



Asymptotic Normality of Q-ary Linear Codes

Minjia Shi, Olivier Rioul, Patrick Solé

► **To cite this version:**

Minjia Shi, Olivier Rioul, Patrick Solé. Asymptotic Normality of Q-ary Linear Codes. IEEE Communications Letters, Institute of Electrical and Electronics Engineers, 2019, 23 (11), pp.1895-1898. hal-02300794

HAL Id: hal-02300794

<https://hal.telecom-paris.fr/hal-02300794>

Submitted on 28 Aug 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Asymptotic Normality of Q -ary Linear Codes

Minjia Shi, *Member, IEEE*, Olivier Rioul *Member, IEEE*, and Patrick Solé *Member, IEEE*

Abstract—Sidel’nikov proved in 1971 that the weight distribution of long binary codes is asymptotically Gaussian. Delsarte sketched in 1975 an extension of this result to Q -ary codes when $Q > 2$. In this note, we complete Delsarte’s proof.

Index Terms—Linear codes, Weight distribution, Berry’s inequality.

I. INTRODUCTION

AN empirical observation made by many coding theorists is that the weight distribution of long Q -ary codes is well approximated by a normal (Gaussian) law (see Fig. 1). A rigorous result was established by Sidel’nikov [1] in the binary case ($Q = 2$), and by Delsarte [2] for $Q > 2$. The result is that the weight distribution of long codes is asymptotically normal when the dual distance of the code is large enough. A friendly exposition of Sidel’nikov’s bound for $Q = 2$ can be found in [3, Chapter 9, § 10], but Delsarte’s proof [2] contains several unproved bounds. In this paper, we give a complete proof with sharper bounds.

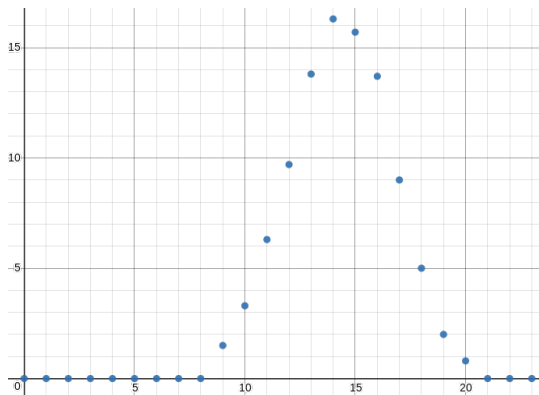


Fig. 1. Weight distribution of a $[23, 17]$ BCH code over \mathbb{F}_3 .

Minjia Shi is with the Key Laboratory of Intelligent Computing & Signal Processing, Ministry of Education, School of Mathematical Sciences of Anhui University, Anhui, 230601, P. R. China (E-mail: smjwcl.good@163.com)

Olivier Rioul is with the LTCI, Télécom Paris, Institut Polytechnique de Paris, Paris, France and with the CMAP, École Polytechnique, Institut Polytechnique de Paris, Palaiseau, France. E-mail: olivier.rioul@telecom-paristech.fr

Patrick Solé is with Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France

This research is supported by National Natural Science Foundation of China (61672036), Excellent Youth Foundation of Natural Science Foundation of Anhui Province (1808085J20) and Academic fund for outstanding talents in universities (g xbjZD03)

This paper is organized as follows. The next section collects basic facts and notations needed in the remainder of the paper. Section III compares the moments of the weight distribution with that of a binomial distribution of parameters (p, q) . Section IV builds on Berry’s inequality that compares the probability distribution function of an arbitrary law with that of the normal law. Section V provides a bound on the centered moments of a binomial law, that is sharper than that in [2]. Section VI simplifies a bound by Delsarte using a technique due to Essen. Section VII derives the main result by combining the bounds of the previous sections.

II. DEFINITIONS AND NOTATIONS

We are given any $[n, k, d]$ Q -ary linear code with length $n \geq 2$ and dual distance (minimal distance of its dual code) $d' > 2$. For $j = 0, 1, \dots, n$ let A_j and A'_j be the number of codewords of weight j in the code and its dual code, respectively. The MacWilliams identities [3] can be written as

$$\sum_{j=0}^n A_j x^j = \frac{1}{Q^{n-k}} \sum_{j=0}^n A'_j (1-x)^j (1+(Q-1)x)^{n-j}. \quad (1)$$

For $x = 1$ we have $\sum_{j=0}^n A_j = \frac{1}{Q^{n-k}} A'_0 Q^n = Q^k$. Hence defining $a_j = \frac{A_j}{Q^k}$ for $j = 0, 1, \dots, n$, we obtain a discrete distribution $\mathbf{a} = (a_0, a_1, \dots, a_n)$ of nonnegative real numbers satisfying $\sum_{j=0}^n a_j = 1$. From (1) its generating function is

$$\sum_{j=0}^n a_j x^j = \frac{1}{Q^n} \sum_{j=0}^n A'_j (1-x)^j (1+(Q-1)x)^{n-j}. \quad (2)$$

The mean $\mu = \sum_{j=0}^n j a_j$ of the weight distribution \mathbf{a} is given by the derivative of (2) at $x = 1$: $\mu = \frac{1}{Q^n} (-A'_1 Q^{n-1} + A'_0 n(Q-1)Q^{n-1}) = n(1 - \frac{1}{Q})$, where we have used that $A'_0 = 1$ and $A'_1 = 0$ (since $d' > 1$). In the remainder of this paper, we use the notations $p = 1 - \frac{1}{Q}$ and $q = \frac{1}{Q}$ which satisfy the relations $p + q = 1$ and $0 \leq q \leq \frac{1}{2} \leq p \leq 1$. Thus $\mu = np$ and (2) becomes

$$\sum_{j=0}^n a_j x^j = \sum_{j=0}^n A'_j q^j (1-x)^j (q + px)^{n-j}. \quad (3)$$

III. BINOMIAL MOMENT COMPARISON

For any distribution $\mathbf{p} = (p_0, p_1, \dots, p_n)$ of nonnegative real numbers satisfying $\sum_{j=0}^n p_j = 1$, we define its k th centered moment as $\mu_k(\mathbf{p}) = \sum_{j=0}^n p_j (j - \mu)^k$. Thus $\mu_0(\mathbf{p}) = 1$, $\mu_1(\mathbf{p}) = 0$ and $\mu_2(\mathbf{p}) = \sigma^2$ is the

variance of \mathbf{p} . The *moment-generating function* of \mathbf{p} is then defined as

$$M_{\mathbf{p}}(t) = \sum_{k=0}^{+\infty} \mu_k(\mathbf{p}) \frac{t^k}{k!} = \sum_{j=0}^n p_j e^{(j-\mu)t}. \quad (4)$$

Lemma 1 ([2, Lemma 4]). *The first d' centered moments of the weight distribution \mathbf{a} coincide with those of the binomial (n, p) distribution \mathbf{b} :*

$$\mu_k(\mathbf{a}) = \mu_k(\mathbf{b}) \quad (k = 0, 1, \dots, d' - 1) \quad (5)$$

where $\mathbf{b} = (b_0, b_1, \dots, b_n)$ is defined by $b_j = \binom{n}{j} p^j q^{n-j}$, $j = 0, 1, \dots, n$.

A proof is given in [2]. We provide here a simplified proof for completeness.

Proof. First we remark that \mathbf{a} and \mathbf{b} have the same mean $\mu = np$. Using (3), the moment-generating function $M_{\mathbf{a}}(t) = \sum_{j=0}^n a_j (e^t)^j e^{-npt}$ of \mathbf{a} is

$$M_{\mathbf{a}}(t) = \sum_{j=0}^n A'_j q^j (e^{-pt} - e^{qt})^j (qe^{-pt} + pe^{qt})^{n-j},$$

while that of \mathbf{b} is

$$M_{\mathbf{b}}(t) = \sum_{j=0}^n \binom{n}{j} (pe^t)^j q^{n-j} e^{-npt} = (qe^{-pt} + pe^{qt})^n.$$

Therefore,

$$M_{\mathbf{a}}(t) = M_{\mathbf{b}}(t) \cdot \sum_{j=0}^n A'_j q^j \left(\frac{e^{-pt} - e^{qt}}{qe^{-pt} + pe^{qt}} \right)^j$$

where $\frac{e^{-pt} - e^{qt}}{qe^{-pt} + pe^{qt}} = \frac{-t+o(t)}{1+o(t)} = O(t)$. Since $A'_0 = 1$ and $A'_1 = A'_2 = \dots = A'_{d'-1} = 0$ it follows that $M_{\mathbf{a}}(t) = M_{\mathbf{b}}(t) \cdot (1 + O(t^{d'}))$, which implies (5). \square

Notice that in particular for $k = 2$, \mathbf{a} and \mathbf{b} have the same variance $\sigma^2 = npq$, hence the same standard deviation $\sigma = \sqrt{npq}$.

IV. BERRY'S INEQUALITY

Any discrete distribution $\mathbf{p} = (p_0, p_1, \dots, p_n)$ of nonnegative real numbers satisfying $\sum_{j=0}^n p_j = 1$, $\sum_{j=0}^n j p_j = \mu$, and $\sum_{j=0}^n p_j (j - \mu)^2 = \sigma^2$ can be seen as a probability distribution of a random variable X having mean μ and standard deviation σ . The corresponding *normalized* variable $\frac{X-\mu}{\sigma}$ has mean = 0, standard deviation = 1 and normalized centered moments

$$m_k(\mathbf{p}) = \mathbb{E} \left[\left(\frac{X-\mu}{\sigma} \right)^k \right] = \sum_{j=0}^n p_j \left(\frac{j-\mu}{\sigma} \right)^k = \frac{\mu_k(\mathbf{p})}{\sigma^k}.$$

Its corresponding *characteristic function* is noted

$$\begin{aligned} \hat{p}(t) &= \mathbb{E} \left(e^{i \frac{X-\mu}{\sigma} t} \right) = \sum_{j=0}^n p_j e^{i \frac{j-\mu}{\sigma} t} = M_{\mathbf{p}} \left(\frac{it}{\sigma} \right) \\ &= \sum_{k=0}^{+\infty} m_k(\mathbf{p}) \frac{(it)^k}{k!}, \end{aligned} \quad (6)$$

while its cumulative *distribution function* is

$$P(x) = \mathbb{P} \left(\frac{X-\mu}{\sigma} \leq x \right) = \sum_{\frac{j-\mu}{\sigma} \leq x} p_j = \sum_{j \leq \mu + \sigma x} p_j.$$

The goal of the present paper is to compare the cumulative weight distribution function $A(x) = \sum_{j \leq \mu + \sigma x} a_j$ to the cumulative distribution function of the standard normal $\mathcal{N}(0, 1)$: $\Phi(x) = \int_{-\infty}^x \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt$. Our starting point is Berry's inequality [4, § XVI.3 Eq. (3.13)]:

$$|A(x) - \Phi(x)| \leq \frac{1}{\pi} \int_{-T}^T \left| \frac{\hat{a}(t) - \hat{\phi}(t)}{t} \right| dt + \frac{24m}{\pi T} \quad (7)$$

valid for any $x \in \mathbb{R}$ and $T > 0$, where $\hat{a}(t) = \sum_{k=0}^{+\infty} m_k(\mathbf{a}) \frac{(it)^k}{k!}$ is the characteristic function associated to the weight distribution \mathbf{a} , $\hat{\phi}(t) = e^{-t^2/2}$ is the characteristic function of the standard normal, and

$$m = \max_{t \in \mathbb{R}} \Phi'(t) = \max_{t \in \mathbb{R}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}}. \quad (8)$$

Lemma 2. *For any even $r < d'$,*

$$\begin{aligned} |A(x) - \Phi(x)| &\leq \frac{1}{\pi} \int_{-T}^T \left| \frac{\hat{b}(t) - \hat{\phi}(t)}{t} \right| dt + \frac{4T^r m_r(\mathbf{b})}{\pi r \cdot r!} \\ &\quad + \frac{24}{\pi \sqrt{2\pi} T}, \end{aligned} \quad (9)$$

where $\hat{\phi}(t) = e^{-t^2/2}$ and

$$\hat{b}(t) = M_{\mathbf{b}} \left(\frac{it}{\sigma} \right) = (qe^{-ipt/\sigma} + pe^{iqt/\sigma})^n \quad (10)$$

is the characteristic function of the normalized binomial (n, p) .

Proof. By Taylor's theorem with Lagrange form of the remainder, for any $t \in \mathbb{R}$,

$$\left| e^{it} - \sum_{k=0}^{r-1} \frac{(it)^k}{k!} \right| \leq \frac{|t|^r}{r}.$$

Thus the characteristic function (6) satisfies

$$\left| \mathbb{E} \left(e^{i \frac{X-\mu}{\sigma} t} \right) - \sum_{k=0}^{r-1} \mathbb{E} \left[\left(\frac{X-\mu}{\sigma} \right)^k \right] \frac{(it)^k}{k!} \right| \leq \mathbb{E} \left[\left| \frac{X-\mu}{\sigma} \right|^r \right] \frac{|t|^r}{r!}.$$

This is

$$\left| \hat{p}(t) - \sum_{k=0}^{r-1} m_k(\mathbf{p}) \frac{(it)^k}{k!} \right| \leq m'_r(\mathbf{p}) \frac{|t|^r}{r!} \quad (11)$$

where $m'_r(\mathbf{p}) = \mathbb{E} \left[\left| \frac{X-\mu}{\sigma} \right|^r \right]$ denotes the *absolute* moment which equals $m_r(\mathbf{p})$ when r is even.

Now, by Lemma 1, for any $r < d'$, $m_k(\mathbf{a}) = m_k(\mathbf{b})$ for all $k = 0, 1, \dots, r$. Therefore, using (11) for distributions \mathbf{a} and \mathbf{b} , for any even $r < d'$,

$$\begin{aligned} |\hat{a}(t) - \hat{b}(t)| &\leq \left| \hat{a}(t) - \sum_{k=0}^{r-1} m_k(\mathbf{a}) \frac{(it)^k}{k!} \right| \\ &\quad + \left| \hat{b}(t) - \sum_{k=0}^{r-1} m_k(\mathbf{b}) \frac{(it)^k}{k!} \right| \\ &\leq m'_r(\mathbf{a}) \frac{|t|^r}{r!} + m'_r(\mathbf{b}) \frac{|t|^r}{r!} = 2m_r(\mathbf{b}) \frac{|t|^r}{r!}. \end{aligned}$$

It follows that

$$\begin{aligned} \frac{1}{\pi} \int_{-T}^T \left| \frac{\hat{a}(t) - \hat{\phi}(t)}{t} \right| dt &\leq \frac{1}{\pi} \int_{-T}^T \left| \frac{\hat{b}(t) - \hat{\phi}(t)}{t} \right| dt \\ &\quad + \frac{1}{\pi} \int_{-T}^T 2m_r(\mathbf{b}) \frac{|t|^r}{|t|r!} dt, \end{aligned} \quad (12)$$

where

$$\begin{aligned} \frac{1}{\pi} \int_{-T}^T 2m_r(\mathbf{b}) \frac{|t|^r}{|t|r!} dt &= \frac{4}{\pi} m_r(\mathbf{b}) \int_0^T \frac{t^{r-1}}{r!} dt \\ &= \frac{4T^r m_r(\mathbf{b})}{\pi r \cdot r!}. \end{aligned} \quad (13)$$

Combining (7), (8), (12), and (13) gives (9). \square

V. A BOUND ON THE BINOMIAL MOMENT

In his proof [2], Delsarte mentions the bound

$$m_r(\mathbf{b}) \leq \left(\frac{epr}{2q} \right)^{r/2} \sqrt{\frac{r}{4\pi}}$$

for even r , without any justification. This bound seems *ad hoc*, as it is not symmetric in the variables p and q . In this section, we give a detailed proof of a different, more symmetric, and sharper bound.

Lemma 3. *For any even $r \geq 0$, we have*

$$m_r(\mathbf{b}) \leq \left(\frac{2}{pq} \right)^{r/2} \frac{r!}{(r/2)!}.$$

Proof. Let X_1, \dots, X_n be independent Bernoulli(p) random variables. Then $X = \sum_{j=1}^n X_j$ follows the binomial (n, p) distribution so that

$$m_r(\mathbf{b}) = \mathbb{E} \left[\left(\frac{X - \mu}{\sigma} \right)^k \right] = \sigma^{-r} \mathbb{E} \left[\left(\sum_{j=1}^n (X_j - p) \right)^r \right].$$

We now symmetrize this expression using n additional independent Bernoulli(p) random variables X'_1, \dots, X'_n , independent of the X_j 's. We have

$$\begin{aligned} \mathbb{E} \left[\left(\sum_{j=1}^n (X_j - p) \right)^r \right] &= \mathbb{E} \left[\left(\sum_{j=1}^n (X_j - \mathbb{E}(X'_j)) \right)^r \right] \\ &= \mathbb{E}_X \left[\left(\mathbb{E}_{X'} \left(\sum_{j=1}^n (X_j - X'_j) \right) \right)^r \right] \\ &\leq \mathbb{E} \left[\left(\sum_{j=1}^n (X_j - X'_j) \right)^r \right] \end{aligned}$$

by Jensen's inequality applied to the function $x \mapsto x^r$, which is convex for even r .

Now the independent random variables $Y_j = \frac{X_j - X'_j}{2}$ ($j = 1, \dots, n$) equal 1 or -1 with probability pq , and equal 0 otherwise. Therefore, their moments are $\mathbb{E}(Y_j^k) = 2pq \leq \frac{1}{2}$ for even k , $\mathbb{E}(Y_j^k) = 0$ otherwise.

Let Z_1, \dots, Z_n be independent standard normal variables. Their moments are $\mathbb{E}(Z_j^k) = (k-1)!! > \frac{1}{2}$ for even k , $\mathbb{E}(Z_j^k) = 0$ otherwise. If we expand the r th power in $\mathbb{E} \left[\left(\sum_{j=1}^n Y_j \right)^r \right]$ to a sum of monomials, any monomial with odd exponents vanishes because $\mathbb{E}(Y_j^k) = 0 = \mathbb{E}(Z_j^k)$ for odd k . The other monomials are nonnegative with moments $\mathbb{E}(Y_j^k) \leq \frac{1}{2} < \mathbb{E}(Z_j^k)$ for even k . Therefore, $\mathbb{E} \left[\left(\sum_{j=1}^n Y_j \right)^r \right]$ is term-by-term dominated by $\mathbb{E} \left[\left(\sum_{j=1}^n Z_j \right)^r \right]$, where $\sum_{j=1}^n Z_j$ follows a normal $\mathcal{N}(0, n)$ distribution:

$$\mathbb{E} \left[\left(\sum_{j=1}^n Y_j \right)^r \right] \leq \mathbb{E} \left[\left(\sum_{j=1}^n Z_j \right)^r \right] = (r-1)!! (\sqrt{n})^r,$$

where $(r-1)!! = 1 \cdot 3 \cdot \dots \cdot (r-3) \cdot (r-1) = \frac{r!}{2^{r/2}(r/2)!}$. Combining the above inequalities gives

$$m_r(\mathbf{b}) \leq \sigma^{-r} 2^r \frac{r!}{2^{r/2}(r/2)!} (\sqrt{n})^r \quad (14)$$

where $\sigma = \sqrt{npq}$. \square

VI. ESSEN'S SIMPLIFIED DERIVATION

To conclude on a bound of (7) it remains to bound the integral term $\frac{1}{\pi} \int_{-T}^T \left| \frac{\hat{b}(t) - \hat{\phi}(t)}{t} \right| dt$ in (9). The derivation is similar to derivations by Essen [5] but can be simplified by exploiting the closed-form expression (10). Delsarte mentions the bound $\sqrt{\frac{3p\pi}{nq}}$ for $T < \frac{\sigma}{p^2+q^2}$ "without giving details" [2]. In this section, we provide a detailed proof of a simplified derivation of an improved bound.

Lemma 4. *For any $T \leq \frac{\sigma}{p^2+q^2}$, we have*

$$\frac{1}{\pi} \int_{-T}^T \left| \frac{\hat{b}(t) - \hat{\phi}(t)}{t} \right| dt < \sqrt{\frac{3p}{nq}}.$$

Proof. From (10) we have

$$\begin{aligned} &\frac{1}{\pi} \int_{-T}^T \left| \frac{\hat{b}(t) - \hat{\phi}(t)}{t} \right| dt \\ &= \frac{1}{\pi} \int_{-T}^T \left| \frac{(qe^{-ipt/\sigma} + pe^{iqt/\sigma})^n - e^{-t^2/2}}{t} \right| dt \quad (15) \\ &= \frac{1}{\pi} \int_{-T/\sigma}^{T/\sigma} \left| (qe^{-ipt} + pe^{iqt})^n - (e^{-pq t^2/2})^n \right| \frac{dt}{|t|} \end{aligned}$$

by the change of variable $t \leftarrow t/\sigma$, where $\sigma^2 = npq$.

Here $qe^{-ipt} + pe^{iqt}$ is the characteristic function of a centered Bernoulli(p) with mean 0 and variance pq , and $e^{-pq t^2/2}$ is the characteristic function of a normal distribution with the same mean 0 and variance pq . From (11) with $r = 3$ we have

$$\left| qe^{-ipt} + pe^{iqt} - 1 + pq \frac{t^2}{2} \right| \leq (qp^3 + pq^3) \frac{|t|^3}{6}. \quad (16)$$

Similarly for the centered normal we have

$$\left| e^{-pqt^2/2} - 1 + pq\frac{t^2}{2} \right| \leq (pq)^2 \frac{t^4}{8}. \quad (17)$$

Now from the factorization $b^n - a^n = (b-a)(b^{n-1} + b^{n-2}a + \dots + ba^{n-2} + a^{n-1})$ we have, for any $a, b \in \mathbb{C}$,

$$|b^n - a^n| \leq nc^{n-1}|b - a| \quad (18)$$

for any constant $c \geq \max(|a|, |b|)$. We apply this inequality to $a = e^{-pqt^2/2}$ and $b = qe^{-ipt} + pe^{iqt}$. First from (16), we have

$$\begin{aligned} & |qe^{-ipt} + pe^{iqt}| \\ & \leq |1 - pq\frac{t^2}{2}| + pq(p^2 + q^2)|t|\frac{t^2}{6} \\ & \leq 1 - pq\frac{t^2}{2} + pq\frac{t^2}{6} = 1 - pq\frac{t^2}{3} \leq e^{-pqt^2/3} \end{aligned} \quad (19)$$

provided that $1 - pq\frac{t^2}{2} \geq 0$, that is, $|t| \leq \sqrt{\frac{2}{pq}}$ and that $|t| \leq \frac{1}{p^2+q^2}$. It is easily seen that $p^2 + q^2 \geq \frac{1}{2}$ and $\frac{pq}{2} \leq \frac{1}{8}$ so that $\frac{1}{p^2+q^2} \leq 2 \leq 2\sqrt{2} \leq \sqrt{\frac{2}{pq}}$. Thus (19) holds for any $|t| \leq \frac{T}{\sigma}$ in (15) provided that we choose $T \leq \frac{\sigma}{p^2+q^2}$. Thus we can take $c = e^{-pqt^2/3} \geq \max(|a|, |b|)$ in (18). Next, from (16) and (17) we have

$$\begin{aligned} |b - a| &= |qe^{-ipt} + pe^{iqt} - e^{-pqt^2/2}| \\ &\leq pq(p^2 + q^2)\frac{|t|^3}{6} + (pq)^2\frac{t^4}{8}. \end{aligned}$$

Combining the above inequalities we obtain

$$\begin{aligned} & \frac{1}{\pi} \int_{-T}^T \left| \frac{\hat{b}(t) - \hat{\phi}(t)}{t} \right| dt \\ & \leq \frac{n}{\pi} \int_{-\frac{T}{\sigma}}^{\frac{T}{\sigma}} \left(pq(p^2 + q^2)\frac{|t|^3}{6} + (pq)^2\frac{t^4}{8} \right) e^{-(n-1)pqt^2/3} \frac{dt}{|t|} \\ & \leq \frac{n}{\pi} \int_{-\infty}^{+\infty} \left(p^2q\frac{t^2}{6} + (pq)^2\frac{|t|^3}{8} \right) e^{-npqt^2/6} dt, \end{aligned}$$

where we have used that $p^2 + q^2 \leq p^2 + pq = p$ since $q \leq p$ and $n-1 \geq \frac{n}{2}$ since $n \geq 2$.

Let T be a centered normal variable of variance $V = \mathbb{E}(T^2) = \frac{3}{npq}$ and of absolute third moment $\mathbb{E}(|T|^3) = \frac{4V^2}{\sqrt{2\pi V}}$. We obtain

$$\begin{aligned} & \frac{1}{\pi} \int_{-T}^T \left| \frac{\hat{b}(t) - \hat{\phi}(t)}{t} \right| dt \\ & \leq \frac{n}{\pi} \sqrt{2\pi V} \left(p^2q \frac{\mathbb{E}(T^2)}{6} + (pq)^2 \frac{\mathbb{E}(|T|^3)}{8} \right) \\ & = \frac{n}{\pi} \sqrt{2\pi V} p^2q \frac{V}{6} + \frac{n}{\pi} (pq)^2 \frac{V^2}{2} \\ & = \frac{n}{\pi} \sqrt{\frac{6\pi}{npq}} p^2q \frac{3}{6npq} + \frac{n}{\pi} (pq)^2 \frac{9}{2(npq)^2} \\ & = \sqrt{\frac{3p}{2q\pi n}} + \frac{9}{2\pi n} < \frac{5}{2} \sqrt{\frac{3p}{2q\pi n}} \\ & = \sqrt{\frac{75p}{8q\pi n}} < \sqrt{\frac{3p}{qn}}, \end{aligned}$$

because

$$\frac{9}{2\pi n} \leq \frac{3\sqrt{3}}{\sqrt{2\pi n}} \sqrt{\frac{3}{2\pi n}} \leq \frac{3\sqrt{3}}{2\sqrt{\pi}} \sqrt{\frac{3p}{2q\pi n}} < \frac{3}{2} \sqrt{\frac{3p}{2q\pi n}}$$

(since $n \geq 2$, $p \geq q$ and $\pi > \frac{25}{8} > 3$). \square

VII. MAIN RESULTS

Theorem 1. *There exists a constant c that depends only on Q such that for all $x \in \mathbb{R}$,*

$$|A(x) - \Phi(x)| < \frac{c}{\sqrt{d'}}. \quad (20)$$

Proof. Combining Lemmas 2, 3 and 4 yields, for any $T \leq \frac{\sqrt{npq}}{p^2+q^2}$ and even $r < d'$

$$\begin{aligned} |A(x) - \Phi(x)| &< \sqrt{\frac{3p}{nq}} + \frac{4T^r}{\pi r \cdot r!} \left(\frac{2}{pq} \right)^{r/2} \frac{r!}{(r/2)!} + \frac{24}{\pi\sqrt{2\pi T}} \\ &\leq \frac{p\sqrt{3}}{qT} + \frac{4T^r}{\pi r \sqrt{\pi r}} \left(\frac{4e}{pqr} \right)^{r/2} + \frac{24}{\pi\sqrt{2\pi T}}, \end{aligned}$$

where we have used $\sqrt{npq} \geq T(p^2 + q^2) \geq T(pq + q^2) = Tq$ as well as Stirling's approximation $(r/2)! \geq \sqrt{\pi r} (r/2)^{r/2} e^{-r/2}$.

Let $T = \sqrt{\alpha \frac{q}{p} r}$ where $\alpha \leq 1$. Since $r \leq n$, we check that $T \leq \sqrt{\frac{q}{p} n} = \frac{\sqrt{npq}}{p} \leq \frac{\sqrt{npq}}{p^2+q^2}$. Therefore, we have

$$\begin{aligned} & |A(x) - \Phi(x)| \\ & < \left(\sqrt{\frac{3p^3}{\alpha q^3}} + \frac{4}{\pi r \sqrt{\pi}} \left(\frac{4e\alpha}{p^2} \right)^{r/2} + \frac{24}{\pi\sqrt{2\pi}} \sqrt{\frac{p}{\alpha q}} \right) \frac{1}{\sqrt{r}}. \end{aligned}$$

The $(r/2)$ th power disappears by setting $\alpha = \frac{p^2}{4e} < 1$. Since $d' > 2$ we can take $r \geq 2$ such that $r \geq d'/2$. We obtain

$$\begin{aligned} |A(x) - \Phi(x)| &< \left(\sqrt{\frac{24ep}{q^3}} + \frac{2\sqrt{2}}{\pi\sqrt{\pi}} + \frac{24}{\pi\sqrt{\pi}} \sqrt{\frac{4e}{pq}} \right) \frac{1}{\sqrt{d'}} \\ &= \left(\sqrt{24e(Q-1)Q} + \frac{2\sqrt{2}}{\pi\sqrt{\pi}} + \frac{24Q}{\pi\sqrt{\pi}} \sqrt{\frac{4e}{Q-1}} \right) \frac{1}{\sqrt{d'}} \end{aligned}$$

which ends the proof. \square

VIII. CONCLUSION

We have established rigorously the asymptotic normality of the weight distribution of Q -ary codes. Casting the topic into the framework of translation invariant association schemes might lead to interesting generalizations like Lee metric codes, rank metric codes, and block designs.

REFERENCES

- [1] V. M. Sidelnikov, "Weight spectrum of binary Bose-Chaudhuri-Hocquenghem codes," *Problemy Peredachi Informatsii*, vol. 7, no. 1, pp. 14–22, Jan.–Mar. 1971.
- [2] P. Delsarte, "Distance distribution of functions over Hamming spaces," in *Philips Research Reports*, vol. 30, no. 1, Feb. 1975, pp. 1–8.
- [3] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. North-Holland, 1977.
- [4] W. Feller, *An Introduction to Probability Theory and Its Applications*. John Wiley & Sons, 1970, vol. II.
- [5] C.-G. Essén, "On the remainder term in the central limit theorem," *Arkiv för Matematik*, vol. 8, no. 2, pp. 7–15, 1968.