Adapting Peters' ISD to the d-split SD problem in the context of SDitH

Kevin Carrier[1] and Jean-Pierre Tillich[2]

[1] ETIS Laboratory, CY Cergy-Paris University kevin.carrier@ensea.fr
[2] Project COSMIQ, Inria de Paris jean-pierre.tillich@inria.fr

In this short document, we analyze an adaptation of the Peters' ISD [Pet10] to the d-split SD problem in the context of SDitH[3]. We do not give all the details about the algorithm or its analysis here but we strongly encourage readers to refer to the cited documents to better understand the context and the formulas given below.

A key recovery attack on SDitH is as hard as solving the d-split Syndrome Decoding problem (SD(d)). Unfortunately, the specification document of SDitH does not take into account that for the regime of parameters which is chosen, SD(d) has actually several solutions: between 31 and 748 depending on the parameters. The security analysis provided in Section 7.1 of the SDitH submission ignores this point. In particular, their analysis is about finding one particular solution instead of any one among many and also, their reduction from SD(1) to SD(d) (see Theorem 6.1 and Appendix A of the supporting document[3]) is not tight in the case at hand. Indeed, the designers of SDitH use apparently the following inequality:

$$T_{SD(d)} \geq \frac{\binom{m/d}{w/d}^d}{\binom{m}{w}} T_{SD(1)} \tag{1}$$

where $T_{SD(d)}$ is the time complexity to solve the SD(d) problem. But when the number of solutions to the SD(d) problem is strictly greater than 0, the factor $\binom{m/d}{w/d}^d / \binom{m}{w}$ is underestimated. Note that in the present document, we do not use the above reduction; we prefer to give an actual attack.

Our adaption of Peters' ISD is quite simple: it consists in considering the splitting in the bet that is made on the searched error vector. In the case of Stern (and its Peters' version), one only has to choose the information set I as $I_1 \cup \cdots \cup I_d$ where I_i is of size $\frac{k}{d}$ and is randomly chosen in the i-th piece of the split positions. Then we address these sets separately and make a bet on each of them.

One iteration of Peters' ISD first consists in performing a Gaussian elimination whose cost only depends on the SD(d) parameters:

$$T_{\text{Gauss}} = \frac{q-1}{q}(m-k)\sum_{i=1}^{m-k}(m-i+2) = \frac{q-1}{2q}(m-k)^2(m+k+3) \tag{2}$$

Then two lists are built and merged with the aim of producing candidates that will then be verified. The sizes of those lists depend on d and are worth:

$$L_1 = \tbinom{\lfloor k/2 \rfloor}{p}(q-1)^p \quad \text{and} \quad L_2 = \tbinom{k-\lfloor k/2 \rfloor}{p}(q-1)^p \qquad \text{if } d=1 \tag{3}$$

$$L_1 = \tbinom{\lfloor k/4 \rfloor}{p/2}\tbinom{\lfloor k/2 \rfloor - \lfloor k/4 \rfloor}{p/2}(q-1)^p \quad \text{and} \quad L_2 = \tbinom{\lfloor k/2 \rfloor - \lfloor k/4 \rfloor}{p/2}\tbinom{k-2\lfloor k/2 \rfloor + \lfloor k/4 \rfloor}{p/2}(q-1)^p \qquad \text{if } d=2 \tag{4}$$

The cost to produce the lists is

$$T_{\text{lists}} = \left(\left(\tfrac{k}{2} - p + 1\right) + L_1 + L_2\right)\ell \tag{5}$$

and the cost to check all the potential solutions given by the collision search is

$$T_{\text{check}} = \frac{q}{q-1}(w-2p+1)2p\left(1+\frac{q-2}{q-1}\right)\frac{L_1 L_2}{q^\ell} \tag{6}$$

On the other hand, the average number of needed iterations is $\frac{1}{P_{\text{succ}} \cdot N_{\text{sols}}}$ where P_{succ} is the probability of finding one particular solution and N_{sols} is the expected number of solutions. Both P_{succ} and N_{sols} depend on d too:

$$P_{\text{succ}} = \frac{L_1 L_2 \binom{m-k-\ell}{w-2p}}{\binom{m}{w}(q-1)^{2p}} \quad \text{and} \quad N_{\text{sols}} = 1 + \frac{\binom{m}{w}}{q^{m-k}} \qquad \text{if } d=1 \tag{7}$$

$$P_{\text{succ}} = \frac{L_1 L_2 \binom{\lfloor (m-k-\ell)/2 \rfloor}{(w-2p)/2}\binom{m-k-\ell-\lfloor (m-k-\ell)/2 \rfloor}{(w-2p)/2}}{\binom{m/2}{w/2}^2 (q-1)^{2p}} \quad \text{and} \quad N_{\text{sols}} = 1 + \frac{\binom{m/2}{w/2}^2}{q^{m-k}} \qquad \text{if } d=2 \tag{8}$$

Finally, the time complexity of the d-split version of the Peters' ISD is

$$T_{\text{SD}(d)}(p,\ell) = \left(\frac{T_{\text{Gauss}} + T_{\text{lists}} + T_{\text{check}}}{P_{\text{succ}} \cdot N_{\text{sols}}}\right) \cdot \log_2(q) \tag{9}$$

An exhaustive search allows to find the parameters p and ℓ which optimize the complexity $T_{\text{SD}(d)}(p,\ell)$. The following table recalls the results given in the specification document of SDitH[3] and gives the actual security level we estimate for the various SDitH parameter sets:

Parameter sets	**d-split SD parameters**					**The results in the specification document of SDitH[3]**					**Our d-split version of Peters' ISD**		
	q	m	k	w	d	p	ℓ	$T_{\text{SD}(1)}$	$\frac{\binom{m/d}{w/d}^d}{\binom{m}{w}} T_{\text{SD}(1)}$	target security	p	ℓ	$T_{\text{SD(d)}}$
`SDitH_L1_gf256`	256	230	126	79	1	1	2	143.46 bits	143.46 bits	143 bits	1	2	**134.61 bits**
`SDitH_L1_gf251`	251	230	126	79	1	1	2	143.45 bits	143.45 bits	143 bits	1	2	**133.90 bits**
`SDitH_L3_gf256`	256	352	193	120	2	2	5	211.15 bits	207.67 bits	207 bits	2	5	**206.16 bits**
`SDitH_L3_gf251`	251	352	193	120	2	2	5	211.09 bits	207.61 bits	207 bits	2	5	**205.02 bits**
`SDitH_L5_gf256`	256	480	278	150	2	2	5	276.02 bits	272.35 bits	272 bits	2	5	**271.30 bits**
`SDitH_L5_gf251`	251	480	278	150	2	2	5	275.96 bits	272.29 bits	272 bits	2	5	**269.81 bits**

References

Pet10. Christiane Peters. Information-set decoding for linear codes over $\mathbb{F}_q$. In *Post-Quantum Cryptography 2010*, volume 6061 of *LNCS*, pages 81–94. Springer, 2010.

[3] `https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/round-1/spec-files/SDitH-spec-web.pdf`