Sidon sets, sum-free sets and linear codes

Ingo Czerwinski* and Alexander Pott*

February 19, 2023

Abstract

The maximum size of a Sidon set in $\mathbb{F}_2^t$ is of research interest for more than 40 years. In order to tackle this problem we recall a fruitful one-to-one correspondence between sum-free Sidon sets and linear codes with minimum distance greater or equal 5. Our main contribution about codes is a new non-existence result for linear codes with minimum distance 5 based on a sharpening of the Johnson bound from Brouwer and Tolhuizen. This gives, on the Sidon set side, an improvement of the general upper bound for the maximal size of a Sidon set. Additionally, we characterise maximal Sidon sets up to dimension 6 and give the exact maximal size of a Sidon set in dimension 7 to 10 with the help of Grassl's codes table.

Keywords Sidon set, sum-free set, maximal size, linear binary code, codes bound.

1 Sidon sets

Sidon introduced B_2-sequences of positive integers in connection with his work on Fourier analysis [7], [8]. Later, Babai and Sós [1] generalised the definition of B_2-sequences to arbitrary groups and called them Sidon sets. In this work, we focus only on Sidon sets in $\mathbb{F}_2^t$.

Definition 1.1. *Let M be a subset of $\mathbb{F}_2^t$. M is called Sidon if $m_1 + m_2 \neq m_3 + m_4$ for all pairwise distinct $m_1, m_2, m_3, m_4 \in M$.*

Let M be a subset of $\mathbb{F}_2^t$. For any k we call

$$\mathcal{S}_k(M) = \{m_1 + \cdots + m_k : m_1, \ldots, m_k \in M\}$$

the *k-sums* of M and

$$\mathcal{S}_k^*(M) = \{m_1 + \cdots + m_k : m_1, \ldots, m_k \in M \text{ pairwise distinct}\}$$

the *k-star-sums* of M. The Sidon property of M can be characterised in terms of 2-star-sums and 3-star-sums by the following equivalent statements:

(a) M is Sidon;

(b) $|\mathcal{S}_2^*(M)| = \binom{|M|}{2}$;

(c) $\mathcal{S}_3^*(M) \cap M = \emptyset$.

The fundamental problem about a Sidon set is the question about its maximal size, which was already discussed about 40 years ago by Babai and Sós [1]. We will denote by $s_{max}(\mathbb{F}_2^t)$ the maximal size of a Sidon set in $\mathbb{F}_2^t$. An upper bound arises directly from the fact that all 2-star-sums of M have to be distinct, hence

$$\binom{|M|}{2} = \frac{|M|(|M|-1)}{2} \leq |\mathbb{F}_2^t \setminus \{0\}|,$$

*Faculty of Mathematics, Otto von Guericke University Magdeburg, 39106 Magdeburg, Germany, (ingo@czerwinski.eu, alexander.pott@ovgu.de)

which is equivalent to

$$s_{max}(\mathbb{F}_2^t) \leq \begin{cases} 2^{\frac{t+1}{2}} & \text{for } t \text{ odd,} \\ \left\lfloor \sqrt{2^{t+1}} + 0.5 \right\rfloor & \text{for } t \text{ even.} \end{cases} \tag{1}$$

This bound will be improved later in Theorem 3.6.

It is natural to ask if every Sidon set can be extended to a Sidon set of maximal size by adding elements. Already in dimension 6 (see Proposition 1.2) this is not true and motivates the following definition for a Sidon set $M \subseteq \mathbb{F}_2^t$: M is called *maximal* if $M = S$ for every Sidon set S with $M \subseteq S \subseteq \mathbb{F}_2^t$.

It is possible to characterise maximal Sidon sets via their 3-star-sums and 3-sums: *Let $M \subseteq \mathbb{F}_2^t$ be a Sidon set with $|M| \geq 3$ (and therefore $t \geq 2$). Then the following statements are equivalent:*

(a) M is maximal;

(b) $\mathcal{S}_3^*(M) \cup M = \mathbb{F}_2^t$;

(c) $\mathcal{S}_3(M) = \mathbb{F}_2^t$.

As the property of being maximal Sidon is invariant under affine permutations we are able to characterise all maximal Sidon sets up to dimension 6.

Proposition 1.2. *Let M be a maximal Sidon set of $\mathbb{F}_2^t$ and $e_1, \ldots, e_t$ be the standard basis of $\mathbb{F}_2^t$. Then M is affine equivalent to*

t=1: $M_1 = \{0, e_1\} = \mathbb{F}_2^1$ *with* $|M_1| = 2$.

t=2: $M_2 = \{0, e_1, e_2\} = \mathbb{F}_2^2 \setminus \{0\}$ *with* $|M_2| = 3$.

t=3: $M_3 = \{0, e_1, e_2, e_3\}$ *with* $|M_3| = 4$.

t=4: $M_4 = \{0, e_1, e_2, e_3, e_4, e_1 + e_2 + e_3 + e_4\}$ *with* $|M_4| = 6$.

t=5: $M_{5a} = \{0, e_1, e_2, e_3, e_4, e_5, e_1 + e_2 + e_3 + e_4\}$; *or*
$M_{5b} = \{0, e_1, e_2, e_3, e_4, e_5, e_1 + e_2 + e_3 + e_4 + e_5\}$;
with $|M_{5a}| = |M_{5b}| = 7$.

t=6: $M_{6a} = M_{5a} \cup \{e_6, e_{i_1} + e_{i_2} + e_5 + e_6\}$; *or*
$M_{6b} = M_{5b} \cup \{e_6, e_{j_1} + e_{j_2} + e_{j_3} + e_6\}$; *or*
$M_{6c} = \{0, e_1, e_2, e_3, e_4, e_5, e_6, e_1 + e_2 + e_3 + e_4 + e_5 + e_6\}$;
with distinct $i_1, i_2 \in \{1, 2, 3, 4\}$, *pairwise distinct* $j_1, j_2, j_3 \in \{1, 2, 3, 4, 5\}$
and $|M_{6a}| = |M_{6b}| = 9 > |M_{6c}| = 8$.

For dimension 7 and 8 we were not able to classify maximal Sidon sets, but we obtain all possible sizes of maximal Sidon sets by computer calculations.

Proposition 1.3. *Let M be a maximal Sidon set of $\mathbb{F}_2^t$. If $t = 7$ then $|M| = 12$ and if $t = 8$ then $|M| \in \{15, 16, 18\}$.*

We have seen that a maximal Sidon set can be characterised via its 3-sums, i.e. a Sidon set M is maximal if and only if $\mathcal{S}_3(M) = \mathbb{F}_2^t$. We obtain the following result about the 4-sums of Sidon sets:

Theorem 1.4. *Let M be a Sidon set of $\mathbb{F}_2^t$. If $|M| > s_{max}(\mathbb{F}_2^{t-1})$ then*

(a) $\mathcal{S}_4(M) = \mathbb{F}_2^t$ and

(b) $\dim\langle M \rangle = \dim\langle \mathcal{S}_2^*(M) \rangle = t$.

2 Sum-free sets

In this section we introduce sum-free sets and give some basic properties. The motivation is two fold. Their structure helps when looking at the maximal size of Sidon sets and we also need them to give a one-to-one correspondence between linear codes with minimum distance greater or equal 5 and sum-free Sidon sets. More on this topic can be found in the survey papers of Green and Ruzsa [5] as well as Tao and Vu [10].

Definition 2.1. *Let M be a subset of $\mathbb{F}_2^t$. M is called sum-free if $m_1+m_2 \neq m_3$ for all $m_1, m_2, m_3 \in M$.*

By definition, 0 is never contained in a sum-free set. We give some basic properties. *Let M be a subset of $\mathbb{F}_2^t$.*

(a) M is sum-free if and only if $\mathcal{S}_2(M) \cap M = \emptyset$.

(b) If M is sum-free then $|M| \leq 2^{t-1}$.

(c) M is sum-free and $|M| = 2^{t-1}$ if and only if $M = H + a$ for a hyperplane $H \leq \mathbb{F}_2^t$ and $a \in \mathbb{F}_2^t \setminus \{0\}$.

We note that the property of being sum-free is invariant under linear permutations but not in general under affine permutations.

For a subset M of $\mathbb{F}_2^t$ and $g \in \mathbb{F}_2^t \setminus M$ it is possible to give exact conditions when $M \cup \{g\}$ keeps the property of being sum-free, Sidon or sum-free Sidon. Here we just state the case when 0 is contained in a Sidon set which is of special interest: *Let M be a subset of $\mathbb{F}_2^t$.*

(a) If M is sum-free Sidon then $M \cup \{0\}$ is Sidon and not sum-free.

(b) If M is Sidon and $0 \in M$ then $M \setminus \{0\}$ is sum-free Sidon.

Therefore it is equivalent to discuss the maximal size of a sum-free Sidon set instead of discussing the maximal size of a Sidon set.

Proposition 2.2. *Let $sfs_{max}(\mathbb{F}_2^t)$ denote the maximal size of a sum-free Sidon set in $\mathbb{F}_2^t$. Then*

$$s_{max}(\mathbb{F}_2^t) = sfs_{max}(\mathbb{F}_2^t) + 1.$$

3 Linear Codes

A (binary) *linear code* $\mathcal{C}$ of *length* n and *dimension* k is a k dimensional vector subspace $\mathcal{C} \subseteq \mathbb{F}_2^n$. Such a code $\mathcal{C}$ is called a $[n,k]$-code and $c \in \mathcal{C}$ is called *code word* of $\mathcal{C}$. If the minimum distance of $\mathcal{C}$ is d, that is the minimum number of non-zero entries of all non-zero code words of $\mathcal{C}$, then $\mathcal{C}$ is called a $[n,k,d]$-code.

We recall a fruitful correspondence between linear codes and additive structures from Cohen and Zémor [3]. Therefore we assume in the rest of this section that $\mathbb{F}_2^n$ is endowed with an ordering of its elements e.g the binary ordering, but all what follows is independent of this ordering.

Let M be a non-empty subset of $\mathbb{F}_2^t \setminus \{0\}$. Then we define its associated linear code $\mathcal{C}_M \subseteq \mathbb{F}_2^{|M|}$ with minimum distance $d \geq 3$ by the $t \times |M|$ check matrix $\mathcal{H}_M = (m)_{m \in M}$ where each $m \in M$ is seen as a column vector. Hence $\mathcal{C}_M = \{c \in \mathbb{F}_2^{|M|} : \mathcal{H}_M \cdot c^\intercal = 0\}$.

Now the other way around. The columns of a check matrix of a $[n,k,d]$-code $\mathcal{C}$ with $d \geq 3$ form a set $M_\mathcal{C} \subseteq \mathbb{F}_2^{n-k} \setminus \{0\}$ with $|M_\mathcal{C}| = n$.

This leads due to Cohen and Zémor [3] to a one-to-one correspondence

$$\begin{array}{ccc} M & \longmapsto & \mathcal{C}_M \\ M_\mathcal{C} & \longmapsfrom & \mathcal{C} \end{array}$$

between properties on the codes side like minimum distance and covering radius to additive properties of the corresponding sets and vice versa. *For example, let M be a non-empty subset of $\mathbb{F}_2^t \setminus \{0\}$ and let $\mathcal{C}_M$ be its associated $[|M|,k,d]$-code then M is sum-free if and only if $d \geq 4$; and M is sum-free Sidon if and only if $d \geq 5$.*

The following Theorem gives details on this correspondence with a focus on Sidon sets and will be used later to prove Proposition 3.3.

Theorem 3.1. *Let M be a non-empty subset of $\mathbb{F}_2^t \setminus \{0\}$ and let $\mathcal{C}_M$ be its $[|M|,k,d]$-code.*

(a) If $|M| \geq s_{max}(\mathbb{F}_2^t)$ then $d = 3$ or $d = 4$.

(b) If M is sum-free Sidon and $|M| \geq s_{max}(\mathbb{F}_2^{t-1})$ then $k = |M| - t$.

(c) If M is sum-free Sidon and $|M| \geq s_{max}(\mathbb{F}_2^{t-1}) + 1$ then $d = 5$.

The following Theorem is another interesting connection between a code property (covering radius) and a Sidon property:

Theorem 3.2. *Let M be a non-empty subset of $\mathbb{F}_2^t \setminus \{0\}$ and let $\mathcal{C}_M$ be its $[|M|, k, d]$-code with covering radius R.*

(a) If M is sum-free Sidon and $|M| \geq s_{max}(\mathbb{F}_2^{t-1})$ then $R = 3$ or $R = 4$.

(b) M is maximal sum-free Sidon (that means we cannot extend it to a larger sum-free Sidon set) if and only if $R = 3$.

We define

$$d_{max}(n, k) = \max\{d : \text{there exists an } [n, k, d]\text{-code}\}.$$

as the *maximal minimum distance* of a code with given length n and dimension k. It is one of the main properties of *optimal* codes and frequently listed as a matrix $(d_{max}(n, k))_{n,k}$. Now we translate Theorem 3.1 to some properties of the subdiagonals of $(d_{max}(n, k))_{n,k}$, namely the entries $(d_{max}(n, n-t))_n$ for a fixed t.

Proposition 3.3. *Let be $n, t \in \mathbb{N}$.*

(a) $d_{max}(n, n-t) = 3$ if and only if $2^{t-1} < n < 2^t$.

(b) $d_{max}(n, n-t) = 4$ if and only if $s_{max}(\mathbb{F}_2^t) \leq n \leq 2^{t-1}$.

(c) $d_{max}(n, n-t) = 5$ if and only if $s_{max}(\mathbb{F}_2^{t-1}) < n < s_{max}(\mathbb{F}_2^t)$.

(d) $d_{max}(n, n-t) \geq 6$ if and only if $n \leq s_{max}(\mathbb{F}_2^{t-1})$.

Due to the importance of non-existence statements for Sidon sets and as well for linear codes we reformulate Proposition 3.3.

Corollary 3.4. *Let be $n, t \in \mathbb{N}$. Then the following statements are equivalent:*

(a) There is no $[n, n-t, 5]$ code.

(b) There is no Sidon set $M \subseteq \mathbb{F}_2^t$ of size $n+1$.

(c) $s_{max}(\mathbb{F}_2^t) \leq n$.

The following result from Brouwer and Tolhuizen [2] is achieved by a sharpening of the Johnson bound. It improves the upper bound (1) for odd dimension.

Theorem 3.5 ([2])**.** *There is no $[n, n-t, 5]$ code for $n = 2^{(t+1)/2} - 2$, hence*

$$s_{max}(\mathbb{F}_2^t) \leq 2^{(t+1)/2} - 2$$

for t odd with $t \geq 7$.

Taking the same arguments as Brouwer and Tolhuizen, we are able to generalise this result to arbitrary $t \geq 6$ and thereby further improve the upper bound (1). This improves also a bound given by Tait and Won [9].

Theorem 3.6. *Let be $t \geq 6$, $\left\lfloor \sqrt{2^{t+1}} + 0.5 \right\rfloor - 4 = 3a + b$ with $a, b \in \mathbb{Z}_{\geq 0}$, $b \leq 2$, $\varepsilon = \sqrt{2^{t+1}} + 0.5 - \left\lfloor \sqrt{2^{t+1}} + 0.5 \right\rfloor \in [0, 1)$ and*

$$\lambda_{a,b,\varepsilon} = \begin{cases} 1 & \text{for } a \text{ odd and } b = 0, \\ 2 & \text{for } a \text{ odd, } b = 1 \text{ and } 0 \leq \varepsilon \leq 1 - \frac{1}{2^{(t-4)/2}}, \\ 1 & \text{for } a \text{ odd, } b = 1 \text{ and } 1 - \frac{1}{2^{(t-4)/2}} < \varepsilon < 1, \\ 2 & \text{for } a \text{ odd and } b = 2, \\ 2 & \text{for } a \text{ even, } b = 0 \text{ and } 0 \leq \varepsilon \leq 0.5, \\ 1 & \text{for } a \text{ even, } b = 0 \text{ and } 0.5 < \varepsilon < 1, \\ 2 & \text{for } a \text{ even, } b = 1 \text{ and } 0 \leq \varepsilon \leq 1 - \frac{1}{2^{(t-5)/2}}, \\ 1 & \text{for } a \text{ even, } b = 1 \text{ and } 1 - \frac{1}{2^{(t-5)/2}} < \varepsilon \leq 1 - \frac{1}{2^{(t+7)/2}}, \\ 0 & \text{for } a \text{ even, } b = 1 \text{ and } 1 - \frac{1}{2^{(t+7)/2}} < \varepsilon < 1, \\ 0 & \text{for } a \text{ even and } b = 2. \end{cases}$$

Then there is no $[n, n-t, 5]$ *code for* $n = \left\lfloor \sqrt{2^{t+1}} + 0.5 \right\rfloor - \lambda_{a,b,\varepsilon}$*, hence*

$$s_{max}(\mathbb{F}_2^t) \leq \begin{cases} 2^{(t+1)/2} - 2 & \text{for } t \text{ odd,} \\ \left\lfloor \sqrt{2^{t+1}} + 0.5 \right\rfloor - \lambda_{a,b,\varepsilon} & \text{for } t \text{ even.} \end{cases} \tag{2}$$

On the codes side, Theorem 3.6 improves for t even and $t \geq 16$ several entries in the codes table of the MinT project [6] (`http://mint.sbg.ac.at/`). Some examples are listed in Corollary 3.7.

Corollary 3.7. *The following codes have maximal minimum distance 4 (instead of 4-5 as given in [6]):*

t	*16*	*18*	*20*	*22*	*24*	*26*
$[n,k]$ *code*	$[360, 344]$	$[723, 705]$	$[1446, 1426]$	$[2895, 2873]$	$[5791, 5767]$, $[5792, 5768]$	$[11583, 11557]$

We finish by giving Table 1 about the maximal size of Sidon sets and related bounds/constructions in small dimensions. The codes bound mentioned in this table arises from Proposition 3.3 (b) and Grassl's codes table [4] (http://codetables.de). Similarly, the codes construction comes from Proposition 3.3 (c) and Grassl's codes table.

Table 1: Maximal size of a Sidon set in $\mathbb{F}_2^t$ and related bounds/constructions.

t	4	5	6	7	8	9	10	11	12	13	14	15
Bound (1)	6	8	11	16	23	32	45	64	91	128	181	256
New bound (2)			10	14	21	30	43	62	90	126	180	254
Codes bound	6	7	9	12	18	24	34	58	89	125	179	254
$s_{max}(\mathbb{F}_2^t)$	6	7	9	12	18	24	34	?	?	?	?	?
Codes construction	6	7	9	12	18	24	34	48	66	82	129	152

References

[1] L. Babai and V. T. Sós, "Sidon sets in groups and induced subgraphs of Cayley graphs," *European Journal of Combinatorics*, vol. 6, no. 2, pp. 101–114, 1985 (cited on page 1).

[2] A. E. Brouwer and L. M. G. M. Tolhuizen, "A sharpening of the Johnson bound for binary linear codes and the nonexistence of linear codes with Preparata parameters," *Designs, Codes and Cryptography*, vol. 3, no. 2, pp. 95–98, May 1993 (cited on page 4).

[3] G. Cohen and G. Zémor, "Subset sums and coding theory," en, in *Structure theory of set addition*, ser. Astérisque 258, D. Jean-Marc, L. Bernard and Y. A. A., Eds., Société mathématique de France, 1999 (cited on page 3).

[4] M. Grassl, *Bounds on the minimum distance of linear codes and quantum codes*, Online available at `http://www.codetables.de`, Accessed on 2022-11-02, 2007 (cited on page 5).

[5] B. Green and I. Z. Ruzsa, "Sum-free sets in abelian groups," *Israel Journal of Mathematics*, vol. 147, no. 1, pp. 157–188, Dec. 2005 (cited on page 2).

[6] R. Schürer and W. C. Schmid, "Mint: A database for optimal net parameters," in *Monte Carlo and Quasi-Monte Carlo Methods 2004*, Springer, 2006, pp. 457–469 (cited on page 5).

[7] S. Sidon, "Ein Satz über trigonometrische Polynome und seine Anwendung in der Theorie der Fourier-Reihen," *Mathematische Annalen*, vol. 106, no. 1, pp. 536–539, Dec. 1932 (cited on page 1).

[8] S. Sidon, "Über die Fourier Konstanten der Funktionen der Klasse L_p für $p > 1$," *Acta Univ. Szeged Sect. Sci. Math*, vol. 7, pp. 175–176, 1935 (cited on page 1).

[9] M. Tait and R. Won, "Improved bounds on sizes of generalized caps in $AG(n, q)$," *SIAM Journal on Discrete Mathematics*, vol. 35, no. 1, pp. 521–531, 2021 (cited on page 4).

[10] T. Tao and V. Vu, "Sum-free sets in groups: A survey," *Journal of Combinatorics*, vol. 8, no. 3, pp. 541–552, 2017 (cited on page 2).