Integer Syndrome Decoding in the Presence of Noise

Vlad Dragoi, Brice Colombier, Pierre-Louis Cayrel, Vincent Grosso

Integer Syndrome Decoding in the Presence of Noise

Vlad-Florin Drăgoi
Faculty of Exact Sciences
Aurel Vlaicu University
Arad, Romania
vlad.dragoi@uav.ro

Brice Colombier
Univ Grenoble Alpes, CNRS
Grenoble INP, TIMA
38000 Grenoble, France
brice.colombier@grenoble-inp.fr

Pierre-Louis Cayrel and Vincent Grosso
Univ Lyon, UJM-Saint-Etienne, CNRS
Laboratoire Hubert Curien UMR 5516
F-42023, Saint-Etienne, France
{pierre.louis.cayrel; vincent.grosso}@univ-st-etienne.fr

***Abstract*—Code-based cryptography received attention after the NIST started the post-quantum cryptography standardization process in 2016. A central NP-hard problem is the binary syndrome decoding problem, on which the security of many code-based cryptosystems lies. The best known methods to solve this problem all stem from the information-set decoding strategy. A recent line of work considers augmented versions of this strategy, with hints provided by side-channel information. In this work, we consider the integer syndrome decoding problem, where the integer syndrome is available but might be noisy. We study how the performance of the decoder is affected by the noise. We provide experimental results on cryptographic parameters for the Classic McEliece and BIKE cryptosystems, which are in the fourth round of the NIST standardization process.**

***Index Terms*—Code-based cryptography, Syndrome decoding problem, Information-set decoding**

I. Introduction

With an increasing practical feasibility of a quantum computer, the threat posed by Shor's algorithm [1] on number theory base cryptosystems grows significantly. To address this threat, NIST began a standardization process in 2016 for post-quantum cryptography. In July 2022, three code-based candidates (McEliece [2], BIKE [3], HQC [4]) qualified for the fourth round.Their security relies on the NP-hardness of the binary syndrome decoding problem (SDP) [5]. Given a parity-check matrix $\boldsymbol{H}$ of a binary linear code, a binary syndrome vector $\boldsymbol{s}^*$ and an integer t, the SDP consists in finding a binary vector $\boldsymbol{x}$ of Hamming weight t such that $\boldsymbol{Hx} = \boldsymbol{s}^*$. The best known strategy to solve the SDP is referred to as *information-set decoding* (ISD). Originally proposed by Prange [6], it has been incrementally refined since [7], [8], [9], [10], [11], [12], [13].The complexity of the ISD method has been used to better tune the parameters of the cryptosystems according to the required security levels [14].

a) Syndrome decoding with hints: Recently, modified versions of the SDP, with additional information, obtained via side-channel analysis, was considered. In [15], authors study the case where parts of the error are known, or only their Hamming weight. The case where the *integer* syndrome $\boldsymbol{s}$ is available, as if the matrix-vector multiplication had been performed in the integer ring instead of the binary finite field, is considered in [16]. One method to obtain the integer syndrome is by laser fault injection attack [17]. The problem one has to solve in this case is the integer syndrome decoding ($\mathbb{N} - \mathsf{SDP}$), where the input is the parity-check matrix $\boldsymbol{H}$, the integer syndrome vector $\boldsymbol{s}$ and the weight of the solution t. The same question is raised, does $\boldsymbol{Hx} = \boldsymbol{s}$ admits a solution of weight t? This problem can be tackled down by means of Integer Linear Programming [17] or probabilistic methods [18]. Another method of obtaining $\boldsymbol{s}$, much more feasible and realistic than laser fault injection, is by side-channel analysis [19]. However, due to physical factors, the estimation of $\boldsymbol{s}$ might not be perfectly accurate. Hence, in the $\mathbb{N} - \mathsf{SDP}$ in the presence of noise, we are given a noisy integer syndrome $\widetilde{\boldsymbol{s}} = \boldsymbol{s} + \epsilon$, where ϵ models the noise. The solution proposed in [19] uses a combination of ISD techniques and the score decoder from [18]. In [19] only simulations are provided to assess the performance of this algorithm.

b) Learning with errors and hints: Not only code-based cryptosystems are vulnerable to such attacks. Similar results were obtained in the context of lattice-based cryptosystems by Bootle et al. [20]. The BLISS cryptosystem was cryptanalysed by means of similar hybrid attacks, where side-channel attacks revealed an Integer version of the Learning with Errors (ILWE). The ILWE problem is the lattice-based equivalent of the N-SDP. However, ILWE was solved with another technique that does not seem to work for N-SDP. Nevertheless, it points out that such scenarios are not only applicable to code-based cryptosystems.

c) Quantitative Group Testing: Quantitative Group Testing (QGT) is an active field of research, lately boosted by the the COVID-19 epidemic. In the QGT we are given a large population out of which some individuals suffer from a disease, and the goal is to identify the infected individuals. Possible applications of QGT go from bio-informatics [21], traffic monitoring [22] and confidential data transfer [23], [24] to machine learning [25], [26]. The N-SDP can be also seen as a QGT in presence of noise. As we shall demonstrate, the algorithm we propose here, solves a noisy QGT instance, by adapting and improving (using coding theory tools, such as

V.-F. Drăgoi was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI-UEFISCDI, project number PN-III-P1-1.1-PD-2019-0285, within PNCDI III.

B. Colombier was supported by the French National Research Agency in the framework of the "Investissements d'avenir" program "ANR-15-IDEX-02" and the LabEx PERSYVAL "ANR-11-LABX-0025-01".

ISD techniques) a recent solution to the classical QGT [18].

Contributions: In this article, we analyze in detail the algorithm proposed in [19], referred to as ISD-score decoder, and provide the following contributions. First, we demonstrate that the ISD-score decoder finds a solution to the $\mathbb{N}-\mathsf{SDP}$ in the presence of noise with high probability, as long as the weight is sufficiently sub-linear in n, more exactly, $t \leq O\left(\frac{n-k}{\log(n-k)}\right)$, where n is the code length and k the dimension. We consider two noise models, *i.e.*, Binomial centered in zero and Bernoulli variables. The ISD-score decoder can tolerate noise levels that are linear in the weight of the solution t. For that we partially build our demonstration on the techniques used in [18]. We incorporate the noise models into these techniques and, by using sharper inequalities, determine a much clearer condition for having a higher probability of success. One consequence of this new method is that when the noise is null and the ISD part is ignored, equivalently the ISD-score decoder boils down to the algorithm proposed in [18], the conditions we obatain for t are larger that those from [18]. This gives a lower bound on the number of syndrome entries required to find a solution, known as the information theoretic bound.

Due to page limitations, all proofs have been omitted, the full version of the article being available on `eprint.iacr`.

II. Preliminaries

Notations: A finite field is denoted by $\mathbb{F}$, and the ring of integers by $\mathbb{Z}$. We write $\mathbb{N}_n^* = \{1,\dots,n\}$ and $\mathbb{Z}_{-n,n} = \{-n,\dots,0,\dots,n\}$. For $p \in [0,1]$ and $n \in \mathbb{N}$ we denote the Bernoulli distribution by $\mathcal{B}er(p)$ and the Binomial distribution by $\mathcal{B}(n,p)$. We denote by $W(x)$ the Lambert W function. Matrices and vectors are written in bold capital, respectively small letters. Also, $\mathrm{HW}(\boldsymbol{c})$ and $\mathsf{Supp}(\boldsymbol{c})$ denotes the Hamming weight and the support of the vector $\boldsymbol{c}$.

Integer syndrome decoding problem: Let us recall the formal definition of $\mathbb{N}-\mathsf{SDP}$ as already stated in [17], [16].

Definition 1 ($\mathbb{N}-\mathsf{SDP}$)**.**
Inputs: $\boldsymbol{H} \in \{0,1\}^{(n-k)\times n}$, $\boldsymbol{s} \in \mathbb{N}^{n-k}$, $t \in \mathbb{N}^*$.
Output: $\boldsymbol{x} \in \{0,1\}^n$, *s.t.* $\boldsymbol{Hx} = \boldsymbol{s}$, *and* $HW(\boldsymbol{x}) = t$.

To define $\mathbb{N}-\mathsf{SDP}$ in the presence of noise as generally as possible, we model the noise $\epsilon = (\epsilon_1,\dots,\epsilon_{n-k})$ as a vector of random variables $\epsilon_i \sim \mathcal{D}$, where $\mathcal{D}$ is a discrete probability distribution. In the $\mathbb{N}-\mathsf{SDP}$ in the presence of noise we are given a noisy syndrome $\widetilde{\boldsymbol{s}} = \boldsymbol{s} + \epsilon$ and the value $\boldsymbol{s}^* = \boldsymbol{s} \pmod 2$ (component-wise).

Definition 2 ($\mathbb{N}-\mathsf{SDP}$ in the presence of noise ϵ)**.**
Inputs: $\boldsymbol{H} \in \{0,1\}^{(n-k)\times n}$, $\widetilde{\boldsymbol{s}} \in \mathbb{Z}^{n-k}$
$\boldsymbol{s}^* \in \{0,1\}^{n-k}$, $t \in \mathbb{N}^*$
Output: $\boldsymbol{x} \in \{0,1\}^n$, *s.t.* $\boldsymbol{Hx} = \boldsymbol{s}^\star$ *with* $HW(\boldsymbol{x}) = t$
$\boldsymbol{s}^\star = \boldsymbol{s} \mod 2$, *and* $\widetilde{\boldsymbol{s}} = \boldsymbol{s} + \epsilon$.

Remark that $\mathbb{N}-\mathsf{SDP}$ in presence of noise is the SDP with additional information. Under certain conditions, we hope that, given $(\boldsymbol{H}, \boldsymbol{s}^*, t, \widetilde{\boldsymbol{s}})$, we can find $\boldsymbol{x}$, solution to the SDP. Also, when the noise is zero we face the classic $\mathbb{N}-\mathsf{SDP}$.

III. ISD-Score decoder

A. Score decoder

The idea of assigning a score to each column was already used for the $\mathbb{N}-\mathsf{SDP}$ in [19]. The goal is to distinguish columns of $\boldsymbol{H}$ in $\mathsf{Supp}(\boldsymbol{x})$ from columns outside $\mathsf{Supp}(\boldsymbol{x})$, where $\boldsymbol{x}$ is a solution. We shall begin by defining a particular efficient score decoder, introduced by [18]. For a better illustration of the nice features of the decoder in the presence of noise, we will express it in function of the noiseless decoder. As we shall see, this method allows not only to derive a particularly simple relation between those two, but also to deduce conditions on the tolerated noise level.

Definition 3. *Let* $\boldsymbol{H} \in \{0,1\}^{(n-k)\times n}, \boldsymbol{s} \in \mathbb{N}^{n-k}$ *and* $t \in \mathbb{Z}^*$ *be the input of* $\mathbb{N}-\mathsf{SDP}$. *Then define the score of a column:* $\forall i \in \mathbb{N}_n^* \quad \psi_i(\boldsymbol{s}) = \sum_{\ell=1}^{n-k}\left(h_{\ell,i}\boldsymbol{s}_\ell + (1-h_{\ell,i})(t-\boldsymbol{s}_\ell)\right))$.

ψ can be seen as a correlation measure between the columns of $\boldsymbol{H}$ and the syndrome, and thus, it reveals those columns that contributed to the syndrome. For the $\mathbb{N}-\mathsf{SDP}$ in the presence of noise we shall use $\psi_i(\widetilde{\boldsymbol{s}})$. The next result, rephrased from [18], expresses the capability of the score decoder to distinguish between columns in the support of the solution vector from columns which are outside the support.

Theorem 1. *Let* $\boldsymbol{H} \in \{0,1\}^{(n-k)\times n}$ *be a random matrix, with* $h_{j,i}$ *independent variables s.t.* $h_{j,i} \sim \mathcal{B}er(\frac{1}{2})$ *and* $\boldsymbol{s} \in \mathbb{N}^{n-k}$ *such that* $\exists\ \boldsymbol{x} \in \{0,1\}^n$ *with* $HW(\boldsymbol{x}) = t$ *satisfying* $\boldsymbol{Hx} = \boldsymbol{s}$. *Then* $\psi_i(\boldsymbol{s})$ *follows the distribution* $\begin{cases} \mathcal{B}((n-k)t, \frac{1}{2}) & , i \notin \mathsf{Supp}(\boldsymbol{x}) \\ \mathcal{B}((n-k)(t-1), \frac{1}{2}) + n - k & , i \in \mathsf{Supp}(\boldsymbol{x}) \end{cases}$.

From Thm. 1, we deduce that for $i \notin \mathsf{Supp}(\boldsymbol{x})$ we have $\mathbb{E}(\psi_i(\boldsymbol{s})) = \frac{(n-k)t}{2}$, while for $i \in \mathsf{Supp}(\boldsymbol{x})$ we have $\mathbb{E}(\psi_i(\boldsymbol{s})) = \frac{(n-k)t}{2} + \frac{n-k}{2}$. This difference in the mean points out that ψ can be a distinguisher.In addition, the variance also differs, a fact that will be used in the tail bounds.

B. Score decoder in the presence of noise

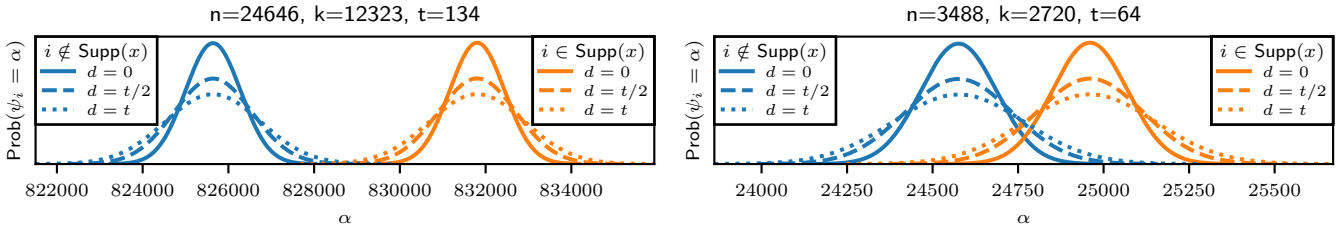
Here, we assume that ϵ_i are i.i.d. random variables, the noise does not depend on the distribution of the entries in $\boldsymbol{H}$ and the distribution $\mathcal{D}$ is symmetric.

Proposition 1 ([19])**.** *For* $j \in \mathbb{Z}_{n-k}^*$ *let* ϵ_j *be i.i.d. discrete random variables following a symmetric distribution over the set* $\mathbb{Z}_{-d,d}$, *s.t.* ϵ_j *and* $h_{i,j}$ *are independent. Then* $\mathsf{Prob}\left(\psi_i(\widetilde{\boldsymbol{s}}) - \psi_i(\boldsymbol{s}) = \alpha\right) = \mathsf{Prob}\left(\sum_{j=1}^{n-k}\epsilon_j = \alpha\right)$.

Remark 1. *To keep* $\psi_i(\widetilde{\boldsymbol{s}}) - \psi_i(\boldsymbol{s})$ *as small as possible, it is not necessary that the* ϵ_i *values are small. Indeed, the* ϵ_i *values may be large, as long as* $\sum_{i=1}^{n-k}\epsilon_i$ *is close to zero.*

For a centered binomial noise, we deduce the following.

Corollary 1. *Let* $d \in \mathbb{N}$. *If* $\epsilon_i \sim -d + \mathcal{B}(2d, \frac{1}{2})$ *then* $\psi_i(\widetilde{\boldsymbol{s}})$ *is a random variable that follows the distribution* $\begin{cases} \mathcal{B}((n-k)(t+2d), \frac{1}{2}) - d(n-k) & i \notin \mathsf{Supp}(\boldsymbol{x}) \\ \mathcal{B}((n-k)(t-1+2d), \frac{1}{2}) - (d-1)(n-k) & i \in \mathsf{Supp}(\boldsymbol{x}) \end{cases}$



(a) Example parameters set of BIKE

(b) Example parameters set of Classic McEliece

Fig. 1: Distribution of ψ_i for $\epsilon \sim -d + \mathcal{B}(2d, \frac{1}{2})$

Moreover, $\mathbb{E}(\psi_i(\widetilde{\boldsymbol{s}})) = \mathbb{E}(\psi_i(\boldsymbol{s}))$ *and* $Var(\psi_i(\widetilde{\boldsymbol{s}})) = Var(\psi_i(\boldsymbol{s})) + (n-k)d/2$.

To maintain the capability to distinguish between positions inside the support and positions outside the support, the noise parameter d from $\mathcal{B}(2d, \frac{1}{2})$ should be restricted.

Corollary 2. *Let* $\epsilon_i \sim -d + \mathcal{B}(2d, \frac{1}{2})$ *and* $g(n,k,t)$ *a positive unbounded function. Then* $\mathsf{Prob}\left(|\psi_i(\widetilde{\boldsymbol{s}}) - \psi_i(\boldsymbol{s})| \leq \sqrt{\frac{d(n-k)g(n,k,y)}{2}}\right) \geq 1 - O(\frac{1}{g(n,k,t)})$. *Moreover, for any* $d \leq \frac{n-k}{8g(n,k,t)}$, $\psi(\widetilde{\boldsymbol{s}})$ *distinguishes positions in* $\mathsf{Supp}(\boldsymbol{x})$ *from positions outside* $\mathsf{Supp}(\boldsymbol{x})$ *w.h.p.*

Fig. 1 shows the distribution of ψ_i values for different levels of noise, ranging from $d = 0$, *i.e.* the noiseless setting, to a very high noise of $\mathcal{B}(2t, \frac{1}{2})$. Notice that the distinguishing capability is much higher for the BIKE parameters [3], as shown in Fig. 1a, than for the Classic McEliece parameters [2], as shown in Fig. 1b.

C. Combining ISD and score decoder

The idea in [19] was to boost the distinguishing capability of the score decoder with ISD-like techniques. To this end, the score decoder is integrated in the "permutation" step of the ISD method. Indeed, this method starts by performing a permutation on the columns of $\boldsymbol{H}$ that will hopefully rearrange the solution in a useful way. More precisely, in the first ISD algorithm, the Prange decoder [6], a "good" permutation ($\boldsymbol{\Pi}$) is one that satisfies $\boldsymbol{\Pi}^{-1}\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{0} \end{pmatrix}$. Hence, the initial system becomes $\boldsymbol{H}\boldsymbol{\Pi}\boldsymbol{\Pi}^{-1}\boldsymbol{x} = \boldsymbol{s}^*$. By Gaussian elimination on $\boldsymbol{H}\boldsymbol{\Pi}$ one can find an invertible matrix $\boldsymbol{A}$ s.t. $\boldsymbol{A}\boldsymbol{H}\boldsymbol{\Pi} = (\boldsymbol{I} \| \boldsymbol{B})$. Hence, the system becomes $(\boldsymbol{I} \| \boldsymbol{B}) \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{0} \end{pmatrix} = \boldsymbol{A}\boldsymbol{s}^*$ which yields $\boldsymbol{x}_1 = \boldsymbol{A}\boldsymbol{s}^*$. In the original ISD methods, permutations are sampled randomly until a "good" one is obtained. Thanks to the extra-information provided by $\boldsymbol{s}$ or $\widetilde{\boldsymbol{s}}$, the function ψ allows to construct a permutation which by no means is random. Indeed, we have seen that ψ, by its nature, allows one to distinguish between positions in $\mathsf{Supp}(\boldsymbol{x})$ and positions outside. Hence, the underlying permutation, hopefully is a "good" permutation. As pointed out in [19], sorting the list of values $\psi_i(\widetilde{\boldsymbol{s}})$ in descending order is equivalent to generating a permutation $\boldsymbol{\Pi}$. Algorithm 1 finds a solution to the $\mathbb{N}-\mathsf{SDP}$ in the presence of noise as long as $\boldsymbol{\Pi}$ is "good" enough.

Algorithm 1 PRANGE SCORE DECODER($\boldsymbol{H}, \boldsymbol{s}, t$)

1: Compute $\boldsymbol{\Pi}$ from the list $\psi_i(\widetilde{\boldsymbol{s}})$
2: Compute $\boldsymbol{A}^*, \boldsymbol{H}^* \leftarrow \mathtt{rref}(\boldsymbol{H}\boldsymbol{\Pi})$
3: **if** $\mathrm{HW}(\boldsymbol{A}^*\boldsymbol{s}^\star) = t$ **then**
4: **return** $\boldsymbol{x} = \boldsymbol{\Pi}\begin{pmatrix} \boldsymbol{A}^*\boldsymbol{s}^\star \\ \boldsymbol{0}_{n-r} \end{pmatrix}$ ▷ $r = \mathtt{rank}(\boldsymbol{A})$

The procedure $\mathtt{rref}(\boldsymbol{H}\boldsymbol{\Pi})$, which stands for "reduced row echelon form", is equivalent to performing a partial Gaussian elimination over $\mathbb{F}_2$. Indeed, there is an $(n-k)\times(n-k)$ non-singular matrix $\boldsymbol{A}^*$ such that, $\boldsymbol{A}^*\boldsymbol{H}\boldsymbol{\Pi} = \begin{bmatrix} \boldsymbol{I_r} \\ \boldsymbol{0_{n-k-r,r}} \end{bmatrix} \| \boldsymbol{B}^*$ where $\boldsymbol{H}\boldsymbol{\Pi} = [\boldsymbol{A} \| \boldsymbol{B}]$ with $\boldsymbol{A}$ a $(n-k)\times r$ matrix satisfying $\boldsymbol{A}^*\boldsymbol{A} = \begin{bmatrix} \boldsymbol{I_r} \\ \boldsymbol{0_{n-k-r,r}} \end{bmatrix}$, and $\boldsymbol{B}^* = \boldsymbol{A}^*\boldsymbol{B}$. In the case of a full rank matrix $\boldsymbol{A}$ we have $\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{I_{n-k}}$. From the description of the algorithm above, the following result can be deduced.

Proposition 2 ([19])**.** PRANGE SCORE DECODER *outputs a valid solution as long as there exists at least one set* $L \subset \mathbb{N}_n^* \setminus \mathsf{Supp}(\boldsymbol{x})$ *with* $\#L \geq n - r$ *such that* $\min\{\psi_i(\widetilde{\boldsymbol{s}}), i \in \mathsf{Supp}(\boldsymbol{x})\} > \max\{\psi_i(\boldsymbol{x}), i \in L\}$.

The overall time complexity of PRANGE SCORE DECODER is $O((n-k)^3)$, since it is dominated by the partial Gaussian elimination, *i.e.* the computation of $\boldsymbol{A}^*$. Since the permutation $\boldsymbol{\Pi}$ might not move all the positions in $\mathsf{Supp}(\boldsymbol{x})$ in the first $n-k$ positions, more powerful ISD methods may be used, *e.g.* Lee-Brickell [7], Stern [8] or Dumer [9]. The idea is to allow a number of δ positions from $\mathsf{Supp}(\boldsymbol{x})$ outside the first $n-k$ positions. This is equivalent to extending PRANGE SCORE DECODER so that it covers error vectors with a more general pattern. The Lee-Brickell score decoder, where δ positions are searched exhaustively, is thus proposed in [19] as a possible solution. When the Lee-Brickell variant is used and $\delta = \mathcal{O}(1), k = \mathcal{O}(n)$, the work factor of the resulting algorithm becomes polynomial in n. An algorithm allowing δ out of t positions from $\mathsf{Supp}(\boldsymbol{x})$ in the last $n-r$ positions will be further called δ-ISD-score decoder or simply ISD-score decoder.

Proposition 3. *The δ-ISD-score decoder outputs a solution as long as $\exists J \subset \mathbb{N}_n$ of cardinality $n-k$ and at most δ indices $i \in \mathsf{Supp}(\boldsymbol{x})$ with values $\psi_i(\widetilde{\boldsymbol{s}}) < \psi_j(\widetilde{\boldsymbol{s}})$ with $j \in J$.*

IV. Success probability of the ISD-Score decoder

A. Main results

The following result gives a condition on the parameters for having a high probability of success for the ISD score decoder on the $\mathbb{N}-\mathsf{SDP}$ in presence of noise.

Theorem 2. *Let $\epsilon_i \sim -d + \mathcal{B}(2d, \frac{1}{2})$. If the interval $\left[\sqrt{\frac{t+2d}{n-k} W\left(\frac{n-t}{n-k-t+\delta+1}\frac{e\sqrt{2}}{\pi}\right)^2}, 1-\sqrt{\frac{t+2d-1}{n-k} W\left(\frac{t}{\delta+1}\frac{2e}{\pi}\right)^2}\right]$ is non-empty, then w.h.p. the ISD-score decoder succeeds in finding a solution.*

To give a more sensitive meaning of our result, we can approximate the value of the Lambert W function by $W(m) = \log m - \log\log m + O(\frac{\log\log m}{\log m})$ when m tends to infinity. Using only the first term we define $I_\beta = \left[\sqrt{\frac{2(t+2d)}{n-k}\log\frac{n-t}{n-k-t+\delta+1}}, 1-\sqrt{\frac{2(t+2d-1)}{n-k}\log\frac{t}{\delta+1}}\right]$. Hence, we deduce the following result.

Proposition 4. *Let $m \to \infty$. If $I_\beta \neq \emptyset$ then the probability of success of the ISD-score decoder is at least*

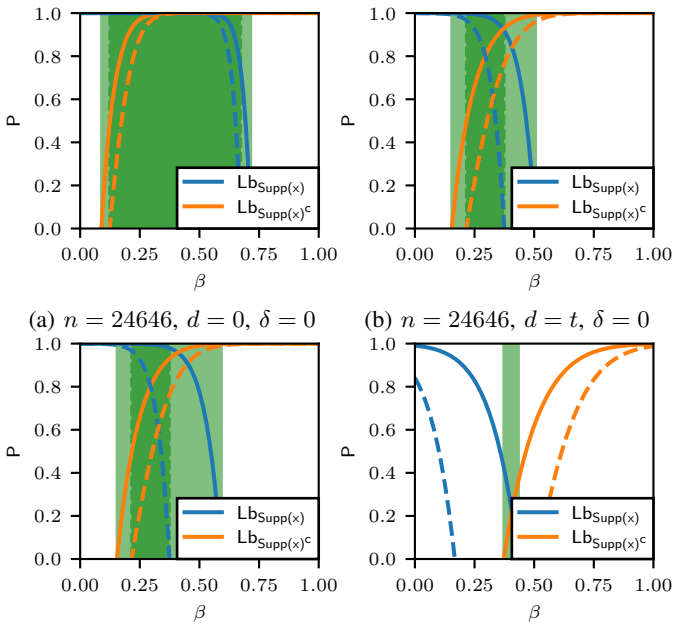
$$\left(1-\frac{e}{2\pi}\frac{1}{\sqrt{\log\frac{n-t}{n-k-t+\delta+1}}}\right)\left(1-\frac{e}{\sqrt{2}\pi}\frac{1}{\sqrt{\log\frac{t}{\delta+1}}}\right).$$

The construction of the interval in Thm. 2 comes from the underlying proof, where two functions depending on a parameter $\beta \in [0,1]$ ($\mathrm{Lb}_{\mathsf{Supp}(\boldsymbol{x})}$, $\mathrm{Lb}_{\mathsf{Supp}(\boldsymbol{x})^c}$) are required to have co-domain $[0,1]$. These two functions are $\mathrm{Lb}_{\mathsf{Supp}(\boldsymbol{x})^c} = 1 - \frac{e(n-t)}{\sqrt{2}\pi\beta(n-k-t+\delta+1)}\sqrt{\frac{t+2d}{n-k}}e^{-\frac{(n-k)\beta^2}{2(t+2d)}}$, $\mathrm{Lb}_{\mathsf{Supp}(\boldsymbol{x})} = 1 - \frac{e.t}{\pi(1-\beta)(\delta+1)}\sqrt{\frac{t+2d-1}{n-k}}e^{-\frac{(n-k)(1-\beta)^2}{2(t+2d-1)}}$.

To fairly compare with state-of-the-art techniques such as the algorithm in [18], which is only valid for the noiseless scenario, we adapted the conditions from [18] to the noise model considered here. This gives two similar functions in β, namely $1-\frac{n-t}{n-k-t}\sqrt{\frac{t+2d}{n-k}}e^{-\frac{(n-k)\beta^2}{2(t+2d)}}$, and $1-t\sqrt{\frac{t+2d-1}{n-k}}e^{-\frac{(n-k)(1-\beta)^2}{2(t+2d-1)}}$. In Fig. 2, we plot the modified functions from [18] (dashed lines) and $\mathrm{Lb}_{\mathsf{Supp}(\boldsymbol{x})}, \mathrm{Lb}_{\mathsf{Supp}(\boldsymbol{x})^c}$ (solid lines). In dark green and light green, the valid interval/region for the adapted functions from [18], and our functions, respectively, is represented. Notice that for all parameter sets and all noise levels considered here, our function offers a larger interval. Hence, this implies that for some sets of parameters, *e.g.,* in Fig. 2d, the interval is empty w.r.t. conditions in [18], while w.r.t. our conditions the interval exists.

B. Information-theoretic bounds

1) Bounding the value of t: To see how large t must be the following estimate can be used.



(a) $n = 24646$, $d = 0$, $\delta = 0$ (b) $n = 24646$, $d = t$, $\delta = 0$

(c) $n = 24646$, $d = t$, $\delta = 3$ (d) $n = 3488$, $d = t$, $\delta = 3$

Fig. 2: Valid β interval from the bounds in [18] (dashed lines) and the proposed ones (solid lines)

Theorem 3 (Upper bound on t)**.** *Let $k \leq n-t+\delta+1-(n-t)(\delta+1)/t$ and $d = ct/2$. Then $I_\beta \neq \emptyset$ as long as we have*

$$t \leq \frac{n-k}{8(1+c)W\left(\frac{n-k}{8(1+c)(\delta+1)}\right)} \tag{1}$$

Moreover, when $n \to \infty$, we have that $t \leq O\left(\frac{n-k}{\log(n-k)}\right)$.

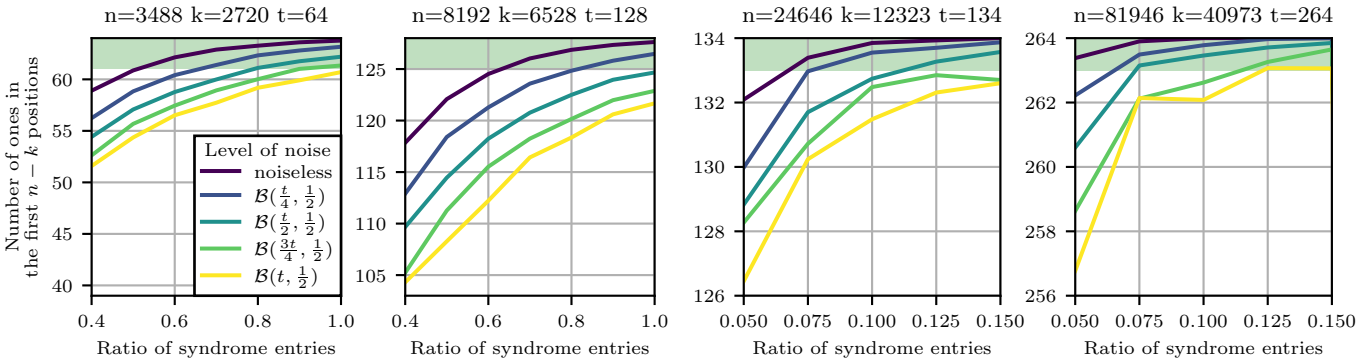
Using the first term approximation for the Lambert W function near infinity we have $t \leq \frac{n-k}{8(1+c)\log\frac{n-k}{8(1+c)(\delta+1)}}$.

Recall that a preliminary condition on d was determined, more exactly, $d \leq \frac{n-k}{8g(n,k,t)}$, where $g(n,k,t)$ is a positive unbounded function. Taking $d = \frac{n-k}{8\log(n-k)} \leq \frac{n-k}{8\log\log(n-k)}$ validates the choice in the hypothesis $d = ct/2$, as per Thm. 3 $t \leq O\left(\frac{n-k}{\log(n-k)}\right)$. Taking into account this condition and the hypothesis of Thm. 3, *i.e.* $d = ct/2$, we deduce the following upper bound on t

$$d = \frac{ct}{2} \leq \frac{n-k}{8\log t} \quad \Rightarrow \quad t\log t \leq \frac{n-k}{4c}. \tag{2}$$

This improves the constant term by $t \leq \frac{n-k}{4cW(\frac{n-k}{4c})}$.

Remark 2. *Notice that we cannot decrease the non-constant factors lower than what we have achieved here. More exactly we need to have at least $t\log\frac{t}{\delta+1} \leq \frac{n-k}{2(1+c)}$ to possibly make the interval from Proposition 4 non-empty. Therefore, the minimum number of syndrome entries required for this algorithm to output a valid solution has to be at least $2(1+c)t\log t/(\delta+1)$. In the noiseless scenario, this becomes $2t\log t$.*



(a) Classic McEliece parameters set and the $[t-3,t]$ interval

(b) BIKE parameters set and the $[t-1,t]$ interval

Fig. 3: Number of ones in the first $n-k$ positions for some of the Classic McEliece and BIKE sets of parameters and different levels of a centered binomial noise.

2) Bounding the required ratio of syndrome entries: The existence of a value such that the ISD-score decoder succeeds in finding a solution using fewer syndrome entries could be deduced. It suffices to replace $(n-k)$ with $\gamma(n-k)$, where $\gamma \in (0,1]$ represents the percentage of syndrome entries required to achieve a high probability. This value can be deduced from Thm. 3. Given $n-k$, the maximum value of t for which the success probability is close to 1 also determines the minimum number of required rows. More exactly, for a fixed value of t and $n-k$, we can compute $\gamma(n-k)$, the value for which t satisfies $8t(1+c)\log\frac{t}{\delta+1} = \gamma(n-k)$. By Thm. 3, with only $\gamma(n-k)$ rows, one can recover a solution of weight at most t w.h.p. Formally, the following holds.

Corollary 3. *Let $d = ct/2$ where c is a constant. The minimum quantity of information required by the ISD-score decoder to find a solution is $4(1+c)t\log\frac{t}{\delta+1}$. Moreover, in the noiseless scenario, the minimum quantity of information becomes $4t\log\frac{t}{\delta+1}$.*

Consequently, the constant term pointed out in Remark 2 may be improved, however not lower than $2(1+c)\log\frac{t}{\delta+1}$.

V. Experimental results

For the simulation we have set the $(n,\ k,\ t)$ parameters according to the specifications of the Classic McEliece [2] and BIKE [3]. We choose two security levels for both schemes to illustrate the performance of the ISD-score decoder.

Our experiments look at the number of syndrome entries required to bring $t-\delta$ ones in the first $n-k$ positions, as dictated by the ISD method. Results are shown in Fig. 3. The green band is the $[t-\delta;t]$ interval for which, given the value of n, exhaustive search for the correct permutation is feasible, *i.e.* $[t-3;t]$ for Classic McEliece and $[t-1;t]$ for BIKE. Let us explain the meaning of the plots, when these are read horizontally. One way this could be read is as the weight of solutions retrieved by the ISD-Score decoder with probability 1. For example, when $n = 8192$ and noise level equal to t we can hope to retrieve solutions of weight at most 122 (which is smaller that the proposed parameters), while for the same length and noise smaller than $t/2$ we can retrieve any solution of weight at most 128 using the ISD-score decoder using $\delta = 3$, or equivalently solutions of weight 125 using the Prange-score decoder. To summarize, except for the case $n = 8192$ with noise levels strictly greater than $t/2$, all the plots suggests that the ISD-score decoder is able to retrieve with high probability a valid solution of weight t in presence of noise.

We can also read the plots vertically. This gives us the ratio of syndrome entries required to find a solution of given weight with high probability. The abscissa of the points of intersection between the curves and the green stripe gives minimum percentage of syndrome entries required in the ISD-score decoder to successfully retrieve a valid solution of weight t. For the BIKE cryptosystem, the ratio of syndrome entries required to bring at least $t-1$ ones in the first $n-k$ positions ranges is as low as $7.5\,\%$ in the noiseless setting. For the Classic McEliece cryptosystem, the ratio of syndrome entries required to bring at least $t-3$ ones in the first $n-k$ positions ranges from $48\,\%$ to $62\,\%$ without noise. In both cases, moderate noise levels are well tolerated, and prove the efficiency of the decoder in these settings.

We have also computed the best theoretical lower bound we could hope for, *i.e.* the percentage of syndrome entries should be at least $\frac{2(1+c)t}{n-k}\log\frac{t}{\delta+1}$. When compared with the experimental results, we noticed that theoretical values are less than $10\,\%$ smaller than the experimental values.

VI. Conclusion

This article evaluated the efficiency of the score decoder for integer syndrome decoding in the presence of noise. We proved that, even in the presence of noise, this decoder is indeed able to successfully bring $t-\delta$ ones in the first $n-k$ positions, as required by the ISD-based methods. We then experimentally validate this capability considering the parameter sets of two post-quantum cryptosystems, Classic McEliece and BIKE. Future works could investigate other types of noise or improve the efficiency of the decoder, bringing it closer to the information-theoretic bound.

REFERENCES

[1] P. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in *FOCS*, S. Goldwasser, Ed., 1994, pp. 124–134.

[2] M. R. Albrecht, D. J. Bernstein, T. Chou, C. Cid, J. Gilcher, T. Lange, V. Maram, I. von Maurich, R. Misoczki, R. Niederhagen, K. G. Paterson, E. Persichetti, C. Peters, P. Schwabe, N. Sendrier, J. Szefer, C. J. Tjhai, M. Tomlinson, and W. Wang, "Classic McEliece," National Institute of Standards and Technology, Tech. Rep., 2020, available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.

[3] N. Aragon, P. Barreto, S. Bettaieb, L. Bidoux, O. Blazy, J.-C. Deneuville, P. Gaborit, S. Gueron, T. Guneysu, C. Aguilar Melchor, R. Misoczki, E. Persichetti, N. Sendrier, J.-P. Tillich, G. Zémor, V. Vasseur, and S. Ghosh, "BIKE," National Institute of Standards and Technology, Tech. Rep., 2020, available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.

[4] C. A. Melchor, N. Aragon, S. Bettaieb, L. Bidoux, O. Blazy, J.-C. Deneuville, P. Gaborit, E. Persichetti, G. Zémor, and I. Bourges, "Hamming quasi-cyclic (hqc)," *NIST PQC Round*, vol. 2, no. 4, p. 13, 2018.

[5] E. Berlekamp, R. McEliece, and H. van Tilborg, "On the inherent intractability of certain coding problems," *IEEE Trans. Inform. Theory*, vol. 24, no. 3, pp. 384–386, May 1978.

[6] E. Prange, "The use of information sets in decoding cyclic codes," *IRE Transactions on Information Theory*, vol. 8, no. 5, pp. 5–9, 1962.

[7] P. J. Lee and E. F. Brickell, "An observation on the security of McEliece's public-key cryptosystem," in *Advances in Cryptology - EUROCRYPT'88*, ser. Lecture Notes in Comput. Sci., vol. 330. Springer, 1988, pp. 275–280.

[8] J. Stern, "A method for finding codewords of small weight," in *Coding Theory and Applications*, ser. Lecture Notes in Comput. Sci., G. D. Cohen and J. Wolfmann, Eds., vol. 388. Springer, 1988, pp. 106–113.

[9] I. Dumer, "Two decoding algorithms for linear codes," *Probl. Inf. Transm.*, vol. 25, no. 1, pp. 17–23, 1989.

[10] A. May, A. Meurer, and E. Thomae, "Decoding random linear codes in $O(2^{0.054n})$," in *Advances in Cryptology - ASIACRYPT 2011*, ser. Lecture Notes in Comput. Sci., D. H. Lee and X. Wang, Eds., vol. 7073. Springer, 2011, pp. 107–124.

[11] A. Becker, A. Joux, A. May, and A. Meurer, "Decoding random binary linear codes in $2^{n/20}$: How $1 + 1 = 0$ improves information set decoding," in *Advances in Cryptology - EUROCRYPT 2012*, ser. Lecture Notes in Comput. Sci. Springer, 2012.

[12] A. May and I. Ozerov, "On computing nearest neighbors with applications to decoding of binary linear codes," in *Advances in Cryptology - EUROCRYPT 2015*, ser. Lecture Notes in Comput. Sci., E. Oswald and M. Fischlin, Eds., vol. 9056. Springer, 2015, pp. 203–228.

[13] L. Both and A. May, "Decoding linear codes with high error rate and its impact for LPN security," in *Post-Quantum Cryptography - 9th International Conference, PQCrypto 2018, Fort Lauderdale, FL, USA, April 9-11, 2018, Proceedings*, ser. Lecture Notes in Computer Science, T. Lange and R. Steinwandt, Eds., vol. 10786. Springer, 2018, pp. 25–46. [Online]. Available: https://doi.org/10.1007/978-3-319-79063-3_2

[14] A. Esser and E. Bellini, "Syndrome decoding estimator," in *IACR International Conference on Practice and Theory of Public-Key Cryptography*, ser. Lecture Notes in Computer Science, G. Hanaoka, J. Shikata, and Y. Watanabe, Eds., vol. 13177. virtual event: Springer, Mar. 2022, pp. 112–141.

[15] A. Horlemann, S. Puchinger, J. Renner, T. Schamberger, and A. Wachter-Zeh, "Information-set decoding with hints," in *International Workshop on Code-Based Cryptography*, ser. Lecture Notes in Computer Science, A. Wachter-Zeh, H. Bartz, and G. Liva, Eds., vol. 13150. Munich, Germany: Springer, jun 2021, pp. 60–83.

[16] V.-F. Dragoi, P.-L. Cayrel, B. Colombier, D. Bucerzan, and S. Hoara, "Solving a modified syndrome decoding problem using integer programming," *INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS and CONTROL*, vol. 15, no. 5, 2020. [Online]. Available: http://univagora.ro/jour/index.php/ijccc/article/view/3920

[17] P. Cayrel, B. Colombier, V. Dragoi, A. Menu, and L. Bossuet, "Message-recovery laser fault injection attack on the classic McEliece cryptosystem," in *Advances in Cryptology - EUROCRYPT 2021 - 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, October 17-21, 2021, Proceedings, Part II*, ser. Lecture Notes in Computer Science, A. Canteaut and F. Standaert, Eds., vol. 12697. Springer, 2021, pp. 438–467.

[18] U. Feige and A. Lellouche, "Quantitative group testing and the rank of random matrices," *CoRR*, vol. abs/2006.09074, 2020. [Online]. Available: https://arxiv.org/abs/2006.09074

[19] B. Colombier, V.-F. Dragoi, P.-L. Cayrel, and V. Grosso, "Message-recovery profiled side-channel attack on the classic McEliece cryptosystem," Cryptology ePrint Archive, Report 2022/125, 2022. [Online]. Available: https://ia.cr/2022/125

[20] J. Bootle, C. Delaplace, T. Espitau, P.-A. Fouque, and M. Tibouchi, "LWE without modular reduction and improved side-channel attacks against bliss," in *Advances in Cryptology – ASIACRYPT 2018*, T. Peyrin and S. Galbraith, Eds. Cham: Springer International Publishing, 2018, pp. 494–524.

[21] C.-C. Cao, C. Li, and X. Sun, "Quantitative group testing-based overlapping pool sequencing to identify rare variant carriers," *BMC Bioinformatics*, vol. 15, no. 195, pp. 1–14, 2014.

[22] C. Wang, Q. Zhao, and C.-N. Chuah, "Group testing under sum observations for heavy hitter detection," in *2015 Information Theory and Applications Workshop (ITA)*, 2015, pp. 149–153.

[23] I. Dinur and K. Nissim, "Revealing information while preserving privacy," in *Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, ser. PODS '03. New York, NY, USA: Association for Computing Machinery, 2003, p. 202–210. [Online]. Available: https://doi.org/10.1145/773153.773173

[24] N. R. Adam and J. C. Worthmann, "Security-control methods for statistical databases: A comparative study," *ACM Comput. Surv.*, vol. 21, no. 4, p. 515–556, dec 1989. [Online]. Available: https://doi.org/10.1145/76894.76895

[25] J. P. Martins, R. Santos, and R. Sousa, *Testing the Maximum by the Mean in Quantitative Group Tests*. Cham: Springer International Publishing, 2014, pp. 55–63. [Online]. Available: https://doi.org/10.1007/978-3-319-05323-3_5

[26] I.-H. Wang, S.-L. Huang, K.-Y. Lee, and K.-C. Chen, "Data extraction via histogram and arithmetic mean queries: Fundamental limits and algorithms," in *2016 IEEE International Symposium on Information Theory (ISIT)*, 2016, pp. 1386–1390.