Key-Recovery Attack on the ASASA Cryptosystem with Expanding S-Boxes

Henri Gilbert(✉), Jérôme Plût, and Joana Treger

ANSSI, Paris, France
henri.gilbert@ssi.gouv.fr

Abstract. We present a cryptanalysis of the ASASA public key cipher introduced at Asiacrypt 2014 [3]. This scheme alternates three layers of affine transformations A with two layers of quadratic substitutions S. We show that the partial derivatives of the public key polynomials contain information about the intermediate layer. This enables us to present a very simple distinguisher between an ASASA public key and random polynomials. We then expand upon the ideas of the distinguisher to achieve a full secret key recovery. This method uses only linear algebra and has a complexity dominated by the cost of computing the kernels of 2^{26} small matrices with entries in $\mathbb{F}_{16}$.

Introduction

A long-standing challenge in asymmetric cryptography is to bring asymmetric cryptography closer to symmetric cryptography by designing public key schemes whose overall structure and elementary operations are similar to those used in mainstream block ciphers such as AES. Solving this appealing but difficult challenge would not only increase the diversity in asymmetric cryptography, but might also help reducing the considerable performance gap between asymmetric cryptography and symmetric cryptography (the latter currently being more efficient by several orders of magnitude). This might as well allow the emergence of symmetric algorithms with some extra features, as for instance symmetric encryption schemes with a secure white-box implementation. Until 2014 however, as far as we know, all attempts of public key scheme designs with block cipher features, *e.g.* [10,15,16], eventually turned out to be weak [1,7,8,13,18].

Asymmetric ASASA Schemes. Some new candidate solutions to the above challenge were proposed in a paper published at Asiacrypt 2014 by Biryukov et al. [3]. One conducting idea for the new designs stems from the observations that: (1) Traditional SPN block ciphers such as AES can be viewed as an alternance of (at least partly secret) affine transformations A and S-box layers S, and generally comprise a substantial number of A rounds (essentially 10 in the

This work was partially supported by the French National Research Agency through the BRUTUS project (contract ANR-14-CE28-0015).

R. Gennaro and M. Robshaw (Eds.): CRYPTO 2015, Part I, LNCS 9215, pp. 475–490, 2015.
DOI: 10.1007/978-3-662-47989-6_23

case of AES-128); as shown by Biryukov and Shamir [4], some efficient generic attacks exist for the ASASA structure with secret S and A layers and small bijective S-boxes. (2) The efforts to design public key schemes with an alternance of A and S layers mainly focused so far on multivariate schemes with an ASA structure, with one single large S-box described by low degree equations over a finite field.[1] Based on the former considerations, the authors of [3] advocate for use of public multivariate schemes with an ASASA structure, *i.e.* with the simplest possible structure for which no generic attack is known in the case of small bijective S-boxes—or more generally of injective S-boxes whose non-zero linear combinations of outputs are not too strongly biased.

More precisely, the authors of [3] proposed a public-key encryption scheme named the *asymmetric ASASA scheme with expanding S-boxes*, conjecturing that it offers a comfortable security margin with respect to the potential lines of attack identified in their security analysis. This scheme uses small input-expanding injective quadratic S-boxes. Since the these S-boxes have a length expansion factor of 2, the whole scheme has a length expansion factor of 4. The standard plaintext and ciphertext length for this scheme are respectively 128 and 512 bits.

While this works focuses exclusively on the ASASA scheme with expanding S-boxes, the same authors also proposed in [3] a second public-key scheme based on the ASASA structure, named the χ*-scheme*. Indeed, this alternative construction makes use of Daemen's bijective quadratic S-box χ based on cellular automata [5] and also used in various recent hash functions. The standard plaintext and ciphertext length is 128 bits. In this χ-scheme, one single large S-box is used at each S layer. In their security analysis though, the authors of [3] consider many attacks on weakened versions of the χ-scheme and conclude that the security margin of the χ-scheme must be lower than that of the expanding scheme. They therefore express some caveats on its security and only "offer it as a cryptanalytic challenge, but not for practical use", unlike the expanding scheme.

In both ASASA public key encryption schemes, quadratic S-boxes are being used. The public key consists of the quartic equations of the encryption function and the private key consists of the specification of the A and S layers and of some *perturbation polynomials*, which are added to a few components of the vector representing the output of the second S layer.[2] Another property of these public key encryption schemes is that they can also be viewed as symmetric ciphers with a decent encryption and decryption speed and the following extra feature: as an alternative to using the secret key to efficiently encrypt, the public key provides

[1] While a noticeable exception is the multivariate scheme $R2$ [14] that contains two S layers with small S-boxes, one weakness highlighted by attacks on $R2$ or its variant $R2^-$ that were eventually discovered is the fact that the $R2$ S-boxes are not injective.

[2] While the role of perturbations is essential in the case of the χ-scheme since its variants without perturbations are reported in [3] to be vulnerable to efficient Gröbner basis attacks, in the case of the expanding scheme, perturbations are mostly introduced to provide some extra resistance against decomposition attacks that could potentially reduce the ASASA structure to the functional composition of two ASA structures.

a slower *strong white-box* implementation of the encryption function — i.e. an obfuscated implementation that is conjectured to prevent that the decryption function be derivable by an adversary who has full access to it.

Our Contribution. In this paper, we present an efficient attack on the ASASA scheme with expanding S-boxes. The starting point for our attack is the analysis of the homogeneous cubic part of the derivatives of the (quartic) polynomials of the public key. We first show that this analysis provides a distinguisher that allows to tell apart the public key of the scheme from a vector of random quartic polynomials. We then describe how to leverage the first information about the secret key provided by the distinguisher to retrieve the intermediate values that lie between the two S-layers, for an equivalent representation of the scheme. At this point, we are essentially left with the problem of solving two quadratic ASA layers. Though generic techniques to solve this problem exist, we give our own algorithm, that is well-adapted to the scheme considered. The overall complexity of the attack is equivalent to at most 2^{26} computations of kernels of matrices of size 64×96 over the finite field $\mathbb{F}_{16}$. We estimate the corresponding computational time to a few CPU-hours, which places this cryptanalysis well within practical limits.

This paper is organised as follows. Section 1 provides a description of the expanding ASASA scheme and presents some useful preliminary results. Section 2 introduces a distinguisher for this scheme that can be used to derive some first information on the secret key. Finally, Sect. 3 shows how to efficiently derive an equivalent secret key from the public key.

1 The ASASA Cryptosystem

1.1 Definition and First Notations

The two asymmetric ASASA schemes of [3] are composed of polynomial transformations over the base field $k = \mathbb{F}_{16}$; they are obtained by alternating three k-affine layers and two non-linear polynomial-based S layers. The ASASA scheme with expanding S-boxes, on which we are focusing, involves S-boxes whose output is twice as big as their input; 32 perturbation polynomials of degree four are also applied just before the last affine transform. More precisely, each S-box maps a 4-tuple of k-values onto an 8-tuple of k-values, defined as degree 2 polynomials over k in the inputs. The resulting scheme, which we simply call the ASASA cryptosystem in the remaining of this paper for simplification purposes, has then degree 4 over k. Going into details, the private key of the ASASA cryptosystem consists of:

- Three uniformly random invertible affine transformations $\mathcal{A}^x$ of k^{32}, $\mathcal{A}^y$ of k^{64}, and $\mathcal{A}^z$ of k^{128};
- Two sets of uniformly random quadratic functions from k^4 to k^8 corresponding to the first and second S-box layer $S^x = \{S^x_{0,0}, \ldots, S^x_{0,7}, \ldots S^x_{7,0}, \ldots S^x_{7,7}\}$, that determines 8 S-boxes $S^x_0, \ldots S^x_7$, and $S^y = \{S^y_{0,0}, \ldots, S^y_{7,0}, \ldots S^y_{15,0}, \ldots, S^y_{15,7}\}$, that determines 16 S-boxes $S^y_0, \ldots S^y_{15}$;

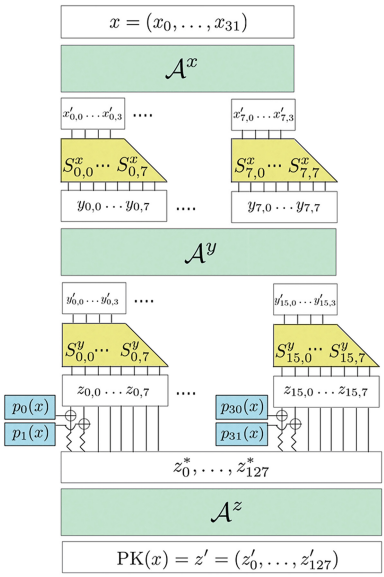

Fig. 1. The ASASA cryptosystem with expanding S-boxes.

– Thirty-two uniformly random quartic perturbation polynomials $p_0, \ldots, p_{31}$ on k^{32}.

The public key and the associated public encryption function are derived from the secret key as illustrated at Fig. 1, following those successive steps:

(i) The plaintext state is the tuple of variables $x = (x_0, \ldots, x_{31}) \in k^{32}$;
(ii) Let $x' = \mathcal{A}^x \cdot x \in k^{32}$;
(iii) Define $y = (y_{0,0}, \ldots, y_{7,7}) \in k^{64}$ as $y_{r,i} = S^x_{r,i}(x'_{r,0}, \ldots, x'_{r,3})$, for $r, i = 0 \ldots 7$;
(iv) Let $y' = \mathcal{A}^y \cdot y \in k^{64}$;
(v) Define $z = (z_{0,0}, \ldots, z_{15,7}) \in k^{128}$ as $z_{r,i} = S^y_{r,i}(y'_{r,0}, \ldots, y'_{r,3})$, for $i = 0 \ldots 7$ and $r = 0 \ldots 15$;
(vi) For $r \in \{0, 15\}$, do $z^*_{8r} \leftarrow z_{r,0} + \mathrm{p}_{2r}(x)$, $z^*_{8r+1} \leftarrow z_{r,1} + \mathrm{p}_{2r+1}(x)$, and $z^*_{8r+i} \leftarrow z_{r,i}$ for $i = 2 \ldots 7$ (the two first components out of 8 contiguous components of z are modified);
(vii) The public key $\mathrm{PK}(x)$ is the vector of 128 polynomials over k, $z' = \mathcal{A}^z \cdot z^* \in (k[x_0, \ldots x_{31}])^{128}$; the public encryption function is the associated function from k^{32} to k^{128}.

Dimension of the Secret and Public Key Spaces. Since the dimensions of various vector spaces are central in our analysis, we compute the size of the secret and public key spaces. We do not find the exact same numbers as the original authors [3, 2.5], although the order of magnitude is the same. An

affine transform on n variables is representable by a matrix of size $n \times (n+1)$; therefore, the three A layers have a key size of $32 \times 33 + 64 \times 65 + 128 \times 129 = 21\,728$ elements of $\mathbb{F}_{16}$. An S-box output is an (inhomogeneous) quadratic polynomial in four variables, and is therefore described by $\binom{6}{2} = 15$ coefficients. (The dimensions of spaces of homogeneous polynomials are as given below in Sect. 1.3 of this paper; inhomogeneous polynomials in n variables correspond bijectively to homogeneous polynomials of the same degree in $n+1$ variables). Therefore, the two S layers have a key size of $24 \times 8 \times 15 = 2\,880$ elements of $\mathbb{F}_{16}$. In total, the secret key size is 24 608 elements of $\mathbb{F}_{16}$, or approximately $2^{13.6}$ bytes of data. This does not, however, count the perturbation polynomials, which occupy a space of $32 \times \binom{36}{4}$ elements, or $2^{19.8}$ bytes of data. The public key is a set of $128 \times \binom{36}{4}$ elements of $\mathbb{F}_{16}$, or $2^{21.8}$ bytes of data.

1.2 Equivalent Simple Keys

As for most multivariate cryptosystems, there are multiple private keys that correspond to a given ASASA public key, and we show here that each secret key is equivalent to a simpler one, that we describe. We also redefine the ASASA system in terms of those "simple" keys, which make our attack easier to explain; we point out that this simplification is purely cosmetic though and that our attack does apply to the ASASA system as described in [3].

First, since for each $r = 0, \ldots, 15$, the two outputs $z_{r,0}$, $z_{r,1}$ of the S-box S^y_r are added to arbitrary perturbation polynomials p_i, we obtain the same public key when replacing p_{2r}, p_{2r+1} by $p_{2r} + z_{r,0}$, $p_{2r+1} + z_{r,1}$ and $S^y_{r,0}$, $S^y_{r,1}$ by zero.

Let $\mathcal{A}^x = A^x + a^x$ and $\mathcal{A}^y = A^y + a^y$ be the decomposition of $\mathcal{A}^x$ and $\mathcal{A}^y$ as their linear part plus their constant. We can actually assume that $\mathcal{A}^x = A^x$ and $\mathcal{A}^y = A^y$, as it is always possible to consider a modified S^y S-box layer where the first addition by a^x and the second addition by a^y are absorbed by the polynomials of S^y. The same goes for $\mathcal{A}^z$, where the addition of the 32 components with index $8r$ and $8r+1$ of a^z can be viewed as part of the perturbation polynomials, so that $S^y_{r,0}$ and $S^y_{r,1}$ are still zero. This shows that we can assume $a^x = a^y = a^z = 0$ and consider from now on A^x, A^y and A^z instead of $\mathcal{A}^x$, $\mathcal{A}^y$ and $\mathcal{A}^z$.

Finally, notice that it is always possible to compose the output of the quadratic map S^x_i by a linear transform, and the corresponding input block of the linear map A^y by its inverse. The same applies to S^y_i and A^z, for $i \neq 8r, 8r+1$, *i.e.* as long as the corresponding 32 zero polynomials of S^y are not affected.

To sum up, the description of the ASASA private key of Sect. 1.1 is equivalent to the following:

- Three uniformly random invertible linear transformations A^x of k^{32}, A^y of k^{64} and A^z of k^{128};
- Two sets of 8×8 and 16×6 uniformly random quadratic functions on k^4, $S^x = \{S^x_{0,0}, \ldots, S^x_{7,7}\}$ and $S^y = \{S^y_{0,0}, \ldots, S^y_{15,5}\}$;
- A set of 32 uniformly random quartic polynomials $p_0, \ldots, p_{31}$ on k^{32}.

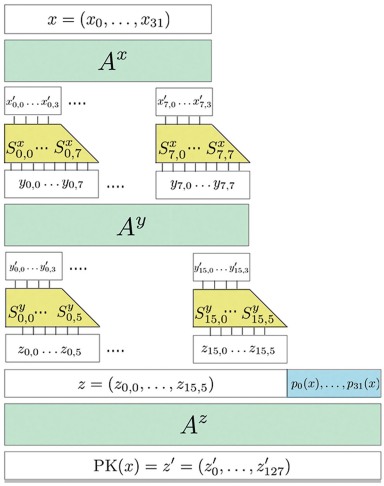

Fig. 2. The ASASA cryptosystem with expanding S-boxes, equivalent representation.

In the remaining of this paper, we shall always consider such *simple private keys*; the corresponding encryption mechanism is illustrated at Fig. 2.

1.3 Notations and Preliminaries

We write $k = \mathbb{F}_{16}$ for the finite field with 16 elements. Throughout this work, we let $q = |k| = 16$.

Homogeneous Polynomials. We write $H_{d,n}$ for the space of homogeneous polynomials of degree d over n variables (over the base field $k = \mathbb{F}_{16}$); it is a vector space of dimension $\binom{n+d-1}{d}$ [17, 2.2.1]. Throughout this paper, we shall usually work with the vector space $H_{d,32}$ of homogeneous polynomials of degree d on the 32 input variables x_i. We shall write H_d instead of $H_{d,32}$.

Let $f(x_1, \ldots, x_n)$ be a polynomial. For any integer d, we write $f_{(d)}$ for the degree d homogeneous part of f.

Vector Spaces. Given subspaces $E \subset H_{d,n}$ and $E' \subset H_{d',n}$, we define $E \cdot E'$ as the subspace of $H_{d+d',n}$ generated by all the products $u \cdot u'$ for $u \in E$ and $u' \in E'$. We also define D as the vector space generated by the 32 derivations $\partial_i = \partial/\partial x_i$. For any vector space $E \subset H_{d,n}$, we define $DE \subset H_{d-1,n}$ as the vector space generated by all $\delta(u)$ for $\delta \in D$ and $u \in E$.

Counting Matrices of a Given Rank. We introduce the following notations, used in the computation of the number of matrices with given size and rank. For any integers $n \geq d$, we define

$$[n,d] = \prod_{i=0}^{d-1} q^n - q^i \qquad \text{and} \qquad \begin{bmatrix} n \\ d \end{bmatrix} = \frac{[n,d]}{[d,d]}. \tag{1}$$

We omit the value q from these notations, since we shall always use $q = 16$. We use the following classical result (see [12, VII.19]).

Proposition 1. *Let m, n, d be three integers.*

(i) There are exactly $[n,d]$ injective maps from k^d to k^n.

(ii) There are exactly $\begin{bmatrix} n \\ d \end{bmatrix}$ subspaces of dimension d of k^n.

(iii) There are exactly $\frac{[n,d]\cdot[m,d]}{[d,d]}$ matrices of size $m \times n$ with rank exactly d. □

Proposition 2. *Let V be a vector space of dimension d over $k = \mathbb{F}_{16}$. A set of n uniformly random, independent vectors $v_i \in V$ generates V with overwhelming probability if $n \geq d + 32$.*

Proof. Let $\pi(n,d)$ be the probability that a random matrix of size $n \times d$ has maximal rank d. We see by Proposition 1 that

$$\pi(n,d) = \frac{[n,d]}{q^{nd}} = \prod_{i=0}^{d-1} 1 - q^{-(n-i)}. \tag{2}$$

For $n > d \gg 0$, since $q^{-n} \ll 1$, we have the asymptotic expansion

$$\pi(n,d) = \exp\left(\sum_{i=0}^{d-1} \log(1 - q^{-(n-i)})\right) \simeq \exp\left(-q^{-n}\frac{q^d - 1}{q-1}\right). \tag{3}$$

For any $\varepsilon > 0$, we find that

$$\pi(n,d) > 1 - \varepsilon \quad \text{iff} \quad n > \log_q \frac{q^d - 1}{q-1} - \log_q \log \frac{1}{1-\varepsilon}. \tag{4}$$

Using $\varepsilon = 2^{-128}$ as our definition of "overwhelming probability", the above condition (4) becomes $n \geq d + 32$.

Expected Behaviour of Derivatives and Product of Random Polynomials. We shall use throughout this work the following heuristics about the derivatives and products of random polynomials.

For $f_1, \ldots, f_r$ uniformly random, independent elements of $H_{d,n}$:

(i) if $r < \frac{1}{n} \dim H_{d-1,n}$, then the nr derivatives $\partial_i f_j$ behave like uniformly random, independent elements of $H_{d-1,n}$;

(ii) if $r < \frac{1}{n} \dim H_{d+1,n}$, then the nr products $x_i f_j$ behave like uniformly random, independent elements of $H_{d+1,n}$.

In particular, according to Proposition 2, we expect that, with overhelming probability, if $nr \leq \dim H_{d-1,n} - 32$, then the $\partial_i f_j$ are free in $H_{d-1,n}$; if $nr \leq \dim H_{d+1,n} - 32$, then the $x_i f_j$ are free in $H_{d+1,n}$. Although giving a detailed proof of these facts would be out of the scope of this work, we obtained an empiric confirmation in the cases of interest to us (namely $n = 32$, and either $d = 4, r = 128$ for derivation, or $d = 2, r = 64$ for multiplication), as well as of the bounds of validity.

2 A Simple Distinguisher

We shall see that the ASASA cryptosystem presents the same flaw as several multivariate cryptosystems, that is, it is possible to distinguish the equations of the public key for the ASASA scheme from random polynomials of the same degree over k. The distinguisher we present here is extremely simple: namely, computing the rank of the matrix of partial derivatives of the polynomials is enough. However, by elaborating on the structure of this distinguisher, we shall explain in Sect. 3 how it is possible to fully recover the secret key.

2.1 Considerations on the Dimension of Vector Spaces Derived from the Public Key

The key observation underlying the distinguisher is that, while the space of homogeneous cubic polynomials $H_3 = H_{3,32}$ has dimension $\binom{34}{3} = 5\,984$, the homogeneous cubic parts of the derivatives of the public key $(\partial \mathrm{PK}_i/\partial x_j)_{(3)}$ actually belong to a much smaller subspace of H_3 which happens to have dimension at most 3 072.

The Case of the ASASA Cryptosystem with No Perturbation Polynomials. As a warm-up, we first consider the encryption of the message $(x_0, \ldots, x_{31})$ under the "non-perturbed" ASASA scheme, *i.e.* the ASASA scheme with no perturbation polynomials. We recall that all intermediate values $\{y_{r,i}\}$, $\{y'_{r,i}\}$, $\{z_{r,i}\}$ and $\{z'_i\}$ introduced at Sect. 1 can be seen as polynomials in the 32 input variables x_i with coefficients in $k = \mathbb{F}_{16}$ (see Fig. 1).

To see that the $(\partial \mathrm{PK}_i/\partial x_j)_{(3)}$, for $i = 0 \ldots 127$, $j = 0 \ldots 31$, belong to a restricted subspace, we consider the second S-box layer S^y; recall that $z_{r,i}$ denotes the i-th output of the S-box S^y_r with input $y'_{r,0}, \ldots y'_{r,3}$. The quadratic polynomials $y'_{r,i}$ may be written $y'_{r,i} = (y'_{r,i})_{(2)} + (y'_{r,i})_{(1)} + (y'_{r,i})_{(0)}$, as the sum of a homogeneous degree 2 part, a linear part and a constant. We therefore see that the homogeneous parts of degree 4 of the polynomials $z_{r,i}$ output by the S-box S^y_r are linear combinations of the terms $(y'_{r,m})_{(2)} \cdot (y'_{r,n})_{(2)}$. Let us write $\partial z_{r,i}/\partial x_j$ for the derivative of the output $z_{r,i}$ along the input variable x_j. There exists coefficients $a_{m,n} \in k$ such that

$$(\partial z_{r,i}/\partial x_j)_{(3)} = \sum_{m,n} a_{m,n} (y'_{r,m})_{(2)} (\partial y'_{r,n}/\partial x_i)_{(1)}. \tag{5}$$

Let $Y'_{(2)} \subset H_{32,2}$ be the vector space spanned by the 64 homogeneous quadratic parts of the polynomials $y'_{r,m}$; it has dimension at most 64. Let also $(DZ)_{(3)}$ be the vector space spanned by the $128 \times 32 = 4\,096$ homogeneous cubic parts of the derivatives of the polynomials $z_{r,i}$. The expression (5) above implies

$$(DZ)_{(3)} \subset Y'_{(2)} \cdot H_1. \tag{6}$$

The vector space $(DZ')_{(3)} = \langle (\partial \mathrm{PK}_i / \partial x_j)_{(3)} \rangle$ being a linear image of $(DZ)_{(3)}$ by A^z, we also have

$$(DZ')_{(3)} \subset Y'_{(2)} \cdot H_1, \tag{7}$$

and

$$\dim (DZ')_{(3)} \leq \dim Y'_{(2)} \cdot H_1 \leq 64 \times 32 = 2\,048. \tag{8}$$

The General ASASA Cryptosystem. For the general ASASA scheme, we have to slightly adapt the result (7) to take into account the perturbation polynomials. We refer the reader to Fig. 2 for the description of ASASA used in this paragraph.

We established in our analysis of the unperturbed scheme that the homogeneous cubic parts of the derivatives of the polynomials $\{z_{r,i}\}$ belong to the vector space $Y'_{(2)} \cdot H_1$, which has dimension 2 048. This still holds for the general ASASA scheme, since up to generation z, the perturbation polynomials do not appear. The next step in the algorithm is the linear transform A^z; its input is the concatenation $z||(p_0, \ldots, p_{31})$, where the first 96 elements are the polynomials of z and the last 32 are the perturbation polynomials. This means that the polynomials of the public key z' are linear combinations of the polynomials of z and the perturbation polynomials $p_0, \ldots p_{15}$. Let DP be the vector space spanned by the homogeneous cubic parts $(\partial p_i / \partial x_j)_{(3)}$ of the derivatives of the perturbation polynomials; it has dimension at most $32 \times 32 = 1\,024$. We necessarily have

$$(DZ')_{(3)} \subset Y'_{(2)} \cdot H_1 \, + \, DP. \tag{9}$$

In terms of dimensions, Eq. (9) implies

$$\dim (DZ')_{(3)} \leq \dim (Y')_{(2)} \cdot H_1 + 1\,024 \leq 3\,072, \tag{10}$$

as claimed.

2.2 The Distinguisher

We now turn the observation of Sect. 2.1 into a working distinguisher.

Proposition 3. *It is possible to distinguish the public key polynomials of the* ASASA *scheme from uniformly random quartic polynomials by computing the rank of a matrix of size* $5\,984 \times 4\,096$ *with coefficients in* $\mathbb{F}_{16}$.

Proof. Let $T = (T_0, \ldots, T_{127})$ be a vector of 128 polynomials that are either uniformly random quartic polynomials, or a (perturbed) ASASA public key PK. We consider the matrix M of size $5\,984 \times 4\,096$ whose columns are the vectors $(\partial T_i/\partial x_j)_{(3)} \in H_{3,32}$, with the usual notations. If the T_i are uniformly random polynomials, then the rank of M is 32 times the rank of the family (T_i); since $128 \times 32 < \dim H_{3,32}$, according to our heuristic, this is usually $32 \times 128 = 4\,096$. If on the contrary T is an ASASA public key, then by (10), the rank of M is at most 3 072.

The distinguisher that returns `ASASA` if the rank of M is $\leq 3\,072$, and `random` otherwise, succeeds with overwhelming probability.[3] □

3 Key Recovery

We now present a secret key recovery attack on an ASASA scheme. The ideas used here are based on the properties already identified in Sect. 2, namely, the space of derivatives of PK contains information about the intermediate values between the two quadratic layers.

To attack the system, we first identify the vector space of quadratic forms manipulated in the middle of the algorithm (as output of the first S layer, and input to the second one). This is the crucial point of the cryptanalysis. It enables us to reduce the problem to two much simpler ASA problems. We then solve each ASA instance in turn. (Note that although we present a specific way to solve to these two ASA instances, it is well-known that ASA instances are weak, and techniques to solve such systems can be found in the literature [2,6,9,11]).

This key-recovery only relies on linear algebra in various spaces of homogeneous polynomials. We refer the reader to Sect. 1.3 for some useful general results in algebra used in the attack; a few other results will be introduced on the fly when needed. For simplicity in this whole Sect. 3, we write Y, instead of $Y'_{(2)}$, for the space generated by the homogeneous quadratic parts of the polynomials of y.

The overall complexity of the attack is about 2^{26} times the computation of the rank of a square matrix of size 64×96 with coefficients in $\mathbb{F}_{16}$. We point out that our method only uses the quadratic terms of the secret quadratic layers; it is therefore also applicable to homogeneous instances of the ASASA cryptosystem.

3.1 Computing the Middle Layer

As already mentioned, our attack uses the same data as the distinguisher. More precisely, the key result was given at Eq. (9): the vector space DZ' of derivatives of PK contains information about the space $Y \cdot H_1$, *i.e.* about the vector space Y of homogeneous quadratic functions produced by the first S layer. However, the

[3] We may also investigate the case of a reinforced ASASA scheme with more perturbation polynomials, *i.e.* with 96 "legitimate" outputs and $p \geq 32$ perturbations. We easily find that our distinguisher works at least for $p \leq 90$. The same bound applies to the key recovery attack of Sect. 3 below.

observed vector space deduced from the public key also contains some unwanted vectors originating in the perturbation polynomials.

To access the space $Y \cdot H_1$ and see beyond the perturbation polynomials, the first step is to construct several subspaces of DZ' including $Y \cdot H_1$. We are then able to recover $Y \cdot H_1$ as the intersection of all those subspaces. In a second step we show how, from this recovered vector space, we compute the space Y itself.

Eliminating the Perturbations. This first steps aims at computing the vector space $Y \cdot H_1$ from the public key.

Recall that a public key PK of the ASASA cryptosystem is given as a vector of 128 polynomials PK_i in the 32 input variables. We define $F_i = (\mathrm{PK}_i)_{(4)}$ as the homogeneous quartic part of the public key, and F as the vector space generated by all F_i.

For any derivation $\delta \in D$ and for any $f \in F$, we saw when describing the distinguisher that

$$\delta f \in Y \cdot H_1 + \delta P, \tag{11}$$

where Y is the vector space generated by the 64 quadratic polynomials $(y_{r,i})_{(2)}$, and P is the vector space generated by the 32 perturbation polynomials $(p_i)_{(4)}$.

Let $\Delta \subset D$ be a vector space of dimension d. By (11), we then have

$$\Delta F \subset Y \cdot H_1 + \Delta P, \tag{12}$$

where the right-hand space has dimension at most $64 \times 32 + 32 \times d = 2\,048 + 32\,d$. The space ΔF is generated by $128\,d$ elements δf_i, for $\delta \in \Delta$ and $i = 0, \ldots, 127$. By Proposition 2, these $128\,d$ elements generate the whole space $Y \cdot H_1 + \Delta P$ as long as $128\,d \geq 2\,048 + 32\,d + 32$, or equivalently, $d \geq 22$.

We now consider a family of m vector spaces $\Delta_1, \ldots, \Delta_m \subset D$, each space Δ_i being of dimension $d = 22$. We know by what precedes that for each one of them, $\Delta_i F = Y \cdot H_1 + \Delta_i P$. This implies that

$$\bigcap_{i=1}^{m} \Delta_i F \quad = \quad Y \cdot H_1 \; + \; \bigcap_{i=1}^{m} \Delta_i P. \tag{13}$$

Since $\dim D = 32$ and $\dim \Delta_i = 22$ for each i, the intersection of two spaces Δ_i generally has dimension 12 (this is always the case if we choose the Δ_i correctly), and likewise the intersection of three such spaces has dimension 2, and for $m \geq 4$, we easily find $\Delta_1, \ldots, \Delta_m$ such that $\Delta_1 \cap \ldots \cap \Delta_m = 0$. This implies that

$$\bigcap_{i=1}^{m} \Delta_i P = 0 \text{ for } m \geq 4. \tag{14}$$

Formula (14) means that the intersection $\bigcap_{i=1}^{m} \Delta_i F$ is then exactly the space $Y \cdot H_1$:

$$\bigcap_{i=1}^{m} \Delta_i F = Y \cdot H_1 \text{ for } m \geq 4. \tag{15}$$

Computing the Middle Terms. This part explains how we recover the 64-dimensional space Y from the space $Y{\cdot}H_1 \subset H_3$ obtained during the previous step.

We first prove a short lemma. Let $V \subset H_2$ be a vector space of dimension d and basis (v_j). The vector space $V{\cdot}H_1 \subset H_3$ is generated by the $32\,d$ elements $x_i\,v_j$. If $d \leq 186$, then $32\,d \leq \dim H_3 - 32$; by Proposition 2, we therefore expect these $32\,d$ elements to be linearly independent in H_3. This implies that $\dim(V{\cdot}H_1) = 32\,d$. In particular, this means that when V has dimension ≤ 186, the correspondence between $(\dim V)$ and $(\dim V{\cdot}H_1)$ behaves, with very high probability, as a strictly increasing function.

We now use this lemma to characterize the space Y. Let $\overline{Y} \subset H_2$ be the vector space of all functions g such that, for all i, $gx_i \in Y{\cdot}H_1$. Trivially, $Y \subset \overline{Y}$ and $\overline{Y}{\cdot}H_1 = Y{\cdot}H_1$. Since $\dim Y = 64 \leq 186$, we are in the conditions of the previous lemma, which implies that $Y = \overline{Y}$ with overwhelming probability.

We may easily compute the space $\overline{Y}$ from $Y{\cdot}H_1$ as follows. For each $i = 0, \ldots, 31$, we define G_i as the subspace of multiples of x_i in $Y{\cdot}H_1$; by definition, $\overline{Y}$ is the intersection of all spaces $x_i^{-1}G_i$,

$$\overline{Y} = \bigcap_{i=0}^{31} x_i^{-1} G_i. \tag{16}$$

3.2 Solving a Quadratic ASA Layer

As already mentioned, there exists generic techniques [2,6,9,11] for inverting a public key in the ASA form. We give our own solution here, as it is simple and works well in the particular case of an ASA layer based on quadratic S-boxes. We shall use this technique twice, once for the inner ASA layer, and then once more for the outer one.

This section and the next one (Sect. 3.3) are mutually independent. We present the inner layer ASA first since it is easier to explain.

Notations. The inner ASA layer is represented as the (known) vector space Y generated by the 64 (unknown) quadratic forms $y_{r,i}$ in the 32 input variables x_i. We restrict ourselves here to the homogeneous part of the $y_{r,i}$, since this case is more difficult to solve, but easier to present.

For each fixed r, the eight functions $y_{r,0}, \ldots, y_{r,7}$ are quadratic forms in the same 4 fixed linear combinations $x'_{r,0}, \ldots, x'_{r,3}$ of the input variables x_i. We write X_r for the vector space generated by the $x'_{r,i}$. We may also decompose the space of differentials $D = \langle \partial/\partial x_i \rangle$ according to the S-boxes; namely, for each r, we define D^r as the set of all elements $\delta \in D$ whose restriction to X_r is zero.

We note that $\dim X_r = 4$ and $\dim D^r = 28$. Therefore, for a given r, an uniformly random element of D belongs to D^r with probability $q^{-4} = 2^{-16}$.

Separating the Inputs of the S-Boxes. We show that we are able to identify the elements of D^r, *i.e.* the differentials that vanish on the inputs of a particular S-box.[4]

For any quadratic form $f \in Y$ and any $\delta \in D$, the function δf is a linear form of $x \in X$; this means that $(\delta f)(x)$ is a trilinear function of (δ, f, x). Therefore, for a fixed value of δ, it is a bilinear function of (f, x). We write F_δ for this bilinear form. It is represented by a matrix of size 64×32 whose coefficients are the $(\delta f_i)(x_j)$, where (f_i) is a basis of Y and (x_j) is the standard basis of X.

Let $f \in Y$; we can write f as a sum $\sum f_s$ for $s = 0, \ldots, 7$, where f_s is a quadratic form on X_s. For any $\delta \in D^r$, we have $\delta f_r = 0$, so that δf is the sum of the δf_s for $s \neq r$. Since each one of these terms uses only the variables from X_s, this means that $(\delta f)(x) = 0$ for $x \in X_r$. Put differently, the kernel of F_δ (here seen as a linear map from X to the dual of Y) contains X_r.

Now let δ be an element of D not belonging to any of the D^r: since the maps (δf_i) are 64 random linear forms on the 32-dimensional space X; by Proposition 2, with overwhelming probability, the intersection of their kernels is zero.

As a result, we see that the rank of the matrix F_δ is always ≤ 28 if δ belongs to at least one of the D^r and 32 with overwhelming probability otherwise. This also provides a test, given two elements δ and ε of D^r and D^s, for the equality $r = s$: since the kernels of the matrices F_δ, F_ε contain respectively X_r and X_s, their intersection is non-trivial when $r = s$.

The Algorithm. We compute the spaces X_r and D^r at the same time, and up to a permutation of $\{0, \ldots, 7\}$, since we do not know the labeling of the S-boxes. For each $r = 0, \ldots, 7$, we keep candidates U_r and V_r for the spaces X_r and D^r; initially, U_r is the whole space X while the space V_r is empty. During the whole algorithm, we have $U_r \supset X_r$ and $V_r \subset D^r$, with both inclusions being equalities at the end. We also note that, at every step of the alorithm, U_r is the intersection of the kernels of all elements of V_r.

We now describe the algorithm. We randomly generate elements δ of D and compute the kernel $K \subset X$ of the matrix F_δ. If this kernel has dimension at least 4, then it intersects non-trivially one of the spaces U_r and the intersection also has dimension at least 4. We then update V_r to $V_r \oplus \langle \delta \rangle$ and U_r to $U_r \cap K$. The algorithm ends when each space V_r has the required dimension 28, as then $V_r = D^r$ as required, and therefore $U_r = X_r$.

Recovering the S-Boxes. Once the spaces X_r are known, computing the S-boxes is easy. We write Y_r for the vector space generated by the 8 outputs of the S-box S_r^x, and X^r for the direct sum of all X_s for $s \neq r$. Then Y_r is exactly the

[4] Since the elements of Y are quadratic forms, their differentials are exactly the associated polar forms. This means that we may represent the derivations as elements of X, using the relation $(\delta f)(x) = f(x + \delta) - f(x) - f(\delta)$; in this view, the space D^r is the direct sum of all the X_s for $s \neq r$. However, we chose to use an explanation based on differentials, since this is both closer to the differential cryptanalysis point of view, and easier to generalize to polynomials of higher degree.

set of elements of Y that vanish on all points of X^r. We may therefore compute Y_r with linear algebra. (Another possibility is to use the derivations spaces D^r, also computed in the previous step, since Y_r is the set of elements $f \in Y$ such that, for all $\delta \in D^r$, $\delta f = 0$). Once bases of both X_r and Y_r are known, we recover the secret functions S_r^x by interpolation.

Complexity. $q^4 \times 28 \times 8 \approx 15{\cdot}10^6$ elements of X, the expected cost for the execution of this algorithm is approximately $2^{23.8}$ times that of the computation of the kernel of a matrix of size 32×64 with entries in $\mathbb{F}_{16}$.

3.3 Solving the Outer ASA Layer

We again use the representation of the middle layer as a 64-dimensional vector space Y of quadratic forms computed in Sect. 3.1. We now determine the output functions F_i as linear combinations of quadratic forms in the elements of Y and the 32 perturbation polynomials p_i.

Computing the Outputs of the S-Boxes. We recall that F is the vector space generated by the homogeneous quartic part of the ASASA public key. This vector space is the direct sum of the 32-dimensional space P generated by the homogeneous quartic parts of the perturbations, and the 96-dimensional vector space Z generated by the outputs of the 16 S-boxes S_r^y. Since the middle layer Y is known as a result of Sect. 3.1, we may compute Z as the intersection $Z = (Y{\cdot}Y) \cap F$.

Reducing to a Quadratic ASA Layer. We now study the 16 S-boxes S_r^y. We already know the 64-dimensional vector space Y of their (quadratic) inputs and the 96-dimensional vector space Z of their (quartic) outputs. Our goal for now is to rewrite each element of Z as an explicit quadratic function of the elements of Y, so as to be able to apply the techniques of Sect. 3.2.[5]

We represent Y by its Hermite normal basis relative to a particular basis of H_2: the first 32 elements of Y have the form $u_i = x_i^2 + \ldots$ for $i = 1, \ldots, 32$, the next 31 elements have the form $v_i = x_1 x_i + \ldots$ for $i = 2, \ldots, 32$, and the last one is $w = x_2 x_3 + \ldots$, where none of the omitted "..." expressions contain either squares or terms $x_1 x_i$ or $x_2 x_3$. In the (unlikely) case where the Hermite normal form of Y does not contain the monomials $x_1^2, \ldots, x_2 x_3$, we may always replace the public key F by its composition $F \circ \sigma$ by a random invertible linear

[5] We note that in the (very unlikely) case where the computation of the space Y performed in Sect. 3.1 returned a space $Y' \supsetneq Y$, the computation performed here will allow us to remove the few extra elements of Y': namely, since the elements of Z are quadratic forms in the elements of Y, the unneeded elements of Y' will not appear in these expressions. This means that, in practice, this step (Sect. 3.3) should be performed before the inner step (Sect. 3.2).

transformation of the input variables, such that $Y \circ \sigma$ has the suitable Hermite normal form.

We now consider a basis of the space Z. Any term of the form $\lambda\, x_i^4$ appearing in a basis element of Z comes, in its expression as a quadratic forms over Y, from a term $\lambda\, u_i^2$. Likewise, any term of the form $\mu\, x_1^3 x_i$ comes from a term $\mu\, u_1 v_i$, and so on.

In this way, we identify the second ASA layer as a quadratic map from Y to Z.

Solving the ASA Problem. The problem we have to solve is now almost identical to the one we solved in Sect. 3.2, except that we now have 16 instead of 8 S-boxes, and 96 instead of 64 quadratic forms.

Applying the previous algorithm to this case thus has a global complexity of approximately $q^4 \times 60 \times 16 \approx 2^{25.9}$ times the cost of computing the kernel of a matrix of size 64×96 with entries in $\mathbb{F}_{16}$. This is the dominant step in the key recovery procedure. We estimate the corresponding computational cost to a few CPU-hours.

3.4 Computing the Inhomogeneous Terms

We just presented an algorithm computing the *homogeneous* part (of degree two) of the quadratic S layers of the ASASA public key. These homogeneous terms represent the largest part of the secret key. Once they are computed, recovering the inhomogeneous terms (of degree one) is quite simple.

Each output S-box has one such linear term, represented by four coefficients; in total, there are therefore $(64+96)\times 4 = 640$ unknown coefficients. We consider the homogeneous parts of degree one and three of the public key PK_i. These functions are linear in the unknown inhomogeneous terms. Since there are exactly $(\dim H_1 + \dim H_3)\times 32 = 192\,512$ such functions, we have enough linear equations to recover all inhomogeneous terms.

Conclusion

We presented a very efficient distinguisher on the main ASASA scheme proposition of [3], that evolved into a full key-recovery algorithm with very reasonable complexity. The complexity of the attack can be approximated by the cost of computing the kernels of 2^{26} matrices of size 64×96 with entries in $\mathbb{F}_{16}$. This cost is well within practical limits. A classical venture to "repairing" a multivariate cryptosystem is to consider the homogeneous variant of the broken scheme. We point out that our cryptanalysis works by considering the homogenous quartic parts of the polynomials of the public key, thus defeats any such attempt. Another possibility would be to reinforce the scheme by adding more perturbation polynomials. However, our attack still works without any modification even for a larger number of perturbations.

References

1. Biham, E.: Cryptanalysis of patarin's 2-round public key system with S boxes (2R). In: Preneel, B. (ed.) EUROCRYPT 2000. LNCS, vol. 1807, pp. 408–416. Springer, Heidelberg (2000)
2. Billet, O., Gilbert, H., Ech-Chatbi, C.: Cryptanalysis of a white box AES implementation. In: Handschuh, H., Hasan, M.A. (eds.) SAC 2004. LNCS, vol. 3357, pp. 227–240. Springer, Heidelberg (2004)
3. Biryukov, A., Bouillaguet, C., Khovratovich, D.: Cryptographic Schemes based on the ASASA structure: black-box, white-box, and public-key (extended abstract). In: Sarkar, P., Iwata, T. (eds.) ASIACRYPT 2014. LNCS, vol. 8873, pp. 63–84. Springer, Heidelberg (2014)
4. Biryukov, A., Shamir, A.: Structural cryptanalysis of ASASA. In: Pfitzmann, B. (ed.) EUROCRYPT 2001. LNCS, vol. 2045, pp. 395–405. Springer, Heidelberg (2001)
5. Daemen, J.: Cipher and hash function design strategies based on linear and differential cryptanalysis. Ph.D. thesis, Doctoral Dissertation, KU Leuven, March 1995
6. De Mulder, Y., Roelse, P., Preneel, B.: Cryptanalysis of the Xiao – Lai white-box AES implementation. In: Knudsen, L.R., Wu, H. (eds.) SAC 2012. LNCS, vol. 7707, pp. 34–49. Springer, Heidelberg (2013)
7. Ding-Feng, Y., Kwok-Yan, L., Zong-Duo, D.: Cryptanalysis of 2R schemes. In: Wiener, M. (ed.) CRYPTO 1999. LNCS, vol. 1666, pp. 315–325. Springer, Heidelberg (1999)
8. Faugère, J.-C., Perret, L.: Cryptanalysis of $2R^-$ schemes. In: Dwork, C. (ed.) CRYPTO 2006. LNCS, vol. 4117, pp. 357–372. Springer, Heidelberg (2006)
9. Goubin, L., Masereel, J.-M., Quisquater, M.: Cryptanalysis of white box DES implementations. In: Adams, C., Miri, A., Wiener, M. (eds.) SAC 2007. LNCS, vol. 4876, pp. 278–295. Springer, Heidelberg (2007)
10. Matsumoto, T., Imai, H.: Public quadratic polynomial-tuples for efficient signature-verification and message-encryption. In: Günther, C.G. (ed.) EUROCRYPT 1988. LNCS, vol. 330, pp. 419–453. Springer, Heidelberg (1988)
11. Michiels, W., Gorissen, P., Hollmann, H.D.L.: Cryptanalysis of a generic class of white-box implementations. In: Avanzi, R.M., Keliher, L., Sica, F. (eds.) SAC 2008. LNCS, vol. 5381, pp. 414–428. Springer, Heidelberg (2009)
12. Newman, M.: Integral Matrices. Academic Press, New York (1972)
13. Patarin, J.: Cryptanalysis of the matsumoto and imai public key scheme of eurocrypt '88. In: Coppersmith, D. (ed.) CRYPTO 1995. LNCS, vol. 963, pp. 248–261. Springer, Heidelberg (1995)
14. Patarin, J.: Asymmetric cryptography with a hidden monomial. In: Koblitz, N. (ed.) CRYPTO 1996. LNCS, vol. 1109, pp. 45–60. Springer, Heidelberg (1996)
15. Patarin, J., Goubin, L.: Asymmetric cryptography with S-Boxes is it easier than expected to design efficient asymmetric cryptosystems? In: Information and Communications, Security, pp. 369–380 (1997)
16. Rijmen, V., Preneel, B.: A family of trapdoor ciphers. In: Biham, E. (ed.) FSE 1997. LNCS, vol. 1267, pp. 139–148. Springer, Heidelberg (1997)
17. Sturmfels, B.: Algorithms in Invariant Theory. Springer Science & Business Media, New York (2008)
18. Wu, H., Bao, F., Deng, R.H., Ye, Q.-Z.: Cryptanalysis of Rijmen-Preneel trapdoor ciphers. In: Ohta, K., Pei, D. (eds.) ASIACRYPT 1998. LNCS, vol. 1514, pp. 126–132. Springer, Heidelberg (1998)