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[1,2*†], Brice Colombier[3†], Pierre-Louis Cayrel[3†], Vincent Grosso[3†]

[1*]Faculty of Exact Sciences, Aurel Vlaicu University of Arad, Romania.
[2]LITIS, University of Rouen Normandy,France.
[3]CNRS Laboratoire Hubert Curien UMR 5516, University of Lyon, UJM-Saint-Etienne, Saint-Etienne,France.

*Corresponding author(s). E-mail(s): vlad.dragoi@uav.ro;
Contributing authors: b.colombier@univ-st-etienne.fr;
pierre.louis.cayrel@univ-st-etienne.fr; vincent.grosso@univ-st-etienne.fr;
†These authors contributed equally to this work.

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, first introduced by Prange in 1962. A recent line of work considers augmented versions of this strategy, with hints typically 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 BIKE and *Classic McEliece* cryptosystems, which are both candidates for the fourth round of the NIST standardization process.

Keywords: Code-based cryptography, Syndrome decoding problem, Information-set decoding

1 Introduction

1.1 Post-quantum cryptography: on its way to become reality

With the practical feasibility of a quantum computer of sufficient capacity getting more and more probable by the day, the threat posed by Shor's algorithm [38] on number theory base cryptosystems grows as well. To address this threat, NIST began a standardization process in 2016 for post-quantum cryptography. The fourth round of this process started in July 2022 when, in the Key Encapsulation Mechanism category, four candidates were submitted. Among them, the *Classic McEliece* [2] and the BIKE [3] cryptosystems are two solutions based on error-correcting codes. Their security relies on the NP-hardness of the binary syndrome decoding problem (SDP) [6]. The SDP is the core hard problem of several cryptographic constructions, e.g., the FSB hash function [4], the SYND stream cipher [25] or the Stern identification scheme [40].

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 $\boldsymbol{H}\boldsymbol{x} = \boldsymbol{s}^*$. There are three main techniques for solving the SDP: statistical decoding [24, 28, 36, 15, 12], information set decoding (ISD) [37, 30, 39, 31, 19, 20, 10, 23, 7, 33, 5, 34, 9] and generalized inverse based decoding [18]. Information Set Decoding was originally proposed by Prange in 1962 [37], and it has been incrementally refined since by Lee and Brickell [30], Stern [39] and, more recently, by May, Meurer and Thomae [33] and by Becker, Joux, May and Meurer [5]. The complexity of the ISD method has been used to better tune the parameters of the cryptosystems [21] according to the required security levels.

1.2 Integer syndrome decoding

One recent line of work considers modified versions of the SDP, for which additional information is available, for instance via side-channel analysis on implementations of the aforementioned cryptosystems. In [27], 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, instead of the binary one, as if the matrix-vector multiplication had been performed in the integer ring instead of the binary finite field, is considered in [17]. One method to obtain the integer syndrome is by laser fault injection attack, as presented in [13]. The problem one has to solve in this case is the integer syndrome decoding, referred to as $\mathbb{N} - \mathsf{SDP}$, 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, whether $\boldsymbol{H}\boldsymbol{x} = \boldsymbol{s}$ admits a solution of weight t. This problem can be tackled down by means of Integer Linear Programming [13] or probabilistic methods [22].

Another method of obtaining an integer syndrome, much more feasible and realistic than laser fault injection, is by side-channel analysis [14]. However, due to physical factors, the integer entries of the syndrome might not be perfectly accurate. Hence, in the resulting problem, 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 as a vector of random variables. The solution proposed in [14] uses a combination of ISD techniques and the score

decoder from [22]. However, only simulations were provided to assess the performance of this proposal and no theoretical evidence was given.

1.3 Related work

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. [8]. 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 $\mathbb{N}-\mathsf{SDP}$. However, ILWE was solved with another technique that does not seem to work for $\mathbb{N}-\mathsf{SDP}$. Nevertheless, it points out that such scenarios extend broader than code-based cryptography.

Quantitative Group Testing

Quantitative Group Testing (QGT) is an active field of research, lately boosted by 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 [11], traffic monitoring [42] and confidential data transfer [16, 1] to machine learning [32, 43]. The $\mathbb{N}-\mathsf{SDP}$ 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 ISD techniques) a recent solution to the classical QGT [22].

1.4 Contributions

In this article, we analyze in detail the algorithm proposed in [14], 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 sub-linear in n, more exactly, $t \leq O\left(\frac{n-k}{\log(n-k)}\right)$, where n is the length of the code and k the dimension. We consider two noise models, present in several schemes/scenarios, *i.e.*, Binomial centered in zero and Bernoulli variables. We demonstrate that 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 [22]. 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 [22], the conditions we propose on the range of t for which the algorithm succeeds is larger than those from [22]. This gives a lower bound on the number of syndrome entries, or the number of rows in the parity-check matrix, required to find a solution, known as the information theoretic bound. A 5 page short version of this article was presented at the Information Theory Workshop (ITW) 2022. We are extending on this short version by providing the full proofs of our results additional comments and a complete section on experimental results.

Outline of the article

In Section 2 we introduce the SDP and its variants, $\mathbb{N}-\mathsf{SDP}$ and $\mathbb{N}-\mathsf{SDP}$ in the presence of noise. We also recall the cryptographic context where these problems occur. Section 3 begins by recalling the score decoder proposed in [14]. Then, it analyzes the distribution of the discriminant function for the $\mathbb{N}-\mathsf{SDP}$ in the presence of noise. The section ends with the description of the ISD-Score decoder. Next, we analyze the success probability of the ISD-Score decoder in Section 4. The theoretical results from this part are being compared with numerical values from our implementation of the algorithm in Section 5. The section also makes a parallel between the efficiency of the ISD-Score decoder and other methods such as ILP. Finally, we conclude the article in Section 6.

2 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 $n \in \mathbb{N}^*$ a random variable X that follows a distribution will be marked as $X \sim \mathcal{B}er(p)$ for the Bernoulli distribution and by $X \sim \mathcal{B}(n,p)$ for the Binomial distribution. We denote by $W(x)$ the Lambert W function. Matrices and vectors are written in bold capital, respectively small letters. We also use $\mathrm{HW}(\boldsymbol{c})$ to denote the Hamming weight of the vector $\boldsymbol{c}$.

Error correcting codes

Let n and k be two positive integers such that $k \leq n$. An $[n,k]$ linear code can be defined as a sub-vector space of dimension k of the vector space $\mathbb{F}^n$. A code can be specified either by its generator matrix $\boldsymbol{G} \in \mathbb{F}^{k \times n}$ (a basis for the code), or by its parity-check matrix $\boldsymbol{H} \in \mathbb{F}^{(n-k)\times n}$ (a basis for the dual code). The minimum distance, or the Hamming distance of a code $\mathcal{C}$, is the minimum of all $\mathrm{HW}(v)$ for $\boldsymbol{v} \in \mathcal{C}, \boldsymbol{v} \neq \boldsymbol{0}$.

One of the main features of linear codes is their ability to decode noisy information/data. Several general decoding strategies exist, the syndrome decoding problem being one of them.

Some variations of the syndrome decoding problem

Let us start by formally defining the binary SDP.

Definition 1 (SDP)**.**

Inputs: $\boldsymbol{H} \in \mathbb{F}_2^{(n-k)\times n}$, $\boldsymbol{s}^* \in \mathbb{F}_2^{n-k}$, $t \in \mathbb{N}^*$.

Output: $\boldsymbol{x} \in \mathbb{F}_2^n$ *s.t.* $\boldsymbol{H}\boldsymbol{x} = \boldsymbol{s}^*$, *and* $HW(\boldsymbol{x}) = t$.

This problem is NP-Complete [6] and, as we shall quickly see, it constitutes the building block of code-based solution for post-quantum cryptography.

Now, a slightly different problem, the $\mathbb{N}-\mathsf{SDP}$ [13, 17], considers matrix-vector multiplication over the ring of integers instead of the binary field $\mathbb{F}_2$. Formally, the problem can be stated as follows.

Definition 2 ($\mathbb{N}-\mathsf{SDP}$)**.**
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 $\epsilon = (\epsilon_1, \ldots, \epsilon_{n-k})$ 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, instead of having access to an instance of the $\mathbb{N}-\mathsf{SDP}$, *i.e.*, $(\boldsymbol{H}, \boldsymbol{s}, t)$, we are given a noisy syndrome $\widetilde{\boldsymbol{s}} = \boldsymbol{s} + \epsilon$ and the value $\boldsymbol{s}^* = \boldsymbol{s} \pmod 2$ (component-wise).

Definition 3 ($\mathbb{N}-\mathsf{SDP}$ 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.* $\boldsymbol{Hx} = \boldsymbol{s}^*$ *with* $HW(\boldsymbol{x}) = t$
$\boldsymbol{s}^* = \boldsymbol{s} \pmod 2$, *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 $\mathbb{N}-\mathsf{SDP}$.

The Niederreiter encryption framework

Both, *Classic McEliece* [2] and BIKE [3], are based on the Niederreiter encryption scheme [35]. Here, we will focus on the encryption algorithm (see Alg. 1).

Algorithm 1 Niederreiter encryption

1: **function** ENCRYPT($\boldsymbol{m}$, pk)
2: Encode $\boldsymbol{m} \to \boldsymbol{x}$ with $\mathrm{HW}(\boldsymbol{x}) = t$
3: Compute $\boldsymbol{s}^* = \boldsymbol{H}_{\mathrm{pub}}\boldsymbol{x}$
4: **return** $\boldsymbol{s}^*$

Recent message recovery attacks are pointing the encryption step, where the ciphertext is obtained from the multiplication of the public parity-check matrix $\boldsymbol{H}_{\mathrm{pub}}$ and the secret error vector $\boldsymbol{x}$. Hence, in [14, 13] the matrix-vector multiplication is targeted as leakage point (line 3 in Algorithm 1). We shall not insist here on the technical details that allow the derivation of the integer syndrome $\boldsymbol{s}$ or the noisy integer syndrome $\widetilde{\boldsymbol{s}}$ from this matrix-vector computation. However, such an exploit is achievable, hence, enabling one to tackle the $\mathbb{N}-\mathsf{SDP}$ or the $\mathbb{N}-\mathsf{SDP}$ in presence of noise, in order to retrieve the secret message. The sets of (n, k, t) parameters defined in [2] and [3] are given in Table 1.

3 ISD-Score decoder

3.1 Score decoder

The idea of assigning a score to each column was already used in for the $\mathbb{N}-\mathsf{SDP}$ in [14]. The objective is to distinguish columns of $\boldsymbol{H}$ in the support of the solution vector from columns which are outside the support. We shall begin by defining a score decoder, as introduced in [22], that proved to be particularly discriminant in the

Table 1: (n, k, t) parameters for *Classic McEliece* and BIKE

	n	k	t
Classic McEliece	3488	2720	64
	4608	3360	96
	6688	5024	128
	8192	6528	128
BIKE	24646	12323	134
	49318	24659	199
	81946	40973	264

context of $\mathbb{N}-\mathsf{SDP}$. 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 4. *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:*

$$\forall i \in \mathbb{N}_n^* \quad \psi_i(\boldsymbol{s}) = \sum_{\ell=1}^{n-k} \left(h_{\ell,i}\boldsymbol{s}_\ell + (1-h_{\ell,i})(t-\boldsymbol{s}_\ell)\right). \tag{1}$$

For the $\mathbb{N}-\mathsf{SDP}$ in the presence of noise we shall use $\psi_i(\widetilde{\boldsymbol{s}})$. The next result, rephrased from [22], 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.

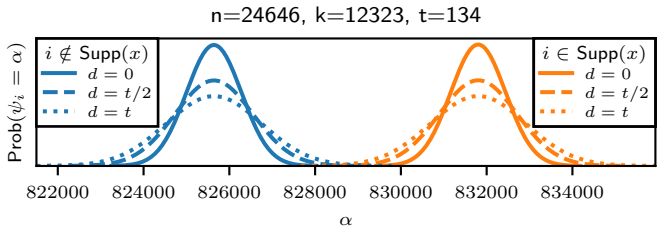
Theorem 1. *Let* $\boldsymbol{H} \in \{0,1\}^{(n-k)\times n}$ *be a random matrix, with distribution given by* $h_{j,i} \sim \mathcal{B}er(\frac{1}{2})$ *and* $\boldsymbol{s} \in \mathbb{N}^{n-k}$ *such that* $\exists\ \boldsymbol{x} \in \{0,1\}^n$ *with* $HW(\boldsymbol{x}) = t$ *satisfying* $\boldsymbol{H}\boldsymbol{x} = \boldsymbol{s}$. *Then*

$$\psi_i(\boldsymbol{s}) \sim \begin{cases} \mathcal{B}((n-k)t, \frac{1}{2}) & , i \notin \mathsf{Supp}(\boldsymbol{x}) \\ \mathcal{B}((n-k)(t-1), \frac{1}{2}) + n - k & , i \in \mathsf{Supp}(\boldsymbol{x}) \end{cases}$$

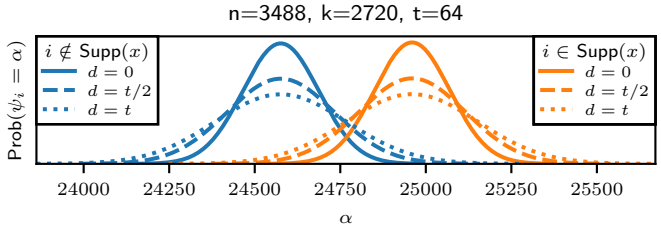
Straightforward from Theorem 1 we have $\mathbb{E}(\psi_i(\boldsymbol{s})) = (n-k)t/2$ for $i \notin \mathsf{Supp}(\boldsymbol{x})$ and $\mathbb{E}(\psi_i(\boldsymbol{s})) = (n-k)t/2 + (n-k)/2$ for $i \in \mathsf{Supp}(\boldsymbol{x})$. The difference in the average value points out that ψ can be a distinguisher between positions in the support and outside the support of the vector $\boldsymbol{x}$. In addition, the variance also differs, fact that will be used in the tail bounds. Moving forward, we will consider the noisy version of this problem in the next section.

3.2 Score decoder in the presence of noise

As in [14], we make some assumptions on the noise considered here, *i.e.*, ϵ_i are independent and identically distributed random variables, the noise does not depend on the distribution of the entries in $\boldsymbol{H}$ and the distribution $\mathcal{D}$ is symmetric.



(a) Example parameters set of BIKE



(b) Example parameters set of *Classic McEliece*

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

Proposition 2 ([14]). *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*

$$\mathsf{Prob}\left(\psi_i(\widetilde{\boldsymbol{s}}) - \psi_i(\boldsymbol{s}) = \alpha\right) = \mathsf{Prob}\left(\sum_{j=1}^{n-k} \epsilon_j = \alpha\right).$$

Proof. Let $Y_{\ell,i} = (2h_{\ell,i} - 1)\epsilon_\ell$. Then we have,

$$\begin{aligned}
\psi_i(\widetilde{\boldsymbol{s}}) &= \sum_{\ell=1}^{n-k} \left(h_{\ell,i}(\widetilde{s_\ell} + (1 - h_{\ell,i})(t - \widetilde{s_\ell}))\right) \\
&= \sum_{\ell=1}^{n-k} \left(h_{\ell,i}(s_\ell + \epsilon_\ell + (1 - h_{\ell,i})(t - s_\ell - \epsilon_\ell))\right) \\
\psi_i(\widetilde{\boldsymbol{s}}) &= \psi_i(\boldsymbol{s}) + \sum_{\ell=1}^{n-k} \underbrace{\left(h_{\ell,i}\epsilon_\ell - (1 - h_{\ell,i})\epsilon_\ell\right)}_{Y_{l,i}}
\end{aligned}$$

For any fixed value of $\ell \in \mathbb{Z}^*_{n-k}$ we have $\mathsf{Prob}(Y_{\ell,i} = \alpha_\ell) = \mathsf{Prob}(\epsilon_\ell = \alpha_\ell)$ for any $\alpha_\ell \in \mathbb{Z}_{-d,d}$ (using the symmetry property and the independence of $h_{\ell,i}$ and ϵ_ℓ). Hence $Y_{\ell,i}$ follows the same distribution as ϵ_ℓ. Thus, $\psi_i(\widetilde{\boldsymbol{s}}) - \psi_i(\boldsymbol{s}) \in \mathbb{Z}_{-(n-k)d,(n-k)d}$ with probability distribution $\mathsf{Prob}(\psi_i(\widetilde{\boldsymbol{s}}) - \psi_i(\boldsymbol{s}) = \alpha) = \mathsf{Prob}\left(\sum_{j=1}^{n-k} \epsilon_j = \alpha\right)$. □

Keeping the difference $\psi_i(\widetilde{\boldsymbol{s}}) - \psi_i(\boldsymbol{s})$ as small as possible resumes to controlling the sum of ϵ_j. The variance of ϵ_j plays a crucial role in the distinguishing capacity of ψ.

Proposition 3. *For any $j \in \mathbb{Z}^*_{n-k}$ let ϵ_j be a discrete random variable satisfying the conditions from Proposition 2 and let $\sigma^2 = Var(\epsilon_j)$. Let $g(n,k,t)$ be a function in the parameters of $\mathbb{N} - \mathsf{SDP}$. Then for any $\alpha > \sigma\sqrt{(n-k)g(n,k,t)}$*

$$\mathsf{Prob}(\psi_i(\widetilde{\boldsymbol{s}}) - \psi_i(\boldsymbol{s}) \geq \alpha) \leq \frac{1}{g(n,k,t)}. \tag{2}$$

Proof. Use Chebyshev inequality for the sum of ϵ_j and the linearity of the variance. □

The case of centered binomial noise

Corollary 4. *Let $d \in \mathbb{N}$ and $\epsilon_i \sim -d + \mathcal{B}(2d, \frac{1}{2})$. Then*

- *for $i \notin \mathsf{Supp}(\boldsymbol{x})$*

$$\psi_i(\widetilde{\boldsymbol{s}}) \sim -d(n-k) + \mathcal{B}\left((n-k)(t+2d), \frac{1}{2}\right)$$

- *for $i \in \mathsf{Supp}(\boldsymbol{x})$*

$$\psi_i(\widetilde{\boldsymbol{s}}) \sim -(d-1)(n-k) + \mathcal{B}\left((n-k)(t-1+2d), \frac{1}{2}\right)$$

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

In particular, we can put $g(n,k,t) = \log\log t$ or $g(n,k,t) = \log\log n$ depending on the wanted speed of convergence.

Figure 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, as shown in Figure 1a, than for the *Classic McEliece* parameters, as shown in Figure 1b.

Bernoulli noise

Proposition 6. *Let $\epsilon_i \sim \mathcal{B}er(\{0,1\}, 1/2)$. Then $\psi_i(\widetilde{\boldsymbol{s}})$ is a random variable that follows the distribution $\begin{cases} \mathcal{B}((n-k)(t+2), \frac{1}{2}) - (n-k) & , i \notin \mathsf{Supp}(\boldsymbol{x}) \\ \mathcal{B}((n-k)(t+1), \frac{1}{2}) & , i \in \mathsf{Supp}(\boldsymbol{x}) \end{cases}$. Moreover, $\mathbb{E}(\psi_i(\widetilde{\boldsymbol{s}})) = \mathbb{E}(\psi_i(\boldsymbol{s}))$ and $Var(\psi_i(\widetilde{\boldsymbol{s}})) = Var(\psi_i(\boldsymbol{s})) + (n-k)/2$.*

Notice that, in the case of a Bernoulli type of noise, the behavior is equivalent to the case of a centered binomial noise. (equivalent to $d = 1$ in Corollary 4). Indeed, the result in Proposition 6 is equivalent to the one given in Corollary 4 with $d = 1$.

3.3 Combining ISD and score decoder

The idea in [14] 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 [37], 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. $\boldsymbol{A}\boldsymbol{H}\boldsymbol{\Pi} = \big(\boldsymbol{I} \| \boldsymbol{B}\big)$. Hence, the system becomes $\big(\boldsymbol{I} \| \boldsymbol{B}\big) \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{0} \end{pmatrix} = \boldsymbol{A}\boldsymbol{s}^*$ 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 the support of $\boldsymbol{x}$ and positions outside. Hence, the underlying permutation, hopefully is a "good" permutation. As pointed out in [14], sorting the list of values $\psi_i(\widetilde{\boldsymbol{s}})$ in descending order is equivalent to generating a permutation $\boldsymbol{\Pi}$. Algorithm 2 finds a solution to the $\mathbb{N} - \mathsf{SDP}$ in the presence of noise as long as $\boldsymbol{\Pi}$ is "good" enough.

Algorithm 2 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** $\mathrm{HW}(\boldsymbol{A}^*\boldsymbol{s}^*) = t$ **then**
4: **return** $\boldsymbol{x} = \boldsymbol{\Pi} \begin{pmatrix} \boldsymbol{A}^*\boldsymbol{s}^* \\ \boldsymbol{0}_{n-r} \end{pmatrix}$ ▷ $r = \mathtt{rank}(\boldsymbol{A})$

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 $\boldsymbol{H}\boldsymbol{\Pi} = [\boldsymbol{A} \parallel \boldsymbol{B}]$ 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 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 7 ([14])**.** 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* $\min\{\psi_i(\widetilde{\boldsymbol{s}}), i \in \mathsf{Supp}(\boldsymbol{x})\} > \max\{\psi_i(\widetilde{\boldsymbol{x}}), i \in L\}$.

The overall time complexity of Prange Score Decoder is $\mathcal{O}((n-k)^3)$, 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 the support of $\boldsymbol{x}$ in the first $n-k$ positions, more powerful ISD methods may be used, *e.g.* Lee-Brickell [30], Stern [39] or Dumer [19]. 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 [14] as a possible solution.

Algorithm 3 Lee-Brickell Score Decoder ([14])

function Lee-Brickell Score Decoder($\boldsymbol{H}, \widetilde{\boldsymbol{s}}, \boldsymbol{s}^*, t$)
  Compute $\boldsymbol{\Pi} \leftarrow$ sort($\boldsymbol{H}, \widetilde{\boldsymbol{s}}, t$)
  Set $\boldsymbol{H\Pi} = [\boldsymbol{A} \parallel \boldsymbol{B}]$
  Compute $\boldsymbol{A}^*, \boldsymbol{H}^* \leftarrow$`rref`($\boldsymbol{H\Pi}$) and $\boldsymbol{B}^* = \boldsymbol{A}^*\boldsymbol{B}$
  Compute $\boldsymbol{s}^{'} = \boldsymbol{A}^*\boldsymbol{s}^*$
  if HW($\boldsymbol{s}^{'}$) == t **then**
    return $\boldsymbol{x} = \boldsymbol{\Pi}(\boldsymbol{s}^{'} \parallel \boldsymbol{0}_k)^t$
  else
    for $i \leftarrow 1, \delta$ **do**
      $S =$ `Gener-Subsets`($\{1, \ldots, k\}$,i)
      for E in S **do**
        $\boldsymbol{x}^{''} \leftarrow$ `Vector`($\{0,1\}, k, E$)
        $\boldsymbol{x}^{'} \leftarrow \boldsymbol{s}^{'} - \boldsymbol{B}^*\boldsymbol{x}^{''}$
        if HW($\boldsymbol{x}^{'}$) == $t - i$ **then**
          return $\left(\boldsymbol{\Pi}(\boldsymbol{x}^{'} \parallel \boldsymbol{x}^{''})^t, \boldsymbol{\Pi}\right)$

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.

Proposition 8. *The δ-ISD-score decoder outputs a valid solution as long as there are 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 a set $J \subset \mathbb{N}_n$ of cardinality $n-k$.*

4 Success probability of the ISD-Score decoder

4.1 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 9. *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 valid solution.*

To prove this theorem we shall use 3 steps. More precisely, we first give an estimation on the tails of the distributions $\psi_i(\widetilde{\boldsymbol{s}})$, then we insert these results into a generic upper bound on the probability of success of the ISD-score decoder, and finally we study the range of parameters for which our conditions are valid.

4.1.1 Tail bounds on the distribution

Firstly we have the following result on the distribution of ψ in the noiseless scenario.
Theorem 10. *Let $\beta \in (0,1)$ and $B_\beta = \frac{(n-k)t}{2} + \frac{\beta(n-k)}{2}$. Then we have*
for $i \notin \mathsf{Supp}(\boldsymbol{x})$

$$\mathsf{Prob}\left(\psi_i(\boldsymbol{s}) \geq B_\beta\right) \leq \frac{e}{\sqrt{2}\pi\beta}\sqrt{\frac{t}{n-k}}e^{-\frac{n-k}{2t}\beta^2}, \tag{3}$$

for $i \in \mathsf{Supp}(\boldsymbol{x})$

$$\mathsf{Prob}\left(\psi_i(\boldsymbol{s}) \leq B_\beta\right) \leq \frac{e}{\pi(1-\beta)}\sqrt{\frac{t-1}{n-k}}e^{-\frac{n-k}{2(t-1)}(1-\beta)^2}. \tag{4}$$

Moving forward, in the case of a binomial noise we have
Theorem 11. *Let $\epsilon_i \sim -d + \mathcal{B}(2d, \frac{1}{2})$, $\beta \in (0,1)$ and B_β as previously defined. Then we have*
for $i \notin \mathsf{Supp}(\boldsymbol{x})$

$$\mathsf{Prob}\left(\psi_i(\widetilde{\boldsymbol{s}}) \geq B_\beta\right) \leq \frac{e}{\sqrt{2}\pi\beta}\sqrt{\frac{t+2d}{n-k}}e^{-\frac{(n-k)\beta^2}{2(t+2d)}}, \tag{5}$$

for $i \in \mathsf{Supp}(\boldsymbol{x})$

$$\mathsf{Prob}\left(\psi_i(\widetilde{\boldsymbol{s}}) \leq B_\beta\right) \leq \frac{e}{\pi(1-\beta)}\sqrt{\frac{t+2d-1}{n-k}}e^{-\frac{(n-k)(1-\beta)^2}{2(t+2d-1)}}. \tag{6}$$

The proof of these theorems is given in Appendix. Let us denote the two upper bounds in Theorem 11 by $\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})}(n,k,t,\beta)$ and $\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})^c}(n,k,t,\beta)$.

4.1.2 A general bound on the success probability using tail estimations

A general theorem regarding the success probability of ISD-score decoder can be stated. For that we suppose that the distribution $\psi_i(\widetilde{\boldsymbol{s}})$ when $i \in \mathsf{Supp}(\boldsymbol{x})$ has to be different from $\psi_i(\widetilde{\boldsymbol{s}})$ when $i \notin \mathsf{Supp}(\boldsymbol{x})$, e.g., it is at least shifted. If not it is obvious that ISD-score decoder can not retrieve a valid solution with high probability.
Theorem 12. *Let $\psi_i(\widetilde{\boldsymbol{s}})$ be random variables and $f(n,k,t,d,B), g(n,k,t,d,B)$ be two functions s.t.*

$$\mathsf{Prob}(\psi_i(\widetilde{\boldsymbol{s}}) \leq B) \leq e^{-f(n,k,t,d,B)} \quad , i \in \mathsf{Supp}(\boldsymbol{x}) \tag{7}$$

$$\mathsf{Prob}(\psi_i(\widetilde{\boldsymbol{s}}) \geq B) \leq e^{-g(n,k,t,d,B)} \quad , i \notin \mathsf{Supp}(\boldsymbol{x}) \tag{8}$$

The ISD-score decoder finds the solution if $\exists B^*$ *s.t.*

- $0 \leq 1 - \frac{t}{\delta+1} e^{-f(n,k,t,d,B^*)} \leq 1$,
- $0 \leq 1 - \frac{n-t}{n-k-t+\delta+1} e^{-g(n,k,t,d,B^*)} \leq 1$,
- $\frac{t}{\delta+1} e^{-f(n,k,t,d,B^*)} + \frac{n-t}{n-k-t+\delta+1} e^{-g(n,k,t,d,B^*)}$ *is close to zero,*

Typically, the theorem gives a sufficient condition for having a high probability of success. Indeed, if one finds a value B_β for which the lower bound tends to 1 then the Score function achieves its goal, namely to distinguish positions in the support of $\boldsymbol{x}$ from those outside it. The proof of this result is given in the Appendix.

Combining the tail bounds on the distribution of $\psi_i(\widetilde{\boldsymbol{s}})$ with the condition on β^* for having a high probability of success enables the following result. Denote $\mathrm{Lb}_{\mathsf{Supp}(\boldsymbol{x})^c} = 1 - \frac{e(n-t)}{\sqrt{2}\pi\beta(n-k-t+\delta+1)} \sqrt{\frac{t+2d}{n-k}} e^{-\frac{(n-k)\beta^2}{2(t+2d)}}$, $\mathrm{Lb}_{\mathsf{Supp}(\boldsymbol{x})} = 1 - \frac{e.t}{\pi(1-\beta)(\delta+1)} \sqrt{\frac{t+2d-1}{n-k}} e^{-\frac{(n-k)(1-\beta)^2}{2(t+2d-1)}}$.

Proposition 13. *Let* $\epsilon_i \sim -d + \mathcal{B}(2d, \frac{1}{2})$. *If* $\exists \beta^* \in (0,1)$ *s.t.* $\mathrm{Lb}_{\mathsf{Supp}(\boldsymbol{x})}, \mathrm{Lb}_{\mathsf{Supp}(\boldsymbol{x})^c} \in [0,1]$ *and* $\mathrm{Lb}_{\mathsf{Supp}(\boldsymbol{x})}\mathrm{Lb}_{\mathsf{Supp}(\boldsymbol{x})^c}$ *is close to 1, then w.h.p. ISD-score decoder succeeds in finding a valid solution.*

Corollary 14. *When* $d = 0$ *and* $\delta = 0$ *the condition on* β^* *simplifies to*

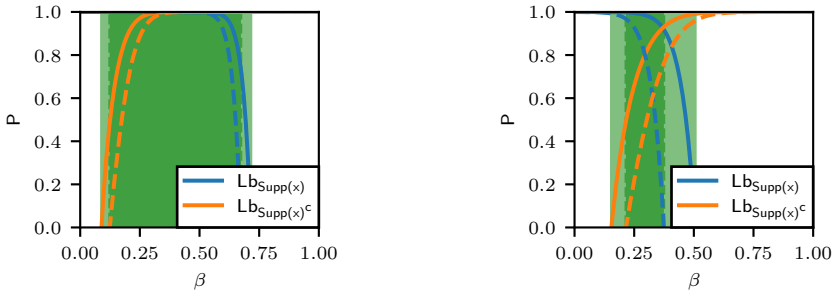
- $0 \leq \frac{et}{\pi(1-\beta)} \sqrt{\frac{t}{n-k}} e^{-\frac{(n-k)(1-\beta)^2}{2t}} \leq 1$,
- $0 \leq \frac{e(n-t)}{(\sqrt{2}\pi\beta)(n-k-t)} \sqrt{\frac{t}{n-k}} e^{-\frac{(n-k)\beta^2}{2t}} \leq 1$,
- $\frac{et}{\pi(1-\beta)} \sqrt{\frac{t}{n-k}} e^{-\frac{(n-k)(1-\beta)^2}{2t}} + \frac{e(n-t)}{(\sqrt{2}\pi\beta)(n-k-t)} \sqrt{\frac{t}{n-k}} e^{-\frac{(n-k)\beta^2}{2t}}$ *is close to zero,*

To fairly compare with state-of-the-art techniques such as the algorithm in [22], which is only valid for the noiseless scenario, we adapted the conditions from [22] 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 Figure 2, we plot the modified functions from [22] (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 [22], and our functions, respectively, are 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 Figure 2d, the interval is empty w.r.t. conditions in [22], while w.r.t. our conditions the interval exists.

4.1.3 Range of valid parameters

Here, we shall determine the conditions on the parameters such that the conditions in Proposition 13 are satisfied. We will begin by determining the existence of β^*. We will need to denote by $W(x)$ the Lambert W function.



(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 [22] (dashed lines) and the proposed ones (solid lines)

Proposition 15. *For any* $\beta \geq \sqrt{\frac{t+2d}{n-k} W\left(\frac{n-t}{n-k-t+\delta+1}\frac{e}{\sqrt{2\pi}}\right)^2}$ *we have that* $\frac{n-t}{n-k-t+\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})^c}(n,k,t,d,\beta) \leq 1$, *and for any* $\beta \leq 1 - \sqrt{\frac{t+2d-1}{n-k} W\left(\frac{t}{\delta+1}\frac{e}{\pi}\right)^2}$ *we have that* $\frac{t}{\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})}(n,k,t,d,\beta) \leq 1$.

Having both functions positive and strictly smaller than 1, at the same time, can be achieved as long the interval defined by the two extreme points, in the previous Proposition is non-empty, i.e.,

$$\sqrt{\frac{t+2d}{n-k} W\left(\frac{n-t}{n-k-t+\delta+1}\frac{e}{\sqrt{2\pi}}\right)^2} \leq 1 - \sqrt{\frac{t+2d-1}{n-k} W\left(\frac{t}{\delta+1}\frac{e}{\pi}\right)^2}$$

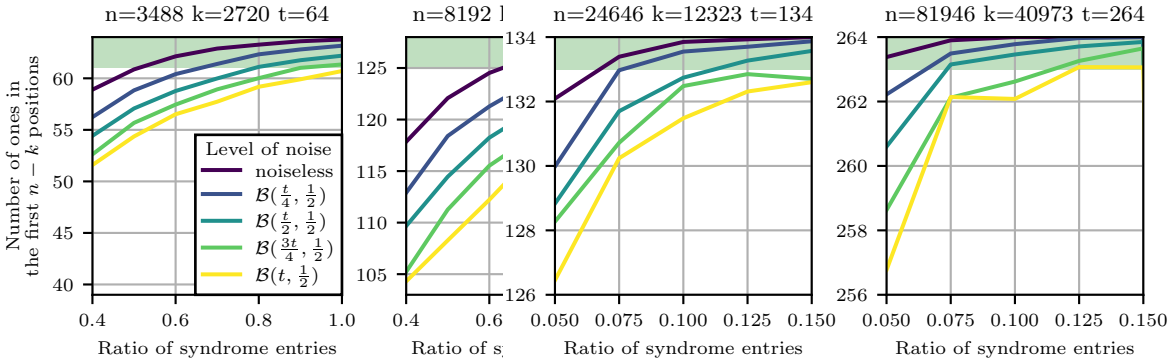
To give a more sensitive meaning of our result, we could approximate the value of the Lambert W function by $W(m) = \log m - \log\log m + \frac{\log\log m}{\log m}$ as m tends to infinity. Using only the first term we define $I_\beta = \left[\sqrt{\frac{2(t+2d)}{n-k}\log\frac{n-t}{n-k-t+\delta+1}}, 1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

Proposition 16. *If $I_\beta \neq \emptyset$ then the probability of success of the ISD-score decoder is at least*

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

Typically, our result gives a sub-interval where the conditions are safely satisfied. When simulations are to be performed, one could solve the inequalities in order to determine a more accurate interval. However, in using the Taylor series of the LambertW function we can deduce the following.

Corollary 17. *Let $f_{n,k,t,\delta} = \frac{n-t}{n-k-t+\delta+1}$ and $f^*_{t,\delta} = \frac{t}{\delta+1}$. The extreme points of the interval where the first two conditions in Theorem 13 are satisfied, converges to $\sqrt{\frac{t+2d}{n-k}\left(2\log f_{n,k,t,\delta} - \log 2\log f_{n,k,t,\delta} + \frac{\log 2\log f_{n,k,t,\delta}}{2\log f_{n,k,t,\delta}}\right)}$, and $1 - \sqrt{\frac{t+2d-1}{n-k}\left(2\log f^*_{t,\delta} - \log 2\log f^*_{t,\delta} + \frac{\log 2\log f^*_{t,\delta}}{2\log f^*_{t,\delta}}\right)}$.*



(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.

4.2 Information-theoretic bounds

4.2.1 Bounding the value of t

To see how large the weight of the error t must be to have a non-empty interval, the following rough estimate can be used.

Theorem 18 (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*

$$t \leq \frac{n-k}{8(1+c)W\left(\frac{n-k}{8(1+c)(\delta+1)}\right)} \tag{9}$$

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

Using a first term approximation for the Lambert W function near infinity, we obtain a threshold on t. More exactly this value can be approximated by $\frac{n-k}{8(1+c)\log\frac{n-k}{8(1+c)(\delta+1)}}$.

Now, recall that we have determined a preliminary condition on d, such that the ψ function can distinguish between positions in the support of the solution and outside it. This condition was $d \leq \frac{n-k}{8\log\log(n-k)}$. Taking a slightly smaller noise level, *e.g.* $d = \frac{n-k}{8\log(n-k)} \leq \frac{n-k}{8\log\log(n-k)}$ validates the choice in the hypothesis $d = ct/2$, as per Theorem 18 $t \leq \mathcal{O}\left(\frac{n-k}{\log(n-k)}\right)$. Taking into account this condition and the hypothesis of Theorem 18, *i.e.* $d = ct/2$, we deduce the following upper bound on t

$$d = \frac{ct}{2} \leq \frac{n-k}{8\log t} \quad \Rightarrow \quad t\log t \leq \frac{n-k}{4c}. \tag{10}$$

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

4.2.2 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 Theorem 18. Typically, given a number of rows $n-k$, the maximum value of t for which the success probability is close enough 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 Theorem 18, with only $\gamma(n-k)$ rows, one can recover a solution of weight at most t with high probability. Formally, the following holds.

Corollary 19. *Let* $d = ct/2$ *where* c *is a constant. Then the minimum quantity of information required by the ISD-score decoder to find a valid 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, we deduce that one could improve the constant term, however, not lower than $2(1+c)\log\frac{t}{\delta+1}$.

5 Experimental results

The following experiments have been carried out on a standard laptop embedding an 8-core processor running at 1.6 GHz and 32 GB of RAM. The ILP solver we used is provided by the Scipy Python package [41] under the `scipy.optimize.linprog` function. The score decoder is implemented using the Numpy Python package [26] to perform matrix computations.

5.1 Success probability and ratio of syndrome entries

For the results presented below, we set the $(n,\ k,\ t)$ parameters according to the specifications of the *Classic McEliece* [2] and BIKE [3] cryptosystems.

The following 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 Figure 3, for both the *Classic McEliece* and the BIKE cryptosystems. 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. The green stripe represents the region corresponding to possible values of δ. The value of δ for the $[t-\delta; t]$ interval is lower for the BIKE cryptosystem since it comes with much larger values of n, making the exhaustive search for the correct permutation much more costly. Conversely, we allow for $\delta = 3$ in the case of *Classic McEliece* since the n values are smaller. 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 than 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 suggest 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 from 4.75 % to 6.5 %. 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 %. 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 comparing the experimental results shown in Figure 3 and Table 2, we observe that theoretical values are around 10 % smaller than the experimental values.

5.2 ILP solver and ISD-score decoder

Percentage of required entries

To compare the ILP solver with the ISD-score decoder we used the parameters for the *Classic McEliece* proposal. We decided to consider only the *Classic McEliece* because the execution time of the ILP solver for the smallest parameters of BIKE exceeded tens of minutes for a single instance of the $\mathbb{N}-\mathsf{SDP}$. Obtaining in a reasonable time a solid statistical evidence of the performance of the ILP solver for BIKE, would assume a much more optimized implementation of the solver, which is not the main purpose of this article. The results for the ILP solver in the noiseless scenario are given in Figure 4. The success rate is computed for ten evenly spaced ratios ranging from 1 to 100 %.

Table 2: Theoretical lower bound on the ratio of syndrome entries necessary for the ISD-score decoder

n	noiseless	$\mathcal{B}(\frac{t}{4}, \frac{1}{2})$	$\mathcal{B}(\frac{t}{2}, \frac{1}{2})$	$\mathcal{B}(\frac{3t}{4}, \frac{1}{2})$	$\mathcal{B}(t, \frac{1}{2})$
Classic McEliece					
3488	0.46	0.58	0.69	0.81	0.92
4608	0.49	0.61	0.73	0.86	0.98
6688	0.53	0.67	0.80	0.93	1.00
8192	0.53	0.67	0.80	0.93	1.00
BIKE					
24646	0.09	0.11	0.14	0.16	0.18
49318	0.07	0.09	0.11	0.13	0.15
81946	0.06	0.08	0.09	0.11	0.13

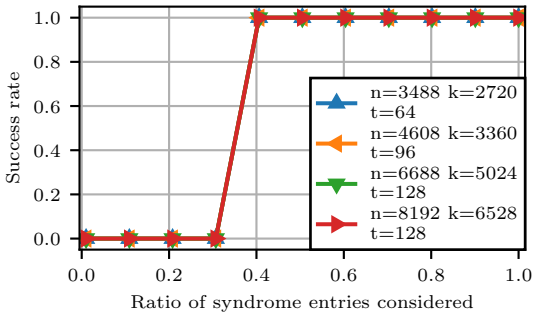


Fig. 4: Success rate of the ILP solver for the $\mathbb{N} - \mathsf{SDP}$ for four sets of parameters and different ratios of syndrome entries considered

We observe that the behavior is the same for all sets of parameters. When considering 30 % of syndrome entries, the ILP solver failed at recovering the error vector ten times out of ten. Conversely, when considering 40 % of syndrome entries, the ILP solver succeeded at recovering the error vector ten times out of ten. Hence, the main drawback of the ILP solver, when compared to the ISD-score decoder, is that the ILP cannot be used when only a small percentage of syndrome entries are known.

Noisy setting

In a noisy setting, the differences between the ILP solver and the ISD-score decoder is even more dramatic. Indeed, the ILP solver either succeeds in finding a valid solution, with t ones in the first t positions, or it fails. Conversely, the ISD-score decoder succeeds if $t - \delta$ ones are in the first $(n - k)$ positions, providing a much larger margin in the noisy setting.

Eventually, the permutation returned by the ISD-score decoder is always better than a random permutation. Therefore, one can always resort to exhaustive search afterwards.

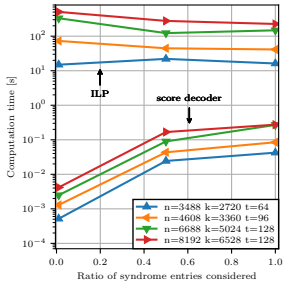


Fig. 5: Computation time of the ILP solver and the Score decoder

Computation time

When comparing the time required by the two algorithms for retrieving a valid solution, we notice a significant gap between the two algorithms. From Figure 5 we can see that it takes less than 0.1 s for the ISD-score decoder, while for the ILP it takes at least 10 s for any of the parameters of the *Classic McEliece* scheme. Broadly speaking, the ILP solver is three orders of magnitude slower than the ISD-score decoder.

6 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.

Acknowledgments.

Appendix A Proof of Theorem 1

Proof. By definition 4 we have that

$$\psi_i(\boldsymbol{s}) = \sum_{\ell=1}^{n-k} \left(h_{\ell,i} s_\ell + (1 - h_{\ell,i})(t - s_\ell)\right). \tag{A1}$$

Let us denote $X_\ell = h_{\ell,i} s_\ell + (1 - h_{\ell,i})(t - s_\ell)$. As $s_\ell = \sum_{j \in \mathsf{Supp}(\boldsymbol{x})} h_{\ell,j}$ we deduce that

$$X_\ell = h_{\ell,i} \sum_{j \in \mathsf{Supp}(\boldsymbol{x})} h_{\ell,j} + (1 - h_{\ell,i})(t - \sum_{j \in \mathsf{Supp}(\boldsymbol{x})} h_{\ell,j}). \tag{A2}$$

If $i \notin \mathsf{Supp}(\boldsymbol{x})$ then

$$X_\ell = \begin{cases} \sum\limits_{j \in \mathsf{Supp}(\boldsymbol{x})} h_{\ell,j} & = s_\ell \quad , \text{ if } h_{\ell,i} = 1 \\ t - \sum\limits_{j \in \mathsf{Supp}(\boldsymbol{x})} h_{\ell,j} & = t - s_\ell \ , \text{ if } h_{\ell,i} = 0 \end{cases}$$

As $s_\ell \sim \mathcal{B}(t, \frac{1}{2})$ we deduce that $X_\ell \sim \mathcal{B}(t, \frac{1}{2})$ for all $i \notin \mathsf{Supp}(\boldsymbol{x})$, and by independence we obtain $\psi_i(\boldsymbol{s}) \sim \mathcal{B}((n-k)t, \frac{1}{2})$, then $\mathbb{E}(\psi_i(\boldsymbol{s})) = \frac{(n-k)t}{2}$ and $Var(\psi_i(\boldsymbol{s})) = \frac{(n-k)t}{4}$.

If $i \in \mathsf{Supp}(\boldsymbol{x})$ we have that s_ℓ and $h_{\ell,i}$ are dependent random variables. Hence we obtain

$$X_\ell = \begin{cases} 1 + \sum\limits_{j \in \mathsf{Supp}(\boldsymbol{x}) \setminus \{i\}} h_{\ell,j} & , \text{ if } h_{\ell,i} = 1 \\ 1 + (t-1) - \sum\limits_{j \in \mathsf{Supp}(\boldsymbol{x}) \setminus \{i\}} h_{\ell,j} & , \text{ if } h_{\ell,i} = 0 \end{cases}$$

As $s_\ell - h_{\ell,i} \sim \mathcal{B}(t-1, \frac{1}{2})$ we deduce that $X_\ell \sim 1 + \mathcal{B}(t-1, \frac{1}{2})$ for all $i \in \mathsf{Supp}(\boldsymbol{x})$, and by independence of the variables X_ℓ we obtain $\psi_i(\boldsymbol{s}) \sim (n-k) + \mathcal{B}((n-k)(t-1), \frac{1}{2})$. □

Appendix B Proof of Corollary 5

Proof. Apply Proposition 3 and Corollary 4 to obtain the the results. In order to determine the upper bound on d, we start by computing the intervals of confidence for $\psi_i(\widetilde{\boldsymbol{s}})$ from Proposition 3. This yields an interval $I_{\widetilde{\boldsymbol{s}}}(i)$ defined by the two extremal points $\mathbb{E}(\psi_i(\widetilde{\boldsymbol{s}})) \pm \sqrt{\frac{d(n-k)g(n,k,t)}{2}}$, i.e.,

- $i \notin \mathsf{Supp}(\boldsymbol{x})$ the points $\frac{(n-k)t}{2} \pm \sqrt{\frac{d(n-k)g(n,k,t)}{2}}$
- $i \in \mathsf{Supp}(\boldsymbol{x})$ the points $\frac{(n-k)t}{2} + \frac{n-k}{2} \pm \sqrt{\frac{d(n-k)g(n,k,t)}{2}}$.

The two intervals are disjoint if we have

$$2\sqrt{\frac{d(n-k)g(n,k,t)}{2}} \leq \frac{n-k}{2} \tag{B3}$$

Hence, we obtain $d \leq \frac{n-k}{8g(n,k,t)}$. □

Appendix C Proof of Theorem 10 and Theorem 11

Let us begin by a useful result.

Lemma 20 ([29]). *Let $X \sim \mathcal{B}(n, \frac{1}{2})$ and $\frac{n}{2} \leq \alpha \leq n$. Then*

$$\mathsf{Prob}(X \geq \alpha) \leq \frac{\alpha + 1}{2\alpha - n + 1} \mathsf{Prob}(X = \alpha). \tag{C4}$$

Lemma 21 ([22]). *Let $X \sim \mathcal{B}(n, \frac{1}{2})$ and $\alpha \leq \frac{n}{2}$. Then*

$$\mathsf{Prob}\left(X = \frac{n}{2} + \alpha\right) \leq \frac{e}{2\pi}\sqrt{\frac{n}{\frac{n^2}{4} - \alpha^2}}e^{-\frac{2\alpha^2}{n}}. \tag{C5}$$

Proposition 22. *Let $X \sim \mathcal{B}(n, \frac{1}{2})$ and $\alpha < n$. Then*

$$\mathsf{Prob}\left(X \geq \frac{n}{2} + \frac{\alpha}{2}\right) \leq \frac{e}{2\pi}\left(1 + \frac{n+1}{\alpha+1}\right)\sqrt{\frac{n}{n^2 - \alpha^2}}e^{-\frac{\alpha^2}{2n}}. \tag{C6}$$

Proof. Use Lemma 20 and 21. □

We can now proceed to the proof of Theorem 10.

Proof. Recall that

$$\psi_i(\boldsymbol{s}) \sim \begin{cases} \mathcal{B}((n-k)t, \frac{1}{2}) & \text{for } i \notin \mathsf{Supp}(\boldsymbol{x}) \\ n - k + \mathcal{B}((n-k)(t-1), \frac{1}{2}) & \text{for } i \in \mathsf{Supp}(\boldsymbol{x}) \end{cases}$$

By Proposition 22, for $i \notin \mathsf{Supp}(\boldsymbol{x})$ we have that

$$\mathsf{Prob}\left(\psi_i(\boldsymbol{s}) \geq B_\beta\right) \leq \frac{\frac{e}{2\pi}\left(1 + \frac{(n-k)t+1}{(n-k)\beta+1}\right)\sqrt{\frac{(n-k)t}{(n-k)^2t^2 - (n-k)^2\beta^2}}}{e^{\frac{(n-k)^2\beta^2}{2(n-k)t}}} \tag{C7}$$

$$\leq \frac{e}{2\pi\beta}\sqrt{\frac{t+\beta}{t-\beta}}\sqrt{\frac{t}{(n-k)}}e^{-\frac{(n-k)\beta^2}{2t}} \tag{C8}$$

$$\leq \frac{e}{\sqrt{2}\pi\beta}\sqrt{\frac{t}{(n-k)}}e^{-\frac{(n-k)\beta^2}{2t}} \tag{C9}$$

For $i \in \mathsf{Supp}(\boldsymbol{x})$ we have that $E(\psi_i(\boldsymbol{s})) = \frac{(n-k)t}{2} + \frac{n-k}{2}$. Hence, by Proposition 22 we obtain that $\mathsf{Prob}\left(\psi_i(\boldsymbol{s}) \leq \frac{(n-k)t}{2} + \frac{(n-k)\beta}{2}\right)$ is upper bounded by

$$\leq \frac{\frac{e}{2\pi}\left(1 + \frac{(n-k)(t-1)+1}{(n-k)(1-\beta)+1}\right)\sqrt{\frac{(n-k)(t-1)}{(n-k)^2(t-1)^2 - (n-k)^2(1-\beta)^2}}}{e^{\frac{(n-k)^2(1-\beta)^2}{2(n-k)(t-1)}}} \tag{C10}$$

$$\leq \frac{e}{2\pi\beta}\sqrt{\frac{t-\beta}{t+\beta+2}}\sqrt{\frac{t-1}{(n-k)}}e^{-\frac{(n-k)(1-\beta)^2}{2(t-1)}} \tag{C11}$$

$$\leq \frac{e}{2\pi\beta}\sqrt{\frac{t-1}{(n-k)}}e^{-\frac{(n-k)\beta^2}{2(t-1)}}. \tag{C12}$$

□

As for Theorem 11 we have:

Proof. Recall that we have

- for $i \notin \mathsf{Supp}(\boldsymbol{x})$

$$\psi_i(\widetilde{\boldsymbol{s}}) \sim -d(n-k) + \mathcal{B}\left((n-k)(t+2d), \frac{1}{2}\right);$$

- for $i \in \mathsf{Supp}(\boldsymbol{x})$

$$\psi_i(\widetilde{\boldsymbol{s}}) \sim -(d-1)(n-k) + \mathcal{B}\left((n-k)(t-1+2d), \frac{1}{2}\right).$$

The proof is thus identical with that of Theorem 10 by simply putting $t' = t + 2d$ when $i \notin \mathsf{Supp}(\boldsymbol{x})$ and $t' = t + 2d - 1$ when $i \in \mathsf{Supp}(\boldsymbol{x})$. □

Appendix D Proof of Theorem 12

Proof. Let X_B denote the number of indices $j \in \mathsf{Supp}(\boldsymbol{x})$ for which $\psi_i(\widetilde{\boldsymbol{s}}) \leq B$, and Y_B the number of indices $j \notin \mathsf{Supp}(\boldsymbol{x})$ for which $\psi_i(\widetilde{\boldsymbol{s}}) \geq B$. The probability of success of our algorithm equals

$$\begin{aligned}
&= \sum_B \mathsf{Prob}(X_B \leq \delta)\mathsf{Prob}(Y_B \leq n-k-t+\delta) \\
&= \sum_B (1 - \mathsf{Prob}(X_B \geq \delta+1)) \cdot (1 - \mathsf{Prob}(Y_B \geq n-k-t+\delta+1)) \\
&\geq \sum_B \left(1 - \frac{t}{\delta+1} e^{-f(n,k,t,d,B)}\right) \cdot \left(1 - \frac{n-t}{n-k-t+\delta+1} e^{-g(n,k,t,d,B)}\right).
\end{aligned}$$

In the last equation we have used Markov's inequality. Also, the last sum is over those values B for which the two terms in the sum are both positive and smaller than 1. Now suppose that a B^* satisfying the required condition exists. Then the probability of success is

$$\begin{aligned}
&\geq \left(1 - \frac{t}{\delta+1} e^{-f(n,k,t,d,B^*)}\right) \cdot \left(1 - \frac{n-t}{n-k-t+\delta+1} e^{-g(n,k,t,d,B^*)}\right) \\
&\geq 1 - \frac{t}{\delta+1} e^{-f(n,k,t,d,B^*)} - \frac{n-t}{n-k-t+\delta+1} e^{-g(n,k,t,d,B^*)}.
\end{aligned}$$

□

Appendix E Range of valid parameters: proofs and comments

The first useful results concerns the monotony of the two upper bounds.

Lemma 23. *The functions* $\frac{t}{\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})}(n,k,t,d,\beta)$ *and* $\frac{n-t}{n-k-t+\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})^c}(n,k,t,d,\beta)$ *in* $\beta \in (0,1)$*, are positive increasing, and positive decreasing, resp.*

Proof. Let $f(n,k,t,d,\beta) = \frac{t}{\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})}(n,k,t,\beta^*)$ and $g(n,k,t,d,\beta) = \frac{n-t}{n-k-t+\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})^c}(n,k,t,d,\beta)$. We have that both functions f,g are positive. We also have

$$\frac{\partial g(n,k,t,d,\beta)}{\partial \beta} = -\frac{(n-k)\beta^2+(t+2d)}{\beta(t+2d)}g(n,k,t,d,\beta)$$
$$\frac{\partial f(n,k,t,d,\beta)}{\partial \beta} = \frac{(n-k)(1-\beta)^2+(t+2d-1)}{(1-\beta)(t+2d-1)}f(n,k,t,d,\beta).$$

Using the fact that f and g are positive we deduce the wanted result. □

Now we can demonstrate Proposition 15.

Proof. Let us consider the limit point β where the two functions equal 1. As the first function is decreasing we then obtain a lower bound on β.

$$\frac{n-t}{n-k-t+\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})^c}(n,k,t,d,\beta) = 1 \tag{E13}$$
$$\frac{n-t}{n-k-t+\delta+1}\frac{e}{\sqrt{2}\pi\beta}\sqrt{\frac{t+2d}{n-k}}e^{-\frac{(n-k)\beta^2}{2(t+2d)}} = 1 \tag{E14}$$
$$\left(\frac{n-t}{n-k-t+\delta+1}\frac{e}{\sqrt{2}\pi}\right)^2\frac{t+2d}{(n-k)\beta^2} = e^{\frac{(n-k)\beta^2}{t+2d}} \tag{E15}$$

By letting $y = \frac{(n-k)\beta^2}{t+2d}$ we have

$$ye^y = \left(\frac{n-t}{n-k-t+\delta+1}\frac{e}{\sqrt{2}\pi}\right)^2, \tag{E16}$$

admitting a real solution $y = W\left(\frac{n-t}{n-k-t+\delta+1}\frac{e}{\sqrt{2}\pi}\right)^2$, where W is the Lambert W function. From this we deduce $\beta = \sqrt{\frac{t+2d}{n-k}W\left(\frac{n-t}{n-k-t+\delta+1}\frac{e}{\sqrt{2}\pi}\right)^2}$. The second function is increasing hence, it gives an upper bound on β.

$$\frac{t}{\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})}(n,k,t,d,\beta) = 1 \tag{E17}$$
$$\frac{t}{\delta+1}\frac{e}{\pi(1-\beta)}\sqrt{\frac{t+2d-1}{n-k}}e^{-\frac{(n-k)(1-\beta)^2}{2(t+2d-1)}} = 1 \tag{E18}$$
$$\left(\frac{t}{\delta+1}\frac{e}{\pi}\right)^2\frac{t+2d-1}{(n-k)(1-\beta)^2} = e^{\frac{(n-k)(1-\beta)^2}{t+2d-1}} \tag{E19}$$

As in the first case we obtain $1-\beta = \sqrt{\frac{t+2d-1}{n-k} W\left(\frac{t}{\delta+1}\frac{e}{\pi}\right)^2}$. □

Proposition 16 gives a slightly weaker condition, however, it helps understanding the order of magnitude of the parameters. Let us demonstrate the result.

Proof. Let $\beta \geq \beta_1 = \sqrt{2\frac{t+2d}{n-k}\log\frac{n-t}{n-k-t+\delta+1}}$. Then the quantity $\frac{n-t}{n-k-t+\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})^c}(n,k,t,d,\beta)$ equals

$$= \frac{n-t}{n-k-t+\delta+1}\frac{e}{\sqrt{2}\pi\beta}\sqrt{\frac{t+2d}{n-k}}e^{-\frac{(n-k)\beta^2}{2(t+2d)}} \tag{E20}$$

$$\leq \frac{n-t}{n-k-t+\delta+1}\frac{e}{2\pi\sqrt{\log\frac{n-t}{n-k-t+\delta+1}}}e^{-\log\frac{n-t}{n-k-t+\delta+1}} \tag{E21}$$

$$= \frac{e}{2\pi}\frac{1}{\sqrt{\log\frac{n-t}{n-k-t+\delta+1}}}. \tag{E22}$$

Let $1-\beta \geq \beta_2 = \sqrt{2\frac{t+2d-1}{n-k}\log\frac{t}{\delta+1}}$. Then the quantity $\frac{t}{\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})}(n,k,t,d,\beta)$ equals

$$= \frac{t}{\delta+1}\frac{e}{\pi(1-\beta)}\sqrt{\frac{t+2d-1}{n-k}}e^{-\frac{(n-k)(1-\beta)^2}{2(t+2d-1)}} \tag{E23}$$

$$\leq \frac{e}{\sqrt{2}\pi}\frac{1}{\sqrt{\log\frac{t}{\delta+1}}}. \tag{E24}$$

From this we deduce

$$\frac{\sqrt{2(t+2d)}}{\sqrt{n-k}}\sqrt{\log\frac{n-t}{n-k-t+\delta+1}} \leq \beta \tag{E25}$$

$$\beta \leq 1 - \frac{\sqrt{2(t+2d-1)}}{\sqrt{n-k}}\sqrt{\log\frac{t}{\delta+1}}. \tag{E26}$$

Now, suppose that $[\beta_1,\beta_2]$ is non-empty and take $\beta^* \in [\beta_1,\beta_2]$. Since $\frac{t}{\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})}(n,k,t,d,\beta)$ is increasing in β, this implies that $\frac{t}{\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})}(n,k,t,d,\beta^*)$ is upper bounded by

$$\frac{t}{\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})}(n,k,t,d,\beta_2) = \frac{e}{\sqrt{2}\pi}\frac{1}{\sqrt{\log\frac{t}{\delta+1}}}. \tag{E27}$$

Also, as $\frac{n-t}{n-k-t+\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})^c}(n,k,t,d,\beta)$ is decreasing in β we have that $\frac{n-t}{n-k-t+\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})^c}(n,k,t,d,\beta^*)$ is upper bounded by

$$\leq \frac{n-t}{n-k-t+\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})^c}(n,k,t,d,\beta_1) \tag{E28}$$

$$\leq \frac{e}{2\pi}\frac{1}{\sqrt{\log\frac{n-t}{n-k-t+\delta+1}}}. \tag{E29}$$

Equations (E27) and (E28) implies that both function $\frac{n-t}{n-k-t+\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})^c}(n,k,t,d,\beta)$, $\frac{t}{\delta+1}\mathrm{Ub}_{\mathsf{Supp}(\boldsymbol{x})}(n,k,t,d,\beta)$ are smaller than 1 in β^* and that the probability of success is at least

$$\left(1-\frac{e}{2\pi}\frac{1}{\sqrt{\log\frac{n-t}{n-k-t+\delta+1}}}\right)\left(1-\frac{e}{\sqrt{2}\pi}\frac{1}{\sqrt{\log\frac{t}{\delta+1}}}\right). \tag{E30}$$

□

The last result to demonstrate from this section is Theorem 18

Proof. Taking the simplified interval for β, the existence of this interval implies

$$\frac{\sqrt{2(t+2d)}}{\sqrt{n-k}}\sqrt{\log\frac{n-t}{n-k-t+\delta+1}} \leq 1-\frac{\sqrt{2(t+2d-1)}}{\sqrt{n-k}}\sqrt{\log\frac{t}{\delta+1}} \tag{E31}$$

$$\sqrt{\log\frac{n-t}{n-k-t+\delta+1}}+\sqrt{\log\frac{t}{\delta+1}} \leq \sqrt{\frac{n-k}{2(t+2d)}} \tag{E32}$$

Using the condition on k we deduce

$$\sqrt{\log\frac{n-t}{n-k-t+\delta+1}} \leq \sqrt{\log\frac{n-t}{n-(n-t+\delta+1-\frac{(n-t)(\delta+1)}{t})-t+\delta+1}} \tag{E33}$$

$$\sqrt{\log\frac{n-t}{n-k-t+\delta+1}} \leq \sqrt{\log\frac{t}{\delta+1}}. \tag{E34}$$

Hence, the following should hold

$$2\sqrt{\log\frac{t}{\delta+1}} \leq \sqrt{\frac{n-k}{2(1+c)t}} \tag{E35}$$

$$\frac{t}{\delta+1}\log\frac{t}{\delta+1} \leq \frac{n-k}{8(1+c)(\delta+1)}, \tag{E36}$$

which is satisfied as long as $t \leq \frac{n-k}{8(1+c)W(\frac{n-k}{8(1+c)(\delta+1)})}$. The initial condition on k implies $t \leq \frac{n-k+2\delta+2-\sqrt{(n-k+2\delta+2)^2-4n(\delta+1)}}{2}$ which is greater than or equal to $\frac{n-k}{8(1+c)W(\frac{n-k}{8(1+c)(\delta+1)})}$.

As for the asymptotic, use one term approximation for the LambertW function near infinity. □

References

[1] Nabil R. Adam and John C. Worthmann. Security-control methods for statistical databases: A comparative study. *ACM Comput. Surv.*, 21(4):515–556, dec 1989.

[2] Martin R. Albrecht, Daniel J. Bernstein, Tung Chou, Carlos Cid, Jan Gilcher, Tanja Lange, Varun Maram, Ingo von Maurich, Rafael Misoczki, Ruben Niederhagen, Kenneth G. Paterson, Edoardo Persichetti, Christiane Peters, Peter Schwabe, Nicolas Sendrier, Jakub Szefer, Cen Jung Tjhai, Martin Tomlinson, and Wen Wang. Classic McEliece. Technical report, National Institute of Standards and Technology, 2020. available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.

[3] Nicolas Aragon, Paulo Barreto, Slim Bettaieb, Loic Bidoux, Olivier Blazy, Jean-Christophe Deneuville, Phillipe Gaborit, Shay Gueron, Tim Guneysu, Carlos Aguilar Melchor, Rafael Misoczki, Edoardo Persichetti, Nicolas Sendrier, Jean-Pierre Tillich, Gilles Zémor, Valentin Vasseur, and Santosh Ghosh. BIKE. Technical report, National Institute of Standards and Technology, 2020. available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.

[4] Daniel Augot, Matthieu Finiasz, and Nicolas Sendrier. A fast provably secure cryptographic hash function. *IACR Cryptol. ePrint Arch.*, 2003:230, 2003.

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

[6] Elwyn Berlekamp, Robert McEliece, and Henk van Tilborg. On the inherent intractability of certain coding problems. *IEEE Trans. Inform. Theory*, 24(3):384–386, May 1978.

[7] Daniel J. Bernstein, Tanja Lange, and Christiane Peters. Smaller decoding exponents: ball-collision decoding. In *Advances in Cryptology - CRYPTO 2011*, volume 6841 of *Lecture Notes in Comput. Sci.*, pages 743–760, 2011.

[8] Jonathan Bootle, Claire Delaplace, Thomas Espitau, Pierre-Alain Fouque, and Mehdi Tibouchi. LWE without modular reduction and improved side-channel attacks against bliss. In Thomas Peyrin and Steven Galbraith, editors, *Advances in Cryptology – ASIACRYPT 2018*, pages 494–524, Cham, 2018. Springer International Publishing.

[9] Leif Both and Alexander May. Decoding linear codes with high error rate and its impact for LPN security. In Tanja Lange and Rainer Steinwandt, editors, *Post-Quantum Cryptography - 9th International Conference, PQCrypto 2018, Fort*

Lauderdale, FL, USA, April 9-11, 2018, Proceedings, volume 10786 of *Lecture Notes in Computer Science*, pages 25–46. Springer, 2018.

[10] Anne Canteaut and Florent Chabaud. A new algorithm for finding minimum-weight words in a linear code: Application to mceliece's cryptosystem and to narrow-sense BCH codes of length 511. *IEEE Trans. Inform. Theory*, 44(1):367–378, 1998.

[11] Chang-Chang Cao, Cheng Li, and Xiao Sun. Quantitative group testing-based overlapping pool sequencing to identify rare variant carriers. *BMC Bioinformatics*, 15(195):1–14, 2014.

[12] Kevin Carrier, Thomas Debris-Alazard, Charles Meyer-Hilfiger, and Jean-Pierre Tillich. Statistical decoding 2.0: Reducing decoding to LPN, 2022.

[13] Pierre-Louis Cayrel, Brice Colombier, Vlad-Florin Dragoi, Alexandre Menu, and Lilian Bossuet. Message-recovery laser fault injection attack on the classic McEliece cryptosystem. In Anne Canteaut and François-Xavier Standaert, editors, *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*, volume 12697 of *Lecture Notes in Computer Science*, pages 438–467. Springer, 2021.

[14] Brice Colombier, Vlad-Florin Dragoi, Pierre-Louis Cayrel, and Vincent Grosso. Message-recovery profiled side-channel attack on the classic McEliece cryptosystem. Cryptology ePrint Archive, Report 2022/125, 2022.

[15] Thomas Debris-Alazard and Jean-Pierre Tillich. Statistical decoding. In *2017 IEEE International Symposium on Information Theory (ISIT)*, pages 1798–1802, 2017.

[16] Irit Dinur and Kobbi Nissim. Revealing information while preserving privacy. In *Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, PODS '03, page 202–210, New York, NY, USA, 2003. Association for Computing Machinery.

[17] Vlad-Florin Dragoi, Pierre-Louis Cayrel, Brice Colombier, Dominic Bucerzan, and Sorin Hoara. Solving a modified syndrome decoding problem using integer programming. *INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS and CONTROL*, 15(5), 2020.

[18] Vlad-Florin Dragoi and Ferucio Laurentiu Tiplea. Generalized-inverse based decoding. Technical report, 2022.

[19] Il'ya Dumer. Two decoding algorithms for linear codes. *Probl. Inf. Transm.*, 25(1):17–23, 1989.

[20] Ilya Dumer. On minimum distance decoding of linear codes. In *Proc. 5th Joint Soviet-Swedish Int. Workshop Inform. Theory*, pages 50–52, Moscow, 1991.

[21] Andre Esser and Emanuele Bellini. Syndrome decoding estimator. In Goichiro Hanaoka, Junji Shikata, and Yohei Watanabe, editors, *IACR International Conference on Practice and Theory of Public-Key Cryptography*, volume 13177 of *Lecture Notes in Computer Science*, pages 112–141, virtual event, March 2022. Springer.

[22] Uriel Feige and Amir Lellouche. Quantitative group testing and the rank of random matrices. *CoRR*, abs/2006.09074, 2020.

[23] Matthieu Finiasz and Nicolas Sendrier. Security bounds for the design of code-based cryptosystems. In M. Matsui, editor, *Advances in Cryptology - ASIACRYPT 2009*, volume 5912 of *Lecture Notes in Comput. Sci.*, pages 88–105. Springer, 2009.

[24] M. P. C. Fossorier, K. Kobara, and H. Imai. Modeling bit flipping decoding based on nonorthogonal check sums with application to iterative decoding attack of mceliece cryptosystem. *IEEE Trans. Inf. Theor.*, 53(1):402–411, jan 2007.

[25] Philippe Gaborit, Cedric Lauradoux, and Nicolas Sendrier. Synd: a fast code-based stream cipher with a security reduction. In *2007 IEEE International Symposium on Information Theory*, pages 186–190, 2007.

[26] Charles R Harris, K Jarrod Millman, Stéfan J Van Der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J Smith, et al. Array programming with numpy. *Nature*, 585(7825):357–362, 2020.

[27] Anna-Lena Horlemann, Sven Puchinger, Julian Renner, Thomas Schamberger, and Antonia Wachter-Zeh. Information-set decoding with hints. In Antonia Wachter-Zeh, Hannes Bartz, and Gianluigi Liva, editors, *International Workshop on Code-Based Cryptography*, volume 13150 of *Lecture Notes in Computer Science*, pages 60–83, Munich, Germany, jun 2021. Springer.

[28] A. Al Jabri. A statistical decoding algorithm for general linear block codes. In Bahram Honary, editor, *Cryptography and Coding*, pages 1–8, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.

[29] Bernhard Klar. Bounds on tail probabilities of discrete distributions. *Probability in the Engineering and Informational Sciences*, 14:161 – 171, 2000.

[30] Pil J. Lee and Ernest F. Brickell. An observation on the security of McEliece's public-key cryptosystem. In *Advances in Cryptology - EUROCRYPT'88*, volume 330 of *Lecture Notes in Comput. Sci.*, pages 275–280. Springer, 1988.

[31] Jeffrey Leon. A probabilistic algorithm for computing minimum weights of large error-correcting codes. *IEEE Trans. Inform. Theory*, 34(5):1354–1359, 1988.

[32] João Paulo Martins, Rui Santos, and Ricardo Sousa. *Testing the Maximum by the Mean in Quantitative Group Tests*, pages 55–63. Springer International Publishing, Cham, 2014.

[33] Alexander May, Alexander Meurer, and Enrico Thomae. Decoding random linear codes in $O(2^{0.054n})$. In Dong Hoon Lee and Xiaoyun Wang, editors, *Advances in Cryptology - ASIACRYPT 2011*, volume 7073 of *Lecture Notes in Comput. Sci.*, pages 107–124. Springer, 2011.

[34] Alexander May and Ilya Ozerov. On computing nearest neighbors with applications to decoding of binary linear codes. In E. Oswald and M. Fischlin, editors, *Advances in Cryptology - EUROCRYPT 2015*, volume 9056 of *Lecture Notes in Comput. Sci.*, pages 203–228. Springer, 2015.

[35] Harald Niederreiter. Knapsack-type cryptosystems and algebraic coding theory. *Problems of Control and Information Theory*, 15(2):159–166, 1986.

[36] R. Overbeck. Statistical decoding revisited. In Lynn Margaret Batten and Reihaneh Safavi-Naini, editors, *Information Security and Privacy*, pages 283–294, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[37] Eugene Prange. The use of information sets in decoding cyclic codes. *IRE Transactions on Information Theory*, 8(5):5–9, 1962.

[38] P.W. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In S. Goldwasser, editor, *FOCS*, pages 124–134, 1994.

[39] Jacques Stern. A method for finding codewords of small weight. In G. D. Cohen and J. Wolfmann, editors, *Coding Theory and Applications*, volume 388 of *Lecture Notes in Comput. Sci.*, pages 106–113. Springer, 1988.

[40] Jacques Stern. A new identification scheme based on syndrome decoding. In Douglas R. Stinson, editor, *Advances in Cryptology - CRYPTO '93, 13th Annual International Cryptology Conference, Santa Barbara, California, USA, August 22-26, 1993, Proceedings*, volume 773 of *Lecture Notes in Computer Science*, pages 13–21. Springer, 1993.

[41] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature methods*, 17(3):261–272, 2020.

[42] Chao Wang, Qing Zhao, and Chen-Nee Chuah. Group testing under sum observations for heavy hitter detection. In *2015 Information Theory and Applications Workshop (ITA)*, pages 149–153, 2015.

[43] I-Hsiang Wang, Shao-Lun Huang, Kuan-Yun Lee, and Kwang-Cheng Chen. Data extraction via histogram and arithmetic mean queries: Fundamental limits and algorithms. In *2016 IEEE International Symposium on Information Theory (ISIT)*, pages 1386–1390, 2016.