Goppa's conjecture and counting points on hyperelliptic curves

David Joyner* and Greg Coy

6-4-2006

A long standing problem has been to develop "good" binary linear codes to be used for error-correction. We show in this paper that the Goppa conjecture regarding "good" binary codes is incompatible with a conjecture on the number of points of hyperelliptic curves over finite fields of odd prime order. This rest of this introduction is devoted to explaining the precise result.

Let C denote a binary block code. The length of C is denoted n and the smallest weight of any non-zero codeword, is denoted d. (Note d is equal to the minimum distance of C, though this is generally not true if C is a non-linear code.) If C is, in addition, linear the dimension of C is denoted k and, in the usual terminology, we say $C \subset \mathbb{F}^n$ is a $[n, k, d]_2$-code, where $\mathbb{F} = GF(2)$ is the field with two elements. For any two $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, let $d(\mathbf{x}, \mathbf{y})$ denote the *Hamming metric*:

$$d(\mathbf{x}, \mathbf{y}) = |\{1 \leq i \leq n \mid x_i \neq y_i\}|. \tag{1}$$

The *weight* w of $\mathbf{v}$ is the number of non-zero entries of $\mathbf{v}$.

We will construct an interesting family of non-linear binary codes C_i which have the property that $\frac{\log_2(|C_i|)}{n}$ has a limit $= \frac{1}{4}$ and $\frac{d}{n}$ has a limit $\geq \frac{1}{4}$. This family of non-linear codes forms the codes which we call long quardatic residue codes (defined precisely later). They are motivated by a clever construction of Bazzi-Mitter [BM].

Denoting the volume of a Hamming sphere of radius r in $\mathbb{F}^n$ by $V(n, r)$, the binary version of the *Gilbert-Varshamov bound* asserts that there is an $[n, k, d]_2$ code C satisfying $k \geq \log_2(\frac{2^n}{V(n,d-1)})$ elements [HP].

Conjecture 1 *(Goppa's conjecture [JV],[G]) The binary version of the Gilbert-Varshamov bound is asymptotically exact.*

A *hyperelliptic curve* X over $GF(p)$ is a polynomial equation of the form $y^2 = h(x)$, where $h(x)$ is a polynomial with coefficients in $GF(p)$ with distinct roots. The number of solutions to $y^2 = h(x)$ mod p, $|X(GF(p))|$, can be related to a character sum (see Remark 1 below), thanks to classical work of Artin, Hasse, and Weil. This

*Department of Mathematics, United States Naval Academy, wdj@usna.edu

formula yields good estimates for $|X(GF(p))|$ in many cases (especially when p is large compared to the degree of h). A long standing problem has been to improve on the trivial estimate when p is small compared to the degree of h. We are interested in a special type of hyperelliptic: For each non-empty subset $S \subset GF(p)$, consider the hyperelliptic curve X_S defined by $y^2 = f_S(x)$ where $f_S(x) = \prod_{a \in S}(x-a)$. Let $|X_S(GF(p))|$ denote the cardinality of the set of all (x, y) satisfying $y^2 = f_S(x)$ plus the number of points at infinity on X_S.

Conjecture 2 *("small cardinality conjecture") For an infinite number of primes p for which $p \equiv 1 \pmod 4$, the number of $GF(p)$-rational points on X_S satisfies, for all subsets $S \subset GF(p)$, the bound $|X_S(GF(p))| \leq 1.57p$.*

This paper proves, using the long quadratic residue codes constructed below, that both of these conjectures cannot be true. We close this introduction with a few open questions which, on the basis of this result, seem natural.

Question 1: For each $p \equiv 1 \pmod 4$ is there an effectively computable subset $S \subset GF(p)$ such that $|X_S(GF(p))| > \frac{3}{2}p$?

Question 2: Does there exist a $c < 2$ such that, for all $p \equiv 1 \pmod 4$ and all $S \subset GF(p)$, we have $|X_S(GF(p))| < c \cdot p$?

(A weaker version of this second question was conjectured in [BM].)

In the direction of these questions, we establish (using a coding theory bound of McEliese-Rumsey-Rodemich-Welsh) the following weaker result: *There exists a constant p_0 having the following property: if $p \equiv 1 \pmod 4$ and $p > p_0$ then there exists a subset $S \subset GF(p)$ for which the bound $|X_S(GF(p))| > c_0 p$ holds.*[1] Here $c_0 = 2 \cdot (1 - h^{-1}(1/4)) \cong 1.399$, where $h(t) = H(\frac{1}{2} - \sqrt{t(1-t)})$ and $H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$.

F. Voloch [V] has kindly allowed the author to include some interesting explicit constructions along these lines at the end of this paper. In particular, his result answers the second question above in the negative. The first question is, as far as we know, still open.

1 Quasi-Quadratic Residue Codes

These are some observations on the interesting paper by Bazzi and Mitter [BM].

If $S \subseteq GF(p)$, let $f_S(x) = \prod_{a \in S}(x-a) \in GF(p)[x]$. Let χ be the quadratic residue character, which is 1 on the set Q quadratic residues in $GF(p)^\times$, -1 on the set N non-quadratic residues, and is 0 on $0 \in GF(p)$. Let $R = \mathbb{F}[x]/(x^p - 1)$ and $r_S \in R$ denotes the polynomial

$$r_S(x) = \sum_{i \in S} x^i,$$

where $S \subseteq GF(p)$. (Note that $r_S^2 = r_{2S}$, where $2S$ is the set of elements $2s \in GF(p)$, for $s \in S$. In particular, since $Q \subseteq GF(p)^\times$ is a subgroup, $r_Q^2 = r_Q$ if and only if

[1] Unfortunately, the method of proof gives no clue of how to compute p_0 or S.

$2 \in Q$ if and only if $p \equiv \pm 1 \pmod 8$ (by the quadratic reciprocity law). Moreover, if $2 \in N$ then $r_Q^2 = r_N$.)

Define the *QQR code* as

$$C_{NQ} = \{(r_N r_S, r_Q r_S) \mid S \subseteq GF(p)\},$$

where N, Q are as above.

These are binary codes of length $2p$ and dimension p (if $p \equiv 3 \pmod 4$) or $p-1$ (if $p \equiv 1 \pmod 4$).

Lemma 1 *(Bazzi-Mitter [BM], Proposition 3.3) Assume* 2 *and* -1 *are non-quadratic residues mod* p *(i.e.* $p \equiv 3 \pmod 8$*).*

If $c = (r_N r_S, r_Q r_S)$ *is a nonzero codeword of the* $[2p, p]$ *binary code* C_{NQ} *then the weight of this codeword can be expressed in terms of a character sum as*

$$wt(c) = p - \sum_{a \in GF(p)} \chi(f_S(a)),$$

if $|S|$ *is even, and*

$$wt(c) = p + \sum_{a \in GF(p)} \chi(f_{GF(p)-S}(a)),$$

if $|S|$ *is odd.*

In fact, looking carefully at their proof, one finds the following result:

Lemma 2 *Let* $c = (r_N r_S, r_Q r_S)$ *be a nonzero codeword of* C_{NQ}*.*

(a) If $|S|$ *is even*

$$wt(c) = p - \sum_{a \in GF(p)} \chi(f_S(a)).$$

(b) If $|S|$ *is odd and* $p \equiv 1 \pmod 4$ *then the weight is*

$$wt(c) = p - \sum_{a \in GF(p)} \chi(f_{GF(p)-S}(a)).$$

(c) If $p \equiv 3 \pmod 4$ *and* $|S|$ *odd both hold then*

$$wt(c) = p + \sum_{a \in GF(p)} \chi(f_{GF(p)-S}(a)).$$

Proof: If $A, B \subseteq GF(p)$ then we claim

$$wt(r_A r_B) = \sum_{k \in GF(p)} \text{parity } |A \cap (k - B)|, \tag{2}$$

where $k - B = \{k - b \mid b \in B\}$. In general it is true that $(\sum_\ell a_\ell x^\ell)(\sum_m b_m x^m) = \sum_n c_n x^n, \;\; c_n = \sum_{\ell+m=n} a_\ell b_m$. Indeed,

$$r_A(x) r_B(x) = \sum_{k \in GF(p)} n_k x^k,$$

where n_k counts the $x \in A$, $y \in B$ such that $x + y = k$. Now

$$n_k \equiv \sum_{x \in A \cap (k-B)} 1 = |A \cap (k - B)| \pmod 2,$$

so (2) is true.

Let $S \subseteq GF(p)$, then we have

$$p - wt(r_Q r_S) - wt(r_N r_S) = \sum_{a \in GF(p)} 1 - \text{ parity } |Q \cap (a - S)| - \text{ parity } |N \cap (a - S)|.$$

Let

$$T_a(S) = 1 - \text{ parity } |Q \cap (a - S)| - \text{ parity } |N \cap (a - S)|.$$

Case 1. If $|S|$ is even and $a \in S$ then $0 \in a - S$ so $|Q \cap (a - S)|$ odd implies that $|N \cap (a - S)|$ is even, since 0 is not included in $Q \cap (a - S)$ or $N \cap (a - S)$. Likewise, $|Q \cap (a - S)|$ even implies that $|N \cap (a - S)|$ is odd. Therefore $T_a(S) = 0$.

Case 2. If $|S|$ is even and $a \notin S$ then parity $|Q \cap (a - S)|$ =parity$|N \cap (a - S)|$. If $|Q \cap (a - S)|$ is even then $T_a(S) = 1$ and if $|Q \cap (a - S)|$ is odd then $T_a(S) = -1$.

Case 3. $|S|$ is odd. We claim that $(a - S)^c = a - S^c$. (Proof: Let $s \in S$ and $\bar{s} \in S^c$. Then $a - s = a - \bar{s} \implies s = \bar{s}$, which is obviously a contradiction. Therefore $(a - S) \cap (a - S^c) = \emptyset$, so $(a - S)^c \supseteq (a - S^c)$. Replace S by S^c to prove the claim.) Also note that

$$Q \cap (a - S) \sqcup Q \cap (a - S^c) = GF(p) \cap Q = Q$$

has $|Q| = \frac{p-1}{2}$ elements ($\sqcup$ denotes disjoint unon). So

$$\text{parity } |Q \cap (a - S)| = \text{parity } |Q \cap (a - S^c)|$$

if and only if $|Q|$ is even and

$$\text{parity } |Q \cap (a - S)| \neq \text{parity } |Q \cap (a - S^c)|$$

if and only if and only if $|Q|$ is odd.

Conclusion.

$$|S| \text{ even}: \; T_a(S) = \prod_{x \in a - S} \left(\frac{x}{p}\right)$$

$$|S| \text{ odd and } p \equiv 3 \pmod 4\colon\ T_a(S) = -T_a(S^c)$$
$$|S| \text{ odd and } p \equiv 1 \pmod 4\colon\ T_a(S) = T_a(S^c)$$

From which the lemma follows. □

Remark 1

- *$|S|$ even: The $\sum_{a\in GF(p)} \chi(f_S(a))$ is equal to $-p-2$ plus the number of $GF(p)$-rational points on the (smooth projective model of the) hyperelliptic curve $X_S:\ y^2 = f_S(x)$. In other words,*

$$\sum_{a\in GF(p)} \chi(f_S(a)) = -p-2+|X_S(GF(p))|.$$

- *$|S|$ odd: The $\sum_{a\in GF(p)} \chi(f_S(a))$ is equal to $-p-1$ plus the number of $GF(p)$-rational points on the (smooth projective model of the) hyperelliptic curve X_S : $y^2 = f_S(x)$. In other words,*

$$\sum_{a\in GF(p)} \chi(f_S(a)) = -p-1+|X_S(GF(p))|.$$

- *The genus of the (smooth projective model of the) curve $y^2 = f_S(x)$, when $|S|$ is even, is $\frac{|S|-2}{2}$. Therefore, Weil's estimate gives in this case*

$$|\sum_{a\in GF(p)} \chi(f_S(a))| \le (|S|-2)q^{1/2}+1,$$

which is trivial if $|S| > q^{1/2}$. (The estimate when S is odd is similar.)

2 Long Quadratic Residue Codes

We now introduce a new code, constructed similarly to the QQR codes discussed above:

$$C = \{(r_N r_S, r_Q r_S, r_N r_{S^c}, r_Q r_{S^c}) \mid S \subseteq GF(p)\}$$

where, for p prime and $T \subseteq GF(p)$,

$$r_T(x) = \sum_{i\in T} x^i.$$

Observe that this code is *non-linear* with respect to the usual coordinate-wise addition.

For any $S \subseteq GF(p)$, let

$$c_S = (r_N r_S, r_Q r_S, r_N r_{S^c}, r_Q r_{S^c})$$

and let

$$v_S = (r_N r_S, r_Q r_S, r_N r_S, r_Q r_S).$$

If $S_1 \Delta S_2$ denotes the symmetric difference between S_1 and S_2 then it is easy to check that

$$c_{S_1} + c_{S_2} = v_{S_1 \Delta S_2}. \tag{3}$$

We know that

$$wt(r_N r_S, r_Q r_S) = \begin{cases} p - \sum_{a \in GF(p)} \left(\frac{f_S(a)}{p}\right), & |S| \text{ even (any } p), \\ p - \sum_{a \in GF(p)} \left(\frac{f_{S^c}(a)}{p}\right), & |S| \text{ odd and } p \equiv 1 \pmod 4, \\ p + \sum_{a \in GF(p)} \left(\frac{f_{S^c}(a)}{p}\right), & |S| \text{ odd and } p \equiv 3 \pmod 4, \end{cases}$$

by Lemma 2.

If $p \equiv 1 \pmod 4$ then

$$\begin{aligned} wt\left(r_N r_S, r_Q r_S, r_N r_{S^c}, r_Q r_{S^c}\right) &= wt\left(r_N r_S, r_Q r_S\right) + wt\left(r_N r_{S^c}, r_Q r_{S^c}\right) \\ &= p - \sum_{a \in GF(p)} \left(\frac{f_S(a)}{p}\right) + p - \sum_{a \in GF(p)} \left(\frac{f_{S^c}(a)}{p}\right) \\ &= 2p - \sum_{a \in GF(p)} \left[\left(\frac{f_S(a)}{p}\right) + \left(\frac{f_{S^c}(a)}{p}\right)\right]. \end{aligned} \tag{4}$$

We have the following trivial estimates:

$$|\sum_{a \in GF(p)} \left(\frac{f_S(a)}{p}\right)| \leq |S^c|$$

and

$$|\sum_{a \in GF(p)} \left(\frac{f_{S^c}(a)}{p}\right)| \leq |S|,$$

therefore

$$|\sum_{a \in GF(p)} \left[\left(\frac{f_S(a)}{p}\right) + \left(\frac{f_{S^c}(a)}{p}\right)\right]| \leq |S^c| + |S| = p.$$

This implies the minimum non-zero weight ρ of C satisfies $\rho \geq p$, when $p \equiv 1 \pmod 4$.

We now compute the size of C. We now prove the *claim*: if $p \equiv 3 \pmod 4$ then the map that sends S to the codeword $(r_N r_S, r_Q r_S, r_N r_{S^c}, r_Q r_{S^c})$ is injective. This implies $|C| = 2^p$. Suppose not, then there are two subsets $S_1,\ S_2 \subseteq GF(p)$ that are mapped to the same codeword. Subtracting, the subset $T = S_1 \Delta S_2$ satisfies $r_Q r_T = r_N r_T = r_Q r_{T^c} = r_N r_{T^c} = 0$. If $|T|$ is even then $0 = (r_Q + r_N) r_T = (r_{GF(p)} - 1) r_T = r_T$. This forces T to be the empty set, so $S_1 = S_2$. Now if $|T|$ is odd then similar reasoning implies that T^c is the empty set. Therefore, $S_1 = \emptyset$ and $S_2 = GF(p)$ or vice versa. This proves the claim.

In case $p \equiv 1 \pmod 4$, *claim*: $|C| = 2^{p-1}$. Again, suppose there are two subsets $S_1,\ S_2 \subseteq GF(p)$ that are mapped to the same codeword. Then the subset $T = S_1 \Delta S_2$

for which $r_Q r_T = r_N r_T = r_Q r_{T^c} = r_N r_{T^c} = 0$. This implies either $T = \emptyset$ or $T = GF(p)$. Therefore, either $S_1 = S_2$ or $S_1 = S_2^c$.

We have proven the following result.

Theorem 1 *The non-linear code C has length $n = 4p$ and minimum non-zero weight $\rho \geq p$. It has size $M = 2^{p-1}$ if $p \equiv 1 \pmod 4$, and size $M = 2^p$ if $p \equiv 3 \pmod 4$.*

2.1 Duality in the LQR code

First, a few observations.

For any $T \subseteq GF(p)$, let $\overline{T}$ denote the set which is $T \cup \{0\}$ if $0 \notin T$, and $T - \{0\}$ if $0 \in T$.

By (3), it suffices to find the minimum non-zero weight of v_S, $S \subseteq GF(p)$. We first find a "duality" relation between v_S and v_{S^c}, and c_S and c_{S^c}.

By (4), it is obvious that if $p \equiv 1 \pmod 4$ then $v_S = v_{S^c}$.

If $p \equiv 1 \pmod 4$ then $r_{GF(p)} r_Q = r_{GF(p)} r_N = 0$, so

$$v_{GF(p)} = (0, 0, 0, 0) \in R_p^4 = \mathbb{F}^n.$$

If $p \equiv 3 \pmod 4$ then $r_{GF(p)} r_Q = r_Q$, $r_{GF(p)} r_N = r_N$, so

$$v_{GF(p)} = (r_N, r_Q, r_N, r_Q).$$

These facts together with (3) imply: if $p \equiv 1 \pmod 4$ then $c_S = c_{S^c}$; if $p \equiv 3 \pmod 4$ then $c_S = c_{S^c} + (r_N, r_Q, r_N, r_Q)$. Consequently, if $p \equiv 3 \pmod 4$ then

$$\begin{aligned} v_{S_1 \Delta S_2} &= c_{S_1} + c_{S_2} \\ &= c_{S_1} + c_{S_2^c} + (r_N, r_Q, r_N, r_Q) \\ &= v_{S_1 \Delta S_2^c} + (r_N, r_Q, r_N, r_Q) = v_{(S_1 \Delta S_2)^c} + (r_N, r_Q, r_N, r_Q). \end{aligned}$$

In particular, if $p \equiv 3 \pmod 4$ then $v_S = v_{S^c} + (r_N, r_Q, r_N, r_Q)$, and this can be more compactly re-expressed as

$$v_S = v_{\overline{S^c}}, \tag{5}$$

provided $p \equiv 3 \pmod 4$.

By (3), we have

$$c_S = c_{S^c}, \tag{6}$$

provided $p \equiv 1 \pmod 4$. Now that we know this, we can write, if $p \equiv 1 \pmod 4$,

$$v_{S_1 \Delta S_2} = c_{S_1} + c_{S_2} = c_{S_1} + c_{S_2^c} = v_{S_1 \Delta S_2^c} = v_{(S_1 \Delta S_2)^c}.$$

In particular,

$$v_S = v_{S^c}, \tag{7}$$

provided $p \equiv 1 \pmod 4$. This is also a consequence of (3) and (6). (So now we have at least three proofs of this fact!)

2.2 Linear version of the LQR code

Let us denote the map $S \longmapsto c_S$ by ϕ and its inverse (which only exists if $p \equiv 3 \pmod 4$) by ψ.

When $p \equiv 3 \pmod 4$, define "addition" $\oplus$ on C by

$$c_{S_1} \oplus c_{S_2} = c_{S_1 \Delta S_2}, \tag{8}$$

for arbitrary subsets of $GF(p)$. It is easy to check that this operation $\oplus$ is well-defined in case $p \equiv 3 \pmod 4$.

The surprising fact is the following result.

Lemma 3 *In case $p \equiv 1 \pmod 4$, (8) is well-defined.*

Proof: Assume $p \equiv 1 \pmod 4$. Recall from the above discussion that ϕ is a 2-to-1 map from the set $2^{GF(p)}$ of subsets of $GF(p)$ to the codewords of C. For all $c \in C$, there is a $S \subseteq GF(p)$ and $c = c_T$ (for $T \subseteq GF(p)$) if and only if $T = S$ or $T = S^c$. We know $c_{S_1^c \Delta S_2} = c_{S_1 \Delta S_2^c} = c_{S_1^c \Delta S_2^c} = c_{(S_1 \Delta S_2)^c}$. This implies $\oplus$ does not depend on the choice in $\phi^{-1}(c_{S_1}) = \{S_1, S_1^c\}$, $\phi^{-1}(c_{S_2}) = \{S_2, S_2^c\}$ made to compute the right-hand side of (8). $\square$

This operation $\oplus$ is commutative, since the symmetric difference operation Δ is symmetric, and it is associative since Δ is associative. Each element is the inverse of itself and the element $c_\emptyset = 0$ is the identity.

Therefore, C is a vector space over $\mathbb{F}$ with the operation $\oplus$. When $p \equiv 3 \pmod 4$, the set of codewords c_S, S a singleton subset of $GF(p)$, forms a basis. When $p \equiv 1 \pmod 4$, the set of codewords c_S, S a singleton subset of $GF(p)$, is linearly dependent (their sum is 0), so do not form a basis. However, if you just omit one (any one – pick your least favorite), you get a basis.

Define the metric $d_\oplus$ as follows. For $c_1, c_2 \in C$, let

$$d_\oplus(c_1, c_2) = wt(c_1 \oplus c_2).$$

Of course, the Hamming metric, denoted d_H to be unambiguous, satisfies $d_H(c_1, c_2) = wt(c_1 + c_2)$.

We make a few remarks comparing $d_\oplus$ to d_H. If $c \in C \cong R_p^4$ is written $c = (p_1(c), p_2(c), p_3(c), p_4(c))$, for polynomials $p_i(c) \in R_p$, then

$$wt(c_1 \oplus c_2) = wt(p_1(c_1 \oplus c_2)) + wt(p_2(c_1 \oplus c_2)) + wt(p_3(c_1 \oplus c_2)) + wt(p_4(c_1 \oplus c_2)),$$

and

$$wt(c_1 + c_2) = wt(p_1(c_1 + c_2)) + wt(p_2(c_1 + c_2)) + wt(p_3(c_1 + c_2)) + wt(p_4(c_1 + c_2)).$$

By definition of C,

$$wt(p_1(c_1 \oplus c_2)) = wt(p_1(c_1 + c_2)), \quad wt(p_2(c_1 \oplus c_2)) = +wt(p_2(c_1 + c_2)).$$

On the other hand, if $c_1 = c_{S_1}$, $c_2 = c_{S_2}$,

$$wt(p_3(c_1 \oplus c_2)) + wt(p_4(c_1 \oplus c_2)) = wt(r_N r_{(S_1 \Delta S_2)^c}) + wt(r_Q r_{(S_1 \Delta S_2)^c})$$

and

$$wt(p_3(c_1 + c_2)) + wt(p_4(c_1 + c_2)) = wt(r_N r_{S_1 \Delta S_2}) + wt(r_Q r_{S_1 \Delta S_2}).$$

The difference between $d_\oplus(c_1, c_2)$ and $d_H(c_1, c_2)$ can now by computed using Lemma 2.

We assert that, with this notion of addition, the "$\oplus$-Hamming metric" on $(C, \oplus)$ is the more "natural" one to use.

This and the previous section prove the following result.

Theorem 2 *The linear code $(C, \oplus)$ has length $n = 4p$ and minimum $\oplus$-distance $d \geq p$. It has size $M = 2^{p-1}$ if $p \equiv 1 \pmod 4$, and size $M = 2^p$ if $p \equiv 3 \pmod 4$.*

2.3 Connection with hyperelliptic curves

The goal of this section is to prove the main result of this paper.

Theorem 3 *If the small cardinality conjecture is true then Goppa's conjecture is false.*

Proof: Recall Goppa's conjecture is that the binary asymptotic Gilbert-Varshamov bound is best possible for any family of binary codes. The asymptotic GV bound states that the rate R is greater than or equal to $1 - H(\delta)$, where $H(\delta) = \delta - \delta \log_2(\delta) - (1 - \delta) \log_2(1 - \delta)$. According to Goppa's conjecture if $R = \frac{1}{4}$ then the best possible δ is $\delta_0 = .215$. This means that the minimum distance of our long quadratic residue code with rate $R = \frac{1}{4}$ satisfies $d < \delta_0 \cdot 4p = .859p$. Recall that the weight of a codeword c in this LQR code is the weight of the 4-tuple of polynomials $(r_N r_S, r_Q r_S, r_N r_{S^c}, r_Q r_{S^c})$. Let us assume $|S|$ is even for simplicity. Then by Lemma 8 and remark 1 we know that the $wt(r_N r_S, r_Q r_S) = 2p + 2 - |X_S(GF(p))|$. Now suppose $|S|$ is odd and $p \equiv 1 \pmod 4$. Then By lemma 8 and remark 1 we know that the $wt(r_N r_S, r_Q r_S) = 2p + 1 - |X_S(GF(p))|$. By the small cardinality conjecture, $wt(r_N r_S, r_Q r_S) \geq 2p + 1 - |X_S(GF(p))| > 2p - 1.57p = .43p$. Therefore $d = \min_S \; wt((r_N r_S, r_Q r_S, r_N r_{S^c}, r_Q r_{S^c})) \geq 2 \cdot min_S wt(r_N r_S, r_Q r_S) > .86p$. This contradicts the estimate above.□

Using the same argument and the first McEliese-Rumsey-Rodemich-Welsh bound ([HP], Theorem 2.10.6), we prove the following unconditional result.

Theorem 4 *The following statement is false. (*) For an infinite number of primes p for which $p \equiv 1 \pmod 4$, the number of $GF(p)$-rational points on X_S satisfies, for all subsets $S \subset GF(p)$, the bound $|X_S(GF(p))| \leq 1.39p$.*

Corollary 1 *There is a constant p_0 (ineffectively computable) having the following property: if $p \equiv 1 \pmod 4$ and $p > p_0$ then there is a subset $S \subset GF(p)$ for which the bound $|X_S(GF(p))| > 1.39p$ holds.*

Proof: If a prime p satisfies (*) then we shall call it "admissible." We show that the statement (*) contradicts the first asymptotic MRRW bound. Indeed, this MRRW bound states that the rate R is less than or equal to $h(\delta) = H(\frac{1}{2} - \sqrt{\delta(1-\delta)})$. This, and the fact that $R = \frac{1}{4}$ for our long QR codes, imply $\delta \leq \delta_0 = h^{-1}(1/4) \cong 0.3004$. Therefore, for all large p (admissible or not), $d \leq \delta_0 \cdot 4p$. On the other hand, if p is admissible and $|X_S(GF(p))| \leq c \cdot p$ (where $c = 1.39$) then by the above argument, $d \geq 4 \cdot (p - \frac{c}{2}p)$. Together, we obtain $1 - \frac{c}{2} \leq \delta_0$, so $c \geq 2 \cdot (1 - h^{-1}(1/4)) \cong 1.399$. This is a contradiction. □

2.4 Some results of Voloch

When $p \equiv 1 \pmod 8$, the following result improves on the lower bound in the above corollary, and holds, in addition, when $p \equiv 3, 7 \pmod 8$.

Lemma 4 *(Voloch) If $p \equiv 1, 3 \pmod 8$ then $|X_Q(GF(p))| = 1.5p + a$, where Q is the set of quadratic residues and a is a small constant, $-\frac{1}{2} \leq a \leq \frac{5}{2}$.*

A similar bound holds if X_Q is replaced by X_N and $p \equiv 1, 3 \pmod 8$ is replaced by $p \equiv 7 \pmod 8$ (in which case 2 is a quadratic residue).

Proof: From Remark 1, we know that if $p \equiv 3 \pmod 8$ (so $|Q|$ is odd):

$$\sum_{a \in GF(p)} \chi(f_Q(a)) = -p - 1 + |X_Q(GF(p))|.$$

Similarly, if $p \equiv 1 \pmod 8$ (so $|Q|$ is even):

$$\sum_{a \in GF(p)} \chi(f_Q(a)) = -p - 2 + |X_Q(GF(p))|.$$

Since $b^{\frac{p-1}{2}} \equiv \chi(b) \pmod p$, we have

$$x^{\frac{p-1}{2}} - 1 = \prod_{a \in Q} (x - a) = f_Q(x), \quad x^{\frac{p-1}{2}} + 1 = \prod_{a \in N} (x - a).$$

In particular, for all $n \in N$,

$$f_Q(n) = \prod_{a \in Q} (n - a) = n^{\frac{p-1}{2}} - 1 \equiv -2 \pmod p.$$

Since $p \equiv 1, 3 \pmod 8$, we have $\chi(-2) = 1$, so $\chi(f_Q(n)) = 1$ for all $n \in N$. It follows that $|X_Q(GF(p))| = \frac{3}{2}p + \chi(f_Q(0)) + \frac{1}{2}$ ($p \equiv 3 \pmod 8$) or $|X_Q(GF(p))| = \frac{3}{2}p + \chi(f_Q(0)) + \frac{3}{2}$ ($p \equiv 1 \pmod 8$). □

Here is an extension of the idea in the above proof. Fix an integer $\ell > 2$. Assuming ℓ divides $p - 1$, there are distinct ℓ-th roots $r_1 = 1$, r_2, ..., r_ℓ in $GF(p)$ for which $x^{p-1} - 1 = \prod_{i=1}^{\ell} (x^{\frac{p-1}{\ell}} - r_i)$. Also, $x^{\frac{p-1}{\ell}} - 1 = \prod_{a \in P_\ell} (x - a) = f_{P_\ell}(x)$, where P_ℓ denotes the set of non-zero ℓ-th powers in $GF(p)$.

Claim: It is possible to find an infinite sequence of primes p satisfying $p \equiv 1 \pmod{\ell}$ and $\chi(r_i - 1) = 1$, for all $2 \leq i \leq \ell$ (where χ denotes the Legendre character mod p). If the claim is true then we will have a lower bound for $|X_{P_\ell}(GF(p))|$ on the order of $(2 - \frac{1}{\ell})p$, by Remark 1. proof of claim: It's a well-known fact in algebraic number theory that $p \equiv 1 \pmod{\ell}$ implies the prime p splits completely in the cyclotomic field $\mathbb{Q}_\ell$ generated by the ℓ-th roots of unity in $\mathbb{C}$, denoted $\tilde{r}_1 = 1$, $\tilde{r}_2$, ..., $\tilde{r}_\ell$. The condition $\chi(r_i - 1) = 1$ means that p splits in the extension of $\mathbb{Q}_\ell$ obtained by adjoining $\sqrt{\tilde{r}_i - 1}$ (here $i = 2, ..., \ell$). By Chebotarev's density theorem there exists infinitely many such p, as claimed. In fact, there are effective versions which give explicit information on computing such p [LO], [Se].

This, together with the previous lemma, proves the following result.

Theorem 5 *(Voloch) If $n \geq 2$ is any fixed integer then there exist primes p and subsets $S \subset GF(p)$ for which $|X_S(GF(p))| = (2 - \frac{1}{n})p + a$, where a is a small constant, $-\frac{1}{2} \leq a \leq \frac{5}{2}$.*

Acknowledgements: We thank Prof. Amin Shokrollahi of the Ecole Polytechnique Fédérale de Lausanne for many helpful remarks on an earlier version and Prof. Felipe Voloch of the University of Texas for allowing us to include his construction above.

References

[BM] L. Bazzi and S. Mitter, *Some constructions of codes from group actions*, preprint, 2001.

[G] V. D. Goppa, *Bounds for codes*, Dokl. Acad. Nauk. 333(1993)423.

[HP] W. C. Huffman and V. Pless, **Fundamentals of error-correcting codes**, Cambridge Univ. Press, 2003.

[HV] T. Helleseth and J. F. Voloch, *Double circulant quadratic residue codes*, IEEE Transactions in Information Theory, to appear, available on the webpage `http://www.ma.utexas.edu/users/voloch/preprint.html`

[JV] T. Jiang and A. Vardy, *Asymptotic improvement of the Gilbert-Varshamov bound on the size of binary codes*, IEEE Trans Info Theory, to appear.

[LO] J. C. Lagarias and A. M. Odlyzko, *Effective versions of the Chebotarev density theorem*, in **Algebraic number fields (L- functions and Galois theory)** A. Fröhlich, ed., Academic Press 1977, 409-464.

[MS] F. MacWilliams and N. Sloane, **The theory of error-correcting codes**, North-Holland, 1977.

[Se] J.-P. Serre, *Quelques applications du théorème de densité de Chebotarev*, Publications Mathmatiques de l'IHS, 54 (1981), p. 123-201
http://www.numdam.org/item?id=PMIHES_1981__54__123_0

[V] F. Voloch, email communication, 5-2006.