# The uncertainty principle over finite fields

Martino Borello, Patrick Solé

# The uncertainty principle over finite fields

Martino Borello

*Université Paris 8, Laboratoire de Géométrie, Analyse et Applications, LAGA, Université Sorbonne Paris Nord, CNRS, UMR 7539, F-93430, Villetaneuse, France*

Patrick Solé

*Aix Marseille University, CNRS, Centrale Marseille, I2M, Marseille, France*

**Abstract**

In this paper we study the uncertainty principle (UP) connecting a function over a finite field and its Mattson-Solomon polynomial, which is a kind of Fourier transform in positive characteristic. Three versions of the UP over finite fields are studied, in connection with the asymptotic theory of cyclic codes. We first show that no finite field satisfies the *strong* version of UP, introduced recently by Evra, Kowalsky, Lubotzky, 2017. A refinement of the *weak* version is given, by using the asymptotic Plotkin bound. A *naive* version, which is the direct analogue over finite fields of the Donoho-Stark bound over the complex numbers, is proved by using the BCH bound. It is strong enough to show that there exist sequences of cyclic codes of length $n$, arbitrary rate, and minimum distance $\Omega(n^\alpha)$ for all $0 < \alpha < 1/2$. Finally, a connection with Ramsey Theory is pointed out.

*Keywords:* uncertainty principle, cyclic codes, Mattson-Solomon polynomial, BCH bound, asymptotically good codes
*2010 MSC:* 43A99, 94B15, 20C05

## 1. Introduction

The uncertainty principle (UP) is a very famous inequality in Physics [7], and Signal Processing [3] (see [16] for a general very recent survey on the UP). It compares the supports of functions and of their complex-valued Fourier transforms. In a paper of 2017 [4], a connection between UP and the asymptotic performance of cyclic codes was pointed out. Note that the

existence of an asymptotically good family of cyclic codes is a problem open for more than half a century [13]. The reference [4] is an attempt to motivate further research into, and eventually solve this very hard problem. In a recent note [11], a connection with Ramsey Theory and the Szemerédi Theorem was derived.

In the present paper, we replace the classical Discrete Fourier transform ([3, §2]) by a vectorial version of the Mattson-Solomon polynomial ([10, Ch.8, §6]). In contrast with all the results in [16], this transform takes its values in a finite field. We study three versions of the UP for this kind of transform.

The *strong* version of the UP over finite fields is defined in [4] by analogy with the bound of [14] for the classical Fourier transform. Exploiting the connection with the theory of MDS codes, we show that no finite fields may satisfy the strong UP.

The *weak* version of the UP is a similar and weakened statement depending on two real parameters $\lambda$ and $\epsilon$. In [4] it is shown that a finite field satisfying this UP enjoys sequences of asymptotically good cyclic codes. Here, we show that, if this version holds over $\mathbb{F}_q$, then $\lambda < \frac{q-1}{q}$.

The third version is the straight analogue of the Donoho-Stark bound of [3] and we call this the *naive* version. It allows us to construct sequences of cyclic codes with nonzero rate and minimum distance growing like a power $\alpha$ of the length with $0 < \alpha < 1/2$.

Finally, with similar arguments, we give an alternate proof of the results of [11], based on the familiar BCH bound and a generalization based on the Hartmann-Tzeng bound on the minimum distance of cyclic codes ([8, Th. 4.5.6]).

The material is organized as follows: the next section collects background material; Section 3 is about the strong version; Section 4 contains numerical results related to the weak version; Section 5 is dedicated to the naive version of UP; Section 6 deals with the Ramsey Theory connection; Section 7 concludes the article. An Appendix building on the naive version shows the existence of cyclic codes of all rates with minimum distance $\Omega(n^\alpha)$, for all $0 < \alpha < 1/2$.

## 2. Background

Throughout the paper, $\mathbb{F}_q$ denotes a finite field of cardinality $q$, where $q$ is a prime power.

*2.1. Linear codes and asymptotics*

The **(Hamming) weight** of $x \in \mathbb{F}_q^n$ is denoted by $w_H(x)$. The minimum nonzero weight $d$ of a linear code is called the **minimum distance**. A **linear code** is a subspace of $\mathbb{F}_q^n$. Its parameters are written as $[n, k, d]$ where $k$ is its **dimension** as an $\mathbb{F}_q$-vector space. If $\mathcal{C}_n$ is a sequence of linear codes of parameters $[n, k_n, d_n]$, the **rate** $R$ and **relative distance** $\delta$ are defined as

$$R := \liminf_{n\to\infty} \frac{k_n}{n} \text{ and } \delta := \liminf_{n\to\infty} \frac{d_n}{n}.$$

A family of codes is said to be **good** iff it contains a sequence with rate and relative distance such that $R \cdot \delta \neq 0$. The **binary entropy function** $H(x)$ of the real variable $x$ is defined (see [10, p.308]) for $0 < x < 1$,

$$H(x) := -x \log_2(x) - (1-x) \log_2(1-x).$$

*2.2. Cyclic Codes*

Consider the quotient ring $R(\mathbb{F}_q, n) := \mathbb{F}_q[x]/(x^n - 1)$. We will identify each class of $R(\mathbb{F}_q, n)$ with the unique polynomial of degree less than $n$ contained in it. The ring $R(\mathbb{F}_q, n)$ is principal, and we denote by $C(f)$ the ideal with generator $f$. It is well-known that every ideal of $R(\mathbb{F}_q, n)$ has a unique monic generator of minimal degree, and this is a divisor of $x^n - 1$. Whenever we will consider and ideal $C(f)$, we will implicitly assume that $f$ is such generator. The polynomials of $R(\mathbb{F}_q, n)$ are in one-to-one correspondence with the vectors of $\mathbb{F}_q^n$, by the map

$$\varphi : f := (f_0, f_1, \ldots, f_{n-1}) \mapsto f(x) := \sum_{i=0}^{n-1} f_i x^i.$$

The (Hamming) weight of a polynomial is the (Hamming) weight of the corresponding vector. A **cyclic code** is an ideal in $R(\mathbb{F}_q, n)$ or its preimage in $\mathbb{F}_q^n$ via $\varphi$.

The **zeros** of $C(f)$ are the roots of $f$ in the algebraic closure of $\mathbb{F}_q$. The dimension of $C(f)$ is $n - \deg(f)$, and $\deg(f)$ equals the number of zeros of $C(f)$ (see for example [10, Chap. 7]). The well-known **BCH-bound** [10, Chap., Th. 8] states that if among the zeros of $f$ there exists $\delta - 1$ consecutive powers of a primitive $n$-th root of unity and $(m, q) = 1$, then the minimum distance of $C(f)$ is at least $\delta$.

*2.3. Mattson-Solomon polynomial*

Let $\zeta$ be a primitive root of unity of order $n$ in the algebraic closure $\overline{\mathbb{F}_q}$ of $\mathbb{F}_q$. The **Mattson-Solomon polynomial** ([10, Chap. 8]) associated with a vector $f := (f_0, f_1, \ldots, f_{n-1})$ is the following polynomial in $\overline{\mathbb{F}_q}[z]$:

$$\hat{f}(z) := \sum_{i=1}^{n} F_i z^{n-i},$$

where $F_i := f(\zeta^i)$ is the evaluation of $f(x)$ in $\zeta^i$. It is sometimes called a discrete Fourier transform of $f$. In the following, we will prefer the vectorial version of the Mattson-Solomon polynomial, which is

$$\hat{f} := (F_1, F_2, \ldots, F_n) = (f(\zeta), f(\zeta^2), \ldots, f(\zeta^n)).$$

*2.4. An invariant of fields*

We introduce here, following [4], the invariant of fields

$$\mu(\mathbb{F}_q, n) := \min\{d(C(f)) + \dim C(f) \mid f \in R(\mathbb{F}_q, n), f \neq 0\}.$$

By the Singleton bound, $\mu(\mathbb{F}_q, n) \leq n+1$ for any $n$. Moreover, equality holds if $n$ is prime and $q$ is a primitive root modulo $n$ or if $q$ is a power of $n$ ([4, Propositions 4.3. and 4.4.]).

**Remark 1.** *Note that $\mu(\mathbb{F}_q, n) = n+1$ if and only if all cyclic codes of length $n$ over $\mathbb{F}_q$ are MDS.*

**Remark 2.** *Note that, as observed in [4], if we consider the complex field $\mathbb{C}$ instead of $\mathbb{F}_q$, then the* ***uncertainty principle for simple cyclic group*** *(proved for example in [6, 14]) may be reformulated as follows: $\mu(\mathbb{C}, p) = p+1$ for any prime $p$.*

In next sections we aim to investigate analogues of the uncertainty principle over finite fields.

## 3. Strong version of UP

The following version of UP is the one stated in [4].

**Definition 1.** *A finite field $\mathbb{F}_q$ satisfies the* **(strong) uncertainty principle** *if, for all primes $p$, we have $\mu(\mathbb{F}_q, p) = p + 1$.*

As we have already mentioned, in [4, Prop. 4.3] it is shown that $\mu(\mathbb{F}_q, p) = p + 1$ if $q$ is primitive mod $p$ (under this hypothesis, there exists exactly 3 cyclic codes of length $p$ over $\mathbb{F}_q$ and of positive dimension, that are all trivial MDS codes). We show below that is almost the only case.

**Theorem 1.** *Assume the MDS conjecture* [10, Res. Prob. 11.4]. *If $q$ is not primitive modulo $p$ and if $p > q + 2$ then $\mu(\mathbb{F}_q, p) \leq p$.*

*Proof.* By the hypothesis we know that there are polynomials $f | x^p - 1$ in $R(\mathbb{F}_q, p)$ such that $1 < \deg(f) < p - 1$. Let $[p, k, d]$ be the parameters of the cyclic code $C(f)$. Let $g \in C(f)$ of weight $d$. The code $C(g) \subseteq C(f)$ is certainly not the repetition code, since $k > 1$. Its parameters $[p, k' \leq k, d]$ satisfy $d + k' \geq \mu(\mathbb{F}_q, p)$ and if, arguing by contradiction, $\mu(\mathbb{F}_q, p) > p$, we see that $d \geq p - k' + 1$, entailing that $C(g)$ is MDS. But we know, by [10, Chapt. 11], that MDS codes of parameters $[N, K, D]$ with $1 < K < N - 1$ only exist for lengths at most $q + 2$. This is the so-called MDS Conjecture that is now proved in many cases [1, 2]. Note that codes of parameters $[N, 1, N]$ and $[N, N - 1, 2]$ exist for all lengths $N$. □

**Remark 3.** *A similar (slightly weaker) result holds unconditionally, since it is well-known that nontrivial MDS codes have length at most $2q - 2$ (see for example [8, Corollary 7.4.4]). So, with the same arguments we can prove that $\mu(\mathbb{F}_q, p) \leq p$ if $q$ is not primitive modulo $p$ and if $p > 2q - 2$.*

**Corollary 1.** *No finite field satisfies the (strong) uncertainty principle.*

*Proof.* Suppose that $\mathbb{F}_q$ satisfies the (strong) uncertainty principle. Then Theorem 1 would imply that all $p > q + 2$ (or eventually $> 2q - 2$, if we refer to Remark 3) are necessarily such that $q$ is primitive modulo $p$. But we know that this is not possible: it is enough to consider all primes $p$ such that $q$ is a quadratic residue (so that $q$ cannot be primitive modulo $p$) and observe that, by quadratic reciprocity, these correspond to $p$ being in some non-empty set of residue classes modulo $4q$ (so that they are infinitely many by Dirichlet's theorem). □

## 4. Weak version of UP

The following is Definition 5.3 in [4].

**Definition 2.** *Let $0 < \epsilon < \lambda \leq 1$. A finite field $\mathbb{F}_q$ satisfies the* $(\epsilon, \lambda)$**-uncertainty principle**, *if*

$$\begin{aligned} \mu(\mathbb{F}_q, p) &> \lambda p & (1) \\ ord_p(q) &< \epsilon p, & (2) \end{aligned}$$

*for infinitely many primes $p$, where $ord_p(q)$ is the order of $q$ in $\mathbb{F}_p^*$.*

In [4, Th. 5.4] it is shown that finite fields satisfying this definition enjoy sequences of asymptotically good cyclic codes. Intuitively, (1) guarantees to get codes with a large minimum distance, whereas (2) guarantees to get codes with a large dimension.

In the following table, we show some values of $\mu(\mathbb{F}_2, p)$, for small primes $p$, omitting those for which 2 is primitive modulo $p$.

| $p$ | 7 | 17 | 23 | 31 | 41 | 43 | 47 | 71 | 73 | 79 | 89 | 97 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mu(\mathbb{F}_2, p)$ | 7 | 14 | 19 | 20 | 30 | 28 | 35 | 47 | 37 | 55 | 45 | 64 |

In the following proposition, we get a restriction on possible values of $\lambda$ for finite fields satisfying the principle above.

**Proposition 1.** *If $\mathbb{F}_q$ satisfies the $(\epsilon, \lambda)$-uncertainty principle then $\lambda < \frac{q-1}{q}$.*

*Proof.* By combining [4, Th. 5.4] with the same argument as in the proof of Theorem 1, we see that under the hypothesis, there are sequences of cyclic codes of length $p$ over $\mathbb{F}_q$, of rate $R$ and relative distance $\delta$ such that

$$p\lambda < \mu(\mathbb{F}_q, p) < p\delta + pR.$$

In particular this implies that

$$\lambda < \min\{\delta + \alpha_q(\delta) \mid \delta \in (0, 1)\},$$

where $\alpha_q(\delta)$ is the largest possible rate of a code of relative distance $\delta$. But we know, by the asymptotic Plotkin bound [8, Th.2.10.2], that

- for $0 < \delta < \frac{q-1}{q}$, we have $\alpha_q(\delta) < 1 - \frac{q\delta}{q-1}$, and

- for $\frac{q-1}{q} \leq \delta < 1$, we have $\alpha_q(\delta) = 0$.

It follows that the function $\delta + \alpha_q(\delta)$ is smaller than or equal to $f(\delta)$ where

- $f(\delta) = 1 - \frac{\delta}{q-1}$, for $0 < \delta < \frac{q-1}{q}$, and
- $f(\delta) = \delta$ for $\frac{q-1}{q} \leq \delta < 1$.

Thus the minimum of $f(\delta)$ for $\delta \in (0,1)$ is met at $\delta = \frac{q-1}{q}$, and equals $\frac{q-1}{q}$. □

## 5. Naive version of UP

We prove here a finite field version of a result due to Donoho et al. [3] in characteristic zero. Throughout this section we assume $(n,q) = 1$ and we let $\zeta$ denote a primitive root of unity of order $n$ in the algebraic closure of $\mathbb{F}_q$.

**Proposition 2** (Naive UP)**.** *For any $f \neq 0$ in $\mathbb{F}_q^n$,*

$$w_H(f) \cdot w_H(\hat{f}) \geq n,$$

*where $\hat{f} := (f(\zeta), f(\zeta^2), \ldots, f(\zeta^n))$ is the vectorial version of the Mattson-Solomon polynomial.*

*Proof.* Let $w := w_H(f)$. By the BCH bound, $\hat{f}$ cannot have $w$ consecutive zeros.

Suppose first that $w$ divides $n$. Partition the set $\{1, \ldots, n\}$ into $n/w$ intervals of consecutive indices of length $w$. In each of these intervals there is at least one index where $\hat{f}$ is nonzero. Thus, we have exhibited $n/w$ nonzeros of $\hat{f}$. The desired inequality follows in that case.

Equality holds if $\hat{f}$ has exactly one nonzero for each interval. Moreover, these nonzeros must be equally spaced, since otherwise there would be more than $w$ consecutive zeros between some pairs of nonzero elements of $\hat{f}$.

If $w$ does not divide $n$, then there is no way of distributing fewer than $\lceil n/w \rceil$ nonzero elements among $n$ places without leaving a gap of at least $w$ consecutive zeros. Thus $w_H(\hat{f}) \geq \lceil n/w \rceil$. □

**Remark:** The constant $n$ is best possible in view of the example of $f$ equal to the all-one vector (in this case $w_H(\hat{f}) = 1$). Note also that a sharper bound

has been very recently proved in [5] by using van Lint-Wilson bound [9].

This can be reformulated in terms of cyclic codes as follows: for any $f \in R(\mathbb{F}_q, n)$, $f \neq 0$,

$$d(C(f)) \cdot \dim C(f) \geq n.$$

This allows to prove the following result, whose proof is technical and relegated to an appendix.

**Theorem 2.** *For every real number $0 < \alpha < 1/2$ there are sequences of cyclic codes of rate $R$ with minimum distance $\Omega(n^\alpha)$.*

**Remark:** for $R \leq 1/2$, the square root bound on the minimum distance of quadratic residue codes (see for example [10, Chap. 16, Th. 1]) gives an explicit construction of cyclic codes with asymptotic minimum distance bounded below by the square root of the length. However, for $R > 1/2$ our result is the best, to our knowledge.

## 6. Connection with Ramsey Theory

In [11] a connection between the uncertainty problem over finite fields and Ramsey Theory is pointed out. We give here a slight generalization, and a reinterpretation in terms of Coding Theory. We require a pair of definitions.

**Definition 3.** *An* **arithmetic progression** *of length $m$ in $\mathbb{Z}/n\mathbb{Z}$ is any subset of the form $\{a + kb \mid k \in \{0, \ldots, m-1\}\}$ with $b \neq 0$.*

**Definition 4.** *The* **Szemerédi function** *$r_m(n)$ is the largest size of a subset of $\mathbb{Z}/n\mathbb{Z}$ not containing an arithmetic progression of length $m$.*

**Proposition 3.** *For $p$ prime such that $(q, p) = 1$, we have*

$$\mu(\mathbb{F}_q, p) \geq \min\{m + p - r_m(p) \mid 1 \leq m \leq p\}.$$

*Proof.* Let $f \in R(\mathbb{F}_q, p)$. If $f$ has weight $m$, then, by BCH bound again, among the zeros of $f$ there cannot be $m$ consecutive powers of an $p$-th primitive root of unity. So $\{i \mid f(\zeta^i) = 0, \zeta^p = 1, \zeta \neq 1\}$ is a subset of $\mathbb{Z}/p\mathbb{Z}$ not containing an arithmetic progression of length $m$. Hence the number of zeros of $f$ is bounded above by $r_m(p)$. Then

$$\begin{aligned}
\mu(\mathbb{F}_q, p) &= \min\{w_H(f) + w_H(\hat{f}) \mid f \in R(\mathbb{F}_q, p), f \neq 0\} \\
&\geq \min\{w_H(f) + p - r_{w_H(f)}(p) \mid f \in R(\mathbb{F}_q, p), f \neq 0\} \\
&\geq \min\{m + p - r_m(p) \mid 1 \leq m \leq p\}.
\end{aligned}$$

□

**Remark:** if $p$ is not prime, Proposition 3 is not true. For example, $\mu(\mathbb{F}_2, 9) = 6$, but $\min\{m + 9 - r_m(9) \mid 1 \leq m \leq 9\} = 8$. This is due to the fact that powers of 9-th primitive root of unity may be 3-rd root of unity.

**Remark:** to our best knowledge, the function $r_m(p)$ is only known for $m$ fixed and $p \to \infty$ [15]. The fact that then $r_m(p) = o(p)$ is the celebrated Szemerédi Theorem. Any result on $r_m(p)$ when $m$ grows proportionally to $p$ with $p \to \infty$ would impact on the UP over finite fields.

We can generalize further by replacing arithmetic progression by their 2D analogues that is to say sets of the shape

$$A(\delta, s) = \{a + kb + rc \mid k \in \{0, \ldots, \delta - 2\} \text{ and } r \in \{0, \ldots, s\}\},$$

for $b, c$ coprime with $n$. Define then the function $r_{\delta,s}(n)$ as the largest size of a subset of $\mathbb{Z}/n\mathbb{Z}$ not containing an $A(\delta, s)$.

**Proposition 4.** *For $p$ prime such that $(q, p) = 1$, we have*

$$\mu(\mathbb{F}_q, p) \geq \min\{\delta + s - 1 + p - r_{\delta,s}(p) \mid \delta \in \{2, \ldots, p\} \text{ and } s \in \{0, \ldots, p - \delta\}\}.$$

*Proof.* The proof is the same as that of Proposition 3, up to the replacement of the BCH bound by the Hartmann-Tzeng bound [10, p. 206]. □

## 7. Conclusion and Open Problems

In reaction to the recent papers [4] and [12], we have considered the uncertainty principle when the Fourier transform takes its values over finite fields. Exploring the connection with MDS codes, we prove that no finite field satisfies the strong version of UP introduced in [4]. The weak version remains conjectural and we prove that it can only hold if $\lambda < \frac{q-1}{q}$. This should not discourage the researchers to try and prove the weak version of UP for some values of $\lambda$ respecting this bound.

The analogue of the DFT UP of Donoho-Stark [3], which we called naive UP, allowed us to construct long cyclic codes of length $n$ and minimum distance $\Omega(n^\alpha)$, where $0 < \alpha < 1/2$. The proof is technical and relegated to an appendix. More suggestively, the arguments used to prove the naive UP

yields an alternative proof of the results of [12], based on the BCH bound on the minimum distance of cyclic codes. A generalization based on the Hartmann-Tzeng bound has been sketched out.

More generally, it would be worthy to generalize all these results to abelian codes.

**Acknowledgement**

The authors are grateful to Alexis Bonnecaze for programming help and to Pieter Moree for the fruitful discussions on the topic.

## References

[1] S. Ball. *On sets of vectors of a finite vector space in which every subset of basis size is a basis*. J. Eur. Math. Soc., **14**, 733–748 (2012).

[2] S. Ball , J. de Beule. *On sets of vectors of a finite vector space in which every subset of basis size is a basis II*. Design, Codes, Crypt., **65**, 5–14 (2012).

[3] D.L. Donoho, P.B. Stark. *Uncertainty principles, and signal recovery*. SIAM J. Appl. Math., **49**, 906–931 (1989).

[4] S. Evra, E. Kowalski, A. Lubotzky. *Good cyclic codes and the uncertainty principle*. Enseign. Math., **63**, 305–332 (2017).

[5] T. Feng, H.D. Hollmann, Q. Xiang. *The shift bound for abelian codes and generalizations of the Donoho-Stark uncertainty principle*. IEEE Trans. Inform. Theory, **65(8)**, 4673–4682 (2019).

[6] D. Goldstein, R.M. Guralnick, I.M. Isaacs. *Inequalities for finite group permutation modules*. Trans. Amer. Math. Soc., **357**, 4017–4042 (2005).

[7] W. Heisenberg, W. *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*. Zeitschrift für Physik , **43(3–4)**, 172–198 (1927).

[8] W.C. Huffman, V. Pless. *Fundamentals of Error Correcting Codes*, Cambridge University Press (2003).

[9] J. H. van Lint, R. M. Wilson. *On the minimum distance of cyclic codes*. IEEE Trans. Inform. Theory, **32**, 23–40 (1986).

[10] F.J. MacWilliams, N.J. A. Sloane. *The theory of Error Correcting Codes*. North-Holland, Amsterdam (1981).

[11] S. Quader, A. Russell, R. Sundaram. *Small-Support Uncertainty Principles on Z/p over Finite Fields*. `https://arxiv.org/abs/1906.05179` (2019).

[12] M. Shi, R. Wu, P. Solé. *Asymptotically Good Additive Cyclic Codes Exist*. IEEE Communications Letters, **22(10)**: 1980–1983 (2018).

[13] C. Martinez-Perez, W. Willems. *Is the class of cyclic codes asymptotically good?*. IEEE Trans. Inform. Theory, **52(2)**, 696–700 (2006).

[14] T. Tao. *An uncertainty principle for cyclic groups of prime order*, Math. Res. Lett., **12**, 121–127 (2005).

[15] T. Tao and Van H Vu. *A*dditive combinatorics. volume 105. Cambridge University Press (2006).

[16] A. Widgerson, Y. Widgerson. *The uncertainty principle: variations on a theme*. Bull. Amer. Math. Soc. (2021).

## Appendix: Proof of Theorem 2

We construct a sequence of $q$-ary cyclic codes of rate $0 < R < 1$, and designed minimum distance. Let $p$ be an arbitrary prime, and write $n = q^p - 1$. If $x$ is an indeterminate then, from finite field theory [10, Chap. 4, Th. 10], we know that

$$x^n - 1 = \prod_{a \in \mathbb{F}_q^*} (x - a) \cdot \prod_{i=1}^{s} f_i,$$

where $f_i$ runs over all irreducible polynomials in $x$ of degree $p$, and where $n = q - 1 + sp$. Let $g_I = \prod_{i \in I} f_i$ with $|I| = s' = \lfloor s(1 - R) \rfloor$. Then the dimension of the cyclic code $C(g_I)$ of generator $g$ is $n - ps'$, and it can be checked that

$$\frac{n - s'p}{n} \to R$$

when $p \to \infty$.

We need to control the intersections of the $C(g_I)$'s when $I$ varies.

**Lemma 1.** *Let $r \neq 0$ be an arbitrary vector in $\mathbb{F}_q^n$, of Hamming weight $\leq n^\alpha$ for some $0 < \alpha < 1$. There at most $\Lambda_n$ polynomials $g_I$ with $|I| = s'$ such that $r \in C(g_I)$, where $\Lambda_n = 2^{\left(\frac{n-n^{1-\alpha}}{p}\right)H(R)}$.*

*Proof.* The number $N$ of indexes $I$ such that $r \in C(g_I)$, equals the number of $s'$-sets of $f_i$'s which divide $r(x)$ in polynomial notation.

Let $\zeta$ be a primitive $n$-th root of unity in the algebraic closure $\mathbb{K}$ of $\mathbb{F}_q$, and $\widehat{r} = (r(\zeta), r(\zeta^2), \ldots, r(\zeta^n))$. We have

$$N \leq \binom{\lfloor \frac{Z(r)}{p} \rfloor}{s'},$$

where

$$Z(r) := |\{\omega \in \mathbb{K} \mid \omega^n = 1 \text{ and } r(\omega) = 0\}| = n - w_H(\widehat{r}).$$

By Proposition 2, $w_H(\widehat{r}) \geq n^{1-\alpha}$, so that $Z(r) \leq n - n^{1-\alpha}$. Thus

$$N \leq \binom{\left\lfloor \frac{n-n^{1-\alpha}}{p} \right\rfloor}{s'} \leq 2^{\left(\frac{n-n^{1-\alpha}}{p}\right)H(R)},$$

where the upper bound is a consequence of [10, Chap. 10, Lemma 8]. □

*Proof of Theorem 2.* The number of possible $g_I$'s is

$$\binom{s}{s'} \sim \frac{2^{sH(R)}}{\sqrt{2\pi sR(1-R)}}$$

for $s \to \infty$, by Stirling's approximation of the factorial.

If this number is greater than the product of $\Lambda_n$ by the volume of the Hamming ball of radius $n^\alpha$ in length $n$, then there are codes $C(g_I)$ with minimum distance greater than $n^\alpha$.

The volume of the Hamming ball of radius $n^\alpha$ is bounded above by

$$(1 + \lfloor n^\alpha \rfloor)\binom{n}{\lfloor n^\alpha \rfloor}(q-1)^{\lfloor n^\alpha \rfloor}$$

(see the proof of [8, Lemma 2.10.3]), which is bounded above by a quantity asymptotically equivalent to

$$\frac{2^{nH(n^{\alpha-1})+\log_2(n^\alpha)+n^\alpha\log_2(q-1)}}{\sqrt{2\pi n^\alpha(1-n^{\alpha-1})}}$$

$$\sim \frac{1}{\sqrt{2\pi}}\cdot n^{-n^\alpha(\alpha-1)+\frac{\alpha}{2}}\cdot e^{n^\alpha-n^{2\alpha-1}}\cdot(q-1)^{n^\alpha}$$

for $n\to\infty$. So, applying Lemma 1, the mentioned inequality happens if

$$2^{\left(\frac{-(q^p-1)^{1-\alpha}+(q-1)}{p}\right)H(R)}\cdot(q^p-1)^{-(q^p-1)^\alpha(\alpha-1)+\frac{\alpha}{2}}\cdot$$

$$\cdot e^{(q^p-1)^\alpha-(q^p-1)^{2\alpha-1}}\cdot(q-1)^{(q^p-1)^\alpha}\cdot\left(\frac{q^p-q}{p}\right)^{1/2}\leq\frac{1}{\sqrt{R(1-R)}}.$$

We can write the last inequality as $e^{f_{\alpha,q,R}(p)}\leq\frac{1}{\sqrt{R(1-R)}}$, with

$$f_{\alpha,q,R}(p)=(1-\alpha)\ln(q^p-1)(q^p-1)^\alpha-\ln(2)H(R)\frac{q^p}{p(q^p-1)^\alpha}+o(pq^{\alpha p})$$

for $p\to\infty$, so that $f_{\alpha,q,R}(p)\to-\infty$ for $p\to\infty$ if $\alpha<1/2$, and it grows to $\infty$ otherwise. □