*Bull. Aust. Math. Soc.* **101** (2020), 157–162
doi:10.1017/S0004972719000637

# ON THE GENERALISATION OF SIDEL'NIKOV'S THEOREM TO $q$-ARY LINEAR CODES

YILUN WEI, BO WU and QIJIN WANG✉

(Received 22 March 2019; accepted 27 March 2019; first published online 27 May 2019)

**Abstract**

We generalise Sidel'nikov's theorem from binary codes to $q$-ary codes for $q > 2$. Denoting by $A(z)$ the cumulative distribution function attached to the weight distribution of the code and by $\Phi(z)$ the standard normal distribution function, we show that $|A(z) - \Phi(z)|$ is bounded above by a term which tends to 0 when the code length tends to infinity.

2010 *Mathematics subject classification*: primary 94B05; secondary 97N20.

*Keywords and phrases*: $q$-ary linear code, estimation, Sidel'nikov's theorem.

## 1. Introduction

For the binary alphabet, it is well known that the cumulative distribution of linear codes can be approximated by a standard normal distribution. If $\mathscr{C}$ is an $[n, k, d]$ binary linear code with weight distribution $(A_0, A_1, \ldots, A_n)$, where $A_j$ is the number of codewords in $\mathscr{C}$ with weight $j$, we define $\mathbf{a} = (a_0, a_1, \ldots, a_n)$, where $a_j = A_j/2^k$. The mean and variance of $\mathbf{a}$ are $\mu(\mathbf{a}) = \sum_{j=0}^{n} j a_j$ and $\sigma^2(\mathbf{a}) = \sum_{j=0}^{n} (\mu(\mathbf{a}) - j)^2 a_j$, respectively. The cumulative distribution function (cdf) associated with $\mathbf{a}$ is $A(z) = \sum_{j \geq \mu(\mathbf{a}) - \sigma(\mathbf{a})z}^{n} a_j$. Let

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\, dt$$

be the cdf of the normal law and let $d'$ be the minimum distance of the dual code $\mathscr{C}^\perp$. Sidel'nikov [4] proved that $|A(z) - \Phi(z)| = O(1/\sqrt{d'})$ for $n$ large when $d' \geq 3$, which means that $\Phi(z)$ can be regarded as an asymptotic approximation of $A(z)$.

For $q$-ary alphabets, Delsarte [1] showed that the cdf $A(z)$ of linear codes can still be approximated by $\Phi(z)$. However, he did not provide a detailed proof. In this paper, we consider the situation of $q$-ary linear codes again, and rigorously prove that the asymptotic relation between $A(z)$ and $\Phi(z)$ still holds. More specifically, we derive the bound $|A(z) - \Phi(z)| \leq C/\sqrt[6]{d'}$, where $C$ is a constant that depends only on $q$.

The third author (corresponding author) is supported by Key Project of Natural Science Research of Anhui Higher Education Institutions (KJ2018A0589).

## 2. Preliminaries

For any real vector $\mathbf{v} = (v_0, v_1, \ldots, v_n)$ with $v_j \geq 0$ and $\sum_{j=0}^{n} v_j = 1$, the mean and variance of $\mathbf{v}$ are defined by

$$\mu(\mathbf{v}) = \sum_{j=0}^{n} j v_j \quad \text{and} \quad \sigma^2(\mathbf{v}) = \sum_{j=0}^{n} (\mu(\mathbf{v}) - j)^2 v_j.$$

The $s$th central moment of $\mathbf{v}$ is

$$\mu_s(\mathbf{v}) = \sum_{j=0}^{n} \left(\frac{\mu(\mathbf{v}) - j}{\sigma(\mathbf{v})}\right)^s v_j.$$

Let $\mathscr{C}$ be an $[n, k, d]$ $q$-ary linear code and $A_j$ the number of codewords in $\mathscr{C}$ with weight $j$. Define $a_j = A_j/q^k$ so that $\mathbf{a} = (a_0, a_1, \ldots, a_n)$ satisfies the conditions above. Hence $\mu(\mathbf{a})$, $\sigma^2(\mathbf{a})$ and $\mu_s(\mathbf{a})$ can be defined. The cdf of $\mathbf{a}$ is given by

$$A(z) = \sum_{j \geq \mu(\mathbf{a}) - \sigma(\mathbf{a})z}^{n} a_j.$$

Let $\mathbf{b} = (b_0, b_1, \ldots, b_n)$, where $b_j = q^{-n}\binom{n}{j}(q-1)^j$. Then $\mathbf{b} \sim B(n, 1 - 1/q)$, $\mathbf{b}$ satisfies the conditions above,

$$\mu(\mathbf{b}) = \frac{(q-1)n}{q}, \quad \sigma^2(\mathbf{b}) = \frac{(q-1)n}{q^2}$$

and

$$\mu_s(\mathbf{b}) = \sum_{j=0}^{n} \left(\frac{\mu(\mathbf{b}) - j}{\sigma(\mathbf{b})}\right)^s b_j = \frac{q^{-n}}{\sqrt{(q-1)n}^s} \sum_{j=0}^{n} [(q-1)n - qj]^s \binom{n}{j}(q-1)^j.$$

We use this notation throughout the paper. Further details on $q$-ary linear codes can be found in [6–10].

## 3. Main result

LEMMA 3.1. *Let $\mathscr{C}$ be an $[n, k, d]$ $q$-ary linear code and $d'$ the minimum distance of $\mathscr{C}^\perp$. For $s = 0, 1, \ldots, d' - 1$,*

$$\mu_s(\boldsymbol{a}) = \mu_s(\boldsymbol{b}).$$

PROOF. Applying the MacWilliams identity for $q$-ary codes [3, Equation (M3), page 257],

$$W_{\mathscr{C}^\perp}(x, y) = \frac{1}{|\mathscr{C}|} W_{\mathscr{C}}(y - x, y + (q-1)x). \tag{3.1}$$

Let $A'_i$ $(i = 0, 1, \ldots, n)$ be the number of codewords with weight $i$ in $\mathscr{C}^\perp$ and let $a'_i$ be the MacWilliams transform of $a_i$ with parameter $\lambda = q - 1$ [1]. Substituting $y = 1$ into (3.1) gives the expansion

$$\sum_{i=0}^{n} A'_i x^i = \frac{1}{q^k} \sum_{i=0}^{n} A_i (1-x)^i [1 + (q-1)x]^{n-i} = \sum_{i=0}^{n} a_i (1-x)^i [1 + (q-1)x]^{n-i}.$$

https://doi.org/10.1017/S0004972719000637 Published online by Cambridge University Press

We then find that, for $i = 0, 1, \dots, n$,

$$a'_i = \frac{A'_i}{q^{n-k}}.$$

Since $a'_1 = \cdots = a'_{d'-1} = 0$, by Delsarte [1, Lemma 4], we have $\mu_s(\mathbf{a}) = \mu_s(\mathbf{b})$ for $s = 0, 1, \dots, d' - 1$. □

Because the definitions of $\mu$, $\sigma^2$ and $\mu_s$ in the $q$-ary case are the same as in the binary case, the formulas from Sidel'nikov's derivation [4, Equations (35)–(40), pages 285–286] remain correct. Setting $r = 2[(d' - 1)/2]$ gives the following lemma.

Lemma 3.2. *For all $T > 0$,*

$$|A(z) - \Phi(z)| \le \frac{1}{\pi}\int_{-T}^{T}\frac{1}{|t|}\bigg|\sum_{s=0}^{\infty}\mu_s(\boldsymbol{b})\frac{(it)^s}{s!} - e^{-t^2/2}\bigg|\,dt + \frac{2}{\pi}\int_{-T}^{T}\mu_r(\boldsymbol{b})\frac{|t|^{r-1}}{r!}\,dt + \frac{24}{T\pi\sqrt{2\pi}}. \tag{3.2}$$

Lemma 3.3. *Let $\mathscr{C}$ be an $[n, k, d]$ $q$-ary linear code and $d'$ the minimum distance of $\mathscr{C}^{\perp}$. If $n \ge 6$ and $d' \ge \frac{1}{2}n + 3$, then $r = 2[(d' - 1)/2]$ satisfies*

$$r \ge \frac{n}{2} \ge \frac{d'}{2}.$$

Proof. From the definition of $r$, together with $n \ge 6$ and $d' \ge \frac{1}{2}n + 3$,

$$r = 2\bigg[\frac{d' - 1}{2}\bigg] \ge 2\bigg(\frac{d' - 3}{2}\bigg) \ge \frac{n}{2} \ge \frac{d'}{2}.$$

This completes the proof. □

We now focus on the right-hand side of (3.2) and give upper bounds for each of the three terms.

Lemma 3.4. *For all $T$ with $0 < T \le T_0 = n^{1/6}/(3\rho^{1/3})$,*

$$\frac{1}{\pi}\int_{-T}^{T}\frac{1}{|t|}\bigg|\sum_{s=0}^{\infty}\mu_s(\boldsymbol{b})\frac{(it)^s}{s!} - e^{-t^2/2}\bigg|\,dt \le \frac{8\rho}{\sqrt{2\pi}}\cdot\frac{1}{\sqrt{n}}$$

*where $0 < \rho < \infty$ and $\rho$ is a constant that depends only on $q$.*

Proof. Define the random variables $\xi \sim B(n, 1 - 1/q)$, $\eta = (E\xi - \xi)/\sqrt{D\xi}$, and let $\varphi_\eta(t)$ be the characteristic function of $\eta$. We find

$$\begin{aligned}\varphi_\eta(t) &= \sum_{j=0}^{n}\exp\bigg(it\cdot\frac{(q-1)n - jq}{\sqrt{n(q-1)}}\bigg)\binom{n}{j}(q-1)^j q^{-n}\\ &= \sum_{s=0}^{\infty}\sum_{j=0}^{n}\bigg(\frac{(q-1)n - jq}{\sqrt{n(q-1)}}\bigg)^s\binom{n}{j}(q-1)^j q^{-n}\frac{(it)^s}{s!}\\ &= \sum_{s=0}^{\infty}\mu_s(\mathbf{b})\frac{(it)^s}{s!}.\end{aligned}$$

https://doi.org/10.1017/S0004972719000637 Published online by Cambridge University Press

Thus the first term on the right-hand side of (3.2) is equal to

$$\frac{1}{\pi}\int_{-T}^{T}\frac{1}{|t|}|\varphi_\eta(t)-e^{-t^2/2}|\,dt.$$

Define independent random variables $\xi_1,\ldots,\xi_n$, where each $\xi_j$ satisfies

$$P(\xi_j=\sqrt{q-1})=\frac{1}{q},\quad P\Big(\xi_j=-\frac{1}{\sqrt{q-1}}\Big)=1-\frac{1}{q}.$$

It is easy to verify that $E\xi_j=0$ and $D\xi_j=1$. If we now set $s_n^2=\sum_{j=1}^n D\xi_j$, then $\eta=s_n^{-1}\sum_{j=1}^n\xi_j$. Define $F_j(x)$ to be the distribution function of $\xi_j$ and

$$\rho_j=\sup_{z>0}\Big(\Big|\int_{-z}^{z}x^3\,dF_j(x)\Big|+z\int_{|x|\geq z}x^2\,dF_j\Big).$$

Observe that $\rho_1=\cdots=\rho_n$, so we can set $\rho=\rho_j$. From the definition of the Riemann–Stieltjes integral, $0<\rho<\infty$ and $\rho$ only depends on $q$. From Esseen [2, Lemma 5], for $|t|\leq T\leq T_0=n^{1/6}/(3\rho^{1/3})$,

$$|\varphi_\eta(t)-e^{-t^2/2}|\leq\frac{4\sum_{j=1}^n\rho_j}{s_n^3}|t|^3e^{-t^2/2}=\frac{4\rho}{\sqrt{n}}|t|^3e^{-t^2/2}.$$

Hence

$$\begin{aligned}\frac{1}{\pi}\int_{-T}^{T}\frac{1}{|t|}|\varphi_\eta(t)-e^{-t^2/2}|\,dt&\leq\frac{4\rho}{\pi\sqrt{n}}\int_{-T}^{T}|t|^2e^{-t^2/2}\,dt\\&\leq\frac{8\rho}{\pi\sqrt{n}}\int_0^{+\infty}t^2e^{-t^2/2}\,dt\\&=\frac{8\rho}{\sqrt{2\pi}}\cdot\frac{1}{\sqrt{n}}.\end{aligned}$$

This completes the proof. □

Lemma 3.5 [5]. *Define a random variable $X\sim B(n,p)$. Then, for all even $r$,*

$$\mu_r(X)=\sigma(X)^{-r}E(X-\mu(X))^r\leq\Big(\frac{2}{p(1-p)}\Big)^{r/2}\cdot\frac{r!}{(r/2)!}.$$

Lemma 3.6. *For all $T>0$ and all even $r$,*

$$\frac{2}{\pi}\int_{-T}^{T}\mu_r(\boldsymbol{b})\frac{|t|^{r-1}}{r!}\,dt\leq\frac{4e^{1/24}T^r}{\pi\sqrt{r}\cdot r!}\cdot\Big(\frac{4rq^2}{e(q-1)}\Big)^{r/2}.$$

Proof. Since $\mathbf{b}\sim B(n,1-1/q)$, by Lemma 3.5,

$$\mu_r(\mathbf{b})\leq\Big(\frac{2q^2}{q-1}\Big)^{r/2}\cdot\frac{r!}{(r/2)!}.\tag{3.3}$$

https://doi.org/10.1017/S0004972719000637 Published online by Cambridge University Press

Substituting (3.3) into the second term on the right-hand side of (3.2),

$$\begin{aligned}\frac{2}{\pi}\int_{-T}^{T}\mu_r(\mathbf{b})\frac{|t|^{r-1}}{r!}\,dt &\le \frac{4T^r}{\pi r\cdot r!}\cdot\Big(\frac{2q^2}{q-1}\Big)^{r/2}\cdot\frac{r!}{(r/2)!}\\ &< \frac{4e^{1/24}T^r}{\pi\sqrt{r}\cdot r!}\cdot\Big(\frac{2q^2}{q-1}\Big)^{r/2}\cdot\Big(\frac{2r}{e}\Big)^{r/2}\\ &= \frac{4e^{1/24}T^r}{\pi\sqrt{r}\cdot r!}\cdot\Big(\frac{4rq^2}{e(q-1)}\Big)^{r/2}. \qquad (3.4)\end{aligned}$$

The observation $\sqrt{2\pi m}(m/e)^m < m! < \sqrt{2\pi m}(m/e)^m e^{1/12m}$ for all $m \in N^*$ has been used in the second inequality in (3.4). □

If we choose a suitable $T$ satisfying $0 < T \le T_0 = n^{1/6}/(3\rho^{1/3})$ and collect all the results above, we reach the following bound for the right-hand side of (3.2), which gives the generalisation of Sidel'nikov's theorem for $q$-ary linear codes.

Theorem 3.7. *Let $\mathscr{C}$ be an $[n,k,d]$ $q$-ary linear code and $d'$ the minimum distance of $\mathscr{C}^\perp$. If $n \ge 6$ and $d' \ge \frac{1}{2}n + 3$, then*

$$|A(z) - \Phi(z)| \le \frac{C}{\sqrt[6]{d'}}$$

*where $C$ is a constant that depends only on $q$.*

Proof. Choose $T$ with $0 < T \le T_0 = n^{1/6}/(3\rho^{1/3})$, so that $\sqrt{n} \ge 27\rho T^3$. Using this along with Lemmas 3.4 and 3.6, the right-hand side of (3.2) is

$$\le \frac{8}{27T^3\sqrt{2\pi}} + \frac{4e^{1/24}T^r}{\pi\sqrt{r}\cdot r!}\cdot\Big(\frac{4rq^2}{e(q-1)}\Big)^{r/2} + \frac{24}{T\pi\sqrt{2\pi}}. \qquad (3.5)$$

Let

$$c = \min\Big(\frac{1}{16}, \frac{(q-1)e}{9\rho^{2/3}}\Big) \quad\text{and}\quad T = \Big(\frac{cr}{(q-1)e}\Big)^{1/2}\cdot\frac{1}{n^{1/3}}.$$

Then $0 < T \le (n/(9\rho^{2/3}))^{1/2}\cdot n^{-1/3} = T_0$ and we can substitute $T$ into (3.5). Finally, from the inequality $(r/e)^r/r! < 1/\sqrt{2\pi r}$ and Lemma 3.3,

$$\begin{aligned}|A(z)-\Phi(z)| &\le \frac{8[(q-1)e]^{3/2}}{27c^{3/2}\sqrt{2\pi}}\cdot\Big(\frac{n^{1/3}}{\sqrt{r}}\Big)^3 + \frac{4e^{1/24}}{\pi\sqrt{r}\cdot r!}\cdot\Big(\frac{r}{e}\Big)^r\cdot\Big(\frac{2qc^{1/2}}{(q-1)n^{1/3}}\Big)^r\\ &\quad + \frac{24[(q-1)e]^{1/2}}{c^{1/2}\pi\sqrt{2\pi}}\cdot\frac{n^{1/3}}{\sqrt{r}}\\ &\le \frac{16[(q-1)e]^{3/2}}{27c^{3/2}\sqrt{2\pi}}\cdot\frac{1}{\sqrt{r}} + \frac{4e^{1/24}}{\pi\sqrt{2\pi}}\cdot\frac{1}{r} + \frac{24[(q-1)e]^{1/2}\cdot 2^{1/3}}{c^{1/2}\pi\sqrt{2\pi}}\cdot\frac{1}{\sqrt[6]{r}}\\ &\le \frac{16[(q-1)e]^{3/2}}{27c^{3/2}\sqrt{\pi}}\cdot\frac{1}{\sqrt{d'}} + \frac{8e^{1/24}}{\pi\sqrt{2\pi}}\cdot\frac{1}{d'} + \frac{24[(q-1)e]^{1/2}}{c^{1/2}\pi\sqrt{\pi}}\cdot\frac{1}{\sqrt[6]{d'}}\\ &= \frac{c_1}{\sqrt{d'}} + \frac{c_2}{d'} + \frac{c_3}{\sqrt[6]{d'}} \le \frac{C}{\sqrt[6]{d'}},\end{aligned}$$

where $C$ is a constant that depends only on $q$. □

https://doi.org/10.1017/S0004972719000637 Published online by Cambridge University Press

## 4. Conclusion and open problem

The estimate $|A(z) - \Phi(z)| = O(1/\sqrt[6]{d'})$ when $n \to \infty$ that we have obtained for $q$-ary codes is coarser than Sidel'nikov's upper-bound $O(1/\sqrt{d'})$ for the binary case. A challenging open problem is to improve the above estimates.

## Acknowledgements

The authors are grateful to Professor Minjia Shi and Patrick Solé for helpful discussions.

## References

[1] P. Delsarte, 'Distance distribution of functions over Hamming spaces', *Philips Res. Rep.* **30** (1975), 1–8.
[2] C. G. Esseen, 'On the remainder term in the central limit theorem', *Ark. Mat.* **8** (1969), 7–15.
[3] W. C. Huffman and V. Pless, *Fundamentals of Error-Correcting Code* (Cambridge University Press, Cambridge, 2003).
[4] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes. I*, North-Holland Mathematical Library, 16 (North-Holland, Amsterdam, 1977).
[5] M. J. Shi, O. Rioul and P. Solé, 'On the asymptotic normality of $Q$-ary linear codes', in preparation.
[6] M. J. Shi, L. Q. Qian, L. Sok, N. Aydin and P. Solé, 'On constacyclic codes over $\mathbb{Z}_4[u]/\langle u^2 - 1\rangle$ and their Gray images', *Finite Fields Appl.* **45** (2017), 86–95.
[7] M. J. Shi, Z. Sepasdar, A. Alahmadi and P. Solé, 'On two weight $\mathbb{Z}_{2^k}$-codes', *Des. Codes Cryptogr.* **86** (2018), 1201–1209.
[8] M. J. Shi, R. S. Wu, L. Q. Qian, L. Sok and P. Solé, 'New classes of $p$-ary few weight codes', *Bull. Malays. Math. Sci. Soc.*, to appear.
[9] M. J. Shi and Y. P. Zhang, 'Quasi-twisted codes with constacyclic constituent codes', *Finite Fields Appl.* **39** (2016), 159–178.
[10] M. J. Shi, H. W. Zhu and P. Solé, 'How many weights can a linear code have?', *Des. Codes Cryptogr.* **87** (2019), 87–95.

YILUN WEI, School of Mathematical Sciences,
Anhui University, Hefei, Anhui, 230601, China
e-mail: ylwei1997@163.com

BO WU, School of Mathematical Sciences,
Anhui University, Hefei, Anhui, 230601, China
e-mail: wubo@ahu.edu.cn

QIJIN WANG, Anhui Xinhua University,
Hefei, Anhui, 230088, China
e-mail: qjwang118@163.com