# Key Recovery Attack on the Partial Vandermonde Knapsack Problem

Dipayan  $Das^1$  and Antoine  $Joux^2$ 

<sup>1</sup>NTT Social Informatics Laboratories, Tokyo, Japan <sup>2</sup>CISPA Helmholtz Center for Information Security, Saarbrücken, Germany dipayan.das@ntt.com, joux@cispa.de

Abstract. The Partial Vandermonde (PV) Knapsack problem is an algebraic variant of the low-density inhomogeneous SIS problem. The problem has been used as a building block for various lattice-based constructions, including signatures (ACNS'14, ACISP'18), encryptions (DCC'15,DCC'20), and signature aggregation (Eprint'20). At Crypto'22, Boudgoust, Gachon, and Pellet-Mary proposed a key distinguishing attack on the PV Knapsack exploiting algebraic properties of the problem. Unfortunately, their attack doesn't offer key recovery, except for worst-case keys.

In this paper, we propose an alternative attack on the PV Knapsack problem which provides key recovery for a much larger set of keys. Like the Crypto'22 attack, it is based on lattice reduction and uses a dimension reduction technique to speed-up the underlying lattice reduction algorithm and enhance its performance. As a side bonus, our attack transforms the PV Knapsack problem into uSVP instances instead of SVP instances in the Crypto'22 attack. This also helps the lattice reduction algorithm, both from a theoretical and practical point of view.

We use our attack to re-assess the hardness of the concrete parameters used in the literature. It appears that many contain a non-negligible fraction of weak keys, which are easily identified and extremely susceptible to our attack. For example, a fraction of  $2^{-19}$  of the public keys of a parameter set from ACISP'18 can be solved in about 30 hours on a moderate server using off-the-shelf lattice reduction. This parameter set was initially claimed to have a 129-bit security against key recovery attack. Its security was reduced to 87-bit security using the distinguishing attack from Crypto'22. Similarly, the ACNS'14 proposal also includes a parameter set containing a fraction of  $2^{-19}$  of weak keys; those can be solved in about 17 hours.

## 1 Introduction

The PV Knapsack problem, previously called the partial Fourier recovery problem, was introduced in [HPS<sup>+</sup>14] as a new lattice-based assumption for postquantum cryptography. The efficiency and rich algebraic properties underlying the PV Knapsack problem make it an attractive choice. As a result, the problem has been used as a building block for various primitives, such as encryptions [HS15,BSS22], signatures [HPS<sup>+</sup>14,LZA18], and aggregatable signatures [DHSS20].

Let  $\mathcal{R}_q := \mathbb{Z}_q[x]/(g(x))$  be a quotient polynomial ring, where g splits linearly over  $\mathbb{Z}_q$  for some prime q. In the literature, g is commonly either  $x^n - 1$  with prime n or  $x^n + 1$  with power-of-two n. In the rest of the paper, we assume that g corresponds to one of these choices. Denote by  $\Omega$  the set of all the primitive roots of g over  $\mathbb{Z}_q$ . Consider  $\Omega_t$ , a uniformly selected random subset of  $\Omega$  of size t. The PV Knapsack problem (informally) states the following:

It is hard to recover a uniform ternary  $f(x) \in \mathcal{R}_q$ , given only the evaluations of f(x) at  $\omega \in \Omega_t$  when  $t \approx \lfloor n/2 \rfloor$ .

The PV Knapsack problem also has a decisional version, which asks to distinguish between the evaluations of arbitrary f(x) and ternary f(x) at  $\omega \in \Omega_t$ .

The main approach to solving the problem has been the lattice reduction algorithms [HPS<sup>+</sup>14]. Recently, the authors of [BGP22] proposed an algebraic method that reduces the cost of the lattice reduction for solving the decisional problem.

The distinguishing attack of [BGP22] doesn't lead to a key recovery attack in general. In Section 5 of [BGP22], key recovery is only obtained for a small number of worst-case keys. Furthermore, their paper states: We note however that this does not fully invalidate the claim made in [LZA18], since the 128 bit-security is claimed against search attackers, and not distinguishing attackers.

This quote is the starting motivation for this work. Indeed, it might be worthwhile – from an attacker's viewpoint – to find a search attack against the PV Knapsack. As far as we know, there are some lattice-based assumptions where the search problem remains intractable, even though the decision problem is easy; one example is the FFI problem [DJ23].

#### 2 Preliminaries

#### 2.1 Notations

For any integer N > 1, we write  $\mathbb{Z}_N$  to denote the ring of integers modulo Nand  $\mathbb{Z}_N^*$  to denote the multiplicative subgroup of its units. In particular, when q is prime,  $\mathbb{Z}_q$  is the finite field with q elements. We assume that q is odd and, in that case, we represent elements of  $\mathbb{Z}_q$  by the unique representative belonging to the interval [-(q-1)/2, (q-1)/2].

We let  $\mathcal{R}_q = \mathbb{Z}_q[x]/(g)$  denote the quotient polynomial ring of  $\mathbb{Z}_q[x]$  by g(x), where g(x) is either  $x^n - 1$  with n a prime or  $x^n + 1$  with n a power of two. We insist that g splits into linear factors over  $\mathbb{Z}_q$ . We denote by  $\Omega$  the set of all primitive roots of g in  $\mathbb{Z}_q$ . When n is a prime,  $\Omega$  contains all roots of g except 1; when n is a power of two,  $\Omega$  contains all roots of g. In both cases, for any  $\omega \in \Omega$  and any root  $\omega'$  of g in  $\mathbb{Z}_q$ ,  $\omega'$  can be written as a power of  $\omega$ , say  $\omega^{i_{\omega'}}$ . In particular, when n is prime  $1 = \omega^0$  in  $\mathbb{Z}_q$ . Note that for a prime value of n, the exponent  $i_{\omega'}$  can take all values in  $\mathbb{Z}_n$ . When n is a power of two, the exponent  $i_{\omega'}$  takes all odd values in  $\mathbb{Z}_{2n}$ . As a consequence, if we exclude the non-primitive root 1, the exponents  $i_{\omega'}$  belong to  $\mathbb{Z}_n^*$  when *n* is prime and  $\mathbb{Z}_{2n}^*$  when *n* is a power of two. To lighten notations, we use U(n) as a shorthand for  $\mathbb{Z}_n^*$  when *n* is prime and  $\mathbb{Z}_{2n}^*$  when *n* is a power of two.

As a consequence, it is convenient to choose an arbitrary primitive root  $\omega_1$  of g in  $\mathbb{Z}_q$  and write:

$$\Omega = \{ \omega_i = \omega_1^i \mid i \in U(n) \}.$$

Remark 1. The condition that g splits in  $\mathbb{Z}_q$  implies that  $q = 1 \mod n$  when n is prime and that  $q = 1 \mod 2n$  when n is a power of two.

To represent polynomial in  $\mathcal{R}_q$ , we use the polynomial basis  $\{1, \boldsymbol{x}, \ldots, \boldsymbol{x}^{n-1}\}$ . Since we are working modulo  $\boldsymbol{g}$ , for any polynomial  $\boldsymbol{f}$  in  $\mathcal{R}_q$ , we can interpret  $\boldsymbol{f}(1/\boldsymbol{x})$  as a polynomial. More precisely, if  $\boldsymbol{f}(\boldsymbol{x}) = f_0 + f_1 \boldsymbol{x} + \cdots + f_{n-1} \boldsymbol{x}^{n-1}$ , we define:

$$\boldsymbol{f}(1/\boldsymbol{x}) = \begin{cases} f_0 + f_1 \boldsymbol{x}^{n-1} + \dots + f_{n-1} \boldsymbol{x} & \text{when } n \text{ is prime} \\ f_0 - f_1 \boldsymbol{x}^{n-1} - \dots - f_{n-1} \boldsymbol{x} & \text{when } n \text{ is a power of two} \end{cases}$$

It is easy to verify that for any root  $\omega$  of g, the evaluation of f(x) at  $\omega^{-1}$  coincides with the evaluation of f(1/x) at  $\omega$ , thus justifying our definition.

Since we have specified a basis for polynomials in  $\mathcal{R}_q$ , we can identify a polynomial with the vector of its coefficients in this basis. We use this identification extensively in the descriptions of the various attacks. For any vector  $\boldsymbol{v}$  (or polynomial using the vector identification), we write  $\|\boldsymbol{v}\|$  (resp.  $\|\boldsymbol{v}\|_{\infty}$ ) to denote the  $\ell_2$  norm (resp.  $\ell_{\infty}$  norm) of  $\boldsymbol{v}$ . We also write  $\boldsymbol{A} = (\boldsymbol{A}_1 | \boldsymbol{A}_2)$  to denote the concatenation of two matrices  $\boldsymbol{A}_1$  and  $\boldsymbol{A}_2$ , with the same number of rows.

#### 2.2 The PV Knapsack problem

Let  $\Omega_t$  be a subset of  $t \leq \lfloor n/2 \rfloor$  distinct random elements from  $\Omega$ . Let f(x) be a polynomial in  $\mathcal{R}_q$  whose coefficients are sampled uniformly at random from the set  $\{-1, 0, 1\}$ .

**Definition 1** (PV Knapsack problem). Given  $\mathcal{R}_q$  and  $\{(\omega, f(\omega)) \mid \omega \in \Omega_t\}$ , recover f(x).

Instead of identifying a PV Knapsack instance by  $\Omega_t$ , it is often simpler to identify it by the corresponding index set  $S_t \subset U(n)$ . The PV Knapsack instance then becomes  $\{(i, f(\omega_1^i) \mid i \in S_t)\}$  for some arbitrary primitive root  $\omega_1$  of g.

Remark 2. When n is prime, we have to assume the evaluation of f at 1 is never included, as this provides a simple distinguishing attack on the PV Knapsack problem. This explains why we choose  $\Omega$  to only contain primitive roots.

#### 2.3 Lattice Reduction

Any (full rank) matrix  $\boldsymbol{B} \in \mathbb{Z}^{n \times n}$  generates a lattice  $\mathcal{L}$  of dimension n, which is the set  $\mathcal{L}(\boldsymbol{B}) = \{\boldsymbol{B}\boldsymbol{z} : \boldsymbol{z} \in \mathbb{Z}^n\}$ . A lattice is called q-ary when it contains  $q\mathbb{Z}^n$  as a sublattice. The volume of the lattice  $\mathcal{L}(\boldsymbol{B})$  is defined as  $\mathsf{Vol} = |\det(\boldsymbol{B})|$ .

The key computational problem involving lattices is to find the shortest nonzero vector (SVP) in the lattice  $\mathcal{L}$ . Minkowski's theorem yields the following upper bounds on the norms of the shortest non-zero vector v of any lattice of dimension n and volume Vol:

$$\|\boldsymbol{v}\|_{\infty} \leq \operatorname{Vol}^{1/n}$$
 and  $\|\boldsymbol{v}\| \leq \sqrt{n} \operatorname{Vol}^{1/n}$ .

**Definition 2 (q-ary Kernel lattice).** Let  $A \in \mathbb{Z}_q^{t \times n}$  be any (full rank) matrix with n > t. We define the q-ary Kernel lattice of A as

$$\mathcal{L}_{\boldsymbol{A},q}^{\perp} = \{ \boldsymbol{v} \in \mathbb{Z}^n : \boldsymbol{A}\boldsymbol{v} = 0 \bmod q \}.$$

If we write  $\mathbf{A} = (\mathbf{A}_1 | \mathbf{A}_2)$ , where  $\mathbf{A}_1 \in \mathbb{Z}_q^{t \times t}, \mathbf{A}_2 \in \mathbb{Z}_q^{t \times n-t}$ , then assuming that  $\mathbf{A}_1$  is invertible,  $\mathcal{L}_{\mathbf{A},q}^{\perp}$  has a basis

$$egin{pmatrix} q oldsymbol{I}_t & -oldsymbol{A}_1^{-1}oldsymbol{A}_2 \ oldsymbol{0} & oldsymbol{I}_{n-t} \end{pmatrix}$$
 .

If  $A_1$  is not invertible, we can simply re-order the columns to make A start with a  $t \times t$  invertible matrix. The lattice  $\mathcal{L}_{A,q}^{\perp}$  has dimension n and volume  $q^t$ . Finding a short vector in this lattice, i.e., a short element in the kernel of A, is usually referred to as the short integer solution (SIS) problem.

Let  $\lambda_i$  denotes the smallest radius of a closed ball containing at least *i* linearly independent vectors in the lattice  $\mathcal{L}$ . If  $\lambda_2 > \gamma \lambda_1$  for some  $\gamma \ge 1$ , then the lattice contains a  $\gamma$ -unique SVP (uSVP) solution.

**Definition 3** (uSVP<sub> $\gamma$ </sub> problem). Given a lattice  $\mathcal{L}$ , with the promise that  $\lambda_2 > \gamma \lambda_1$  for  $\gamma \geq 1$ , the uSVP<sub> $\gamma$ </sub> problem asks to find  $\boldsymbol{v}$  such that  $\|\boldsymbol{v}\| = \lambda_1$ . The  $\gamma$  is referred to as the uniqueness gap of the uSVP problem.

In order to find short solutions in a lattice, we rely on the lattice reduction algorithms. LLL [LLL82] is a polynomial time algorithm, but only gives an exponential approximation solution. For cryptanalysis, it is often required for better solutions, which is done using stronger (and so slower) lattice reduction algorithms. For our purpose, we use the implementation of the Blockwise Korkine-Zolotarev (BKZ) algorithm [SE91] given in the fplll software [The23].

According to the analysis of [GN08], the uSVP problem, with uniqueness gap  $\gamma$  in dimension *n*, can be solved using a lattice reduction algorithm that achieves a root Hermite factor close to  $\delta = \gamma^{1/n}$ . In particular, when  $\gamma$  is large enough, the value of  $\delta$  becomes achievable with practical lattice reduction [ADPS16,AGVW17]. With high enough values of  $\delta$ , the uSVP problem becomes efficiently solvable.

#### **3** Previous Attacks

In this section, we briefly describe the attacks that were considered in prior works. Based on the approach of the attacks, we can characterise them as primal and dual attacks in the context of the PV Knapsack problem.

#### 3.1 Direct Primal Attack [HPS<sup>+</sup>14]

The problem can be expressed as a structured variant of the low-density inhomogeneous SIS (or LWE) problem by expressing the evaluation  $f(\omega)$  in terms of powers of  $\omega$ , which is stated below.

Given a (partial) Vandermonde matrix  $\boldsymbol{V} \in \mathbb{Z}_q^{t \times n}$  (with rows generated by powers of  $\omega$  for  $\omega \in \Omega_t$ ) and  $\boldsymbol{b} \in \mathbb{Z}_q^t$  (with elements  $\boldsymbol{f}(\omega)$ ), find  $\boldsymbol{f}$  with  $\|\boldsymbol{f}\|_{\infty} \leq 1$  such that

$$\boldsymbol{V}\boldsymbol{f} = \boldsymbol{b} \bmod q, \tag{1}$$

The authors proposed the strategy of finding the uSVP solution (following Kannan's embedding [LSL13]) on the kernel lattice

$$\mathcal{L}_{oldsymbol{V}',q}^{\perp} = \{oldsymbol{v} \in \mathbb{Z}^{n+1}: oldsymbol{V}'oldsymbol{v} = oldsymbol{0} mmod q\}$$

where  $\mathbf{V}' = (\mathbf{V}|\mathbf{b})$ . Note that,  $(\mathbf{f}|-1)^T$  is a vector in the lattice  $\mathcal{L}_{\mathbf{V}'q}^{\perp}$ , which is a solution to the uSVP problem<sup>1</sup>. In practice, this direct attack is used as a baseline to choose the parameters, with the understanding that they should be selected to ensure that finding the uSVP solution remains intractable both on classical and quantum computers.

#### 3.2 Dual Attack [BGP22]

Here, we give a simplified version of the attack proposed in [BGP22]. For this purpose, we need to restate the PV Knapsack problem as an instance of the Bounded Distance Decoding (BDD) problem in the following manner. Let  $\boldsymbol{z} \in \mathbb{Z}^n$  be any solution to the system of linear equations satisfying  $\boldsymbol{V}\boldsymbol{z} = \boldsymbol{b} \mod q$ , then the PV Knapsack problem asks to find  $\boldsymbol{u} \in \mathcal{L}_{\boldsymbol{V},q}^{\perp}$  (i.e.,  $\boldsymbol{V}\boldsymbol{u} = \boldsymbol{0} \mod q$ ) such that  $\|\boldsymbol{u}-\boldsymbol{z}\|_{\infty} \leq 1$ , i.e., the vector  $\boldsymbol{u}-\boldsymbol{z} = \boldsymbol{f}$ . Algebraically, the element  $\boldsymbol{u}(\boldsymbol{x})$  belongs to the ideal  $I_{\Omega_t}$  of  $\mathcal{R}_q$ , where  $I_{\Omega_t} = \prod_{\omega \in \Omega_t} (\boldsymbol{x}-\omega)$ . So, the PV Knapsack problem can be considered a BDD problem in the ideal  $I_{\Omega_t}$ .

Let u'(x) be an element in the ideal  $I_{\Omega \setminus \Omega_t} = \prod_{\omega' \in \Omega \setminus \Omega_t} (x - \omega')$  with "somewhat" small norm.<sup>2</sup> Then the product u'(x)z(x) = u'(x)f(x) in  $\mathcal{R}_q$  is expected

<sup>&</sup>lt;sup>1</sup> The ternary choice of  $\boldsymbol{f}$  makes the vector  $(\boldsymbol{f}|-1)^T$  extremely short in the lattice  $\mathcal{L}_{\boldsymbol{V}',q}^{\perp}$ . As a consequence, with high probability, it is significantly shorter than any vector that should normally occur in a lattice with this dimension and volume, which justifies the uniqueness assumption on the shortest vector.

<sup>&</sup>lt;sup>2</sup> The elements of the ideal  $I_{\Omega \setminus \Omega_t}$  correspond to the vectors of the *q*-ary Kernel lattice generated by the Vandermonde matrix with rows powers of  $\omega'$  for  $\omega' \in \Omega \setminus \Omega_t$ . Note that the lattice doesn't contain any unusually short vector, so we can expect the shortest vector to have a norm predicted by Minkowski's theorem.

to be small<sup>3</sup> (with coefficients  $\langle q/4 \rangle$ ) for a PV Knapsack instance – a highly unlikely event for a uniform instance. This gives a key distinguishing attack. The cost of the dual attack depends on finding the small u'(x), which can be improved using the algebraic methods.

Let  $\Omega_{2t_1}$  be the largest subset of  $\Omega_t$  that remains invariant under the computation of inverses. In other words,  $\Omega_{2t_1}$  contains  $t_1$  pairs  $(\omega, \omega^{-1})$  with both  $\omega$  and  $\omega^{-1}$  in  $\Omega_t$ . This set  $\Omega_{2t_1}$  is easily constructed by removing any element of  $\Omega_t$  whose inverse is not in  $\Omega_t$ . Thus,  $2t_1 \leq t$  and  $t_1 \leq \lfloor t/2 \rfloor$ . By our choice of  $\boldsymbol{g}$ , all the roots in  $\Omega$  can be paired with their inverse. As a consequence, the complement set  $\Omega \setminus \Omega_{2t_1}$  is also made of such pairs.

This symmetry can be leveraged to find a small element in the ideal  $I_{\Omega \setminus \Omega_{2t_1}}$ , by looking for a small polynomial  $u'(x) \in I_{\Omega \setminus \Omega_{2t_1}}$  with the extra requirement that:

for all 
$$\omega \in \Omega$$
 :  $\boldsymbol{u}'(\omega) = \boldsymbol{u}'(\omega^{-1}).$ 

This is easily achieved by creating u'(x) using a basis of halved dimension obtained from the symmetrisation of  $\{1, x, \ldots, x^{\lfloor n/2 \rfloor}\}$ . For such a polynomial, when x is a root, so is 1/x. Thus we can guarantee that u' vanishes on  $\Omega_{2t_1}$ using only  $t_1$  linear conditions. As a consequence, a small u'(x) can be found using lattice reduction in a (Kernel) lattice of reduced dimension.

Furthermore, if  $t_1$  is not too small, the PV Knapsack problem still reduces to a BDD problem in the ideal  $I_{\Omega_{2t_1}}$ . The condition of  $t_1$  not being too small comes from considering the volume of the lattice, which decreases with  $t_1$  and needs to be large enough for the reduction to BDD work. When  $t_1$  is sufficient, considering the product u'(x)z(x) again gives a distinguishing attack. The authors of [BGP22] (experimentally) show that this occurs with non-negligible probability and thus improves the cost of solving the decisional PV Knapsack problem.

As an extension of this attack, and for some choices of n, one can also aim at exploiting higher order symmetries to reduce to a lattice problem of even smaller dimension. Unfortunately, in general, this reduces the number of evaluations at the roots after symmetrisation too much. So the reduction to BDD no longer works. However, if  $\Omega_t$  can be adversarially chosen, we obtain a degraded version of the PV Knapsack problem. This is called a *worst-case*  $\Omega_t$  in [BGP22]. In this worst-case,  $\Omega_t$  contains a large subset  $\Omega_{rt_0}$ , which remains invariant under a symmetry of order r (instead of 2). <sup>4</sup> With such a forced symmetry inside, this allows the PV Knapsack to remain a BDD instance in the ideal  $I_{\Omega_{rt_0}}$ .

Like before, the set  $\Omega \setminus \Omega_{rt_0}$  also remains invariant under the transformation, so the problem of finding a short solution u' in the ideal  $I_{\Omega/\Omega_{rt_0}}$  can be reduced

<sup>&</sup>lt;sup>3</sup> Note that, the product  $\boldsymbol{u}\boldsymbol{u}' = \boldsymbol{0}$  in  $\mathcal{R}_q$ . For the prime *n* case, we need to include the factor  $(\boldsymbol{x} - 1)$  in the product  $\boldsymbol{u}'(\boldsymbol{x})\boldsymbol{z}(\boldsymbol{x})$  to make  $\boldsymbol{u}\boldsymbol{u}' = \boldsymbol{0}$  in  $\mathcal{R}_q$ . Furthermore, the choice of  $\boldsymbol{g}$  (sparse with small coefficient) leads to O(n) coefficient growth of the product  $\boldsymbol{u}'(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x})$ .

<sup>&</sup>lt;sup>4</sup> For example, the power of two n allows the use of a subgroup of maximal order r = n/2, while the prime n of the form  $n - 1 = 0 \mod 3$  from the parameters of [HPS<sup>+</sup>14] allows the use of a subgroup of order up to r = (n - 1)/3.

to a lattice of (very) small dimension. Since finding an SVP solution is known in such a small dimension, e.g., using LLL lattice reduction algorithm, then, (hopefully) the product  $u'(x)z(x) \in \mathcal{R}_q$  has all coefficients < q/2 in absolute value (i.e., no wrap-around modulo q happens for the product polynomial), which also gives a key recovery attack.

Because of this worst-case, it appears that the uniformly random choice of  $\Omega_t$  makes more sense in the definition of the PV Knapsack problem. This approach is used in [BSS22,LZA18], while [HPS<sup>+</sup>14] doesn't explicitly mention the choice of  $\Omega_t$ . In the rest of the paper, we concentrate on the key recovery attack for a uniformly random  $\Omega_t$ .

#### 4 Our Contribution

Our main goal is to find an alternative dimension reduction strategy working with the primal attack instead of the dual attack. Indeed, the primal attack corresponds to a uSVP instance, which is believed to be comparatively easier to solve than an SVP instance, both in theory [LSL13], and in practice [GN08,ADPS16,AGVW17].

We achieve this goal by proposing a new dimension reduction primal attack on the PV Knapsack problem. For this, we exploit the symmetries of the ring  $\mathcal{R}_q$  in a new way. This allows us to solve several PV Knapsack instances from the literature in a reasonable time, faster than what was previously thought to be possible.

As in [BGP22], we consider the largest subset  $\Omega_{2t_1}$  of  $\Omega_t$  that remains invariant under the computation of inverses. For any  $\omega$  in  $\Omega_{2t_1}$ , we know the evaluation of  $\boldsymbol{f}$  both at  $\omega$  and  $\omega^{-1}$ . Hence we can compute  $\boldsymbol{f}(\omega) \pm \boldsymbol{f}(\omega^{-1})$ . This gives  $t_1$ distinct evaluations of the two polynomials  $\boldsymbol{f}(\boldsymbol{x}) \pm \boldsymbol{f}(1/\boldsymbol{x})$  at  $\omega \in \Omega_{2t_1}$ . We aim to recover  $\boldsymbol{f}(\boldsymbol{x}) \pm \boldsymbol{f}(1/\boldsymbol{x})$  as uSVP solutions from lattices of smaller dimensions and do the linear algebra to recover the secret  $\boldsymbol{f}(\boldsymbol{x})$ .

Let  $\psi_{+}(x) = f(x) + f(1/x)$  and  $\psi_{-}(x) = f(x) - f(1/x)$ . The polynomials  $\psi_{+}(x)$ , and  $\psi_{-}(x)$  can be generated by a basis of order  $n_{+} = \lceil n/2 \rceil$  and  $n_{-} = \lfloor n/2 \rfloor$ , respectively. These bases are easy to compute from the polynomial basis. Also, if f(x) has coefficients in the set  $\{-1, 0, 1\}, \psi_{\pm}(x)$  has coefficients in the set  $\{-2, -1, 0, 1, 2\}$ . Then the PV Knapsack problem reduces to two independent problems of finding  $\psi_{\pm}(x)$  from  $t_1$  evaluations. This can be achieved by recovering uSVP solutions in lattices of dimensions  $n_{\pm}$ .

There are a few important observations from the above attack.

1. The cost of recovering  $\psi_{\pm}$  as a uSVP solution (using lattice reduction algorithm) depends on the volume of the lattice in reduced dimension. The volume is proportional to the number of distinct evaluations  $t_1$ , which makes the problem easier as  $t_1$  increases. When  $\Omega_t$  is randomly chosen, the value of  $t_1$  is randomised. If the system is used by many users, each one with its own set  $\Omega_t$ , some of them will pick weak keys, i.e., weak sets  $\Omega_t$ , which are easier to attack because of their larger value of  $t_1$ . To analyse our attack, two ingredients are needed: an attack that works when  $t_1$  is large enough

and a probability analysis of this weak-key event.

- 2. Note that, when a PV Knapsack instance is given, an adversary can compute  $t_1$  easily. This only requires reading  $\Omega_t$  to detect pairs of the form  $(\omega, \omega^{-1})$ . As a consequence, the adversary can focus on the keys that are easy enough to attack with lattice techniques. This can be done, for example, by using LWE estimators [APS15,DDGR20] before starting the attack.
- 3. Since the two uSVP problems for finding  $\psi_{\pm}$  are independent, the corresponding lattice reductions can be performed in parallel. Hence, the running time of the attack is directly obtained by estimating the cost of the largest of the two uSVP instances.
- 4. We also study symmetries of order > 2 and their application to a direct attack to solve PV Knapsack. Unfortunately, for random choices of  $\Omega_t$ , it turns out that the symmetry of order 2 is optimal for the parameters proposed in the literature.

In Section 5, we formally describe the attack sketched above. In Section 6, we provide experimental results that indicate that several proposed instances of the PV Knapsack problem can be solved in practice. In Section 7, we give a generalized version of the attack using symmetries of higher order. We hope that despite their inefficiency for the random case, their analysis can be of independent interest.

## 5 Proposed Attack

In this section, we propose a new key recovery attack on the PV Knapsack problem. The key idea is to use symmetry in a new way, thanks to the following lemma.

**Lemma 1.** Let  $f(\mathbf{x})$  be any polynomial in  $\mathcal{R}_q$ , then  $\psi_{\pm}(\mathbf{x}) = f(\mathbf{x}) \pm f(1/\mathbf{x})$ can be generated by a basis of order  $n_{\pm}$ , where  $n_{\pm} = \lceil n/2 \rceil$  and  $n_{\pm} = \lfloor n/2 \rfloor$ . Moreover, if the coefficients of  $f(\mathbf{x})$  are sampled uniformly at random from the set  $\{-1, 0, 1\}$ , then the expected squared-norm of  $\psi_{\pm}(\mathbf{x})$  is upper-bounded by  $4n_{\pm}/3$  in the new basis representation.

Proof. The mapping

$$\boldsymbol{x}^i \rightarrow \boldsymbol{x}^i + 1/\boldsymbol{x}^i$$
 for  $0 \leq i \leq |n/2|$ 

is well-defined. Hence, by linearity, the polynomial  $\psi_+(x) = f(x) + f(1/x)$  can be generated by a basis of order  $n_+$ , as required. In particular, for prime n, since  $1/x = x^{n-1}$ ,  $\psi_+(x)$  is generated by the basis

$$\{2, (\boldsymbol{x} + \boldsymbol{x}^{n-1}), \dots, (\boldsymbol{x}^{\lfloor n/2 \rfloor} + \boldsymbol{x}^{\lfloor n/2 \rfloor + 1})\}$$

For power of two n, since  $1/x = -x^{n-1}, \psi_+(x)$  is generated by the basis

$$\{2, (\boldsymbol{x} - \boldsymbol{x}^{n-1}), \dots, (\boldsymbol{x}^{(n/2-1)} - \boldsymbol{x}^{(n/2+1)})\}$$

Similarly, the mapping

$$\boldsymbol{x}^i \to \boldsymbol{x}^i - 1/\boldsymbol{x}^i$$
 for  $1 \le i \le |n/2|$ 

is well-defined. Hence, by linearity, the polynomial  $\psi_{-}(x) = f(x) - f(1/x)$  can be generated by a basis of order  $n_{-}$ , as required. In particular, for prime n,  $\psi_{-}(x)$  is generated by the basis

$$\{(\boldsymbol{x}-\boldsymbol{x}^{n-1}),\ldots,(\boldsymbol{x}^{\lfloor n/2 \rfloor}-\boldsymbol{x}^{\lfloor n/2 \rfloor+1})\}$$

For power of two  $n, \psi_{-}(x)$  is generated by the basis

$$\{(x+x^{n-1}),\ldots,(x^{n/2-1}+x^{n/2+1}),2x^{n/2}\}$$

If individual coefficients of f are uniformly sampled from  $\{-1, 0, 1\}$ , then sums of symmetric coefficients  $f_i + f_{n-i}$  are in  $\{-2, -1, 0, 1, 2\}$  and follow the probability distribution given in Table 1.

| $f_i + f_{n-i}$ | 0   | 1   | -1  | 2   | -2  |
|-----------------|-----|-----|-----|-----|-----|
| Prob.           | 3/9 | 2/9 | 2/9 | 1/9 | 1/9 |

Table 1: Probability distribution of  $f_i + f_{n-i}$ 

Now, if f(x) is sampled uniformly with ternary coefficients, most coefficients of  $\psi_{\pm}(x)$  follow the distribution of  $f_i + f_{n-i}$ . The exceptions being the special coefficients associated to 2 and  $2x^{n/2}$  which follow the initial uniform distribution in  $\{-1, 0, 1\}$  and have a lower expectation of their squares. Hence, by linearity of expectations, the expected squared-norm of  $\psi_{\pm}$  in the new basis representation is upper-bounded by  $4n_{\pm}/3$ .

This allows us to design a new low-density inhomogeneous SIS problem corresponding to the evaluation of  $\psi_{\pm}$  at  $t_1$  values. In order to do this, let us create a matrix  $W_{\pm}$  with  $t_1$  rows and  $n_{\pm}$  columns, whose entries are the evaluations of each of the  $n_{\pm}$  monomials at an arbitrary choice of  $t_1$  representatives for the pairs  $(\omega, \omega^{-1})$  that occur in  $\Omega_{2t_1}$ . We also create a vector  $\mathbf{b}_{\pm}$  whose coefficients are the known evaluations of  $\psi_{\pm}$  at each of the representative. With these notations, we look for a solution of:

$$\boldsymbol{W}_{\pm}\boldsymbol{\psi}_{\pm} = \boldsymbol{b}_{\pm} \bmod q. \tag{2}$$

Following the same strategy as in the direct primal attack, we search for a short vector in the kernel lattice:

$$\mathcal{L}_{\boldsymbol{W}_{\pm}',q}^{\perp} = \{ \boldsymbol{v} \in \mathbb{Z}^{n+1} : \boldsymbol{W}_{\pm}' \boldsymbol{v} = \boldsymbol{0} \bmod q \}$$

where  $W'_{\pm} = (W_{\pm}|b_{\pm})$ . As before,  $(\psi_{\pm}|-1)^T$  is a very short vector in the lattice and we expect that it yields a uSVP solution.

#### 5.1 Analysis of the new attack

As we already mentioned, to analyse the attack, we need two ingredients. First, given  $t_1$ , the number of pairs in  $\Omega_t$ , we need to estimate the cost of successfully conducting the uSVP computation. Second, for a random set  $\Omega_t$ , we need to compute the probability of occurrence of a given value of  $t_1$ . Since from a public key,  $t_1$  can be computed extremely efficiently, this probability directly corresponds to the fraction of users that can be attacked with the corresponding uSVP problem.

**Cost of the uSVP resolution.** Thanks to the analysis of [GN08], we know that the cost of solving a uSVP problem mostly depends on the root Hermite factor that can be computed from the uniqueness gap  $\gamma$ . We recall that this factor is  $\delta = \gamma^{1/n}$ .

In our attack, we do not really have a promise problem. However, since the lattices we consider come from a cryptographic problem, we can follow a standard heuristic approach and assume that they behave as randomly as it can. More precisely, both for the direct primal attack of [HPS<sup>+</sup>14] and for our new attack with dimension reduction, we consider a lattice in which a vector of short length is guaranteed. The heuristic we use is to consider that other (linearly independent) vectors in the lattice have a length which can be estimated from Minkowski's bound. In other words, given its volume V and dimension d, we estimate the value of  $\lambda_2$  to be  $\sqrt{d} V^{1/d}$ . To estimate  $\lambda_1$ , we use the squareroot of the expected squared-norm. Putting the two estimations together, it just remains to compute  $\gamma = \lambda_2/\lambda_1$  and take its d-th root to obtain the corresponding  $\delta$ .

Recall, in the (full) primal attack, the PV Knapsack gives a uSVP instance in dimension n+1, and volume  $q^t$ . We also have a short vector of expected squared-norm 2n/3+1. As a consequence, the corresponding root Hermite factor can be estimated by:

$$\delta_{\text{full}} = \left(\frac{\sqrt{n+1} \, q^{t/(n+1)}}{\sqrt{2n/3+1}}\right)^{1/(n+1)}$$

Similarly, in our attack, we get two lattices of dimensions  $n_{\pm} + 1$  and volume  $q^{t_1}$ . In that case, the expected squared-norm of the shortest vector is  $4n_{\pm}/3 + 1$ . Thus, we get an estimation of:

$$\delta_{\pm \text{new}} = \left(\frac{\sqrt{n_{\pm} + 1} \, q^{t_1/(n_{\pm} + 1)}}{\sqrt{4n_{\pm}/3 + 1}}\right)^{1/(n_{\pm} + 1)}$$

Following [GN08], we need to compare  $\delta_{\text{full}}$  and  $\delta_{\pm \text{new}}$  to know when the new attack beats the full primal attack. To do the comparison, we slightly simplify the expression, replacing  $n_{\pm}$  by n/2 and any instance of n+1 by n or  $n_{\pm}+1$  by  $n_{\pm}$ . After the simplification, we expect the new attack to become faster as soon as:

$$\left(\sqrt{3/2}q^{t/n}\right)^{1/n} < \left(\sqrt{3/4}q^{2t_1/n}\right)^{2/n}$$

Ignoring the small constants, this happens when  $4t_1 > t$ .

This estimation is somewhat pessimistic. Indeed, the dimension of the lattice also counts when using lattice reduction, so even when the root Hermite factors are equal, the newer lattice should be easier to reduce due to its smaller dimension.

**Distribution of**  $t_1$ . To study the probability distribution of  $t_1$ , we perform a standard combinatorial analysis. The total number of sets  $\Omega_t$  of t elements chosen from the primitive roots is:

$$\binom{2\lfloor n/2\rfloor}{t}.$$

When  $t_1$  is a fixed integer in  $[0, \lfloor t/2 \rfloor]$ , to choose a set of size t with exactly  $t_1$  pairs, we need to take  $t_1$  pairs from the  $\lfloor n/2 \rfloor$ , followed by  $t - 2t_1$  unpaired elements in the remaining pairs. Thus, the total number of possibilities is:

$$\binom{\lfloor n/2 \rfloor}{t_1} \binom{\lfloor n/2 \rfloor - t_1}{t - 2t_1} 2^{t - 2t_1}$$

As a consequence, the probability of getting  $t_1$  for a random  $\Omega_t$  is:

$$\pi_1(t_1) = \frac{\binom{\lfloor n/2 \rfloor}{t_1} \binom{\lfloor n/2 \rfloor - t_1}{t - 2t_1} 2^{t - 2t_1}}{\binom{2 \lfloor n/2 \rfloor}{t}}$$

When t = n/2, the distribution of the values of  $t_1$  is strongly concentrated around t/4, which is precisely the tipping point between the direct primal attack and our new attack. This is illustrated by Figure 1. However, we see that for this typical case,  $t_1$  can deviate from t/4. This explains the existence of weak instances vulnerable to our attack.

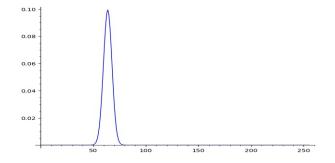


Fig. 1:  $\pi_1(t_1)$  for n = 512, t = 256

#### 6 Experimental Results

In this section, we analyse the effect of our attack on the concrete hardness of the problem used in the literature. We ran all our experiments on an Intel Xeon CPU E5-2683 v4 @ 2.10GHz 1200 MHz processor. The attack only depends on the value of  $t_1$ , and not on the choice of elements in  $\Omega_{2t_1}$ . So to perform experiments on our attack, we first fix a value of  $t_1$ . Then we sample a uniform primitive root  $\omega_1$  of g and characterise  $\Omega_{2t_1}$  by a random index set  $S_{t_1} \subset U(n)$  of size  $t_1$  (distinct up to negation). The lattice reduction algorithms are performed in parallel with fplll software [The23].

The running time of a BKZ lattice reduction algorithm is exponential on the blocksize. In [CN11], the authors experimentally observed that most of the progress is made in the initial rounds of the BKZ reductions for the (relatively) large blocksize. In our experiments, the running time is the time taken by the lattice reduction algorithm to discover the secret.

Since it is not feasible to run lattice reductions for every parameter, following the common practice, we use LWE estimators [APS15,DDGR20] to predict the running time of several instances. The LWE estimators heuristically predict the lattice reduction strength (which is characterised by the block size of the BKZ algorithm) required to find the secret in the primal attack.

|             | HPSSW1 | HPSSW2 | HPSSW3  | HPSSW4     |
|-------------|--------|--------|---------|------------|
| n           | 433    | 577    | 769     | 1153       |
| t           | 200    | 280    | 386     | 600        |
| q           | 775937 | 743177 | 1047379 | 968521     |
| $\lambda$   | << 62  | << 80  | 76      | $\geq 130$ |
| $\lambda^*$ | 47     | 52     | 63      | 87         |

Table 2: Parameters:  $PASS_{RS}$  [HPS<sup>+</sup>14]

## 6.1 **PASS<sub>RS</sub>** Signature from $[HPS^+14]$

In this paper, the authors proposed  $\mathsf{PASS}_{\mathsf{RS}}$  signature scheme following the Fiat-Shamir with aborts strategy on the hardness of the PV Knapsack problem. The scheme is defined for the prime *n* case of the problem. The proposed parameters are given in Table 2.

The  $\lambda$  in the Table is the claimed bit security in the proposal. The  $\lambda^*$  is the re-evaluated bit security in the direct primal attack using the LWE estimator [APS15],<sup>5</sup> except for HPSSW1.

 $<sup>^5</sup>$  As asked in  $\ [APS15],$  we include the commit value used in this paper, which is fd4a460.

|       | Direct primal attack: $\lambda^* = 52$ , BKZ block size 73 |                       |                |                |  |  |  |  |  |
|-------|--|-----------------------|----------------|----------------|--|--|--|--|--|
| $t_1$ | Prob.  | Running time (in hrs) | BKZ block size | Bits operation |  |  |  |  |  |
|       | $2^{-11}$  | 115                   | 65             | $2^{50}$       |  |  |  |  |  |
| 84    | $2^{-13}$  | 54                    | 60             | $2^{49}$       |  |  |  |  |  |
| 86    | $2^{-16}$  | 51                    | 60             | $2^{48}$       |  |  |  |  |  |
| 88    | $2^{-19}$  | 17                    | 60             | $2^{46}$       |  |  |  |  |  |
| 90    | $2^{-23}$  | 6                     | 58             | $2^{45}$       |  |  |  |  |  |

Table 3: Experimental results of our attack on the weak keys of HPSSW2.

| Dire  | ect pri   | mal attack: $\lambda^*$ = | = 63, BKZ block size 112 |
|-------|-----------|---------------------------|--------------------------|
| $t_1$ |           | BKZ block size            |                          |
| 110   | $2^{-8}$  | 112                       | $2^{61}$                 |
|       | $2^{-12}$ | 106                       | $2^{60}$                 |
|       | $2^{-13}$ | 103                       | $2^{59}$                 |
|       | $2^{-16}$ | 100                       | $2^{58}$                 |
|       | $2^{-19}$ | 95                        | $2^{56}$                 |
| 125   | $2^{-27}$ | 87                        | $2^{54}$                 |
|       | $2^{-32}$ | 84                        | $2^{53}$                 |
| 130   | $2^{-37}$ | 80                        | $2^{52}$                 |

Table 4: Predicted cost of our attack on the weak keys of HPSSW3 using the LWE estimator [APS15].

For HPSSW1, the bit security is achieved experimentally; we recovered the secret within 25 hours ( $2^{47}$ -bits operation) using BKZ block size 55. For this reason, we have excluded it in our attack analysis.

We ran experiments of our attack on the HPSSW2 weak keys. The experimental results are given in Table 3. For the other parameters, we use the LWE estimator from [APS15], the details are given in Table 4 and Table 5.

| Dire  | ect pri   | mal attack: $\lambda^*$ = | = 87, BKZ block size 200 |
|-------|-----------|---------------------------|--------------------------|
| $t_1$ | Prob.     | BKZ block size            | 1                        |
| 167   | $2^{-6}$  | 196                       | $2^{85}$                 |
| 170   | $2^{-7}$  | 191                       | $2^{83}$                 |
| 172   |           | 188                       | $2^{82}$                 |
| 177   | $2^{-12}$ | 180                       | $2^{80}$                 |
|       | $2^{-17}$ | 172                       | $2^{78}$                 |
|       |           | 165                       | $2^{76}$                 |
| 192   |           | 158                       | $2^{74}$                 |
| 198   | $2^{-39}$ | 151                       | $2^{72}$                 |

Table 5: Predicted cost of our attack on the weak keys of HPSSW4 using the LWE estimator [APS15].

|             | LZA1                  | LZA2         |
|-------------|-----------------------|--------------|
| n           | 512                   | 1024         |
| t           | 256<br>$2^{16} \pm 1$ | 512          |
| q           | $2^{16} + 1$          | $2^{16} + 1$ |
| $\lambda$   | 129                   | 198          |
| $\lambda^*$ | 54                    | 99           |

Table 6: Parameters: Signature scheme [LZA18]

#### 6.2 Signature scheme from [LZA18]

In this paper, the authors proposed a signature scheme on the hardness of the PV Knapsack problem following the  $\mathsf{PASS}_{\mathsf{RS}}$  signature scheme, but for the power-of-two *n* case. The proposed parameters are given in Table 6. The  $\lambda^*$  is computed using the LWE estimator [APS15].

Remark 3. Because of the huge difference between  $\lambda$  and  $\lambda^*$ , it is important to look for the source of the discrepancy. The best explanation we found is that the analysis in [LZA18] apparently considers the dimension of the lattice in the direct primal attack as n + t + 1 (Section 4 [LZA18]), instead of n + 1.

We ran experiments of our attack on the LZA1 weak keys. The experimental results are given Table 7. For LZA2, we use the LWE estimator from [APS15], the details are given in Table 8.

|       | Direct primal attack: $\lambda^* = 54$ , BKZ block size 83 |  |    |          |  |  |  |  |  |  |
|-------|--|--|----|----------|--|--|--|--|--|--|
| $t_1$ | Prob.  | cob. Running time (in hrs) BKZ block size Bits operation |    |          |  |  |  |  |  |  |
| 80    | $2^{-15}$  | 117  | 70 | $2^{50}$ |  |  |  |  |  |  |
| 83    | $2^{-19}$  | 30   | 60 | $2^{48}$ |  |  |  |  |  |  |
| 85    | $2^{-23}$  | 9.5  | 60 | $2^{46}$ |  |  |  |  |  |  |
|       |  | 8  | 60 | $2^{45}$ |  |  |  |  |  |  |
| 90    | $2^{-34}$  | 7.5  | 57 | $2^{45}$ |  |  |  |  |  |  |

Table 7: Experimental results of our attack on the weak keys of LZA1.

## 6.3 PASSEncrypt, PVRegevEncrypt schemes from [BSS22]

In this paper, the authors proposed PASSEncrypt, PVRegevEncrypt encryption schemes based on the hardness of the PV Knapsack problem. The schemes are defined for the power-of-two n case of the problem. While PASSEncrypt is a modified version of the encryption scheme proposed in [HS15], PVRegevEncrypt is a (partial) Vandermonde variant of the Regev-style encryption scheme. The

| Dire | ect pri   | mal attack: $\lambda^*$ = | = 99, BKZ block size 243 |
|------|-----------|---------------------------|--------------------------|
|      |           | BKZ block size            | Bits operation           |
|      | $2^{-8}$  | 237                       | $2^{97}$                 |
|      | $2^{-11}$ | 233                       | $2^{94}$                 |
| 148  | $2^{-13}$ | 228                       | $2^{93}$                 |
| 151  | $2^{-15}$ | 222                       | $2^{91}$                 |
| 154  | $2^{-19}$ | 216                       | $2^{90}$                 |
| 157  | $2^{-23}$ | 210                       | $2^{88}$                 |
|      | $2^{-31}$ | 199                       | $2^{85}$                 |
| 166  | $2^{-37}$ | 194                       | $2^{84}$                 |

Table 8: Predicted cost of our attack on the weak keys of LZA2 using the LWE estimator [APS15].

|                           | BSS1  | BSS2  |
|---------------------------|-------|-------|
| n                         | 1024  | 2048  |
| t                         | 512   | 1024  |
| q                         | 12289 | 12289 |
| $\lambda_Q = \lambda_Q^*$ | 79    | 188   |

Table 9: Parameters: PASSEncrypt, PVRegevEncrypt [BSS22]

proposed parameters are given in Table 9.

The concrete hardness of the parameters is computed using the LWE Leaky estimator [DDGR20]. The BKZ algorithm with block size  $\beta$  uses an SVP oracle in dimension  $\beta$ ; the running time is evaluated using the core SVP hardness, which is only the cost of one call to an SVP oracle in dimension  $\beta$ . They further considered one SVP call cost  $2^{0.265\beta}$  using a quantum algorithm. We also used the same estimation model for analysing the hardness of the weak keys. The details are given in Table 10 and Table 11.

## 7 Symmetries of higher order

In this section, we illustrate a generalized version of our attack to symmetries of higher order by going to order 3. It is straightforward to go to other orders. The case of order 3 naturally arises from the concrete parameters of [HPS<sup>+</sup>14]. Indeed, they use a prime n satisfying  $n - 1 = 0 \mod 3$  to do the fast Fourier transformation.

**Lemma 2.** Let n be a prime satisfying  $n-1 = 0 \mod 3$ , and let  $\theta$  be an element of order 3 in U(n), i.e.,  $\theta^3 = 1$  in U(n). For any polynomial  $\mathbf{f}(\mathbf{x}) \in \mathcal{R}_q = \mathbb{Z}_q[\mathbf{x}]/(\mathbf{x}^n - 1), \, \psi_1(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}(\mathbf{x}^{\theta}) + \mathbf{f}(\mathbf{x}^{\theta^2})$  can be generated by a basis of order  $n_{\theta} = \lceil n/3 \rceil$ . Moreover, if the coefficients of  $\mathbf{f}(\mathbf{x})$  are sampled uniformly

| Dire | ect pri   | mal attack: $\lambda_Q$ = | $=\lambda_Q^* = 79$ , BKZ block size 298 |
|------|-----------|---------------------------|--|
|      |           | BKZ block size            | Quantum core $SVP$ cost                  |
|      | $2^{-9}$  | 290                       | $2^{77}$                                 |
|      | $2^{-13}$ | 280                       | $2^{74}$                                 |
|      | $2^{-14}$ | 276                       | $2^{73}$                                 |
| 153  | $2^{-18}$ | 269                       | $2^{71}$                                 |
| 155  | $2^{-20}$ | 264                       | $2^{70}$                                 |
|      | $2^{-24}$ | 258                       | $2^{68}$                                 |
|      | $2^{-30}$ | 250                       | $2^{66}$                                 |
| 165  | $2^{-35}$ | 244                       | $2^{64}$                                 |

Table 10: Predicted cost of our attack on the weak keys of BSS1 using the LWE Leaky estimator [DDGR20].

| Dire  | ect pri   | mal attack: $\lambda_Q$ : | $=\lambda_Q^* = 188$ , BKZ block size 710 |
|-------|-----------|---------------------------|---|
| $t_1$ | Prob.     | BKZ block size            |   |
| 266   |           | 692                       | $2^{183}$                                 |
| 270   |           | 682                       | $2^{180}$                                 |
| 276   |           | 668                       | $2^{177}$                                 |
|       | $2^{-12}$ | 653                       | $2^{173}$                                 |
| 288   | $2^{-15}$ | 640                       | $2^{169}$                                 |
| 296   | $2^{-22}$ | 622                       | $2^{164}$                                 |
| 300   | $2^{-26}$ | 614                       | $2^{162}$                                 |
| 304   |           | 605                       | $2^{160}$                                 |
| 308   | $2^{-35}$ | 597                       | $2^{158}$                                 |

Table 11: Predicted cost of our attack on the weak keys of BSS2 using the LWE Leaky estimator [DDGR20].

at random from the set  $\{-1, 0, 1\}$ , then the expected squared-norm of  $\psi_1(x)$  is upper-bounded by  $2n_{\theta}$  in the new basis representation.

*Proof.* Let a be any primitive element of the group U(n) (i.e., a generator of U(n)). Note that there are  $\phi(\phi(n))$  many such elements, where  $\phi(.)$  is Euler phi-function; we can pick any of those. Let k = (n-1)/3 and  $\theta = a^k \in U(n)$ . Then  $\theta$  is an element of order 3.

Note that the mapping

$$\boldsymbol{x}^{a^i} \rightarrow \boldsymbol{x}^{a^i} + \boldsymbol{x}^{a^i\theta} + \boldsymbol{x}^{a^i\theta^2}$$
 for  $0 \le i \le k-1$ 

is well-defined, since U(n) is a disjoint union of:

 $\{a^i \mid 0 \le i \le k-1\}, \quad \{a^i\theta \mid 0 \le i \le k-1\}, \text{ and } \{a^i\theta^2 \mid 0 \le i \le k-1\}.$ 

Hence, by linearity, the polynomial  $\psi_1(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{x}^{\theta}) + \boldsymbol{f}(\boldsymbol{x}^{\theta^2})$  is generated by the basis  $\{3, \boldsymbol{x}^{a^i} + \boldsymbol{x}^{a^i\theta} + \boldsymbol{x}^{a^i\theta^2}\}$  for  $0 \leq i \leq k-1$  of order  $n_{\theta} = 1+k$ , as claimed.

If individual coefficients of  $\mathbf{f}$  are uniformly sampled from  $\{-1, 0, 1\}$ , then sums of symmetric coefficients  $f_{a^i} + f_{a^i\theta} + f_{a^i\theta^2}$  are in  $\{-3, -2, -1, 0, 1, 2, 3\}$ and follow the probability distribution given in Table 12. So the coefficients of  $\psi_1(\mathbf{x})$  follow the distribution of  $f_{a^i} + f_{a^i\theta} + f_{a^i\theta^2}$ , except for the special coefficient associated to 3, which has a lower expectation of the square. Hence, by linearity of expectations, the expected squared-norm of  $\psi_1$  in the new basis representation is upper-bounded by  $2n_{\theta}$ .

| $f_{a^i} + f_{a^i\theta} + f_{a^i\theta^2}$ | 0    | 1    | -1   | 2    | -2   | 3    | -3   |
|---|------|------|------|------|------|------|------|
| Prob.                                       | 7/27 | 6/27 | 6/27 | 3/27 | 3/27 | 1/27 | 1/27 |

Table 12: Probability distribution of  $f_{a^i} + f_{a^i\theta} + f_{a^i\theta^2}$ 

However, this only gives us one polynomial  $\psi_1$  in reduced dimension, which is essentially the equivalent of  $\psi_+$  in the order 2 attack. We cannot directly construct an equivalent of  $\psi_-$ , so we use a different approach to get two other polynomials in reduced dimension.

Let us define  $f_2(x) = xf(x)$ ,  $f_3(x) = x^2f(x)$ , and  $\psi_2(x) = f_2(x) + f_2(x^{\theta}) + f_2(x^{\theta^2})$ ,  $\psi_3(x) = f_3(x) + f_3(x^{\theta}) + f_3(x^{\theta^2})$ . Then, if the coefficients of f(x) are sampled uniformly at random from the ternary set, each  $\psi_i$  has an expected squared-norm bounded by  $2n_{\theta}$  in the new basis representation. Indeed, the choice of g makes the coefficients of f,  $f_2$ , and  $f_3$  only different shifts in the polynomial basis representation. As a result, by linearity and from the distribution of the sums of symmetric coefficients, each  $\psi_i$  provides the same expected squared-norm in the new basis representation.

For the PV Knapsack problem, let  $\Omega_{3t_2}$  is the largest subset of  $\Omega_t$  that remains invariant under the transformation of  $\theta$ . In other words,  $\Omega_{3t_2}$  contains  $t_2$  triplets  $(\omega, \omega^{\theta}, \omega^{\theta^2})$  with all  $\omega$ ,  $\omega^{\theta}$ , and  $\omega^{\theta^2}$  in  $\Omega_t$ , where  $t_2 \leq \lfloor t/3 \rfloor$ . For any  $\omega$  in  $\Omega_{3t_2}$ , we know the evaluations of f at  $\omega$ ,  $\omega^{\theta}$ , and  $\omega^{\theta^2}$ . Hence we can compute  $t_2$  distinct evaluations of each of the polynomials  $\psi_i(\mathbf{x})$  at  $\omega \in \Omega_{3t_2}$ . This follows by writing  $\psi_2(\mathbf{x}) = \mathbf{x} f(\mathbf{x}) + \mathbf{x}^{\theta} f(\mathbf{x}^{\theta}) + \mathbf{x}^{\theta^2} f(\mathbf{x}^{\theta^2})$  and  $\psi_3(\mathbf{x}) = \mathbf{x}^2 f(\mathbf{x}) + \mathbf{x}^{2\theta} f(\mathbf{x}^{\theta}) + \mathbf{x}^{2\theta^2} f(\mathbf{x}^{\theta^2})$ .

This also allows to design a low-density inhomogeneous SIS problems to solve PV Knapsack problem. We create a matrix  $W_{\theta}$  with  $t_2$  rows and  $n_{\theta}$  columns, whose entries are the evaluations of each of the  $n_{\theta}$  monomials at an arbitrary choice of  $t_2$  representatives for the triplets in  $\Omega_{3t_2}$ . We create a vector  $\mathbf{b}_i$  whose coefficients are the known evaluations of  $\psi_i$  at each representative. We look for a solution of:

$$\boldsymbol{W}_{\boldsymbol{\theta}}\boldsymbol{\psi}_i = \boldsymbol{b}_i \bmod q. \tag{3}$$

Like before, we search for a short vector in the Kernel lattice

$$\mathcal{L}_{oldsymbol{W}_i',q}^{\perp} = \{oldsymbol{v} \in \mathbb{Z}^{n+1}: oldsymbol{W}_i'oldsymbol{v} = oldsymbol{0} mmod q\}$$

where  $\mathbf{W}'_i = (\mathbf{W}_{\theta}|\mathbf{b}_i)$ , and we expect  $(\boldsymbol{\psi}_i|-1)^T$  yields a uSVP solution. The knowledge of each  $\boldsymbol{\psi}_i$  gives  $n_{\theta}$  (independent) linear equations of the (unknown) coefficients of  $f(\boldsymbol{x})$ . So by doing linear algebra, we recover  $f(\boldsymbol{x})$ .

**Distribution of**  $t_2$ . The total number of sets  $\Omega_t$  of t elements chosen from the primitive roots is:

 $\binom{3\lfloor n/3\rfloor}{t}.$ 

When  $t_2$  is a fixed integer in  $[0, \lfloor t/3 \rfloor]$ , to choose a set of size t with exactly  $t_2$  triplets, we need to take  $t_2$  triplets from the set of  $\lfloor n/3 \rfloor$  triplets, followed by  $t - 3t_2$  non-triplets from the remaining triplets. Now, a non-triplet element can come as a combination of both pair and unpair. Thus, the total number of possibilities is:

$$\binom{\lfloor n/3 \rfloor}{t_2} \sum_{i=0}^{s} \binom{\lfloor n/3 \rfloor - t_2}{i} \binom{\lfloor n/3 \rfloor - t_2 - i}{t - 3t_2 - 2i} 3^{t - 3t_2 - i}$$

where  $s = \min\{\lfloor (t - 3t_2)/2 \rfloor, \lfloor n/3 \rfloor - t_2\}$ . So the probability of getting  $t_2$  for a random  $\Omega_t$  is:

$$\pi_2(t_2) = \frac{\binom{\lfloor n/3 \rfloor}{t_2} \sum_{i=0}^{s} \binom{\lfloor n/3 \rfloor - t_2}{i} \binom{\lfloor n/3 \rfloor - t_2 - i}{t_- 3t_2 - 2i} 3^{t-3t_2 - i}}{\binom{3 \lfloor n/3 \rfloor}{t}}.$$

**Comparison with symmetries of order 2.** We first would like to note that the worst-case keys, which are fully symmetric of higher order, clearly outperforms the order 2 symmetry attack. This is even clearer if one adversarially selects a key with symmetry of order 3 but no symmetry of order 2.

We keep this in mind; we now aim to compare the higher order symmetry with the order 2 symmetry for randomly selected keys. Let us start by comparing concrete examples of the attacks with symmetries of order 2 and 3.

For HPSSW2, when the value of  $t_2 = 42$ , we have  $\pi_2(t_2) = 2^{-26}$ . In this case, we recovered the secret in 111 hours (2<sup>49</sup>-bits operation) using BKZ block size 68. With  $\pi_1(t_1) = 2^{-26}$ , we get the value of  $t_1 = 92$ . In this case, we recovered the secret in 6.5 hours (2<sup>45</sup>-bits operation) using BKZ block size 58. Unfortunately, we never recovered the secret for a smaller value of  $t_2$  running lattice reductions for 7 days.

Similarly, we can do a comparison for HPSSW3, HPSSW4 by using the LWE estimator [APS15], it is shown in Figure 2.

On these three examples, it is clear that the order 2 attack performs better than the order 3 version. To understand why, let us consider a variant of the PV Knapsack problem, where the number of evaluation points t is close to pn, instead of n/2. Here p is an element in (0, 1).

In that case, the direct attack involves a lattice of dimension n+1 and volume  $q^{pn}$ . For the order 2 attack, we estimate the average of pairs to be  $p^2 n/2$ . As

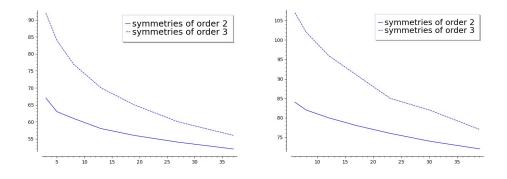


Fig. 2: Comparison: Predicted bits operation vs  $\pi_i(t_i)^{-1}$  (in log<sub>2</sub> scale) for the weak keys of HPSSW3 and HPSSW4.

a consequence, the attack involves a lattice of dimension  $\lceil n/2 \rceil + 1$  and volume  $q^{p^2 n/2}$ . For the order 3, we estimate the average of triplets to be  $p^3 n/3$  and get an attack involving dimension  $\lceil n/3 \rceil + 1$  and volume  $q^{p^3 n/3}$ . Ignoring constants, we can compare the root Hermite factors of the three attacks by looking at the three numbers:

$$p$$
,  $2p^2$ , and  $3p^3$ 

The case p = 1/2 that we previously considered is the crossover point between the direct attack and the order 2 symmetry attack. Similarly, the crossover point between the direct attack and the order 3 symmetry attack is  $p = 1/\sqrt{3} \approx 0.58$ . Finally, the crossover point between the order 2 and order 3 attacks is  $p = 2/3 \approx 0.67$ .

As a consequence, the higher symmetries only become worthwhile for random keys when the number of evaluation points in the PV Knapsack problem is much larger than what appears in practical parameters. We also see that by reducing the number of evaluation points below n/2, one can circumvent the gain provided by our main attack with symmetries of order 2.

Yet, since checking for symmetries is really fast, it cannot hurt a dedicated adversary to check their existence before launching the lattice reduction part of the attack.

## Acknowledgements

The main part of this work was done while Dipayan Das was affiliated with CISPA. This work has been supported by the European Union's H2020 Programme under grant agreement number ERC-669891.

## References

ADPS16. Erdem Alkim, Léo Ducas, Thomas Pöppelmann, and Peter Schwabe. Postquantum key exchange - A new hope. In Thorsten Holz and Stefan Savage, editors, 25th USENIX Security Symposium, USENIX Security 16, Austin, TX, USA, August 10-12, 2016, pages 327–343. USENIX Association, 2016.

- AGVW17. Martin R. Albrecht, Florian Göpfert, Fernando Virdia, and Thomas Wunderer. Revisiting the expected cost of solving uSVP and applications to LWE. In Tsuyoshi Takagi and Thomas Peyrin, editors, Advances in Cryptology ASIACRYPT 2017 23rd International Conference on the Theory and Applications of Cryptology and Information Security, Hong Kong, China, December 3-7, 2017, Proceedings, Part I, volume 10624 of Lecture Notes in Computer Science, pages 297–322. Springer, 2017.
- APS15. Martin R. Albrecht, Rachel Player, and Sam Scott. On the concrete hardness of learning with errors. J. Math. Cryptol., 9(3):169–203, 2015.
- BGP22. Katharina Boudgoust, Erell Gachon, and Alice Pellet-Mary. Some easy instances of Ideal-SVP and implications on the partial Vandermonde Knapsack problem. In Yevgeniy Dodis and Thomas Shrimpton, editors, Advances in Cryptology - CRYPTO 2022 - 42nd Annual International Cryptology Conference, Santa Barbara, CA, USA, August 15-18, 2022, Proceedings, Part II, volume 13508 of Lecture Notes in Computer Science, pages 480– 509. Springer, 2022.
- BSS22. Katharina Boudgoust, Amin Sakzad, and Ron Steinfeld. Vandermonde meets Regev: public key encryption schemes based on partial Vandermonde problems. *Des. Codes Cryptogr.*, 90(8):1899–1936, 2022.
- CN11. Yuanmi Chen and Phong Q. Nguyen. BKZ 2.0: Better lattice security estimates. In Dong Hoon Lee and Xiaoyun Wang, editors, Advances in Cryptology - ASIACRYPT 2011 - 17th International Conference on the Theory and Application of Cryptology and Information Security, Seoul, South Korea, December 4-8, 2011. Proceedings, volume 7073 of Lecture Notes in Computer Science, pages 1–20. Springer, 2011.
- DDGR20. Dana Dachman-Soled, Léo Ducas, Huijing Gong, and Mélissa Rossi. LWE with side information: Attacks and concrete security estimation. In Daniele Micciancio and Thomas Ristenpart, editors, Advances in Cryptology -CRYPTO 2020 - 40th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2020, Proceedings, Part II, volume 12171 of Lecture Notes in Computer Science, pages 329–358. Springer, 2020.
- DHSS20. Yarkin Doröz, Jeffrey Hoffstein, Joseph H. Silverman, and Berk Sunar. MMSAT: A scheme for multimessage multiuser signature aggregation. *IACR Cryptol. ePrint Arch.*, 2020.
- DJ23. Dipayan Das and Antoine Joux. On the hardness of the Finite Field Isomorphism problem. In Carmit Hazay and Martijn Stam, editors, Advances in Cryptology EUROCRYPT 2023 42nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Lyon, France, April 23-27, 2023, Proceedings, Part V, volume 14008 of Lecture Notes in Computer Science, pages 343–359. Springer, 2023.
- GN08. Nicolas Gama and Phong Q. Nguyen. Predicting lattice reduction. In Nigel P. Smart, editor, Advances in Cryptology - EUROCRYPT 2008, 27th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Istanbul, Turkey, April 13-17, 2008. Proceedings, volume 4965 of Lecture Notes in Computer Science, pages 31-51. Springer, 2008.
- HPS<sup>+</sup>14. Jeffrey Hoffstein, Jill Pipher, John M. Schanck, Joseph H. Silverman, and William Whyte. Practical signatures from the partial Fourier recovery

problem. In Ioana Boureanu, Philippe Owesarski, and Serge Vaudenay, editors, Applied Cryptography and Network Security - 12th International Conference, ACNS 2014, Lausanne, Switzerland, June 10-13, 2014. Proceedings, volume 8479 of Lecture Notes in Computer Science, pages 476–493. Springer, 2014.

- HS15. Jeffrey Hoffstein and Joseph H. Silverman. Pass-encrypt: a public key cryptosystem based on partial evaluation of polynomials. Des. Codes Cryptogr., 77(2-3):541–552, 2015.
- LLL82. Arjen K Lenstra, Hendrik Willem Lenstra, and László Lovász. Factoring polynomials with rational coefficients. *Mathematische annalen*, 261(ARTICLE):515–534, 1982.
- LSL13. Laura Luzzi, Damien Stehlé, and Cong Ling. Decoding by embedding: Correct decoding radius and DMT optimality. *IEEE Trans. Inf. Theory*, 59(5):2960–2973, 2013.
- LZA18. Xingye Lu, Zhenfei Zhang, and Man Ho Au. Practical signatures from the partial Fourier recovery problem revisited: A provably-secure and Gaussian-distributed construction. In Willy Susilo and Guomin Yang, editors, Information Security and Privacy - 23rd Australasian Conference, ACISP 2018, Wollongong, NSW, Australia, July 11-13, 2018, Proceedings, volume 10946 of Lecture Notes in Computer Science, pages 813–820. Springer, 2018.
- SE91. Claus-Peter Schnorr and M. Euchner. Lattice basis reduction: Improved practical algorithms and solving subset sum problems. In Lothar Budach, editor, Fundamentals of Computation Theory, 8th International Symposium, FCT '91, Gosen, Germany, September 9-13, 1991, Proceedings, volume 529 of Lecture Notes in Computer Science, pages 68–85. Springer, 1991.
- The23. The FPLLL Development Team. fplll, a lattice reduction library, Version: 5.4.5. Available at https://github.com/fplll/fplll, 2023.