QUANTUM REDUCTION OF FINDING SHORT CODE VECTORS TO THE DECODING PROBLEM

THOMAS DEBRIS–ALAZARD[1,2], MAXIME REMAUD[1,3], AND JEAN-PIERRE TILLICH[1]

ABSTRACT. We give a quantum reduction from finding short codewords in a random linear code to decoding for the Hamming metric. This is the first time such a reduction (classical or quantum) has been obtained. Our reduction adapts to linear codes Stehlé-Steinfield-Tanaka-Xagawa re-interpretation of Regev's quantum reduction from finding short lattice vectors to solving the Closest Vector Problem. The Hamming metric is a much coarser metric than the Euclidean metric and this adaptation has needed several new ingredients to make it work. For instance, in order to have a meaningful reduction it is necessary in the Hamming metric to choose a very large decoding radius and this needs in many cases to go beyond the radius where decoding is always unique. Another crucial step for the analysis of the reduction is the choice of the errors that are being fed to the decoding algorithm. For lattices, errors are usually sampled according to a Gaussian distribution. However, it turns out that the Bernoulli distribution (the analogue for codes of the Gaussian) is too much spread out and cannot be used, as such, for the reduction with codes. This problem was solved by using instead a truncated Bernoulli distribution.

1. INTRODUCTION

Code-based Cryptography. Many cryptosystems such as public-key encryption schemes [McE78, Ale11, MTSB12], authentication protocols [Ste93] or pseudorandom generators [FS96] are built relying on the hardness of finding the closest codeword, a task called *decoding*. In the case of a random linear code, which is the standard case, this problem can be expressed as follows:

Problem 1 ($\mathsf{DP}(q,n,k,t)$)**.** *The decoding problem with parameters $q,n,k,t \in \mathbb{N}$ is defined as:*

- *Given: $(\mathbf{G}, \mathbf{uG} + \mathbf{e})$ where $\mathbf{G} \in \mathbb{F}_q^{k \times n}$ and $\mathbf{u} \in \mathbb{F}_q^k$ are sampled uniformly at random over their domain and $\mathbf{e} \in \mathbb{F}_q^n$ over the words of weight t,*
- *Find: $\mathbf{e}$.*

This problem really corresponds to decoding the k-dimensional vector space $\mathcal{C}$ (*i.e.*, the code) generated by the rows of $\mathbf{G}$:

$$(1) \qquad \mathcal{C} \stackrel{\text{def}}{=} \left\{ \mathbf{uG} \colon \mathbf{u} \in \mathbb{F}_q^k \right\}.$$

We are given the noisy codeword $\mathbf{c} + \mathbf{e}$ where $\mathbf{c}$ belongs to $\mathcal{C}$ and we are asked to find the error $\mathbf{e}$ (or what amounts to the same, the original codeword $\mathbf{c}$). This problem for random codes has been studied for a long time and despite many efforts on this issue, the best algorithms are exponential in the codelength n in the regime where t and k are linear in n.

[1] INRIA

[2] LABORATOIRE LIX, ÉCOLE POLYTECHNIQUE, INSTITUT POLYTECHNIQUE DE PARIS, 1 RUE HONORÉ D'ESTIENNE D'ORVES, 91120 PALAISEAU CEDEX

[3] ATOS QUANTUM LAB

E-mail addresses: thomas.debris@inria.fr, maxime.remaud@atos.net, jean-pierre.tillich@inria.fr.

Usually this decoding problem is considered, for the Hamming weight $|\mathbf{e}| \stackrel{\text{def}}{=} \sharp\{i \in [\![1, n]\!], \quad e_i \neq 0\}$, in the regime where the code rate $R \stackrel{\text{def}}{=} \frac{k}{n}$ is fixed in $(0, 1)$ and $q = 2$, but there are also other interesting parameters for cryptographic applications. For instance, the Learning Parity with Noise problem (LPN) corresponds to $\mathsf{DP}(2, n, k, t)$ where n is the number of samples, k the length of the secret while the error is sampled according to a Bernoulli distribution of fixed rate t/n. As the number of samples in LPN is unlimited, this problem really corresponds to decoding a code of rate arbitrarily close to 0.

While the security of many code-based cryptosystems relies on the hardness of the decoding problem, it can also be based on finding a "short" codeword (as in [MTSB12] or in [AHI$^+$17, BLVW19, YZW$^+$19] to build collision resistant hash functions), a problem which is stated as follows.

Problem 2 ($\mathsf{SCP}(q, n, k, w)$). *The short codeword problem with parameters $q, n, k, w \in \mathbb{N}$ is defined as:*

- *Given: $\mathbf{H} \in \mathbb{F}_q^{(n-k) \times n}$ which is sampled uniformly at random,*
- *Find: $\mathbf{c} \in \mathbb{F}_q^n$ such that $\mathbf{H}\mathbf{c}^{\mathsf{T}} = \mathbf{0}$ and the weight of $\mathbf{c}$ belongs to $[\![1, w]\!]$.*

Here we are looking for a non-zero codeword $\mathbf{c}$ of weight $\leq w$ in the k-dimensional code $\mathcal{C}$ defined by the so-called parity-check matrix $\mathbf{H}$, namely:

$$\mathcal{C} \stackrel{\text{def}}{=} \left\{ \mathbf{c} \in \mathbb{F}_q^n \colon \mathbf{H}\mathbf{c}^{\mathsf{T}} = \mathbf{0} \right\}.$$

Decoding and looking for short codewords are problems that have been conjectured to be extremely close. They have been studied for a long time [Pra62, Ste88, Dum89, MMT11, BJMM12, MO15, BM18, CDMT22], and for instance in the regime of parameters where the code rate R is fixed in $(0, 1)$, the best algorithms for solving them are the same (namely Information Set Decoding). A reduction from decoding to the problem of finding short codewords is known but in an LPN context [AHI$^+$17, BLVW19, YZW$^+$19, DR22]. However, even in an LPN context, no reduction is known in the other direction. These problems can be viewed in some sense as a code version of the LWE and SIS problems respectively in lattice-based cryptography [Reg09]. Our contribution in this article is precisely to give the code-based version of this reduction, namely a quantum reduction from finding short codewords to decoding. This problem was open for quite some time. To simplify the statements, we will state it in the regime of parameters where the rate R is fixed in $(0, 1)$, but actually it also works in the LPN setting (but needs to be adapted in several places where we use exponential bounds in n).

There is a fundamental difficulty of reducing the research of low weight codewords to decoding a linear code which is due to the fact that the nature of these two problems is very different. Decoding concentrates on a region of parameters where there is typically just one solution, whereas finding low weight codewords concentrates on a region of parameters where there are solutions (and typically an exponential number of solutions). This makes these problems inherently very different. This was also the case for the reduction of SIS to LWE and the fact that we can have a reduction from one to another by looking for quantum reductions instead of classical reductions was really a breakthrough at that time.

Parameter range for DP and SCP. An important parameter for the reduction is the decoding distance parameter t. The largest value of t for which the decoding problem is ensured to have a unique solution is equal to $\left\lfloor \frac{d_{\min} - 1}{2} \right\rfloor$ where $d_{\min} \stackrel{\text{def}}{=} \min\{d(\mathbf{c}, \mathbf{c}') \colon \mathbf{c} \in \mathcal{C}, \ \mathbf{c}' \in \mathcal{C}, \ \mathbf{c} \neq \mathbf{c}'\}$ is the minimum distance of $\mathcal{C}$ (which depends of course on the metric $d(\cdot, \cdot)$ that is considered). Standard probabilistic arguments can be used to show that the minimum distance of a random linear code (code $\mathcal{C}$ obtained as in Equation (1) by a generator matrix $\mathbf{G}$ chosen uniformly at random in $\mathbb{F}_q^{k \times n}$)

is with very high probability equal, up to an additive constant, to the *Gilbert-Varshamov distance* $d_{\mathrm{GV}}(n,k)$ (or simply d_{GV} if there is no ambiguity). It is defined (for all translation invariant metrics) for a code of dimension k and length n, as the largest integer t for which

$$q^k B_t \leq q^n \tag{2}$$

where B_t is the size[(1)] of a ball of radius t. It is generally convenient to consider the normalized Gilbert-Varshamov distance defined as $\delta_{\mathrm{GV}}(n,k) \stackrel{\text{def}}{=} \frac{d_{\mathrm{GV}}(n,k)}{n}$. For the Hamming metric, we have

$$\delta_{\mathrm{GV}}(n,k) = h_q^{-1}\left(1 - \frac{k}{n}\right) + O\left(\frac{1}{n}\right) \quad \text{where}$$

$h_q(x) \stackrel{\text{def}}{=} -x \log_q\left(\frac{x}{q-1}\right) - (1-x)\log_q(1-x)$ and h_q^{-1} is its inverse, ranging over $\left[0, \frac{q-1}{q}\right]$.

This Gilbert-Varshamov distance also happens to quantify the region where we *typically* have unique decoding. More precisely, it turns out that for many metrics of interest, including the Hamming metric, the same probabilistic arguments also show that the solution to the decoding problem is unique with probability $1 - 2^{-\Omega(n)}$ and when n goes to infinity for fixed positive ε, $t \leq (1-\varepsilon)d_{\mathrm{GV}}(n,k)$.

The best algorithms for solving the decoding problem have exponential complexity in n as soon as t is linear in n and the code rate R is bounded away from 0 and 1. This is in contrast to the short codeword problem which becomes easy when the weight w is above a certain range. The reason is that it is easy to produce codewords of small weight by using the fact that the code is a vector space of dimension k. Thus we can just produce codewords with $k-1$ entries equal to 0 by solving a linear system which gives good candidates for having a small weight. It is straightforward that this strategy produces in polynomial time, for instance with the Hamming metric, codewords of weight $\approx \omega_{\text{easy}}(n,k)n$ where

$$\omega_{\text{easy}}(n,k) \stackrel{\text{def}}{=} \frac{q-1}{q}\left(1 - \frac{k}{n}\right) \tag{3}$$

Obtaining larger weights is also readily obtained by choosing only part of the $k-1$ entries to be equal to 0. It should be noted that below $\omega_{\text{easy}}(n,k)$ the best known algorithms for solving this problem have all exponential complexity for a fixed rate R and a fixed ratio $\omega = \frac{w}{n}$.

Regev's quantum reduction strategy adapted to coding theory. In [Reg05] (see also the extended version [Reg09]) Regev showed how to transform a random oracle solving the decoding problem in a lattice into a quantum algorithm outputting a rather small vector in the dual lattice. Our aim is to show here that the natural translation of this approach in coding theory gives an algorithm that outputs a rather small vector in the dual code. Roughly speaking Regev's approach relies on a fundamental result about the Fourier transform.

Proposition 1. *Consider an Abelian group G and a function $f : G \mapsto \mathbb{C}$ that is constant on the cosets of a subgroup H of G. Then the Fourier transform $\widehat{f}$ is constant on the dual subgroup $H^\perp$.*

This innocent looking fact, together with the fact that the quantum Fourier transform (QFT for short) can be performed in polylog time when the group G is Abelian, is arguably the key to several remarkable quantum algorithms solving in polynomial time the period finding in a vectorial Boolean function [Sim94], the factoring problem [Sho94] or the discrete logarithm problem [Sho94]. All of these problems can be rephrased in terms of the hidden Abelian subgroup problem, where one is given such a function f that is constant (and distinct) on the cosets of an unknown subgroup H and one is asked to recover H. This is achieved by

---

[(1)]Note that because of the translation invariance, this size does not depend on the center.

(i) creating the uniform superposition $\frac{1}{\sqrt{|G|}} \sum_{x \in G} |x\rangle |f(x)\rangle$,

(ii) measuring the second register and discarding it, yielding a quantum state of the form $\frac{1}{\sqrt{|H|}} \sum_{h \in H} |x + h\rangle$,

(iii) applying the QFT to it yielding a superposition of elements in the dual subgroup $H^\perp$ (and therefore gaining information on H in this way).

Proposition 1 is used in a similar way in Regev's reduction. Translating Regev's reduction in coding theory would use this framework by considering that the linear code $\mathcal{C}$ we want to decode plays the role of the aforementioned H. From now on we will assume that this code is of dimension k and length n over $\mathbb{F}_q$. The algorithm would basically look as follows for reducing the search of small codewords in the dual code $\mathcal{C}^\perp = \left\{ \mathbf{c}^\perp \in \mathbb{F}_q^n \colon \mathbf{c} \cdot \mathbf{c}^\perp = 0, \; \forall \mathbf{c} \in \mathcal{C} \right\}$ (where $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i y_i$ is the standard inner product in $\mathbb{F}_q^n$) to decoding errors of weight t in $\mathcal{C}$.

Step 1. Use a quantized version of the decoding algorithm to prepare the state

$$\frac{1}{\sqrt{Z}} \sum_{\mathbf{c} \in \mathcal{C}, \mathbf{e} \in \mathbb{F}_q^n} \pi_{\mathbf{e}} |\mathbf{c} + \mathbf{e}\rangle$$

where Z is a normalizing constant and $(|\pi_{\mathbf{e}}|^2)_{\mathbf{e}}$ is a probability distribution on errors that concentrate around the weight t we are able to decode. This is done

(i) by preparing first a superposition of codewords and errors,

$$\frac{1}{\sqrt{Z}} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{\mathbf{e} \in \mathbb{F}_q^n} \pi_{\mathbf{e}} |\mathbf{c}\rangle |\mathbf{e}\rangle \,,$$

(ii) then adding the second register to the first one to get the entangled state

$$\frac{1}{\sqrt{Z}} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{\mathbf{e} \in \mathbb{F}_q^n} \pi_{\mathbf{e}} |\mathbf{c} + \mathbf{e}\rangle |\mathbf{e}\rangle$$

(iii) and finally disentangling it thanks to a quantized version of the decoding algorithm, which from $\mathbf{c} + \mathbf{e}$ recovers $\mathbf{e}$ and subtracts it from the second register to get the state

$$\frac{1}{\sqrt{Z}} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{\mathbf{e} \in \mathbb{F}_q^n} \pi_{\mathbf{e}} |\mathbf{c} + \mathbf{e}\rangle |\mathbf{0}\rangle \,.$$

Step 2. Apply the QFT on $\mathbb{F}_q^n$ to obtain a superposition of elements $\mathbf{c}^\perp$ in the dual code

$$\sum_{\mathbf{c}^\perp \in \mathcal{C}^\perp} \alpha_{\mathbf{c}^\perp} |\mathbf{c}^\perp\rangle \,.$$

Step 3. Measure the register to output $\mathbf{c}^\perp$ of rather small norm in $\mathcal{C}^\perp$.

The second step is a direct consequence of Proposition 1. The last one raises the issue of whether or not the QFT concentrates the weight of the vector output by this algorithm on weights t' for which finding a codeword in $\mathcal{C}^\perp$ is not known to be easy, as it is the case for Regev's reduction on lattices equipped with the Euclidean metric.

On the difficulty of translating Regev's reduction to the Hamming metric. This thread of research has been pioneered by Yilei Chen [Che] and later on in [CV], where basically the following approach was taken. The natural analog in the Hamming metric case of the Gaussian noise model used in Regev's reduction [Reg09] is the q-ary symmetric channel. Its associated quantum state is given by

$$\left| \pi^{\mathrm{SC}} \right\rangle \overset{\text{def}}{=} \sum_{\mathbf{e} \in \mathbb{F}_q^n} (1-\tau)^{\frac{n-|\mathbf{e}|}{2}} \left(\frac{\tau}{q-1} \right)^{\frac{|\mathbf{e}|}{2}} |\mathbf{e}\rangle = \left(\sqrt{1-\tau} \, |0\rangle + \sum_{\alpha \in \mathbb{F}_q^*} \sqrt{\frac{\tau}{q-1}} \, |\alpha\rangle \right)^{\otimes n}$$

where $|\mathbf{e}|$ stands for the Hamming weight of $\mathbf{e}$, n for the length of $\mathbf{e}$ and τ is the crossover probability of the q-ary symmetric channel. Indeed, measuring such a state yields an error distributed like a q-ary symmetric channel of crossover probability τ. In both cases (be it for the Gaussian noise or the q-ary symmetric channel), the Fourier transform yields a dual noise which is again Gaussian or q-ary symmetric respectively and the quantum state corresponding to the error is a product state which considerably simplifies the computation. In the case of the q-ary symmetric noisy chanel, applying the QFT on $\left|\pi^{\mathrm{SC}}\right\rangle$ yields the quantum state

$$\sum_{\mathbf{e}\in\mathbb{F}_q^n}(1-\tau^\perp)^{\frac{n-|\mathbf{e}|}{2}}\left(\frac{\tau^\perp}{q-1}\right)^{\frac{|\mathbf{e}|}{2}}|\mathbf{e}\rangle = \left(\sqrt{1-\tau^\perp}\,|0\rangle + \sum_{\alpha\in\mathbb{F}_q^*}\sqrt{\frac{\tau^\perp}{q-1}}\,|\alpha\rangle\right)^{\otimes n} \stackrel{\mathrm{def}}{=} \widehat{|\pi^{\mathrm{SC}}\rangle}$$

where (see Fact 1)

$$\tau^\perp \stackrel{\mathrm{def}}{=} \frac{\left(\sqrt{(q-1)(1-\tau)}-\sqrt{\tau}\right)^2}{q}.$$

This new quantum state represents a q-ary symmetric channel of parameter $\tau^\perp$. If we measure $\left|\pi^{\mathrm{SC}}\right\rangle$ we get relative weights $\approx \tau$ whereas if we measure $\widehat{|\pi^{\mathrm{SC}}\rangle}$ we get relative weights around $\tau^\perp$.

It would thus be tempting to conclude that the "ideal" version of the algorithm presented above will output dual codewords (in Step 3) of relative Hamming weight $\approx \tau^\perp < \omega_{\mathrm{easy}}$, *i.e.*, in the regime where there is a chance that it is difficult to produce such words. However, this natural approach runs into the following problem. The parameter τ of the Bernoulli noise has to be chosen so that the typical error weight τn is equal to or slightly below the weight t we can decode. Such a τ is therefore at most the relative Gibert-Varshamov δ_{GV}. However [CV] proved that in this case the most likely relative weight we measure at Step 3 is *typically zero* if $\tau^\perp < \omega_{\mathrm{easy}}$. In other words, the straightforward application of Regev's approach to coding theory fails to give a useful reduction.

We will give in Remark 1 another explanation for the failure of this approach. It can be summarized by saying that the Bernoulli noise model is not concentrated enough on its typical weight τn.

Our approach. To tackle this issue, it would therefore be natural to choose the most concentrated noise model on the weight t, namely:

$$\left|\pi^{\mathrm{unif}}\right\rangle = \sum_{\mathbf{e}\,:\,|\mathbf{e}|=t}\frac{1}{\sqrt{S_t}}\,|\mathbf{e}\rangle$$

where S_t is the cardinality of the sphere of radius t in the Hamming metric, *i.e.*, $S_t = (q-1)^t\binom{n}{t}$. Understanding of which weight w is the outcome after measuring the state $\sum_{\mathbf{c}^\perp\in\mathcal{C}^\perp}\alpha_{\mathbf{c}^\perp}\left|\mathbf{c}^\perp\right\rangle$ in Step 3 is more difficult in the constant weight error model than in the Bernoulli noise model. In particular, it involves properties of Krawtchouk polynomials. However, it can be shown that when $\omega \stackrel{\mathrm{def}}{=} \frac{w}{n}$ lies in a whole interval $\left[\tau^\perp, \tau_+^\perp\right]$ where $\tau_+^\perp \stackrel{\mathrm{def}}{=} \frac{\left(\sqrt{(q-1)(1-\tau)}+\sqrt{\tau}\right)^2}{q}$, we have many points where the probability of measuring a word of weight w is actually $\frac{1}{\mathsf{poly}(n)}$. The "dual" weight distribution is not really concentrated on a single value but spread over a large interval. This would provide a useful reduction when we use a decoding algorithm that succeeds on a non-negligible set of inputs.

Unfortunately, to be relevant in a cryptographic context, we must consider the case where decoding succeeds only for a potentially very low probability ε, the aim being to turn our decoding algorithm into an algorithm that produces a low weight codeword from the dual with some probability $\mathsf{poly}(\varepsilon)$. This cannot be obtained with the uniform distribution on the sphere of radius t. Indeed, the "ideal" version of the algorithm we presented before (where we assume we always succeed with our decoding algorithm) describes a state we obtain in Step 2 that is not completely orthogonal

(the scalar product is bounded from below by a quantity $\mathsf{poly}(\varepsilon)$) to the "real" state after applying this approximate decoding process and the QFT. If we were to measure this state directly (starting from the uniform noise model over the sphere of radius t), we would only be sure to measure a word of some relative weight lying in the interval $\left[\tau^\perp, \tau_+^\perp\right]$ with probability $\mathsf{poly}(\varepsilon)$, since the "ideal" state concentrates its relative weight distribution in all this interval. In this way, we cannot ensure the measurement of a dual codeword of smallest possible relative weight, namely $\tau^\perp$, it could be $1/2 \in \left[\tau^\perp, \tau_+^\perp\right]$. We really need here a distribution that is rather sharply concentrated around the decoding radius of our decoding algorithm, but whose Fourier transform is also sharply concentrated around a certain weight.

So far, two noise models have been considered for the reduction to work, each with an advantage and a drawback:

- the q-ary symmetric noise $\left|\pi^{\mathrm{SC}}\right\rangle$ is not concentrated enough on its typical weight τn *but* its dual noise is sufficiently concentrated on $\tau^\perp n$,
- the uniform noise $\left|\pi^{\mathrm{unif}}\right\rangle$ on the sphere of radius τn is sufficiently concentrated *but* its dual noise is spread out on the whole interval $\left(\tau^\perp n, \tau_+^\perp n\right)$.

Interestingly, the issues with these two distributions are opposite. Fortunately, it turns out that we have a natural noise model to get a best-of-both-worlds model: *truncating* the q-ary symmetric noisy channel. More precisely, consider the following noise model (for some small enough constant $\varepsilon > 0$),

$$\left|\pi^{\mathrm{Trunc}}\right\rangle = \frac{1}{\sqrt{N}} \sum_{\substack{\mathbf{e}: \\ |\mathbf{e}| \in [(1-\varepsilon)t, (1+\varepsilon)t]}} (1 - \tau)^{\frac{n - |\mathbf{e}|}{2}} \left(\frac{\tau}{q - 1}\right)^{\frac{|\mathbf{e}|}{2}} |\mathbf{e}\rangle$$

where N is a normalizing factor. This noise model solves our above issues with $\left|\pi^{\mathrm{SC}}\right\rangle$ and $\left|\pi^{\mathrm{unif}}\right\rangle$ because it verifies our two constraints for the reduction to work:

- (i) its weight distribution is sufficiently concentrated around τn,
- (ii) its dual noise after applying the Fourier transform is concentrated on the relative weight $\tau^\perp n$.

Contrary to (i), assertion (ii) may seem unclear. It relies on the following equality as we will show in Lemma 11

$$\left\| \left|\pi^{\mathrm{Trunc}}\right\rangle - \left|\pi^{\mathrm{SC}}\right\rangle \right\| = 2^{-\Omega(n)}$$

where $\|\cdot\|$ stands for the norm of the Hilbert space in which the quantum states are embedded. Therefore, applying the Fourier transform (which is an isometry for $\|\cdot\|$) on $\left|\pi^{\mathrm{Trunc}}\right\rangle$ will yield a quantum state which is $2^{-\Omega(n)}$-close of $\widehat{\left|\pi^{\mathrm{SC}}\right\rangle}$.

With our approach and the truncated q-ary symmetric channel, we transform through the QFT a decoding algorithm correcting τn errors into an algorithm outputting with non-negligible probability words of weight $\approx \tau^\perp n$ in the dual code. The distance $\tau^\perp$ is clearly a decreasing function of τ and the issue is now whether or not there exists a $\tau < \delta_{\mathrm{GV}}(n, k)$ (this is the biggest value for which we can hope that decoding is successful with probability $1 - o(1)$) such that $\tau^\perp < \omega_{\mathrm{easy}}(n, n - k)$ (here we want to find short codewords in the dual code $\mathcal{C}^\perp$ which is of dimension $n - k$), since this would yield a useful reduction. It turns out that in many cases we have to choose $\tau > \delta_{\mathrm{GV}}(n, k)/2$, meaning that we are not in the regime where decoding necessarily has at most one solution. This complicates the proof of the reduction somewhat since with a quantized version of the decoding algorithm, we will not be able to produce at Step 1 the state $\frac{1}{\sqrt{Z}} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{\mathbf{e} \in \mathbb{F}_q^n} \pi_{\mathbf{e}} |\mathbf{c} + \mathbf{e}\rangle$ (since decoding fails for some $\mathbf{e}$) but we will show that as long as $\tau < \delta_{\mathrm{GV}}(n, k)$, we will get a state close to it. This will be enough for our purpose.

By putting all these ingredients together, we are able to prove the following result.

Theorem (informal). *The short codeword problem* $\mathsf{SCP}(q, n, n - k, w)$ *reduces to the decoding problem* $\mathsf{DP}(q, n, k, t)$ *for* $w = \tau^\perp n + O(1)$ *where*

$$\tau \overset{def}{=} \frac{t}{n} \qquad and \qquad \tau^\perp \overset{def}{=} \frac{\left(\sqrt{(q-1)(1-\tau)} - \sqrt{\tau}\right)^2}{q}.$$

It will turn out that for $q = 2$ (see Section 4) we can find for any rate $R = \frac{k}{n}$ in $(0, 1)$ a $t < d_{\mathrm{GV}}(n, k)$ for which the corresponding w is below $\omega_{\mathrm{easy}}(n, n - k)n$ (the reduction is useful in this case). It also corresponds to parameter ranges that are relevant for certain cryptographic applications (see [AFS05] whose security relies on the SCP problem in a parameter range which is covered by our reduction and [Ste93] which is an identification scheme whose security actually relies on the decoding problem just below the Gilbert-Varshamov distance). Unfortunately, this is not true anymore when $q \geq 5$, where there is always a range for R for which w is above $\omega_{\mathrm{easy}}(n, n - k)n$, for any choice of $t < d_{\mathrm{GV}}(n, k)$: the reduction becomes useless in this case. Roughly speaking, when q grows, the Hamming metric gets coarser (we have only $n + 1$ different values for the metric on $\mathbb{F}_q^n$, whereas the size of the ambient space gets bigger) and this results in the range of values of R for which this reduction is useful becoming smaller.

Considering other metrics. The whole approach we have followed here (properly choosing the error distribution and going beyond the unique decoding radius for decoding if necessary) can of course be adapted to other metrics. It is easy for instance to apply it to the rank metric which is becoming increasingly popular in code-based cryptography, see for instance [ABD+19, AAB+19, BCG+19, BGHM20]. This metric is even coarser than the Hamming metric: on $\mathbb{F}_q^{m \times n}$ there are only $1 + \min(m, n)$ different values for the rank weight (given a matrix, it is defined as its rank). In this case, as we will see, the reduction is always useless (*i.e.*, reduces to weights which are always easy to produce for a random linear code).

2. Notation and some useful facts

General Notation. For a and b integers with $a \leq b$, we denote by $[\![a, b]\!]$ the set of integers $\{a, a + 1, \dots, b\}$. We extend this notation when a and b are not integers to the set of integers in $[a, b]$. Vectors are in *row notation* and they will be written with bold letters (such as $\mathbf{e}$). Uppercase bold letters are used to denote matrices (such as $\mathbf{H}$). Notation $\mathscr{S}_t$ is the sphere of radius t around 0 in $\mathbb{F}_q^n$ (for a metric $|\cdot|$ that will be clear from the context) and S_t is its cardinality. In all this paper, $\mathsf{poly}(n)$ denotes a quantity which is an $O(n^a)$ for some constant a.

Subspaces and Gaussian binomial coefficient. $F \leq \mathbb{F}_q^n$ denotes that F is a subspace of $\mathbb{F}_q^n$. When $\mathbf{E}$ is a matrix, we denote its rank by $|\mathbf{E}|$.

The following asymptotic expressions will be extremely useful. For the rank metric, S_ℓ is equal to the number of matrices in $\mathbb{F}_q^{m \times n}$ of rank ℓ. From this interpretation we obtain

$$S_\ell = \prod_{j=0}^{\ell-1}(q^m - q^j)\begin{bmatrix} n \\ \ell \end{bmatrix}_q \quad \text{with} \quad \begin{bmatrix} n \\ t \end{bmatrix}_q = \begin{cases} \displaystyle\prod_{i=0}^{t-1} \frac{q^n - q^i}{q^t - q^i} & \text{if } t \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

It follows by using the fact that $\prod_{i=1}^{\infty}(1 - q^{-i})$ is some constant depending on q, that as $n \to +\infty$

$$\begin{bmatrix} n \\ \ell \end{bmatrix}_q = \Theta\left(q^{\ell(n-\ell)}\right), \tag{4}$$

$$S_\ell = \Theta\left(q^{\ell(m+n-\ell)}\right). \tag{5}$$

From (5) we deduce that

$$(6) \qquad \frac{S_{u+1}}{S_u} = \Theta\left(\frac{q^{(u+1)(m+n-u-1)}}{q^{u(m+n-u)}}\right) = \Theta\left(q^{m+n-2u-1}\right).$$

Quantum Fourier Transform (QFT). The Quantum Fourier Transform $\widehat{|\psi\rangle}$ of a state $|\psi\rangle \overset{\text{def}}{=} \sum_{\mathbf{x}\in\mathbb{F}_q^n} \alpha_{\mathbf{x}} |\mathbf{x}\rangle$ is defined by using the characters $\chi_{\mathbf{y}}$ of the additive group $\mathbb{F}_q^n$ (there are as many characters as there are elements in $\mathbb{F}_q^n$ and we assume that the characteristic of $\mathbb{F}_q$ is the prime p with $q = p^s$) as

$$\widehat{|\psi\rangle} \overset{\text{def}}{=} \frac{1}{\sqrt{q^n}} \sum_{\mathbf{y}\in\mathbb{F}_q^n} \widehat{\alpha_{\mathbf{y}}} |\mathbf{y}\rangle \qquad \text{with} \qquad \widehat{\alpha_{\mathbf{y}}} \overset{\text{def}}{=} \sum_{\mathbf{x}\in\mathbb{F}_q^n} \alpha_{\mathbf{x}} \chi_{\mathbf{y}}(\mathbf{x})$$

where

$$\chi_{\mathbf{y}}(\mathbf{x}) \overset{\text{def}}{=} e^{\frac{2i\pi \operatorname{Tr}(\mathbf{x}\cdot\mathbf{y})}{p}}, \quad \text{with}$$

$$\mathbf{x}\cdot\mathbf{y} \overset{\text{def}}{=} \sum_{i=1}^{n} x_i y_i \quad \text{with } \mathbf{x} = (x_i)_{i=1}^n \text{ and } \mathbf{y} = (y_i)_{i=1}^n \quad \text{and}$$

$$\operatorname{Tr}(a) \overset{\text{def}}{=} a + a^p + a^{p^2} + \cdots + a^{p^{s-1}}.$$

A useful computation is the QFT of a state representing the q-ary symmetric channel of crossover probability τ:

Fact 1. *Let $\left|\pi^{SC}\right\rangle \overset{def}{=} \left(\sqrt{1-\tau}|0\rangle + \sum_{\alpha\in\mathbb{F}_q^*} \sqrt{\frac{\tau}{q-1}} |\alpha\rangle\right)^{\otimes n}$, where $\tau \in \left[0, \frac{q-1}{q}\right]$, then*

$$(7) \qquad \widehat{|\pi^{SC}\rangle} = \left(\sqrt{1-\tau^{\perp}}|0\rangle + \sum_{\alpha\in\mathbb{F}_q^*} \sqrt{\frac{\tau^{\perp}}{q-1}} |\alpha\rangle\right)^{\otimes n}, \quad where$$

$$(8) \qquad \tau^{\perp} \overset{def}{=} \frac{\left(\sqrt{(q-1)(1-\tau)} - \sqrt{\tau}\right)^2}{q}.$$

Proof. Let $|\psi\rangle \overset{\text{def}}{=} \sqrt{1-\tau}|0\rangle + \sum_{\alpha\in\mathbb{F}_q^*} \sqrt{\frac{\tau}{q-1}} |\alpha\rangle$. Then, it is readily verified that

$$\widehat{|\pi^{SC}\rangle} = \left(\widehat{|\psi\rangle}\right)^{\otimes n} \qquad \text{where} \qquad \widehat{|\psi\rangle} = \frac{1}{\sqrt{q}}\left(\beta_0 + \sum_{y\in\mathbb{F}_q^*} \beta_y |y\rangle\right) \qquad \text{with}$$

$$\beta_0 = \sqrt{1-\tau} + (q-1)\sqrt{\frac{\tau}{q-1}} \qquad \text{and} \qquad \beta_y = \sqrt{1-\tau} + \sqrt{\frac{\tau}{q-1}}\sum_{x\in\mathbb{F}_q^*} \chi_y(x).$$

It is easy to verify that for any $y \in \mathbb{F}_q^*$ we have for any $x_0 \in \mathbb{F}_q^*$:

$$\sum_{x\in\mathbb{F}_q} \chi_y(x) = \sum_{x\in\mathbb{F}_q} \chi_y(x_0\cdot x) = \chi_y(x_0)\sum_{x\in\mathbb{F}_q} \chi_y(x)$$

Since there exists $x_0 \in \mathbb{F}_q^*$ such that $\chi_y(x_0) \neq 1$, we deduce that $\sum_{x \in \mathbb{F}_q} \chi_y(x) = 0$. Now since $\sum_{x \in \mathbb{F}_q} \chi_y(x) = 1 + \sum_{x \in \mathbb{F}_q^*} \chi_y(x)$, we have $\sum_{x \in \mathbb{F}_q^*} \chi_y(x) = -1$ for any y in $\mathbb{F}_q^*$ and therefore

$$
\begin{aligned}
\frac{\beta_y}{\sqrt{q}} &= \sqrt{\frac{1-\tau}{q}} - \sqrt{\frac{\tau}{q(q-1)}} \\
&= \sqrt{\frac{(q-1)(1-\tau)}{q(q-1)}} - \sqrt{\frac{\tau}{q(q-1)}} \\
&= \frac{\sqrt{\frac{(q-1)(1-\tau)}{q}} - \sqrt{\frac{\tau}{q}}}{\sqrt{q-1}} \\
&= \frac{\sqrt{\tau^\perp}}{\sqrt{q-1}}.
\end{aligned}
$$

It is then clear that $\frac{\beta_0}{\sqrt{q}} = \sqrt{1-\tau^\perp}$ from $\sum_{y \in \mathbb{F}_q} |\beta_y|^2 = 1$, concluding the proof. $\qquad\square$

The dual code $\mathcal{C}^\perp$ of a linear code $\mathcal{C}$ over $\mathbb{F}_q$ is easily seen to be defined equivalently from the characters or from the inner product $\mathbf{x} \cdot \mathbf{y}$ as follows:

$$
\begin{aligned}
\mathcal{C}^\perp &\stackrel{\text{def}}{=} \left\{ \mathbf{y} \in \mathbb{F}_q^n : \forall \mathbf{c} \in \mathcal{C}, \ \chi_{\mathbf{y}}(\mathbf{c}) = 1 \right\} \\
&= \left\{ \mathbf{y} \in \mathbb{F}_q^n : \forall \mathbf{c} \in \mathcal{C}, \ \mathbf{y} \cdot \mathbf{c} = 0 \right\}.
\end{aligned}
$$

3. Quantum Reduction from Sampling Short Codewords to Decoding

3.1. **A general result.** We assume here that we have a probabilistic algorithm $\mathscr{A}$ that solves (sometimes) the decoding problem at distance t. Its inputs are a generator matrix $\mathbf{G} \in \mathbb{F}_q^{k \times n}$ of a code $\mathcal{C} \subseteq \mathbb{F}_q^n$ (*i.e.*, $\mathcal{C} = \{ \mathbf{u}\mathbf{G} : \mathbf{u} \in \mathbb{F}_q^k \}$) and a noisy codeword $\mathbf{c} + \mathbf{e}$ where $\mathbf{c}$ belongs to $\mathcal{C}$. We denote by $\mathbf{r} \in \mathbb{F}_2^\ell$ the internal coins of $\mathscr{A}$. It outputs with a certain probability ε, the "right" $\mathbf{e}$ when being fed with $\mathbf{c} + \mathbf{e}$ where $\mathbf{c}$ and $\mathbf{e}$ are uniformly chosen at random in $\mathcal{C}$ and among the errors of weight t respectively:

$$
(9) \qquad \varepsilon \stackrel{\text{def}}{=} \mathbb{P}_{\mathbf{G},\mathbf{c},\mathbf{e},\mathbf{r}} \left(\mathscr{A}(\mathbf{G}, \mathbf{c}+\mathbf{e}, \mathbf{r}) = \mathbf{e} \right).
$$

The quantum reduction starts by building the initial superposition

$$
\frac{1}{\sqrt{2^\ell q^k}} \sum_{\mathbf{e} \in \mathbb{F}_q^n} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{\mathbf{r} \in \mathbb{F}_2^\ell} \pi_{\mathbf{e}} \ket{\mathbf{e}} \ket{\mathbf{c}} \ket{\mathbf{r}}
$$

where $\ket{\pi} \stackrel{\text{def}}{=} \sum_{\mathbf{e} \in \mathbb{F}_q^n} \pi_{\mathbf{e}} \ket{\mathbf{e}}$ is some quantum superposition of errors. In practice, we will voluntarily omit to write the internal coins of $\mathscr{A}$ for the sake of readability and will therefore start from

$$
\frac{1}{\sqrt{q^k}} \sum_{\mathbf{e} \in \mathbb{F}_q^n} \sum_{\mathbf{c} \in \mathcal{C}} \pi_{\mathbf{e}} \ket{\mathbf{e}} \ket{\mathbf{c}}
$$

For the same reason, we will write $\varepsilon = \mathbb{P}_{\mathbf{G},\mathbf{c},\mathbf{e}} \left(\mathscr{A}(\mathbf{G}, \mathbf{c}+\mathbf{e}) = \mathbf{e} \right)$.

The quantum algorithm that gives the reduction can then be described as follows.

Algorithm of the quantum reduction.

Initial state preparation	$=$	$\dfrac{1}{\sqrt{q^k}} \displaystyle\sum_{\mathbf{e}\in\mathbb{F}_q^n}\sum_{\mathbf{c}\in\mathcal{C}} \pi_{\mathbf{e}} \ket{\mathbf{e}}\ket{\mathbf{c}}$
adding $\mathbf{e}$ to $\mathbf{c}$:	$\mapsto$	$\dfrac{1}{\sqrt{q^k}} \displaystyle\sum_{\mathbf{e}\in\mathbb{F}_q^n}\sum_{\mathbf{c}\in\mathcal{C}} \pi_{\mathbf{e}} \ket{\mathbf{e}}\ket{\mathbf{c}+\mathbf{e}}$
(10) applying $\mathscr{A}$:	$\overset{\mathscr{A}}{\mapsto}$	$\dfrac{1}{\sqrt{q^k}} \displaystyle\sum_{\mathbf{e}\in\mathbb{F}_q^n}\sum_{\mathbf{c}\in\mathcal{C}} \pi_{\mathbf{e}} \ket{\mathbf{e}-\mathscr{A}(\mathbf{G},\mathbf{c}+\mathbf{e})}\ket{\mathbf{c}+\mathbf{e}} \overset{\text{def}}{=} \ket{\psi_{\mathscr{A}}}$
(11) QFT on the 2nd register:	$\mapsto$	$\widehat{\ket{\psi_{\mathscr{A}}}}$
(12) measuring the whole state:	$\mapsto$	$\ket{\mathbf{e}}\ket{\mathbf{c}^{\perp}}$

We will now give a general theorem about an algorithm of this kind and will show that it produces a codeword of the dual code $\mathcal{C}^{\perp}$ of some weight u with probability $\mathsf{poly}(\varepsilon)$ when certain conditions are met.

Theorem 1. *Assume that $\ket{\pi}$ is radial and non-negative, i.e., $\pi_{\mathbf{e}} = f(|\mathbf{e}|)$ for some non-negative function f. Assume that $\widehat{\ket{\pi}} = \sum_{\mathbf{e}\in\mathbb{F}_q^n} \widehat{\pi_{\mathbf{e}}}\ket{\mathbf{e}}$ is radial too*[(2)] *and let $\widehat{f}(w) = \widehat{\pi_{\mathbf{e}}}$ for any element $\mathbf{e}$ of $\mathbb{F}_q^n$ of weight w. Furthermore, assume that there exists an interval $\mathscr{W} \subseteq \llbracket 0,n \rrbracket$ such that:*

(C1)   *(Concentration of π)* $\qquad\qquad\qquad\qquad\qquad\qquad \dfrac{\braket{\pi|\mathbf{1}}^2}{q^{n-k}} = 2^{-\Omega(n)},$

(C2)   *(Exponentially many dual codewords of weight $u \in \mathscr{W}$)* $\qquad \displaystyle\sum_{u\in\mathscr{W}} \dfrac{q^k}{S_u} = 2^{-\Omega(n)},$

(C3)   *(Concentration of the dual distribution $\widehat{\pi}$ on $\mathscr{W}$)* $\qquad \displaystyle\sum_{u\in\mathscr{W}} S_u \left|\widehat{f}(u)\right|^2 = 1 - 2^{-\Omega(n)}$

with $\ket{\mathbf{1}}$ being the (unnormalized) superposition of errors : $\ket{\mathbf{1}} \overset{def}{=} \sum_{\mathbf{e}\in\mathbb{F}_q^n}\ket{\mathbf{e}}$.

Suppose that there exists an algorithm $\mathscr{A}$ solving the decoding problem $\mathsf{DP}(q,n,k,t)$ with success probability ε. Then, there exists a quantum algorithm which takes as input a generator matrix $\mathbf{G} \in \mathbb{F}_q^{k\times n}$ of $\mathcal{C}$ and outputs a codeword of weight $u \in \mathscr{W}$ in $\mathcal{C}^{\perp}$ with probability greater than $\dfrac{p_t^2 \varepsilon^3}{16} - O(p_t^4 \varepsilon^5) - 2^{-\Omega(n)} - O\left(q^{-\min(k,n-k)}\right)$ where $p_t \overset{def}{=} \sum_{\mathbf{e}\,:\,|\mathbf{e}|=t} |\pi_{\mathbf{e}}|^2 = S_t f(t)^2$.

Remark 1.

- *We will use this theorem for Hamming and rank metrics, but it can be applied to* any *metric for which the Fourier transform is radially preserving.*
- *When $\pi_{\mathbf{e}}$ is non-negative, Condition (C1) basically requires the probability distribution on $\mathbb{F}_q^n$, $\mu \overset{def}{=} (\pi_{\mathbf{e}}^2)_{\mathbf{e}\in\mathbb{F}_q^n}$, to be sufficiently concentrated. This quantity can be expressed as $q^k(1 - H^2(\mu, U))^2$, where U stands for the uniform distribution over $\mathbb{F}_q^n$ and $H(\mathbf{p},\mathbf{q}) \overset{def}{=} \sqrt{1 - \sum_i \sqrt{p_i q_i}}$ is the Hellinger distribution between two probability distributions $\mathbf{p}$ and $\mathbf{q}$ defined over a same probability space.*
  (1) It is clearly maximal for the uniform probability distribution over $\mathbb{F}_q^n$, and

---

[(2)]In other words, we assume that the Fourier transform is radially preserving. This property depends on the characters chosen to define the Fourier transform and the metric. Recall that a radial function is a function which is constant on spheres centered around 0. This property clearly holds for functions $f\colon \mathbb{F}_q^n \to \mathbb{C}$ with the characters chosen here and the Hamming metric. We give this more general statement in order to apply it in other cases of interest, for instance the rank metric.

(2) on the other hand, when considering the Hamming metric and $|\pi\rangle = \left|\pi^{SC}\right\rangle$, we have

$$\frac{\left\langle\pi^{SC}|\mathbf{1}\right\rangle^2}{q^{n-k}} \;=\; \frac{q^k}{q^n}\left|\sum_{\mathbf{y}\in\mathbb{F}_q^n}\chi_{\mathbf{y}}(\mathbf{0})\pi_{\mathbf{y}}\right|^2 \;=\; q^k\left|\widehat{f}(0)\right|^2 \;=\; q^k(1-\tau^\perp)^n.$$

It can be verified that there is no way to choose τ such that at the same time:
- *(i) $\tau n \leq d_{\mathrm{GV}}(n,k)$ (otherwise there is no hope to decode correctly most of the time),*
- *(ii) $\tau^\perp \leq \omega_{\mathrm{easy}}$ (otherwise finding codewords in $\mathcal{C}^\perp$ of weight $\tau^\perp n$ is easy),*
- *(iii) $q^k(1-\tau^\perp)^n = o(1)$.*

The quantity $\frac{\left\langle\pi^{SC}|\mathbf{1}\right\rangle^2}{q^{n-k}}$ is just too big, or in other words, the distribution of μ is too much spread out and not concentrated enough on its typical weight.

- *Condition ($C2$) expresses that u lies in a subset of values for which a random $[n, n-k]$-code has an exponential expected number of codewords. Indeed, the expected number of codewords of weight u is equal to $\frac{S_u}{q^k}$.*

- *Finally, Condition ($C3$) expresses that the dual probability distribution $\widehat{\mu} \stackrel{def}{=} (|\widehat{\pi}_{\mathbf{e}}|^2)_{\mathbf{e}\in\mathbb{F}_q^n}$ is almost completely supported on $\mathcal{W}$ up to an exponentially small vanishing term.*

3.2. **Outline of the proof of Theorem 1.** Let us first give a general outline of the proof before detailing each step.

Step 1. We prove that after applying $\mathscr{A}$ in the reduction, $|\psi_{\mathscr{A}}\rangle$ is close enough to the "disentangled" state

$$(13) \qquad |\psi_{\mathrm{ideal}}\rangle \stackrel{\mathrm{def}}{=} \frac{1}{\sqrt{Z}}\sum_{\mathbf{e}\in\mathbb{F}_q^n}\sum_{\mathbf{c}\in\mathcal{C}}\pi_{\mathbf{e}}\,|\mathbf{0}_n\rangle\,|\mathbf{c}+\mathbf{e}\rangle$$

where Z is a normalizing constant.

Step 2. We then analyze the effect of the QFT on the "ideal state" $|\psi_{\mathrm{ideal}}\rangle$ and a subsequent measurement of it. We namely prove that measuring it produces a codeword $\mathbf{c}^\perp \in \mathcal{C}^\perp$ of weight u with probability $\left|\widehat{f}(u)\right|^2$, up to a normalizing factor.

Step 3. Then we prove that the number of codewords of weight u in $\mathcal{C}^\perp$ is typically very close to $\frac{S_u}{q^k}$. With Step 2 and the assumptions of Theorem 1, we infer that the probability of observing a dual codeword of weight in the set $\mathcal{W}$ after measuring $\widehat{|\psi_{\mathrm{ideal}}\rangle}$ is exponentially close to 1.

Step 4. We upper-bound the statistical distance between the probability distribution of the states after measuring $\widehat{|\psi_{\mathscr{A}}\rangle}$ and $\widehat{|\psi_{\mathrm{ideal}}\rangle}$ respectively by using Step 1 and the properties of the trace distance given in Fact 2 below.

Let us give more details about these steps.

Step 1. For this purpose, we use the trace distance between quantum states (as in [SSTX09] where this has been used in the lattice setting). It is defined as follows :

$$(14) \qquad D_{\mathrm{tr}}(|\phi\rangle, |\psi\rangle) \stackrel{\mathrm{def}}{=} \sqrt{1 - |\langle\phi|\psi\rangle|^2}.$$

This distance meets the following properties that will prove useful in our context:

Fact 2.

(I) *It can never increase after a quantum evolution [NC16, §9,Th. 9.1];*

(II) *The pair of probability distributions (p_m, q_m) of the measurement outcome m of any quantum measurement performed on the pair of states $(|\phi\rangle, |\psi\rangle)$ satisfies* [NC16, §9,Th. 9.2]

(15) $$D_{\mathrm{stat}}(p_m, q_m) \leq D_{\mathrm{tr}}(|\phi\rangle, |\psi\rangle)$$

where D_{stat} is the statistical distance (also called the total variation distance) between two probability distributions. It is defined by:

$$D_{\mathrm{stat}}(p, q) \stackrel{def}{=} \frac{1}{2} \sum_{x \in \mathscr{X}} |p(x) - q(x)|$$

where p and q are two discrete probability distributions on $\mathscr{X}$.

With this notion we can prove that

Proposition 2. *With probability greater than $1 - \frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}} - O\left(q^{-\min(k,n-k)}\right)$ over the choices of $\mathbf{G}$ we have:*

$$D_{\mathrm{tr}}(|\psi_{\mathscr{A}}\rangle, |\psi_{ideal}\rangle) \leq \sqrt{1 - \frac{p_t^2}{2}\varepsilon_{\mathbf{G}}^2},$$

where $\varepsilon_{\mathbf{G}}$ is the probability that $\mathscr{A}$ returns the right error $\mathbf{e}$ when the input matrix is $\mathbf{G}$, i.e.,

(16) $$\varepsilon_{\mathbf{G}} \stackrel{def}{=} \mathbb{P}_{\mathbf{c},\mathbf{e}}\left(\mathscr{A}(\mathbf{G}, \mathbf{c} + \mathbf{e}) = \mathbf{e}\right).$$

The proof of this result follows immediately from three lemmas (whose proof is in Appendix A.1). The first one bounds the trace distance in terms of $\varepsilon_{\mathbf{G}}$ and Z, and the second one gives a tight upper-bound on the expected value of Z for a related probabilistic model. The latter is used to derive the third one which is fundamental and states that it is very unlikely for Z to be much greater than the "natural" constant q^k:

Lemma 1. *We have:*

$$D_{\mathrm{tr}}(|\psi_{\mathscr{A}}\rangle, |\psi_{ideal}\rangle) \leq \sqrt{1 - \frac{q^k p_t^2}{Z}\varepsilon_{\mathbf{G}}^2}.$$

Lemma 2. *Assume that $\mathcal{C}$ is chosen by uniformly drawing at random a parity-check matrix $\mathbf{H}$ for it. We have:*

(17) $$\mathbb{E}(Z) \leq q^k \left(1 + \frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}}\right).$$

Lemma 3. *Let $\eta > 0$. We have:*

$$\mathbb{P}_{\mathbf{G}}(Z > q^k(1 + \eta)) \leq \frac{1}{\eta}\frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}} + O\left(q^{-\min(k,n-k)}\right).$$

Proposition 2 immediately follows by using $\eta = 1$ in Lemma 3 and plugging this bound on Z in Lemma 1. The quantity q^k is the natural value for Z since it is what we can expect when all the $\mathbf{c} + \mathbf{e}$ terms (taking all $\mathbf{c}$ in $\mathcal{C}$ and all typical $\mathbf{e}$) are different. The constant Z increases precisely when there are many collisions for the $\mathbf{c} + \mathbf{e}$ terms. However, in this case, we do not expect to be able to solve the decoding problem anymore.

Step 2. More precisely, we prove that

Lemma 4. *If the Fourier transform is radially preserving, meaning that it transforms a radial function into a radial function, then after measuring $|\widehat{\psi_{ideal}}\rangle$ we obtain a state $|\mathbf{0}_n\rangle|\mathbf{c}^{\perp}\rangle$ with $\mathbf{c}^{\perp} \in \mathcal{C}^{\perp}$ of weight u with probability $\frac{q^{2k}}{Z}N_u^{\perp}\left|\widehat{f}(u)\right|^2$ where $\widehat{f}(u) \stackrel{def}{=} \widehat{\pi}_{\mathbf{e}}$ for an arbitrary $\mathbf{e}$ of weight u and $N_u^{\perp}$ is the number of codewords of weight u in $\mathcal{C}^{\perp}$.*

The proof is given in Appendix A.2.

Step 3. This step consists in quantifying how close to 1 the probability of observing a dual codeword of weight in the set $\mathcal{W}$ after measuring $|\widehat{\psi_{\text{ideal}}}\rangle$ is. More specifically, we have

Proposition 3. *Under the assumptions made in Theorem 1, the probability of obtaining a codeword $\mathbf{c}^\perp \in \mathcal{C}^\perp$ of weight $u \in \mathcal{W}$ when measuring $|\widehat{\psi_{ideal}}\rangle$ is $\geq 1 - \alpha(\pi)$ for a proportion $\geq 1 - \beta(\pi)$ of matrices $\mathbf{G}$, where:*

$$\alpha(\pi) \quad \stackrel{def}{=} \quad \sum_{u \in \mathcal{W}} \left(\frac{q^k}{S_u}\right)^{1/4} + \sqrt{\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}} - 2^{-\Omega(n)},$$

$$\beta(\pi) \quad \stackrel{def}{=} \quad (q-1) \sum_{u \in \mathcal{W}} \sqrt{\frac{q^k}{S_u}} + \sqrt{\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}} + O\left(q^{-\min(k, n-k)}\right).$$

This proposition is proved in Appendix A.3.

Step 4. We first prove the following point

Lemma 5. *Call $\mathcal{G}$ the set of "good matrices" $\mathbf{G} \in \mathbb{F}_q^{k \times n}$ that satisfy at the same time:*

   (i) $\varepsilon_\mathbf{G} \geq \varepsilon/2$ *(where ε and $\varepsilon_\mathbf{G}$ are defined in Equations (9) and (16)),*
   (ii) $Z \leq 2q^k$.

The proportion of good matrices is at least $\varepsilon/2 - \delta(\pi)$ where $\delta(\pi) \stackrel{def}{=} \frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}} + O\left(q^{-\min(k, n-k)}\right)$.

Proof. By definition,

$$\varepsilon = \frac{1}{q^{kn}} \sum_{\mathbf{G} \in \mathbb{F}_q^{k \times n}} \varepsilon_\mathbf{G}.$$

Let $\mathcal{B}$ be the set of matrices $\mathbf{G}$ that are not good, namely for which (a) $\varepsilon_\mathbf{G} < \varepsilon/2$ or (b) $Z > 2q^k$. By Lemma 3, the density of matrices verifying (b) is smaller than $\delta(\pi)$. Therefore,

$$\varepsilon \leq \frac{1}{q^{kn}} \sum_{\mathbf{G} \notin \mathcal{B}} 1 + \delta(\pi) \frac{\varepsilon}{2} \leq \frac{1}{q^{kn}} \sum_{\mathbf{G} \notin \mathcal{B}} 1 + \delta(\pi) + \frac{\varepsilon}{2}$$

which concludes the proof. $\qquad\square$

We use this lemma to prove that the statistical distance between the weight distributions obtained by measuring $|\widehat{\psi_\mathscr{A}}\rangle$ and $|\widehat{\psi_{\text{ideal}}}\rangle$ cannot be too far away:

Lemma 6. *Let P, respectively Q, be the distribution of the weights $|\mathbf{c}^\perp|$ of the state $|\mathbf{e}\rangle |\mathbf{c}^\perp\rangle$ obtained by measuring the state $|\widehat{\psi_\mathscr{A}}\rangle$, respectively $|\widehat{\psi_{ideal}}\rangle$. We have*

$$D_{\text{stat}}(P, Q) \leq 1 - \frac{p_t^2 \varepsilon^3}{16} + O\left(p_t^4 \varepsilon^5\right) + \delta(\pi).$$

Proof. Let,

$$P_\mathbf{G}(u) \quad \stackrel{def}{=} \quad \mathbb{P}_{\mathbf{c}, \mathbf{e}} \left(\text{measuring } |\mathbf{c}^\perp\rangle \text{ of weight } u \text{ in the 2nd register of } |\widehat{\psi_\mathscr{A}}\rangle \text{ for a code choice } \mathbf{G}\right)$$

$$Q_\mathbf{G}(u) \quad \stackrel{def}{=} \quad \mathbb{P}_{\mathbf{c}, \mathbf{e}} \left(\text{measuring } |\mathbf{c}^\perp\rangle \text{ of weight } u \text{ in the 2nd register of } |\widehat{\psi_{\text{ideal}}}\rangle \text{ for a code choice } \mathbf{G}\right)$$

We start the proof by noticing that

$$
\begin{aligned}
D_{\text{stat}}(P, Q) &= \frac{1}{2} \sum_u |P(u) - Q(u)| = \frac{1}{2} \sum_u \left| \sum_{\mathbf{G} \in \mathbb{F}_q^{k \times n}} \frac{1}{q^{kn}} (P_{\mathbf{G}}(u) - Q_{\mathbf{G}}(u)) \right| \\
&\leq \frac{1}{q^{kn}} \sum_{\mathbf{G} \in \mathbb{F}_q^{k \times n}} \frac{1}{2} \sum_u |P_{\mathbf{G}}(u) - Q_{\mathbf{G}}(u)| \\
&= \frac{1}{q^{kn}} \sum_{\mathbf{G} \in \mathbb{F}_q^{k \times n}} D_{\text{stat}}(P_{\mathbf{G}}, Q_{\mathbf{G}}) \\
&= \sum_{\mathbf{G} \in \mathscr{G}} \frac{D_{\text{stat}}(P_{\mathbf{G}}, Q_{\mathbf{G}})}{q^{kn}} + \sum_{\mathbf{G} \notin \mathscr{G}} \frac{D_{\text{stat}}(P_{\mathbf{G}}, Q_{\mathbf{G}})}{q^{kn}} \\
&\leq \sum_{\mathbf{G} \in \mathscr{G}} \frac{D_{\text{tr}}(|\psi_{\mathscr{A}}\rangle, |\psi_{\text{ideal}}\rangle)}{q^{kn}} + \sum_{\mathbf{G} \notin \mathscr{G}} \frac{1}{q^{kn}} \qquad \text{(by Equation 15)} \\
&\leq \sum_{\mathbf{G} \in \mathscr{G}} \frac{\sqrt{1 - \frac{p_t^2 \varepsilon^2}{4}}}{q^{kn}} + \sum_{\mathbf{G} \notin \mathscr{G}} \frac{1}{q^{kn}} \qquad \text{(by Proposition 2)} \\
&\leq \sqrt{1 - \frac{p_t^2 \varepsilon^2}{4}} (\varepsilon/2 - \delta(\pi)) + 1 - \varepsilon/2 + \delta(\pi) \qquad \text{(by Lemma 5)} \\
&\leq (\varepsilon/2 - \delta(\pi)) \left(1 - \frac{p_t^2 \varepsilon^2}{8} + O\left(p_t^4 \varepsilon^4\right) \right) + 1 - \varepsilon/2 + \delta(\pi) \\
&\leq 1 - \frac{p_t^2 \varepsilon^3}{16} + O\left(p_t^4 \varepsilon^5\right) + \delta(\pi)
\end{aligned}
$$

which concludes the proof. $\qquad\square$

3.3. **Proof of Theorem 1.** We are now ready to prove Theorem 1. By Proposition 3 we know that

$$
\sum_{u \in \mathscr{W}} Q(u) \geq (1 - \alpha(\pi))(1 - \beta(\pi)) \geq 1 - \alpha(\pi) - \beta(\pi).
$$

But now we have the following computation,

$$
\begin{aligned}
\sum_{u \in \mathscr{W}} P(u) &\geq \sum_{u \in \mathscr{W}} Q(u) - D_{\text{stat}}(P, Q) \\
&\geq 1 - \alpha(\pi) - \beta(\pi) - 1 + \frac{p_t^2 \varepsilon^3}{16} - O\left(p_t^4 \varepsilon^5\right) - \delta(\pi) \\
&= \frac{p_t^2 \varepsilon^3}{16} - O\left(p_t^4 \varepsilon^5\right) - \alpha(\pi) - \beta(\pi) - \delta(\pi)
\end{aligned}
$$

which concludes the proof by definition of $\alpha(\pi)$, $\beta(\pi)$ and $\delta(\pi)$.

3.4. **Application to the Hamming metric.** The assumptions of Theorem 1 will be satisfied for the Hamming metric for weights u close to $\tau^\perp n$ (where $\tau^\perp$ is given in Equation (8)) and we will prove that

Theorem 2. *Suppose that there exists an algorithm $\mathscr{A}$ solving with success probability ε the decoding problem $\mathsf{DP}(q, n, k, t)$ at Hamming distance $1 \leq t \stackrel{def}{=} \tau n \leq (1 - \delta) d_{\text{GV}}(n, k)$ for any arbitrary $\delta > 0$. Then, there exists a quantum algorithm which takes as input a generator matrix $\mathbf{G} \in \mathbb{F}_q^{k \times n}$ of a code $\mathcal{C} \subseteq \mathbb{F}_q^n$ and outputs $\mathbf{c}^\perp \in \mathcal{C}^\perp$ of weight $u \in [\![(1 - \alpha)\tau^\perp n, (1 + \alpha)\tau^\perp n]\!]$ (where α is any arbitrary constant > 0) with probability over a uniform choice of $\mathbf{G}$ given by a*

$\Omega\left(\frac{\varepsilon^3}{n} - O\left(\frac{\varepsilon^5}{n^2}\right) - 2^{-\Omega(n)}\right)$ *where:*

$$(18) \qquad\qquad \tau^\perp \stackrel{def}{=} \frac{1}{q}\left(\sqrt{(q-1)(1-\tau)} - \sqrt{\tau}\right)^2.$$

The proof of this theorem relies on Theorem 1, for a suitable choice of quantum state $|\pi\rangle$. This is done by choosing $|\pi\rangle = \left|\pi^{\text{Trunc}}\right\rangle$ which represents a truncated q-ary symmetric channel of crossover probability τ. All its weights are in an interval $[\![(1-\eta)t, (1+\eta)t]\!]$ where η is some positive constant which will be chosen later on. More precisely, let us first define the (untruncated) quantum state representing the q-ary symmetric channel of crossover probability τ:

$$\left|\pi^{\text{SC}}\right\rangle \stackrel{def}{=} \left(\sqrt{1-\tau}\,|0\rangle + \sqrt{\tau/(q-1)}\sum_{\alpha\in\mathbb{F}_q^*}|\alpha\rangle\right)^{\otimes n}.$$

Indeed, since $\left|\pi^{\text{SC}}\right\rangle$ can also be written as

$$\left|\pi^{\text{SC}}\right\rangle = \sum_{\mathbf{e}\in\mathbb{F}_q^n}\pi_{\mathbf{e}}\,|\mathbf{e}\rangle \quad\text{with}\quad \pi_{\mathbf{e}} \stackrel{def}{=} \sqrt{(1-\tau)^{n-|\mathbf{e}|}\left(\frac{\tau}{q-1}\right)^{|\mathbf{e}|}},$$

measuring $\left|\pi^{\text{SC}}\right\rangle$ mimics the error we have in a q-ary symmetric channel of crossover probability τ, *i.e.*,

$$\mathbb{P}(\text{measurement outputs } \mathbf{e}) = |\pi_{\mathbf{e}}|^2 = \left(\frac{\tau}{q-1}\right)^{|\mathbf{e}|}(1-\tau)^{n-|\mathbf{e}|}.$$

We will be interested in the truncated version given by

$$\left|\pi^{\text{Trunc}}\right\rangle \stackrel{def}{=} \sum_{\substack{\mathbf{e}\in\mathbb{F}_q^n:\\ |\mathbf{e}|\in[\![(1-\eta)t,(1+\eta)t]\!]}}\pi_{\mathbf{e}}^{\text{Trunc}}\,|\mathbf{e}\rangle \quad\text{with}\quad \pi_{\mathbf{e}}^{\text{Trunc}} \stackrel{def}{=} \begin{cases} \frac{\pi_{\mathbf{e}}}{\sqrt{N}} & \text{if } |\mathbf{e}|\in[\![(1-\eta)t, (1+\eta)t]\!] \\ 0 & \text{otherwise}\end{cases}$$

where N is the normalizing constant given by

$$(19) \qquad\qquad N \stackrel{def}{=} \sum_{\substack{\mathbf{e}\in\mathbb{F}_q^n\\ |\mathbf{e}|\in[\![(1-\eta)t,(1+\eta)t]\!]}}|\pi_{\mathbf{e}}|^2.$$

It will be helpful to notice that for all $\eta > 0$, N is exponentially close to 1:

Lemma 7. *For all $\eta > 0$, we have*

$$N = 1 - 2^{-\Omega(n)}.$$

Proof. Notice that by Equation (19),

$$(20) \quad 1 - N = \sum_{\substack{\mathbf{e}\in\mathbb{F}_q^n\\ |\mathbf{e}|\notin[\![(1-\eta)t,(1+\eta)t]\!]}}|\pi_{\mathbf{e}}|^2 = \mathbb{P}_{\mathbf{e}}(|\mathbf{e}|\notin[\![(1-\eta)t, (1+\eta)t]\!])$$

$$= \mathbb{P}_{\mathbf{e}}(|\mathbf{e}| < (1-\eta)\tau n) \;+\; \mathbb{P}_{\mathbf{e}}(|\mathbf{e}| > (1+\eta)\tau n)$$

where $|\mathbf{e}|$ is the sum of n independent (binary) Bernoulli random variables of parameter τ. Therefore, by Hoeffding's inequality, we have for all $\eta > 0$,

$$\mathbb{P}_{\mathbf{e}}(|\mathbf{e}| < (1-\eta)\tau n) \le e^{-2\eta^2\tau^2 n} \quad\text{and}\quad \mathbb{P}_{\mathbf{e}}(|\mathbf{e}| > (1+\eta)\tau n) \le e^{-2\eta^2\tau^2 n}$$

which concludes the proof by plugging this in Equation (20). $\qquad\square$

Theorem 2 is proved by showing that $\left|\pi^{\text{Trunc}}\right\rangle$ satisfies all the requirements of Theorem 1 when η is small enough.

Step 1: Verification of Condition ($C1$). This amounts to proving the following lemma

Lemma 8. *For $\eta > 0$ small enough, we have,*

$$\frac{\left\langle \pi^{Trunc}|\mathbf{1}\right\rangle^2}{q^{n-k}} = 2^{-\Omega(n)}.$$

Before proving this result, it will be helpful to notice that:

Lemma 9. *If $u \leq (1-\delta)d_{\mathrm{GV}}(n,k)$ for some $\delta > 0$, then*

$$\frac{S_u}{q^{n-k}} = 2^{-\Omega(n)}.$$

Proof. Recall that the size B_u of the Hamming ball of radius u is of the form

$$B_u = q^{n\,h_q(\mu)(1+o(1))}$$

where $\mu \stackrel{\text{def}}{=} u/n$. From this we obtain

$$
\begin{aligned}
\frac{S_u}{q^{n-k}} &\leq \frac{B_u}{B_{d_{\mathrm{GV}}}} && \text{(since } S_u \leq B_u \text{ and } B_{d_{\mathrm{GV}}} \leq q^{n-k}) \\
&\leq q^{n(h_q(\mu)-h_q(\delta_{\mathrm{GV}})+o(1))} \\
&\leq q^{n(h_q((1-\delta)\delta_{\mathrm{GV}})-h_q(\delta_{\mathrm{GV}})+o(1))}.
\end{aligned}
$$

We finish the proof by noticing that $h_q((1-\delta)\delta_{\mathrm{GV}}) - h_q(\delta_{\mathrm{GV}}) < 0$. $\qquad\square$

We are now ready to prove Lemma 8.

Proof of Lemma 8. We have the following computation,

$$
\begin{aligned}
\left\langle \pi^{\mathrm{Trunc}}|\mathbf{1}\right\rangle &= \sum_{\substack{\mathbf{e}\in\mathbb{F}_q^n \\ |\mathbf{e}|\in[\![(1-\eta)t,(1+\eta)t]\!]}} \pi_{\mathbf{e}}^{\mathrm{Trunc}} \\
&= \sum_{r\in[\![(1-\eta)t,(1+\eta)t]\!]} \sum_{\substack{\mathbf{e}\in\mathbb{F}_q^n \\ |\mathbf{e}|=r}} \frac{\pi_{\mathbf{e}}}{\sqrt{N}} \\
&= \frac{1}{\sqrt{N}} \sum_{r\in[\![(1-\eta)t,(1+\eta)t]\!]} \sqrt{\binom{n}{r}(q-1)^r} \sqrt{\binom{n}{r}(q-1)^r (1-\tau)^{n-r}\left(\frac{\tau}{q-1}\right)^r} \\
&\leq \frac{1}{\sqrt{N}} \sum_{r\in[\![(1-\eta)t,(1+\eta)t]\!]} \sqrt{\binom{n}{r}(q-1)^r}
\end{aligned}
$$

(21)

where in the last line we used that $\binom{n}{u}(q-1)^u\left(\frac{\tau}{q-1}\right)^u (1-\tau)^{n-u} \leq 1$ for any $u \in [\![0,n]\!]$. Therefore, using Equation (21), we have the following computation

$$
\begin{aligned}
\left\langle \pi^{\mathrm{Trunc}}|\mathbf{1}\right\rangle^2 &\leq \frac{1}{N}\left(\sum_{r=\lceil(1-\eta)t\rceil}^{\lfloor(1+\eta)t\rfloor} \sqrt{\binom{n}{r}(q-1)^r}\right)\left(\sum_{r'=\lceil(1-\eta)t\rceil}^{\lfloor(1+\eta)t\rfloor} \sqrt{\binom{n}{r'}(q-1)^{r'}}\right) \\
&\leq \frac{(n+1)^2}{N} \max_{r,r'\in[\![(1-\eta)t,(1+\eta)t]\!]} \sqrt{\binom{n}{r}\binom{n}{r'}(q-1)^r(q-1)^{r'}} \quad \text{(since } \sharp[\![(1-\eta)t,(1+\eta)t]\!] \leq n+1) \\
&\leq \frac{(n+1)^2}{N}\binom{n}{(1+\eta)t}(q-1)^{(1+\eta)t} \quad \text{(the max is reached for } r=r'=(1+\eta)t) \\
&= \frac{(n+1)^2}{N} S_{(1+\eta)t}
\end{aligned}
$$

Using this last inequality and Lemmas 7 and 9, we obtain

$$\frac{\left\langle \pi^{\text{Trunc}} | \mathbf{1} \right\rangle^2}{q^{n-k}} = O\left((n+1)^2 \, \frac{S_{(1+\eta)t}}{q^{n-k}} \right) = 2^{-\Omega(n)}$$

where we choose η small enough such that $(1+\eta)t \le (1-\delta')d_{\text{GV}}(n,k)$ for some $\delta' > 0$ (recall that by assumption $t \le (1-\delta)d_{\text{GV}}(n,k)$ for $\delta > 0$). $\qquad\qquad\square$

Remark 2. *As explained in the introduction and Remark 1, Lemma 8 is not satisfied by $\left| \pi^{SC} \right\rangle$. Here truncating the error distribution is essential to verify this concentration lemma.*

Step 2: Verification of Conditions (C2) and (C3). We prove here that the Conditions (C2) and (C3) of Theorem 1 are met by $\left| \pi^{\text{Trunc}} \right\rangle$. This results from a combination of arguments: (i) these two conditions are met for $\left| \pi^{SC} \right\rangle$ (ii) $\left| \pi^{SC} \right\rangle$ and $\left| \pi^{\text{Trunc}} \right\rangle$ are very close and so are $\widehat{\left| \pi^{SC} \right\rangle}$ and $\widehat{\left| \pi^{\text{Trunc}} \right\rangle}$ (because they are obtained from the first pair by applying a QFT, which is unitary). More precisely we are going to prove that

Lemma 10. *Let $t^\perp \overset{\text{def}}{=} \frac{(\sqrt{(q-1)(n-t)} - \sqrt{t})^2}{q}$ and $\alpha > 0$ be some constant small enough. We let $\widehat{\left| \pi^{\text{Trunc}} \right\rangle} = \sum_{\mathbf{e} \in \mathbb{F}_q^n} \widehat{\pi_{\mathbf{e}}^{\text{Trunc}}} \left| \mathbf{e} \right\rangle$. This state is radial and we let $\widehat{f^{\text{Trunc}}}(w) = \widehat{\pi_{\mathbf{e}}^{\text{Trunc}}}$ for any element $\mathbf{e}$ of $\mathbb{F}_q^n$ of Hamming weight w. We have*

$$p_t = \Omega\left(\frac{1}{\sqrt{n}} \right),$$

$$\forall u \in [\![t^\perp(1-\alpha), t^\perp(1+\alpha)]\!], \quad \frac{q^k}{S_u} = 2^{-\Omega(n)} \quad and \quad \sum_{u=\lceil t^\perp(1-\alpha) \rceil}^{\lfloor t^\perp(1+\alpha) \rfloor} S_u \left| \widehat{f^{\text{Trunc}}}(u) \right|^2 = 1 - 2^{-\Omega(n)}.$$

As explained above, to prove this result we will rely on the following lemma

Lemma 11. *For all $\eta > 0$, we have*

$$(22) \qquad \left\| \left| \pi^{\text{Trunc}} \right\rangle - \left| \pi^{SC} \right\rangle \right\| = 2^{-\Omega(n)} \quad and \quad \left\| \widehat{\left| \pi^{\text{Trunc}} \right\rangle} - \widehat{\left| \pi^{SC} \right\rangle} \right\| = 2^{-\Omega(n)}$$

Proof. We have the following computation,

$$\left\| \left| \pi^{\text{Trunc}} \right\rangle - \left| \pi^{SC} \right\rangle \right\|^2 = \sum_{\mathbf{e} \in \mathbb{F}_q^n} (\pi_{\mathbf{e}} - \pi_{\mathbf{e}}^{\text{Trunc}})^2$$

$$= \sum_{\substack{\mathbf{e} \in \mathbb{F}_q^n: \\ |\mathbf{e}| \notin [\![(1-\eta)t, (1+\eta)t]\!]}} \pi_{\mathbf{e}}^2 + \sum_{\substack{\mathbf{e} \in \mathbb{F}_q^n: \\ |\mathbf{e}| \in [\![(1-\eta)t, (1+\eta)t]\!]}} \left(\pi_{\mathbf{e}} - \frac{\pi_{\mathbf{e}}}{\sqrt{N}} \right)^2$$

$$= 1 - N + \frac{(1 - \sqrt{N})^2}{N} \sum_{\substack{\mathbf{e} \in \mathbb{F}_q^n \\ |\mathbf{e}| \in [\![(1-\eta)t, (1+\eta)t]\!]}} \pi_{\mathbf{e}}^2 \quad \text{(by Equation (19))}$$

$$= 1 - N + (1 - \sqrt{N})^2 \quad \text{(by Equation (19))}$$

$$\le 2^{-\Omega(n)} \quad \text{(by Lemma (7)).}$$

The second relation follows, since the QFT is an isometry with respect to $\|\cdot\|$. $\qquad\qquad\square$

With this lemma at hand, we are ready to prove Lemma 10.

Proof of Lemma 10. By definition,

$$p_t = \frac{1}{N} \sum_{\mathbf{e}\,:\,|\mathbf{e}|=t} (1-\tau)^{n-t} \left(\frac{\tau}{q-1}\right)^t = \frac{\binom{n}{t}(q-1)^t\, q^{nh_q(\tau)}}{N} = \Omega\left(\frac{1}{\sqrt{n}}\right)$$

where in the last equality we used Lemma 7 and Stirling's formula.

The equality $\frac{q^k}{S_u} = 2^{-\Omega(n)}$ is verified when u is sufficiently close to $t^\perp$ (α small enough), because it is readily verified that there exists some constant $\beta > 0$ such that $t^\perp \geq (1+\beta)d_{\mathrm{GV}}(n, n-k)$. This together with $t^\perp \leq \frac{(q-1)n}{q}$ implies that $\frac{q^k}{S_u} = 2^{-\Omega(n)}$ for any u in $[\![t^\perp(1-\alpha), t^\perp(1+\alpha)]\!]$.

The untruncated distribution $\left|\pi^{\mathrm{SC}}\right\rangle = \sum_{\mathbf{e}} \pi_{\mathbf{e}} \left|\mathbf{e}\right\rangle$ is radial, and so is its Fourier transform $\widehat{\left|\pi^{\mathrm{SC}}\right\rangle} = \sum_{\mathbf{e}} \widehat{\pi}_{\mathbf{e}} \left|\mathbf{e}\right\rangle$. We let $\widehat{f}(u) = \widehat{\pi}_{\mathbf{e}}$ where $\mathbf{e}$ is any $\mathbf{e} \in \mathbb{F}_q^n$ of Hamming weight u. We notice that

$$\sum_{u=\lceil t^\perp(1-\alpha)\rceil}^{\lfloor t^\perp(1+\alpha)\rfloor} S_u \left|\widehat{f}(u)\right|^2 = \sum_{\mathbf{e}\in\mathbb{F}_q^n\,:\,|\mathbf{e}|\in[(1-\alpha)t^\perp,(1+\alpha)t^\perp]} |\widehat{\pi}_{\mathbf{e}}|^2 = \mathbb{P}_{\mathbf{e}}\left(|\mathbf{e}| \in \left[(1-\alpha)t^\perp n, (1+\alpha)t^\perp n\right]\right),$$

where $\tau^\perp = \frac{t^\perp}{n}$ and $|\mathbf{e}|$ is the sum of n independent (binary) Bernoulli random variables of parameter $\tau^\perp$. Therefore, by using Hoeffding's bound again we obtain that

$$\tag{23} \sum_{u\in[(1-\alpha)t^\perp,(1+\alpha)t^\perp]} S_u \left|\widehat{f}(u)\right|^2 = 1 - 2^{-\Omega(n)}.$$

meaning that $\widehat{f}$ concentrates around vectors of weight $t^\perp$. Consider the projection of $\widehat{\left|\pi^{\mathrm{SC}}\right\rangle}$ and $\widehat{\left|\pi^{\mathrm{Trunc}}\right\rangle}$ on the space spanned by the states $\left|\mathbf{e}\right\rangle$ for $|\mathbf{e}| \in [\![(1-\alpha)t^\perp, (1+\alpha)t^\perp]\!]$:

$$\widetilde{\left|\pi^{\mathrm{SC}}\right\rangle} \overset{\mathrm{def}}{=} \sum_{\mathbf{e}\in\mathbb{F}_q^n\,:\,|\mathbf{e}|\in[(1-\alpha)t^\perp,(1+\alpha)t^\perp]} \widehat{\pi}_{\mathbf{e}} \left|\mathbf{e}\right\rangle$$

$$\widetilde{\left|\pi^{\mathrm{Trunc}}\right\rangle} \overset{\mathrm{def}}{=} \sum_{\mathbf{e}\in\mathbb{F}_q^n\,:\,|\mathbf{e}|\in[(1-\alpha)t^\perp,(1+\alpha)t^\perp]} \widehat{\pi_{\mathbf{e}}^{\mathrm{Trunc}}} \left|\mathbf{e}\right\rangle$$

Since a projection can only reduce the norm, we have

$$\tag{24} \left\|\widetilde{\left|\pi^{\mathrm{SC}}\right\rangle} - \widetilde{\left|\pi^{\mathrm{Trunc}}\right\rangle}\right\| \leq \left\|\widehat{\left|\pi^{\mathrm{SC}}\right\rangle} - \widehat{\left|\pi^{\mathrm{Trunc}}\right\rangle}\right\| = 2^{-\Omega(n)}.$$

We deduce from the triangle inequality that

$$\tag{25} \left\|\widetilde{\left|\pi^{\mathrm{Trunc}}\right\rangle}\right\| \geq \left\|\widetilde{\left|\pi^{\mathrm{SC}}\right\rangle}\right\| - \left\|\widetilde{\left|\pi^{\mathrm{SC}}\right\rangle} - \widetilde{\left|\pi^{\mathrm{Trunc}}\right\rangle}\right\|,$$

and then from Equations (24) and (23) $\left(\text{which says } \left\|\widetilde{\left|\pi^{\mathrm{SC}}\right\rangle}\right\|^2 = 1 - 2^{-\Omega(n)}\right)$ that $\left\|\widetilde{\left|\pi^{\mathrm{Trunc}}\right\rangle}\right\| \geq 1 - 2^{-\Omega(n)}$. Since $\left\|\widetilde{\left|\pi^{\mathrm{Trunc}}\right\rangle}\right\| \leq \left\|\widehat{\left|\pi^{\mathrm{Trunc}}\right\rangle}\right\| = 1$ we finally obtain

$$\left\|\widetilde{\left|\pi^{\mathrm{Trunc}}\right\rangle}\right\| = 1 - 2^{-\Omega(n)}.$$

This directly implies $\sum_{u\in[\![(1-\alpha)t^\perp,(1+\alpha)t^\perp]\!]} S_u \left|\widehat{f^{\mathrm{Trunc}}}(u)\right|^2 = \left\|\widetilde{\left|\pi^{\mathrm{Trunc}}\right\rangle}\right\|^2 = 1 - 2^{-\Omega(n)}.$ $\qquad\square$

Proof of Theorem 2. This immediately follows from Lemmas 8 and 10 which show that the relevant assumptions of Theorem 1 are verified for the choice $\left|\pi\right\rangle = \left|\pi^{\mathrm{Trunc}}\right\rangle$ for η small enough.

3.5. **Application to the rank metric.** The assumptions of Theorem 1 will also be satisfied in the context of codes $\mathcal{C} \subseteq \mathbb{F}_q^{m \times n}$ embedded with the rank metric (given some matrix in $\mathbb{F}_q^{m \times n}$, its weight is defined as its rank). The Gilbert-Varshamov distance $d_{\mathrm{GV}}(m, n, k)$ is defined in a similar way, but it depends on three parameters here and corresponds to the largest radius t of a ball in the rank metric for which

$$q^k B_t \leq q^{m \times n}.$$

We will be able to prove that

Theorem 3. *Suppose that there exists an algorithm $\mathscr{A}$ solving with success probability ε the decoding problem at rank distance $1 \leq t < d_{\mathrm{GV}}(m, n, k)$ where $m \geq n$. Then, there exists a quantum algorithm which takes as input a generator matrix of $\mathcal{C} \subseteq \mathbb{F}_q^{m \times n}$ and outputs $\mathbf{c}^\perp \in \mathcal{C}^\perp$ of weight $u \in (t^\perp - \eta n, t^\perp]$ where $t^\perp \overset{def}{=} n - t$ (for any arbitrary constant $\eta > 0$) with probability over a uniform choice of the generator matrix given by a $\Omega\left(\varepsilon^3 - O\left(\varepsilon^5\right) - 2^{-\Omega(n)}\right)$.*

Remark 3. *Our assumption that $m \geq n$ can be done without loss of generality. In the case where $n > m$ we can just consider the transposed code $\mathcal{C}^\mathsf{T} \overset{def}{=} \{\mathbf{M}^\mathsf{T} : \mathbf{M} \in \mathcal{C}\}$: taking the transpose is a linear automorphism and can be used to transform any algorithm decoding $\mathcal{C}$ into an algorithm decoding $\mathcal{C}^\mathsf{T}$ with the same complexity.*

As in the Hamming case, Theorem 3 will be a consequence of Theorem 1. Therefore we first have to choose appropriately a quantum state $|\pi\rangle$ that will model the noise distribution.

Step 1 : Choosing $|\pi\rangle$. Let,

$$(26) \qquad |\pi\rangle \overset{\text{def}}{=} \frac{1}{\sqrt{N}} \sum_{\substack{V \leq \mathbb{F}_q^n \\ \dim V = t}} |\pi_V\rangle \quad \text{where}$$

$$(27) \qquad |\pi_U\rangle \overset{\text{def}}{=} \left(\frac{1}{\sqrt{q^{\dim U}}} \sum_{\mathbf{u} \in U} |\mathbf{u}\rangle\right)^{\otimes m}$$

and N is a normalizing constant. $|\pi_U\rangle$ can be viewed as a uniform superposition of matrices whose rows belong to U. These matrices have rank at most $\dim U$ and $|\pi\rangle$ is close to a uniform distribution of all matrices of rank t. We also have the following alternative description for $|\pi\rangle$.

Lemma 12. *$|\pi\rangle$ is radial, i.e., we may write $|\pi\rangle$ as*

$$|\pi\rangle = \sum_{\mathbf{E} \in \mathbb{F}_q^{m \times n} : \, |\mathbf{E}| \leq t} \pi_{\mathbf{E}} |\mathbf{E}\rangle$$

with $|\mathbf{E}\rangle \overset{def}{=} |\mathbf{E}_1\rangle \otimes \cdots \otimes |\mathbf{E}_m\rangle$ where the $\mathbf{E}_i$'s denote the rows of $\mathbf{E}$ and $\pi_{\mathbf{E}} = f(|\mathbf{E}|)$ where

$$f(u) = \begin{cases} \dfrac{\left[\begin{smallmatrix} n-u \\ t-u \end{smallmatrix}\right]_q}{\sqrt{q^{mt} N}} & \text{if } u \leq t, \\ 0 & \text{otherwise.} \end{cases}$$

This lemma is proved in Appendix B.1, as well as the following one, which gives estimations for N and p_t in order to apply Theorem 1.

Lemma 13. *We have:*

$$N = \Theta\left(\begin{bmatrix} n \\ t \end{bmatrix}_q\right) \quad \text{and} \quad p_t = \Theta\left(1\right).$$

Step 2: Verification that $\widehat{|\pi\rangle}$ is radial. The following proposition states that $\widehat{|\pi\rangle}$ has actually the same form as $|\pi\rangle$ where t is replaced by $n - t$:

Proposition 4. *We have,*

$$\widehat{|\pi\rangle} = \frac{1}{\sqrt{N}} \sum_{\substack{W \le \mathbb{F}_q^n \\ \dim W = n-t}} |\pi_W\rangle .$$

Proof. We apply the QFT on $|\pi\rangle$, defined in Equation (26). It gives,

$$\widehat{|\pi\rangle} = \frac{1}{\sqrt{N}} \sum_{\substack{V \le \mathbb{F}_q^n \\ \dim V = t}} \left(\frac{1}{\sqrt{q^{n+t}}} \sum_{\mathbf{y} \in \mathbb{F}_q^n} \left(\sum_{\mathbf{v} \in V} \chi_{\mathbf{y}}(\mathbf{v}) \right) |\mathbf{y}\rangle \right)^{\otimes m}$$

By distinguishing the cases where $\mathbf{y} \in V^\perp$ (the dual of V with the standard inner product) or not:

$$\widehat{|\pi\rangle} = \frac{1}{\sqrt{N}} \sum_{\substack{V \le \mathbb{F}_q^n \\ \dim V = t}} \left(\frac{1}{\sqrt{q^{n+t}}} \sum_{\mathbf{y} \in V^\perp} q^t |\mathbf{y}\rangle \right)^{\otimes m} = \frac{1}{\sqrt{N}} \sum_{\substack{W \le \mathbb{F}_q^n \\ \dim W = n-t}} \left(\frac{1}{\sqrt{q^{n-t}}} \sum_{\mathbf{y} \in W} |\mathbf{y}\rangle \right)^{\otimes m}$$

which concludes the proof. $\qquad\square$

We can now straightforwardly apply Lemma 12 on $\widehat{|\pi\rangle}$ and obtain

Lemma 14. *The state $\widehat{|\pi\rangle}$ is radial and can be written as $\sum_{\mathbf{E} \in \mathbb{F}_q^{m \times n} : |\mathbf{E}| \le n-t} \widehat{\pi_{\mathbf{E}}} |\mathbf{E}\rangle$. If we let $\widehat{f}(u) \overset{def}{=} \widehat{\pi_{\mathbf{E}}}$ for any $\mathbf{E} \in \mathbb{F}_q^{m \times n}$ of rank u, we have*

$$\widehat{f}(u) = \begin{cases} \dfrac{\left[\begin{smallmatrix} n-u \\ n-t-u \end{smallmatrix} \right]_q}{\sqrt{q^{m(n-t)}N}} & \text{if } u \le n-t, \\ 0 & \text{otherwise.} \end{cases}$$

Step 3: Verification of Conditions ($C1$), ($C2$) and ($C3$). This is achieved in the following Lemmas that are proved in Appendix B.2.

Lemma 15. *We have,*

$$\frac{S_t}{q^{mn-k}} = q^{-\Omega(n)} \quad and \quad \frac{\langle \pi | \mathbf{1} \rangle^2}{q^{mn-k}} = q^{-\Omega(n)}.$$

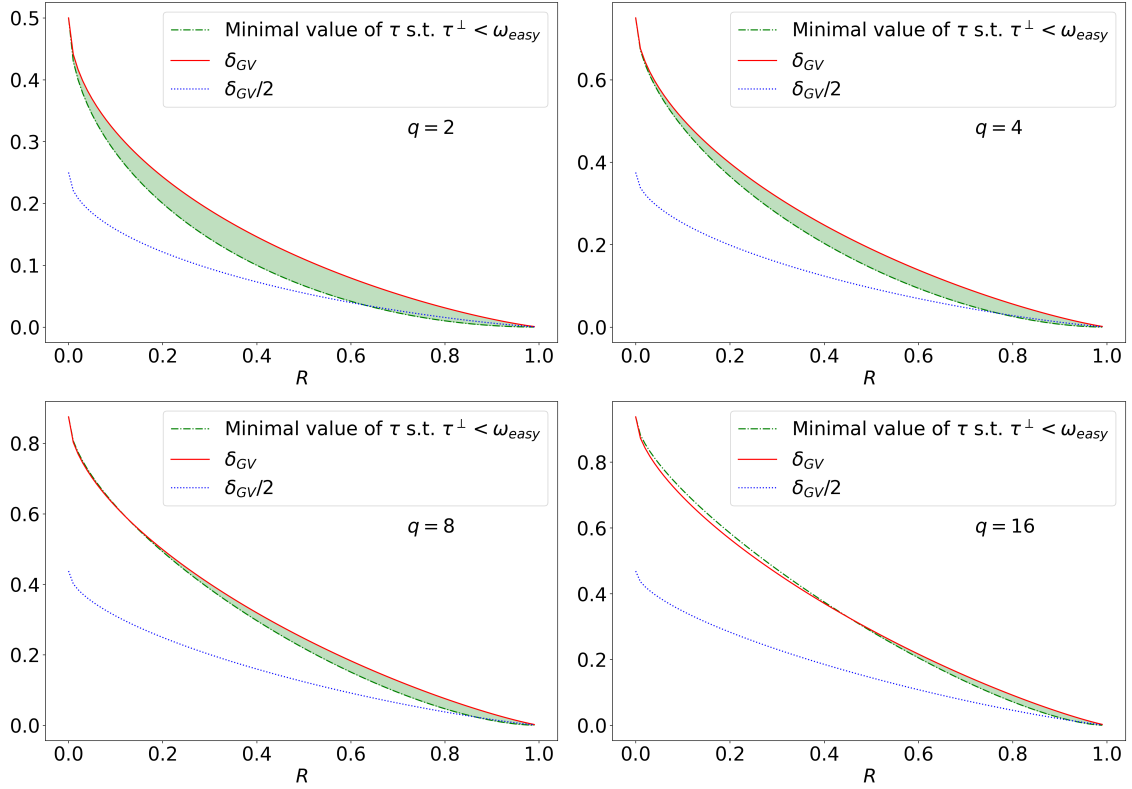**Lemma 16.** *For any $\eta > 0$, we have,*

$$\forall u \in [\![(1-\eta)n - t, n-t]\!], \ \frac{q^k}{S_u} = q^{-\Omega(n)} \quad and \quad \sum_{u \in [\![(1-\eta)n-t, n-t]\!]} S_u \left| \widehat{f}(u) \right|^2 = 1 - q^{-\Omega(n)}.$$

Proof of Theorem 3. This follows from Lemmas 12, 14, 15 and 16 that allow to apply Theorem 1, completing the proof.

4. About the usefulness of our reduction.

It is now interesting to look at the parameters for which our reduction is useful for both the Hamming and rank metrics.

FIGURE 1. Range of values for τ as function of R.

4.1. **Hamming case.** A lower-bound on τ is obtained with the following arguments. First, if one wants to compute dual codewords (via the quantum measure) for which no poly-time algorithm is known, one has to ensure that $\tau^\perp < \omega_{\text{easy}}(n, n-k) = \frac{q-1}{q}\frac{k}{n}$. But notice that $\tau \mapsto \tau^\perp$ is a decreasing involution on $\left[0, \frac{1}{2}\right]$. Therefore, for the reduction to be meaningful, it is necessary that

$$(28) \quad \tau > \omega_{\text{easy}}(n, n-k)^\perp = \frac{1}{q}\left(\sqrt{(q-1)\left(1 - \frac{q-1}{q}\frac{k}{n}\right)} - \sqrt{\frac{q-1}{q}\frac{k}{n}}\right)^2$$

$$= \frac{q-1}{q^2}\left(\sqrt{q - (q-1)\frac{k}{n}} - \sqrt{\frac{k}{n}}\right)^2$$

Furthermore, according to Theorem 2, the relative decoding distance τ has to verify

$$\tau < \frac{d_{\text{GV}}(n,k)}{n} = h_q^{-1}\left(1 - \frac{k}{n}\right) + O\left(\frac{1}{n}\right) = \delta_{\text{GV}}(n,k) + O\left(\frac{1}{n}\right).$$

Roughly speaking, it gives the tightest upper-bound for which we can expect to correctly decode with an overwhelming probability. Combining this with (28) leads to a whole interval in which τ needs to lie for the reduction to work and be meaningful. In Figure 1, we draw (asymptotically in n) this range of values of τ as function of $R \stackrel{\text{def}}{=} \frac{k}{n}$ for different values of q (the green area). Two important remarks can be made from these graphs:

  (1) the range of interesting values for τ (*i.e.,* values such that we solve a hard instance of SCP) shrinks as q grows and depending on R;

(2) the lower bound on τ corresponding to $\tau^\perp < \omega_{\text{easy}}(n, n - k)$ is almost always above $\delta_{\text{GV}}(n, k)/2$, meaning that in order to solve a hard instance of SCP, it is a necessity that τ goes beyond the unique decoding radius.

4.2. **Rank Case.** It turns out that unfortunately, for the rank metric (which is coarser than the Hamming metric), we always reduce decoding t errors to finding dual codewords of weight $t^\perp = n - t$ where $t^\perp$ belongs to a range of values for which it is always easy to find codewords of this weight as we now show.

To verify this point, consider a linear code $\mathcal{C} \subseteq \mathbb{F}_q^{m \times n}$ of dimension K (with $m \geq n$). It is easy to find short codewords if they are above a certain range. To produce codewords of small weight, we use the fact that the dual code is a vector space of dimension $nm - K$. Thus, we can just produce codewords with $nm - K - 1$ entries equal to 0 that will be good candidates for having a small weight by solving a linear system. The entries are chosen so as to fill columns with zeroes. It is straightforward that this strategy produces in polynomial time codewords of weight $\approx Rn$ (since in our case $n \leq m$) where R is the rate of $\mathcal{C}$ defined by $R \stackrel{\text{def}}{=} \frac{K}{mn}$.

Notice now that $t^\perp$ is a decreasing function of the decoding distance t. The largest value for which we can hope to decode is the Gilbert-Varshamov distance $d_{\text{GV}}(m, n, K)$. The relative Gilbert-Varshamov distance $\delta_{\text{GV}}(m, n, K) \stackrel{\text{def}}{=} \frac{d_{\text{GV}}(m,n,K)}{n}$ satisfies the relation

$$R = 1 - \delta_{\text{GV}}(1 + \nu - \delta_{\text{GV}})$$

where $\nu \stackrel{\text{def}}{=} \frac{m}{n} \geq 1$. However, we have (where $\delta_{\text{GV}}^\perp$ is defined as $t^\perp/n$ when $t/n = \delta_{\text{GV}}(m, n, K)$)

$$\frac{t^\perp}{n} \geq \delta_{\text{GV}}^\perp = 1 - \delta_{\text{GV}} = R + \delta_{\text{GV}}(\nu - \delta_{\text{GV}}) \geq R.$$

In other words, we are always in a regime where finding codewords of relative weight $t^\perp/n$ is easy.

5. Concluding Remarks

Considering other metrics. The whole approach we have followed here (properly choosing the error distribution and, if necessary, going beyond the unique decoding radius for decoding) can of course be adapted to other metrics than those we have considered here (Hamming and rank). For instance, it would be interesting to investigate it also for metrics like the Lee metric (more or less the L_1 norm version of the Euclidean metric on $\mathbb{Z}_q^n$) which has also begun to find its way in code-based cryptography [HTW20] and should have a behavior closer to the Euclidean metric if the size of the alphabet grows with the code length.

Devising quantum algorithms for producing low weight codewords. Interestingly, the very same quantum reduction was recently used in [CLZ22] to devise a polynomial-time algorithm for finding moderately low weight codewords for the L_∞ norm in a regime of parameters for which no known polynomial time algorithms exist to perform this task. The code rate regime considered there (vanishing with the codelength) avoids the technicalities we needed here (namely truncating the error distribution). Again, this work emphasizes the role of the dual noise model obtained by applying the Fourier transform.

The dual error distribution. Moving to this quantum setting allows us to define in a natural way a dual error distribution. Indeed, consider the quantum state representing the noise added to the codeword

$$|\pi\rangle \stackrel{\text{def}}{=} \sum_{\mathbf{e} \in \mathbb{F}_q^n} \pi_{\mathbf{e}} |\mathbf{e}\rangle.$$

$(|\pi_{\mathbf{e}}|^2)_{\mathbf{e} \in \mathbb{F}_q^n}$ is a probability distribution of the error and its quantum Fourier transform $\widehat{|\pi\rangle} \stackrel{\text{def}}{=} \sum_{\mathbf{e} \in \mathbb{F}_q^n} \widehat{\pi}_{\mathbf{e}} |\mathbf{e}\rangle$ gives a "dual" probability distribution of the noise $(|\widehat{\pi}_{\mathbf{e}}|^2)_{\mathbf{e} \in \mathbb{F}_q^n}$ which we view as the dual noise distribution. Note that a given noise distribution $(p_{\mathbf{e}})_{\mathbf{e} \in \mathbb{F}_q^n}$ may be represented by several different quantum states $\sum_{\mathbf{e} \in \mathbb{F}_q^n} \pi_{\mathbf{e}} |\mathbf{e}\rangle$ since it is sufficient that $p_{\mathbf{e}} = |\pi_{\mathbf{e}}|^2$ for all $\mathbf{e}$. The choice of $\pi_{\mathbf{e}} = \sqrt{p_{\mathbf{e}}}$ is natural but not canonical and is used to define what we call the dual noise distribution. Interestingly enough, this dual noise distribution seems to capture very fundamental quantities in coding theory as we will explain in the following paragraph. Moreover, contrarily to the dual noise considered in [For01, §IX], which is also obtained by taking the Fourier transform and dualizing the sum-product algorithm and does not always yield a probability distribution when applied to a probability distribution, we define here a dual noise distribution that is always a probability distribution.

About the dual distance $t^{\perp} = \tau^{\perp}n$. This dual distribution also allows us to define a notion of "dual distance". We namely consider here probability distributions (such as the Bernoulli noise) that are concentrated around a certain weight s and such that the dual distribution concentrates around another weight as well that can be considered as the "dual weight" $s^{\perp}$. In the case of the Bernoulli noise of parameter τ, the dual distribution is again a Bernoulli noise distribution, of parameter $\tau^{\perp}$. In other words, the dual distance $t^{\perp}$ can be defined as $\tau^{\perp}n$ when $t = \tau n$. Note that the choice we made for defining the dual distribution, namely that $\pi_{\mathbf{e}} = \sqrt{p_{\mathbf{e}}}$ (where $p_{\mathbf{e}}$ is the probability that $\mathbf{e}$ is drawn from a Bernoulli distribution of parameter τ), minimizes the weight $t^{\perp}$ around which the dual distribution converges. This gives the strongest reduction in our case.

Such an interpretation also holds in the lattice based setting (and the Euclidean distance) [Reg05, SSTX09] when the noise is distributed over $\mathbb{R}^n$ according to a Gaussian distribution $D_s(\mathbf{x}) = \frac{e^{-\pi \|\mathbf{x}\|^2 / s^2}}{s^n}$ concentrated around the distance $t = s\sqrt{\frac{n}{2\pi}}$. The dual noise is obtained by taking the Fourier transform of $\sqrt{D_s}$ and squaring the result. The dual noise is a centered Gaussian

$$t^{\perp} \stackrel{\text{def}}{=} \frac{1}{2s}\sqrt{\frac{n}{2\pi}} = \frac{n}{4\pi t}$$

where in the last equality we used that $t = s\sqrt{\frac{n}{2\pi}}$.

The same also holds in the case of the rank metric, where we define a quantum state representing the noise $|\pi\rangle$ whose rank weight concentrates around the weight t we can decode. After applying the quantum Fourier transform, $\widehat{|\pi\rangle}$ is a superposition of elements whose rank weight concentrates around $t^{\perp} \stackrel{\text{def}}{=} n - t$ (when we consider codes of length n).

This notion of dual distance has an intriguing connection with a very fundamental and old issue which is still open for an overwhelming majority of metric spaces: what is the largest packing density for a given packing radius? The answer is generally not even known asymptotically. One of the most powerful technique which gives the best known bounds in the case of the Hamming metric or for the Euclidean metric over $\mathbb{R}^n$ is obtained through a linear programming approach [Del72, DL98], see [MRRW77, Lev79]. Interestingly enough, the first linear programming bound in coding theory [MRRW77] or the one [Lev79, CE03] for sphere packing in $\mathbb{R}^n$ can be be rephrased in terms of our dual distance and the Gilbert-Varshamov distance. Strictly speaking, the Gilbert-Varshamov is a coding theoretic notion, but it has also an analogue for $\mathbb{R}^n$, see [DDRT22] for instance. For this purpose, let us express the Gilbert-Varshamov distance as a function $d_{\text{GV}}(\delta)$ of the packing density $\delta \stackrel{\text{def}}{=} \frac{|\mathcal{C}|}{|\mathcal{E}|}$ where $|\mathcal{C}|$ is the code size and $|\mathcal{E}|$ the ambient space size. It is the supremum of the radius t of a ball B_t such that

$$|B_t| \leq \frac{1}{\delta}.$$

The Gilbert-Varshamov distance for $\mathbb{R}^n$ and a given packing density δ can be defined similarly as the supremum of the radius t of a ball B_t such that

$$\mathrm{Vol}(B_t) \leq \frac{1}{\delta}.$$

With this notion, the first linear programming bound in coding theory [MRRW77] or the one [Lev79] for sphere packing in $\mathbb{R}^n$ can be expressed as

$$(29) \qquad t^\perp \geq d_{\mathrm{GV}}(\delta^*)(1 + o(1)),$$

where δ^* is the dual packing density: $\delta^* = \frac{1}{\delta}$ in the case of $\mathbb{R}^n$ and $\delta^* = \frac{1}{\delta|\mathcal{E}|}$ in coding theoretic setting[3]. This indeed gives an upper-bound on t since $t^\perp$ is a decreasing function of t. This suggests that there might be a direct proof of (29) relying only on the way $t^\perp$ is defined and could yield new bounds on the packing density/minimum distance of a code for the metrics where linear programming bounds are not known.

---

[3] This notion is related to the notion of dual lattice or dual code. Indeed in the code case, if the density $\frac{|\mathcal{C}|}{|\mathcal{E}|}$ of a linear code $\mathcal{C}$ is δ, the density of the dual code $\mathcal{C}^\perp$ is δ^* since $\frac{|\mathcal{C}^\perp|}{|\mathcal{E}|} = \frac{|\mathcal{C}^\perp|\cdot|\mathcal{C}|}{|\mathcal{E}|\cdot|\mathcal{C}|} = \frac{|\mathcal{E}|}{|\mathcal{E}|\cdot|\mathcal{C}|} = \frac{1}{\delta|\mathcal{E}|}$. Similarly, for $\mathbb{R}^n$ if the density of a lattice Λ is δ the density of the dual lattice $\Lambda^* \overset{\mathrm{def}}{=} \{\mathbf{y} \in \mathbb{R}^n : <\mathbf{x}, \mathbf{y}> \in \mathbb{Z}, \forall \mathbf{x} \in \Lambda\}$ is $\frac{1}{\delta} = \delta^*$.

APPENDIX A. PROOF OF THEOREM 1

A.1. **Step 1: Proof of Lemmas 1 and 3.** Let us recall Lemma 1 first:

Lemma 1. *We have:*

$$D_{\mathrm{tr}}(|\psi_{\mathscr{A}}\rangle, |\psi_{ideal}\rangle) \leq \sqrt{1 - \frac{q^k p_t^2}{Z} \varepsilon_{\mathbf{G}}^2}.$$

Proof. Let $\mathscr{G}$ be the set of $(\mathbf{c}, \mathbf{e})$'s that correspond to inputs of weight t to $\mathscr{A}$ that are correctly decoded:

$$\mathscr{G} \stackrel{\mathrm{def}}{=} \{(\mathbf{c}, \mathbf{e}) \in \mathcal{C} \times \mathscr{S}_t \colon \mathscr{A}(\mathbf{G}, \mathbf{c} + \mathbf{e}) = \mathbf{e}\}.$$

Let us recall that

$$|\psi_{\mathscr{A}}\rangle = \frac{1}{\sqrt{q^k}} \sum_{\mathbf{e} \in \mathbb{F}_q^n} \sum_{\mathbf{c} \in \mathcal{C}} \pi_{\mathbf{e}} |\mathbf{e} - \mathscr{A}(\mathbf{G}, \mathbf{c} + \mathbf{e})\rangle |\mathbf{c} + \mathbf{e}\rangle \quad \text{and} \quad |\psi_{\mathrm{ideal}}\rangle = \frac{1}{\sqrt{Z}} \sum_{\mathbf{e} \in \mathbb{F}_q^n} \sum_{\mathbf{c} \in \mathcal{C}} \pi_{\mathbf{e}} |\mathbf{0}_n\rangle |\mathbf{c} + \mathbf{e}\rangle$$

From this we deduce by using the non-negativity of $\pi_{\mathbf{e}}$ that

$$\begin{aligned}
\langle \psi_{\mathscr{A}} | \psi_{\mathrm{ideal}} \rangle &\geq \frac{1}{\sqrt{q^k Z}} \sum_{(\mathbf{c}, \mathbf{e}) \in \mathscr{G}} \pi_{\mathbf{e}}^2 \\
&= \sqrt{\frac{q^k}{Z}} S_t f(t)^2 \frac{\sharp \mathscr{G}}{q^k S_t} \\
&= \sqrt{\frac{q^k}{Z}} p_t \, \varepsilon_{\mathbf{G}}.
\end{aligned}$$

□

Remark 4. *Here we do not have as in the lattice case* [SSTX09] *to make the assumption that the decoder is "strongly solution independent". In our case we can indeed have a uniform superposition over all the codewords and we can just use the way our error probability is defined, namely as the ratio* $\frac{\sharp \mathscr{G}}{q^k S_t}$.

All the probabilistic results of this section are easier to prove if, instead of choosing a code $\mathcal{C}$ by picking uniformly at random a generator matrix $\mathbf{G}$ for it, we slightly change the probabilistic model by picking uniformly at random a parity-check matrix $\mathbf{H} \in \mathbb{F}_q^{(n-k) \times n}$ for it, *i.e.,*

$$\mathcal{C} = \{\mathbf{x} \in \mathbb{F}_q^n \colon \mathbf{H}\mathbf{x}^{\mathsf{T}} = \mathbf{0}\}.$$

We will denote $\mathbb{P}_{\mathbf{G}}$ and $\mathbb{P}_{\mathbf{H}}$ respectively the probabilities in the initial model and the probabilities in the new model. The two probability distributions are closely related: the first model always produces linear codes of dimension $\leq k$ and codes of dimension $= k$ with probability $1 - O(q^{-(n-k)})$ whereas the second model always produces linear codes of dimension $\geq k$ and codes of dimension $= k$ with probability $1 - O(q^{-k})$. This relationship is expressed by the following lemma.

Lemma 17. *Let $\mathcal{E}$ be an ensemble of linear codes of length n in $\mathbb{F}_q$. We have*

$$\mathbb{P}_{\mathbf{G}}(\mathcal{E}) \leq \mathbb{P}_{\mathbf{H}}(\mathcal{E}) + O\left(q^{-\min(k, n-k)}\right).$$

With this new probabilistic model, the expected value of Z is given by:

Lemma 2. *Assume that $\mathcal{C}$ is chosen by uniformly drawing at random a parity-check matrix $\mathbf{H}$ for it. We have:*

$$(17) \qquad \mathbb{E}(Z) \leq q^k \left(1 + \frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}\right).$$

Proof. Computing $\mathbb{E}_{\mathbf{H}}(Z)$ with this alternate probabilistic model is straightforward. We have

$$
\begin{aligned}
Z &= \left\| \sum_{\mathbf{c}\in\mathcal{C},\mathbf{e}\in\mathbb{F}_q^n} \pi_{\mathbf{e}} \left|\mathbf{0}_n\right\rangle \left|\mathbf{c}+\mathbf{e}\right\rangle \right\|^2 \\
&= q^k \sum_{\mathbf{e}\in\mathbb{F}_q^n} |\pi_{\mathbf{e}}|^2 + \sum_{\substack{(\mathbf{c},\mathbf{e})\neq(\mathbf{c}',\mathbf{e}'):\\ \mathbf{c}+\mathbf{e}=\mathbf{c}'+\mathbf{e}'}} \pi_{\mathbf{e}}\pi_{\mathbf{e}'} \\
&= q^k \left(1 + \sum_{\mathbf{e}\neq\mathbf{e}'\,:\,\mathbf{H}(\mathbf{e}-\mathbf{e}')^{\mathsf{T}}=\mathbf{0}} \pi_{\mathbf{e}}\pi_{\mathbf{e}'} \right)
\end{aligned}
$$

where $\mathbf{H}$ is an arbitrary-parity check matrix for $\mathcal{C}$. Let

$$
X \overset{\mathrm{def}}{=} \sum_{\mathbf{e}\neq\mathbf{e}'\,:\,\mathbf{H}(\mathbf{e}-\mathbf{e}')^{\mathsf{T}}=\mathbf{0}} \pi_{\mathbf{e}}\pi_{\mathbf{e}'}.
$$

The point of the probabilistic model where the parity-check matrix $\mathbf{H}$ is uniformly drawn at random is that for non-zero element $\mathbf{x} \in \mathbb{F}_q^n$ we have

$$
\mathbb{P}_{\mathbf{H}}(\mathbf{x}\in\mathcal{C}) = \mathbb{P}_{\mathbf{H}}(\mathbf{H}\mathbf{x}^{\mathsf{T}}=\mathbf{0}) = \frac{1}{q^{n-k}}.
$$

From this we deduce

$$
\begin{aligned}
\mathbb{E}_{\mathbf{H}}(X) &= \sum_{\mathbf{e}\neq\mathbf{e}'} \pi_{\mathbf{e}}\pi_{\mathbf{e}'}\mathbb{P}_{\mathbf{H}}((\mathbf{e}-\mathbf{e}')\in\mathcal{C}) \\
&= \sum_{\mathbf{e}\neq\mathbf{e}'} \frac{\pi_{\mathbf{e}}\pi_{\mathbf{e}'}}{q^{n-k}} \\
&\leq \sum_{\mathbf{e},\mathbf{e}'} \frac{\pi_{\mathbf{e}}\pi_{\mathbf{e}'}}{q^{n-k}} \\
&= \frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}}.
\end{aligned}
$$

From this inequality we conclude the proof. $\qquad\square$

With the help of these two lemmas, we can upper-bound the probability for Z to be bigger than $q^k(1+\eta)$ for any $\eta > 0$ and prove Lemma 3, that we recall:

Lemma 3. *Let $\eta > 0$. We have:*

$$
\mathbb{P}_{\mathbf{G}}(Z > q^k(1+\eta)) \leq \frac{1}{\eta}\frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}} + O\left(q^{-\min(k,n-k)}\right).
$$

Proof. We take back the notation from the proof of Lemma 2. We have

$$
\begin{aligned}
\mathbb{P}_{\mathbf{G}}(Z > q^k(1+\eta)) &= \mathbb{P}_{\mathbf{G}}(X > \eta) \\
&\leq \mathbb{P}_{\mathbf{H}}(X > \eta) + O\left(q^{-\min(k,n-k)}\right) \quad \text{(by Lemma 17)} \\
&\leq \frac{1}{\eta}\mathbb{E}_{\mathbf{H}}(X) + O\left(q^{-\min(k,n-k)}\right) \quad \text{(Markov inequality)} \\
&\leq \frac{1}{\eta}\frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}} + O\left(q^{-\min(k,n-k)}\right)
\end{aligned}
$$

which concludes the proof. $\qquad\square$

A.2. **Step 2: Proof of Lemma 4.** If we apply a QFT on the second register of $|\psi_{\text{ideal}}\rangle$ (given in Equation (13)), we obtain:

$$\widehat{|\psi_{\text{ideal}}\rangle} \stackrel{\text{def}}{=} (\mathbf{I} \otimes \text{QFT} \otimes \mathbf{I}) |\psi_{\text{ideal}}\rangle = \frac{q^k}{\sqrt{Z}} \sum_{\mathbf{c}^\perp \in \mathcal{C}^\perp} \widehat{\pi}_{\mathbf{c}^\perp} |\mathbf{0}_n\rangle |\mathbf{c}^\perp\rangle,$$

where $\widehat{|\pi\rangle} = \sum_{\mathbf{e} \in \mathbb{F}_q^n} \widehat{\pi}_{\mathbf{e}} |\mathbf{e}\rangle$ is the QFT of $|\pi\rangle$. We use this remark to prove Lemma 4, that we recall:

Lemma 4. *If the Fourier transform is radially preserving, meaning that it transforms a radial function into a radial function, then after measuring $\widehat{|\psi_{ideal}\rangle}$ we obtain a state $|\mathbf{0}_n\rangle |\mathbf{c}^\perp\rangle$ with $\mathbf{c}^\perp \in \mathcal{C}^\perp$ of weight u with probability $\frac{q^{2k}}{Z} N_u^\perp \left|\widehat{f}(u)\right|^2$ where $\widehat{f}(u) \stackrel{def}{=} \widehat{\pi}_{\mathbf{e}}$ for an arbitrary $\mathbf{e}$ of weight u and $N_u^\perp$ is the number of codewords of weight u in $\mathcal{C}^\perp$.*

Proof. For $\mathbf{e} \in \mathbb{F}_q^n$, let

$$|\mathbf{1}_{\mathcal{C}+\mathbf{e}}\rangle \stackrel{\text{def}}{=} \sum_{\mathbf{c} \in \mathcal{C}} |\mathbf{c} + \mathbf{e}\rangle.$$

We have

$$
\begin{aligned}
\widehat{|\mathbf{1}_{\mathcal{C}+\mathbf{e}}\rangle} &= \sum_{\mathbf{c} \in \mathcal{C}} \frac{1}{\sqrt{q^n}} \sum_{\mathbf{y} \in \mathbb{F}_q^n} \chi_{\mathbf{y}}(\mathbf{c} + \mathbf{e}) |\mathbf{y}\rangle \\
&= \frac{1}{\sqrt{q^n}} \sum_{\mathbf{y} \in \mathbb{F}_q^n} \chi_{\mathbf{y}}(\mathbf{e}) \sum_{\mathbf{c} \in \mathcal{C}} \chi_{\mathbf{y}}(\mathbf{c}) |\mathbf{y}\rangle \\
(30) \qquad &= \frac{q^k}{\sqrt{q^n}} \sum_{\mathbf{c}^\perp \in \mathcal{C}^\perp} \chi_{\mathbf{c}^\perp}(\mathbf{e}) |\mathbf{c}^\perp\rangle \quad \text{(since } \sum_{\mathbf{c} \in \mathcal{C}} \chi_{\mathbf{y}}(\mathbf{c}) = 0 \text{ if } \mathbf{y} \notin \mathcal{C}^\perp \text{ and } q^k \text{ otherwise)}
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\widehat{\left| \sum_{\mathbf{e} \in \mathbb{F}_q^n, \mathbf{c} \in \mathcal{C}} \pi_{\mathbf{e}} |\mathbf{c} + \mathbf{e}\rangle \right\rangle} &= \sum_{\mathbf{e} \in \mathbb{F}_q^n} \pi_{\mathbf{e}} \widehat{|\mathbf{1}_{\mathcal{C}+\mathbf{e}}\rangle} \\
&= \frac{q^k}{\sqrt{q^n}} \sum_{\mathbf{e} \in \mathbb{F}_q^n} \pi_{\mathbf{e}} \sum_{\mathbf{c}^\perp \in \mathcal{C}^\perp} \chi_{\mathbf{c}^\perp}(\mathbf{e}) |\mathbf{c}^\perp\rangle \\
&= q^k \sum_{\mathbf{c}^\perp \in \mathcal{C}^\perp} \frac{1}{\sqrt{q^n}} \sum_{\mathbf{e} \in \mathbb{F}_q^n} \pi_{\mathbf{e}} \chi_{\mathbf{c}^\perp}(\mathbf{e}) |\mathbf{c}^\perp\rangle \\
&= q^k \sum_{\mathbf{c}^\perp \in \mathcal{C}^\perp} \widehat{\pi}_{\mathbf{c}^\perp} |\mathbf{c}^\perp\rangle.
\end{aligned}
$$

It follows that

$$\widehat{|\psi_{\text{ideal}}\rangle} = \frac{q^k}{\sqrt{Z}} \sum_{\mathbf{c}^\perp \in \mathcal{C}^\perp} \widehat{\pi}_{\mathbf{c}^\perp} |\mathbf{0}_n\rangle |\mathbf{c}^\perp\rangle.$$

After measurement we get a state $|\mathbf{0}_n\rangle |\mathbf{c}^\perp\rangle$ with probability $\frac{q^{2k}}{Z} \left|\widehat{f}(|\mathbf{c}^\perp|)\right|^2$. By summing over all $\mathbf{c}^\perp \in \mathcal{C}^\perp$ of weight u, we conclude the proof. □

A.3. **Step 3 : Proof of Proposition 3.** We first need for this a good estimation of $\widehat{|\psi_{\text{ideal}}\rangle}$'s amplitudes. This will be a consequence of the following lemma.

Lemma 18. *If the generator matrix* $\mathbf{G}$ *of a code* $\mathcal{C}$ *is chosen uniformly at random in* $\mathbb{F}_q^{k \times n}$ *then the number* $N_u^\perp$ *of codewords of weight* u *in* $\mathcal{C}^\perp$ *satisfies*

$$\mathbb{P}_\mathbf{G}\left(\left| N_u^\perp - \frac{S_u}{q^k} \right| \geq \left(\frac{S_u}{q^k} \right)^{3/4} \right) \leq (q-1)\sqrt{\frac{q^k}{S_u}}.$$

Proof. Let $\mathbb{1}_\mathbf{x}$ be the indicator function of the event "$\mathbf{x} \in \mathcal{C}^\perp$". By definition,

$$(31) \qquad\qquad\qquad N_u^\perp = \sum_{\mathbf{x} \in \mathscr{S}_u} \mathbb{1}_\mathbf{x}.$$

We have $\mathbb{E}(\mathbb{1}_\mathbf{x}) = \mathbb{P}_\mathbf{G}(\mathbf{x} \in \mathcal{C}^\perp) = \frac{1}{q^k}$, implying that $\mathbb{E}(N_u^\perp) = \frac{S_u}{q^k}$. By using Bienaymé-Tchebychevs inequality, we obtain:

$$\mathbb{P}\left(\left| N_u^\perp - \frac{S_u}{q^k} \right| \geq a \right) \leq \frac{\mathbf{Var}(N_u^\perp)}{a^2}$$

$$= \frac{1}{a^2}\left(\sum_{\mathbf{x} \in \mathscr{S}_u} \mathbf{Var}(\mathbb{1}_\mathbf{x}) + \sum_{\substack{\mathbf{x},\mathbf{y} \in \mathscr{S}_u \\ \mathbf{x} \neq \mathbf{y}}} \mathbb{E}(\mathbb{1}_\mathbf{x}\mathbb{1}_\mathbf{y}) - \mathbb{E}(\mathbb{1}_\mathbf{x})\mathbb{E}(\mathbb{1}_\mathbf{y}) \right)$$

$$\leq \frac{1}{a^2}\left(\sum_{\mathbf{x} \in \mathscr{S}_u} \mathbb{E}(\mathbb{1}_\mathbf{x}) + \sum_{\substack{\mathbf{x},\mathbf{y} \in \mathscr{S}_u \\ \mathbf{x} \neq \mathbf{y}}} \mathbb{E}(\mathbb{1}_\mathbf{x}\mathbb{1}_\mathbf{y}) - \mathbb{E}(\mathbb{1}_\mathbf{x})\mathbb{E}(\mathbb{1}_\mathbf{y}) \right)$$

$$(32) \qquad = \frac{1}{a^2}\left(\frac{S_u}{q^k} + \sum_{\substack{\mathbf{x},\mathbf{y} \in \mathscr{S}_u \\ \mathbf{x} \neq \mathbf{y}}} \mathbb{E}(\mathbb{1}_\mathbf{x}\mathbb{1}_\mathbf{y}) - \mathbb{E}(\mathbb{1}_\mathbf{x})\mathbb{E}(\mathbb{1}_\mathbf{y}) \right)$$

where we used that $\mathbf{Var}(\mathbb{1}_\mathbf{x}) \leq \mathbb{E}(\mathbb{1}_\mathbf{x}^2) = \mathbb{E}(\mathbb{1}_\mathbf{x})$. Let us now upper-bound the second term of the inequality. It is readily verified that:

$$\mathbb{E}(\mathbb{1}_\mathbf{x}\mathbb{1}_\mathbf{y}) = \begin{cases} 1/q^k & \text{if } \mathbf{x} \text{ and } \mathbf{y} \text{ are colinear,} \\ 1/q^{2k} & \text{otherwise.} \end{cases}$$

Therefore, we deduce that:

$$\sum_{\substack{\mathbf{x},\mathbf{y} \in \mathscr{S}_u \\ \mathbf{x} \neq \mathbf{y}}} \mathbb{E}(\mathbb{1}_\mathbf{x}\mathbb{1}_\mathbf{y}) - \mathbb{E}(\mathbb{1}_\mathbf{x})\mathbb{E}(\mathbb{1}_\mathbf{y}) = \sum_{\mathbf{x} \in \mathscr{S}_u} \sum_{\substack{\mathbf{y} \in \mathscr{S}_u \backslash \mathbf{x}: \\ \text{colinear to } \mathbf{x}}} \frac{1}{q^k} - \frac{1}{q^{2k}}$$

$$\leq \sum_{\mathbf{x} \in \mathscr{S}_u} \sum_{\substack{\mathbf{y} \in \mathscr{S}_u \backslash \mathbf{x}: \\ \text{colinear to } \mathbf{x}}} \frac{1}{q^k}$$

$$(33) \qquad\qquad\qquad\qquad \leq \frac{(q-2)S_u}{q^k}$$

It gives by plugging (33) in (32):

$$\mathbb{P}_\mathbf{G}\left(\left| N_u^\perp - \frac{S_u}{q^k} \right| \geq a \right) \leq \frac{1}{a^2}\left(\frac{S_u}{q^k} + \frac{(q-2)S_u}{q^k} \right) = \frac{(q-1)S_u}{a^2 q^k}$$

which concludes the proof by choosing $a = \left(\frac{S_u}{q^k} \right)^{3/4}$. $\qquad\qquad\qquad\qquad\qquad \square$

We are ready to prove Proposition 3 which we now recall:

Proposition 3. *Under the assumptions made in Theorem 1, the probability of obtaining a codeword $\mathbf{c}^{\perp} \in \mathcal{C}^{\perp}$ of weight $u \in \mathscr{W}$ when measuring $|\widehat{\psi_{ideal}}\rangle$ is $\geq 1 - \alpha(\pi)$ for a proportion $\geq 1 - \beta(\pi)$ of matrices $\mathbf{G}$, where:*

$$\alpha(\pi) \quad \stackrel{def}{=} \quad \sum_{u \in \mathscr{W}} \left(\frac{q^k}{S_u} \right)^{1/4} + \sqrt{\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}} - 2^{-\Omega(n)},$$

$$\beta(\pi) \quad \stackrel{def}{=} \quad (q-1) \sum_{u \in \mathscr{W}} \sqrt{\frac{q^k}{S_u}} + \sqrt{\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}} + O\left(q^{-\min(k, n-k)} \right).$$

Proof. Let $\mathscr{Q}$ be the quantum algorithm starting from $|\psi\rangle$ which computes $(\mathbf{I} \otimes \mathrm{QFT} \otimes \mathbf{I}) |\psi\rangle$. This algorithm succeeds when measuring a dual codeword $\mathbf{c}^{\perp} \in \mathcal{C}^{\perp}$ of weight $u \in \mathscr{W}$. When starting with $|\psi_{\mathrm{ideal}}\rangle$, the probability of success of $\mathscr{Q}$ is equal to $\sum_{u \in \mathscr{W}} \frac{q^{2k} N_u^{\perp}}{Z} \left| \widehat{f}(u) \right|^2$ by Lemma 4. Let

$$\mathscr{B} \quad \stackrel{def}{=} \quad \left\{ \mathbf{G} \in \mathbb{F}_q^{k \times n} : Z > q^k \left(1 + \sqrt{\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}} \right) \right\} \quad \text{and}$$

$$\mathcal{E}_u \quad \stackrel{def}{=} \quad \left\{ \mathbf{G} \in \mathbb{F}_q^{k \times n} : \left| N_u^{\perp} - \frac{S_u}{q^k} \right| \geq \left(\frac{S_u}{q^k} \right)^{3/4} \right\}.$$

By Lemmas 3 and 18 we have that

$$\mathbb{P}\left(\mathbf{G} \in \mathscr{B} \cup \bigcup_{u \in \mathscr{W}} \mathcal{E}_u \right) \leq \beta(\pi) = (q-1) \sum_{u \in \mathscr{W}} \sqrt{\frac{q^k}{S_u}} + \sqrt{\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}} + O\left(q^{-\min(k, n-k)} \right).$$

Therefore, for a proportion $\geq 1 - \beta(\pi)$ of codes (over matrices $\mathbf{G}$):

(i) $Z \leq q^k \left(1 + \sqrt{\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}} \right)$ and $Z \geq q^k$ (this is true for any $\mathbf{G}$ as $\pi_{\mathbf{e}} \geq 0$ for any $\mathbf{e}$),

(ii) for all u in $\mathscr{W}$, $\left| \frac{q^k N_u^{\perp}}{S_u} - 1 \right| \leq \left(\frac{q^k}{S_u} \right)^{1/4}$.

We deduce that for a proportion $\geq 1 - \beta(\pi)$ of codes and for all u in $\mathscr{W}$:

$$\frac{1 - \left(\frac{q^k}{S_u} \right)^{1/4}}{1 + \sqrt{\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}}} \leq \frac{q^k N_u^{\perp}}{S_u} \frac{q^k}{Z} \leq 1 + \left(\frac{q^k}{S_u} \right)^{1/4}.$$

This implies that for a proportion $\geq 1 - \beta(\pi)$ of codes and for all u in $\mathscr{W}$ we have that

$$\frac{1 - \sum_{u \in \mathscr{W}} \left(\frac{q^k}{S_u} \right)^{1/4}}{1 + \sqrt{\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}}} \leq \frac{q^{2k} N_u^{\perp}}{S_u Z} \leq 1 + \sum_{u \in \mathscr{W}} \left(\frac{q^k}{S_u} \right)^{1/4},$$

from which we deduce that under the same conditions we also have

$$(34) \quad S_u \left| \widehat{f}(u) \right|^2 \left(\frac{1 - \sum_{u \in \mathscr{W}} \left(\frac{q^k}{S_u} \right)^{1/4}}{1 + \sqrt{\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}}} \right)$$

$$\leq \sum_{u \in \mathscr{W}} \frac{q^{2k} N_u^{\perp}}{Z} \left| \widehat{f}(u) \right|^2 \leq S_u \left| \widehat{f}(u) \right|^2 \left(1 + \sum_{u \in \mathscr{W}} \left(\frac{q^k}{S_u} \right)^{1/4} \right).$$

Now,

$$\frac{1 - \sum_{u \in \mathscr{W}} \left(\frac{q^k}{S_u}\right)^{1/4}}{1 + \sqrt{\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}}} \geq 1 - \sum_{u \in \mathscr{W}} \left(\frac{q^k}{S_u}\right)^{1/4} - \sqrt{\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}}$$

Therefore, by plugging this in Equation (34) we have for a proportion $\geq 1 - \beta(\pi)$ of codes:

$$(35) \qquad (1 - \delta) \sum_{u \in \mathscr{W}} S_u \left|\widehat{f}(u)\right|^2 \leq \sum_{u \in \mathscr{W}} \frac{q^{2k} N_u^\perp}{Z} \left|\widehat{f}(u)\right|^2 \leq (1 + \delta) \sum_{u \in \mathscr{W}} S_u \left|\widehat{f}(u)\right|^2$$

where $\delta \stackrel{\text{def}}{=} \sum_{u \in \mathscr{W}} \left(\frac{q^k}{S_u}\right)^{1/4} + \sqrt{\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{n-k}}}$. We finish the proof by applying Lemma 4: we namely know that after measuring the state obtained by $\mathscr{Q}$, we obtain a dual codeword of weight u in $\mathscr{W}$ with probability $\sum_{u \in \mathscr{W}} \frac{q^{2k} N_u^\perp}{Z} \left|\widehat{f}(u)\right|^2$. $\qquad \square$

Appendix B. Proof of Theorem 3

B.1. **Step 1 (choice of $|\pi\rangle$): Proofs of Lemmas 12 and 13.** The following lemma will be very helpful in what follows.

Lemma 19. [BCN89, §9.3, Lem. 9.3.2] *Let V be a subspace of dimension s, then there are exactly $q^{(t-\ell)(s-\ell)} \begin{bmatrix} n-s \\ t-\ell \end{bmatrix}_q \begin{bmatrix} s \\ \ell \end{bmatrix}_q$ subspaces W of dimension t such that $\dim(V \cap W) = \ell$.*

Let us now recall Lemma 12:

Lemma 12. *$|\pi\rangle$ is radial, i.e., we may write $|\pi\rangle$ as*

$$|\pi\rangle = \sum_{\mathbf{E} \in \mathbb{F}_q^{m \times n} : |\mathbf{E}| \leq t} \pi_{\mathbf{E}} |\mathbf{E}\rangle$$

with $|\mathbf{E}\rangle \stackrel{def}{=} |\mathbf{E}_1\rangle \otimes \cdots \otimes |\mathbf{E}_m\rangle$ where the $\mathbf{E}_i$'s denote the rows of $\mathbf{E}$ and $\pi_{\mathbf{E}} = f(|\mathbf{E}|)$ where

$$f(u) = \begin{cases} \frac{\begin{bmatrix} n-u \\ t-u \end{bmatrix}_q}{\sqrt{q^{mt} N}} & \text{if } u \leq t, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Let U be the $\mathbb{F}_q$-space generated by the $\mathbf{E}_i$'s. We denote by u the dimension of U. We have

$$\pi_{\mathbf{E}} = \frac{1}{\sqrt{q^{mt} N}} \, \sharp \left\{ V \leq \mathbb{F}_q^n : \dim V = t \text{ and } U \subseteq V \right\} = \frac{\begin{bmatrix} n-u \\ t-u \end{bmatrix}_q}{\sqrt{q^{mt} N}},$$

where we used Lemma 19 for the last equality. It concludes the proof. $\qquad \square$

Another asymptotic expression for N and an estimate for p_t are given by:

Lemma 13. *We have:*

$$N = \Theta\left(\begin{bmatrix} n \\ t \end{bmatrix}_q\right) \quad \text{and} \quad p_t = \Theta(1).$$

This lemma will be a consequence of the following lemmas.

Lemma 20. *For any $V, W \leq \mathbb{F}_q^n$ such that $V \neq W$ and $\dim V = \dim W = t$ we have,*

$$\langle \pi_V | \pi_W \rangle = q^{m(\dim(V \cap W) - t)}.$$

Proof. Recall that,

$$|\pi_U\rangle = \left(\frac{1}{\sqrt{q^{\dim U}}} \sum_{\mathbf{u} \in U} |\mathbf{u}\rangle\right)^{\otimes m}$$

Therefore we have,

$$\langle \pi_V | \pi_W \rangle = \left(\frac{1}{q^t} \sum_{\mathbf{v} \in V} \sum_{\mathbf{w} \in W} \langle \mathbf{v} | \mathbf{w} \rangle\right)^m = \left(\frac{1}{q^t} \sharp(V \cap W)\right)^m$$

which concludes the proof. □

Lemma 21. *We have,*

$$\sum_{\substack{V \leq \mathbb{F}_q^n \\ \dim V = t}} \sum_{\substack{W \leq \mathbb{F}_q^n \\ \dim W = t \\ W \neq V}} \langle \pi_V | \pi_W \rangle = O\left(\begin{bmatrix} n \\ t \end{bmatrix}_q\right).$$

Proof. We have

$$\sum_{\substack{V \leq \mathbb{F}_q^n \\ \dim V = t}} \sum_{\substack{W \leq \mathbb{F}_q^n \\ \dim W = t \\ W \neq V}} \langle \pi_V | \pi_W \rangle = \sum_{\substack{V \leq \mathbb{F}_q^n \\ \dim V = t}} \sum_{\substack{W \leq \mathbb{F}_q^n \\ \dim W = t \\ W \neq V}} q^{m(\dim(V \cap W) - t)}$$

$$= \sum_{\substack{V \leq \mathbb{F}_q^n \\ \dim V = t}} \sum_{\ell=0}^{t-1} \sum_{\substack{W \leq \mathbb{F}_q^n \\ \dim W = t \\ \dim(W \cap V) = \ell}} \frac{1}{q^{m(t-\ell)}}$$

$$\text{(36)} \qquad = \sum_{\substack{V \leq \mathbb{F}_q^n \\ \dim V = t}} \sum_{\ell=0}^{t-1} \frac{1}{q^{m(t-\ell)}} q^{(t-\ell)^2} \begin{bmatrix} t \\ \ell \end{bmatrix}_q \begin{bmatrix} n-t \\ t-\ell \end{bmatrix}_q$$

$$= \begin{bmatrix} n \\ t \end{bmatrix}_q \sum_{\ell=0}^{t-1} q^{(t-\ell-m)(t-\ell)} \begin{bmatrix} t \\ \ell \end{bmatrix}_q \begin{bmatrix} n-t \\ t-\ell \end{bmatrix}_q$$

where in Equation (36) we used Lemma 19. Now, there exists some constant $c > 0$ such that:

$$\begin{bmatrix} n \\ \ell \end{bmatrix}_q \leq cq^{\ell(n-\ell)}.$$

Then, for some constant $C > 0$,

$$\sum_{\substack{V \leq \mathbb{F}_q^n \\ \dim V = t}} \sum_{\substack{W \leq \mathbb{F}_q^n \\ \dim W = t \\ W \neq V}} \langle \pi_V | \pi_W \rangle \leq C \begin{bmatrix} n \\ t \end{bmatrix}_q \sum_{\ell=0}^{t-1} q^{(t-\ell-m)(t-\ell)+\ell(t-\ell)+(t-\ell)(n-2t+\ell)}$$

$$= C \begin{bmatrix} n \\ t \end{bmatrix}_q \sum_{\ell=0}^{t-1} q^{(t-\ell)(t-\ell-m+\ell+n-2t+\ell)}$$

$$= C \begin{bmatrix} n \\ t \end{bmatrix}_q \sum_{\ell=0}^{t-1} q^{(t-\ell)(-t+\ell-m+n)}$$

$$\leq C \begin{bmatrix} n \\ t \end{bmatrix}_q \sum_{\ell=0}^{t-1} q^{-(t-\ell)^2} \quad \text{(since } n \leq m)$$

which concludes the proof. □

We are now ready to prove Lemma 13.

Proof of Lemma 13. By definition of N we have:

$$N = \left\| \sum_{\substack{V \leq \mathbb{F}_q^n \\ \dim V = t}} \pi_V \right\|^2 = \sum_{\substack{V \leq \mathbb{F}_q^n \\ \dim V = t}} \|\pi_V\|^2 + \sum_{\substack{V \leq \mathbb{F}_q^n \\ \dim V = t}} \sum_{\substack{W \leq \mathbb{F}_q^n \\ \dim W = t \\ W \neq V}} \langle \pi_V, \pi_W \rangle$$

and by definition of π_V:

$$\sum_{\substack{V \leq \mathbb{F}_q^n \\ \dim V = t}} \|\pi_V\|^2 = \begin{bmatrix} n \\ t \end{bmatrix}_q \left(\frac{1}{q^t} \sum_{v \in V} 1 \right)^m = \begin{bmatrix} n \\ t \end{bmatrix}_q$$

This concludes the proof that $N = \begin{bmatrix} n \\ t \end{bmatrix}_q$ by using Lemma 21. Now, by definition of p_t, we have:

$$
\begin{aligned}
p_t &= S_t f(t)^2 \\
&= S_t \frac{\begin{bmatrix} n-t \\ 0 \end{bmatrix}_q^2}{q^{mt} N} \quad \text{(by Lemma 12)} \\
&= \frac{S_t}{\Theta(S_t)} \quad \text{(by using the estimate for } N \text{ and Equation (5))}
\end{aligned}
$$

allowing us to conclude that $p_t = \Theta(1)$. $\qquad\square$

B.2. **Step 3: Proofs of Lemmas 15 and 16.** Recall Lemma 15 first:

Lemma 15. *We have,*

$$\frac{S_t}{q^{mn-k}} = q^{-\Omega(n)} \quad and \quad \frac{\langle \pi | \mathbf{1} \rangle^2}{q^{mn-k}} = q^{-\Omega(n)}.$$

Proof. In order to prove the first equation, note that $d_{\mathrm{GV}}(m,n,k)$ is defined such that $\frac{S_{d_{\mathrm{GV}}}}{q^{mn-k}} \leq 1$. From this and Equation (6) we deduce

$$\frac{S_{d_{\mathrm{GV}}-1}}{q^{mn-k}} = \frac{S_{d_{\mathrm{GV}}-1}}{S_{d_{\mathrm{GV}}}} \frac{S_{d_{\mathrm{GV}}}}{q^{mn-k}} \leq \Theta\left(q^{-(m+n-2d_{\mathrm{GV}}-1)}\right) = \Theta\left(q^{-\Omega(n)}\right)$$

where the last equality follows from the fact that (see for instance [Loi06])

$$(37) \qquad d_{\mathrm{GV}}(n,m,k) = \frac{m+n-\sqrt{(m-n)^2+4k}}{2}(1+o(1)).$$

Now, for the second equation, we have

$$
\begin{aligned}
\frac{\langle \pi | \mathbf{1} \rangle^2}{q^{mn}} &= \left| \widehat{f}(0) \right|^2 \\
&= \frac{\begin{bmatrix} n \\ n-t \end{bmatrix}_q^2}{q^{m(n-t)} N} \quad \text{(by Lemma 14)} \\
&= \Theta\left(\frac{q^{mt} \begin{bmatrix} n \\ t \end{bmatrix}_q}{q^{mn}} \right) \quad \text{(By Lemma (13))} \\
&= \Theta\left(\frac{S_t}{q^{mn}} \right) \quad \text{(By Equations (4) and (5))}
\end{aligned}
$$

Thanks to the first equation, we complete the proof. $\qquad\square$

The second lemma we need is recalled here:

Lemma 16. *For any $\eta > 0$, we have,*

$$\forall u \in [\![(1-\eta)n - t, n - t]\!], \; \frac{q^k}{S_u} = q^{-\Omega(n)} \quad and \quad \sum_{u \in [\![(1-\eta)n-t,n-t]\!]} S_u \left| \widehat{f}(u) \right|^2 = 1 - q^{-\Omega(n)}.$$

Proof. The first equation can be proved in the same way as Lemma 15. For the second identity, first notice (by Lemma 14) that

$$\sum_{u \in [\![n-t-\eta n, n-t]\!]} S_u \left| \widehat{f}(u) \right|^2 = 1 - \sum_{u < n-t-\eta n} S_u \left| \widehat{f}(u) \right|^2.$$

Now, we have the following computation:

$$\sum_{u < n-t-\eta n} S_u \left| \widehat{f}(u) \right|^2 = \sum_{u < n-t-\eta n} S_u \frac{\left[\genfrac{}{}{0pt}{}{n-u}{n-u-t} \right]_q^2}{N \, q^{m(n-t)}} \quad \text{(by Lemma 15)}$$

$$= \Theta \left(\sum_{u < n-t-\eta n} S_u \frac{\left[\genfrac{}{}{0pt}{}{n-u}{n-u-t} \right]_q^2}{\left[\genfrac{}{}{0pt}{}{n}{t} \right]_q q^{m(n-t)}} \right) \quad \text{(by Lemma 13)}$$

$$= \Theta \left(\sum_{u < n-t-\eta n} q^{u(m+n-u)} \frac{q^{2(n-t-u)t}}{q^{t(n-t)} \, q^{m(n-t)}} \right) \quad \text{(By Equations (4) and (5))}$$

$$= \Theta \left(q^{\max_{u < n-t-\eta n} u(m+n-u-2t)} \, q^{(t-m)(n-t)} \right)$$

Let $g(u) \stackrel{\text{def}}{=} u(m + n - u - 2t)$. Then, $g'(u) = m + n - 2(t + u) \geq 0$ as $t + u \leq n \leq m$. Therefore, g is an increasing function and by setting $u = n - t - \eta n$, we obtain

$$\sum_{u < n-t-\eta n} S_u \left| \widehat{f}(u) \right|^2 \leq n \, \Theta \left(q^{(n-t-\eta n)(m-t+\eta n)} \, q^{(t-m)(n-t)} \right)$$

$$= n \, \Theta \left(q^{(n-t)(t-m+m-t+\eta n)} \, q^{-\eta n(m-t+\eta n)} \right)$$

$$= n \, \Theta \left(q^{-\eta n(m-t+\eta n-n+t)} \right)$$

$$= n \, \Theta \left(q^{-\eta n(m-n+\eta n)} \right)$$

$$= q^{-\Omega(n)}$$

as $n \leq m$ by assumption. It concludes the proof.

$\square$

References

[AAB+19]   Carlos Aguilar Melchor, Nicolas Aragon, Slim Bettaieb, Loïc Bidoux, Olivier Blazy, Jean-Christophe Deneuville, Philippe Gaborit, Gilles Zémor, Alain Couvreur, and Adrien Hauteville. Rank quasi cyclic (RQC). Second round submission to the NIST post-quantum cryptography call, April 2019.

[ABD+19]   Nicolas Aragon, Olivier Blazy, Jean-Christophe Deneuville, Philippe Gaborit, Adrien Hauteville, Olivier Ruatta, Jean-Pierre Tillich, Gilles Zémor, Carlos Aguilar Melchor, Slim Bettaieb, Loïc Bidoux, Magali Bardet, and Ayoub Otmani. ROLLO (merger of Rank-Ouroboros, LAKE and LOCKER). Second round submission to the NIST post-quantum cryptography call, March 2019.

[AFS05]   Daniel Augot, Matthieu Finiasz, and Nicolas Sendrier. A family of fast syndrome based cryptographic hash functions. In *Ed Dawson, Serge Vaudenay (editors). Progress cryptology-Mycrypt First international conference on cryptology Malaysia, ISBN 978-3-540-28938-8*, volume 3715 of *LNCS*, pages 64–83, Kuala Lumpur, Malaysia, September 2005. Springer.

[AHI+17]   Benny Applebaum, Naama Haramaty, Yuval Ishai, Eyal Kushilevitz, and Vinod Vaikuntanathan. Low-complexity cryptographic hash functions. In *ITCS*, volume 67 of *LIPIcs*, pages 7:1–7:31. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.

[Ale11]   Michael Alekhnovich. More on average case vs approximation complexity. *Computational Complexity*, 20(4):755–786, 2011.

[BCG+19]   Emanuele Bellini, Florian Caullery, Philippe Gaborit, Marc Manzano, and Víctor Mateu. Improved Veron identification and signature schemes in the rank metric. In *Proc. IEEE Int. Symposium Inf. Theory - ISIT 2019*, volume abs/1903.10212, pages 1872–1876, Paris, France, July 2019. IEEE.

[BCN89]   Andries E. Brouwer, Arjeh M. Cohen, and Arnold Neumaier. *Distance-Regular Graphs*. Number 18 in Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics. Springer Verlag Berlin Heidelberg, 1989.

[BGHM20]   Emanuele Bellini, Philippe Gaborit, Alexandros Hasikos, and Víctor Mateu. Enhancing code based zero-knowledge proofs using rank metric. In Stephan Krenn, Haya Shulman, and Serge Vaudenay, editors, *Cryptology and Network Security - 19th International Conference, CANS 2020, Vienna, Austria, December 14-16, 2020, Proceedings*, volume 12579 of *Lecture Notes in Computer Science*, pages 570–592. Springer, 2020.

[BJMM12]   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*, LNCS. Springer, 2012.

[BLVW19]   Zvika Brakerski, Vadim Lyubashevsky, Vinod Vaikuntanathan, and Daniel Wichs. Worst-case hardness for LPN and cryptographic hashing via code smoothing. In Yuval Ishai and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2019 - 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Darmstadt, Germany, May 19-23, 2019, Proceedings, Part III*, volume 11478 of *LNCS*, pages 619–635. Springer, 2019.

[BM18]   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 2018*, volume 10786 of *LNCS*, pages 25–46, Fort Lauderdale, FL, USA, April 2018. Springer.

[CDMT22]   Kevin Carrier, Thomas Debris-Alazard, Charles Meyer-Hilfiger, and Jean-Pierre Tillich. Statistical decoding 2.0: Reducing decoding to LPN. In *Advances in Cryptology - ASIACRYPT 2022*, LNCS. Springer, 2022.

[CE03]   Henry Cohn and Noam Elkies. New upper bounds on sphere packings I. *Ann. of Math. (2)*, 157(2):689–714, 2003.

[Che]   Yilei Chen. personal communication.

[CLZ22]   Yilei Chen, Qipeng Liu, and Mark Zhandry. Quantum algorithms for variants of average-case lattice problems via filtering. In Orr Dunkelman and Stefan Dziembowski, editors, *Advances in Cryptology - EUROCRYPT 2022 - 41st Annual International Conference on the Theory and Applications of Cryptographic Techniques, Trondheim, Norway, May 30 - June 3, 2022, Proceedings, Part III*, volume 13277 of *LNCS*, pages 372–401. Springer, 2022.

[CV]   Yilei Chen and June Vuong. Quantum reduction from binary SIS to LPN and more via generalized lattice tail bounds. preprint.

[DDRT22]   Thomas Debris-Alazard, Léo Ducas, Nicolas Resch, and Jean-Pierre Tillich. Smoothing codes and lattices: Systematic study and new bounds. *CoRR*, abs/2205.10552, 2022.

[Del72]   Philippe Delsarte. Bounds for unrestricted codes. *Philips Res. Rep.*, 27:272–289, 1972.

[DL98]   Philippe Delsarte and Vladimir Iossifovitch Levenshtein. Association schemes and coding theory. *IEEE Trans. Inform. Theory*, 44(6):2477–2504, 1998.

[DR22]   Thomas Debris-Alazard and Nicolas Resch. Worst and average case hardness of decoding via smoothing bounds. preprint, December 2022. eprint.

[Dum89]   Il'ya Dumer. Two decoding algorithms for linear codes. *Probl. Inf. Transm.*, 25(1):17–23, 1989.

[For01]   G. David Jr. Forney. Codes on graphs: Normal realizations. *IEEE Trans. Inform. Theory*, 47(2):520–548, 2001.

[FS96]   Jean-Bernard Fischer and Jacques Stern. An efficient pseudo-random generator provably as secure as syndrome decoding. In Ueli Maurer, editor, *Advances in Cryptology - EUROCRYPT'96*, volume 1070 of *LNCS*, pages 245–255. Springer, 1996.

[HTW20]   Anna-Lena Horlemann-Trautmann and Violetta Weger. Information set decoding in the lee metric with applications to cryptography. *Advances in Mathematics of Communications*, 0, 2020. online version, to appear.

[Lev79]   Vladimir Iossifovitch Levenshtein. Bounds for packings in $n-$dimensional Euclidean space. *Soviet Math. Dokl.*, 20:417–421, 1979.

[Loi06]   Pierre Loidreau. Properties of codes in rank metric, 2006.

[McE78]   Robert J. McEliece. *A Public-Key System Based on Algebraic Coding Theory*, pages 114–116. Jet Propulsion Lab, 1978. DSN Progress Report 44.

[MMT11]   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 *LNCS*, pages 107–124. Springer, 2011.

[MO15]   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 *LNCS*, pages 203–228. Springer, 2015.

[MRRW77]   R. J. McEliece, E. R. Rodemich, H. Rumsey, and L. R. Welch. New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities. *IEEE Trans. Inform. Theory*, 23(2):157–166, 1977.

[MTSB12]   Rafael Misoczki, Jean-Pierre Tillich, Nicolas Sendrier, and Paulo S. L. M. Barreto. MDPC-McEliece: New McEliece variants from moderate density parity-check codes, 2012.

[NC16]   Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information (10th Anniversary edition)*. Cambridge University Press, 2016.

[Pra62]   Eugene Prange. The use of information sets in decoding cyclic codes. *IRE Transactions on Information Theory*, 8(5):5–9, 1962.

[Reg05]   Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*, pages 84–93, 2005.

[Reg09]   Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. Extended version of [Reg05], dated May 2009, 2009.

[Sho94]   Peter W. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In S. Goldwasser, editor, *FOCS*, pages 124–134, 1994.

[Sim94]   Daniel R. Simon. On the power of quantum computation. In *35th Annual Symposium on Foundations of Computer Science, Santa Fe, New Mexico, USA, 20-22 November 1994*, pages 116–123. IEEE Computer Society, 1994.

[SSTX09]   Damien Stehlé, Ron Steinfeld, Keisuke Tanaka, and Keita Xagawa. Efficient public key encryption based on ideal lattices. In Mitsuru Matsui, editor, *Advances in Cryptology - ASIACRYPT 2009, 15th International Conference on the Theory and Application of Cryptology and Information Security, Tokyo, Japan, December 6-10, 2009. Proceedings*, volume 5912 of *LNCS*, pages 617–635. Springer, 2009.

[Ste88]   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 *LNCS*, pages 106–113. Springer, 1988.

[Ste93]   Jacques Stern. A new identification scheme based on syndrome decoding. In D.R. Stinson, editor, *Advances in Cryptology - CRYPTO'93*, volume 773 of *LNCS*, pages 13–21. Springer, 1993.

[YZW+19]   Yu Yu, Jiang Zhang, Jian Weng, Chun Guo, and Xiangxue Li. Collision resistant hashing from subexponential learning parity with noise. In *ASIACRYPT (2)*, volume 11922 of *Lecture Notes in Computer Science*, pages 3–24. Springer, 2019.