The Quantum Decoding Problem : Tight Achievability Bounds and Application to Regev's Reduction

Agathe Blanvillain[1] and André Chailloux[1] and Jean-Pierre Tillich[1]
Inria de Paris, {agathe.blanvillain,andre.chailloux,jean-pierre.tillich}@inria.fr

Abstract

We consider the quantum decoding problem. It consists in recovering a codeword given a superposition of noisy versions of this codeword. By measuring the superposition, we get back to the classical decoding problem. It appears for the first time in Chen, Liu and Zhandry's work showing a quantum advantage for the Short Integer Solution (SIS) problem for the l_∞ norm. In a recent paper, Chailloux and Tillich proved that when we have a noise following a Bernoulli distribution, the quantum decoding problem can be solved in polynomial time and is therefore easier than classical decoding for which the best known algorithms have an exponential complexity. They also give an information theoretic limit for the code rate at which this problem can be solved which turns out to be above the Shannon limit.

In this paper, we generalize the last result to all memoryless noise models. We also show similar results in the rank metric case which corresponds to a noise model which is not memoryless. We analyze the Pretty Good Measurement, from which we derive an information theoretic limit for this problem. By using the algorithm for the quantum decoding problem together with Regev's reduction, we derive a quantum algorithm sampling codewords from the dual code according to a probability distribution which is the dual of the original noise. It turns out that at the information theoretic limit, we get the most likely nonzero codeword of the dual code. When the distribution is a decreasing function of the weight, we find minimal nonzero codewords. Note that Regev's reduction used together with classical decoding is much less satisfying since it is not able to output those minimum weight codewords.

Contents

1 Introduction

1.1 State of the art

Post quantum cryptography. Post-quantum cryptography's aim is to deploy cryptographic schemes that are secure against quantum computers. Some of the main proposals for post-quantum cryptography are based on variants of decoding problems, be they lattice based or code based. This is clear in code-based cryptography where the overwhelming majority of proposals is based on decoding either in the Hamming or in the rank metric. In lattice based cryptography, the Learning With Errors (LWE) problem [Reg05] is ubiquitous and is nothing but decoding a linear code over a large alphabet subject to a discrete Gaussian noise. Most of quantum algorithms [Ber10, KT17, Laa16, CL21] for attacking these post-quantum cryptographic schemes are quantum extensions of the best classical algorithms, using simple quantum algorithmic tools such as Grover's algorithm [Gro96], or more advanced tools such as quantum walks [AAKV01, MNRS11]. While these algorithms are important for assessing the security of post-quantum cryptographic schemes, they intrinsically fail to provide exponential speedups or indications of strong quantum advantage.

Regev's reduction. Recently, a new family of quantum algorithms was developed inspired by Regev's reduction [Reg05]. Originally, Regev's reduction is a quantum complexity reduction from the Learning with Errors problem to the Short Integer Solution (SIS) problem. More precisely, a quantum algorithm for the SIS problem is devised from a (classical or quantum) algorithm for the LWE problem. This can be rephrased in coding theoretic terms. With the help of a decoding algorithm operating in the regime where there is generally at most a unique solution, an algorithm outputting a rather short dual codeword is obtained, where short is measured in terms of the Euclidean distance. Here the code $\mathscr{C}$ which is decoded is a generic linear code and so is the dual code $\mathscr{C}^\perp$, where we recall that

$$\mathscr{C} = \{\boldsymbol{s}\mathbf{G} : \boldsymbol{s} \in \mathbb{F}_q^k\} \subseteq \mathbb{F}_q^n \quad ; \quad \mathscr{C}^\perp = \{\boldsymbol{y} \in \mathbb{F}_q^n : \forall \boldsymbol{c} \in \mathscr{C},\ \boldsymbol{y} \cdot \boldsymbol{c} = 0\}.$$

More precisely Regev's algorithm reduces the short codeword problem

Problem 1 (Short Codeword Problem SCP(q, n, k, w))**.** *Take q, n, k, w positive integers.* $\mathbf{H}$ *is chosen uniformly at random in* $\mathbb{F}_q^{(n-k)\times n}$.
Given: $\mathbf{H}$,
Goal: *find* $\boldsymbol{c} \in \mathbb{F}_q^n \backslash \{0\}$ *which satisfies* $\mathbf{H}\boldsymbol{c}^\intercal = 0$ *and* $|\boldsymbol{c}| \leq w$.

to the decoding problem described by

Problem 2 (Decoding Problem DP(q, n, k, p))**.** *Take q, n, k positive integers and p a probability distribution over $\mathbb{F}_q^n$. Take a generator matrix* $\mathbf{G}$ *in* $\mathbb{F}_q^{k\times n}$*, a codeword* $\boldsymbol{c} = \boldsymbol{u}\mathbf{G}$ *of the code* $\mathscr{C}$ *generated by* $\mathbf{G}$ *where* $\boldsymbol{u}$ *is chosen uniformly at random in* $\mathbb{F}_q^k$.
Given: $(\mathbf{G}, \boldsymbol{c} + \boldsymbol{e})$, *where* $\boldsymbol{e}$ *is sampled according to* p,
Goal: *recover* $\boldsymbol{c}$.

Here $|\boldsymbol{c}|$ is the weight of the codeword $\boldsymbol{c} = c_1, \ldots, c_n$. In Regev's reduction it is the Euclidean weight $|\boldsymbol{c}| = \sqrt{\sum_{i=1}^n c_i^2}$ where the sum is performed over the integers and c_i is taken in the integer interval $\left\{-\frac{q-1}{2}, \cdots, \frac{q-1}{2}\right\}$ (in the case where q is odd and $\left\{-\frac{q}{2}, \cdots, \frac{q-2}{2}\right\}$ otherwise). Note that in both cases we have a random linear code that we either want to decode or for which we look for a short codeword. The originality of this reduction is that it uses a decoding algorithm which works with a decoding radius where the solution is unique (this can be viewed as the injective regime) to get a short codeword for a weight w where there are exponentially many codewords of this weight (this can be viewed as the surjective regime). It took a while to get an analogue of this reduction for the Hamming metric [DRT24] where to get a non trivial reduction, it is mandatory in certain parameter regimes to consider a decoding algorithm which works for a larger decoding radius where the solution of the decoding algorithm is typically unique but not always unique.

Regev's reduction makes crucially use of the quantum Fourier transform (QFT). It maps the quantum state $\sum_{\boldsymbol{x}\in\mathbb{F}_q^n} f(\boldsymbol{x})|\boldsymbol{x}\rangle$ to $\sum_{\boldsymbol{x}\in\mathbb{F}_q^n} \widehat{f}(\boldsymbol{x})|\boldsymbol{x}\rangle$ where $\widehat{f}$ is the classical Fourier transform of f. The latter is defined by using the characters $\chi_{\boldsymbol{y}}$ of $\mathbb{F}_q^n$ (see §2.1), the Fourier transform over $\mathbb{F}_q^n$ being defined by $\hat{f}(\boldsymbol{y}) \triangleq \frac{1}{\sqrt{q^n}} \sum_{\boldsymbol{x}\in\mathbb{F}_q^n} \chi_{\boldsymbol{y}}(\boldsymbol{x}) f(\boldsymbol{x})$. For this we first fix a function $f : \mathbb{F}_q^n \to \mathbb{C}$ with $\|f\|_2 = 1$ so that $|f|^2$ will serve as a probability distribution of the errors we feed in the decoder. Regev's reduction consists in performing the following operations:

$$
\begin{aligned}
&\text{Initial state preparation:} && \frac{1}{\sqrt{|\mathscr{C}|}} \sum_{\boldsymbol{c}\in\mathscr{C}} \sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}\rangle|\boldsymbol{e}\rangle \\
&\text{adding } \boldsymbol{c} \text{ to } \boldsymbol{e}: && \mapsto \quad \frac{1}{\sqrt{|\mathscr{C}|}} \sum_{\boldsymbol{c}\in\mathscr{C}} \sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}\rangle|\boldsymbol{c}+\boldsymbol{e}\rangle \\
&\text{erase the 1st register by decoding } \boldsymbol{c}: && \mapsto \quad |\boldsymbol{0}\rangle \frac{1}{\sqrt{|\mathscr{C}|}} \sum_{\boldsymbol{c}\in\mathscr{C},\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}+\boldsymbol{e}\rangle \\
&\text{apply the QFT:} && \mapsto \quad \frac{1}{\sqrt{Z}} \sum_{\boldsymbol{d}\in\mathscr{C}^\perp} \hat{f}(\boldsymbol{d})|\boldsymbol{d}\rangle \\
&\text{measuring the whole state:} && \mapsto \quad |\boldsymbol{d}\rangle \text{ with } \boldsymbol{d} \in \mathscr{C}^\perp.
\end{aligned}
$$

The idea of Regev's algorithm is to make use of a classical decoder recovering $\boldsymbol{c}$ (and therefore $\boldsymbol{e}$) from $\boldsymbol{c}+\boldsymbol{e}$. To be able to perform this operation, we are necessarily in the aforementioned injective regime of decoding. Such an algorithm, if it exists, allows us to "erase" the $|\boldsymbol{c}\rangle$ register when applied coherently. The resulting state $\frac{1}{\sqrt{|\mathscr{C}|}} \sum_{\boldsymbol{c}\in\mathscr{C},\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}+\boldsymbol{e}\rangle = \sum_{\boldsymbol{y}\in\mathbb{F}_q^n} \alpha_{\boldsymbol{y}}|\boldsymbol{y}\rangle$ is periodic in the sense that its amplitudes $\alpha_{\boldsymbol{y}}$ clearly satisfy $\alpha_{\boldsymbol{y}} = \alpha_{\boldsymbol{y}+\boldsymbol{c}}$ for any $\boldsymbol{c}$ in $\mathscr{C}$. This explains why when we apply the quantum Fourier transform to it, we only get non zero amplitudes on the orthogonal space $\mathscr{C}^\perp$. Measuring the final state will give a dual codeword according to a probability distribution proportional to $|\widehat{f}|^2$ restricted on the dual code. If $|f(\boldsymbol{e})|^2$ concentrates its weight on small $\boldsymbol{e}$'s, then it turns out that in many examples of interest $|\hat{f}(\boldsymbol{d})|^2$ also concentrates on rather small $\boldsymbol{d}$'s. This gives a way to sample codewords in $\mathscr{C}^\perp$ of rather small weight/norm.

It is worthwhile to note that by using the same decoding algorithm with a slight twist in Regev's reduction, we can construct $\sum_{\boldsymbol{y}\in\boldsymbol{v}+\mathscr{C}^\perp} \widehat{f}(\boldsymbol{y})|\boldsymbol{y}\rangle$ for any $\boldsymbol{v} \in \mathbb{F}_q^n$. Indeed, Regev's reduction is slightly modified as follows

$$
\begin{aligned}
&\frac{1}{\sqrt{|\mathscr{C}|}} \sum_{\boldsymbol{c}\in\mathscr{C},\boldsymbol{e}\in\mathbb{F}_q^n} \chi_{\boldsymbol{c}}(-\boldsymbol{v}) f(\boldsymbol{e})|\boldsymbol{c}\rangle|\boldsymbol{c}+\boldsymbol{e}\rangle \\
\xrightarrow{\text{Decoder}} &\frac{1}{\sqrt{|\mathscr{C}|}} \sum_{\boldsymbol{c}\in\mathscr{C},\boldsymbol{e}\in\mathbb{F}_q^n} \chi_{\boldsymbol{c}}(-\boldsymbol{v}) f(\boldsymbol{e})|\boldsymbol{c}+\boldsymbol{e}\rangle \\
\xrightarrow{\text{QFT}} &\frac{q^n}{\sqrt{|\mathscr{C}|}} \sum_{\boldsymbol{y}\in\boldsymbol{v}+\mathscr{C}^\perp} \widehat{f}(\boldsymbol{y})|\boldsymbol{y}\rangle.
\end{aligned}
$$

By measuring the resulting state $\sum_{\boldsymbol{y}\in\boldsymbol{v}+\mathscr{C}^\perp} \widehat{f}(\boldsymbol{y})|\boldsymbol{y}\rangle$ we obtain an element in $\boldsymbol{v}+\mathscr{C}^\perp$ such that $\boldsymbol{y}$ is relatively small. In other words, we solve the variant of the decoding problem for the dual code $\mathscr{C}^\perp$ for which we just have to find for any word $\boldsymbol{v}$ of the ambient space a sufficiently close codeword. This is generally the decoding problem as it appears in rate distortion theory [CT91, Chap 10]. It might be also called "decoding in the surjective regime", where we are not asked to find the closest codeword, but a close enough codeword. In other words, Regev's reduction can be tweaked to transform a classical decoding algorithm working in the injective regime into a quantum decoding algorithm working in the surjective regime. Recently this approach has been applied on structured code families such as Reed-Solomon codes [JSW+24, CT25] for which there exist efficient decoding algorithms. This approach

yields a polynomial time algorithm for decoding the dual Reed Solomon code in the surjective regime. For suitable choices of the function f this seems to yield a quantum polynomial time algorithm for decoding in cases where the best known classical algorithms take exponential time, thus yielding a problem which is a good candidate for quantum advantage.

Improving Regev's reduction by quantum decoding. In principle, this reduction gives a quantum algorithm for producing in a code low weight codewords. As a rule of thumb, the greater the decoding radius is for the decoding algorithm used for erasing the first register, the lower the weight of the codewords is after measuring in the last step. However, this reduction is somewhat unsatisfactory in the sense that even if we take the best decoding algorithm for the Hamming distance, we are far for obtaining the minimum (Hamming) weight codewords in the dual code [DRT24]. There has been a recent trend to improve this reduction by a slight variation of the step consisting in erasing the first register [CLZ22]. Indeed, instead of trying to recover c from $c+e$, the new approach is to observe that the state at the second step can be written as

$$\frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c}\in\mathscr{C}}\sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}\rangle|\boldsymbol{c}+\boldsymbol{e}\rangle = \frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c}\in\mathscr{C}}|\boldsymbol{c}\rangle|\psi_{\boldsymbol{c}}\rangle, \quad \text{where } |\psi_{\boldsymbol{c}}\rangle \stackrel{\triangle}{=} \sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}+\boldsymbol{e}\rangle.$$

$|\psi_{\boldsymbol{c}}\rangle$ can be viewed as a superposition of all noisy versions of the codeword $\boldsymbol{c}$ and to recover $\boldsymbol{c}$ and then subtract it to erase the first register, we might as well recover $\boldsymbol{c}$ from this noisy superposition $|\psi_{\boldsymbol{c}}\rangle = \sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}+\boldsymbol{e}\rangle$. In other words we have to solve the following *quantum decoding problem* (which is called S$-|$LWE$\rangle$ in [CLZ22])

Problem 3 (Quantum Decoding Problem QDP(q,n,k,f))**.** *Take q,n,k positive integers and $f:\mathbb{F}_q^n\to\mathbb{C}$ so that $\|f\|_2=1$.*

Take $\mathbf{G}\leftarrow\mathbb{F}_q^{k\times n}$, $\boldsymbol{u}\leftarrow\mathbb{F}_q^k$, and $\boldsymbol{c}=\boldsymbol{u}\mathbf{G}$, $|\psi_{\boldsymbol{c}}\rangle=\sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}+\boldsymbol{e}\rangle$.

Given $(\mathbf{G},|\psi_{\boldsymbol{c}}\rangle)$, the goal is to recover $\boldsymbol{c}$.

Note that this problem reduces to the classical decoding problem DP(q,n,k,p) where $p(\boldsymbol{x})=|f(\boldsymbol{x})|^2$ by measuring the quantum input state. However, this quantum decoding problem might turn out to be intrinsically easier than classical decoding. Indeed, for an f which has a product structure $f(\boldsymbol{e})=\prod_{i=1}^n g(e_i)$, Chen, Liu and Zhandry showed in [CLZ22] that one can use a well chosen quantum measurement on each qudit of $|\psi_{\boldsymbol{c}}\rangle$ (note that in this case $|\psi_{\boldsymbol{c}}\rangle$ has the form $\bigotimes_{i=1}^n\sum_{e_i\in\mathbb{F}_q} g(e_i)|c_i+e_i\rangle$) yielding partial information on the symbols c_i of the codeword $\boldsymbol{c}$. This in turn can be used together with the Arora-Ge algorithm [AG11] to recover $\boldsymbol{c}$. With a well chosen function g which is such that its Fourier transform $\hat{g}$ has support included in $\{-\frac{q-1}{2}+c,\cdots,\frac{q-1}{2}-c\}$ for some constant c, one obtains after performing the rest of Regev's reduction (namely applying the QFT after erasing the first register and measuring at the end) codewords $\boldsymbol{y}$ in $\mathscr{C}^\perp$ which are of infinite norm $\|\boldsymbol{y}\|_\infty\leq\frac{q-1}{2}-c$. It is then shown in [CLZ22] that this results for constant c in a quantum polynomial time algorithm performing this task of producing rather low weight codewords for the infinite norm in a regime of parameters for which all known classical algorithms have subexponential complexity. In other words, this gives an excellent candidate for providing a quantum advantage. This research thread has been further explored in [CHL+23, DFS24].

Understanding in depth the Quantum Decoding Problem is therefore a very important question. Probably the most fundamental issue in this respect is to ask is for a given family $(f_n)_{n\geq1}$ of noise functions what is the largest achievable rate R for which as soon as the rate $\frac{k}{n}$ is below R there exists a quantum algorithm (of potentially unbounded complexity) that solves this problem with high probability. The authors of [CT24] give the largest achievable rate in the case f corresponds to a q-ary symmetric channel, namely when $f(\boldsymbol{e})=\sqrt{(1-p)^{n-|\boldsymbol{e}|}\left(\frac{p}{q-1}\right)^{|\boldsymbol{e}|}}$ for some p which is the crossover probability of the q-ary symmetric channel. In this work, we generalize the above result to any product error function $f_n=g^{\otimes n}$ giving generic achievability and matching non-achievability bounds. Contrarily to the non-achievability results of [CT24]

which hold only for random codes, we prove the relevant non-achievability result for *any code*. While our work share some techniques with [CT24], we use more information theoretic tools which allow us to cleanly phrase our results in terms of the q-ary entropy of the function $|\widehat{f}|^2$. We now present our contributions more concretely.

1.2 Contributions

The capacity of the classical-quantum channel. The mapping

$$\begin{aligned} \alpha \in \mathbb{F}_q \quad &\to \quad |\psi_\alpha\rangle \stackrel{\triangle}{=} \sum_{u \in \mathbb{F}_q} g(u)|\alpha + u\rangle \\ \boldsymbol{c} = (c_1, \cdots, c_n) \quad &\to \quad |\psi_{\boldsymbol{c}}\rangle = |\psi_{c_1}\rangle \otimes \cdots \otimes |\psi_{c_n}\rangle \end{aligned}$$

for g of norm 1 can be considered as a classical-quantum (c-q) pure state channel [Hol12, §5.1]. The maximum rate at which we can transmit information on such a channel is given by the Holevo capacity C_χ of the channel (see [Hol12, Theorem 5.4]) which in the case of our pure state channel takes the form

$$C_\chi = \sup_{p_\alpha} S\left(\sum_{\alpha \in \mathbb{F}_q} p_\alpha \rho_\alpha\right) \tag{1}$$

where p_α is a probability distribution over $\mathbb{F}_q$, $\rho_\alpha \stackrel{\triangle}{=} |\psi_\alpha\rangle\langle\psi_\alpha|$ and $S(\rho)$ is the von Neumann entropy of the quantum state ρ (see §2.3). It is readily seen (see §3.1) that the Holevo capacity of this c-q channel is given by

$$C_\chi = H_q(|\widehat{g}|^2), \tag{2}$$

where H_q is the q-ary entropy (see Definition 4 for the definition of the q-ary entropy H_q if needed). This follows on the spot by noticing that the capacity of this c-q channel is the same as the one that maps α to the quantum Fourier transform $|\widehat{\psi_\alpha}\rangle$ of $|\psi_\alpha\rangle$. It is easy to see that the Holevo capacity of this channel is nothing but $S\left(\frac{1}{q}\sum_{\alpha \in \mathbb{F}_q} \rho'_\alpha\right)$ where $\rho'_\alpha \stackrel{\triangle}{=} |\widehat{\psi_\alpha}\rangle\langle\widehat{\psi_\alpha}|$. The matrix $\frac{1}{q}\sum_{\alpha \in \mathbb{F}_q} \rho'_\alpha$ is nothing but the diagonal matrix with entries the $|\widehat{g}(\alpha)|^2$. This explains (2). Therefore, $H_q(|\widehat{g}|^2)$ is an upper-bound on the achievability rate for our decoding problem, since the issue in our case is whether or not random linear codes achieve the Holevo capacity.

Achievability result for random linear codes. This turns out to be the case and the fact that random linear codes achieve the Holevo capacity of this c-q channel basically comes from the fact that the supremum in (1) is attained for a uniform distribution p_α, *i.e.* the Holevo capacity C_χ of this c-q channel is nothing but the symmetric Holevo capacity of the channel that is

$$\begin{aligned} C_\chi \quad &= \quad S(\rho), \text{ where} \\ \rho \quad &\stackrel{\triangle}{=} \quad \frac{1}{q}\sum_{\alpha \in \mathbb{F}_q} \rho_\alpha. \end{aligned}$$

It turns out in our case that the optimal decoding strategy for linear codes (meaning here the one that maximizes the average probability of success) is given by the Pretty Good Measurement (PGM) which is a generic measurement for quantum state discrimination. This measurement was already studied in [CT24] in the case f corresponded to a q-ary symmetric channel. We generalize this result to treat the case of a general function f which shows that the Holevo capacity of this channel is indeed attained by linear codes

Theorem 1. *Let $q \geq 2$ be a prime power, $R \in (0,1)$ and $g : \mathbb{F}_q \to \mathbb{C}$ such that $\|g\|_2 = 1$. If $R < H_q(|\widehat{g}|^2)$ then there exists a quantum algorithm that solves $\mathrm{QDP}(q, n, \lfloor Rn \rfloor, g^{\otimes n})$ with probability $1 - \mathrm{negl}(n)$.*

Non-achievability result in the general case. The fact that the Holevo capacity of the c-q channel is equal to $H_q(|\widehat{g}|^2)$ already shows that this achievability result is really tight and that it holds in a more general setting: it does not need the generator matrix of the linear code to be chosen uniformly at random, it does not even need the code to be linear. To state this impossibility result, we are going to define a more general QDP problem and will overload the notation as follows

Problem 4 (Quantum Decoding Problem QDP($\mathscr{C}, f$))**.** *Take a subset $\mathscr{C}$ of $\mathbb{F}_q^n$ and $f : \mathbb{F}_q^n \to \mathbb{C}$ so that $\|f\|_2 = 1$. We choose a codeword $\boldsymbol{c}$ uniformly at random in $\mathscr{C}$ and let $|\psi_{\boldsymbol{c}}\rangle = \sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}+\boldsymbol{e}\rangle$. Given $(\mathscr{C}, |\psi_{\boldsymbol{c}}\rangle)$, the goal is to recover $\boldsymbol{c}$.*

This problem is more general than the previous one and even in this case if we choose the code in the best possible way we can not avoid the following non-achievability result.

Theorem 2. *Let $g : \mathbb{F}_q \to \mathbb{C}$ with $\|g\|_2 = 1$. Let R be a fixed constant in $(0, 1)$ such that $R > H_q(|\widehat{g}|^2)$. For any quantum algorithm and any code $\mathscr{C}$ with q^{Rn} codewords the probability p_{succ} to solve $\mathrm{QDP}(\mathscr{C}, g^{\otimes n})$ satisfies*

$$p_{succ} \leq q^{n(H_q(|\widehat{g}|^2)-R+\frac{1}{n^{1/3}})} + \mathrm{negl}(n) = \mathrm{negl}(n).$$

The strong converse theorem [ON00] already shows that $p_{\mathrm{succ}} = \mathrm{negl}(n)$, however the proof technique of Theorem 2 is of independent interest. The proof for the case at hand is much shorter. We will make use of the particular form of the quantum channel states ρ_α at hand. In terms of proof techniques, we look again at the quantum Fourier transforms of our states $\widehat{|\psi_{\boldsymbol{c}}\rangle}$ and, by using the notion of typical sets, we can show that all of the states $\widehat{|\psi_{\boldsymbol{c}}\rangle}$ approximately lie in the same subspace of dimension $q^{nH_q(|\widehat{g}|^2)}$. The condition $H_q(|\widehat{g}|^2) < R$ implies $\frac{q^{nH_q(|\widehat{g}|^2)}}{q^{Rn}} = \mathrm{negl}(n)$. Since there are q^{Rn} different codewords, we will be able to recover any $\boldsymbol{c}$ from $\widehat{|\psi_{\boldsymbol{c}}\rangle}$ only with negligible probability. Formally, we perform this argument using quantum information theory techniques which are inspired by conditional quantum min-entropy notions (see [KRS09] for example).

A quantum information theoretic proof of the discrete Hirschman uncertainty principle. With our two theorems, we have a good understanding of the information theoretic difficulty of the Quantum Decoding Problem. Clearly there is a huge gain in terms of the noise that can be corrected with the quantum version of the problem instead of measuring and solving the classical decoding problem: the quantum decoding problem can be solved up to a rate R satisfying $R < H_q(|\widehat{g}|^2)$ whereas solving the classical problem can be achieved only up to the classical capacity of the additive channel with probability distribution $|g|^2$. The classical capacity is equal to $1 - H_q(|g|^2)$ and there is in general a significant gap between the c-q capacity $H_q(|\widehat{g}|^2)$ of the channel and the classical capacity $1 - H_q(|g|^2)$. In any case, the fact that the c-q capacity can not be smaller than the classical channel capacity we get after measuring the pure state $|\psi_{\boldsymbol{c}}\rangle$ output by the c-q channel shows the discrete Hirschman uncertainty principle [DCT91, Th. 23]

Proposition 1 (Discrete Hirschman uncertainty principle)**.** *For any function $g : \mathbb{F}_q \mapsto \mathbb{C}$ satisfying $\|g\|_2 = 1$ we have*

$$H_q(|g|^2) + H_q(|\widehat{g}|^2) \leq 1$$

Note that this gives for the first time a (quantum) information theoretic proof of this result [DCT91, Th. 23] which was first stated and proved in a slightly more general form in [DCT91] by completely other means.

The power of Regev's reduction with Pretty Good Measurement. Our fourth contribution –and this is the main contribution of this article– is that this significant gap translates into giving much more power to Regev's reduction. As observed in [CT23, §6], there is no hope of having a generic Regev's reduction. However in the case of Regev's reduction based on using the PGM and a Bernoulli noise, the situation is a little bit better, since with a slight tweak of the PGM, [CT24] obtained an algorithm which outputs dual codewords whose weight can be as low as the minimum Hamming distance of the dual code at the information theoretic limit where the PGM works. We generalize this result here. To state this result, we introduce two probability distributions

Definition 1. *We denote by p the probability distribution on $\mathbb{F}_q^n$ equal to $(|\widehat{g}|^2)^{\otimes n}$ and by r the conditional probability distribution of p given that we are in $(\mathscr{C}^\perp)^*$, i.e.*

$$r(\boldsymbol{x}) = \begin{cases} 0 & \text{if } \boldsymbol{x} \notin (\mathscr{C}^\perp)^* \\ \dfrac{p(\boldsymbol{x})}{\sum\limits_{\boldsymbol{y}\in(\mathscr{C}^\perp)^*} p(\boldsymbol{y})} & \text{else} \end{cases}, \text{ well defined when } \sum_{\boldsymbol{y}\in(\mathscr{C}^\perp)^*} p(\boldsymbol{y}) \neq 0.$$

In our general setting, a tweaked version of the PGM can be used in Regev's reduction in order to sample non zero dual codewords according to the probability distribution r. This allows to sample codewords of rather large probability p, since we will show:
(i) the dual codewords $\boldsymbol{c}$ we sample are typically such that $p(\boldsymbol{c}) \approx q^{-nH_q(|\widehat{g}|^2)}$,
(ii) we cannot do much better than this since it turns out that for most dual codes of rate $1-R$ there is no dual codeword $\boldsymbol{c}$ of probability $p(\boldsymbol{c})$ greater than a quantity which is about q^{-nR}.

Since the limiting rate at which the PGM works is precisely $H_q(|\widehat{g}^2|)$, we see that at this limit we obtain an algorithm outputting dual codewords $\boldsymbol{c}$ whose probability $p(\boldsymbol{c})$ is close to be maximal, and this maximum is about q^{-nR}. More precisely, we have

Theorem 3. *Let $g : \mathbb{F}_q \to \mathbb{C}$ with $\|g\|_2 = 1$. Let R be a real number in $(0, H_q(|\widehat{g}|^2))$. Let $f : \mathbb{F}_q^n \to \mathbb{C}$ be so that $f = g^{\otimes n}$. Let $\mathscr{C}$ be a linear code on $\mathbb{F}_q^n$ with generator matrix $\mathbf{G}$ drawn uniformly at random in $\mathbb{F}_q^{\lfloor Rn \rfloor \times n}$. For any $\epsilon > 0$:*

- *Algorithm 4.1 outputs codewords $\boldsymbol{c} \in (\mathscr{C}^\perp)^*$ such that $p(\boldsymbol{c}) \geq q^{-n(H_q(|\widehat{g}^2|)-\epsilon)}$ with probability $1 - \mathrm{negl}(n)$.*
- *The probability over $\mathbf{G}$ that there exists a codeword $\boldsymbol{c}$ in $(\mathscr{C}^\perp)^*$ such that $p(\boldsymbol{c})$ is larger than $q^{-n(R-\epsilon)}$ is a negligible function of n.*

There are cases where the mapping associating to $\boldsymbol{x} \in \mathbb{F}_q^n$ the quantity $\log \frac{p(\mathbf{0})}{p(\boldsymbol{x})} = n \log |\widehat{g}(0)|^2 - \sum_{i=1}^n \log |\widehat{g}(x_i)|^2$ is a weight $|\cdot|$ that defines a metric as $d(\boldsymbol{x}, \boldsymbol{y}) = |\boldsymbol{y} - \boldsymbol{x}|$. The "q-ary symmetric channel" considered in [CT24] is basically an example of this kind. There, we had $g(0) = \sqrt{1-p}$ and $g(\alpha) = \sqrt{\frac{p}{q-1}}$ for $\alpha \neq 0$. In this case, the most likely element (according to the probability distribution p) is the minimum weight codeword in $(\mathscr{C}^\perp)^*$ since p is in this case obviously a decreasing function of the weight $|\cdot|$. In other words, we obtain in this case through Regev's reduction and the PGM a way to obtain a minimum weight codeword in $(\mathscr{C}^\perp)^*$ at the limiting rate where the PGM works (*i.e.* at the Holevo capacity of the c-q channel). This is in strong contrast with using a classical decoder instead in Regev's reduction as shown in [DRT24]. In the last case, we are really far away from getting a minimum weight codeword at the limit where classical decoding works when applied to the noise corresponding to a q-ary symmetric channel. In some sense, this means that in order to get a tight Regev reduction, it is really a quantum decoder (*i.e.* an algorithm solving the quantum decoding problem) which is desirable rather than a classical decoder.

The rank metric case. In all our proofs we make heavily use of the notion of typical set associated to the probability distribution $p \stackrel{\triangle}{=} |\widehat{f}|^2$. This notion applies to more general probability distributions than those corresponding to memoryless c-q channels. We have given a more general form, namely Theorems 4 and 5 giving a tight achievability in the case where for the probability distribution p the elements of probability $\approx q^{H_q(p)}$ carry almost all the probability of the space (meaning that they form a typical set in the usual sense). In this case again, the critical rate marking the distinction between achievability of decoding and non achievability corresponds to the entropy per symbol $\frac{H_q(p)}{n}$. We illustrate this behavior in a case which is of interest in code based cryptography which corresponds to the rank metric. This alternative in code-based cryptography to the usual Hamming metric has lead to submissions to the NIST competition for post-quantum cryptography [ABD+19, AAB+24] which compare favorably to the corresponding submissions in Hamming metric. It can be used to obtain for instance efficient public key encryption or key exchange schemes [ABD+19, ADG+24], signature schemes [AAB+24, BCF+25], or somewhat homomorphic encryption schemes [ADG25]. In this case, the distribution of p does not correspond to a memoryless channel. Nevertheless, we give a tight achievability result in this case (see Theorems 6 and 7) as a corollary of Theorems 4 and 5. The probability distribution p used in this case is a decreasing function of the rank weight. We also give a more general form of Theorem 3 showing that Regev's reduction when applied to a variant of the PGM yields elements of the dual code of maximal probability at the limit where the PGM works. This more general result applies also to the rank metric case and shows that again we get at the limit where the PGM works elements in the dual code of minimum rank weight.

2 Notation and Preliminaries

2.1 Notation

General notation. The finite field with q elements is denoted by $\mathbb{F}_q$. The set of nonnegative integers is denoted by $\mathbb{N}$. The cardinality of a finite set $\mathcal{E}$ is denoted by $|\mathcal{E}|$. For a vector space W and subset V of it, we write that $V \leq W$ iff V is a subspace of W. A *negligible function* is a function $\eta : \mathbb{N} \to \mathbb{R}$ such that for every positive integer a there exists an integer N such that for all $n \geq N$ we have $\eta(n) \leq \frac{1}{n^a}$. When for a function $f : \mathbb{N} \to \mathbb{R}$ we write that $f(n) = \mathrm{negl}(n)$, this means that f is a negligible function.

Vector and matrices. For a Hermitian matrix $\mathbf{M}$ we write that $\mathbf{M} \succcurlyeq 0$ when $\mathbf{M}$ is positive semidefinite. Vectors are row vectors as is standard in the coding community and $\boldsymbol{x}^\intercal$ (resp. $\mathbf{M}^\intercal$) denotes the transpose of a vector (resp. of a matrix). In particular, vectors will always be denoted by bold small letters and matrices with bold capital letters. The inner product between two vectors $\boldsymbol{x} = (x_i)_{1 \leq i \leq n}$ and $\boldsymbol{y} = (y_i)_{1 \leq i \leq n}$ of $\mathbb{F}_q^n$ is denoted by $\boldsymbol{x} \cdot \boldsymbol{y} \stackrel{\triangle}{=} \sum_{i=1}^n x_i y_i$. $\mathbf{I}$ denotes the identity matrix. Its size will be clear from the context.

The classical and quantum Fourier transform on $\mathbb{F}_q^n$. In this article, we will use the quantum Fourier transform on $\mathbb{F}_q^n$. It is based on the characters of the group $(\mathbb{F}_q^n, +)$ which are defined as follows (for more details see [LN97, Chap 5, §1], in particular a description of the characters in terms of the trace function is given in [LN97, Ch. 5, §1, Th. 5.7]).

Definition 2. *Fix $q = p^s$ for a prime integer p and an integer $s \geq 1$. The characters of $\mathbb{F}_q$ are the functions $\chi_y : \mathbb{F}_q \to \mathbb{C}$ indexed by elements $y \in \mathbb{F}_q$ defined as follows*

$$\begin{aligned} \chi_y(x) &\stackrel{\triangle}{=} e^{\frac{2i\pi \operatorname{tr}(xy)}{p}}, \quad \textit{with} \\ \operatorname{tr}(a) &\stackrel{\triangle}{=} a + a^p + a^{p^2} + \cdots + a^{p^{s-1}}. \end{aligned}$$

We extend the definition to vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}_q^n$ as follows:

$$\chi_{\boldsymbol{y}}(\boldsymbol{x}) \stackrel{\triangle}{=} \prod_{i=1}^{n} \chi_{y_i}(x_i) = e^{\frac{2i\pi \operatorname{tr}(\boldsymbol{x}\cdot\boldsymbol{y})}{p}}.$$

When q is prime, we have $\chi_y(x) = e^{\frac{2i\pi xy}{q}} = \omega_q^{xy}$, where $\omega_q \stackrel{\triangle}{=} e^{\frac{2i\pi}{q}}$.

Definition 3. *Let $f : \mathbb{F}_q^n \to \mathbb{C}$. We define $\widehat{f} : \mathbb{F}_q^n \to \mathbb{C}$ such that $\widehat{f}(\boldsymbol{y}) = \frac{1}{\sqrt{q^n}} \sum_{\boldsymbol{x}\in\mathbb{F}_q^n} \chi_{\boldsymbol{y}}(\boldsymbol{x}) f(\boldsymbol{x})$.*

The way we normalize the Fourier transform $\widehat{f}$ of a function f ensures that $\|f\|_2 = \|\widehat{f}\|_2$. The quantum Fourier transform over $\mathbb{F}_q^n$ is the quantum unitary such that for each $f : \mathbb{F}_q^n \to \mathbb{C}$ such that $\|f\|_2 = 1$, we have

$$\mathrm{QFT}_{\mathbb{F}_q^n}\left(\sum_{\boldsymbol{x}\in\mathbb{F}_q^n} f(\boldsymbol{x})|\boldsymbol{x}\rangle\right) = \sum_{\boldsymbol{x}\in\mathbb{F}_q^n} \widehat{f}(\boldsymbol{x})|\boldsymbol{x}\rangle.$$

Linear codes. A linear code over $\mathbb{F}_q$ of length n is a subspace of $\mathbb{F}_q^n$. We say that it is an $[n, k]$ code when its dimension is k. It can be specified by a generating matrix $\mathbf{G} \in \mathbb{F}_q^{k\times n}$, in which case $\mathscr{C} = \{\boldsymbol{s}\mathbf{G} : \boldsymbol{s} \in \mathbb{F}_q^k\}$. We also define $\mathscr{C}^* = \mathscr{C}\backslash\{\mathbf{0}\}$. The dual code $\mathscr{C}^\perp$ which is defined as $\mathscr{C}^\perp = \{\boldsymbol{y} \in \mathbb{F}_q^n : \boldsymbol{y}\cdot\boldsymbol{c} = 0\}$ can also be described as

$$\mathscr{C}^\perp = \{\boldsymbol{y} \in \mathbb{F}_q^n : \chi_{\boldsymbol{y}}(\boldsymbol{c}) = 1,\ \forall \boldsymbol{c} \in \mathscr{C}\}. \tag{3}$$

When the code length n is a product $n = a \cdot b$ we can arrange the entries of a vector $\boldsymbol{x} = (x_i)_{0\le i<n} \in \mathbb{F}_q^n$ in a matrix $\mathbf{X} = \mathrm{Mat}(\boldsymbol{x}) \in \mathbb{F}_q^{a\times b}$. This is obtained by defining the entry $X_{i,j}$ in the i-th row and the j-th column of $\mathbf{X}$ as $X_{i,j} = x_{i\cdot b+j}$ where $i \in \{0, \cdots, a-1\}$, $j \in \{0, \cdots, b-1\}$. This allows us to define the rank weight on $\mathbb{F}_q^n$ as

$$|\boldsymbol{x}|_{\mathrm{rk}} \stackrel{\triangle}{=} \mathrm{rank}(\mathrm{Mat}(\boldsymbol{x})).$$

2.2 Information Theory

We first recall the notion of entropy

Definition 4. *Let p be a probability distribution on a finite alphabet $\mathcal{A}$. Its q-ary entropy is defined as*

$$H_q(p) \stackrel{\triangle}{=} -\sum_{x\in\mathcal{A}} p(x)\log_q p(x).$$

We will use this notion to define what is a typical set for a probability distribution defined over $\mathbb{F}_q^n$:

Definition 5 (typical set)**.** *The typical set $\mathcal{T}_\epsilon^{(n)}(p)$ corresponding to the probability distribution p over $\mathbb{F}_q^n$, gap ϵ where ϵ is a positive number and bounding functions $A(\epsilon)$ and $B(\epsilon)$ is defined by*

$$\mathcal{T}_\epsilon^{(n)}(p) = \{\boldsymbol{y} \in \mathbb{F}_q^n : A(\epsilon) \le p(\boldsymbol{y}) \le B(\epsilon)\}.$$

We will generally drop the dependence in p which will be clear from the context and simply write $\mathcal{T}_\epsilon^{(n)}$. The defect $\delta(\epsilon, n)$ is the function of ϵ defined by $1 - p(\mathcal{T}_\epsilon^{(n)})$.

A direct consequence of this definition is that we have

Lemma 1. *For all $\epsilon > 0$ and positive integer n we have*

$$(1-\delta(\epsilon,n))\frac{1}{B(\epsilon)} \leq |\mathcal{T}_\epsilon^{(n)}| \leq \frac{1}{A(\epsilon)}.$$

The rationale behind this definition is that for many families of probability distributions over $\mathbb{F}_q^n$ we have $p(\mathcal{T}_\epsilon^{(n)}) = 1 - o(1)$ for any fixed ϵ as n tends to infinity for bounding functions $A(\epsilon)$ and $B(\epsilon)$ which are close to each other. In particular it holds for product probability distributions $p = r^{\otimes n}$ for fixed r as is well known

Proposition 2. *[CT06, Section 3.1]*

Let $r : \mathbb{F}_q \to [0,1]$ be a probability function on $\mathbb{F}_q$. Let $p = r^{\otimes n}$ and $\mathcal{T}_\epsilon^{(n)}$ be the associated ϵ-typical set for some $n \in \mathbb{N}$ and $\epsilon > 0$ corresponding to the bounding functions $A(\epsilon) = q^{-n(H_q(r)+\epsilon)}$ and $B(\epsilon) = q^{-n(H_q(r)-\epsilon)}$.

1. *$\delta(\epsilon, n) \leq e^{\frac{-2n\epsilon^2}{K}}$ for some constant K that depends only on r. In particular, when $\epsilon = \frac{1}{n^{1/3}}$, then, $\delta(\epsilon, n) = \mathrm{negl}(n)$.*
2. *$(1-\delta(\epsilon,n))q^{n(H_q(r)-\epsilon)} \leq |\mathcal{T}_\epsilon^{(n)}| \leq q^{n(H_q(r)+\epsilon)}$.*

Proof. Fix $\mathcal{T}_\epsilon^{(n)}$ for some $r : \mathbb{F}_q \to [0,1], n \in \mathbb{N}$ and $\epsilon > 0$. Let $p = r^{\otimes n}$. Let $\boldsymbol{x} = (x_1, \cdots, x_n)$ be a random variable on $\mathbb{F}_q^n$ distributed according to p. We have

$$\begin{aligned}\delta(\epsilon,n) &= \Pr\left[\boldsymbol{x} \notin \mathcal{T}_\epsilon^{(n)}\right] \\ &= \Pr\left[\left|\sum_{i=1}^{n} -\log_q(r(x_i)) - nH_q(r)\right| \geq n\epsilon\right]\end{aligned}$$

Let $X_1, \ldots, X_n$ be random i.i.d. variables such that

$$\Pr[X_i = \alpha] = \begin{cases} -\log_q(r(\alpha)) & \text{if } r(\alpha) \neq 0 \\ 0 & \text{if } r(\alpha) = 0 \end{cases}$$

Let $S_n = \sum_{i=1}^n X_i$. Notice that $\mathbb{E}[X_i] = H_q(r)$ hence $\mathbb{E}[S_n] = nH_q(r)$. Also, notice that $0 \leq X_i \leq K$ where $K = \max_{\alpha : r(\alpha) \neq 0} \{-\log_q(r(\alpha))\}$. Using Hoeffding's inequality[1], we write

$$\delta(\epsilon,n) = \Pr(S_n - \mathbb{E}[S_n] \geq t) \leq e^{-\frac{2t^2}{nK^2}} = e^{-\frac{2\epsilon^2 n}{K^2}}$$

The second point follows directly from Lemma 1. □

We will also need the following results for the variance and the expectation of the cardinality of the intersection of a random linear code on $\mathbb{F}_q^n$ with an arbitrary subset of $\mathbb{F}_q^n$.

Lemma 2. *(Intersection Expectation and Variance lemma) [Bar97, Lemma 1.1] Let $\mathscr{C}$ be a random linear code on $\mathbb{F}_q^n$ whose parity-check matrix $\mathbf{H}$ is chosen uniformly at random in $\mathbb{F}_q^{(n-k)\times n}$. Let E be an arbitrary subset of $\mathbb{F}_q^n$. Then*

$$\frac{|E|}{q^{n-k}} \leq \mathbb{E}_{\mathbf{H}}(|\mathscr{C} \cap E|) \leq \frac{|E|-1}{q^{n-k}} + 1$$

$$\mathrm{Var}_{\mathbf{H}}(|\mathscr{C} \cap E|) \leq (q-1)\mathbb{E}_{\mathbf{H}}(|\mathscr{C} \cap E|)$$

[1] For any i.i.d random variables $X_1, \ldots X_n$ such that $\Pr(a_i \leq X_i \leq b_i) = 1$, for any $t > 0$, $\Pr(S_n - \mathbb{E}[S_n] \geq t) \leq \exp\left(-\frac{2t^2}{M}\right)$ with $M = \sum_{i=1}^n (b_i - a_i)^2$ and $S_n = \sum_{i=1}^n X_i$.

2.3 Quantum information theory

Definition 6 (von Neumann entropy)**.** *The von Neumann entropy $S(\rho)$ of a quantum state ρ living in a d-dimensional Hilbert space is defined as $S(\rho) \stackrel{\triangle}{=} -\sum_{i=1}^{d} \lambda_i \log_d \lambda_i$ where $\rho = \sum_{i=1}^{d} \lambda_i |e_i\rangle\langle e_i|$ is the spectral decomposition of ρ.*

Definition 7. *For two quantum states ρ, σ, we define their trace distance $D(\rho, \sigma) \stackrel{\triangle}{=} \frac{1}{2}\operatorname{tr}(|\rho - \sigma|)$ where $|A|$ is the positive square root of $A^\dagger A$.*

Claim 1 ([NC10])**.** *If $\rho = |\phi\rangle\langle\phi|$ and $\sigma = |\psi\rangle\langle\psi|$, then $D(\rho, \sigma) = \sqrt{1 - |\langle\phi|\psi\rangle|^2}$.*

Claim 2 ([NC10])**.** *$D(\rho, \sigma) = \max_P\{\operatorname{tr}(P\rho - P\sigma)\}$, where the maximum is over all positive operators $P \preccurlyeq I$.*

Claim 3. *Let A, B be two positive semidefinite matrices. We have $\operatorname{tr}(AB) \geq 0$. As a direct corollary, for all A, B and C hermitian matrices such that $A \succcurlyeq 0$ and $B \preccurlyeq C$, we have that $\operatorname{tr}(AB) \leq \operatorname{tr}(AC)$.*

Proof. We write $\operatorname{tr}(AB) = \operatorname{tr}(B^{1/2}AB^{1/2}) \geq 0$, where the inequality comes from the fact that $B^{1/2}AB^{1/2}$ is also positive semidefinite. □

2.4 The Pretty Good Measurement

Definition 8. *Let J be a set of quantum pure states $\{|\psi_j\rangle\}_{1 \leq j \leq |J|}$. Let $\rho = \sum_{j=1}^{|J|} |\psi_j\rangle\langle\psi_j|$. The Pretty Good Measurement associated to this set is the POVM $\{M_j\}_{1 \leq j \leq |J|}$ with*

$$M_j = \rho^{-1/2}|\psi_j\rangle\langle\psi_j|\rho^{-1/2}$$

From now on, we will refer to the Pretty Good Measurement as PGM.

One can check that we have indeed $\sum_j \rho^{-1/2}|\psi_j\rangle\langle\psi_j|\rho^{-1/2} = \mathbf{I}$. In our case, it turns out that the PGM is actually optimal for attaining the highest average probability of success.

Proposition 3. *Assume that the $|\psi_j\rangle$ are equally probable, independent and form a geometrically uniform set of states [EF01, § VIII, A.], in the sense that there exists an Abelian group G of unitaries such that $\{|\psi_j\rangle,\ j \in J\} = \{U|\psi_{j_0}\rangle,\ U \in G\}$ for some $j_0 \in J$. Then the PGM applied to the $|\psi_j\rangle$'s has the largest average probability of success.*

This was proved in [BKMH97, prop. 1] in the case of an Abelian group generated by a single element and the more general form stated here is given in [EF01, §VIII, C, Th. 4]. It is also a consequence of [SKIH98, Th. 1].

3 (Non-)Achievability Results

Our goal is here to first give a proof of the achievability result, namely Theorem 1. This will be obtained by computing the probability of success of the Pretty Good Measurement when applied to a random linear code. It will show that as soon as the rate of the linear code is below $H_q(|\widehat{g}|)^2$, then we can solve the quantum decoding problem with high probability. On the other hand, the non-achievability result, namely Theorem 2, will apply to *any* code and shows that as soon as the rate of the code is above this limit, we can not hope to solve the quantum decoding problem with non-negligible probability anymore. Finally, we will apply these two results in the case of rank metric where we get similar results as for the Hamming metric [CT24], where the limit of achievability corresponds to sample dual codewords of minimum rank distance.

3.1 Notation and Preliminaries

It turns out that for both results it will be very helpful to apply the quantum Fourier transform to the states $|\psi_{\boldsymbol{c}}\rangle$ that we want to distinguish in the quantum decoding problem. In the whole section, $\mathscr{C}$ denotes the code that we want to decode and f denotes the associated noise distribution over $\mathbb{F}_q^n$. It is such that $\|f\|_2 = 1$. The input states to the decoding problem are denoted by $|\psi_{\boldsymbol{c}}\rangle$ for $\boldsymbol{c} \in \mathscr{C}$, *i.e.*

$$|\psi_{\boldsymbol{c}}\rangle \stackrel{\triangle}{=} \sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}+\boldsymbol{e}\rangle.$$

Note that these states are just shifts of the "noise state" $\sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{e}\rangle$ which explains why the quantum Fourier transform applied to the $|\psi_{\boldsymbol{c}}\rangle$'s have all the same amplitudes up to a phase term as shown by

Lemma 3. *We have*

$$\widehat{|\psi_{\boldsymbol{c}}\rangle} = \sum_{\boldsymbol{y}\in\mathbb{F}_q^n} \widehat{f}(\boldsymbol{y})\chi_{\boldsymbol{c}}(\boldsymbol{y})|\boldsymbol{y}\rangle. \tag{4}$$

Proof. (4) follows on the spot from observing that

$$\begin{aligned}
\widehat{|\psi_{\boldsymbol{c}}\rangle} &= \mathrm{QFT}_{\mathbb{F}_q^n}\left(\sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}+\boldsymbol{e}\rangle\right) \\
&= \frac{1}{\sqrt{q^n}}\sum_{\boldsymbol{y}\in\mathbb{F}_q^n}\left(\sum_{\boldsymbol{e}\in\mathbb{F}_q^n}\chi_{\boldsymbol{c}+\boldsymbol{e}}(\boldsymbol{y})f(\boldsymbol{e})\right)|\boldsymbol{y}\rangle \\
&= \sum_{\boldsymbol{y}\in\mathbb{F}_q^n}\underbrace{\frac{1}{\sqrt{q^n}}\sum_{\boldsymbol{e}\in\mathbb{F}_q^n}\chi_{\boldsymbol{e}}(\boldsymbol{y})f(\boldsymbol{e})}_{\widehat{f}(\boldsymbol{y})}\chi_{\boldsymbol{c}}(\boldsymbol{y})|\boldsymbol{y}\rangle \\
&= \sum_{\boldsymbol{y}\in\mathbb{F}_q^n}\widehat{f}(\boldsymbol{y})\chi_{\boldsymbol{c}}(\boldsymbol{y})|\boldsymbol{y}\rangle.
\end{aligned}$$

□

The capacity of the classical-quantum pure state channel in the memoryless case. The Fourier transform turns out to be really helpful for computing the capacity of the c-q channel capacity. Indeed we have

Proposition 4. *We assume that $f = g^{\otimes n}$ where g is a function of L_2 norm 1 defined over $\mathbb{F}_q$. The Holevo capacity of the c-q channel mapping any $\boldsymbol{c} \in \mathbb{F}_q^n$ to the pure state $|\psi_{\boldsymbol{c}}\rangle$ is equal to $H_q\left(|\widehat{g}|^2\right)$.*

Proof. The Holevo capacity $C_{\mathcal{H}}$ of this pure state channel is equal to $\sup_{p_\alpha} S\left(\sum_{\alpha\in\mathbb{F}_q} p_\alpha\rho_\alpha\right)$ where

$$\begin{aligned}
\rho_\alpha &\stackrel{\triangle}{=} |\psi_\alpha\rangle\langle\psi_\alpha| \text{ with} \\
|\psi_\alpha\rangle &\stackrel{\triangle}{=} \sum_{u\in\mathbb{F}_q} g(u)|\alpha+u\rangle.
\end{aligned}$$

For $\beta \in \mathbb{F}_q$ we denote by U_β the unitary mapping $|\alpha\rangle$ to $|\alpha+\beta\rangle$ for any $\alpha \in \mathbb{F}_q$. Let us denote by $(p_\alpha)_{\alpha\in\mathbb{F}_q}$ the probability distribution that attains the supremum of $S\left(\sum_{\alpha\in\mathbb{F}_q} p_\alpha\rho_\alpha\right)$. If

there are several distributions that attain the supremum, we choose arbitrarily one of them. Let us observe that for any $\beta \in \mathbb{F}_q$

$$\begin{aligned} C_{\mathcal{H}} = S\left(\sum_{\alpha\in\mathbb{F}_q} p_\alpha \rho_\alpha\right) &= S\left(U_\beta \left(\sum_{\alpha\in\mathbb{F}_q} p_\alpha \rho_\alpha\right) U_\beta^\dagger\right) \\ &= S\left(\sum_{\alpha\in\mathbb{F}_q} p_\alpha U_\beta |\psi_\alpha\rangle\langle\psi_\alpha| U_\beta^\dagger\right) \\ &= S\left(\sum_{\alpha\in\mathbb{F}_q} p_\alpha \rho_{\alpha+\beta}\right) \\ &= S\left(\sum_{\alpha\in\mathbb{F}_q} q_\alpha^\beta \rho_\alpha\right) \end{aligned} \tag{5}$$

where q^β is the probability distribution over $\mathbb{F}_q$ defined by $q_\alpha^\beta = p_{\alpha-\beta}$. Let $\sigma_\beta \stackrel{\triangle}{=} \sum_{\alpha\in\mathbb{F}_q} q_\alpha^\beta \rho_\alpha$. The above can be therefore rewritten $\forall \beta \in \mathbb{F}_q$, $C_{\mathcal{H}} = S(\sigma_\beta)$. By using now the concavity of the von Neumann entropy we deduce that

$$C_{\mathcal{H}} = \frac{1}{q}\sum_{\beta\in\mathbb{F}_q} S(\sigma_\beta) \leq S\left(\frac{1}{q}\sum_{\beta\in\mathbb{F}_q}\sigma_\beta\right) = S\left(\frac{1}{q}\sum_{\beta\in\mathbb{F}_q}\sum_{\alpha\in\mathbb{F}_q} q_\alpha^\beta \rho_\alpha\right) = S\left(\frac{1}{q}\sum_{\alpha\in\mathbb{F}_q}\rho_\alpha\right).$$

We deduce that the Holevo capacity is equal to $S(\rho)$ where $\rho \stackrel{\triangle}{=} \frac{1}{q}\sum_{\alpha\in\mathbb{F}_q}\rho_\alpha$. It is convenient here to express these density matrices in the Fourier basis, since ρ is diagonal in this basis:

$$\begin{aligned} S(\rho) &= S\left(\frac{1}{q}\sum_{\alpha\in\mathbb{F}_q} |\psi_\alpha\rangle\langle\psi_\alpha|\right) \\ &= S\left(\frac{1}{q}\sum_{\alpha\in\mathbb{F}_q} |\widehat{\psi_\alpha}\rangle\langle\widehat{\psi_\alpha}|\right) \\ &= S\left(\frac{1}{q}\sum_{\alpha\in\mathbb{F}_q}\sum_{\beta\in\mathbb{F}_q}\sum_{\gamma\in\mathbb{F}_q} \chi_\alpha(\beta)\widehat{g}(\beta)\overline{\chi_\alpha(\gamma)\widehat{g}(\gamma)}|\beta\rangle\langle\gamma|\right) \quad \text{(by using Lemma 3)} \\ &= S\left(\frac{1}{q}\sum_{\alpha\in\mathbb{F}_q}\sum_{\beta\in\mathbb{F}_q}\sum_{\gamma\in\mathbb{F}_q} \chi_\alpha(\beta-\gamma)\widehat{g}(\beta)\overline{\widehat{g}(\gamma)}|\beta\rangle\langle\gamma|\right) \\ &= S\left(\sum_{\beta\in\mathbb{F}_q} |\widehat{g}(\beta)|^2|\beta\rangle\langle\beta|\right) \qquad (6) \\ &= H_q(|\widehat{g}|^2). \end{aligned}$$

(6) is consequence of the fact

$$\sum_{\alpha\in\mathbb{F}_q}\chi_\alpha(\beta-\gamma) = \begin{cases} 0 & \text{if } \beta \neq \gamma \\ q & \text{otherwise.} \end{cases}$$

□

Probability of success of the Pretty Good Measurement. Applying a unitary transform does not change the probability of success of pretty good measurement and therefore the probability

of success of the pretty good measurement $\mathrm{P}_{\mathrm{PGM}}$ for recovering $\boldsymbol{c}$ from $|\psi_{\boldsymbol{c}}\rangle$ is equal to the probability of success of the pretty good measurement for recovering $\boldsymbol{c}$ from $|\widehat{\psi_{\boldsymbol{c}}}\rangle$. Let us first rewrite a little bit the expression of $|\widehat{\psi_{\boldsymbol{c}}}\rangle$ by using the fact that $\chi_{\boldsymbol{c}}(\boldsymbol{e})$ is constant on the cosets of $\mathscr{C}^\perp$ for any $\boldsymbol{c}$ in $\mathscr{C}$ since $\chi_{\boldsymbol{c}}(\boldsymbol{c}^\perp)=1$ for any $\boldsymbol{c}$ in $\mathscr{C}$ and $\boldsymbol{c}^\perp$ in $\mathscr{C}^\perp$. For any $\boldsymbol{s}\in\mathbb{F}_q^n/C^\perp$ we fix an arbitrary element $\boldsymbol{u_s}$ in $\boldsymbol{s}+\mathscr{C}^\perp$. Therefore

$$\begin{aligned}
|\widehat{\psi_{\boldsymbol{c}}}\rangle &= \sum_{\boldsymbol{e}\in\mathbb{F}_q^n} \widehat{f}(\boldsymbol{e})\chi_{\boldsymbol{c}}(\boldsymbol{e})|\boldsymbol{e}\rangle \\
&= \sum_{\boldsymbol{s}\in\mathbb{F}_q^n/\mathscr{C}^\perp} \sum_{\boldsymbol{e}\in\boldsymbol{s}+\mathscr{C}^\perp} \widehat{f}(\boldsymbol{e})\chi_{\boldsymbol{c}}(\boldsymbol{e})|\boldsymbol{e}\rangle \\
&= \sum_{\boldsymbol{s}\in\mathbb{F}_q^n/\mathscr{C}^\perp} \chi_{\boldsymbol{c}}(u_{\boldsymbol{s}}) \sum_{\boldsymbol{e}\in\boldsymbol{s}+\mathscr{C}^\perp} \widehat{f}(\boldsymbol{e})|\boldsymbol{e}\rangle \\
&= \sum_{\boldsymbol{s}\in\mathbb{F}_q^n/\mathscr{C}^\perp} \chi_{\boldsymbol{c}}(u_{\boldsymbol{s}})|W_{\boldsymbol{s}}\rangle,
\end{aligned}$$

$$\text{where } |W_{\boldsymbol{s}}\rangle \stackrel{\triangle}{=} \sum_{\boldsymbol{e}\in\boldsymbol{s}+\mathscr{C}^\perp} \widehat{f}(\boldsymbol{e})|\boldsymbol{e}\rangle.$$

Proposition 5. *The PGM succeeds to recover $\boldsymbol{c}$ from $|\widehat{\psi_{\boldsymbol{c}}}\rangle$ with probability $P_{PGM}=\frac{1}{q^k}\left(\sum_{\boldsymbol{s}\in\mathbb{F}_q^n/\mathscr{C}^\perp} Z_{\boldsymbol{s}}\right)^2$ where $Z_{\boldsymbol{s}}\stackrel{\triangle}{=}\||W_{\boldsymbol{s}}\rangle\|$*

Proof. To get the probability of success of the PGM, we need to define the operators of the measurement. For that, we will use the same technique as in [CT23, §5.1]. The operator of the PGM associated to the state $|\widehat{\psi_{\boldsymbol{c}}}\rangle$ is defined as follows:

$$M_{\boldsymbol{c}} = \rho^{-1/2}|\widehat{\psi_{\boldsymbol{c}}}\rangle\langle\widehat{\psi_{\boldsymbol{c}}}|\rho^{-1/2} \text{ with } \rho=\sum_{\boldsymbol{c}\in\mathscr{C}}|\widehat{\psi_{\boldsymbol{c}}}\rangle\langle\widehat{\psi_{\boldsymbol{c}}}|$$

We write

$$\rho=\sum_{\boldsymbol{c}\in\mathscr{C}}|\widehat{\psi_{\boldsymbol{c}}}\rangle\langle\widehat{\psi_{\boldsymbol{c}}}| = q^k\sum_{\boldsymbol{s}\in\mathbb{F}_q^n/\mathscr{C}^\perp} Z_{\boldsymbol{s}}^2|\widetilde{W_{\boldsymbol{s}}}\rangle\langle\widetilde{W_{\boldsymbol{s}}}|$$

where $|\widetilde{W_{\boldsymbol{s}}}\rangle=\dfrac{1}{Z_{\boldsymbol{s}}}|W_{\boldsymbol{s}}\rangle$. Then, $\rho^{-1/2}=\dfrac{1}{\sqrt{q^k}}\sum_{\boldsymbol{s}\in\mathbb{F}_q^n/\mathscr{C}^\perp}\dfrac{1}{Z_{\boldsymbol{s}}}|\widetilde{W_{\boldsymbol{s}}}\rangle\langle\widetilde{W_{\boldsymbol{s}}}|$.
We get

$$\begin{aligned}
\rho^{-1/2}|\widehat{\psi_c}\rangle &= \frac{1}{\sqrt{q^k}}\sum_{\boldsymbol{s},\boldsymbol{s}'\in\mathbb{F}_q^n/\mathscr{C}^\perp}\frac{Z_{\boldsymbol{s}'}}{Z_{\boldsymbol{s}}}\chi_{\boldsymbol{c}}(u_{\boldsymbol{s}})|\widetilde{W_{\boldsymbol{s}}}\rangle\langle\widetilde{W_{\boldsymbol{s}}}|\widetilde{W_{\boldsymbol{s}'}}\rangle \\
&= \frac{1}{\sqrt{q^k}}\sum_{\boldsymbol{s}\in\mathbb{F}_q^n/\mathscr{C}^\perp}\chi_{\boldsymbol{c}}(u_{\boldsymbol{s}})|\widetilde{W_{\boldsymbol{s}}}\rangle
\end{aligned}$$

using the fact that the $|\widetilde{W_{\boldsymbol{s}}}\rangle$ are pairwise orthogonal.
Then, we can write

$$\begin{aligned}
\langle\widehat{\psi_{\boldsymbol{c}}}|\rho^{-1/2}|\widehat{\psi_{\boldsymbol{c}}}\rangle &= \frac{1}{\sqrt{q^k}}\sum_{\boldsymbol{s},\boldsymbol{s}'\in\mathbb{F}_q^n/\mathscr{C}^\perp}\chi_{\boldsymbol{c}}(u_{\boldsymbol{s}'}-u_{\boldsymbol{s}})\langle\widetilde{W_{\boldsymbol{s}}}|\widetilde{W_{\boldsymbol{s}'}}\rangle \\
&= \frac{1}{\sqrt{q^k}}\sum_{\boldsymbol{s}\in\mathbb{F}_q^n/\mathscr{C}^\perp}\langle\widetilde{W_{\boldsymbol{s}}}|W_{\boldsymbol{s}}\rangle
\end{aligned}$$

Finally, we get the probability of success of the PGM:

$$\begin{aligned} p_{\boldsymbol{c}} &= \langle \widehat{\psi_{\boldsymbol{c}}} | M_{\boldsymbol{c}} | \widehat{\psi_{\boldsymbol{c}}} \rangle \\ &= \langle \widehat{\psi_{\boldsymbol{c}}} | \rho^{-1/2} | \widehat{\psi_{\boldsymbol{c}}} \rangle \langle \widehat{\psi_{\boldsymbol{c}}} | \rho^{-1/2} | \widehat{\psi_{\boldsymbol{c}}} \rangle \\ &= \frac{1}{q^k} \left| \sum_{\boldsymbol{s} \in \mathbb{F}_q^n / \mathscr{C}^\perp} \langle \widetilde{W_{\boldsymbol{s}}} | W_{\boldsymbol{s}} \rangle \right|^2 = \frac{1}{q^k} \left(\sum_{\boldsymbol{s} \in \mathbb{F}_q^n / \mathscr{C}^\perp} Z_{\boldsymbol{s}} \right)^2 \end{aligned}$$

□

Nice family of typical sets. All the results of this article rely heavily on the notion of typical sets. The following definition will turn out to be very useful in the rest of the paper.

Definition 9 (nice family of typical sets)**.** *Let $(n_i)_{i\geq 1}$ be an increasing set of positive integers and let $(p_i)_{i\geq 1}$ be a family of probability distributions where for every positive integer i, p_i is a probability distribution over $\mathbb{F}_q^{n_i}$. We say that the sequence $(p_i)_{i\geq 1}$ admits a nice family of typical sets $\left\{\mathcal{T}_\epsilon^{(n_i)}(p_i), i \in \mathbb{N}^*, \epsilon > 0\right\}$ if there exist a constant H, two functions α and β defined over $\mathbb{R}^+$ and ranging over $\mathbb{R}^+$ whose limit at 0 is equal to 0 and two positive constants K_1 and K_2 such that the defect function δ and the bounding functions $A_i(\epsilon)$ and $B_i(\epsilon)$ of the typical set $\mathcal{T}_\epsilon^{(n_i)}$ satisfy*

1. *$\delta(\epsilon, n_i) = \mathrm{negl}(i)$ for every $\epsilon > 0$*
2. *$A_i(\epsilon) \geq K_1 q^{-n_i(H+\alpha(\epsilon))}$ for every positive integer i and any $\epsilon > 0$,*
3. *$B_i(\epsilon) \leq K_2 q^{-n_i(H-\beta(\epsilon))}$ for every positive integer i and any $\epsilon > 0$.*

For the proof of the strong converse of the achievability result, it will be helpful to allow the ϵ involved in the definition to go to 0 as i tends to infinity and still require that the defect function is a negligible function of i. This gives the following definition

Definition 10 (nice family of typical sets in the strong sense)**.** *Let $(n_i)_{i\geq 1}$ be an increasing set of positive integers and let $(p_i)_{i\geq 1}$ be a family of probability distributions where for every positive integer i, p_i is a probability distribution over $\mathbb{F}_q^{n_i}$. We say that the sequence $(p_i)_{i\geq 1}$ admits a nice family of typical sets in the strong sense $\left\{\mathcal{T}_{\epsilon(i)}^{(n_i)}(p_i), i \in \mathbb{N}^*\right\}$ where ϵ is a function from $\mathbb{N}^*$ to $\mathbb{R}^+$ where $\epsilon(i)$ tends to 0 as i tends to infinity if there exist a constant H, two functions α and β defined over $\mathbb{R}^+$ and ranging over $\mathbb{R}^+$ whose limit at 0 is equal to 0 and two positive constants K_1 and K_2 such that the defect function δ and the bounding functions $A_i(\epsilon)$ and $B_i(\epsilon)$ of the typical set $\mathcal{T}_\epsilon^{(n_i)}$ satisfy*

1. *$\delta(\epsilon(i), n_i) = \mathrm{negl}(i)$,*
2. *$A_i(\epsilon) \geq K_1 q^{-n_i(H+\alpha(\epsilon))}$ for every positive integer i and any $\epsilon > 0$,*
3. *$B_i(\epsilon) \leq K_2 q^{-n_i(H-\beta(\epsilon))}$ for every positive integer i and any $\epsilon > 0$.*

Product probability distributions $p_i = p^{\otimes n_i}$ where p is a probability distribution over $\mathbb{F}_q$ admit a nice family of typical sets as shown by the the standard definition of a typical set recalled in Proposition 2. It corresponds to $H = H_q(p)$, $K_1 = K_2 = 1$, $\alpha(\epsilon) = \beta(\epsilon) = \epsilon$. We also notice that for a sequence of probability distributions admitting a nice family of typical sets the asymptotic entropy per symbol $\lim_{i\to\infty} \frac{H_q(p_i)}{n_i}$ exists and is equal to H:

Proposition 6. *For a nice family of typical sets, the quantity H appearing in its definition is necessarily unique and the entropy per symbol $\frac{H_q(p_i)}{n_i}$ has a limit when i tends to infinity:*

$$\lim_{i\to\infty} \frac{H_q(p_i)}{n_i} = H.$$

Proof. Fix $\epsilon > 0$ and a nonnegative integer i and $\delta = \delta(\epsilon, n_i)$. From the definition of a typical set we show that the entropy $H_q(p_i)$ is bounded by

$$-(1-\delta)\log_q(B_i(\epsilon)) \leq H_q(p_i) \leq -(1-\delta)\log_q(A_i(\epsilon)) + \delta n_i, \tag{7}$$

The lower bound just comes from the lower bound

$$H_q(p_i) \geq -\sum_{\boldsymbol{x}\in\mathcal{T}_\epsilon^{(n_i)}} \log_q(p_i(\boldsymbol{x}))p_i(\boldsymbol{x}) \geq -\log_q(B_i(\epsilon))p_i\left(\mathcal{T}_\epsilon^{(n_i)}\right) = -(1-\delta)\log_q(B_i(\epsilon)).$$

The upper bound follows from

$$\begin{aligned}
H_q(p_i) &= -\sum_{\boldsymbol{x}\in\mathcal{T}_\epsilon^{(n_i)}} \log_q(p_i(\boldsymbol{x}))p_i(\boldsymbol{x}) - \sum_{\boldsymbol{x}\in\mathbb{F}_q^{n_i}\setminus\mathcal{T}_\epsilon^{(n_i)}} \log_q(p_i(\boldsymbol{x}))p_i(\boldsymbol{x}) \\
&\leq -(1-\delta)\log_q(A_i(\epsilon)) + \delta \sum_{\boldsymbol{x}\in\mathbb{F}_q^{n_i}\setminus\mathcal{T}_\epsilon^{(n_i)}} -\log_q(p_i(\boldsymbol{x}))q_i(\boldsymbol{x}) \ \ (\text{where } q_i(\boldsymbol{x}) = \frac{p_i(\boldsymbol{x})}{\delta}) \\
&\leq -(1-\delta)\log_q(A_i(\epsilon)) + \delta\left(\sum_{\boldsymbol{x}\in\mathbb{F}_q^{n_i}\setminus\mathcal{T}_\epsilon^{(n_i)}} -\log_q(q_i(\boldsymbol{x}))q_i(\boldsymbol{x})\right) + \delta\log_q(\delta) \\
&\leq -(1-\delta)\log_q(A_i(\epsilon)) + \delta n_i.
\end{aligned}$$

The last inequality comes from the fact that q_i can be interpreted as a probability distribution over $\mathbb{F}_q^{n_i}\setminus\mathcal{T}_\epsilon^{(n_i)}$ and therefore $-\log_q(q_i(\boldsymbol{x}))q_i(\boldsymbol{x}) = H_q(q_i)$ and by using that the q-ary entropy of a random variable taking M different values is upper-bounded by $\log_q(M)$.

In the case of a nice typical set $\mathcal{T}_\epsilon^{(n_i)}$ (7) gives

$$(1-\mathrm{negl}(i))\left(H - \beta(\epsilon) - \frac{\log_q K_2}{n_i}\right) \leq \frac{H_q(p_i)}{n_i} \leq (1-\mathrm{negl}(i))\left(H + \alpha(\epsilon) - \frac{\log_q K_1}{n_i}\right) + \mathrm{negl}(i).$$

Letting ϵ going to zero shows that

$$(1-\mathrm{negl}(i))\left(H - \frac{\log_q K_2}{n_i}\right) \leq \frac{H_q(p_i)}{n_i} \leq (1-\mathrm{negl}(i))\left(H - \frac{\log_q K_1}{n_i}\right) + \mathrm{negl}(i).$$

Then letting i go to infinity shows that $\frac{H_q(p_i)}{n_i}$ has a limit equal to H. □

Another way to put this definition is to say that a sequence of probability distributions admitting a nice family of typical sets satisfies some form of the asymptotic equipartition property (AEP): almost all the mass of the probability distribution p_i concentrates on $\mathcal{T}_\epsilon^{(n_i)}$ for every $\epsilon > 0$ as i tends to infinity, these sets have size $\approx q^{n_i H}$ and the probabilities of the elements in this subset is close to $q^{n_i H}$ up to a multiplicative constant of the form $q^{n_i\gamma(\epsilon)}$ where $\gamma(\epsilon)$ tends to zero with ϵ.

3.2 Achievability Result

We are going to prove here the following theorem:

Theorem 1. *Let $q \geq 2$ be a prime power, $R \in (0,1)$ and $g : \mathbb{F}_q \to \mathbb{C}$ such that $\|g\|_2 = 1$. If $R < H_q(|\widehat{g}|^2)$ then there exists a quantum algorithm that solves $\mathrm{QDP}(q, n, \lfloor Rn \rfloor, g^{\otimes n})$ with probability $1 - \mathrm{negl}(n)$.*

As we would like to prove a similar result for the rank metric later on, we are going to prove a more general result here, which will enable us to prove both results.

Theorem 4. *Let $q \geq 2$ be a prime power. We assume that we have an increasing sequence of positive integers $(n_i)_{i\geq 1}$ and a family of functions $(f_i)_{i\geq 1}$ with $f_i : \mathbb{F}_q^{n_i} \to \mathbb{C}$ with $\|f_i\|_2 = 1$ such that the sequence $(p_i)_{i\geq 1}$ defined by $p_i \stackrel{\triangle}{=} |\widehat{f_i}|^2$ admits a nice family of typical sets. We let $H \stackrel{\triangle}{=} \lim_{i\to\infty} \frac{H_q(p_i)}{n_i}$. For every $R \in (0, H)$ the PGM solves $\mathrm{QDP}(q, n_i, \lfloor Rn_i \rfloor, f_i)$ with probability $1 - \mathrm{negl}(i)$.*

Fix $R \in (0, H)$ and let $k_i = \lfloor Rn_i \rfloor$. From Proposition 5 we know that the average probability $\overline{\mathrm{P_{PGM}}}$ over the randomness of codes (obtained by choosing a generator matrix $\mathbf{G}$ of $\mathscr{C}$ at random) is given by

$$\overline{\mathrm{P_{PGM}}} = \frac{1}{q^{k_i}} \mathbb{E}_{\mathbf{G}} \left(\sum_{\boldsymbol{s} \in \mathbb{F}_q^n / \mathscr{C}^\perp} Z_{\boldsymbol{s}} \right)^2 \quad \text{where} \tag{8}$$

$$Z_{\boldsymbol{s}} \stackrel{\triangle}{=} \sqrt{\sum_{\boldsymbol{e} \in \boldsymbol{s} + \mathscr{C}^\perp} \left| \widehat{f_i}(\boldsymbol{e}) \right|^2}. \tag{9}$$

We are going to show that $\overline{\mathrm{P_{PGM}}}$ is equal to $1 - \mathrm{negl}(i)$. This implies the theorem on the spot. To prove this, we will consider a random variable $\widetilde{Z_{\boldsymbol{s}}}(\epsilon)$ defined for every $\epsilon > 0$ which is related to $Z_{\boldsymbol{s}}$ which is such that $Z_{\boldsymbol{s}} \geq \widetilde{Z_{\boldsymbol{s}}}(\epsilon)$, but $Z_{\boldsymbol{s}} \approx \widetilde{Z_{\boldsymbol{s}}}(\epsilon)$ and $\widetilde{Z_{\boldsymbol{s}}}(\epsilon)$ has a much smaller variance than $Z_{\boldsymbol{s}}$, which allows us to prove concentration by using the second moment method. This random variable is defined by

$$\widetilde{Z_{\boldsymbol{s}}}(\epsilon) \stackrel{\triangle}{=} \sqrt{\sum_{\boldsymbol{y} \in \mathcal{T}_\epsilon^{(n_i)} \cap (\boldsymbol{s} + \mathscr{C}^\perp)} \left| \widehat{f_i}(\boldsymbol{y}) \right|^2}$$

Lemma 4. *For any $\epsilon > 0$, any positive integer i and $\boldsymbol{s} \in \mathbb{F}_q^{n_i}$, we have $\mathbb{E}_{\mathbf{G}} \left\{ \widetilde{Z_{\boldsymbol{s}}}(\epsilon)^2 \right\} \geq \frac{1-\delta(\epsilon,n_i)}{q^{k_i}}$*

Proof.

$$\begin{aligned}
\mathbb{E}_{\mathbf{G}} \left\{ \widetilde{Z_{\boldsymbol{s}}}(\epsilon)^2 \right\} &= \mathbb{E}_{\mathbf{G}} \left\{ \sum_{\boldsymbol{y} \in \mathcal{T}_\epsilon^{(n_i)} \cap (\boldsymbol{s} + \mathscr{C}^\perp)} p_i(\boldsymbol{y}) \right\} \\
&= \sum_{\boldsymbol{y} \in \mathcal{T}_\epsilon^{(n_i)}} p(\boldsymbol{y}) \Pr_{\mathbf{G}}(\boldsymbol{y} \in (\boldsymbol{s} + \mathscr{C}^\perp)) \\
&\geq \frac{1}{q^{k_i}} \sum_{\boldsymbol{y} \in \mathcal{T}_\epsilon^{(n_i)}} p(\boldsymbol{y}) \qquad (10) \\
&= \frac{1 - \delta(\epsilon, n_i)}{q^{k_i}} \text{ (by using the definition of } \delta(\epsilon, n_i)).
\end{aligned}$$

Inequality (10) follows from the fact that $\Pr_{\mathbf{G}}(\boldsymbol{y} \in (\boldsymbol{s} + \mathscr{C}^\perp)) = 1/q^{k_i}$ when $\boldsymbol{y} \neq \boldsymbol{s}$ and is equal to 1 if $\boldsymbol{y} = \boldsymbol{s}$. □

Remark 1. *Notice that the only reason why there is an inequality in the expression for $\mathbb{E} \left\{ \widetilde{Z_{\boldsymbol{s}}}(\epsilon)^2 \right\}$ is related to the possibility that $\boldsymbol{s}$ belongs to $\mathcal{T}_\epsilon^{(n_i)}$. If $\boldsymbol{s}$ does not belong to $\mathcal{T}_\epsilon^{(n_i)}$ we have an equality instead and in any case if we let $\mathscr{T} \stackrel{\triangle}{=} \mathcal{T}_\epsilon^{(n_i)} \setminus \{\boldsymbol{s}\}$ we have*

$$\mathbb{E} \left\{ \sum_{\boldsymbol{y} \in \mathscr{T} \cap (\boldsymbol{s} + \mathscr{C}^\perp)} \left| \widehat{f_i}(\boldsymbol{y}) \right|^2 \right\} = \begin{cases} \frac{1-\delta(\epsilon,n_i)}{q^{k_i}} & \text{if } \boldsymbol{s} \notin \mathcal{T}_\epsilon^{(n_i)} \\ \frac{1-\delta(\epsilon,n_i)-p(\boldsymbol{s})}{q^{k_i}} & \text{if } \boldsymbol{s} \in \mathcal{T}_\epsilon^{(n_i)} \end{cases} \tag{11}$$

The variance of $\widetilde{Z_{\boldsymbol{s}}}(\epsilon)^2$ on the other hand is upper-bounded by

Lemma 5. *For any $\epsilon > 0$, any positive integer i and $\boldsymbol{s} \in \mathbb{F}_q^{n_i}$, we have with the assumptions of Theorem 4*

$$\mathrm{Var}\left[\widetilde{Z_{\boldsymbol{s}}}(\epsilon)^2\right] \leq \frac{qK_2^2}{K_1} \cdot \frac{q^{-n_i(H-2\beta(\epsilon)-\alpha(\epsilon))}}{q^{k_i}},$$

where K_1, K_2 and $\alpha(\epsilon)$, $\beta(\epsilon)$ are respectively the constants and the functions appearing in Definition 5 of a nice family of typical sets.

Proof.

$$\begin{aligned}
\mathrm{Var}\left[\widetilde{Z_{\boldsymbol{s}}}(\epsilon)^2\right] &= \mathrm{Var}\left(\sum_{\boldsymbol{y}\in\mathcal{T}_\epsilon^{(n_i)}\cap(\boldsymbol{s}+\mathscr{C}^\perp)} p(\boldsymbol{y})\right) \\
&= \sum_{\boldsymbol{y}\in\mathcal{T}_\epsilon^{(n_i)}} (p(\boldsymbol{y}))^2 \,\mathrm{Var}(\mathbb{1}_{\boldsymbol{y}\in(\boldsymbol{s}+\mathscr{C}^\perp)}) + \sum_{\substack{\boldsymbol{x},\boldsymbol{y}\in\mathcal{T}_\epsilon^{(n_i)} \\ \boldsymbol{x}\neq\boldsymbol{y}}} p(\boldsymbol{x})p(\boldsymbol{y})\mathrm{Cov}(\mathbb{1}_{\boldsymbol{x}\in(\boldsymbol{s}+\mathscr{C}^\perp)}, \mathbb{1}_{\boldsymbol{y}\in(\boldsymbol{s}+\mathscr{C}^\perp)})
\end{aligned}$$

Here,

$$\begin{aligned}
\mathrm{Var}(\mathbb{1}_{\boldsymbol{y}\in(\boldsymbol{s}+\mathscr{C}^\perp)}) &= \frac{1}{q^{k_i}}\left(1-\frac{1}{q^{k_i}}\right) \quad \text{if } \boldsymbol{y}\neq\boldsymbol{s} \\
&= 0 \quad \text{else.}
\end{aligned}$$

When either $\boldsymbol{x}-\boldsymbol{s}$ or $\boldsymbol{y}-\boldsymbol{s}$ is zero:

$$\mathrm{Cov}(\mathbb{1}_{\boldsymbol{x}\in(\boldsymbol{s}+\mathscr{C}^\perp)}, \mathbb{1}_{\boldsymbol{y}\in(\boldsymbol{s}+\mathscr{C}^\perp)}) = 0.$$

When $\boldsymbol{x}-\boldsymbol{s}$ and $\boldsymbol{y}-\boldsymbol{s}$ are colinear (but none of them is zero),

$$\begin{aligned}
\mathrm{Cov}(\mathbb{1}_{\boldsymbol{x}\in(\boldsymbol{s}+\mathscr{C}^\perp)}, \mathbb{1}_{\boldsymbol{y}\in(\boldsymbol{s}+\mathscr{C}^\perp)}) &= \Pr(\boldsymbol{x}\in(\boldsymbol{s}+\mathscr{C}^\perp), \boldsymbol{y}\in(\boldsymbol{s}+\mathscr{C}^\perp)) - \Pr(\boldsymbol{x}\in(\boldsymbol{s}+\mathscr{C}^\perp))\Pr(\boldsymbol{y}\in(\boldsymbol{s}+\mathscr{C}^\perp)) \\
&= \frac{1}{q^{k_i}} - \frac{1}{q^{2k_i}} \\
&= \frac{1}{q^{k_i}}\left(1-\frac{1}{q^{k_i}}\right)
\end{aligned}$$

and when $\boldsymbol{x}$ and $\boldsymbol{y}$ are not colinear

$$\begin{aligned}
\mathrm{Cov}(\mathbb{1}_{\boldsymbol{x}\in(\boldsymbol{s}+\mathscr{C}^\perp)}, \mathbb{1}_{\boldsymbol{y}\in(\boldsymbol{s}+\mathscr{C}^\perp)}) &= \Pr(\boldsymbol{x}\in(\boldsymbol{s}+\mathscr{C}^\perp), \boldsymbol{y}\in(\boldsymbol{s}+\mathscr{C}^\perp)) - \Pr(\boldsymbol{x}\in(\boldsymbol{s}+\mathscr{C}^\perp))\Pr(\boldsymbol{y}\in(\boldsymbol{s}+\mathscr{C}^\perp)) \\
&= \frac{1}{q^{2k_i}} - \frac{1}{q^{2k_i}} \\
&= 0
\end{aligned}$$

So the variance can be upper-bounded by

$$\begin{aligned}
\mathrm{Var}\left[\widetilde{Z_{\boldsymbol{s}}}(\epsilon)^2\right] &\leq \frac{1}{q^{k_i}}\left(1-\frac{1}{q^{k_i}}\right)\left(\sum_{\boldsymbol{y}\in\mathcal{T}_\epsilon^{(n_i)}} (p(\boldsymbol{y}))^2 + \sum_{\substack{\boldsymbol{x},\boldsymbol{y}\in\mathcal{T}_\epsilon^{(n_i)} \\ \boldsymbol{x}\neq\boldsymbol{y} \text{ with } \boldsymbol{x} \text{ and } \boldsymbol{y} \text{ colinear}}} p(\boldsymbol{x})p(\boldsymbol{y})\right) \\
&= \frac{1}{q^{k_i}}\left(1-\frac{1}{q^{k_i}}\right) \sum_{\substack{\boldsymbol{x},\boldsymbol{y}\in\mathcal{T}_\epsilon^{(n_i)} \\ \boldsymbol{x} \text{ and } \boldsymbol{y} \text{ colinear}}} p(\boldsymbol{x})p(\boldsymbol{y}) \\
&= \frac{1}{q^{k_i}}\left(1-\frac{1}{q^{k_i}}\right) \sum_{\boldsymbol{y}\in\mathcal{T}_\epsilon^{(n_i)}} \left(\sum_{\substack{\boldsymbol{x}\in\mathcal{T}_\epsilon^{(n_i)} \\ \boldsymbol{x},\boldsymbol{y} \text{ colinear}}} p(\boldsymbol{x})p(\boldsymbol{y})\right) \\
&\leq \frac{|\mathcal{T}_\epsilon^{(n_i)}|}{q^{k_i}} qK_2^2 q^{-2n_i(H-\beta(\epsilon))} \\
&\leq \frac{qK_2^2}{K_1} \frac{q^{-n_i(H-2\beta(\epsilon)-\alpha(\epsilon))}}{q^{k_i}}, \qquad (12)
\end{aligned}$$

where in the last inequality we used Lemma 1 together with the definition of a nice family of typical sets

$$|\mathcal{T}_\epsilon^{(n_i)}| \leq \frac{1}{A_i(\epsilon)} \leq \frac{q^{n_i(H+\alpha(\epsilon))}}{K_1}.$$

□

These two lemmas yield the following concentration result

Lemma 6. *Let $\gamma \stackrel{\triangle}{=} H - R$. If ϵ is such that $2\beta(\epsilon) + \alpha(\epsilon) \leq \gamma/2$, then*

$$\Pr_{\mathbf{G}}\left[\widetilde{Z_{\boldsymbol{s}}}(\epsilon)^2 \geq \frac{1}{q^{k_i}}(1 - \delta(\epsilon, n_i) - q^{-\frac{\gamma n_i}{8}})\right] = 1 - \mathrm{negl}(i).$$

Proof. We use Bienaymé-Tchebychev's inequality for $X \stackrel{\triangle}{=} \widetilde{Z_{\boldsymbol{s}}}(\epsilon)^2$ and obtain for $\Delta > 0$:

$$\Pr\left(|X - \mathbb{E}(X)| \geq \Delta\right) \leq \frac{\mathrm{Var}(X)}{\Delta^2}. \tag{13}$$

To simplify expressions denote by K the constant $K \stackrel{\triangle}{=} \frac{qK_2^2}{K_1}$. Let us rewrite a little bit the upper bound we have on $\mathrm{Var}(X)$ derived from Lemma 5. We have

$$\begin{aligned}
\mathrm{Var}(X) &\leq \frac{qK_2^2}{K_1} \cdot \frac{q^{-n_i(H-2\beta(\epsilon)-\alpha(\epsilon))}}{q^{k_i}} \\
&= K\frac{q^{-n_i(H-R-2\beta(\epsilon)-\alpha(\epsilon))}}{q^{2k_i}} \\
&\leq K\frac{q^{-\frac{\gamma n_i}{2}}}{q^{2k_i}}.
\end{aligned} \tag{14}$$

We choose

$$\Delta = \frac{q^{-\frac{\gamma n_i}{8}}}{q^{k_i}}.$$

By using (14) we obtain

$$\begin{aligned}
\frac{\mathrm{Var}(X)}{\Delta^2} &\leq K \cdot \frac{q^{2k_i} \cdot q^{\frac{\gamma n_i}{4}} q^{-\frac{\gamma n_i}{2}}}{q^{2k_i}} \\
&= K \cdot q^{-\frac{\gamma n_i}{4}} \\
&= \mathrm{negl}(i).
\end{aligned}$$

We obtain by using this inequality in (13) and replacing Δ by its value:

$$\Pr_{\mathbf{G}}\left[\widetilde{Z_{\boldsymbol{s}}}(\epsilon)^2 \geq \frac{1}{q^{k_i}}(1 - \delta(\epsilon, n_i) - q^{-\frac{\gamma n_i}{8}})\right] = 1 - \mathrm{negl}(i).$$

□

A slight variation of this concentration result will be needed later on. It says that

Lemma 7. *Let $\gamma \stackrel{\triangle}{=} H - R$. If ϵ is such that $2\beta(\epsilon) + \alpha(\epsilon) \leq \gamma/2$, then if we let $\mathscr{T} \stackrel{\triangle}{=} \mathcal{T}_\epsilon^{(n_i)} \setminus \{\boldsymbol{s}\}$ we have*

$$\Pr_{\mathbf{G}}\left[\left|\sum_{\boldsymbol{y} \in \mathscr{T} \cap (\boldsymbol{s} + \mathscr{C}^\perp)} \left|\widehat{f}_i(\boldsymbol{y})\right|^2 - E\right| \leq \frac{q^{-\frac{\gamma n_i}{8}}}{q^{k_i}}\right] = 1 - \mathrm{negl}(i),$$

where E is the expected value of $\sum_{\boldsymbol{y} \in \mathscr{T} \cap (\boldsymbol{s} + \mathscr{C}^\perp)} \left|\widehat{f}_i(\boldsymbol{y})\right|^2$. It is given by the following expression

$$E = \begin{cases} \frac{1-\delta(\epsilon,n_i)}{q^{k_i}} & \text{if } \boldsymbol{s} \notin \mathcal{T}_\epsilon^{(n_i)} \\ \frac{1-\delta(\epsilon,n_i)-p(\boldsymbol{s})}{q^{k_i}} & \text{if } \boldsymbol{s} \in \mathcal{T}_\epsilon^{(n_i)} \end{cases}$$

Proof. We use for the expected value of $\sum_{\boldsymbol{y}\in\mathscr{T}\cap(\boldsymbol{s}+\mathscr{C}^\perp)}\left|\widehat{f}_i(\boldsymbol{y})\right|^2$ its expression given in (11). The upper-bound on the variance of $\sum_{\boldsymbol{y}\in\mathscr{T}\cap(\boldsymbol{s}+\mathscr{C}^\perp)}\left|\widehat{f}_i(\boldsymbol{y})\right|^2$ is clearly the same as the upper-bound on the variance of $\widetilde{Z_{\boldsymbol{s}}}(\epsilon)^2$ given in Lemma 5. This yields immediately Lemma 7 similarly to what was done for obtaining Lemma 6 in the lower bound of expectation of the random variable $\sum_{\boldsymbol{y}\in\mathscr{T}\cap(\boldsymbol{s}+\mathscr{C}^\perp)}\left|\widehat{f}_i(\boldsymbol{y})\right|^2$ to which we apply Bienaymé-Tchebychev's inequality. □

We are ready now to prove Theorem 4.

Proof of Theorem 4. Since $Z_{\boldsymbol{s}}^2 \geq \widetilde{Z_{\boldsymbol{s}}}(\epsilon)^2$ for any $\epsilon > 0$, we obtain by using Lemma 6 by choosing ϵ such that $2\beta(\epsilon) + \alpha(\epsilon) \leq \gamma/2$,

$$\Pr_{\mathbf{G}}\left[Z_{\boldsymbol{s}}^2 \geq \frac{1-\delta(\epsilon,n_i)-q^{-\frac{\gamma n_i}{8}}}{q^{k_i}}\right] = 1-\mathrm{negl}(i)$$

which gives $\Pr_{\mathbf{G}}\left[Z_{\boldsymbol{s}} \geq \sqrt{\frac{1}{q^{k_i}}(1-\delta(\epsilon,n_i)-q^{-\frac{\gamma n_i}{8}})}\right] = 1-\mathrm{negl}(i)$. Since $\delta(\epsilon,n_i) = \mathrm{negl}(i)$, this implies for all $\boldsymbol{s} \in \mathbb{F}_q^{n_i}/\mathscr{C}^\perp$:

$$\mathbb{E}_{\mathbf{G}}\left[Z_{\boldsymbol{s}}\right] \geq \frac{1}{\sqrt{q^{k_i}}}(1-\mathrm{negl}(i))$$

Then, $\mathbb{E}_{\mathbf{G}}\left[\sum_{\boldsymbol{s}\in\mathbb{F}_q^{n_i}/\mathscr{C}^\perp} Z_{\boldsymbol{s}}\right] \geq \sqrt{q^{k_i}}\left(1-\mathrm{negl}(i)\right).$

We use Jensen's inequality $\mathbb{E}_{\mathbf{G}}(X^2) \geq (\mathbb{E}_{\mathbf{G}}(X))^2$ in order to obtain:

$$\overline{\mathrm{P_{PGM}}} = \frac{1}{q^{k_i}}\left(\mathbb{E}_{\mathbf{G}}\left\{\sum_{\boldsymbol{s}\in\mathbb{F}_q^{n_i}/\mathscr{C}^\perp} Z_{\boldsymbol{s}}\right\}^2\right) \geq \frac{1}{q^{k_i}}\left(\mathbb{E}\left\{\sum_{\boldsymbol{s}\in\mathbb{F}_q^{n_i}/\mathscr{C}^\perp} Z_{\boldsymbol{s}}\right\}\right)^2 \geq 1-\mathrm{negl}(i).$$

□

With these results at hand we can now prove Theorem 1.

Proof of Theorem 1. In this case $f_i = g^{\otimes i}$ and we take $n_i = i$, $H = H_q(|\widehat{g}|^2)$, $A_i(\epsilon) = q^{-i(H+\epsilon)}$, $B_i(\epsilon) = q^{-i(H-\epsilon)}$. Proposition 2 shows that $\delta(\epsilon,n) = \mathrm{negl}(n)$ for all $\epsilon > 0$. This shows that we can apply Theorem 4 to this case, from which we derive immediately Theorem 1. □

3.3 Strong converse

We are now going to prove that whatever code we take in the QDP problem, be it random as in the statement of the QDP problem or chosen carefully and be it even non-linear, then we cannot solve the QDP problem with non negligible probability as soon as the rate is just slightly above $H_q(|\widehat{g}|^2)$. More precisely, we are going to prove

Theorem 2. *Let $g : \mathbb{F}_q \to \mathbb{C}$ with $\|g\|_2 = 1$. Let R be a fixed constant in $(0,1)$ such that $R > H_q(|\widehat{g}|^2)$. For any quantum algorithm and any code $\mathscr{C}$ with q^{Rn} codewords the probability p_{succ} to solve $\mathrm{QDP}(\mathscr{C}, g^{\otimes n})$ satisfies*

$$p_{succ} \leq q^{n(H_q(|\widehat{g}|^2)-R+\frac{1}{n^{1/3}})} + \mathrm{negl}(n) = \mathrm{negl}(n).$$

As for achievability results, we would like to prove a similar result for the rank metric later on. Then, we are going to prove a more general result here, which will enable us to prove both results :

Theorem 5. *We assume that we have an increasing sequence of positive integers $(n_i)_{i\geq 1}$ and a family of functions $(f_i)_{i\geq 1}$ with $f_i : \mathbb{F}_q^{n_i} \to \mathbb{C}$ with $\|f_i\|_2 = 1$ such that the sequence $(p_i)_{i\geq 1}$ defined by $p_i \stackrel{\triangle}{=} |\widehat{f}_i|^2$ admits a nice family of typical sets in the strong sense.*

Let R be a fixed constant in $(0,1)$ such that $R > H$. For any quantum algorithm and any code $\mathscr{C}_i \in \mathbb{F}_q^{n_i}$ with at least q^{Rn_i} codewords the probability p_{succ} to solve $\mathrm{QDP}(\mathscr{C}_i, f_i)$ satisfies

$$p_{succ} = \mathrm{negl}(i).$$

The proof of this statement will rely on several statements. The first one uses the fact that for any $\boldsymbol{c}$ in $\mathscr{C}_i$ the state $|\widehat{\psi_{\boldsymbol{c}}}\rangle$ which is equal to $\sum_{\boldsymbol{y}\in\mathbb{F}_q^{n_i}} \widehat{f}_i(\boldsymbol{y})\chi_{\boldsymbol{c}}(\boldsymbol{y})|\boldsymbol{y}\rangle$ by Lemma 3 is extremely close to the state $|\widetilde{\psi}_{\boldsymbol{c}}\rangle$ which is obtained from $|\widehat{\psi_{\boldsymbol{c}}}\rangle$ by keeping in the previous sum only those corresponding to a $\boldsymbol{y}$ that are in the typical set $\mathcal{T}_{\epsilon(i)}^{(n_i)}$ associated to $|\widehat{f}_i|^2$.

Lemma 8. *Let i be a positive integer. For $\boldsymbol{c} \in \mathscr{C}$, we let $|\widetilde{\psi}_{\boldsymbol{c}}\rangle \stackrel{\triangle}{=} \frac{1}{\sqrt{1-\delta(\epsilon(i),n_i)}} \sum_{\boldsymbol{y}\in\mathcal{T}_{\epsilon(i)}^{(n_i)}} \widehat{f}_i(\boldsymbol{y})\chi_{\boldsymbol{c}}(\boldsymbol{y})|\boldsymbol{y}\rangle$, where $\delta(\epsilon(i),n_i) \stackrel{\triangle}{=} 1 - \sum_{\boldsymbol{y}\in\mathcal{T}_{\epsilon(i)}^{(n_i)}} |\widehat{f}_i(\boldsymbol{y})|^2$. These states are normalized and we have $\langle\widehat{\psi}_{\boldsymbol{c}}|\widetilde{\psi}_{\boldsymbol{c}}\rangle = \sqrt{1-\delta(\epsilon(i),n_i)}$.*

Proof. This is a direct consequence of

$$\||\widetilde{\psi}_{\boldsymbol{c}}\rangle\|_2^2 = \frac{1}{1-\delta(\epsilon(i),n_i)} \sum_{\boldsymbol{y}\in\mathcal{T}_{\epsilon(i)}^{(n_i)}} |\widehat{f}_i(\boldsymbol{y})|^2 = 1.$$

The second statement is proved in a similar fashion

$$\langle\widehat{\psi}_{\boldsymbol{c}}|\widetilde{\psi}_{\boldsymbol{c}}\rangle = \frac{\sum_{\boldsymbol{y}\in\mathcal{T}_{\epsilon(i)}^{(n_i)}} |\widehat{f}(\boldsymbol{y})^2|}{\sqrt{1-\delta(\epsilon(i),n_i)}} = \frac{1-\delta(\epsilon(i),n_i)}{\sqrt{1-\delta(\epsilon(i),n_i)}} = \sqrt{1-\delta(\epsilon(i),n_i)}.$$

□

The point is that these states $|\widetilde{\psi}_{\boldsymbol{c}}\rangle$ live in a much smaller space which is of dimension the size of the typical set. There is a general upper bound on how we can distinguish such states which is given by

Lemma 9. *Assume that we have N states living in a space of dimension K. The average probability of distinguishing them is upper bounded by $\frac{K}{N}$.*

Proof. Let $|\phi_1\rangle, \cdots, |\phi_N\rangle$ be those states and let $M_1, \cdots, M_N$ be a POVM distinguishing those states. The average probability of success p_{succ} for this POVM is given by

$$p_{\mathrm{succ}} = \frac{1}{N} \sum_{i=1}^{N} \mathrm{tr}\left(M_i |\phi_i\rangle\langle\phi_i|\right).$$

Now let V be the space of dimension K containing the $|\phi_i\rangle$'s and let Π_V be the orthogonal projector onto V, that is $\Pi_V \stackrel{\triangle}{=} \sum_{i=1}^{K} |\boldsymbol{e}_i\rangle\langle\boldsymbol{e}_i|$ where $\{|\boldsymbol{e}_1\rangle, \cdots, |\boldsymbol{e}_K\rangle\}$ is an orthonormal basis of V. Since each $|\phi_i\rangle$ has support in V, all these states satisfy

$$|\phi_i\rangle\langle\phi_i| \preceq \Pi_V. \tag{15}$$

By using Claim 3 with (15) we obtain

$$\begin{aligned} p_{\text{succ}} &= \frac{1}{N}\sum_{i=1}^{N} \text{tr}\left(M_i|\phi_i\rangle\langle\phi_i|\right) \\ &\leq \frac{1}{N}\sum_{i=1}^{N} \text{tr}\left(M_i\Pi_V\right) \text{ (by Claim 3 and (15))} \\ &= \frac{1}{N}\,\text{tr}\left(\sum_{i=1}^{N} M_i\Pi_V\right) \\ &= \frac{1}{N}\,\text{tr}(\Pi_V) \text{ (since } \sum_{i=1}^{N} M_i = I) \\ &= \frac{K}{N} \text{ (since } \text{tr}(\Pi_V) = K). \end{aligned}$$

□

The last ingredient tells us that we can not hope to have a much better success probability if instead of feeding our distinguisher with the $|\widetilde{\psi}_{\boldsymbol{c}}\rangle$ we use the $|\psi_{\boldsymbol{c}}\rangle$ since we have in general

Lemma 10. *Let $|\phi_1\rangle, \cdots, |\phi_N\rangle$ and $|\widetilde{\phi}_1\rangle, \cdots, |\widetilde{\phi}_N\rangle$ be $2N$ quantum states all living in the same Hilbert space. Let*

$$\delta \stackrel{\triangle}{=} \max_{i\in\{1,\cdots,N\}} \sqrt{1-|\langle\phi_i|\widetilde{\phi}_i\rangle|^2} = \max_{i\in\{1,\cdots,N\}} D\left(|\phi_i\rangle\langle\phi_i|, |\widetilde{\phi}_i\rangle\langle\widetilde{\phi}_i|\right).$$

Let p be the average success probability of distinguishing the $|\phi_i\rangle$'s by a certain POVM $\{M_1, \cdots, M_N\}$. Let $\widetilde{p}$ be the average success probability of distinguishing the $|\widetilde{\phi}_i\rangle$'s by the same POVM. We have

$$p \leq \widetilde{p} + \delta.$$

Proof. We write

$$\begin{aligned} p &= \frac{1}{N}\sum_{i=1}^{N} \text{tr}\left(M_i|\phi_i\rangle\langle\phi_i|\right) \\ &\leq \frac{1}{N}\sum_{i=1}^{N} \text{tr}\left(M_i|\widetilde{\phi}_i\rangle\langle\widetilde{\phi}_i|\right) + \delta \qquad \text{(by using Claim 2)} \\ &= \widetilde{p} + \delta. \end{aligned}$$

□

We are ready now to prove Theorem 5.

Proof of Theorem 5. We recall that for any positive integer i

1. $\sum_{\boldsymbol{y}\in\mathcal{T}_{\epsilon(i)}^{(n_i)}} |\widehat{f}_i(\boldsymbol{y})|^2 = \sum_{\boldsymbol{y}\in\mathcal{T}_{\epsilon(i)}^{(n_i)}} |\widehat{f}_i(\boldsymbol{y})\chi_{\boldsymbol{c}}(\boldsymbol{y})|^2 = 1 - \delta(\epsilon(i), n_i)$ with $\delta(\epsilon(i), n_i) = \text{negl}(i)$.
2. $|\mathcal{T}_{\epsilon(i)}^{(n_i)}| \leq q^{n_i(H+o(1))}$.

For the last point we used Lemma 1 which says that $|\mathcal{T}_{\epsilon(i)}^{(n_i)}| \leq \frac{1}{A_i(\epsilon(i))}$ and the fact that $A_i(\epsilon(i)) = q^{-n_i(H+o(1))}$. Now, consider a quantum POVM $\{M_{\boldsymbol{c}}\}_{\boldsymbol{c}\in\mathscr{C}_i}$. Let p_{succ} be the average probability to solve the QDP problem using this measurement whereas $\widetilde{p}$ is the average

probability of distinguishing $\boldsymbol{c}$ with the same POVM when fed instead of $|\widehat{\psi}_{\boldsymbol{c}}\rangle$ the state $|\widetilde{\psi}_{\boldsymbol{c}}\rangle$. By using Lemma 10 we deduce that

$$p_{\text{succ}} \leq \widetilde{p} + \sqrt{\delta(\epsilon(i), n_i)}, \tag{16}$$

since $\langle\widehat{\psi}_{\boldsymbol{c}}|\widetilde{\psi}_{\boldsymbol{c}}\rangle = \sqrt{1-\delta(\epsilon(i),n_i)}$ for every $\boldsymbol{c} \in \mathscr{C}$ by Lemma 8. By applying Lemma 9 to the q^{Rn_i} states $|\widetilde{\phi}_{\boldsymbol{c}}\rangle$ for $\boldsymbol{c} \in \mathscr{C}_i$ we obtain

$$\begin{aligned}
\widetilde{p} &\leq \frac{\left|\mathcal{T}^{(n_i)}_{\epsilon(i)}\right|}{q^{Rn_i}} \\
&\leq \frac{q^{n_i(H+o(1))}}{q^{Rn_i}} \\
&\leq q^{n_i(H-R+o(1))} \\
&= \text{negl}(i).
\end{aligned}$$

By using this upper-bound in (16) we obtain

$$p \leq \text{negl}(i) + \sqrt{\delta(\epsilon(i), n_i)} = \text{negl}(i),$$

where we used Proposition 2 in the last equality to write that $\delta(\epsilon(i), n_i) = \text{negl}(i)$. □

Theorem 2 is a direct corollary of this theorem. Indeed in this case we take

1. $n_i = i$ and $f_i = g^{\otimes i}$.
2. $H = H_q(|\widehat{g}|^2)$
3. The typical set is defined with the bounding functions $A_i(\epsilon) = q^{-i(H+\epsilon)}$ and $B_i(\epsilon) = q^{-i(H-\epsilon)}$.

We then choose ϵ as a function of i of the form $\epsilon(i) = \frac{1}{i^{1/3}}$ and by applying Proposition 2 we verify that the defect function $\delta(\epsilon(i), n_i)$ satisfies $\delta(\epsilon(i), n_i) = \text{negl}(i)$.

3.4 Application to the rank metric

In all this subsection, the codelength n is of the form $n = a \cdot b$ for some integers a and b and we view elements of $\mathbb{F}_q^n$ as matrices in $\mathbb{F}_q^{a\times b}$ as explained in §2.1. We will assume that $a \geq b$. If this is not the case, we will just swap the role of rows and columns in the explanations that follow. Unlike the error model we considered in the previous subsections, the rank metric case corresponds in a natural way to an error model which is not memoryless. There are several ways of defining an error distribution on $\mathbb{F}_q^n$ which would make sense for cryptographic applications. A uniform distribution over all errors of $\mathbb{F}_q^n$ of a given rank weight t would for instance be relevant here. This would complicate the computations of the Fourier transform of the error distribution. A simple way to approximate this distribution is given in [DRT24, §III.E]. It corresponds to an error amplitude distribution $f_t^{a,b}$ defined by

$$\sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f_t^{a,b}(\boldsymbol{e})|\boldsymbol{e}\rangle \triangleq \frac{1}{\sqrt{Z}} \sum_{\substack{V:V\leq\mathbb{F}_q^b \\ \dim V = t}} |\pi_V\rangle \quad \text{where} \tag{17}$$

$$|\pi_V\rangle \triangleq \left(\frac{1}{\sqrt{q^{\dim V}}} \sum_{\boldsymbol{v}\in V} |v\rangle\right)^{\otimes a}, \tag{18}$$

and Z is a normalizing constant which makes the state appearing on the right of (17) to be a valid quantum state *i.e.* a state of norm 1. The state $|\pi_V\rangle$ can be seen as a uniform superposition of vectors $\boldsymbol{x} \in \mathbb{F}_q^n$ whose matrix form $\text{Mat}(\boldsymbol{y})$ gives all matrices formed by rows

taken from the vector subspace V. As shown in [DRT24, Lemma 3.21] $f(\boldsymbol{e})$ is radial, meaning that it only depends on the rank weight of $\boldsymbol{e}$ and we have

$$f_t^{a,b}(\boldsymbol{e}) = \begin{cases} \dfrac{\left[\begin{smallmatrix} b-|\boldsymbol{e}|_{\mathrm{rk}} \\ t-|\boldsymbol{e}|_{\mathrm{rk}} \end{smallmatrix}\right]_q}{\sqrt{q^{at}Z}} & \text{if } |\boldsymbol{e}|_{\mathrm{rk}} \leqslant t \\ 0 & \text{else} \end{cases}$$

with $\left[\begin{smallmatrix} b \\ t \end{smallmatrix}\right]_q = \begin{cases} \prod_{i=0}^{t-1} \frac{q^b-q^i}{q^t-q^i} & \text{if } t \leqslant b \\ 0 & \text{else} \end{cases}$ being the Gaussian binomial coefficient and the normalizing constant is shown to satisfy (see [DRT24, Lemma 3.22]) $Z = \Theta\left(\left[\begin{smallmatrix} b \\ t \end{smallmatrix}\right]_q\right)$. We will generally drop the dependence in a, b which will be clear from the context and simply write $f_t(e)$.
Moreover the Gaussian binomial coefficient and the size S_t of the sphere of radius t for the rank metric over $\mathbb{F}_q^n$ (that is the vectors $\boldsymbol{y}$ in $\mathbb{F}_q^n$ such that $\mathrm{Mat}(\boldsymbol{y})$ is of rank t) satisfy

Fact 1. *[DRT24, Section II.B]* $\left[\begin{smallmatrix} b \\ t \end{smallmatrix}\right]_q = \Theta(q^{t(b-t)})$, $S_t = \Theta\left(q^{t(a+b-t)}\right)$, $|f_t(\boldsymbol{e})|^2 = \Theta\left(q^{(b-t)(t-2u)-at}\right)$. *The constants hidden inside Θ can be taken to be absolute constants that do not depend on a, b or t.*

It follows directly from [DRT24, Prop. 3.23] that

$$\widehat{f_t} = f_{b-t}. \tag{19}$$

The function f_t is radial, meaning that the value $f_t(\boldsymbol{e})$ only depends on the rank weight $|\boldsymbol{e}|_{\mathrm{rk}}$ of $\boldsymbol{e}$. We will be interested in the dual distribution $p \stackrel{\triangle}{=} \left|\widehat{f_t}\right|^2 = |f_{b-t}|^2$. This probability distribution concentrates around the rank weight $n-t$ as shown in Appendix B.2 of [DRT24]

Fact 2. *For any u in $\{0, \cdots, n-t\}$ we have*

$$\tilde{p}(u) = \Theta\left(q^{-(a+t)(b-t)+2tv}\right) \tag{20}$$

$$\sum_{\boldsymbol{e}:|\boldsymbol{e}|_{rk}=u} p(\boldsymbol{e}) = \tilde{p}(u)S_u = \Theta\left(q^{v(b-a)-v^2}\right) \tag{21}$$

where $v \stackrel{\triangle}{=} b-t-u$ and $\tilde{p}(|\boldsymbol{e}|_{rk}) \stackrel{\triangle}{=} p(\boldsymbol{e})$ for any $\boldsymbol{e} \in \mathbb{F}_q^n$, where the constant hidden in Θ notation is an absolute constant that does not depend on a, b or q.

The tails of p are readily seen to verify

Lemma 11. *Let $\epsilon > 0$.*

$$\sum_{u=0}^{\lfloor (1-\epsilon)b \rfloor - t} S_u \tilde{p}(u) = O\left(q^{-b^2\varepsilon^2}\right).$$

Proof. We have

$$\begin{aligned} \sum_{u=0}^{\lfloor (1-\epsilon)b \rfloor - t} S_u \tilde{p}(u) &= \Theta\left(\sum_{v=\lceil b\epsilon \rceil}^{b-t} q^{v(b-a)-v^2}\right) \quad \text{(by Fact 2)} \\ &= O\left(q^{-b^2\varepsilon^2} \sum_{i=0}^{\infty} q^{-i}\right) \\ &= O\left(q^{-b^2\varepsilon^2}\right). \end{aligned}$$

□

The typical set is defined as

Definition and properties of the set $\mathcal{T}_\epsilon^{(n)}$. We let

$$\mathcal{T}_\epsilon^{(n)} \stackrel{\triangle}{=} \left\{ \boldsymbol{e} \in \mathbb{F}_q^n : b(1-\epsilon) - t \leq |\boldsymbol{e}|_{\mathrm{rk}} \leq b - t \right\}. \tag{22}$$

Here, $p(\boldsymbol{e})$ is clearly a decreasing function of the rank weight $|\boldsymbol{e}|_{\mathrm{rk}}$. This shows that this is indeed a typical set as specified in Definition 5. From (20) we infer that the bounding functions $A(\epsilon)$ and $B(\epsilon)$ satisfy

$$A(\epsilon) = \Theta\left(q^{-nH}\right) \tag{23}$$

$$B(\epsilon) = \Theta\left(q^{-n\left(H - \frac{2t}{a}\epsilon\right)}\right) \quad \text{where} \tag{24}$$

$$H \stackrel{\triangle}{=} (1 + t/a)(1 - t/b). \tag{25}$$

From Lemma 11 we deduce that

$$\delta(\epsilon, n) = O\left(q^{-b^2\epsilon^2}\right). \tag{26}$$

Remark 2. *The quantity H can really be viewed as a very good approximation of the entropy per symbol for the distribution p. It is straightforward to use Fact 2 to show that the entropy per symbol $\tilde{H} \stackrel{\triangle}{=} \frac{H_q(p)}{n}$ is of the form $\tilde{H} = H + O(1/a)$.*

$$\begin{aligned}
H_q(p) &= -\sum_{u=0}^{b-t} \log_q(\tilde{p}(u))\tilde{p}(u)S_u \\
&= -\sum_{v=0}^{b-t} \left(-(a+t)(b-t) + 2tv + \Theta(1)\right)\tilde{p}(b-t-v)S_{b-t-v} \quad \textit{(by (20))} \\
&= (a+t)(b-t) - \Theta(1) - 2t\sum_{v=0}^{b-t} v\tilde{p}(b-t-v)S_{b-t-v} \quad \textit{(since } \sum_{v=0}^{b-t} \tilde{p}(b-t-v)S_{b-t-v} = 1\textit{)} \\
&= (a+t)(b-t) + O(b).
\end{aligned}$$

The last point follows from $\sum_{v=0}^{b-t} v\tilde{p}(b-t-v)S_{b-t-v} = O(1)$ which can be derived by using Fact 2 similarly to what is done in Lemma 11.

From there, we give the **achievability result**:

Theorem 6. *Let $q \geq 2$ be a prime power. We assume that there exist three increasing sequences $(a_i)_{i\geq 1}$, $(b_i)_{i\geq 1}$, $(t_i)_{i\geq 1}$ such that $a_i \geq b_i$ for any positive integer i and the sequence $(1 + t_i/a_i)(1 - t_i/b_i)$ tends to a constant $H \in (0, 1)$.*

For every $R \in (0, H)$ the PGM solves $\mathrm{QDP}(q, a_i b_i, \lfloor R a_i b_i \rfloor, f_{t_i}^{a_i,b_i})$ with probability $1 - \mathrm{negl}(i)$.

Proof. From (26) it appears that $\delta(\epsilon, n_i) = \mathrm{negl}(i)$ for any $\epsilon > 0$. Using (24), we can take $\alpha(\epsilon) = 0$, $\beta(\epsilon) = 2\epsilon$ and the assumptions of Theorem 4 are therefore verified. □

Non-achievability follows also immediately from Theorem 5

Theorem 7. *Let $q \geqslant 2$ be a prime power. We assume that there exist three increasing sequences $(a_i)_{i\geq 1}$, $(b_i)_{i\geq 1}$, $(t_i)_{i\geq 1}$ such that $a_i \geq b_i$ for any positive integer i and the sequence $(1 + t_i/a_i)(1 - t_i/b_i)$ tends to a constant $H \in (0, 1)$. Let R be a constant in $(H, 1]$. Then for any quantum algorithm and any family of codes $(\mathscr{C}_i)_{i\geq 1}$ where $\mathscr{C}_i \in \mathbb{F}_q^{n_i}$ contains at least q^{Rn_i} codewords with $n_i \stackrel{\triangle}{=} a_i \cdot b_i$, the probability p_{succ} to solve $\mathrm{QDP}(\mathscr{C}_i, f_{t_i}^{a_i,b_i})$ satisfies*

$$p_{succ} = \mathrm{negl}(i).$$

Proof. We define the gap function ϵ as $\epsilon(i) = \frac{1}{b_i}$. From (26) it appears that $\delta(\epsilon(i), n_i) = \mathrm{negl}(i)$. The fact that $A(\epsilon(i)) = q^{-n_i(H+o(1))}$ as i tends to infinity follows from (23) and the fact that $(1 + t_i/a_i)(1 - t_i/b_i)$ tends to H.

□

4 Sampling Dual Codewords

The algorithm for finding short codewords given in [CT24] applies strictly speaking to the Bernoulli noise and the Hamming metric. However, the algorithm detailed in the full version of the paper [CT23, §6.1] can be used with any noise distribution. It is based on a slight variation of Regev's quantum algorithm which allowed to reduce the SIS problem to LWE. Roughly speaking, Regev's approach [Reg05] revisited by Chen, Liu and Zhandry in [CLZ22] consists in constructing a superposition of dual codewords $\sum_{\boldsymbol{c}\in\mathscr{C}^\perp} \hat{f}(\boldsymbol{c})|\boldsymbol{c}\rangle$ by solving the quantum decoding problem for $\mathscr{C}$ and noise amplitude function f. More precisely, it consists in first preparing the uniform superposition of codewords tensored with a superposition of errors $\frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c}\in\mathscr{C}}|\boldsymbol{c}\rangle\otimes\sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{e}\rangle$ and adding the first register to the second one to create the entangled state $\frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c}\in\mathscr{C},\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}\rangle|\boldsymbol{c}+\boldsymbol{e}\rangle$. This state can be rewritten as $\frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c}\in\mathscr{C}}|\boldsymbol{c}\rangle|\psi_{\boldsymbol{c}}\rangle$ with $|\psi_{\boldsymbol{c}}\rangle \stackrel{\triangle}{=} \sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}+\boldsymbol{e}\rangle$. If we solve the quantum decoding problem for $\mathscr{C}$ and f, we can in principle recover $\boldsymbol{c}$ and erase the first register to get $\sum_{\boldsymbol{c}\in\mathscr{C},\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}+\boldsymbol{e}\rangle$ in the second register. This state $\sum_{\boldsymbol{y}\in\mathbb{F}_q^n}\alpha_{\boldsymbol{y}}|\boldsymbol{y}\rangle$ is periodic with periods the codewords of $\mathscr{C}$. We mean here that $\alpha_{\boldsymbol{y}} = \alpha_{\boldsymbol{y}+\boldsymbol{c}}$ for any $\boldsymbol{y}\in\mathbb{F}_q^n$ and any $\boldsymbol{c}\in\mathscr{C}$. When applying the quantum Fourier transform to it, we obtain precisely the state we are after, namely up to a normalizing factor $\sum_{\boldsymbol{c}\in\mathscr{C}^\perp}\hat{f}(\boldsymbol{c})|\boldsymbol{c}\rangle$. The whole process is summarized by the following sequence of operations

Algorithm 4 : Regev's algorithm for sampling dual codewords.

Initial state preparation:		$\|\phi_0\rangle$	$=$	$\frac{1}{\sqrt{\|\mathscr{C}\|}}\sum_{\boldsymbol{c}\in\mathscr{C}}\sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})\|\boldsymbol{c}\rangle\|\boldsymbol{e}\rangle$
adding $\boldsymbol{c}$ to $\boldsymbol{e}$:	$\mapsto$	$\|\phi_1\rangle$	$=$	$\frac{1}{\sqrt{\|\mathscr{C}\|}}\sum_{\boldsymbol{c}\in\mathscr{C}}\sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})\|\boldsymbol{c}\rangle\|\boldsymbol{c}+\boldsymbol{e}\rangle$
			$=$	$\frac{1}{\sqrt{\|\mathscr{C}\|}}\sum_{\boldsymbol{c}\in\mathscr{C}}\|\boldsymbol{c}\rangle\|\psi_{\boldsymbol{c}}\rangle$
erase the 1st register by decoding $\|\psi_{\boldsymbol{c}}\rangle$:	$\mapsto$	$\|\phi_3\rangle$	$=$	$\frac{1}{\sqrt{\|\mathscr{C}\|}}\sum_{\boldsymbol{c}\in\mathscr{C}}\|\boldsymbol{0}\rangle\|\psi_{\boldsymbol{c}}\rangle$
			$=$	$\|\boldsymbol{0}\rangle\frac{1}{\sqrt{\|\mathscr{C}\|}}\sum_{\boldsymbol{c}\in\mathscr{C},\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})\|\boldsymbol{c}+\boldsymbol{e}\rangle$
apply the QFT:	$\mapsto$	$\|\phi_4\rangle$	$=$	$\frac{1}{\sqrt{Z}}\sum_{\boldsymbol{c}\in\mathscr{C}^\perp}\hat{f}(\boldsymbol{c})\|\boldsymbol{c}\rangle$
measuring the whole state:	$\mapsto$			$\|\boldsymbol{y}\rangle$ with $\boldsymbol{y}\in\mathscr{C}^\perp$.

When we measure the resulting state we sample dual codewords with respect to a probability distribution proportional to $|\hat{f}|^2$ and supported in $\mathscr{C}^\perp$. This is the general picture, however using a decoding algorithm which is not perfect, meaning that from the knowledge of $|\psi_{\boldsymbol{c}}\rangle$ we do not always recover $\boldsymbol{c}$, adds some difficulties because we do not recover exactly $\frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c}\in\mathscr{C},\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}+\boldsymbol{e}\rangle$ after the decoding step. The case of the Bernoulli noise was treated in [CT24], this allows to obtain codewords of small Hamming weight. At the limit where the quantum decoding problem can still be solved, we get codewords at the Gilbert Varshamov distance, in other words typically minimum weight codewords. We are going to generalize this result to more general noise amplitude distributions f and we will show that

- Regev's approach used in conjunction with the PGM for solving the quantum decoding problem allows to sample dual codewords according to a distribution supported on $\mathscr{C}^\perp \setminus$

$\{\mathbf{0}\}$ and proportional to $|\hat{f}|^2$ on this support in the tractability regime;

- at the limit of the tractability regime, we essentially sample the most likely non zero dual codewords according to the aforementioned a probability distribution.

4.1 The sampling algorithm

In the whole section $\mathscr{C}$ denotes a linear code of length n over $\mathbb{F}_q$ whose generator matrix $\mathbf{G}$ is drawn uniformly at random in $\mathbb{F}_q^{k\times n}$, meaning that the code is of dimension k with probability $1-o(1)$ for a fixed rate $R=\frac{k}{n}$ bounded away from 1. Roughly speaking, in order to avoid destroying the second register carrying $|\psi_{\boldsymbol{c}}\rangle$ and still measuring most of the time $\boldsymbol{c}$, we apply the PGM coherently on the second register carrying $|\boldsymbol{c}\rangle$ and write the output on the third register. The value of the third register is then subtracted to the first register and we reverse at that point the PGM between the second and third register. As we have seen in the previous section it is more convenient to express the PGM in this case in the Fourier basis. We apply therefore prior to applying the PGM the quantum Fourier transform on the second register carrying $|\psi_{\boldsymbol{c}}\rangle$. This is basically what is done in Algorithm of Section 6.3 in [CT23], which we recall now

Algorithm 4.1 : Algorithm of the quantum reduction with PGM.

Initial state preparation:
$$|\phi_0\rangle = \frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c}\in\mathscr{C}}\sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}\rangle|\boldsymbol{e}\rangle|\mathbf{0}\rangle$$

adding $\boldsymbol{c}$ to $\boldsymbol{e}$:
$$\mapsto |\phi_1\rangle = \frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c}\in\mathscr{C}}\sum_{\boldsymbol{e}\in\mathbb{F}_q^n} f(\boldsymbol{e})|\boldsymbol{c}\rangle|\boldsymbol{c}+\boldsymbol{e}\rangle|\mathbf{0}\rangle$$
$$= \frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c}\in\mathscr{C}}|\boldsymbol{c}\rangle|\psi_{\boldsymbol{c}}\rangle|\mathbf{0}\rangle$$

apply the QFT on $|\psi_{\boldsymbol{c}}\rangle$:
$$\mapsto |\phi_2\rangle = \frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c}\in\mathscr{C}}|\boldsymbol{c}\rangle|\widehat{\psi_{\boldsymbol{c}}}\rangle|\mathbf{0}\rangle$$

applying the PGM coherently:
$$\mapsto |\phi_3\rangle = \frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c}\in\mathscr{C}}|\boldsymbol{c}\rangle\sum_{\boldsymbol{c}'\in\mathscr{C}}\alpha_{\boldsymbol{c},\boldsymbol{c}'}|Y_{\boldsymbol{c}'}\rangle|\boldsymbol{c}'\rangle$$
$$\text{with } |\widehat{\psi_{\boldsymbol{c}}}\rangle = \sum_{\boldsymbol{c}'\in\mathscr{C}}\alpha_{\boldsymbol{c},\boldsymbol{c}'}|Y_{\boldsymbol{c}'}\rangle$$

erase $\boldsymbol{c}$:
$$\mapsto |\phi_4\rangle = \frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c},\boldsymbol{c}'\in\mathscr{C}}\alpha_{\boldsymbol{c},\boldsymbol{c}'}|\boldsymbol{c}-\boldsymbol{c}'\rangle|Y_{\boldsymbol{c}'}\rangle|\boldsymbol{c}'\rangle$$

reverse PGM:
$$\mapsto |\phi_5\rangle = \frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c},\boldsymbol{c}'\in\mathscr{C}}\alpha_{\boldsymbol{c},\boldsymbol{c}'}|\boldsymbol{c}-\boldsymbol{c}'\rangle|Y_{\boldsymbol{c}'}\rangle|\mathbf{0}\rangle$$
$$= \sqrt{\mathrm{P}_{\mathrm{PGM}}}|\mathbf{0}\rangle\left(\frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c}\in\mathscr{C}}|Y_{\boldsymbol{c}}\rangle\right)+\sum_{\boldsymbol{c},\boldsymbol{c}'\neq\boldsymbol{c}}\alpha_{\boldsymbol{c},\boldsymbol{c}'}|\boldsymbol{c}-\boldsymbol{c}'\rangle|Y_{\boldsymbol{c}'}\rangle$$

measure 1st register, if $|\mathbf{0}\rangle$:
$$\mapsto |\phi_6\rangle = \frac{1}{\sqrt{|\mathscr{C}|}}\sum_{\boldsymbol{c}\in\mathscr{C}}|Y_{\boldsymbol{c}}\rangle = \frac{1}{n_0}\sum_{\boldsymbol{y}\in\mathscr{C}^\perp}\hat{f}(\boldsymbol{y})|\boldsymbol{y}\rangle \text{ w.p. } \mathrm{P}_{\mathrm{PGM}}$$

measuring the whole state:
$$\mapsto |\boldsymbol{y}\rangle \text{ with } \boldsymbol{y}\in\mathscr{C}^\perp.$$

As shown in [CT23, §6.3] in the case of the Bernoulli distribution there are choices of the code parameters k and n for which we measure $|\mathbf{0}\rangle$ with probability $1-o(1)$. This can be avoided by using a slight tweak of this algorithm as shown in [CT23, §6.3.2].

4.2 A slight tweak in the algorithm which avoids measuring the zero codeword

The idea is to tweak the algorithm in order to get as the last state before the final measurement the state $|U_0\rangle$ rather than $|\phi_6\rangle = \frac{1}{n_0}\sum_{\boldsymbol{y}\in\mathscr{C}^\perp}\hat{f}(\boldsymbol{y})|\boldsymbol{y}\rangle$, where

$$|U_0\rangle = \frac{\sum_{\boldsymbol{y}\in\mathscr{C}^\perp:\boldsymbol{y}\neq\boldsymbol{0}}\widehat{f}(\boldsymbol{y})|\boldsymbol{y}\rangle}{\|\sum_{\boldsymbol{y}\in\mathscr{C}^\perp:\boldsymbol{y}\neq\boldsymbol{0}}\widehat{f}(\boldsymbol{y})|\boldsymbol{y}\rangle\|}.$$

In order to achieve this, we replace the PGM by a measurement in a basis defined by the $|Z_{\boldsymbol{c}}\rangle$'s which basically correspond to the $|Y_{\boldsymbol{c}}\rangle$ where we removed the $|\boldsymbol{0}\rangle$ component. In other words, we define for all $\boldsymbol{c}\in\mathscr{C}$ the states

$$|Z_{\boldsymbol{c}}\rangle = \frac{1}{\sqrt{q^k}}\left(|U_0\rangle + \sum_{\boldsymbol{s}\neq\boldsymbol{0}}\chi_{\boldsymbol{c}}(u_{\boldsymbol{s}})|\widetilde{W_{\boldsymbol{s}}}\rangle\right).$$

These states are readily seen to be orthogonal. However projecting in this basis does not make the measurement complete. This can be circumvented by adding to the basis the $|\boldsymbol{0}\rangle$ state. We add an extra outcome $\boldsymbol{u}\in\mathbb{F}_q^n\backslash\mathscr{C}$, define $\mathscr{S} = \mathscr{C}\cup\{\boldsymbol{u}\}$ and set $|Z_{\boldsymbol{u}}\rangle = |\boldsymbol{0}\rangle$. We therefore have a projective measurement $\{|Z_{\boldsymbol{y}}\rangle\}_{\boldsymbol{y}\in\mathscr{S}}$. Again, we have $\langle\widehat{\psi_{\boldsymbol{c}}}|Z_{\boldsymbol{c}}\rangle \geq \sqrt{\mathrm{P}_{\mathrm{PGM}}} - \frac{1}{\sqrt{q^k}}$ and independent of $\boldsymbol{c}$ so w.p. at least $\left(\sqrt{\mathrm{P}_{\mathrm{PGM}}} - \frac{1}{\sqrt{q^k}}\right)^2$, we get the state $|U_0\rangle$. Then, if we measure this state $|U_0\rangle$, we will get a dual codeword $\boldsymbol{y}$ with probability $q(\boldsymbol{y})$ given by

Definition 11. *We define the probability of measuring the state $\boldsymbol{y}\in(\mathscr{C}^\perp)^*$ as*

$$q(\boldsymbol{y}) = \frac{|\hat{f}(\boldsymbol{y})|^2}{\sum_{\boldsymbol{z}\in(\mathscr{C}^\perp)^*}|\hat{f}(\boldsymbol{z})|^2}$$

All non zero elements of $\mathbb{F}_q^n$ have the same probability of belonging to $\mathscr{C}^\perp$, namely q^{-k}. This suggests that the probability q should take almost all its mass in $\mathcal{T}_\epsilon^{(n)}\cap(\mathscr{C}^\perp)^*$ since $p = |\hat{f}|^2$ takes almost all its mass on $\mathcal{T}_\epsilon^{(n)}$. We will show that this holds as long as we are in the achievability region under very mild assumptions on the typical set. We are now going to prove that for a family of nice typical sets we have

Theorem 8. *Let $q\geq 2$ be a prime power. We assume that we have an increasing sequence of positive integers $(n_i)_{i\geq 1}$ and a family of functions $(f_i)_{i\geq 1}$ with $f_i:\mathbb{F}_q^{n_i}\to\mathbb{C}$ with $\|f_i\|_2 = 1$ such that the sequence $(p_i)_{i\geq 1}$ defined by $p_i \stackrel{\triangle}{=} |\widehat{f_i}|^2$ admits a nice family of typical sets not containing 0. We let $H\stackrel{\triangle}{=}\lim_{i\to\infty}\frac{H_q(p_i)}{n_i}$. Let R be a real number chosen in $(0,H)$. Consider a family $(\mathscr{C}_i)_{i\geq 1}$ of random linear codes where $\mathscr{C}_i$ is chosen with a generator matrix chosen uniformly at random in $\mathbb{F}_q^{\lfloor Rn_i\rfloor\times n_i}$. Then, for any $\epsilon>0$ the algorithm of the quantum reduction with PGM using the measurement basis $\{|Z_{\boldsymbol{y}}\rangle\}_{\boldsymbol{y}\in\mathscr{S}}$ described above, outputs codewords of $\mathscr{C}_i^\perp$ in the typical set $\mathcal{T}_\epsilon^{(n_i)}$ with probability $1-\mathrm{negl}(i)$ as i tends to infinity.*

A straightforward corollary of this proposition is the following result

Proposition 7. *Let $g:\mathbb{F}_q\to\mathbb{C}$ with $\|g\|_2 = 1$ and $f:\mathbb{F}_q^n\to\mathbb{C}$ such that $f = g^{\otimes n}$ where n is an arbitrary positive integer. Let R be a real number in $(0, H_q(|\widehat{g}|^2))$. Choose a linear code $\mathscr{C}$ at random by picking its generator matrix uniformly at random in $\mathbb{F}_q^{\lfloor Rn\rfloor\times n}$. Then, for any $\epsilon>0$ the algorithm of the quantum reduction with PGM using the measurement basis $\{|Z_{\boldsymbol{y}}\rangle\}_{\boldsymbol{y}\in\mathscr{S}}$ described above, outputs codewords of $\mathscr{C}^\perp$ in the typical set $\mathcal{T}_\epsilon^{(n)}$ with probability $1-\mathrm{negl}(n)$ as n tends to infinity.*

Proof. This is an application of Theorem 8 where $f_i = g^{\otimes i}$ and we take $n_i = i$, $H = H_q(|\widehat{g}|^2)$, $A_i(\epsilon) = q^{-i(H+\epsilon)}$, $B_i(\epsilon) = q^{-i(H-\epsilon)}$. Proposition 2 shows that $\delta(\epsilon, n) = \mathrm{negl}(n)$ for all $\epsilon > 0$. Moreover, we can always exclude 0 from the typical set, it stays a nice typical set (it satisfies the three conditions of the definition). There is only one case where we can not do this, which corresponds to the case where $\widehat{g}(0) = 1$. However in this case $H = 0$ and we apply the theorem to the empty set. □

The case of the rank metric is also straightforward.

Proposition 8. *Let $q \geqslant 2$ be a prime power. We assume that there exist three increasing sequences $(a_i)_{i\geq 1}$, $(b_i)_{i\geq 1}$, $(t_i)_{i\geq 1}$ such that $a_i \geq b_i$ for every positive integer i and the sequence $(1 + t_i/a_i)(1 - t_i/b_i)$ tends to a constant $H \in (0, 1)$. Let R be a constant in $(0, H)$. Consider a family $(\mathscr{C}_i)_{i\geq 1}$ of random linear codes where $\mathscr{C}_i$ is chosen with a generator matrix chosen uniformly at random in $\mathbb{F}_q^{\lfloor Rn_i \rfloor \times n_i}$. Then, for any $\epsilon > 0$ the algorithm of the quantum reduction with PGM using the measurement basis $\{|Z_{\boldsymbol{y}}\rangle\}_{\boldsymbol{y}\in\mathscr{S}}$ described above, outputs codewords of $\mathscr{C}_i^{\perp}$ in the typical set $\mathcal{T}_\epsilon^{(n_i)}$ with probability $1 - \mathrm{negl}(i)$ as i tends to infinity.*

Proof. The 0 codeword is clearly excluded from the typical set when ϵ is small enough. From (26) we know that $\delta(\epsilon, n_i) = \mathrm{negl}(i)$ for any $\epsilon > 0$. (24) shows that we can take $\alpha(\epsilon) = 0$, $\beta(\epsilon) = 2\epsilon$ and the assumptions of Theorem 8 are verified. □

Remark 3. *It is insightful that at the limit when $R = H$ and large i we have*

$$R \approx (1 + t_i/a_i)(1 - t_i/b_i).$$

The Gilbert-Varshamov distance $t_{dGV}(a_i, b_i, 1 - R)$ of rate $1 - R$ and length $n_i = a_i b_i$ (which is the rate of the dual code of $\mathscr{C}_i$) corresponding to the rank metric viewing vectors of $\mathbb{F}_q^{n_i}$ as matrices in $\mathbb{F}_q^{a_i \times b_i}$ is the largest integer t such that

$$Ra_ib_i > a_it + t(b_i - t). \tag{27}$$

This corresponds to the minimum rank distance of a random linear code of rate $1 - R$ over $\mathbb{F}_q^{a_i b_i}$ viewed as matrix code over $\mathbb{F}_q^{a_i \times b_i}$. On the other hand since we measure with the PGM elements of the typical set we get at the limit where $\epsilon = 0$ elements of rank weight $b_i - t_i$. It is straightforward to verify that if $R = (1 + t_i/a_i)(1 - t_i/b_i)$ then $Ra_ib_i = (a_i + t_i)(b_i - t_i)$ which gives $Ra_ib_i = a_i(b_i - t_i) + (b_i - t_i)(b_i - (b_i - t_i))$. In other words, when we compare this to (27), $b_i - t_i$ is precisely the minimum rank distance of $\mathscr{C}_i^{\perp}$, and we obtain at the limit where the PGM works elements of minimum rank distance. The reason for this behavior will become apparent in the following subsection.

Let us prove now Theorem 8. Before proving this proposition a few lemmas will be very helpful. We let

$$\begin{aligned} X &\triangleq \sum_{\boldsymbol{y}\in\mathcal{T}_\epsilon^{(n_i)}\cap(\mathscr{C}^{\perp})^*} p(\boldsymbol{y}) \\ Y &\triangleq \sum_{\boldsymbol{y}\in(\mathscr{C}^{\perp})^*} p(\boldsymbol{y}). \end{aligned}$$

Note that the probability of outputting an element of $(\mathscr{C}^{\perp})^* \cap \mathcal{T}_\epsilon^{(n)}$ is nothing but the ratio $\frac{X}{Y}$. We expect that X and Y concentrate around their expectation. The following lemmas will use similar ideas as in the proof of Theorem 1. In particular we can set $\boldsymbol{s} = 0$ in Lemma 7 and get

Lemma 12. *Let $\gamma \triangleq H - R$. If ϵ is such that $2\beta(\epsilon) + \alpha(\epsilon) \leq \gamma/2$, then we have for any positive integer i*

$$\Pr_{\mathbf{G}}\left[X \geq \frac{1 - \delta(\epsilon, n_i) - K_2 q^{-n_i(H+\beta(\epsilon))} - q^{-\frac{\gamma n_i}{8}}}{q^{k_i}}\right] = 1 - \mathrm{negl}(i),$$

where $k_i \triangleq \lfloor Rn_i \rfloor$.

On the other hand the tails of Y above its expectation are bounded by

Lemma 13. *Let* $\gamma \triangleq H - R$. *We choose* ϵ *a positive constant such that* $2\beta(\epsilon) + \alpha(\epsilon) \leq \gamma/2$. *We have* $\Pr\left(Y/X \leq 1 + \sqrt{\delta(\epsilon, n_i)}\right) = 1 - \mathrm{negl}(i)$.

Proof. It will be helpful to decompose Y as:

$$\begin{aligned} Y &= \sum_{\boldsymbol{y} \in (\mathscr{C}^\perp)^*} p(\boldsymbol{y}) \\ &= \underbrace{\sum_{\boldsymbol{y} \in (\mathscr{C}^\perp)^* \cap \mathcal{T}_\epsilon^{(n_i)}} p(\boldsymbol{y})}_{X} + \underbrace{\sum_{\boldsymbol{y} \in (\mathscr{C}^\perp)^* \setminus \mathcal{T}_\epsilon^{(n_i)}} p(\boldsymbol{y})}_{\triangleq Z}. \end{aligned}$$

This implies that

$$\Pr\left(Y/X \leq 1 + \sqrt{\delta(\epsilon, n_i)}\right) = \Pr\left(Z/X \leq \sqrt{\delta(\epsilon, n_i)}\right). \tag{28}$$

We define now the following events

$$\begin{aligned} \mathscr{E}_1 &\triangleq \left\{ X \geq \frac{1 - \delta(\epsilon, n_i) - K_2 q^{-n_i(H+\beta(\epsilon))} - q^{-\frac{\gamma n_i}{8}}}{q^{k_i}} \right\} \\ \mathscr{E}_2 &\triangleq \left\{ Z \leq \frac{\sqrt{\delta(\epsilon, n_i)}}{(1 - \delta(\epsilon, n_i) - K_2 q^{-n_i(H+\beta(\epsilon))} - q^{-\frac{\gamma n_i}{8}}) q^{k_i}} \right\}. \end{aligned}$$

We have

$$\Pr\left(Z/X \leq \sqrt{\delta(\epsilon, n_i)}\right) \geq \Pr(\mathscr{E}_1 \cap \mathscr{E}_2) = \Pr(\mathscr{E}_1) + \Pr(\mathscr{E}_2) - \Pr(\mathscr{E}_1 \cup \mathscr{E}_2) \geq \Pr(\mathscr{E}_1) + \Pr(\mathscr{E}_2) - 1. \tag{29}$$

The first term is handled by Lemma 12, we have

$$\Pr\left(X \geq \frac{1 - \delta(\epsilon, n_i) - K_2 q^{-n_i(H+\beta(\epsilon))} - q^{-\frac{\gamma n_i}{8}}}{q^{k_i}} \right) = 1 - \mathrm{negl}(i). \tag{30}$$

The second probability $\Pr(\mathscr{E}_2)$ is treated by the Markov inequality. We observe that

$$\begin{aligned} \mathbb{E}\{Z\} &= \mathbb{E}\left\{ \sum_{\boldsymbol{y} \in (\mathscr{C}^\perp)^* \setminus \mathcal{T}_\epsilon^{(n_i)}} p(\boldsymbol{y}) \right\} \\ &\leq \sum_{\boldsymbol{y} \notin \mathcal{T}_\epsilon^{(n_i)}} p(\boldsymbol{y}) \Pr(\boldsymbol{y} \in (\mathscr{C}^\perp)^*) \\ &= \frac{1}{q^{k_i}} \sum_{\boldsymbol{y} \notin \mathcal{T}_\epsilon^{(n_i)}} p(\boldsymbol{y}) \\ &= \frac{\delta(\epsilon, n_i)}{q^{k_i}}. \end{aligned} \tag{31}$$

Let $p \triangleq \delta(\epsilon, n_i) + K_2 q^{-n_i(H+\beta(\epsilon))} + q^{-\frac{\gamma n_i}{8}}$. By using Markov's inequality we get

$$\begin{aligned} \Pr\left(Z \geq \frac{\sqrt{\delta(\epsilon, n_i)}}{(1-p) q^{k_i}} \right) &\leq \frac{\mathbb{E}\{Z\}}{\frac{\sqrt{\delta(\epsilon, n_i)}}{(1-p) q^{k_i}}} && (32) \\ &\leq (1-p)\sqrt{\delta(\epsilon, n_i)} && (33) \\ &= \mathrm{negl}(i), && (34) \end{aligned}$$

where (32) follows directly from Markov's inequality, (33) follows from the upper bound (31) on the expectation of Z and (34) follows from the fact that family of typical sets is nice. We conclude the proof by using (30) and (34) in (29). □

By using these two lemmas we obtain directly Theorem 8 as shown by

Proof of Theorem 8. Let ϵ be any positive real such that $2\beta(\epsilon)+\alpha(\epsilon) \leq \gamma/2$. The probability of measuring an element of $\mathcal{T}_\epsilon^{(n_i)}$ is given by the ratio X/Y. By using Lemma 13 we know that X/Y is greater than $\frac{1}{1+\sqrt{\delta(\epsilon,n_i)}}$ with probability $1-\mathrm{negl}(i)$ over the choice of the generator matrix $\mathbf{G}$. This leads to the proof of the theorem by observing that (i) the probability of measuring an element of $\mathcal{T}_\epsilon^{(n_i)}$ is increasing with ϵ, therefore if this property holds for a given $\epsilon_0 > 0$, it holds for any $\epsilon \geq \epsilon_0$, (ii) we can choose ϵ arbitrarily close to 0 satisfying $2\beta(\epsilon)+\alpha(\epsilon) \leq \gamma/2$. □

In other words, we obtain with this quantum reduction algorithm a way of obtaining elements in $\mathcal{T}_\epsilon^{(n_i)}$ whose probability is in the range $\left[q^{-n_i(H+\alpha(\epsilon))}, q^{-n(H-\beta(\epsilon))}\right]$. It turns out that when the difference $H-R$ is small we can not have much better (namely elements of bigger probability) as shown by

Proposition 9. *Let $q \geq 2$ be a prime power, R be a fixed real number in $(0,1)$. We choose a random code $\mathscr{C}$ in $\mathbb{F}_q^n$ by choosing a generator matrix in $\mathbb{F}_q^{\lfloor Rn \rfloor \times n}$ uniformly at random. Let p a probability distribution on $\mathbb{F}_q^n$. Let ϵ be a fixed positive real number. The probability over $\mathbf{G}$ that there exists a non zero codeword $\boldsymbol{c}$ in $\mathscr{C}^\perp$ such that $p(\boldsymbol{c}) > q^{-n(R-\epsilon)}$ is a negligible function of n.*

Proof. For $\epsilon > 0$ let $\mathcal{B}_\epsilon^{(n)} \stackrel{\triangle}{=} \left\{ \boldsymbol{y} \in \left(\mathbb{F}_q^n\right)^*,\ p(\boldsymbol{y}) > q^{-n(R-\epsilon)} \right\}$. Let $k \stackrel{\triangle}{=} \lfloor Rn \rfloor$. From Lemma 2, we have $\mathbb{E}\left(|\mathcal{B}_\epsilon^{(n)} \cap (\mathscr{C}^\perp)^*|\right) = \frac{|\mathcal{B}_\epsilon^{(n)}|}{q^k}$. Since for any $\epsilon > 0$, $p(\mathcal{B}_\epsilon^{(n)}) \leq 1$ and $p(\mathcal{B}_\epsilon^{(n)}) > |\mathcal{B}_\epsilon^{(n)}| q^{-n(R-\epsilon)}$ by definition of $\mathcal{B}_\epsilon^{(n)}$ we deduce that $|\mathcal{B}_\epsilon^{(n)}| < q^{n(R-\epsilon)}$. Therefore

$$\begin{aligned}
\mathbb{E}\left(|\mathcal{B}_\epsilon^{(n)} \cap (\mathscr{C}^\perp)^*|\right) &< \frac{q^{n(R-\epsilon)}}{q^k} \\
&= q^{-\epsilon n + O(1)} \\
&= \mathrm{negl}(n).
\end{aligned}$$

Since $\Pr\left(\mathcal{B}_\epsilon^{(n)} \cap (\mathscr{C}^\perp)^* \neq \emptyset\right) \leq \mathbb{E}\left(|\mathcal{B}_\epsilon^{(n)} \cap (\mathscr{C}^\perp)^*|\right)$ we deduce that $\Pr\left(\mathcal{B}_\epsilon^{(n)} \cap (\mathscr{C}^\perp)^* \neq \emptyset\right) = \mathrm{negl}(n)$. □

This lemma together with Proposition 7 leads for instance in the memoryless case to the following theorem.

Theorem 3. *Let $g : \mathbb{F}_q \to \mathbb{C}$ with $\|g\|_2 = 1$. Let R be a real number in $(0, H_q(|\widehat{g}|^2))$. Let $f : \mathbb{F}_q^n \to \mathbb{C}$ be so that $f = g^{\otimes n}$. Let $\mathscr{C}$ be a linear code on $\mathbb{F}_q^n$ with generator matrix $\mathbf{G}$ drawn uniformly at random in $\mathbb{F}_q^{\lfloor Rn \rfloor \times n}$. For any $\epsilon > 0$:*

- *Algorithm 4.1 outputs codewords $\boldsymbol{c} \in (\mathscr{C}^\perp)^*$ such that $p(\boldsymbol{c}) \geq q^{-n(H_q(|\widehat{g}^2|)-\epsilon)}$ with probability $1-\mathrm{negl}(n)$.*
- *The probability over $\mathbf{G}$ that there exists a codeword $\boldsymbol{c}$ in $(\mathscr{C}^\perp)^*$ such that $p(\boldsymbol{c})$ is larger than $q^{-n(R-\epsilon)}$ is a negligible function of n.*

4.3 Distribution associated to a metric

We study an application of the previous results when the probability distribution is associated to an additive metric on $\mathbb{F}_q^n$. What we mean here is a metric defined from a metric $|\cdot|_{\mathbb{F}_q}$ over $\mathbb{F}_q$ as $|\boldsymbol{x}|_{\mathbb{F}_q^n} = \sum_{i=1}^{n} |x_i|_{\mathbb{F}_q}$. The distribution p associated to this metric on $\mathbb{F}_q^n$ is defined by $p(\boldsymbol{x}) = q^{-\lambda|\boldsymbol{x}|_{\mathbb{F}_q^n}}$ for a unique $\lambda > 0$. Let $g : \mathbb{F}_q \to \mathbb{C}$ with $\|g\|_2 = 1$ such that $f = g^{\otimes n}$. We recall that $p = |\widehat{f}|^2$ and we define $r = |\widehat{g}|^2$.

Lemma 14. *For each metric on $\mathbb{F}_q$, there exists a unique $\lambda > 0$ such that $r(\alpha)$ defined by $r(\alpha) = q^{-\lambda|\alpha|_{\mathbb{F}_q}}$ is a probability distribution .*

Proof. Let $F(\lambda) = \sum_{\alpha\in\mathbb{F}_q} q^{-\lambda|\alpha|_{\mathbb{F}_q}}$. Observe that $F(0) = \sum_{\alpha\in\mathbb{F}_q} q^{0|\alpha|_{\mathbb{F}_q}} = q$ and $\lim_{\lambda\to\infty} F(\lambda) = 0$. Since F is continuous and decreasing over $[0, \infty[$, we deduce that there exists a unique $\lambda > 0$ such that $F(\lambda) = 1$, which finishes the proof.

□

This probability distribution is a product probability distribution $p^{\perp}(\boldsymbol{x}) = \prod_{i=1}^{n} r(x_i)$ with $r(x_i) = q^{-\lambda|x_i|_{\mathbb{F}_q}}$. To apply Algorithm 4.1, we can construct the amplitude function f defined on $\mathbb{F}_q^n$ such that $|\widehat{f}|^2 = p$ by choosing $\widehat{f} = \sqrt{p}$ which leads to $f(\boldsymbol{x}) = \frac{1}{\sqrt{q^n}} \sum_{\boldsymbol{y}\in\mathbb{F}_q^n} \chi_{\boldsymbol{y}}(-\boldsymbol{x})\sqrt{p(\boldsymbol{y})}$.

Using Theorem 3, it turns out that when R approaches its limit $H_q(|\widehat{g}|^2)$ at which we might hope to decode, Algorithm 4.1 outputs non-zero codewords in $\mathscr{C}^{\perp}$ of probability larger than $q^{-nH_q(|\widehat{g}|^2)}$ with probability $1 - o(1)$ and there exists no non-zero codeword in $\mathscr{C}^{\perp}$ of smaller weight. Then, as the distribution p is a decreasing function of $|\boldsymbol{x}|_{\mathbb{F}_q^n}$, we obtain non-zero codewords of nearly minimum weight.

Theorem 9. *Let $g : \mathbb{F}_q \to \mathbb{C}$ with $\|g\|_2 = 1$ such that $f = g^{\otimes n}$. Let $\mathscr{C}$ be a linear code on $\mathbb{F}_q^n$. For any $\epsilon > 0$:*

- *We assume that there exists $\gamma > 0$ such that $H_q(|\widehat{g}|^2) - R > \gamma$. Algorithm 4.1 outputs non-zero codewords $\boldsymbol{x} \in \mathscr{C}^{\perp}$ such that $|\boldsymbol{x}|_{\mathbb{F}_q^n} \geqslant \frac{n(H_q(|\widehat{g}|^2)+\epsilon)}{\lambda}$ with probability $1 - o(1)$ as n tends to infinity.*
- *The probability that there exists a non-zero codeword $\boldsymbol{x}$ in $\mathscr{C}^{\perp}$ such that $|\boldsymbol{x}|_{\mathbb{F}_q^n} \geqslant \frac{n(R-\epsilon)}{\lambda}$ is $o(1)$.*

Proof of the Theorem. Using Theorem 3, in the first part, we know that if there exists $\gamma > 0$ such that $H_q(|\widehat{g}|^2) - R > \gamma$, then for any $\epsilon > 0$, Algorithm 4.1 outputs non-zero codewords $\boldsymbol{x}$ in $\mathscr{C}^{\perp}$ of probability larger than $q^{-n(H_q(|\widehat{g}|^2)+\epsilon)}$ with probability $1 - o(1)$ as n tends to infinity. Here, since $p(\boldsymbol{x}) = q^{-\lambda|\boldsymbol{x}|_{\mathbb{F}_q^n}}$, we get vectors $\boldsymbol{x} \in \mathbb{F}_q^n$ such that $|\boldsymbol{x}|_{\mathbb{F}_q^n} \geqslant \frac{n(H_q(|\widehat{g}|^2)+\epsilon)}{\lambda}$ with probability $1 - o(1)$.

Then, the second part claims that for any $\epsilon' > 0$, the probability over $\mathbf{G}$ that there exists a non-zero codeword in $\mathscr{C}^{\perp}$ of probability larger than $q^{-n(R-\epsilon')}$ goes to 0 as n tends to infinity. Here, we get that the probability that there exists no nonzero codeword $\boldsymbol{x}$ in $\mathscr{C}^{\perp}$ such that $|\boldsymbol{x}|_{\mathbb{F}_q^n} \geqslant \frac{n(R-\epsilon')}{\lambda}$ is $1 - o(1)$.

□

References

[AAB+24] Gora Adj, Nicolas Aragon, Stefano Barbero, Magali Bardet, Emanuele Bellini, Loïc Bidoux, Jesús-Javier Chi-Domínguez, Victor Dyseryn, Andre Esser, Thibauld

Feneuil, Philippe Gaborit, Romaric Neveu, Matthieu Rivain, Luis Rivera-Zamarripa, Carlo Sanna, Jean-Pierre Tillich, Javier Verbel, and Floyd Zweydinger. MIRATH. NIST Round 2 submission to the Additional Call for Signature Schemes, October 2024.

[AAKV01] Dorit Aharonov, Andris Ambainis, Julia Kempe, and Umesh Vazirani. Quantum walks on graphs. *Proceedings of ACM Symposium on Theory of Computation (STOC'01)*, pages 50–59, 2001.

[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-Ourobouros, LAKE and LOCKER). Second round submission to the NIST post-quantum cryptography call, March 2019.

[ADG+24] Nicolas Aragon, Victor Dyseryn, Philippe Gaborit, Pierre Loidreau, Julian Renner, and Antonia Wachter-Zeh. LowMS: a new rank metric code-based KEM without ideal structure. *Des. Codes Cryptogr.*, 92(4):1075–1093, 2024.

[ADG25] Carlos Aguilar Melchor, Victor Dyseryn, and Philippe Gaborit. Somewhat homomorphic encryption based on random codes. *Des. Codes Cryptogr.*, 93(6):1645–1669, 2025.

[AG11] Sanjeev Arora and Rong Ge. New algorithms for learning in presence of errors. In Luca Aceto, Monika Henzinger, and Jiří Sgall, editors, *Automata, Languages and Programming*, volume 6755 of *LNCS*, pages 403–415. Springer Berlin Heidelberg, 2011.

[Bar97] Alexander Barg. Complexity issues in coding theory. *Electronic Colloquium on Computational Complexity*, October 1997.

[BCF+25] Loïc Bidoux, Jesús-Javier Chi-Domínguez, Thibauld Feneuil, Philippe Gaborit, Antoine Joux, Matthieu Rivain, and Adrien Vinçotte. RYDE: a digital signature scheme based on rank syndrome decoding problem with MPC-in-the-Head paradigm. *Des. Codes Cryptogr.*, 93(5):1451–1486, 2025.

[Ber10] Daniel J. Bernstein. Grover vs. McEliece. In Nicolas Sendrier, editor, *Post-Quantum Cryptography 2010*, volume 6061 of *LNCS*, pages 73–80. Springer, 2010.

[BKMH97] Masashi Ban, Keiko Kurokawa, Rei Momose, and Osamu Hirota. Optimum measurements for discrimination among symmetric quantum states and parameter estimation. *International Journal of Theoretical Physics*, 36:1269–1288, June 1997.

[CHL+23] Yilei Chen, Zihan Hu, Qipeng Liu, Han Luo, and Yaxin Tu. LWE with quantum amplitudes: Algorithm, hardness, and oblivious sampling. Cryptology ePrint Archive, Paper 2023/1498, 2023.

[CL21] André Chailloux and Johanna Loyer. Lattice sieving via quantum random walks. In Mehdi Tibouchi and Huaxiong Wang, editors, *Advances in Cryptology - ASIACRYPT 2021 - 27th International Conference on the Theory and Application of Cryptology and Information Security, Singapore, December 6-10, 2021, Proceedings, Part IV*, volume 13093 of *Lecture Notes in Computer Science*, pages 63–91. Springer, 2021.

[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.

[CT91] Thomas M. Cover and Joy A. Thomas. *Information Theory*. Wiley Series in Telecommunications. Wiley, 1991.

[CT06] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory 2nd Edition (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, July 2006.

[CT23] André Chailloux and Jean-Pierre Tillich. The quantum decoding problem. IACR Cryptology ePrint Archive, Report2023/1686, October 2023.

[CT24] André Chailloux and Jean-Pierre Tillich. The quantum decoding problem. In Frédéric Magniez and Alex Bredariol Grilo, editors, *19th Conference on the Theory of Quantum Computation, Communication and Cryptography, TQC 2024, September 9-13, 2024, Okinawa, Japan*, volume 310 of *LIPIcs*, pages 6:1–6:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024.

[CT25] André Chailloux and Jean-Pierre Tillich. Quantum advantage from soft decoders. 2025. Arxiv quant-ph, https://arxiv.org/abs/2411.12553.

[DCT91] Amir Dembo, Thomas M. Cover, and Joy A. Thomas. Information theoretic inequalities. *IEEE Trans. Inform. Theory*, 37(6):1501–1518, 1991.

[DFS24] Thomas Debris-Alazard, Pouria Fallahpour, and Damien Stehlé. Quantum oblivious lwe sampling and insecurity of standard model lattice-based snarks. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, page 423–434, New York, NY, USA, 2024. Association for Computing Machinery.

[DRT24] Thomas Debris-Alazard, Maxime Remaud, and Jean-Pierre Tillich. Quantum reduction of finding short code vectors to the decoding problem. *IEEE Trans. Inform. Theory*, 70(7):5323–5342, 2024.

[EF01] Yonina C. Eldar and G. David Jr. Forney. On quantum detection and the square-root measurement. *IEEE Trans. Inform. Theory*, 47(3):858–872, 2001.

[Gro96] Lov Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC)*, pages 212–219, 1996.

[Hol12] Alexander S. Holevo. *Quantum systems, channels, information*, volume 16 of *Studies in Mathematical Physics*. De Gruyter, 2012.

[JSW+24] Stephen P. Jordan, Noah Shutty, Mary Wootters, Adam Zalcman, Alexander Schmidhuber, Robbie King, Sergei V. Isakov, and Ryan Babbush. Optimization by decoded quantum interferometry. *arXiv preprint arXiv:2408.08292*, 2024.

[KRS09] Robert Konig, Renato Renner, and Christian Schaffner. The operational meaning of min- and max-entropy. *IEEE Transactions on Information Theory*, 55(9):4337–4347, 2009.

[KT17] Ghazal Kachigar and Jean-Pierre Tillich. Quantum information set decoding algorithms. In *Post-Quantum Cryptography 2017*, volume 10346 of *LNCS*, pages 69–89, Utrecht, The Netherlands, June 2017. Springer.

[Laa16] Thijs Laarhoven. *Search problems in cryptography*. PhD thesis, Eindhoven University of Technology, 2016.

[LN97] Rudolf Lidl and Harald Niederreiter. *Finite fields*, volume 20 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 1997. With a foreword by P. M. Cohn.

[MNRS11] Frédéric Magniez, Ashwin Nayak, Jérémie Roland, and Miklos Santha. Search via quantum walk. *SIAM Journal on Computing (SICOMP)*, 40(1):142–164, 2011.

[NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[ON00] Tomohiro Ogawa and Hiroshi Nagaoka. Strong converse and Stein's lemma in quantum hypothesis testing. *IEEE Trans. Inform. Theory*, 46(7):2428–2433, 2000.

[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.

[SKIH98] Masahide Sasaki, Kentaro Kato, Masayuki Izutsu, and Osamu Hirota. Quantum channels showing superadditivity in classical capacity. *Phys. Rev. A*, 58:146–158, Jul 1998.