# QUANTUM REDUCTION OF FINDING SHORT CODE VECTORS TO THE DECODING PROBLEM

THOMAS DEBRIS–ALAZARD, MAXIME REMAUD, AND JEAN-PIERRE TILLICH

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 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 can not be used for the reduction with codes. Instead we choose here the uniform distribution over errors of a fixed weight and bring in orthogonal polynomials tools to perform the analysis and an additional amplitude amplification step to obtain the aforementioned result.

## 1. Introduction

**Code-based Cryptography.** Many cryptosystems 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

**Definition 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 decode the code $\mathcal{C}$ which is the $k$-dimensional vector space generated by the rows of $\mathbf{G}$:

$$\mathcal{C}\stackrel{\text{def}}{=}\{\mathbf{uG}:\ \mathbf{u}\in\mathbb{F}_q^k\}, \tag{1}$$

*i.e.* 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 $t$, $k$ and $n-k$ are all linear in $n$.

Usually this decoding problem is considered 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 $\mathsf{DP}(q,n,k,t)$ 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.

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

- *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 $\mathbf{Hc}^\intercal = \mathbf{0}$ and the weight of $\mathbf{c}$ belongs to $(0,w]$*

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:

$$\mathcal{C} \stackrel{\text{def}}{=} \left\{\mathbf{c} \in \mathbb{F}_q^n : \mathbf{Hc}^\intercal = \mathbf{0}\right\}.$$

Decoding and looking for short codewords are problems that have been conjectured for a long time to be extremely close. They have been studied for a long time [Pra62, Ste88, Dum89, MMT11, BJMM12, MO15, BM18], and for instance in the regime of parameters where the rate $R = \frac{k}{n}$ 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]. 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]. A breakthrough there was to obtain a quantum reduction from SIS to LWE [Reg05, SSTX09]. 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 $n$).

**Parameter range for $\mathsf{DP}$ and $\mathsf{SCP}$.** An important parameter for the reduction is the distance decoding $t$. 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 $d_{\min} \stackrel{\text{def}}{=} \min\{d(\mathbf{c},\mathbf{c}') : \mathbf{c} \in \mathcal{C},\ \mathbf{c}' \in \mathcal{C},\ \mathbf{c} \neq \mathbf{c}'\}$ 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 (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 $d_{\mathrm{GV}}$ if there is no ambiguity). It is defined (for all metrics) for a code of dimension $k$ and length $n$, as the largest integer $t$ for which

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

where $B_t$ is the size of a ball of radius $t$. For the Hamming metric we have

$$\begin{aligned} \delta_{\mathrm{GV}}(n,k) &= h_q^{-1}\left(1-\frac{k}{n}\right) + O\left(\frac{1}{n}\right) \quad \text{where} \\ h_q(x) &\stackrel{\text{def}}{=} -x\log_q\left(\frac{x}{q-1}\right) - (1-x)\log_q(1-x) \ \text{ and } h_q^{-1} \text{ its inverse ranging over } [0,\frac{q-1}{q}] \end{aligned}$$

and $\delta_{\mathrm{GV}}(n,k)$ is the normalized Gilbert-Varshamov distance defined as $\delta_{\mathrm{GV}}(n,k) \stackrel{\text{def}}{=} \frac{d_{\mathrm{GV}}(n,k)}{n}$.
This Gilbert-Varshamov distance turns out to quantify also the region where we *typically* have unique decoding. More precisely, it turns out that the same probabilistic arguments also show

that the solution of the decoding problem is unique with probability $1-2^{-\Omega(n)}$ as long as $t \leq (1-\varepsilon)d_{\mathrm{GV}}(n,k)$ as $n$ goes to infinity for fixed positive $\varepsilon$.

Whereas the best algorithms for solving the decoding have exponential complexity in $n$ as soon as $t$ is linear in $n$ and the code rate $R \stackrel{\text{def}}{=} \frac{k}{n}$ is bounded away from 0 and 1, this is not true for the short codeword problem which becomes easy when the weight $w$ is above a certain range. The reason for this, 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 small weight. It is straightforward that this strategy produces in polynomial time, for instance with the Hamming metric, codewords of weight $\approx \omega_{\mathrm{easy}}(n,k)n$ where

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

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 $\omega_{\mathrm{easy}}(n,k)$ 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 $H^{\perp}$.*

Arguably this innocent looking fact (together with the fact that the Fourier transform can be performed in polylog time when the group $G$ is Abelian) is 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 :
(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 Fourier transform 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 $\mathcal{C}^{\perp} = \{\mathbf{c}^{\perp} \in \mathbb{F}_q^n : \mathbf{c}\cdot\mathbf{c}^{\perp} = 0,\ \forall \mathbf{c}^{\perp} \in \mathcal{C}\}$ (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

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

where $Z$ is a normalizing constant and $(|\pi_{\mathbf{e}}|^2)_{\mathbf{e}}$ is a probability distribution on errors that concentrates 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) disentangle it by 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 quantum Fourier transform on $\mathbb{F}_q^n$ to obtain a superposition of elements $\mathbf{c}^\perp$ in the dual code

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

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

The second point is a direct consequence of Proposition 1. The last point raises the issue of whether or not the Fourier transform concentrates the weight output by this algorithm on weights $t'$ for which finding a codeword in $\mathcal{C}^\perp$ is not known to be easy, as 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 already been pursued in the binary case by Yilei Chen [Che] and later on in [CV], where basically the following approach was taken. The natural analogue of the Gaussian noise model used in Regev's reduction [Reg09] in the case of the Hamming metric is the i.i.d. Bernoulli model on each coordinate, i.e. $|\pi_{\mathbf{e}}|^2 = (p/(q-1))^{|\mathbf{e}|}(1-p)^{n-|\mathbf{e}|}$ where $|\mathbf{e}|$ stands for the Hamming weight of $\mathbf{e}$, $n$ the length of $\mathbf{e}$ and $p$ the parameter of the Bernoulli noise. Both distributions can be expressed in terms of the heat kernel operator (the usual Laplacian in the case of the Gaussian noise and the discrete Laplacian for the case of $\mathbb{F}_q^n$), in both cases the Fourier transform yields a dual noise which is itself Gaussian or Bernoulli based and the quantum state corresponding to the error is a product state which simplifies a great deal the computation. However, it has been realized that this natural approach hits a wall. The problem is the following: we begin to choose the parameter $p$ of the Bernoulli noise, so that the typical weight $pn$ of an error $\mathbf{e}$ is equal to or slightly below the weight $t$ we can decode. It turns out that the most likely weight we measure at Step 3 is always zero if we want to have a chance that the dual codeword we measure has normalized weight $< \omega_{\text{easy}}$, i.e. is in the regime where there is a chance that it is difficult to produce such words. In other words, the straightforward application of Regev's approach to coding theory fails to give a useful reduction.

We give in Section C of the appendix an explanation for the failure of this approach. It can be summarized by saying that the Bernoulli noise model is not enough concentrated on the typical weight $pn$. Intuitively what is going wrong in the case of the Bernoulli noise model, can be explained by bringing in the quantum state

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

which represents in some sense the noise we add to the codeword.

In the case of an i.i.d Bernoulli noise of parameter $p$ (i.e. a $q$-ary symmetric channel of crossover probability $p$) we have

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

Applying the quantum Fourier transform to this state yields a quantum state

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

$$\text{where } p^\perp \stackrel{\text{def}}{=} \frac{\left(\sqrt{(q-1)(1-p)} - \sqrt{p}\right)^2}{q}$$

The issue is to understand which is the most likely weight we get when we measure the quantum state at Step 3. This should be the integer $w$ which maximizes the probability $p_w$ to measure a dual codeword $\mathbf{c}^\perp$ of weight $w$ which is equal to (see Lemma 4 in Section A of the appendix)

$$p_w = \frac{q^{2k}}{Z} \sum_{\mathbf{c}^\perp\in\mathcal{C}^\perp : |\mathbf{c}^\perp|=w} |\widehat{\pi}_{\mathbf{c}^\perp}|^2 \tag{4}$$

where $\widehat{|\pi\rangle} = \sum_{\mathbf{e}\in\mathbb{F}_q^n} \widehat{\pi}_{\mathbf{e}}\,|\mathbf{e}\rangle$. In our case, $\widehat{\pi}$ is a radial function (namely $\widehat{\pi}_{\mathbf{e}} = f(|\mathbf{e}|)$ for some function $f$). Therefore we have (for $|\mathbf{e}| = w$): $p_w = \frac{q^{n+k}}{Z}\,\frac{N_w^\perp}{q^{n-k}}\,|\widehat{\pi}_{\mathbf{e}}|^2$ where $N_w^\perp$ is the number of codewords of weight $w$ in $\mathcal{C}^\perp$. The density of codewords of Hamming weight $w$ in a random code $\mathcal{C}^\perp$ is about the same as the density of elements of weight $w$ in the whole space $\mathbb{F}_q^n$, namely $\frac{1}{q^{n-k}}\,N_w^\perp \approx \frac{(q-1)^w\binom{n}{w}}{q^n}$. However there is one notable exception, namely when $w = 0$ where the density of codewords is $\frac{1}{q^{n-k}}$ instead of $\frac{1}{q^n}$. In other words, if $\widehat{\pi}_{\mathbf{0}}$ is too big, it is in 0 that $p_w$ is maximal. For $w > 0$, we expect that the maximum of $w$ is attained for weights where the probability of measuring a weight $w$ is maximal when we measure directly the quantum state $\widehat{|\pi\rangle}$, in this case the most likely outcome is a weight $\approx p^\perp n$. We miss to measure this quantity since in our case $|\widehat{\pi}_{\mathbf{0}}|^2$ is just too large. It is readily seen that

$$\widehat{\pi}_{\mathbf{0}} = \frac{\langle\pi|\mathbf{1}\rangle}{\sqrt{q^n}}$$

(where $\mathbf{1}$ is the all one vector and $\langle|\rangle$ the hermitian product) and that more or less the best we can do to minimize this quantity if we stick to (i) radial probability distributions $|\pi_{\mathbf{e}}|^2$ (i.e. depending only on the weight of $\mathbf{e}$ –which is a natural constraint), (ii) distributions that concentrate almost all their probability around weight $t$, and (iii) nonnegative $\pi_{\mathbf{e}}$'s[(1)] is to actually concentrate the whole distribution on the weight $t$.

**Our approach.** This is precisely what we have done by choosing

$$|\pi\rangle = \sum_{\mathbf{e}:|\mathbf{e}|=t} \frac{1}{\sqrt{S_t}}\,|\mathbf{e}\rangle$$

where $S_t$ is the cardinality of the sphere of radius $t$ in the Hamming metric, namely $(q-1)^t\binom{n}{t}$. Making this choice actually complicates rather significantly the reasoning. Understanding which weight $w$ maximizes $p_w$ is much more involved (it involves in particular rather delicate properties

[(1)] This restriction simplifies significantly the proof in several crucial places. Furthermore if we ask ourselves whether in the Bernoulli case we can improve upon the choice of $\pi_{\mathbf{e}}$ by keeping the equality $|\pi_{\mathbf{e}}|^2 = (p/(q-1))^{|\mathbf{e}|}(1-p)^{n-|\mathbf{e}|}$ and choosing $\pi_{\mathbf{e}}$ in the complex numbers, it turns out that the choice of positivity for the $\pi_{\mathbf{e}}$ that was made leads to a dual distribution $\widehat{|\pi\rangle}$ which is concentrated on the smallest possible weight, namely $p^\perp n$ here.

of Krawtchouk polynomials). However, it will turn out that when $\omega \stackrel{\text{def}}{=} \frac{w}{n}$ lies in a whole interval starting precisely at $\tau^\perp \stackrel{\text{def}}{=} \frac{\left(\sqrt{(q-1)(1-\tau)}-\sqrt{\tau}\right)^2}{q}$ where $\tau \stackrel{\text{def}}{=} \frac{t}{n}$, we have many points where $p_w$ is actually $\frac{1}{\mathsf{poly}(n)}$. The weight distribution is in this case not really concentrated on a single value but is spread out on a large interval (but with the property that there are weights $w$ close to $\tau^\perp n$ for which $p_w = \frac{1}{\mathsf{poly}(n)}$). In other words, the previous Bernoulli noise model (with probability $\tau$) now captures with its most likely weight outcome $\tau^\perp n$ when we measure $\widehat{|\pi\rangle}$ directly, weights which are rather likely to be output at Step 3 (and 0 is now not anymore the most likely outcome).

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

Moreover by building upon the proof technique of [SSTX09] we can show a reduction which is more relevant to cryptography. We consider that we have a decoding algorithm which is only successful for some potentially very small probability $\varepsilon$ and we want to turn it into an algorithm outputting a word of weight $\approx \tau^\perp n$ in the dual code $\mathcal{C}^\perp$ with some probability $\mathsf{poly}(\varepsilon)$. The "ideal" version of the algorithm that we have presented before (where we assume that we are always successful with our decoding algorithm) only describes a state that we get at Step 2 which 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 + quantum Fourier step. If we were to measure this state directly at that point we would not be sure to measure with probability $\mathsf{poly}(\varepsilon)$ a word of weight $\approx \tau^\perp n$. To ensure this, we have to apply a quantum amplification step that would produce in the ideal case a state which is concentrated on a certain weight $w$ close to $\tau^\perp n$. It is based on the fact that for a random code we know with probability $1-2^{-\Omega(n)}$ with exponential precision the probability of measuring a weight $w$ close to $\tau^\perp n$. Putting all these ingredients together we are able to prove the following result.

**Theorem 1** (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*

$$\begin{aligned} \tau &\stackrel{def}{=} \frac{t}{n} \\ \tau^\perp &\stackrel{def}{=} \frac{\left(\sqrt{(q-1)(1-\tau)}-\sqrt{\tau}\right)^2}{q} \end{aligned}$$

It will turn out that for $q=2$ (see Section 3) we can find for any rate $R=\frac{k}{n}$ in $(0,1)$ a $t < d_{\text{GV}}(n,k)$ for which the corresponding $w$ is below $\omega_{\text{easy}}(n,n-k)n$ (the reduction is useful in this case). Unfortunately, this is not true anymore when $q \geq 5$, where there is always a range for $R$ for which $w$ is above $\omega_{\text{easy}}(n,n-k)n$ and this for any choice of $t < d_{\text{GV}}(n,k)$: 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 reflects in the fact that the range of values of $R$ where this reduction is useful gets

smaller. The whole approach that we have followed here (properly choosing the error distribution, if needed go beyond the unique decoding radius for decoding, and apply a subsequent amplification step if needed) can of course be adapted to other metrics. It is easy for instance to apply it for the rank metric [DRT21] which becomes increasingly popular in code-based cryptography, see for instance [ABD+19, AAB+19, BCG+19, BGHM20]. This metric is even coarser: on $\mathbb{F}_q^{m\times n}$ there are only $1+\min(m,n)$ different values for the metric. In this case, it can be verified that the reduction is always useless (i.e. reduces to weights which are always easy to produce for a random linear code). However, it should be interesting to investigate it 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.

**Notation.** For $a$ and $b$ integers with $a \leq b$, we denote by $[\![a,b]\!]$ the set of integers $\{a, a+1, \ldots, b\}$. 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 $\mathbf{H}$). $\mathscr{S}_t$ 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. $\mathsf{poly}(n)$ denotes a quantity which is an $O(n^a)$ for some constant $a$.

## 2. Quantum Reduction from Sampling Short Codewords to Decoding

2.1. **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:* $\mathcal{C} = \{\mathbf{uG} : \mathbf{u} \in \mathbb{F}_q^k\}$) and a noisy codeword $\mathbf{c}+\mathbf{e}$ where $\mathbf{c}$ belongs to $\mathcal{C}$. We denote by $\mathbf{w} \in \{0,1\}^\ell$ the internal coins of $\mathscr{A}$. It outputs with a certain probability $\varepsilon$, the "right" $\mathbf{e}$ when being fed with $\mathbf{c}+\mathbf{e}$ where $\mathbf{c}$ is chosen uniformly at random in $\mathcal{C}$ and $\mathbf{e}$ is chosen uniformly at random among the errors of weight $t$:

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

We can implement $\mathscr{A}$ quantumly in the following way: it maps the quantum state $|\mathbf{e}\rangle|\mathbf{c}+\mathbf{e}\rangle|\mathbf{w}\rangle$ to $|\mathbf{e}-\mathscr{A}(\mathbf{G},\mathbf{c}+\mathbf{e},\mathbf{w})\rangle|\mathbf{c}+\mathbf{e}\rangle|\mathbf{w}\rangle$. The quantum reduction starts by building the initial superposition

$$|\psi_0\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2^\ell q^k}} \sum_{\mathbf{e}\in\mathbb{F}_q^n} \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{w}\in\mathbb{F}_2^\ell} \pi_{\mathbf{e}} |\mathbf{e}\rangle|\mathbf{c}\rangle|\mathbf{w}\rangle$$

where $|\pi\rangle \stackrel{\text{def}}{=} \sum_{\mathbf{e}\in\mathbb{F}_q^n} \pi_{\mathbf{e}}|\mathbf{e}\rangle$ is some quantum superposition of errors. The quantum algorithm which gives the reduction can be described as follows.

**Algorithm of the quantum reduction.**

| | | | |
|---|---|---|---|
| | Initial state preparation | $=$ | $\frac{1}{\sqrt{2^\ell q^k}} \sum_{\mathbf{e}\in\mathbb{F}_q^n} \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{w}\in\mathbb{F}_2^\ell} \pi_{\mathbf{e}} \lvert\mathbf{e}\rangle\lvert\mathbf{c}\rangle\lvert\mathbf{w}\rangle$ |
| | adding $\mathbf{e}$ to $\mathbf{c}$: | $\mapsto$ | $\frac{1}{\sqrt{2^\ell q^k}} \sum_{\mathbf{e}\in\mathbb{F}_q^n} \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{w}\in\mathbb{F}_2^\ell} \pi_{\mathbf{e}} \lvert\mathbf{e}\rangle\lvert\mathbf{c}+\mathbf{e}\rangle\lvert\mathbf{w}\rangle$ |
| (5) | applying $\mathscr{A}$: | $\stackrel{\mathscr{A}}{\mapsto}$ | $\frac{1}{\sqrt{2^\ell q^k}} \sum_{\mathbf{e}\in\mathbb{F}_q^n} \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{w}\in\mathbb{F}_2^\ell} \pi_{\mathbf{e}} \lvert\mathbf{e}-\mathscr{A}(\mathbf{G},\mathbf{c}+\mathbf{e},\mathbf{w})\rangle\lvert\mathbf{c}+\mathbf{e}\rangle\lvert\mathbf{w}\rangle$ |
| (6) | QFT on the 2nd register: | $\mapsto$ | $\left\lvert\psi_{\mathscr{A}}^{\text{QFT}}\right\rangle$ |
| (7) | Amplification of amplitudes: | $\mapsto$ | $\left\lvert\psi_{\mathscr{A}}^{\text{Ampl}}\right\rangle$ |
| (8) | measuring the whole state: | $\mapsto$ | $\lvert\mathbf{e}\rangle\left\lvert\mathbf{c}^\perp\right\rangle\lvert\mathbf{w}\rangle$ |

We will now give a general theorem about an algorithm $\mathscr{A}$ of this kind and will show that it succeeds with probability $\mathsf{poly}(\varepsilon)$ to output a codeword of the dual code $\mathcal{C}^\perp$ of some weight $u$ by using only a polynomial number of calls to $\mathscr{A}$ when certain conditions are met.

**Theorem 2.** *Assume that* $|\pi\rangle = \sum_{\mathbf{e}\in\mathbb{F}_q^n} |\mathbf{e}\rangle$ *is radial and nonnegative, i.e* $\pi_{\mathbf{e}} = f(|\mathbf{e}|)$ *for some function* $f$ *and* $\pi_{\mathbf{e}} \geq 0$ *for all* $\mathbf{e} \in \mathbb{F}_q^n$. *Let* $p_t \stackrel{def}{=} \sum_{\mathbf{e}:|\mathbf{e}|=t} = S_t f(t)^2$. $\widehat{|\pi\rangle} = \sum_{\mathbf{e}\in\mathbb{F}_q^n} \widehat{\pi}_{\mathbf{e}} |\mathbf{e}\rangle$ *is radial too and we let* $f^\perp(u) = \widehat{\pi}_{\mathbf{e}}$ *where* $\mathbf{e}$ *is any element of* $\mathbb{F}_q^n$ *of Hamming weight* $u$. *Furthermore, assume that :*

$$\frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}} = 2^{-\Omega(n)}, \quad \frac{q^k}{S_u} = 2^{-\Omega(n)} \quad \text{and} \quad S_u|f^\perp(u)|^2 = \Omega\left(\frac{1}{\mathsf{poly}(n)}\right) \text{ for some } 1 \leq u \leq n.$$

*with* $|\mathbf{1}\rangle$ *being the (unnormalized) superposition of all errors :* $|\mathbf{1}\rangle \stackrel{def}{=} \sum_{\mathbf{e}\in\mathbb{F}_q^n} |\mathbf{e}\rangle$.

*Suppose that there exists an algorithm* $\mathscr{A}$ *solving the decoding problem with success probability* $\varepsilon$. *Then, there exists a quantum algorithm making only a polynomial number of calls to* $\mathscr{A}$ *and to additional elementary* 1 *or* 2 *qubit gates 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* $\mathcal{C}^\perp$ *with probability bigger than* $\frac{p_t^2\varepsilon^3}{16} - O(p_t^4\varepsilon^5) - 2^{-\Omega(n)} - O\left(q^{-\min(k,n-k)}\right)$.

**Remark 1.** *This theorem is stated here for the Hamming metric, but actually it can be applied to* any *metric for which the Fourier transform is radially preserving: it also applies for instance to the rank metric.*

This theorem will follow from a sequence of lemmas. Roughly speaking the proof uses the following steps.

Step 1. The first one explains that after applying $\mathscr{A}$ in the previous reduction we get a state

$$|\psi_{\mathscr{A}}\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2^\ell q^k}} \sum_{\mathbf{e}\in\mathbb{F}_q^n} \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{w}\in\mathbb{F}_2^\ell} \pi_{\mathbf{e}} |\mathbf{e} - \mathscr{A}(\mathbf{G}, \mathbf{c}+\mathbf{e}, \mathbf{w})\rangle |\mathbf{c}+\mathbf{e}\rangle |\mathbf{w}\rangle$$

which is sufficiently close to the "disentangled" state

$$|\psi_{\text{ideal}}\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{Z}} \sum_{\mathbf{e}\in\mathbb{F}_q^n} \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{w}\in\mathbb{F}_2^\ell} \pi_{\mathbf{e}} |\mathbf{0}_n\rangle |\mathbf{c}+\mathbf{e}\rangle |\mathbf{w}\rangle \tag{9}$$

where $Z$ is a normalizing constant ensuring that the quantum state is indeed valid (*i.e.* is of norm 1).

Step 2. We then analyze the effect of the Fourier transform on the "ideal state" $|\psi_{\text{ideal}}\rangle$, which gives a quantum state $\left|\psi_{\text{ideal}}^{\text{QFT}}\right\rangle$, and we study a subsequent measure of it. We namely prove that measuring it would output a codeword $\mathbf{c}^\perp$ in $\mathcal{C}^\perp$ of weight $u$ with probability $\frac{1}{\mathsf{poly}(n)}$.

Step 3. The last amplitude step is a unitary transform $\mathscr{U}$ which when applied to $\left|\psi_{\text{ideal}}^{\text{QFT}}\right\rangle$ would output a state $\left|\psi_{\text{ideal}}^{\text{Ampl}}\right\rangle$ whose norm is concentrated up to a negligible $2^{-\Omega(n)}$ term on codewords $\mathbf{c}^\perp$ of weight $u$. We prove in this step how this can be achieved by making only a polynomial number of calls to $\mathscr{A}$ under the assumptions of Theorem 2 and finish the proof by proving that if we apply $\mathscr{U}$ to $\left|\psi_{\mathscr{A}}^{\text{QFT}}\right\rangle$ then the conclusion of the theorem follows.

The reason why we use an amplification step is that because of the assumption on the success probability of our decoding algorithm we can only say that $\left|\psi_{\mathscr{A}}^{\text{QFT}}\right\rangle$ is at trace distance $\leq 1-\eta$ of $\left|\psi_{\text{ideal}}^{\text{QFT}}\right\rangle$ for some $\eta > 0$ depending polynomially on $\varepsilon$. This implies that the distribution of

outcomes after measurement between these states will be at statistical distance $\leq 1-\eta$ but even if measuring $\left|\psi_{\text{ideal}}^{\text{QFT}}\right\rangle$ would yield a codeword $\mathbf{c}^\perp$ in $\mathcal{C}^\perp$ of weight $u$ with probability $\frac{1}{\mathsf{poly}(n)}$, measuring $\left|\psi_{\mathscr{A}}^{\text{QFT}}\right\rangle$ might give such a codeword with probability 0. This is not true anymore if we were to reason now on $\left|\psi_{\text{ideal}}^{\text{Ampl}}\right\rangle$. It is at trace distance $\leq 1-\eta$ from $\left|\psi_{\mathscr{A}}^{\text{Ampl}}\right\rangle$, but now measuring $\left|\psi_{\mathscr{A}}^{\text{Ampl}}\right\rangle$ yields necessarily a codeword $\mathbf{c}^\perp$ in $\mathcal{C}^\perp$ of weight $u$ with probability $\geq \eta - 2^{\Omega(n)}$ since the norm of $\left|\psi_{\text{ideal}}^{\text{Ampl}}\right\rangle$ is exponentially concentrated around such codewords and the trace distance between the outcomes of measuring $\left|\psi_{\mathscr{A}}^{\text{Ampl}}\right\rangle$ and $\left|\psi_{\text{ideal}}^{\text{Ampl}}\right\rangle$ is also $\leq 1-\eta$.
The detailed proofs of these steps are given in Section A of the appendix.

2.2. **Application to the Hamming metric.** The assumptions of Theorem 2 will be satisfied for the Hamming metric for a weight $u$ close to $\tau^\perp n$ and we will be able to prove that

**Theorem 3.** *Suppose that there exists an algorithm $\mathscr{A}$ solving with success probability $\varepsilon$ the decoding problem at Hamming distance $1 \leq t \stackrel{def}{=} \tau n \leq \min\left(\frac{n}{q}, d_{\text{GV}}(n,k)(1-\delta)\right)$ for some $\delta > 0$. Then, there exists a quantum algorithm making only a polynomial number of calls to $\mathscr{A}$ and to additional elementary 1 or 2 qubit gates which takes as input $\mathbf{G} \in \mathbb{F}_q^{k\times n}$, a generator matrix of $\mathcal{C}$, and which outputs $\mathbf{c}^\perp \in \mathcal{C}^\perp$ of weight $u \stackrel{def}{=} (\tau^\perp + o(1))n$ with probability (over a uniform choice of $\mathbf{G}$) bigger than $\frac{\varepsilon^3}{16} - O(\varepsilon^5) - 2^{-\Omega(n)}$ where:*

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

This theorem is proved in Section B of the appendix.

**Remark 2.** *The term $\frac{n}{q}$ which appears in the upper-bound for the range of values for which we can apply our reduction is of no importance in the most important case, namely in the binary case ($q=2$) and for larger values of $q$, $[1, \frac{n}{q}]$ contains a significant part of the values of $t$ for which the reduction is useful (see Section 3). This term $\frac{n}{q}$ comes from the way we estimated Krawtchouk polynomials in the interval formed by their first and last zeros. We could have used* [IS98] *but this would involve lengthy computations (but would improve the $\frac{n}{q}$ term to $\frac{q-1}{q}n$). However the method we used, relying on a generalization of* [KS21, Prop. 25] *avoids a lot of computations and is much shorter.*

## 3. About the usefulness of our reduction.

It is now interesting to look at the parameters for which this reduction is useful. In the example of Fig. 1, when $q=2$, when we take a code of rate $R=\frac{1}{2}$, we can see that $\tau^\perp$ is always above $\omega_{easy}$ when $\tau < \delta_{GV}(R)/2$, hence the necessity of going beyond the unique decoding radius. An important point- at some point when $\tau$ goes beyond $\delta_{GV}(R)/2$, $\tau^\perp$ reaches $\omega_{easy}$ and goes below as $\tau$ goes to $\delta_{GV}(R)$, meaning that we are able with our reduction to get a codeword of a weight in the range where the problem is not known to be easy (above $\delta_{GV}(1-R)$ which corresponds to the smallest possible relative weight where there might be a solution and below $\omega_{easy}$). As $q$ grows, it appears that the range of values $\tau$ for which the reduction is useful shrinks. We can see that by taking a look at the cases where $q=7$ or $q=57$. In the latter, for example, the reduction is never useful.

Now, if we plot $\tau^\perp$ against $R$ when $\tau = \delta_{GV}(R)$ (Fig. 2, when $q=2$), the range of values for which the short codeword problem is hard becomes more apparent. In fact, it is always between $\delta_{GV}(1-R)$ (below which there is no non-trivial codeword) and $\omega_{easy}$ (above which the problem

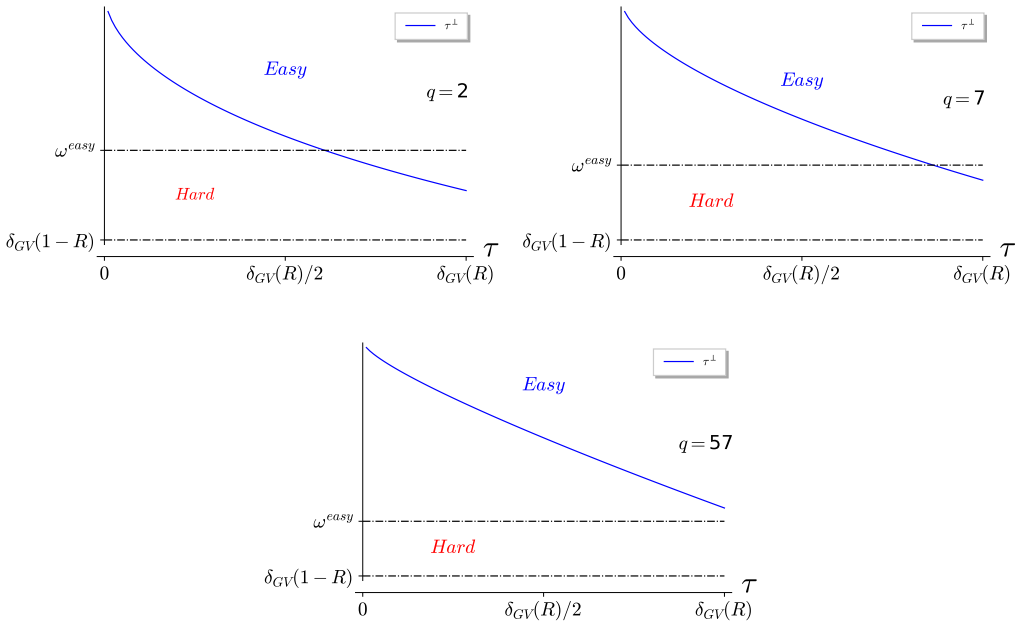


FIGURE 1. $\tau^\perp$ as function of $\tau$ for a rate $R = 1/2$.

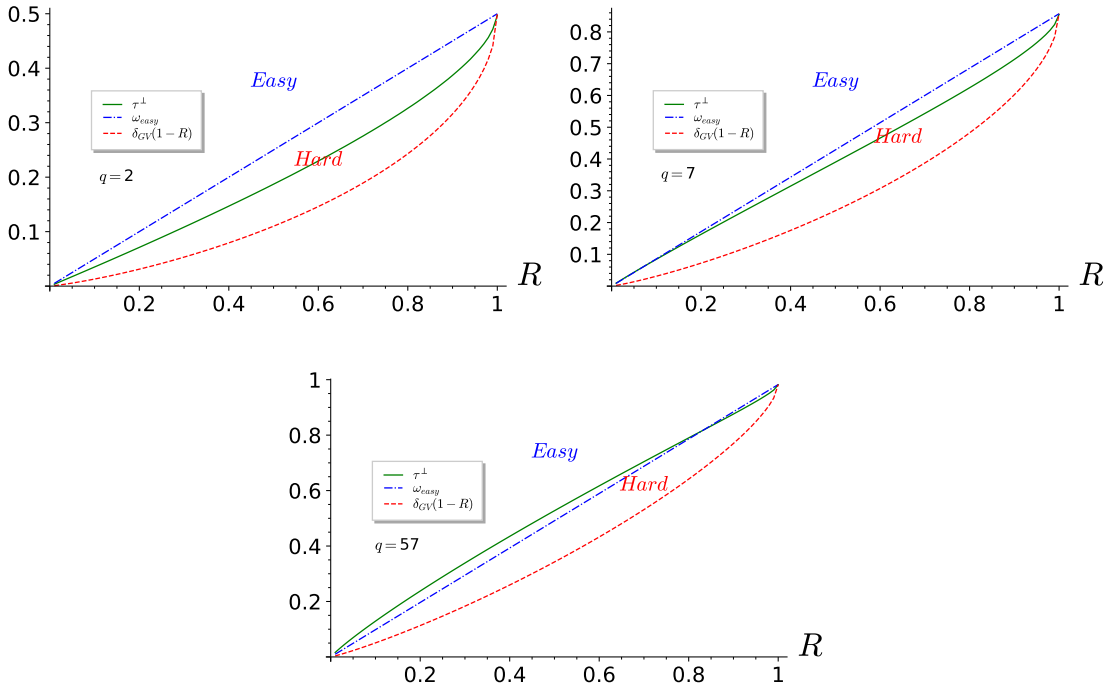


FIGURE 2. Optimal $\tau^\perp$ as function of $R$.

is easy). As $q$ grows, when we plot $\tau^\perp$ for $q = 7$ and $q = 57$, we see that the range of values of $R$ for which $\tau^\perp$ is in the hard zone of the problem shrinks.

## Appendix A. Proof of Theorem 2

A.1. **Step 1.** We are going to measure the distance between quantum states as in [SSTX09] by using the trace distance that is closely connected to the (classical) statistical distance. The trace distance is defined as follows for pure quantum states:

$$D_{\mathrm{tr}}(|\phi\rangle, |\psi\rangle) \stackrel{\mathrm{def}}{=} \sqrt{1 - |\langle\phi|\psi\rangle|^2} \tag{10}$$

Two important properties of this distance are:

(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]

$$D_{\mathrm{stat}}(p_m, q_m) \leq D_{\mathrm{tr}}(|\phi\rangle, |\psi\rangle) \tag{11}$$

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

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

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

In our case, the trace distance between $|\psi_{\mathscr{A}}\rangle$ and $|\psi_{\mathrm{ideal}}\rangle$ is upper-bounded in the following lemma.

**Lemma 1.** *Let $\varepsilon_{\mathbf{G}}$ be the probability that $\mathscr{A}$ returns the right error $\mathbf{e}$ when the input matrix is $\mathbf{G}$, i.e.*

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

*We have*

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

*Proof.* Let $\mathscr{G}$ be the set of $(\mathbf{c}, \mathbf{e}, \mathbf{w})$'s that correspond to inputs to $\mathscr{A}$ that are correctly decoded:

$$\mathscr{G} \stackrel{\mathrm{def}}{=} \left\{ (\mathbf{c}, \mathbf{e}, \mathbf{w}) \in \mathcal{C} \times \mathscr{S}_t \times \mathbb{F}_2^\ell : \mathscr{A}(\mathbf{G}, \mathbf{c} + \mathbf{e}, \mathbf{w}) = \mathbf{e} \right\}.$$

We clearly have ($\pi_{\mathbf{e}} \geq 0$):

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

□

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 (*i.e.* $\mathcal{C} \stackrel{\mathrm{def}}{=} \{\mathbf{uG} : \mathbf{u} \in \mathbb{F}_q^k\}$) we change slightly 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.*

$$\mathcal{C} = \{\mathbf{x} \in \mathbb{F}_q^n : \mathbf{H}\mathbf{x}^\intercal = \mathbf{0}\}.$$

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 $1 - O\left(q^{-(n-k)}\right)$

whereas the second model always produces linear codes of dimension $\geq k$ and codes of dimension $= k$ with probability $1 - O\left(q^{-k}\right)$. This relationship is expressed by the following lemma.

**Lemma 2.** *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 we can easily upper-bound the probability that $Z$ is bigger than $2^\ell q^k(1+\eta)$ for any $\eta > 0$.

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

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

*Proof.* This is obtained through Markov's inequality by computing $\mathbb{E}_{\mathbf{H}}(Z)$. We namely have

$$\begin{aligned}
Z &= \left\| \sum_{\mathbf{c}\in\mathcal{C},\mathbf{e}\in\mathbb{F}_q^n,\mathbf{w}\in\mathbb{F}_2^\ell} \pi_{\mathbf{e}} \left|\mathbf{0}_n\right\rangle \left|\mathbf{c}+\mathbf{e}\right\rangle \left|\mathbf{w}\right\rangle \right\|^2 \\
&= 2^\ell \left\| \sum_{\mathbf{c}\in\mathcal{C},\mathbf{e}\in\mathbb{F}_q^n} \pi_{\mathbf{e}} \left|\mathbf{c}+\mathbf{e}\right\rangle \right\|^2 \\
&= 2^\ell \left( q^k \sum_{\mathbf{e}\in\mathbb{F}_q^n} \pi_{\mathbf{e}}^2 + \sum_{\substack{(\mathbf{c},\mathbf{e})\neq(\mathbf{c}',\mathbf{e}'):\\ \mathbf{c}+\mathbf{e}=\mathbf{c}'+\mathbf{e}'}} \pi_{\mathbf{e}}\pi_{\mathbf{e}'} \right) \\
&= 2^\ell q^k \left( 1 + \sum_{\mathbf{e}\neq\mathbf{e}':\mathbf{H}(\mathbf{e}-\mathbf{e}')^\intercal=\mathbf{0}} \pi_{\mathbf{e}}\pi_{\mathbf{e}'} \right)
\end{aligned}$$

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

$$X \stackrel{\text{def}}{=} \sum_{\mathbf{e}\neq\mathbf{e}':\mathbf{H}(\mathbf{e}-\mathbf{e}')^\intercal=0} \pi_{\mathbf{e}}\pi_{\mathbf{e}'}.$$

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

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

From this we deduce

$$\begin{aligned}
\mathbb{E}_{\mathbf{H}}(X) &= \sum_{\mathbf{e}\neq\mathbf{e}'} \pi_{\mathbf{e}}\pi_{\mathbf{e}'} \mathbb{P}_{\mathbf{H}}((\mathbf{e}-\mathbf{e}') \in \mathcal{C}) \\
&= \sum_{\mathbf{e}\neq\mathbf{e}'} \frac{\pi_{\mathbf{e}}\pi_{\mathbf{e}'}}{q^{n-k}} \\
&\leq \sum_{t,t'} \frac{f(t)f(t')\ S_t\ S_{t'}}{q^{n-k}} \\
&= \frac{\left(\sum_t f(t)\ S_t\right)^2}{q^{n-k}} \\
&= \frac{\langle \pi|\mathbf{1}\rangle^2}{q^{n-k}}.
\end{aligned}$$

Therefore,

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

□

By putting Lemmas 1 and 3 (with $\eta = 1$) together we immediately obtain the following proposition.

**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:*

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

A.2. **Step 2.** Recall that the quantum Fourier transform $\widehat{|\psi\rangle}$ of a state $|\psi\rangle \stackrel{\text{def}}{=} \sum_{\mathbf{x}\in\mathbb{F}_q^n} \alpha_{\mathbf{x}} |\mathbf{x}\rangle$ 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 $q = p^s$)

$$\begin{aligned}\chi_{\mathbf{y}}(\mathbf{x}) &\stackrel{\text{def}}{=} e^{\frac{2i\pi\,\mathrm{Tr}(\mathbf{x}\cdot\mathbf{y})}{p}} \quad \text{where}\\ \mathbf{x}\cdot\mathbf{y} &\stackrel{\text{def}}{=} \sum_{i=1}^{n} x_i y_i \ \text{ with } \mathbf{x} = (x_i)_{i=1}^n \text{ and } \mathbf{y} = (y_i)_{i=1}^n\\ \mathrm{Tr}(a) &\stackrel{\text{def}}{=} a + a^p + a^{p^2} + \cdot + a^{p^{s-1}}\\ \widehat{|\psi\rangle} &\stackrel{\text{def}}{=} \frac{1}{\sqrt{q^n}} \sum_{\mathbf{x},\mathbf{y}\in\mathbb{F}_q^n} \alpha_{\mathbf{x}}\chi_{\mathbf{y}}(\mathbf{x}) |\mathbf{y}\rangle\end{aligned}$$

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 inner product $\mathbf{x}\cdot\mathbf{y}$ or from the characters as follows

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

If we apply the unitary $\mathbf{U}$ corresponding to the quantum Fourier transform on the second register of $|\psi_{\mathrm{ideal}}\rangle$ (given in Equation (9)) it is readily seen that we obtain:

$$\left|\psi_{\mathrm{ideal}}^{\mathrm{QFT}}\right\rangle \stackrel{\text{def}}{=} (\mathbf{Id}\otimes\mathbf{U}\otimes\mathbf{Id})\,|\psi_{\mathrm{ideal}}\rangle = \frac{q^k}{\sqrt{Z}} \sum_{\mathbf{c}^{\perp}\in\mathcal{C}^{\perp}} \sum_{\mathbf{w}\in\mathbb{F}_2^{\ell}} \widehat{\pi}_{\mathbf{c}^{\perp}} |\mathbf{0}_n\rangle \left|\mathbf{c}^{\perp}\right\rangle |\mathbf{w}\rangle,$$

where $\widehat{|\pi\rangle} = \sum_{\mathbf{e}\in\mathbb{F}_q^n} \widehat{\pi}_{\mathbf{e}} |\mathbf{e}\rangle$ is the quantum Fourier transform of $|\pi\rangle$.

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

$$\begin{aligned}
\widehat{|\mathbf{1}_{\mathcal{C}+\mathbf{e}}\rangle} &= \sum_{\mathbf{c}\in\mathcal{C}} \frac{1}{\sqrt{q^n}} \sum_{\mathbf{y}\in\mathbb{F}_q^n} \chi_{\mathbf{y}}(\mathbf{c}+\mathbf{e}) \,|\mathbf{y}\rangle \\
&= \frac{1}{\sqrt{q^n}} \sum_{\mathbf{y}\in\mathbb{F}_q^n} \chi_{\mathbf{y}}(\mathbf{e}) \sum_{\mathbf{c}\in\mathcal{C}} \chi_{\mathbf{y}}(\mathbf{c}) \,|\mathbf{y}\rangle \\
&= \frac{q^k}{\sqrt{q^n}} \sum_{\mathbf{c}^\perp\in\mathcal{C}^\perp} \chi_{\mathbf{c}^\perp}(\mathbf{e}) \left|\mathbf{c}^\perp\right\rangle \quad (\text{since } \sum_{\mathbf{c}\in\mathcal{C}} \chi_{\mathbf{y}}(\mathbf{c}) = 0 \text{ if } \mathbf{y}\notin\mathcal{C}^\perp \text{ and } q^k \text{ else}) \qquad (12)
\end{aligned}$$

Therefore

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

The lemma directly follows from this last equation. □

A.3. **Step 3.** For this we first need to have a good estimation of $\left|\psi_{\text{ideal}}^{\text{QFT}}\right\rangle$'s amplitudes. This will be a consequence of the following lemma.

**Lemma 5.** *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*

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

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

$$N_u^\perp = \sum_{\mathbf{x}\in\mathscr{S}_u} \mathbb{1}_x \qquad (13)$$

It is clear that $\mathbb{E}(\mathbb{1}_{\mathbf{x}}) = \mathbb{P}(\mathbf{x} \in \mathcal{C}^{\perp}) = \frac{1}{q^k}$. Therefore, $\mathbb{E}(N_u^{\perp}) = \frac{S_u}{q^k}$. By using Bienaymé-Tchebychev's inequality, we obtain:

$$\begin{aligned}
\mathbb{P}\left(\left|N_u^{\perp} - \frac{S_u}{q^k}\right| \geq a\right) &\leq \frac{\mathbf{Var}(N_u^{\perp})}{a^2} \\
&= \frac{1}{a^2}\left(\sum_{\mathbf{x}\in\mathscr{S}_u} \mathbf{Var}(\mathbb{1}_{\mathbf{x}}) + \sum_{\substack{\mathbf{x},\mathbf{y}\in\mathscr{S}_u \\ \mathbf{x}\neq\mathbf{y}}} \mathbb{E}(\mathbb{1}_{\mathbf{x}}\mathbb{1}_{\mathbf{y}}) - \mathbb{E}(\mathbb{1}_{\mathbf{x}})\mathbb{E}(\mathbb{1}_{\mathbf{y}})\right) \\
&\leq \frac{1}{a^2}\left(\sum_{\mathbf{x}\in\mathscr{S}_u} \mathbb{E}(\mathbb{1}_{\mathbf{x}}) + \sum_{\substack{\mathbf{x},\mathbf{y}\in\mathscr{S}_u \\ \mathbf{x}\neq\mathbf{y}}} \mathbb{E}(\mathbb{1}_{\mathbf{x}}\mathbb{1}_{\mathbf{y}}) - \mathbb{E}(\mathbb{1}_{\mathbf{x}})\mathbb{E}(\mathbb{1}_{\mathbf{y}})\right) \\
&= \frac{1}{a^2}\left(\frac{S_u}{q^k} + \sum_{\substack{\mathbf{x},\mathbf{y}\in\mathscr{S}_u \\ \mathbf{x}\neq\mathbf{y}}} \mathbb{E}(\mathbb{1}_{\mathbf{x}}\mathbb{1}_{\mathbf{y}}) - \mathbb{E}(\mathbb{1}_{\mathbf{x}})\mathbb{E}(\mathbb{1}_{\mathbf{y}})\right) \qquad (14)
\end{aligned}$$

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:

$$\begin{aligned}
\sum_{\substack{\mathbf{x},\mathbf{y}\in\mathscr{S}_u \\ \mathbf{x}\neq\mathbf{y}}} \mathbb{E}(\mathbb{1}_{\mathbf{x}}\mathbb{1}_{\mathbf{y}}) - \mathbb{E}(\mathbb{1}_{\mathbf{x}})\mathbb{E}(\mathbb{1}_{\mathbf{y}}) &= \sum_{\mathbf{x}\in\mathscr{S}_u} \sum_{\substack{\mathbf{y}\in\mathscr{S}_u\setminus\mathbf{x}: \\ \text{colinear to } \mathbf{x}}} \frac{1}{q^k} - \frac{1}{q^{2k}} \\
&\leq \sum_{\mathbf{x}\in\mathscr{S}_u} \sum_{\substack{\mathbf{y}\in\mathscr{S}_u\setminus\mathbf{x}: \\ \text{colinear to } \mathbf{x}}} \frac{1}{q^k} \\
&\leq \frac{(q-2)S_u}{q^k} \qquad (15)
\end{aligned}$$

It gives by plugging (15) in (14) :

$$\begin{aligned}
\mathbb{P}\left(\left|N_u^{\perp} - \frac{S_u}{q^k}\right| \geq a\right) &\leq \frac{1}{a^2}\left(\frac{S_u}{q^k} + \frac{(q-2)S_u}{q^k}\right) \\
&= \frac{(q-1)S_u}{a^2 q^k}
\end{aligned}$$

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

The point of this lemma is that it will turn out that (where $d_{\mathrm{GV}}^{+}(n, n-k)$ denotes the largest integer $t$ such that $q^{n-k} \cdot S_t \geq q^n$) we will choose a weight $u$ such that it is above $d_{\mathrm{GV}}(n, n-k)$ and below $d_{\mathrm{GV}}^{+}(n, n-k)$ and then for many metrics the term $\frac{q^k}{S_u}$ is exponentially small. Therefore, the probability of measuring $\mathbf{c}^{\perp} \in \mathcal{C}^{\perp}$ of weight $u$ after measuring $\left|\psi_{\text{ideal}}^{\text{QFT}}\right\rangle$ is known up to an exponentially small factor.

Recall now that quantum amplitude amplification techniques (with a classical additional tweak) [BHMT02, §2.1] enable to turn a quantum algorithm, that performs no measurement and which succeeds with probability $p$, into a quantum algorithm that succeeds with probability *exactly* one in roughly $\sqrt{p}$ iterations. The same result algorithms carries over almost verbatim to give an

algorithm working with probability very close to 1 if $p$ is only known with good precision. Let us give here the corresponding statement with the corresponding proof for the reader's convenience.

**Lemma 6.** *Let $\mathscr{B}$ be any quantum algorithm that performs no measurement. We suppose that measuring the output of $\mathscr{B}$ gives a solution with probability $p$ (in this case we say that $\mathscr{B}$ succeeds). Furthermore, we suppose that $p$ is unknown but we know $q = \Omega\left(\frac{1}{\mathsf{poly}(n)}\right)$ such that:*

$$p \in ((1-\delta)q, (1+\delta)q). \tag{16}$$

*Then, there exists a quantum algorithm that runs a $\mathsf{poly}(n)$ number of times $\mathscr{B}$ and $\mathscr{B}^{-1}$ (and uses $\mathsf{poly}(n)$ other gates) such that measuring the output of it gives a solution with probability $\geq 1 - \mathsf{poly}(n)O(\delta^2)$.*

*Proof.* From any quantum algorithm that succeeds with probability $a$, it is easy to build (by adding a qubit and a rotation) a new quantum algorithm that succeeds with probability $\alpha a$ for any chosen $\alpha \in [0,1]$ [BHMT02, §2.1]. By using this tweak, we can construct an algorithm $\mathscr{B}'$ outputting a solution of $\mathscr{B}$ with probability $p' = \alpha p$ for an $\alpha \in [0,1]$ that we will choose later on. Let $\theta \stackrel{\text{def}}{=} \arcsin\sqrt{\alpha p}$. By hypothesis the output of $\mathscr{B}'$ is equal to:

$$\sin\theta\,|G\rangle + \cos\theta\,|B\rangle$$

where $|G\rangle$ denotes the quantum superposition of solutions and $|B\rangle$ a state orthogonal to it. Now by making $T$ steps of amplification, we obtain the following quantum state:

$$\sin((2T+1)\theta)\,|G\rangle + \cos((2T+1)\theta)\,|B\rangle\,.$$

The probability of measuring a solution is given by $\sin^2((2T+1)\theta)$. We choose $\alpha$ as the largest $\alpha \in [0,1]$ such that

$$T \stackrel{\text{def}}{=} \frac{\pi}{4\rho} - \frac{1}{2} \in \mathbb{Z}^+ \quad \text{where } \rho \stackrel{\text{def}}{=} \arcsin\sqrt{\alpha q}. \tag{17}$$

Clearly $T = \mathsf{poly}(n)$.

Let us compute now the success probability $p_{\text{succ}}$ after $T$ steps of amplification. We have the following computation:

$$\begin{aligned}
p_{\text{succ}} &= \sin^2((2T+1)\theta) \\
&= \sin^2\left((2T+1)\rho + (2T+1)(\theta-\rho)\right) \\
&= \cos^2\left((2T+1)(\theta-\rho)\right) \quad \text{(by Equation (17))} \\
&\geq 1 - \left((2T+1)(\theta-\rho)\right)^2 \\
&\geq 1 - \mathsf{poly}(n)(\theta-\rho)^2 \\
&\geq 1 - \mathsf{poly}(n)\delta^2.
\end{aligned}$$

□

We will make the following assumption in the whole section from now on.

**Assumption 1.** *The error distribution $\pi = (\pi_{\mathbf{e}})_{\mathbf{e}\in\mathbb{F}_q^n}$ and $1 \leq u \leq n$ verify:*

$$\frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}} = 2^{-\Omega(n)}, \quad \frac{q^k}{S_u} = 2^{-\Omega(n)} \quad \text{and} \quad S_u|f^{\perp}(u)|^2 = \Omega\left(\frac{1}{\mathsf{poly}(n)}\right).$$

We have now the following proposition.

**Proposition 3.** *If the Fourier transform is radially preserving, then under Assumption 1 it exists a quantum algorithm such that when starting from $|\psi_{\text{ideal}}\rangle$, then for a proportion $\geq 1-\beta(\pi)$ of matrices $\mathbf{G}$, the probability of obtaining a codeword $\mathbf{c}^{\perp}$ of weight $u$ in $\mathcal{C}^{\perp}$ when measuring is greater than or equal to $1-\alpha(\pi)$ where:*

$$\begin{aligned}\alpha(\pi) &\stackrel{def}{=} \widetilde{O}\left(\left(\frac{q^k}{S_u}\right)^{1/4}+\sqrt{\frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}}}\right)^2\\ \beta(\pi) &\stackrel{def}{=} (q-1)\sqrt{\frac{q^k}{S_u}}+\sqrt{\frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}}}+O\left(q^{-\min(k,n-k)}\right)\end{aligned}$$

*Proof.* Let $\mathscr{B}$ be the quantum algorithm starting from $|\psi\rangle$ which computes $(\mathbf{Id}\otimes\mathbf{U}\otimes\mathbf{Id})|\psi\rangle$. This algorithm succeeds when measuring a dual codeword $\mathbf{c}^{\perp}\in\mathcal{C}^{\perp}$ of weight $u$. When starting with $|\psi_{\text{ideal}}\rangle$, the probability of success of $\mathscr{B}$ is equal to $\frac{2^\ell q^{2k}N_u^{\perp}}{Z}|f^{\perp}(u)|^2$ by Lemma 4. Let,

$$\mathscr{E}\stackrel{\text{def}}{=}\left\{\mathbf{G}\in\mathbb{F}_q^{k\times n}\,:\,Z>2^\ell q^k\left(1+\sqrt{\frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}}}\right)\text{ or }\left|N_u^{\perp}-\frac{S_u}{q^k}\right|\geq\left(\frac{S_u}{q^k}\right)^{3/4}\right\}.$$

By Lemmas 3 and 5 we have that $\mathbb{P}(\mathbf{G}\in\mathscr{E})\leq\beta(\pi)=(q-1)\sqrt{\frac{q^k}{S_u}}+\sqrt{\frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}}}+O\left(q^{-\min(k,n-k)}\right)$. Therefore, for a proportion $\geq 1-\beta(\pi)$ of codes (over matrices $\mathbf{G}$):

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

(*ii*) $\left|\frac{q^kN_u^{\perp}}{S_u}-1\right|\leq\left(\frac{q^k}{S_u}\right)^{1/4}$.

Therefore, we have for a proportion $\geq 1-\beta(\pi)$ of codes:

$$S_u|f^{\perp}(u)|^2\left(\frac{1-\left(\frac{q^k}{S_u}\right)^{1/4}}{1+\sqrt{\frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}}}}\right)\leq\frac{2^\ell q^{2k}N_u^{\perp}}{Z}|f^{\perp}(u)|^2\leq S_u|f^{\perp}(u)|^2\left(1+\left(\frac{q^k}{S_u}\right)^{1/4}\right). \tag{18}$$

Now,

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

Therefore, by plugging this in Equation (18) we have for a proportion $\geq 1-\beta(\pi)$ of codes:

$$\frac{2^\ell q^{2k}N_u^{\perp}}{Z}|f^{\perp}(u)|^2\in S_u|f^{\perp}(u)|^2\,(1-\delta,1+\delta)\quad\text{where}\quad\delta\stackrel{\text{def}}{=}\left(\frac{q^k}{S_u}\right)^{1/4}+\sqrt{\frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}}} \tag{19}$$

By Assumption (1), $S_u|f^{\perp}(u)|^2=\Omega\left(\frac{1}{\mathsf{poly}(n)}\right)$, $\delta=2^{-\Omega(n)}$ while $\frac{2^\ell q^{2k}N_u^{\perp}}{Z}|f^{\perp}(u)|^2$ denotes the success probability of $\mathscr{B}$. Therefore to conclude the proof we apply Lemma 6.

□

A.4. **Proof of Theorem 2. Notation.** From now on, $\left|\psi_{\text{ideal}}^{\text{Ampl}}\right\rangle$ (*resp.* $\left|\psi_{\mathscr{A}}^{\text{Ampl}}\right\rangle$) will denote the quantum state after applying the quantum amplification algorithm of Proposition 3 on $\left|\psi_{\text{ideal}}^{\text{QFT}}\right\rangle$ (*resp.* $\left|\psi_{\mathscr{A}}^{\text{QFT}}\right\rangle$).

With this notation we are ready now to prove Theorem 2 which we now recall

**Theorem 2.** *Assume that* $|\pi\rangle = \sum_{\mathbf{e}\in\mathbb{F}_q^n} |\mathbf{e}\rangle$ *is radial and nonnegative, i.e* $\pi_{\mathbf{e}} = f(|\mathbf{e}|)$ *for some function* $f$ *and* $\pi_{\mathbf{e}} \geq 0$ *for all* $\mathbf{e} \in \mathbb{F}_q^n$*. Let* $p_t \stackrel{def}{=} \sum_{\mathbf{e}:|\mathbf{e}|=t} = S_t f(t)^2$*.* $\widehat{|\pi\rangle} = \sum_{\mathbf{e}\in\mathbb{F}_q^n} \widehat{\pi}_{\mathbf{e}} |\mathbf{e}\rangle$ *is radial too and we let* $f^\perp(u) = \widehat{\pi}_{\mathbf{e}}$ *where* $\mathbf{e}$ *is any element of* $\mathbb{F}_q^n$ *of Hamming weight* $u$*. Furthermore, assume that :*

$$\frac{\langle\pi|\mathbf{1}\rangle^2}{q^{n-k}} = 2^{-\Omega(n)}, \quad \frac{q^k}{S_u} = 2^{-\Omega(n)} \quad \text{and} \quad S_u|f^\perp(u)|^2 = \Omega\left(\frac{1}{\mathsf{poly}(n)}\right) \text{ for some } 1 \leq u \leq n.$$

*with* $|\mathbf{1}\rangle$ *being the (unnormalized) superposition of all errors :* $|\mathbf{1}\rangle \stackrel{def}{=} \sum_{\mathbf{e}\in\mathbb{F}_q^n} |\mathbf{e}\rangle$*.*
*Suppose that there exists an algorithm* $\mathscr{A}$ *solving the decoding problem with success probability* $\varepsilon$*. Then, there exists a quantum algorithm making only a polynomial number of calls to* $\mathscr{A}$ *and to additional elementary* 1 *or* 2 *qubit gates 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* $\mathcal{C}^\perp$ *with probability bigger than* $\frac{p_t^2\varepsilon^3}{16} - O(p_t^4\varepsilon^5) - 2^{-\Omega(n)} - O\left(q^{-\min(k,n-k)}\right)$*.*

We start the proof by proving the following point

**Lemma 7.** *Call* $\mathscr{G}$ *the set of "good matrices"* $\mathbf{G} \in \mathbb{F}_q^{k\times n}$ *that satisfy at the same time:*
*(i)* $\varepsilon_{\mathbf{G}} \geq \varepsilon/2$*,*
*(ii)* $Z \leq 2^{\ell+1}q^k$
*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 $\mathscr{B}$ be the set of matrices $\mathbf{G}$ that are not good, namely for which $(a)$ $\varepsilon_{\mathbf{G}} < \varepsilon/2$ or $(b)$ $Z > 2^{\ell+1}q^k$. By Lemma 3, the density of matrices verifying $(b)$ is smaller than $\delta(\pi)$. Therefore,

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

which concludes the proof. □

We use this lemma to prove that the statistical distance between the distributions of weights $\left|\mathbf{c}^\perp\right|$ we obtain by measuring $\left|\psi_{\mathscr{A}}^{\text{Ampl}}\right\rangle$ can not be to far away from the distribution of weights when we measure the state $\left|\psi_{\text{ideal}}^{\text{Ampl}}\right\rangle$:

**Lemma 8.** *Let* $P$*, respectively* $Q$*, be the distribution of the weights* $|\mathbf{c}^\perp|$ *of the state* $|\mathbf{e}\rangle\left|\mathbf{c}^\perp\right\rangle|\mathbf{w}\rangle$ *obtained by measuring the state* $\left|\psi_{\mathscr{A}}^{\text{Ampl}}\right\rangle$*, respectively* $\left|\psi_{ideal}^{\text{Ampl}}\right\rangle$*. 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

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

We start the proof by noticing that

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

□

We are now ready to prove Theorem 2.

*Proof of Theorem 2.* By Proposition 3 we know that

$$Q(u) \geq 1-\alpha(\pi))(1-\beta(\pi)) \geq 1-\alpha(\pi)-\beta(\pi).$$

But now we have the following computation,

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

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

## Appendix B. Proof of Theorem 3

Recall that we have chosen $|\pi\rangle$ as

$$|\pi\rangle = \frac{1}{\sqrt{\binom{n}{t}(q-1)^t}} \sum_{\mathbf{e}\in\mathbb{F}_q^n:|\mathbf{e}|=t} |\mathbf{e}\rangle\,. \tag{20}$$

This time, as long as $t$ is below $(1-\delta)d_{\text{GV}}(n,k)$ for an arbitrary $\delta>0$ the term $\frac{\langle\pi,\mathbf{1}\rangle^2}{q^{n-k}}$ is exponentially small in $n$ as shown by

**Lemma 9.** *If $t \leq (1-\delta)d_{\mathrm{GV}}(n,k)$ then*

$$\frac{\langle \pi, \mathbf{1}\rangle^2}{q^{n-k}} = \frac{S_t}{q^{n-k}} = q^{\alpha(R,\delta)n(1+o(1))} \quad \textit{where} \quad \alpha(R,\delta) < 0.$$

*Proof.* Recall that the size $B_t$ of the Hamming ball of radius $t$ is of the form

$$B_t = q^{n\ h_q(\tau)(1+o(1))}$$

where $h_q(x) = -(1-x)\log_q(1-x) - x\log_q\left(\frac{x}{q-1}\right)$ and $\tau = t/n$. We obtain from this

$$\begin{aligned}
\frac{\langle \pi, \mathbf{1}\rangle^2}{q^{n-k}} &= \frac{S_t}{q^{n-k}} \\
&\leq \frac{B_t}{B_{d_{\mathrm{GV}}}} \qquad (\text{since } S_t \leq B_t \text{ and } B_{d_{\mathrm{GV}}} \leq q^{n-k}) \\
&\leq q^{n(h_q(\tau) - h_q(\delta_{\mathrm{GV}}) + o(1))} \\
&\leq q^{n(h_q((1-\delta)\delta_{\mathrm{GV}}) - h_q(\delta_{\mathrm{GV}}) + o(1))}
\end{aligned}$$

with $\delta_{\mathrm{GV}} \stackrel{\mathrm{def}}{=} d_{\mathrm{GV}}(n,k)/n$. We finish the proof by noticing that $h_q((1-\delta)\delta_{\mathrm{GV}}) - h_q(\delta_{\mathrm{GV}}) < 0$. □

The Fourier transform $\widehat{|\pi\rangle}$ of $|\pi\rangle$ can be expressed in terms of Krawtchouk polynomials as follows.

**Lemma 10.**

$$\widehat{|\pi\rangle} = \frac{1}{\sqrt{q^n}} \sum_{\mathbf{y}\in\mathbb{F}_q^n} \frac{K_t(|\mathbf{y}|)}{\sqrt{\binom{n}{t}(q-1)^t}} |\mathbf{y}\rangle$$

*where*

$$K_t(X; q, n) \stackrel{def}{=} \sum_{j=0}^{t} (-1)^j \binom{X}{j}\binom{n-X}{t-j}(q-1)^{t-j}$$

*is the Krawtchouk polynomial of order $n$, parameter $q$ and degree $t \in [\![0, n]\!]$* [(2)].

*Proof.* By definition of the Fourier transform we have

$$\widehat{|\pi\rangle} = \frac{1}{\sqrt{q^n}} \sum_{\mathbf{y}\in\mathbb{F}_q^n} \frac{1}{\sqrt{\binom{n}{t}(q-1)^t}} \sum_{\substack{\mathbf{e}\in\mathbb{F}_q^n \\ |\mathbf{e}|=t}} \chi_{\mathbf{y}}(\mathbf{e}) |\mathbf{y}\rangle . \tag{21}$$

The sum of characters $\sum_{\substack{\mathbf{e}\in\mathbb{F}_q^n \\ |\mathbf{e}|=t}} \chi_{\mathbf{y}}(\mathbf{e})$ can be expressed as a Krawtchouk polynomial evaluation (see for instance Lemma 5.3.1 in [vL99, §5.3]):

$$\sum_{\substack{\mathbf{e}\in\mathbb{F}_q^n \\ |\mathbf{e}|=t}} \chi_{\mathbf{y}}(\mathbf{e}) = K_t(|\mathbf{y}|).$$

To finish the proof we just have to substitute the character sum for this expression in (21). □

Therefore, for the error distribution $\pi$ we have:

$$S_u |f^{\perp}(u)|^2 = \frac{1}{q^n}\binom{n}{u}(q-1)^u \frac{K_t(u)^2}{(q-1)^t\binom{n}{t}}. \tag{22}$$

To apply Theorem 2 to the Hamming case it remains now to understand how the Krawtchouk polynomial evaluations behave. In particular, if we find some $u$ such that:

$$S_u |f^{\perp}(u)|^2 = \Omega\left(\frac{1}{\mathsf{poly}(n)}\right) \quad \text{and} \quad \frac{q^k}{S_u} = 2^{-\Omega(n)},$$

[(2)] To simplify notation, since $q$ and $n$ are clear here from the context: they are respectively the field size and the length of the codes we consider, we will drop the dependency in $q$ and $n$ and simply write $K_t(X)$.

we would obtain a reduction from finding a dual codeword of weight $u$ to decoding at distance $t$. For our reduction to be meaningful we need to find the smallest $u$ as possible. It is roughly given by the first root of $K_t$ as we prove in what follows.

Let us start by some basic facts on roots of Krawtchouk polynomials.

(1) For $1 \leq t \leq n$, the polynomial $K_t$ has $t$ distinct real roots on $(0, n)$.
(2) Let $x_1$ be the first root of $K_t$. If $1 \leq t \leq n(q-1)/q$, we have [Lev95, Equation (128)]:

$$\frac{x_1}{n} = \tau^\perp + o(1). \tag{23}$$

(3) The distance between any two consecutive roots of $K_t$ is an $o(n)$

**Proposition 4.** *Between any two consecutive roots of $K_t$, where $1 \leq t \leq \frac{n}{q}$, there exists $u$ such that:*

$$\frac{1}{q^n}\binom{n}{u}(q-1)^u \frac{K_t(u)^2}{(q-1)^t\binom{n}{t}} = \Omega\left(\frac{1}{n^5}\right) \tag{24}$$

The proof of this proposition is given in Appendix D. It is a generalization for any $q$ of [KS21, Corollary 26] which corresponds to the case where $q = 2$.

We are now ready to prove Theorem 3. We are going to apply Theorem 2 with the error distribution $\pi$ that we just introduced. Our proof essentially consists in verifying each hypothesis of Theorem 2.

*Proof of Theorem 3.* First, $\widehat{|\pi\rangle}$ is clearly radial (its expression is given in Lemma 10).

By Lemma 9, $\frac{\langle \pi, \mathbf{1}\rangle^2}{q^{n-k}} = \frac{S_t}{q^{n-k}} = q^{-\Omega(n)}$ as $t \leq (1-\delta)d_{\mathrm{GV}}(n,k)$ by assumption of Theorem 3.

Now, by Proposition 4, there exists $u$ between the two first consecutive roots of $K_t$ such that $S_u|f^\perp(u)|^2 = \frac{1}{q^n}\binom{n}{u}(q-1)^u \frac{K_t(u)^2}{(q-1)^t\binom{n}{t}} = \Omega\left(\frac{1}{n^5}\right)$.

We have $|u - x_1| = o(n)$ as the distance between any two consecutive roots of $K_t$ is an $o(n)$. Therefore, by Equation (23): $\frac{u}{n} = \tau^\perp + o(1)$. Furthermore we deduce that $\log_2 q^k/S_u = n\big(R - h_q(\tau^\perp) + o(1)\big)$. Now it can be verified that $h_q(\tau^\perp) > R$ for $R \in (0,1)$ and therefore that $q^k/S_u$ is negligible which concludes that proof. □

## Appendix C. A first failed attempt

To apply Theorem 2, a natural choice for performing the reduction of searching short codewords in $\mathcal{C}^\perp$ to decoding $t$ errors in a linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ would be to choose a quantum state $|\pi\rangle$ that at the same time

(i) concentrates most of its norm on vectors of weight $\approx t$,
(ii) is radially symmetric,
(iii) would ideally be a separable quantum state $|\pi\rangle = |\psi\rangle^{\otimes n}$ which would simplify the computation of the Fourier transform a great deal.

All these requirements would lead to define the quantum state $|\pi\rangle$ as the separable state

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

Since $|\pi^{\mathrm{try}}\rangle$ has also the following form

$$\left|\pi^{\mathrm{try}}\right\rangle = \sum_{\mathbf{e}\in\mathbb{F}_q^n}(1-\tau)^{\frac{n-|\mathbf{e}|}{2}}(\tau/(q-1))^{\frac{|\mathbf{e}|}{2}}\,|\mathbf{e}\rangle\,,$$

measuring this state $|\pi^{\mathrm{try}}\rangle$ really mimics the error we have in a $q$-ary symmetric channel of crossover $\tau$, i.e.

$$\mu_\tau(\mathbf{e}) \stackrel{\mathrm{def}}{=} \mathbb{P}(\text{measure outputs } \mathbf{e}) = (1-\tau)^{n-|\mathbf{e}|}(\tau/(q-1))^{|\mathbf{e}|} \tag{25}$$

where $\mathbf{e}$ is any vector in $\mathbb{F}_q^n$. It is straightforward to compute the Fourier transform of this state to obtain

$$\widehat{|\pi^{\mathrm{try}}\rangle} = \left(\sqrt{1-\tau^\perp}\,|0\rangle + \sqrt{\tau^\perp/(q-1)}\sum_{\alpha\in\mathbb{F}_q^*}|\alpha\rangle\right)^{\otimes n}$$

In other words, the Fourier transform maps the quantum state "representing" the $q$-ary symmetric channel of crossover probability $\tau$ to a quantum state "representing" the $q$-ary symmetric channel of crossover probability $\tau^\perp$. This suggests that the quantum reduction outlined earlier reduces searching for codewords of weight $\approx t^\perp \stackrel{\mathrm{def}}{=} \tau^\perp n$ to decoding $t$ errors in $\mathcal{C}$. Unfortunately, the fundamental quantity appearing in Theorem 2 which is $\frac{\langle\pi^{\mathrm{try}},\mathbf{1}\rangle^2}{q^{n-k}}$ is not negligible at all. Indeed we observe that

$$\begin{aligned}\frac{\langle\pi^{\mathrm{try}},\mathbf{1}\rangle^2}{q^{n-k}} &= \frac{q^k}{q^n}\left|\sum_{\mathbf{y}\in\mathbb{F}_q^n}\chi_{\mathbf{y}}(\mathbf{0})\pi^{\mathrm{try}}_{\mathbf{y}}\right|^2\\ &= q^k|f^\perp(0)|^2\\ &= q^k(1-\tau^\perp)^n\end{aligned}$$

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

The reason of this behavior can be traced back to the fact that the quantity $\frac{\langle\pi^{\mathrm{try}},\mathbf{1}\rangle^2}{q^{n-k}} = q^k(1-\tau^\perp)^n$ is just too big. Notice that

$$\frac{\langle\pi^{\mathrm{try}},\mathbf{1}\rangle^2}{q^n} = (1-H^2(\mu_\tau,U))^2$$

where $\mu_\tau$ is the probability distribution on $\mathbb{F}_q^n$ obtained from measuring $|\pi\rangle$ as defined in (25) (*i.e.* corresponding to the error distribution of a $q$-ary symmetric channel of crossover probability $\tau$), $U$ is the uniform distribution on $\mathbb{F}_q^n$ and $H(\mathbf{p},\mathbf{q})$ stands for the Hellinger distance between two discrete probabilities $p$ and $q$ defined over a same probability space:

$$H(p,q) \stackrel{\mathrm{def}}{=} \sqrt{1-\sum_i\sqrt{p_i,q_i}}.$$

In other words the distribution of $\mu_\tau$ is too much spread out and we need a distribution which is much more concentrated around the weight $t$ for which we assume to have a decoder for $\mathcal{C}$.

## Appendix D. Proof of Proposition 4

Our aim in this section is to prove the following proposition

**Proposition 4.** *Between any two consecutive roots of $K_t$, where $1 \le t \le \frac{n}{q}$, there exists $u$ such that:*

$$\frac{1}{q^n}\binom{n}{u}(q-1)^u\,\frac{K_t(u)^2}{(q-1)^t\binom{n}{t}} = \Omega\left(\frac{1}{n^5}\right) \tag{24}$$

This proposition relies on [KS21, Proposition 25] which is a general result about orthogonal polynomial for a positive discrete $\mu$ over $\{0,\dots,n\}$. Let $(P_0,\dots,P_n)$ be the family of orthogonal polynomials with respect to the inner product $\langle f,g\rangle \stackrel{\text{def}}{=} \sum_{i=0}^{n} f(i)g(i)\mu(i)$.

**Proposition 5** ([KS21, Proposition 25])**.** *Let $s>0$. Let the roots of $P_s$ be $y_1<\cdots<y_s$. Assume that $y_1\geq 1$ and that $y_s\leq n-1$, and that the distance between any two consecutive roots is at least 2. Assume also that the ratios $\frac{\mu(j)}{\mu(j+1)}$ and their inverses are uniformly bounded by some $K>0$. Then, for any $1\leq k\leq s-1$ the $\ell_2$ norm of $P_s$[3] is attained between $y_k$ and $y_{k+1}$, up to a factor of at most $O(\sqrt{K}n^2)$.*

The point is that Krawtchouk polynomials $K_t$'s are orthogonal polynomials (over $\{0,\dots,n\}$) for the measure $\mu(j)=\frac{(q-1)^j\binom{n}{j}}{q^n}$. Their $\ell_2$-norm is given by:

$$\sqrt{\sum_{i=0}^{n} K_t^2(i)\mu(i)} = \sqrt{(q-1)^t\binom{n}{t}}. \tag{26}$$

Furthermore their smallest (*resp.* largest) root is $\geq 1$ (*resp.* $\leq n-1$), at least for $n$ large enough (see [Lev95, Corollary 6.1]). Therefore to prove Proposition 4 (by just applying Proposition 5) it just remains to prove that the roots of $K_t$ are at distance $\geq 2$.

**Lemma 11.** *When $1\leq t\leq \frac{n}{q}$, the roots of $K_t$ are at distance $\geq 2$.*

This lemma will be a consequence of the following theorem.

**Theorem 4** ([Kra03, Theorem 2])**.** *Let $P(x)$ be a discrete orthogonal polynomial, corresponding to an orthogonality measure supported on a subset of integers. Suppose that $P$ satisfies*

$$P(x+1)=b(x)P(x)-c(x)P(x-1) \tag{27}$$

*and has all its roots $x_i$ in the open interval $I$. Then for any $i$, $|x_i-x_{i-1}|\geq 1$ provided $c(x)>0$ for $x\in I$. If in addition, $b(x)>0$ on $I$, then for any $i$, $|x_i-x_{i-1}|\geq 2$.*

*Proof of Lemma 11.* We are going to apply Theorem 4. First, Krawtchouk polynomials verify the following equation:

$$(q-1)(n-x)K_t(x+1)=((q-1)(n-x)+x-qt)K_t(x)-xK_t(x-1) \tag{28}$$

All roots of $K_t$ lie in the interval $(0,n)$. Let,

$$\forall x\in(0,n),\quad \begin{cases} b(x) & \stackrel{\text{def}}{=} \frac{(q-1)(n-x)+x-qt}{(q-1)(n-x)} \\ c(x) & \stackrel{\text{def}}{=} \frac{x}{(q-1)(n-x)} \end{cases}$$

Clearly $c(x)>0$ and $b(x)>0$ (as $q\geq 2$ and $t\leq n/q$) on $(0,n)$ which concludes the proof of this lemma by applying Theorem 4. □

We are now ready to prove Proposition 4.

*Proof of Proposition 4.* The roots of $K_t$ are at distance $\geq 2$ and its smallest (*resp.* largest) root is $\geq 1$ (*resp.* $\leq n-1$). Furthermore, $\frac{\mu(j)}{\mu(j+1)}=\frac{1}{q-1}\frac{j+1}{n-j}$. Therefore, the ratios $\frac{\mu(j)}{\mu(j+1)}$ and their

[3] The $\ell_2$ norm of $P_s$ is defined as $|P_s|_2 \stackrel{\text{def}}{=} \sqrt{\sum_{i=0}^{n} P_s(i)^2\mu(i)}$.

inverses are uniformly bounded by $(q-1)n$. By applying Proposition 5, between each consecutive roots of $K_t$, there exists $u$ such that

$$K_t(u)^2\mu(u) = |K_t|_2^2\ \Omega\left(\frac{1}{(q-1)n\ n^4}\right) \quad \Longrightarrow \quad K_t(u)^2\frac{\binom{n}{u}(q-1)^u}{q^n} = \binom{n}{t}(q-1)^t\ \Omega\left(\frac{1}{n^5}\right)$$

which concludes the proof.

□

## 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, Bardet Magali, and Ayoub Otmani. ROLLO (merger of Rank-Ouroboros, LAKE and LOCKER). Second round submission to the NIST post-quantum cryptography call, March 2019.

[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.

[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.

[BHMT02] Gilles Brassard, Peter Høyer, Michele Mosca, and Alain Tapp. Quantum amplitude amplification and estimation. In *Quantum computation and information (Washington, DC, 2000)*, volume 305 of *Contemp. Math.*, pages 53–74. Amer. Math. Soc., Providence, RI, 2002.

[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.

[Che] Yilei Chen. personal communication.

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

[DRT21] Thomas Debris-Alazard, Maxime Rémaud, and Jean-Pierre Tillich. Reducing the short codeword to decoding in the rank metric. preprint, 2021.

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

[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.

[IS98] Mourad E.H. Ismail and Plamen Simeonov. Strong asymptotics for Krawtchouk polynomials. *Journal of Computational and Applied Mathematics*, pages 121–144, 1998.

[Kra03] Ilia Krasikov. Discrete analogues of the laguerre inequality. *Analysis and Applications*, 01(02):189–197, 2003.

[KS21] Naomi Kirshner and Alex Samorodnitsky. A moment ratio bound for polynomials and some extremal properties of krawchouk polynomials and hamming spheres. *IEEE Trans. Inform. Theory*, 67(6):3509–3541, 2021.

[Lev95] Vladimir I. Levenshtein. Krawtchouk polynomials and universal bounds for codes and designs in hamming spaces. *IEEE Trans. Inf. Theory*, 41(5):1303–1321, 1995.

[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.

[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.

[vL99] Jacobus Hendricus van Lint. *Introduction to coding theory*. Graduate texts in mathematics. Springer, 3rd edition edition, 1999.

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

Inria, Saclay

*Email address*: thomas.debris@inria.fr

Atos Quantum Lab, Les Clayes-sous-Bois and Inria de Paris, Paris 75102

*Email address*: maxime.remaud@atos.net

Inria de Paris, Paris 75012

*Email address*: jean-pierre.tillich@inria.fr