}{3} \cdot \binom{\frac{4-\sqrt{3}}{3} o}{\frac{2-\sqrt{3}}{3} o}^{\omega}\Big),$$

The lower bound on the logarithm of complexity is

$$\log Comp \geq \omega \cdot \frac{4-\sqrt{3}}{3} \cdot H\Big(\frac{2-\sqrt{3}}{4-\sqrt{3}}\Big) \cdot o \approx 0.4\ \omega \cdot o. \tag{14}$$

In practice the lower bound (14) is pretty bad: by setting $\omega = 3$ one needs to have o around 70 to guarantee the security level of 80 bits. So for practical security tighter bounds are needed.

Remark 3. The main reason why our lower bound is so bad is the use of the oracle in the proof of Theorem 1. This oracle gives an attacker on average the values of $\frac{o}{\alpha+1}$ variables and therefore makes the system much easier to solve.

Note that we assumed that the attacker in the direct attack proceeds by first fixing v variables to concrete values and then solving an $o \times o$ system. Here we implicitly assumed that the attacker solves this $o \times o$ system "directly". There are other possibilities. For example, it has been shown to be a good practice (especially when the underlying coefficient field is not too large) first to guess at a couple of variables and then proceed with solving, e.g. with a Gröbner basis algorithm. Recent results, also in context of the UOV, on a "hybrid" approach [4] indicate that one may actually improve a bit on the complexity estimates above[3]. Still, we believe that the complexity estimates above grasp the essence of the problem, namely that on average one deals with an exponential-time algorithm. Therefore, we do not use the improved strategies here to derive more accurate lower bounds. Potentially, one may even try to proceed without the initial fixing of v variables. This may be possible if one uses e.g. a SAT-solver approach, see e.g. [1]. By this approach we do not

[3] In the sense that the attack is more efficient

need to cut down our variety to make things work; a SAT-algorithm is able to find one solution of a system directly. SAT-solvers may be quite efficient for sparse systems over $GF(2)$. Note that here we are dealing with larger fields, rather than $GF(2)$, and there methods of SAT-solving are not so well understood. Moreover, complexity of such algorithms is hard to estimate due to rich heuristics employed there. Therefore, we do not attempt to include analysis based on SAT-solver in this paper.

6 Conclusion and future work

In this paper we presented a theoretical reasoning on why breaking UOV systems directly is on average at least as hard as solving quadratic systems with a random quadratic part. This reasoning is based on the assumption that the complexity of solving an affine system is determined by its homogeneous part of the highest degree, which we believe to be a very plausible assumption. It would be interesting to test this assumption further, e.g. by using the mutant concept, [19].

As an immediate future work we see investigating the question whether similar results may be obtained for other trapdoors, e.g. Rainbow [9] and enSTS [24]. A far more reaching question for the UOV systems would be to see under which assumptions (if any) finding a decomposition of the form $\mathcal{P} = \mathcal{Q} \circ \mathcal{T}$ can be reduced to some problem that is believed to be hard. Existence of an efficient decomposition finding for the balanced variant makes funding such a reduction a very challenging task.

Acknowledgements

The first two authors would like to thank Enrico Thomae and Christopher Wolf for fruitful discussions and helpful comments.

References

1. Bard, A.: Algebraic Cryptanalysis, Springer, 2009.
2. Braeken, A., Wolf, C., Preneel, B.: A Study of the Security of Unbalanced Oil and Vinegar Signature Schemes, Topics in Cryptology CT-RSA 2005.
3. Bardet, M., Faugère, J.-C., Salvy, B., Yang, B.-Y.: Asymptotic Behaviour of the Degree of Regularity of Semi-Regular Polynomial Systems. In Proceedings of MEGA 2005, Eighth International Symposium on Effective Methods in Algebraic Geometry,2005.
4. Bettale, L., Faugère, J.-C., Perret, L.: Hybrid approach for solving multivariate systems over finite fields. Journal of Math. Cryptology, pp. 177–197, 2009.
5. Cao, W., Niw, X., Hu, L., Tang, X., Ding, J.: Cryptanalysis of Two Quartic Encryption Schemes and One Improved MFE Scheme, In PQCrypto 2010, LNCS 6061, pp. 41–60, 2010.
6. Courtois, N.: Nicolas Courtois, Short Signatures, Provable Security, Generic Attacks and Computational Security of Multivariate Polynomial Schemes such as HFE, Quartz and Sflash, available at eprint 2004/143, 2004.
7. Ding, J., Gower, J. E., Schmidt, D.: Multivariate Public Key Cryptosystems. Springer, 2006.
8. Ding, J., Hu, L., Nie, X., Li, J., Wagner, J.: High Order Linearization Equation (HOLE) Attack on Multivariate Public Key Cryptosystems. In PKC 2007, LNCS 4450, pp. 233-248, 2007.
9. Ding J., Schmidt D.: Rainbow, a new multivariate polynomial signature scheme. In Ioannidis, J., Keromytis, A.D., Yung, M. (eds.) ACNS 2005. LNCS vol. 3531, pp. 164–175 Springer, Heidelberg (2005)
10. Ding, J., Yang, B.-Y.: Multivariate Public Key Cryptography. In Bernstein, D.J., Buchmann, J., Dahmen, E. (Eds.): Post-Quantum Cryptography, Springer, 2009.
11. Ding, J., Yang, B.-Y., Chen, C.-H. O., Chen, M.-S., and Cheng, C.M.: New Differential-Algebraic Attacks and Reparametrization of Rainbow. In: LNCS 5037, pp.242–257, Springer, Heidelberg (2005)
12. Faugère, J.-C.: A new efficient algorithm for computing Grbner bases without reduction to zero (F5). Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation ISSAC (T. Mora, ed.), pp. 75–83. ACM Press, July 2002.

13. Faugère, J.-C., Joux, A.: Algebraic Cryptanalysis of Hidden Field Equations (HFE) using Gröbner Bases. In CRYPTO 2003, LNCS 2729, pp. 44-60, 2003.
14. Faugère, J.-C., Perret, L.: On the Security of UOV. In Proceedings of SCC 2008, pp. 103–109, 2008.
15. Kipnis, A., Patarin, L., Goubin, L.: Unbalanced Oil and Vinegar Schemes. In: Stern, J. (ed.) EUROCRYPT 1999. LNCS vol. 1592, pp. 206–222 Springer, Heidelberg (1999)
16. Kipnis, A., Shamir, A.: Cryptanalysis of the Oil and Vinegar Signature scheme. In: Krawzyck, H. (ed.) CRYPTO 1998, LNCS vol. 1462, pp. 257–266 Springer, Heidelberg (1998)
17. Matsumoto, T., Imai, H.: Public quadratic polynomial-tuples for efficient signature verification and message-encryption. In EUROCRYPT 1988, LNCS 330, pp. 419–545, 1988.
18. Mohamed, S.E.M., Ding J., Buchmann J.,Werner F.: Algebraic Attack on the MQQ Public Key Cryptosystem, CANS 2009, LNCS 5888, pp. 392–401, 2009.
19. Mohamed, S.E.M., Cabarcas D., Ding J., Buchmann J., Bulygin S.: MXL3: An efficient algorithm for computing Gröbner bases of zero-dimensional ideals. In proceedings of ICISC 2009, 2010.
20. Patarin, J,: Cryptanalysis of the Matsumoto and Imai public key scheme of Eurocrypt'88. In CRYPTO 1995, LNCS 963, pp.248-261, 1995.
21. Patarin, J,: Hidden Field Equations (HFE) and Isomorphism of Polynomials (IP): two new families of asymmetric algorithms. In EUROCRYPT 1996, LNCS 1070, pp. 33-48, 1996.
22. Patarin, J,: The oil and vinegar signature scheme, presented at the Dagstuhl Workshop on Cryptography (September 97)
23. Petzoldt, A., Bulygin S., Buchmann J.: A Multivariate Signature Scheme with a Partially Cyclic Public Key. In Proceedings of SCC 2010, pp. 229–235, 2010.
24. Tsujii, S., Gotaishi, M., Tadaki, K., Fujita, R.: Proposal of a Signature Scheme based on STS Trapdoor. In Proceedings of PQCrypto'2010, LNCS 6061, 2010.
25. Wang, X., Feng, F., Wang, X., Wang, Q.: A More Secure MFE Multivariate Public Key Encryption Scheme. International Journal of Computer Science and Applications 6(3), pp. 1–9, 2009.
26. Wang, L., Yang, B., Hu, Y., Lai, F.: A Medium-Field Multivariate Public Key Encryption Scheme. In CT-RSA 2006, LNCS 3860, pp. 132-149, 2006.
27. Yang, B.-Y., Chen J.-M.: All in the XL family: Theory and practice. In LNCS 3506 pp. 67-86. Springer, Heidelberg (2007)