

# The fluctuations in the number of points on a family of curves over a finite field

par MAOSHENG XIONG

**ABSTRACT.** Let  $l \geq 2$  be a positive integer,  $\mathbb{F}_q$  a finite field of cardinality  $q$  with  $q \equiv 1 \pmod{l}$ . In this paper, inspired by [7, 4, 5] and using a slightly different method, we study the fluctuations in the number of  $\mathbb{F}_q$ -points on the curve  $C_F$  given by the affine model  $C_F : Y^l = F(X)$ , where  $F$  is drawn at random uniformly from the set of all monic  $l$ -th power-free polynomials  $F \in \mathbb{F}_q[X]$  of degree  $d$  as  $d \rightarrow \infty$ . The method also enables us to study the fluctuations in the number of  $\mathbb{F}_q$ -points on the same family of curves arising from the set of monic irreducible polynomials.

## 1. Introduction

Given a finite field  $\mathbb{F}_q$  of cardinality  $q$  and a monic square-free polynomial  $F \in \mathbb{F}_q[X]$  of degree  $d \geq 3$ , we get a smooth projective hyperelliptic curve  $C_F$  with the affine model

$$C_F : Y^2 = F(X)$$

having genus  $g = (d - 2)/2$  when  $d$  is even and  $g = (d - 1)/2$  when  $d$  is odd. The number of (affine)  $\mathbb{F}_q$ -points on  $C_F$  can be expressed as  $q + S(F)$ , where  $S(F)$  is the character sum

$$S(F) = \sum_{x \in \mathbb{F}_q} \chi(F(x))$$

and  $\chi$  is the quadratic character of  $\mathbb{F}_q^\times$  (with the convention that  $\chi(0) = 0$ ). In an interesting paper ([7]) Kurlberg and Rudnick investigated the fluctuations in the number of (affine)  $\mathbb{F}_q$ -points on  $C_F$  or more precisely the value of  $S(F)$  when  $F$  is drawn at random from the set of all monic square-free polynomials  $F \in \mathbb{F}_q[X]$  of degree  $d$ . They found that

- (i) For  $q$  fixed and the genus  $g \rightarrow \infty$ ,  $S(F)$  is distributed asymptotically as a sum of  $q$  independent identically distributed (i.i.d.) trinomial random variables  $\{X_i\}_{i=1}^q$ , i.e., random variables taking values in  $0, \pm 1$  with probabilities  $1/(q+1)$ ,  $1/2(1+q^{-1})$  and  $1/2(1+q^{-1})$ , respectively.
- (ii) When both  $g \rightarrow \infty$  and  $q \rightarrow \infty$ ,  $S(F)/\sqrt{q}$  has a Gaussian value distribution with mean zero and variance unity.

These results complement the well-known theorem due to Katz and Sarnak [8, 9], which states that, if the genus  $g$  is fixed and  $q \rightarrow \infty$ , then  $S(F)/\sqrt{q}$  is distributed as the trace of a random matrix in the group  $\mathrm{USp}(2g)$  of  $2g \times 2g$  unitary symplectic matrices. Showing consistency with (ii), if both  $q, g \rightarrow \infty$  with  $q \rightarrow \infty$  first, then  $S(F)/\sqrt{q}$  is distributed as that of the trace of a random matrix in  $\mathrm{USp}(2g)$  as  $g \rightarrow \infty$ , which is known to be a standard Gaussian by a theorem of Diaconis and Shahshahani [6]. Related to this work [7], problems of similar flavor with various arithmetic and geometric applications have been considered before by Larsen [12], Knizhnerman and Sokolinskii [10, 11] and Bergström [1]. Recently, extending the results of Kurlberg and Rudnick [7], Bucur, David, Feigon and Lalín in a series of two beautiful papers [4, 5] successfully obtain interesting results on the distribution of the trace of the Frobenius endomorphism  $\mathrm{Frob}_C$  over moduli spaces of cyclic  $l$ -fold covers of genus  $g$  when  $g \rightarrow \infty$ . Interested reader may refer to their papers [4, 5] for more details and for other results related with the subject.

The proofs of [7, 4, 5] are similar and are based on an ingenious counting argument. The main purpose of this paper is to give a slightly treatment of the proof. We start with the observation that, in writing

$$S(\chi, F) = \sum_{x \in \mathbb{F}_q} \chi_x(F),$$

where  $\chi_x(F) = \chi(F(x))$  for each  $F \in \mathbb{F}_q[X]$ , then  $\chi_x : \mathbb{F}_q[X] \rightarrow \mathbb{C}$  is a Dirichlet character of order  $l$  modulo  $X - x$ . Our strategy is to study the distribution of  $S(\chi, F)$  by manipulating appropriate character sums, which in term can be treated by using various tools such as the Riemann hypothesis for algebraic curves over finite fields [14], the Möbius function and other arithmetic functions. The results of [7, 4, 5] then can be derived directly. Our proofs follow the ideas of [7, 4, 5], however, the properties of character sums will be used in an essential way.

Building upon this idea, let  $l \geq 2$  be any positive integer such that  $q \equiv 1 \pmod{l}$  and denote by  $\mathcal{F}_{d,l} \subset \mathbb{F}_q[X]$  the set of monic  $l$ -th power-free polynomials of degree  $d$ , we investigate the fluctuations in the number of affine  $\mathbb{F}_q$ -points on the curve  $C_F$  given by the affine model

$$(1.1) \quad C_F : Y^l = F(X),$$

where  $F$  is drawn at random uniformly from the set  $\mathcal{F}_{d,l}$ . Denote by  $C_F^0(\mathbb{F}_q)$  the set of affine  $\mathbb{F}_q$ -points on  $C_F$ .

**Theorem 1.1.** (1). If  $q$  is fixed and  $d \rightarrow \infty$ , then as  $F$  ranges over all elements in  $\mathcal{F}_{d,l}$ , the limiting distribution of the value  $\#C_F^0(\mathbb{F}_q) - q$  is that of a sum of  $q$  i.i.d random variables  $\{Y_i\}_{i=1}^q$ , where each  $Y_i$  takes values  $0, -1, l-1$  with probabilities  $\left(1 - \frac{1-q^{-1}}{1-q^{-l}}, \frac{l-1}{l} \frac{1-q^{-1}}{1-q^{-l}}, \frac{1}{l} \frac{1-q^{-1}}{1-q^{-l}}\right)$  respectively.

(2). If  $d, q$  both tend to infinity, then as  $F$  ranges over all elements in  $\mathcal{F}_{d,l}$ , the limiting distribution of the value  $(\#C_F^0(\mathbb{F}_q) - q)/\sqrt{q(l-1)}$  is a standard Gaussian with mean zero and variance one.

If  $q$  is fixed and  $d$  tend to infinity, or  $q$  and  $d$  both tend to infinity in such a way that  $d \geq \frac{q(2l-1)}{l-1}$ , we have a more precise statement.

**Theorem 1.2.** Let the random variables  $\{Y_i\}_{i=1}^q$  be as in Theorem 1.1. Then for any  $s \in \mathbb{Z}$ , we have

$$\frac{\#\{F \in \mathcal{F}_{d,l} : \#C_F^0(\mathbb{F}_q) = q + s\}}{\#\mathcal{F}_{d,l}} = \text{Prob} \left( \sum_{i=1}^q Y_i = s \right) \left( 1 + O \left( 2^q q^{-(1-\frac{1}{l})d + (1-\frac{1}{l})q} \right) \right).$$

One of the benefits of our method is its flexibility: it enables us to consider such statistics for other families of curves whenever similar estimates on the character sums apply. As another example, we study the fluctuations of  $\#C_F^0(\mathbb{F}_q)$  for the same family of curves as  $F$  arises from  $\mathcal{P}_d \subset \mathbb{F}_q[X]$ , the set of monic irreducible polynomials of degree  $d$ .

**Theorem 1.3.** (1). If  $q$  is fixed and  $d \rightarrow \infty$ , then as  $F$  ranges over all elements in  $\mathcal{P}_d$ , the limiting distribution of the value  $\#C_F^0(\mathbb{F}_q) - q$  is that of a sum of  $q$  i.i.d random variables  $\{Y_i\}_{i=1}^q$ , where each  $Y_i$  takes values  $-1, l-1$  with probabilities  $(1 - \frac{1}{l}, \frac{1}{l})$  respectively.

(2). If  $d, q$  both tend to infinity, then as  $F$  ranges over all elements in  $\mathcal{P}_d$ , the limiting distribution of the value  $(\#C_F^0(\mathbb{F}_q) - q)/\sqrt{q(l-1)}$  is a standard Gaussian with mean zero and variance one.

If  $q$  is fixed and  $d$  tend to infinity, or  $q$  and  $d$  both tend to infinity in such a way that  $d \geq 4q$ , we have a more precise statement.

**Theorem 1.4.** Let the random variables  $\{Y_i\}_{i=1}^q$  be as in Theorem 1.3. Then for any  $s \in \mathbb{Z}$ , we have

$$\frac{\#\{F \in \mathcal{P}_d : \#C_F^0(\mathbb{F}_q) = q + s\}}{\#\mathcal{P}_d} = \text{Prob} \left( \sum_{i=1}^q Y_i = s \right) \left( 1 + O \left( 2^q q^{(2q-d)/2} \right) \right).$$

We remark that first, if  $l = 2$ , Theorems 1.1 and 1.2 reduces to (i) and (ii) obtained by Kurlberg and Rudnick mentioned above. For a general  $l$ , Theorems 1.2 and 1.4 are analogous to [5, Theorems 1.1 and 1.4] obtained by Bucur, David, Feigon and Lalín in terms of the Frobenius endomorphism. Moreover, denote by  $C_F(\mathbb{F}_q)$  the set of  $\mathbb{F}_q$ -points on the curve  $C_F$  given in (1.1) (i.e., including the points at infinity). For  $F \in \mathcal{F}_{d,l}$  or  $F \in \mathcal{P}_d$ , we have

$$\#C_F(\mathbb{F}_q) = \#C_F^0(\mathbb{F}_q) + \begin{cases} 1 : & d \not\equiv 0 \pmod{l}, \\ l : & d \equiv 0 \pmod{l}, \end{cases}$$

so Theorems 1.1–1.4 can be translated as statements about the distribution of  $\#C_F(\mathbb{F}_q)$  for  $F \in \mathcal{F}_{d,l}$  and  $F \in \mathcal{P}_d$  as  $d \rightarrow \infty$ , and the results depend

on  $d \equiv 0 \pmod{l}$  or not. It may be interesting know to what happens for these two familes  $\mathcal{F}_{d,l}$  and  $\mathcal{P}_d$  if  $d$  is fixed and  $q$  goes to infinity instead.

In the above theorems and in all results below, the implied constants in the notation “ $O$ ” and “ $\ll$ ” are absolute.

**Acknowledgment.** The author would like to express his gratitude to Prof. Wen-Ching Winnie Li for bringing this problem to his attention. The author also thanks the anonymous referee for many valuable suggestions.

## 2. Preliminaries

In this section we collect several standard results which will be used later. We use Rosen [13] as a general reference.

**2.1.** The (partial) zeta function of the rational function field is

$$(2.1) \quad Z(U) := \prod_P \left(1 - U^{\deg P}\right)^{-1}, \quad |U| < q^{-1},$$

the product over all irreducible monic polynomials (“primes”) in  $\mathbb{F}_q[X]$ . By the fundamental theorem of arithmetic in  $\mathbb{F}_q[X]$ ,  $Z(U)$  can be expressed as a sum over all monic polynomials:

$$Z(U) = \sum_{F \text{ monic}} U^{\deg F},$$

and hence

$$(2.2) \quad Z(U) = (1 - qU)^{-1}.$$

Denote by  $\mathcal{V}_d \subset \mathbb{F}_q[X]$  the set of monic polynomials of degree  $d \geq 0$ . We use the Möbius function to pick out the  $l$ -th power-free polynomials via the formula

$$\sum_{A^l | F} \mu(A) = \begin{cases} 1 & : F \text{ is } l\text{-th power-free} \\ 0 & : \text{otherwise} \end{cases}$$

where we sum over all monic polynomials  $A$  whose  $l$ -th power divides  $F$ . Hence

$$\sum_{d \geq 0} \#\mathcal{F}_{d,l} U^d = \sum_{d \geq 0} \sum_{F \in \mathcal{V}_d} \sum_{A^l | F} \mu(A) U^{\deg F}.$$

Writing  $F = A^l F'$ , we have

$$\begin{aligned} \sum_{d \geq 0} \#\mathcal{F}_{d,l} U^d &= \sum_A \mu(A) U^{l \deg A} \sum_F U^{\deg F} \\ &= \prod_P \left(1 - U^{l \deg P}\right) \prod_P \left(1 - U^{\deg P}\right)^{-1}, \end{aligned}$$

where in the above equations and all results below,  $A, F$  denote monic polynomials and  $P$  is reserved for monic irreducible polynomials.

Using (2.1) and (2.2) we obtain

$$\sum_{d \geq 0} \# \mathcal{F}_{d,l} U^d = Z(U)/Z(U^l) = (1 - qU^l) (1 - qU)^{-1}.$$

Expanding the right hand side as a power series in terms of  $U$  and equating the coefficients on both sides, we find

$$(2.3) \quad \# \mathcal{F}_{d,l} = q^d (1 - q^{1-l}), \quad d \geq l.$$

**Lemma 2.1.** Suppose that  $\psi : \mathbb{F}_q[X] \rightarrow \mathbb{C}$  is a non-trivial Dirichlet character modulo  $h \in \mathbb{F}_q[X]$ . Then for any  $d \geq 1$ ,

$$\left| \sum_{F \in \mathcal{F}_{d,l}} \psi(F) \right| \ll q^{\frac{d}{l} + (1 - \frac{1}{l}) \deg h}.$$

*Proof.* If  $d \leq l$ , then the statement of Lemma 2.1 is trivial. Now suppose  $d \geq l$ . We write

$$\sum_{F \in \mathcal{F}_{d,l}} \psi(F) = \sum_{F \in \mathcal{V}_d} \psi(F) \sum_{A^l | F} \mu(A) = \sum_{\deg A \leq d/l} \mu(A) \psi(A)^l \sum_{\deg F = d - l \deg A} \psi(F),$$

where the sums are over monic polynomials. It is easy to see that

$$\sum_{\deg F = n} \psi(F) = 0, \quad n \geq \deg h,$$

hence we have

$$\begin{aligned} \left| \sum_{F \in \mathcal{F}_{d,l}} \psi(F) \right| &= \left| \sum_{(d+1-\deg h)/l \leq \deg A \leq d/l} \mu(A) \psi(A)^l \sum_{\deg F = d - l \deg A} \psi(F) \right| \\ &\leq \sum_{(d+1-\deg h)/l \leq n \leq d/l} q^{d - ln + n} \ll q^{\frac{d}{l} + (1 - \frac{1}{l}) \deg h}. \end{aligned}$$

This completes the proof of Lemma 2.1.  $\square$

**Lemma 2.2.** Suppose that  $h \in \mathbb{F}_q[X]$  is a polynomial with  $\deg h = m \geq 1$ . Then for any  $d \geq 1$ , we have

$$\sum_{\substack{F \in \mathcal{F}_{d,l} \\ \gcd(F,h)=1}} 1 = q^d (1 - q^{1-l}) \prod_{P|h} \frac{1 - q^{-\deg P}}{1 - q^{-l \deg P}} + O\left(q^{\frac{d}{l} + (1 - \frac{1}{l})m}\right).$$

*Proof.* We may assume that  $d \geq m$ . First we compute

$$\sum_{\substack{F \in \mathcal{V}_d \\ \gcd(F,h)=1}} 1 = \sum_{F \in \mathcal{V}_d} \sum_{D|F, D|h} \mu(D) = \sum_{D|h} \mu(D) \sum_{\deg F = d - \deg D} 1.$$

This in turn gives

$$(2.4) \quad \sum_{\substack{F \in \mathcal{V}_d \\ \gcd(F, h)=1}} 1 = \sum_{D|h} \mu(D) q^{d-\deg D} = q^d \prod_{P|h} (1 - q^{-\deg P}).$$

Next

$$\sum_{\substack{F \in \mathcal{F}_{d,l} \\ \gcd(F, h)=1}} 1 = \sum_{\substack{F \in \mathcal{V}_d \\ \gcd(F, h)=1}} \sum_{A^l | F} \mu(A) = \sum_{\substack{\deg A \leq d/l \\ \gcd(A, h)=1}} \mu(A) \sum_{\substack{\deg Q = d-l \deg A \\ \gcd(Q, h)=1}} 1.$$

We find that

$$I_2 = \sum_{\substack{\deg A > (d-m)/l \\ \gcd(A, h)=1}} \sum_{\substack{\deg Q = d-l \deg A \\ \gcd(Q, h)=1}} 1 \leq \sum_{n > (d-m)/l} q^n q^{d-ln} \ll q^{\frac{d}{l} + (1-\frac{1}{l})m}.$$

On the other hand, using (2.4) we have

$$I_1 = \sum_{\substack{\deg A \leq (d-m)/l \\ \gcd(A, h)=1}} \mu(A) \sum_{\substack{\deg Q = d-l \deg A \\ \gcd(Q, h)=1}} 1 = \sum_{\substack{\deg A \leq (d-m)/l \\ \gcd(A, h)=1}} \mu(A) q^{d-l \deg A} \prod_{P|h} (1 - q^{-\deg P}),$$

and this gives us

$$I_1 = \sum_{\substack{A \\ \gcd(A, h)=1}} \mu(A) q^{d-l \deg A} \prod_{P|h} (1 - q^{-\deg P}) + O(I_2).$$

The main term can be rewritten as

$$q^d \prod_{P|h} (1 - q^{-\deg P}) \prod_{\gcd(P, h)=1} (1 - q^{-l \deg P}),$$

which is

$$q^d \prod_{P|h} (1 - q^{-\deg P}) \prod_{P|h} (1 - q^{-l \deg P})^{-1} (1 - q^{1-l}),$$

by appealing to (2.1) and (2.2). Since

$$\sum_{\substack{F \in \mathcal{F}_{d,l} \\ \gcd(F, h)=1}} 1 = I_1 + O(I_2),$$

this completes the proof of Lemma 2.2.  $\square$

**2.2.** Denote by  $\mathcal{P}_d \subset \mathbb{F}_q[X]$  the set of monic irreducible polynomials of degree  $d \geq 1$ . The prime number theorem for polynomials [13] states that

$$(2.5) \quad \#\mathcal{P}_d = \frac{q^d}{d} \left( 1 + O\left(q^{-d/2}\right) \right).$$

The following result is also standard, based on a deep result of Weil ([14]), the analogue of the Riemann hypothesis for function fields over a finite field.

**Lemma 2.3.** Let  $\psi : \mathbb{F}_q[X] \rightarrow \mathbb{C}$  be a non-trivial Dirichlet character modulo  $Q$  in  $\mathbb{F}_q[X]$ , then

$$\left| \sum_{P \in \mathcal{P}_d} \psi(P) \right| \ll \frac{\deg(Q)}{d} q^{d/2}.$$

### 3. Proofs of Theorem 1.1 and Theorem 1.3

We first prove a general result, then Theorem 1.1 and Theorem 1.3 can be derived directly. The idea of the proof is similar to that of [7, 4, 5], though it is presented in a slightly different way via character sums.

Let  $l \geq 2$  be a positive integer such that  $q \equiv 1 \pmod{l}$ . Denote  $\zeta_l = \exp(2\pi i/l)$ . We fix a non-trivial character  $\chi : \mathbb{F}_q^\times \rightarrow \mathbb{C}$  of order  $l$ . For each  $x \in \mathbb{F}_q$ , let  $\chi_x : \mathbb{F}_q[X] \rightarrow \mathbb{C}$  be the Dirichlet character given by

$$\chi_x(F) := \chi(F(x)), \quad F \in \mathbb{F}_q[X].$$

For any  $U \subset \mathbb{F}_q$ , denote

$$g(U) := \prod_{x \in U} (X - x).$$

For the curve  $C_F$  given by the affine model (1.1), denote by  $C_F^0(\mathbb{F}_q)$  the set of the affine  $\mathbb{F}_q$ -points on  $C_F$ . It is known that

$$(3.1) \quad \#C_F^0(\mathbb{F}_q) = q + \sum_{j=1}^{l-1} \sum_{x \in \mathbb{F}_q} \chi_x^j(F).$$

For each  $d$ , there is a finite subset  $\mathcal{X}_d \subset \mathbb{F}_q[X]$ , on which we assign the uniform probability measure, so that  $\{\chi_x\}_{x \in \mathbb{F}_q}$  can be viewed as  $q$  random variables. We assume that there exist  $C, \epsilon > 0$  and  $0 \leq \gamma_q \leq 1$  such that

(a). For any non-trivial Dirichlet character  $\psi : \mathbb{F}_q[X] \rightarrow \mathbb{C}$  modulo  $h \in \mathbb{F}_q[X]$ , we have

$$\frac{1}{\#\mathcal{X}_d} \sum_{F \in \mathcal{X}_d} \psi(F) \leq q^{-\epsilon d + C \deg h}.$$

(b). For any  $U \subset \mathbb{F}_q$ ,

$$\frac{1}{\#\mathcal{X}_d} \sum_{\substack{F \in \mathcal{X}_d \\ \gcd(F, g(U))=1}} 1 = \gamma_q^{\#U} + O\left(q^{-\epsilon d + C\#U}\right).$$

**Theorem 3.1.** For each  $d$ , suppose that  $\mathcal{X}_d \subset \mathbb{F}_q[X]$  satisfies the conditions (a) and (b). Then

(1). For  $q$  fixed and  $d \rightarrow \infty$ , on  $\mathcal{X}_d$ ,  $\#C_F^0(\mathbb{F}_q)$  is distributed asymptotically as a sum of  $q$  i.i.d. random variables  $\{Y_x\}_{x \in \mathbb{F}_q}$ , where for each  $x$ ,  $Y_x$  takes the values  $0, -1, l-1$  with probabilities  $\left(1 - \gamma_q, \frac{(l-1)\gamma_q}{l}, \frac{\gamma_q}{l}\right)$  respectively.

(2). Moreover, if  $\lim_{q \rightarrow \infty} \gamma_q = \gamma > 0$ , then as  $q, d \rightarrow \infty$ , on  $\mathcal{X}_d$ , the limiting distribution of  $\frac{\#C_F^0(\mathbb{F}_q) - q}{\sqrt{q(l-1)\gamma}}$  is a standard Gaussian with mean zero and variance one.

*Proof.* For any vector of nonnegative integers  $\mathbf{r} = (r_x)_{x \in \mathbb{F}_q}$ , denote

$$n(\mathbf{r}) = \min \left\{ \sum_{x \in \mathbb{F}_q} r_x, q \right\}, \quad U(\mathbf{r}) = \{x \in \mathbb{F}_q : r_x > 0\}.$$

Let

$$M_{\mathbf{r}}(\chi, \mathbb{F}_q, \mathcal{X}_d) := \frac{1}{\#\mathcal{X}_d} \sum_{F \in \mathcal{X}_d} \prod_{x \in \mathbb{F}_q} \chi_x^{r_x}(F).$$

If  $r_x \not\equiv 0 \pmod{l}$  for some  $x \in \mathbb{F}_q$ , then  $\prod_{x \in \mathbb{F}_q} \chi_x^{r_x}(F)$  is a non-trivial Dirichlet character modulo  $h = g(U(\mathbf{r}))$  with  $\deg h \leq n(\mathbf{r})$ ; If  $r_x \equiv 0 \pmod{l}$  for any  $x \in \mathbb{F}_q$ , then  $\prod_{x \in \mathbb{F}_q} \chi_x^{r_x}(F)$  is a trivial Dirichlet character modulo  $h = g(U(\mathbf{r}))$  with  $\deg h \leq n(\mathbf{r})$ , and

$$\sum_{F \in \mathcal{X}_d} \prod_{x \in \mathbb{F}_q} \chi_x^{r_x}(F) = \sum_{\substack{F \in \mathcal{X}_d \\ \gcd(F, g(U(\mathbf{r})))=1}} 1.$$

Hence the conditions (a) and (b) can be summarized as

$$M_{\mathbf{r}}(\chi, \mathbb{F}_q, \mathcal{X}_d) = \begin{cases} O(q^{-\epsilon d + Cn(\mathbf{r})}) & r_x \not\equiv 0 \pmod{l} \exists x \in \mathbb{F}_q, \\ \gamma_q^{\#U(\mathbf{r})} + O(q^{-\epsilon d + Cn(\mathbf{r})}) & r_x \equiv 0 \pmod{l} \forall x \in \mathbb{F}_q. \end{cases}$$

For any nonnegative integer  $k$ , we consider the  $k$ -th moment

$$M_k(d, q) = \frac{1}{\#\mathcal{X}_d} \sum_{F \in \mathcal{X}_d} \left( \frac{\sum_{x \in \mathbb{F}_q} \sum_{j=1}^{l-1} \chi_x^j(F)}{\sqrt{q}} \right)^k.$$



We can expand

$$(3.3) \quad \left( \sum_{x \in \mathbb{F}_q} \sum_{j=1}^{l-1} \chi_x^j(F) \right)^k = \sum_{\mathbf{r}=(r_x)_{x \in \mathbb{F}_q}} a(\mathbf{r}) \prod_{x \in \mathbb{F}_q} \chi_x^{r_x}(F),$$

where on the right hand side the sum is over all vectors of nonnegative integers  $\mathbf{r} = (r_x)_{x \in \mathbb{F}_q}$  such that  $n(\mathbf{r}) \leq \sum_x r_x \leq k(l-1)$ , and  $a(\mathbf{r})$ 's are nonnegative combinatorial constants such that

$$(3.4) \quad \sum_{\mathbf{r}=(r_x)_{x \in \mathbb{F}_q}} a(\mathbf{r}) = (l-1)^k q^k.$$

Using (3.3) we find that

$$M_k(d, q) = q^{-k/2} \sum_{\mathbf{r}=(r_x)_{x \in \mathbb{F}_q}} a(\mathbf{r}) \frac{1}{\#\mathcal{X}_d} \sum_{F \in \mathcal{X}_d} \prod_{x \in \mathbb{F}_q} \chi_x^{r_x}(F).$$

Applying (3.2) and (3.4) we obtain

$$M_k(d, q) = q^{-k/2} \sum_{\substack{\mathbf{r}=(r_x)_{x \in \mathbb{F}_q} \\ (***)}} a(\mathbf{r}) \gamma_q^{\#U(\mathbf{r})} + O\left(l^k q^{-\epsilon d + \frac{k}{2} + C \min\{k(l-1), q\}}\right),$$

where the extra condition  $(***)$  means that  $r_x \equiv 0 \pmod{l}$  for any  $x \in \mathbb{F}_q$ .

On the other hand, let  $\{X_x\}_{x \in \mathbb{F}_q}$  be i.i.d. random variables, taking value 0 with probability  $1 - \gamma_q$  and each value  $\zeta_l^j$ ,  $1 \leq j \leq l$  with equal probability  $\gamma_q/l$ , we have for each positive integer  $\lambda > 0$ ,

$$(3.5) \quad \mathbb{E}\left(X_x^\lambda\right) = \begin{cases} 0 & \lambda \not\equiv 0 \pmod{l} \\ \gamma_q & \lambda \equiv 0 \pmod{l} \end{cases}, \quad x \in \mathbb{F}_q.$$

Expanding  $M_k = \mathbb{E}\left\{\left(\frac{\sum_{x \in \mathbb{F}_q} \sum_{j=1}^{l-1} X_x^j}{\sqrt{q}}\right)^k\right\}$  in the same way as for  $M_k(q, d)$ ,

we see that

$$M_k = q^{-k/2} \sum_{\substack{\mathbf{r}=(r_x)_{x \in \mathbb{F}_q} \\ (***)}} a(\mathbf{r}) \gamma_q^{\#U(\mathbf{r})},$$

where the condition  $(***)$  is the same as in the expression of  $M_k(q, d)$ . All other terms become zero because of independence of  $X_x$ 's and the identities (3.5). We conclude that for any nonnegative integer  $k$ ,

$$M_k(q) = \mathbb{E}\left\{\left(\frac{\sum_{x \in \mathbb{F}_q} \sum_{j=1}^{l-1} X_x^j}{\sqrt{q}}\right)^k\right\} \left(1 + O\left(l^k q^{-\epsilon d + k + C \min\{k(l-1), q\}}\right)\right).$$

Finally, denote

$$Y_x = \sum_{j=1}^{l-1} X_x^j, \quad \forall x \in \mathbb{F}_q.$$

It is easy to see that  $\{Y_x\}_{x \in \mathbb{F}_q}$  are  $q$  i.i.d. random variables and for any  $x \in \mathbb{F}_q$ ,

$$\begin{cases} \text{Prob}(Y_x = 0) &= 1 - \gamma_q, \\ \text{Prob}(Y_x = -1) &= \frac{(l-1)\gamma_q}{l}, \\ \text{Prob}(Y_x = l-1) &= \frac{\gamma_q}{l}. \end{cases}$$

From [3, Section 30] and the relation (4.4) we know that as  $d \rightarrow \infty$ , on the probability space  $\mathcal{X}_d$ , the value  $\#C_F^0(\mathbb{F}_q) - q$  is distributed asymptotically as  $\sum_{x \in \mathbb{F}_q} Y_x$ , and as  $d, q \rightarrow \infty$ , since  $\mathbb{E}(Y_x) = 0$ ,  $\text{Var}(Y_x) = (l-1)\gamma_q$  and  $\gamma_q \rightarrow \gamma > 0$  as  $q \rightarrow \infty$ , the limiting distribution of the normalized sum  $\frac{\#C_F^0(\mathbb{F}_q) - q}{\sqrt{q(l-1)\gamma}}$  is a standard Gaussian with mean zero and variance one. This completes the proof of Theorem 3.1.  $\square$

Now we can prove Theorem 1.1 and Theorem 1.3.

**Proofs of Theorem 1.1 and Theorem 1.3.** For  $\mathcal{F}_{d,l}$ , from (2.3), Lemma 2.1 and Lemma 2.2 in Section 2, we see that  $\mathcal{F}_{d,l}$ 's satisfy the conditions (a) and (b) with

$$\epsilon = 1 - \frac{1}{l}, \quad C = 1 - \frac{1}{l}, \quad \gamma_q = \frac{1 - q^{-1}}{1 - q^{-l}},$$

and  $\gamma_q \rightarrow 1$  as  $q \rightarrow \infty$ . For  $\mathcal{P}_d$ , since  $P \in \mathcal{P}_d$  is irreducible of degree  $d$ , if  $d \geq 2$ , then  $\gcd(P, g(U)) = 1$  for any  $U \subset \mathbb{F}_q$ . So the condition (b) is automatically satisfied with  $\gamma_q = 1$ . Moreover, from (2.5) and Lemma 2.3, we find that condition (a) is also satisfied with

$$\epsilon = \frac{1}{2}, \quad C = 1.$$

Then Theorem 1.1 and Theorem 1.3 follow from Theorem 3.1 directly.  $\square$

#### 4. Proofs of Theorem 1.2 and Theorem 1.4

We also prove a general result first, and Theorem 1.2 and Theorem 1.4 can be derived directly.

Kurlberg and Rudnick proved a similar result in [7], and their idea has been used by Bucur, David, Feigon and Lalín in [4, 5] to obtain various interesting results. We follow their ideas, however, our proof is based on properties of character sums.

**Theorem 4.1.** For each  $d$ , suppose that  $\mathcal{X}_d \subset \mathbb{F}_q[X]$  satisfies the conditions (a) and (b) as in Theorem 3.1. Denote

$$T_l = \left\{ \zeta_l^j : 1 \leq j \leq l \right\} \cup \{0\}.$$

Then for any vector  $(s_x)_{x \in \mathbb{F}_q} \in T_l^q$ , we have

$$\text{Prob}_{\mathcal{X}_d}(\chi_x = s_x, \forall x \in \mathbb{F}_q) = \text{Prob}(X_x = s_x, \forall x \in \mathbb{F}_q) \left(1 + O\left(2^q q^{-\epsilon d + Cq}\right)\right),$$

where  $\{X_x\}_{x \in \mathbb{F}_q}$  are i.i.d. random variables and for each  $x \in \mathbb{F}_q$ ,  $X_x$  takes value 0 with probability  $1 - \gamma_q$  and each value  $\zeta_l^j$ ,  $1 \leq j \leq l$  with equal probability  $\frac{\gamma_q}{l}$ .

*Proof.* For  $(s_x)_{x \in \mathbb{F}_q} \in T_l^q$ , we need to compute

$$L = \text{Prob}_{\mathcal{X}_d}(\chi_x = s_x, \forall x \in \mathbb{F}_q).$$

Let

$$A = \{x \in \mathbb{F}_q : s_x = 0\}, \quad B = \mathbb{F}_q / A.$$

Write

$$L = \frac{1}{\#\mathcal{X}_d} \# \left\{ F \in \mathcal{X}_d : \begin{array}{ll} \chi_x(F) = 0, & \forall x \in A \\ \chi_x(F) = s_x \neq 0, & \forall x \in B \end{array} \right\}.$$

It is easy to see that

$$1 - \chi_x^l(F) = \begin{cases} 1 & : \chi_x(F) = 0, \\ 0 & : \chi_x(F) \neq 0, \end{cases}$$

and for any  $s_x \in \{\zeta_l^j : 1 \leq j \leq l\}$ ,

$$\frac{1}{l} \sum_{r_x=1}^l (\chi_x(F) s_x^{-1})^{r_x} = \begin{cases} 1 & : \chi_x(F) = s_x, \\ 0 & : \chi_x(F) \neq s_x. \end{cases}$$

Hence

$$L = \frac{1}{\#\mathcal{X}_d} \sum_{F \in \mathcal{X}_d} \prod_{x \in A} (1 - \chi_x^l(F)) \prod_{x \in B} \frac{1}{l} \sum_{r_x=1}^l (\chi_x(F) s_x^{-1})^{r_x}.$$

For any  $U \subset \mathbb{F}_q$ , denote

$$\chi_U = \prod_{x \in U} \chi_x.$$

We can expand

$$(4.1) \quad \prod_{x \in A} (1 - \chi_x^l(F)) = \sum_{A' \subset A} (-1)^{\#A'} \chi_{A'}^l(F),$$

where the sum is over all sets  $A'$  with  $A' \subset A$ , and

$$(4.2) \quad \prod_{x \in B} \frac{1}{l} \sum_{r_x=1}^l (\chi_x(F) s_x^{-1})^{r_x} = \frac{1}{l^{\#B}} \sum_{\substack{1 \leq r_x \leq l \\ \forall x \in B}} \left( \prod_{x \in B} s_x^{-r_x} \right) \left( \prod_{x \in B} \chi_x^{r_x}(F) \right).$$

Using (4.1) and (4.2) and changing the order of summation we obtain

$$L = \frac{1}{l^{\#B}} \sum_{A' \subset A} (-1)^{\#A'} \sum_{\substack{1 \leq r_x \leq l \\ \forall x \in B}} \left( \prod_{x \in B} s_x^{-r_x} \right) \frac{1}{\#\mathcal{X}_d} \sum_{F \in \mathcal{X}_d} \chi_{A'}^l \prod_{x \in B} \chi_x^{r_x}(F).$$

If for some  $x \in B$ ,  $r_x \neq l$ , then  $\chi_{A'}^l \prod_{x \in B} \chi_x^{r_x}$  is a non-trivial Dirichlet character, and from the condition (a),

$$\frac{1}{\#\mathcal{X}_d} \sum_{F \in \mathcal{X}_d} \chi_{A'}^l \prod_{x \in B} \chi_x^{r_x}(F) \ll q^{-\epsilon d + Cq}.$$

The total contribution from such cases is bounded by

$$\ll \frac{1}{l^{\#B}} \sum_{A' \subset A} 1 \sum_{\substack{1 \leq r_x \leq l \\ \forall x \in B}} q^{-\epsilon d + Cq} \leq 2^q q^{-\epsilon d + Cq}.$$

The main contribution in  $L$  comes from the case that  $r_x = l$  for all  $x \in B$ , that is

$$(4.3) \quad \frac{1}{l^{\#B}} \sum_{A' \subset A} (-1)^{\#A'} \frac{1}{\#\mathcal{X}_d} \sum_{F \in \mathcal{X}_d} \chi_{A'}^l \chi_B^l(F).$$

From the condition (b), we find that

$$\frac{1}{\#\mathcal{X}_d} \sum_{F \in \mathcal{X}_d} \chi_{A'}^l \chi_B^l(F) = \gamma_q^{\#A' + \#B} + O\left(q^{-\epsilon d + Cq}\right).$$

Therefore

$$\begin{aligned} L &= \frac{1}{l^{\#B}} \sum_{A' \subset A} (-1)^{\#A'} \gamma_q^{\#A' + \#B} + O\left(2^q q^{-\epsilon d + Cq}\right) \\ &= \frac{1}{l^{\#B}} \gamma_q^{\#B} (1 - \gamma_q)^{\#A} + O\left(2^q q^{-\epsilon d + Cq}\right), \end{aligned}$$

as we collect the error terms together. We conclude that

$$\text{Prob}_{\mathcal{X}_d}(\chi_x = s_x, \forall x \in \mathbb{F}_q) = \left(\frac{\gamma_q}{l}\right)^{\#B} (1 - \gamma_q)^{\#A} + O\left(2^q q^{-\epsilon d + Cq}\right),$$

where

$$A = \{x \in \mathbb{F}_q : s_x = 0\}, \quad B = \mathbb{F}_q / A.$$

On the other hand, let  $\{X_x\}_{x \in \mathbb{F}_q}$  be  $q$  i.i.d random variables such that for any  $x \in \mathbb{F}_q$ ,  $X_x$  takes value 0 with probability  $1 - \gamma_q$  and each value  $\zeta_l^j$ ,  $1 \leq j \leq l$  with equal probability  $\frac{\gamma_q}{l}$ . It is easy to see that

$$\text{Prob}(X_x = s_x, \forall x \in \mathbb{F}_q) = \prod_{x \in \mathbb{F}_q} \text{Prob}(X_x = s_x) = \left(\frac{\gamma_q}{l}\right)^{\#B} (1 - \gamma_q)^{\#A}.$$

Therefore for any  $(s_x)_{x \in \mathbb{F}_q} \in T_l^q$ ,

$$\text{Prob}_{\mathcal{X}_d}(\chi_x = s_x, \forall x \in \mathbb{F}_q) = \text{Prob}(X_x = s_x, \forall x \in \mathbb{F}_q) + O\left(2^q q^{-\epsilon d + Cq}\right).$$

Noting that as  $d, q \rightarrow \infty$ ,

$$\text{Prob}_{\mathcal{X}_d}(\chi_x = s_x, \forall x \in \mathbb{F}_q) = 0 \text{ if and only if } \text{Prob}(X_x = s_x, \forall x \in \mathbb{F}_q) = 0,$$

this completes the proof of Theorem 4.1.  $\square$

Now we can prove Theorem 1.2 and Theorem 1.4.

**Proofs of Theorem 1.2 and Theorem 1.4.** As we know,  $\mathcal{F}_{d,l}$ 's satisfy the conditions (a) and (b) with

$$\epsilon = 1 - \frac{1}{l}, \quad C = 1 - \frac{1}{l}, \quad \gamma_q = \frac{1 - q^{-1}}{1 - q^{-l}}.$$

Hence from Theorem 4.1,

$$\text{Prob}_{\mathcal{X}_d}(\chi_x = s_x, \forall x \in \mathbb{F}_q) = \text{Prob}(X_x = s_x, \forall x \in \mathbb{F}_q) \left(1 + O\left(2^q q^{-\epsilon d + Cq}\right)\right),$$

where  $\{X_x\}_{x \in \mathbb{F}_q}$  are  $q$  i.i.d. random variables such that for any  $x$ ,  $X_x$  takes value 0 with probability  $1 - \gamma_q$  and each value  $\zeta_l^j$ ,  $1 \leq j \leq l$  with equal probability  $\frac{\gamma_q}{l}$ .

Since

$$\#C_F^0(\mathbb{F}_q) = q + \sum_{j=1}^{l-1} \sum_{x \in \mathbb{F}_q} \chi_x^j(F),$$

for any  $s \in \mathbb{Z}$ , we find that

$$\text{Prob}_{\mathcal{X}_d}(\#C_F^0(\mathbb{F}_q) - q = s) = \sum_{\substack{\sum_{j=1}^{l-1} \sum_x s_x^j = s \\ s_x \in T_l, \forall x}} \text{Prob}_{\mathcal{X}_d}(\chi_x = s_x, \forall x \in \mathbb{F}_q).$$

By Theorem 4.1, we have

$$\begin{aligned} \text{Prob}_{\mathcal{X}_d}(\#C_F^0(\mathbb{F}_q) - q = s) &= \sum_{\substack{\sum_{j=1}^{l-1} \sum_x s_x^j = s \\ s_x \in T_l, \forall x}} \text{Prob}(X_x = s_x, \forall x \in \mathbb{F}_q) \left(1 + O\left(2^q q^{-\epsilon d + Cq}\right)\right) \\ &= \text{Prob}\left(\sum_{j=1}^{l-1} \sum_{x \in \mathbb{F}_q} X_x^j = s\right) \left(1 + O\left(2^q q^{-\epsilon d + Cq}\right)\right). \end{aligned}$$

Denoting

$$Y_x = \sum_{j=1}^{l-1} X_x^j, \quad x \in \mathbb{F}_q,$$

this completes the proof of Theorem 1.2.

Theorem 1.4 can be proved similarly, noting that  $\mathcal{P}_d$ 's satisfy the conditions (a) and (b) with

$$\epsilon = \frac{1}{2}, \quad C = 1, \quad \gamma_q = 1. \quad \square$$

### References

- [1] J. BERGSTRÖM, *Equivariant counts of points of the moduli spaces of pointed hyperelliptic curves*, preprint, <http://arxiv.org/abs/math/0611813v1>, 2006.
- [2] P. BILLINGSLEY, *Probability and Measure*, third ed., Wiley Ser. Probab. Math. Stat., John Wiley & Sons Inc., Ney Youk, 1995, A Wiley-Interscience Publication.
- [3] P. BILLINGSLEY, *Probability and Measure*, third ed., Wiley Ser. Probab. Math. Stat., John Wiley & Sons Inc., Ney Youk, 1995, A Wiley-Interscience Publication.
- [4] A. BUCUR, C. DAVID, B. FEIGON, M. LALÍN, *Statistics for traces of cyclic trigonal curves over finite fields*, International Mathematics Research Notices (2010), 932–967.
- [5] A. BUCUR, C. DAVID, B. FEIGON, M. LALÍN, *Biased statistics for traces of cyclic  $p$ -fold covers over finite fields*, to appear in Proceedings of Women in Numbers, Fields Institute Communications.
- [6] P. DIACONIS, M. SHAHSHAHANI, *On the eigenvalues of random matrices*, in: Studies in Applied Probability, J. Appl. Probab. **31A** (1994), 49–62.
- [7] P. KURLBERG, Z. RUDNICK, *The fluctuations in the number of points on a hyperelliptic curve over a finite field*, J. Number Theory Vol. 129 **3** (2009), 580–587.
- [8] N. M. KATZ, P. SARNAK, *Random Matrices, Frobenius Eigenvalues, and Monodromy*, Amer. Math. Soc. Colloq. Publ., vol. 45, American Mathematical Society, Providence, RI, 1999.
- [9] N. M. KATZ, P. SARNAK, *Zeroes of zeta functions and symmetry*, Bull. Am. Math. Soc. **36** (1999), 1–26.
- [10] L. A. KNIZHNERMAN, V. Z. SOKOLINSKII, *Some estimates for rational trigonometric sums and sums of Legendre symbols*, Uspekhi Mat. Nauk **34** (3 (207)) (1979), 199–200.
- [11] L. A. KNIZHNERMAN, V. Z. SOKOLINSKII, *Trigonometric sums and sums of Legendre symbols with large and small absolute values*, in: Investigations in Number Theory, Saratov. Gos. Univ., Saratov, 1987, 76–89.
- [12] M. LARSEN, *The normal distribution as a limit of generalized sato-tate measures*, preprint.
- [13] M. ROSEN, *Number theory in function fields*, Graduate Texts in Mathematics, **210**, Springer-Verlag, New York, 2002.
- [14] A. WEIL, *Sur les Courbes Algébriques et les Variétés qui s'en Déduisent*, Publ. Inst. Math. Univ. Strasbourg **7** (1945), Hermann et Cie., Paris, 1948. iv+85 pp.

Maosheng Xiong  
 Department of Mathematics  
 Hong Kong University of Science and Technology  
 Clear Water Bay, Kowloon  
 P. R. China  
*E-mail* : [mamsxiong@ust.hk](mailto:mamsxiong@ust.hk)