# Optimal self-orthogonal binary linear codes

Sascha Kurz*
Dept. of Mathematics, University of Bayreuth, 95440 Bayreuth, Germany
email: sascha.kurz@uni-bayreuth.de

November 17, 2026

**Abstract**

We complete the determination of the minimum possible length of eight-dimensional binary self-orthogonal codes with given minimum distance. For dimension nine we give exact values and bounds for minimum distance at most 64. By exploiting the geometric reformulation of linear codes as multisets of points, we fully redo the classification for dimension at most seven in a simplified and computer-free manner. As a by-product we prove the non-existence of $[93, 9, 44]_2$ codes and show several divisibility results for putative optimal nine-dimensional binary codes.

**Keywords:** self-orthogonal codes · linear codes · Griesmer bound · Galois geometry · residual codes
**2020 Mathematics Subject Classification:** Primary 94B05; Secondary 51E21, 94B27, 94B65

## 1 Introduction

A linear code is called self-orthogonal if it is contained in its dual code w.r.t. the standard inner product. They have been studied since the theory of error-correcting codes transitioned from an engineering tool into a mathematical discipline, see e.g. [66]. There are e.g. connections to lattices [77], designs [3, 75], groups [21, 23], DNA codes [69, Section 8], and quantum error-correcting codes [17]. We call a linear $[n, k, d]_q$ code (length) optimal if there does not exist a linear $[n', k, d]_q$ code with $n' < n$. By $n_q(k, d)$ we denote the minimum possible length of an $[n, k, d]_q$ code. The function $n_2(k, \cdot)$ is completely known for all $k \leq 8$ [11]. The majority of the current knowledge on $n_q(k, d)$ can be looked up in the online table [35]. If we additionally require that the linear code has to be self-orthogonal we use the notation $\tilde{n}_q(k, d)$ for the minimum possible length. In terms of distance optimality, $\tilde{n}_2(4, \cdot)$ was completely determined in [61]. In a sequence of papers this was extended to $\tilde{n}_2(k, \cdot)$ for all $k \leq 7$, see [67] and the references cited therein.

*I am grateful to Hirotaka Shimizu for several remarks on an earlier draft.

Here we completely determine $\tilde{n}_2(8,\cdot)$, cf. [55] where the optimal parameters of 8-dimensional binary codes corresponding to additive quaternary codes were recently determined. For the binary case dimension 8 currently seems to be a frontier since still many values of $n_2(9,d)$ are unknown. So, in Theorem 5 and Corollary 6 we give a few exact values and bounds for $\tilde{n}_2(9,d)$ when $d \leq 64$. See e.g. [1, 64, 70] for some very recent activity in this direction. While the problem can be formulated for arbitrary field sizes $q$, here we mainly restrict ourselves to $q = 2$. For recent constructions of self-orthogonal linear codes over $\mathbb{F}_q$ with $q \geq 3$ see e.g. [60, 63, 80, 82, 84]. Full classifications of self-orthogonal codes with small parameters have been obtained in [13] for $q = 2$ and in [12] for $q \in \{3, 4\}$. There is a well known correspondence between linear codes and multisets of points in projective spaces [29], see e.g. [58] for a survey on known results for linear codes and their geometric reformulations. However, in the context of self-orthogonal codes we are not aware of papers extensively using these methods. So, we will do so in this paper.

The remaining part is structured as follows. In Section 2 we introduce the necessary preliminaries before we determine $\tilde{n}_2(k,\cdot)$ for all $k \leq 7$ in Section 3. The determination of $\tilde{n}_2(8,\cdot)$ is presented in Section 4, see Theorem 2. In Section 5 we state and improve the current knowledge on $\tilde{n}_2(9,d)$ for $d \leq 64$, see Theorem 5. As a by-product we prove the non-existence of $[93, 9, 44]_2$ codes in Theorem 4 and show several divisibility results for putative optimal nine-dimensional binary codes in Proposition 10. Generator matrices of examples found by computer searches are listed in Appendix A.

# 2 Preliminaries

In this section we introduce the basic notation and first auxiliary results. We start with the coding theoretic formulation in Subsection 2.1. The well-known reformulation of linear codes as multisets of points in projective geometries, see e.g. [76, §1.1.2] or [29], is briefly outlined in Subsection 2.2. For the reader's sake we provide proofs whenever they are short, instructive, or we did not find our precise formulation in the literature, see e.g. the characterization results for binary self-orthogonal codes in Proposition 1 and Lemma 4. However, we claim no priority at all. It is well known that the minimum possible length of a linear code with given dimension, field size, and minimum distance $d$ is given by the so-called Griesmer bound, see Inequality (1) for all sufficiently large $d$ and attained by the Solomon–Stiffler construction [36, 73]. If we additionally require the linear codes to be self-orthogonal, we have to slightly modify the Griesmer bound, see Inequality (5), and the attaining construction, see Subsection 2.2. In order to make the paper self-contained we provide more technical details on the Griesmer bound and the Solomon–Stiffler construction in Subsection 2.3.

## 2.1 Coding theoretic notation

Let $\mathbb{F}_q$ denote the field with $q$ elements and $[n] := \{1, \ldots, n\}$. A *linear* $[n,k]_q$ code $C$ is a $k$-dimensional subspace of $\mathbb{F}_q^n$. If the rowspace of $G \in \mathbb{F}_q^{k\times n}$ equals $C$, we speak of a *generator matrix*. The parameter $n$ is called the *length* and the parameter $k$ is called the *dimension* of $C$. Elements of $C$ are called *codewords*. For any $c = (c_1, \ldots, c_n) \in \mathbb{F}_q^n$ the *support* of $c$ is defined as $\operatorname{supp}(c) := \{i \in [n] : c_i \neq 0\}$ and the size of the support $|\operatorname{supp}(c)|$ is called the *Hamming weight* $\operatorname{wt}(c)$ of $c$. Multiplying all entries at some position $1 \leq i \leq n$ by some $\lambda \in \mathbb{F}_q^* := \mathbb{F}_q \backslash \{0\}$ does not change the weights of the codewords, which is also true for permutations of the columns. In those cases we call the corresponding codes *equivalent*.[1] Having columns of generator matrices in mind, we also call $g \in \mathbb{F}_q^k$ and $\lambda g$ for some $\lambda \in \mathbb{F}_q^*$ equivalent. The *minimum (Hamming) distance* of a linear code $C$ is defined as the smallest non-zero Hamming weight $\operatorname{wt}(c)$ over all codewords $c \in C$ and the support of $C$ is given by $\operatorname{supp}(C) = \cup_{c\in C} \operatorname{supp}(c)$. We call $|\operatorname{supp}(C)|$ the *effective length* and of *full length* if $|\operatorname{supp}(C)| = n$. If $C$ is an $[n,k]_q$ code with minimum distance at least $d$[2], we speak of an $[n,k,d]_q$ code. By $n_q(k,d)$ we denote the minimum possible length of an $[n,k,d]_q$ code, which is lower bounded by the so-called *Griesmer bound* [36, 73]

$$n_q(k,d) \geq \sum_{i=0}^{k-1} \left\lceil \frac{d}{q^i} \right\rceil =: g_q(k,d). \tag{1}$$

If $n = g_q(k,d)$, then we speak of a *Griesmer code*. They exist for all sufficiently large $d$ [73], so that $n_q(k,d) = g_q(k,d)$ and the determination of $n_q(k,d)$ as a function of $d$ becomes a finite problem. For the current state of knowledge on bounds for $n_q(k,d)$ we refer to the online table [35]. Interpreted as columns of a generator matrix, a $k$-dimensional *simplex code* $\mathcal{S}_q(k)$ over $\mathbb{F}_q$ is given by the $[k]_q := \frac{q^k-1}{q-1}$ non-equivalent vectors in $\mathbb{F}_q^k \backslash \{0\}$ and the *first order Reed–Muller code* $\mathcal{RM}_q(1,k)$ is given by the $q^k$ vectors in $\mathbb{F}_q^{k+1}$ starting with a 1. These are $\left[[k]_q, k, q^{k-1}\right]_q$ and $\left[q^k, k+1, q^k - q^{k-1}\right]_q$ codes, respectively, see e.g. [43]. A linear code is called $\Delta$-*divisible* if the weights of all codewords are divisible by $\Delta$. E.g., $\mathcal{S}_q(k)$ and $\mathcal{RM}_q(1,k)$ are Griesmer codes and $q^{k-1}$-divisible for all $k \geq 1$.

**Theorem 1.** *([79, Theorem 1], [57], or [32] for $q = 2$) If $C$ is an $[n,k,d]_p$ Griesmer code, where $p$ is a prime, and $p^r$ divides $d$, then $C$ is $p^r$-divisible.*

For the divisibility of Griesmer codes when the field size $q$ is a prime power see e.g. [25, 26, 83]. In Lemma 26 we will show a divisibility criterion for binary non-Griesmer codes. For each $S \subseteq [n]$ by $C_S$ we denote the linear code consisting of all elements of an $[n,k]_q$ code $C$ restricted to the coordinates in $S$. With this the *residual code* $\operatorname{Res}(C;c)$ of $C$ w.r.t. a codeword $c \in C$ is given

[1] For non-prime fields $\mathbb{F}_q$ also the application of a field automorphism to the elements in all coordinate positions yields codes that are equivalent in the same sense.

[2] We use this convention to avoid the notation of an $[n,k,\geq d]_q$ code at several places. If needed in this paper, the existence of a codeword of weight $d$ will always follow from $n < n_q(k,d+1)$.

by $C_{[n]\setminus \mathrm{supp}(c)}$ and has length $n - \mathrm{wt}(c)$ as well as a dimension of at most $k-1$. If $\mathrm{wt}(c) < \frac{dq}{q-1}$ the dimension is exactly $k-1$. Unless we explicitly say otherwise, here a residual code is always meant w.r.t. a non-zero codeword.

**Lemma 1.** *(E.g. [27, 30]) If an $[n,k,d]_q$ code $C$ contains a codeword $c$ of weight $w$ with $w < \frac{dq}{q-1}$, then the residual code of $C$ with respect to $c$ is an $\left[n-w, k-1, d-w+\left\lceil \frac{w}{q} \right\rceil\right]_q$-code. The residual code has full length iff $C$ has.*

Some authors speak of the *one-step-Griesmer bound*. The exclusion of codeword weights $w$ via $n - w < n_q\left(k-1, d-w+\left\lceil \frac{w}{q} \right\rceil\right)$ and Lemma 1 is called *residual code argument*. The residual code of a Griesmer code w.r.t. a minimum weight codeword is a Griesmer code again, see e.g. [8, Lemma 29], [24], or Lemma 7. Applying Lemma 1 with minimum weight codewords recursively yields the Griesmer bound (1). The residual code of a codeword in a $q^r$-divisible $[n,k,d]_q$ code is $q^{r-1}$-divisible if $k, r \geq 1$, see e.g. [52, 78] for surveys on divisible codes. For $x, y \in \mathbb{F}_q^n$ the *scalar product* is given by $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$. With this the *dual* of an $[n,k]_q$ code $C$ is given by the $[n, n-k]_q$ code $C^\perp = \left\{y \in \mathbb{F}_q^n : \langle x, y\rangle = 0 \, \forall x \in C\right\}$. The minimum distance of the dual code is called the *dual minimum distance* and denoted by $d^\perp$. We call a linear code *projective* if $d^\perp \geq 3$. If $C \subseteq C^\perp$ then $C$ is called *self-orthogonal* (SO) and *self-dual* if $C = C^\perp$. So, an $[n,k]_q$ code $C$ is self-orthogonal iff $\langle c, c'\rangle = 0$ for all (possibly equal) pairs of (non-zero) codewords $c, c' \in C$. By $\tilde{n}_q(k,d)$ we denote the minimum possible length of an $[n,k,d]_q$ SO code. Clearly, we have $\tilde{n}_q(k,d) \geq n_q(k,d)$, which we call the *coding lower bound*.

Here we will mainly consider *binary codes*, i.e. $q = 2$.[3] Let $C$ be an $[n,k,d]_2$ code. We speak of *even*, *doubly-even*, and *triply-even* codes if $C$ is 2-, 4-, or 8-divisible, respectively. A code that is even but not doubly-even is called *singly-even*. Moreover, we have

$$\mathrm{wt}(c + c') = \mathrm{wt}(c) + \mathrm{wt}(c') - 2 \cdot |\mathrm{supp}(c) \cap \mathrm{supp}(c')| \tag{2}$$

for each $c, c' \in C$. Using the *Schur product* (Hadamard product) $c \star c'$, i.e. componentwise multiplication, we have

$$|\mathrm{supp}(c) \cap \mathrm{supp}(c')| = \mathrm{wt}(c \star c'). \tag{3}$$

If $c, c' \in C\setminus\{0\}$ with $c \neq c'$ and $c'_r$ denotes the codeword corresponding to $c'$ in $\mathrm{Res}(C; c)$, then we have

$$\mathrm{wt}(c'_r) = \mathrm{wt}(c') - |\mathrm{supp}(c) \cap \mathrm{supp}(c')| = \mathrm{wt}(c') - \mathrm{wt}(c \star c'). \tag{4}$$

Let us now characterize binary SO codes. Clearly, an $[n,1]_2$ code is SO iff it is even.

[3] Note that e.g. self-orthogonality is not preserved in general by monomial transforms for $q > 3$, see e.g. the generator matrices $\begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}$ and $\begin{pmatrix} 1 & 4 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}$ over $\mathbb{F}_5$. So, there e.g. does not exist an equivalent geometric reformulation in terms of multisets of points, see Subsection 2.2.

**Proposition 1.** *Let $C$ be an $[n,k]_2$ code, where $k \geq 2$. Then, the following statements are equivalent:*

*(A) $C$ is SO.*

*(B) $C$ is even and the doubly-even codewords form a subcode of dimension $k-1$ or $k$.*

*(C) All residual codes are even.*

*(D) The rows of a generator matrix $G$ have even weights and are pairwise orthogonal.*

*Proof.* Let us first assume that $C$ is SO, i.e. property (A). Since $\langle c, c\rangle = 0$ for every codeword we have $\mathrm{wt}(c) \equiv 0 \pmod 2$, i.e. $C$ is even. For two different codewords $\langle c, c'\rangle = 0$ implies $|\mathrm{supp}(c) \cap \mathrm{supp}(c')| = \mathrm{wt}(c \star c') \equiv 0 \pmod 2$. Using Equation (2) we conclude that the sum of two doubly-even codewords and the sum of two singly-even codewords is always doubly-even, i.e. the set of doubly-even codewords forms a subcode of dimension at least $k-1$. For a codeword $c'_r \in \mathrm{Res}(C;c)$ corresponding to $c' \in C$ Equation (4) gives $\mathrm{wt}(c'_r) \equiv 0 \pmod 2$, i.e., the residual code is even.

Now let us assume that $C$ satisfies (B). Since $C$ is even we have $\langle c, c\rangle = 0$. For each pair of different codewords $c, c' \in C$ property (B) implies that the number of singly-even codewords of $\{c, c', c'+c\}$ is either zero or two. Thus, Equation (2) implies $|\mathrm{supp}(c) \cap \mathrm{supp}(c')| = \mathrm{wt}(c \star c') \equiv 0 \pmod 2$, i.e. $\langle c, c'\rangle = 0$.

Let us now assume that $C$ satisfies (C). Due to Equation (4) it suffices to show that $C$ is even in order to conclude that $C$ is SO. From Equation (4) we conclude $\mathrm{wt}(c') \equiv \mathrm{wt}(c \star c') \equiv \mathrm{wt}(c) \pmod 2$ for all pairs of different codewords $c, c' \in C\backslash\{0\}$, i.e. the non-zero weights are either all even or all odd. The latter case is impossible for $k \geq 2$ due to Equation (2).

The implication (D) $\Rightarrow$ (A) follows from the bilinearity of the scalar product. The other direction is obvious. ∎

Property (B) is a refinement of the commonly stated implications that each SO binary code is even and that each doubly-even binary code is SO. Similarly, for each binary code Equation (2) implies that the set of even codewords forms a subcode of dimension at least $k-1$ in each binary $[n,k]_2$ code. The implication (A) $\Rightarrow$ (B) is e.g. mentioned in [22, Section II]. For the other direction see e.g. [59, Proposition 3.2]. Lemma 1 and property (C) directly imply the following improvement of Lemma 1 for binary SO codes.

**Lemma 2.** *If an $[n,k,d]_2$ SO code $C$ with $k \geq 2$ contains a codeword $c$ of weight $w$ with $w < 2d$ and $d - \frac{w}{2} \equiv 1 \pmod 2$, then the residual code of $C$ with respect to $c$ is an $\left[n-w, k-1, d-\frac{w}{2}+1\right]_2$-code.*

For $w = d$ Lemma 2 was shown in [61, Lemma 3.1] and [67, Lemma 3.7]. Since the residual of a Griesmer $[n,k,d]_2$ code w.r.t. a minimum weight codeword is a Griesmer $[n-d, k-1, \frac{d}{2}]_2$ code, Lemma 2 implies

$$\tilde{n}_2(k,d) \geq \begin{cases} g_2(k,d') & : \ d' \equiv 0 \pmod 4, \\ g_2(k,d')+1 & : \ d' \equiv 2 \pmod 4 \end{cases} =: \tilde{g}_2(k,d), \tag{5}$$

where $d' := 2\left\lceil \frac{d}{2} \right\rceil$ and we assume $k \geq 2$. We call Inequality (5) the *Griesmer SO bound* and remark that it is attained with equality for all sufficiently large $d$,[4] see [67] or Subsection 2.2. Thus, the determination of $\tilde{n}_2(k, \cdot)$ becomes a finite problem for each dimension $k$. The exclusion of codeword weights $w$ via $n - w < n_2\left(k-1, d - \frac{w}{2} + 1\right)$ and Lemma 2 is called *residual SO code argument*. Since the residual codes are even, we can deduce further exclusion criteria for large weights.

**Lemma 3.** *Let $C$ be an $[n,k,d]_2$ SO code with $k \geq 2$. If $C$ is full-length, then there cannot be a codeword of weight $n-1$. If $C$ is projective, then there cannot be a codeword of weight $n-2$.*

**Proposition 2.** *([65], MacWilliams Identities) Let $C$ be an $[n,k,d]_2$ code and $C^\perp$ be the dual code of $C$. Let $A_i(C)$ and $B_i(C)$ be the number of codewords of weight $i$ in $C$ and $C^\perp$, respectively. With this, we have*

$$\sum_{j=0}^{n} K_i(j) A_j(C) = 2^k B_i(C), \quad 0 \leq i \leq n \tag{6}$$

*where*

$$K_i(j) = \sum_{s=0}^{n} (-1)^s \binom{n-j}{i-s}\binom{j}{s}, \quad 0 \leq i \leq n$$

*are the binary Krawtchouk polynomials. We will simplify the notation to $A_i$ and $B_i$ whenever $C$ is clear from the context.*

We call the sequence $\left(i^{A_i}\right)_{i>0}$, where we leave out entries with a zero exponent, the *weight distribution* of $C$. Whenever we speak of the first $l$ MacWilliams identities, we mean Equation (6) for $0 \leq i \leq l-1$. Adding the non-negativity constraints $A_i, B_i \geq 0$ and $A_i = 0$ for forbidden weights $i$, we obtain a linear program where we can maximize or minimize certain quantities, which is called the *linear programming method for linear codes*. If there is no feasible solution, then there is no linear code with the specified parameters. Adding additional equations or inequalities strengthens the formulation. E.g. we can add integer rounding cuts. For an $[n,k,d]_2$ code of full length, i.e. $B_1 = 0$, we can rewrite the first four MacWilliams identities to

$$\sum_{i>0} A_i = 2^k - 1, \tag{7}$$

$$\sum_{i} iA_i = 2^{k-1} n, \tag{8}$$

$$\sum_{i} i^2 A_i = 2^{k-1} \cdot (B_2 + n(n+1)/2), \tag{9}$$

$$\sum_{i} i^3 A_i = 2^{k-2} \cdot \left(3(B_2 n - B_3) + n^2(n+3)/2\right). \tag{10}$$

[4] As mentioned before, the same property holds for linear codes over $\mathbb{F}_q$ and the ordinary Griesmer bound. Other instances of Griesmer-type bounds with this strong asymptotic equivalence are given by additive codes [49] and linear codes w.r.t. the $b$-symbol metric [53].

**Proposition 3.** *([28, Proposition 5], cf. [71]) Let $C$ be an $[n,k,d]$-code with all weights divisible by $\Delta := 2^a$ and let $(A_i)_{i=0,1,\ldots,n}$ be the weight distribution of $C$. Put*

$$\begin{aligned} \alpha &:= \min\{k-a-1, a+1\}, \\ \beta &:= \lfloor (k-a+1)/2 \rfloor, \text{ and} \\ \delta &:= \min\{2\Delta i \mid A_{2\Delta i} \neq 0 \wedge i > 0\}. \end{aligned}$$

*Then the integer*

$$T := \sum_{i=0}^{\lfloor n/(2\Delta) \rfloor} A_{2\Delta i}$$

*satisfies the following conditions.*

1. *$T$ is divisible by $2^{\lfloor (k-1)/(a+1) \rfloor}$.*
2. *If $T < 2^{k-a}$, then $T = 2^{k-a} - 2^{k-a-t}$ for some integer $t$ satisfying $1 \leq t \leq \max\{\alpha, \beta\}$. Moreover, if $t > \beta$, then $C$ has an $[n, k-a-2, \delta]$-subcode and if $t \leq \beta$, it has an $[n, k-a-t, \delta]$-subcode.*
3. *If $T > 2^k - 2^{k-a}$, then $T = 2^k - 2^{k-a} + 2^{k-a-t}$ for some integer $t$ satisfying $0 \leq t \leq \max\{\alpha, \beta\}$. Moreover, if $a = 1$, then $C$ has an $[n, k-t, \delta]$-subcode. If $a > 1$, then $C$ has an $[n, k-1, \delta]$-subcode unless $t = a+1 \leq k-a-1$, in which case it has an $[n, k-2, \delta]$-subcode.*

A special and well-known subcase is that the number of even weight codewords in an $[n,k]_2$ code is either $2^{k-1}$ or $2^k$.

The weight enumerator can be generalized to the split weight enumerator based on a partition of the coordinates [72]. For a linear programming method based on the split weight enumerator we refer e.g. to [44]. Here we will apply the split weight enumerator for a partition given by the support of some codeword and its complement.

## 2.2 Geometric notation

The set of all subspaces of $\mathbb{F}_q^r$, ordered by the incidence relation $\subseteq$, is called $(r-1)$-*dimensional projective geometry over* $\mathbb{F}_q$ and denoted by $\mathrm{PG}(r-1,q)$. Employing this algebraic notion of dimension instead of the geometric one, we will use the term $i$-space to denote an $i$-dimensional subspace of $\mathbb{F}_q^r$. To highlight the important geometric interpretation of subspaces we will call 1-, 2-, and $(r-1)$-spaces points, lines, and hyperplanes, respectively. For two subspaces $S$ and $S'$ we write $S \leq S'$ if $S$ is contained in $S'$. Moreover, we say that $S$ and $S'$ are *incident* iff $S \leq S'$ or $S \geq S'$. Let $[i]_q := \frac{q^i-1}{q-1}$ denote the number of points of an arbitrary $i$-space in $\mathrm{PG}(r-1,q)$ where $r \geq i$. A *premultiset of points* $\mathcal{M}$ of $\mathrm{PG}(r-1,q)$ is a mapping from the set of points of $\mathrm{PG}(r-1,q)$ to $\mathbb{Z}$. The value $\mathcal{M}(P)$ is called the *multiplicity* of $P$. We extend this notation additively to each subspace $S$ via $\mathcal{M}(S) = \sum_{P \leq S} \mathcal{M}(P)$. So, the cardinality of $\mathcal{M}$ equals $\mathcal{M}(V)$, where $V$ denotes the ambient space. For a hyperplane $H$ the number of points of $\mathcal{M}$ in $H$ is given by $\mathcal{M}(H)$. If

$\mathcal{M}(P) \in \mathbb{N}$ for each point $P$, we say that $\mathcal{M}$ is a *multiset of points*. For each subspace $S$ we denote its *characteristic function* by $\chi_S$, i.e. $\chi_S(P) = 1$ if $P \leq S$ and $\chi_S(P) = 0$ otherwise.

Starting from a generator matrix $G \in \mathbb{F}_q^{k\times n}$ of an $[n,k,d]_q$ code $C$ we can associate a multiset of points $\mathcal{M}$ in $\mathrm{PG}(k-1,q)$ by considering the spans of the columns of $G$. The cardinality of $\mathcal{M}$ is $\#\mathcal{M} = n$ iff $C$ is of full length. For the other direction we can build up a generator matrix $G$ of $C$ by taking $\mathcal{M}(P)$ copies of a representing vector for every point $P$. We call $\mathcal{M}$ in $\mathrm{PG}(r-1,q)$ *spanning* if the points with positive multiplicity, called the *support*, span the entire ambient space. The dimension of the associated linear code equals $r$ iff $\mathcal{M}$ is spanning. Let $C$ be a full length $[n,k]_q$ code and $\mathcal{M}$ in $\mathrm{PG}(k-1,q)$ its associated spanning multiset of points. We remark that the ambiguities in these constructions result in equivalent linear codes or multisets of points, see e.g. [76, §1.1.2] or [29] for more details. As an example we mention that $\mathcal{S}_q(k)$ and $\mathcal{RM}_q(1,k-1)$ correspond to $\chi_V$ and $\chi_V - \chi_H$, where $V$ denotes the ambient space $\mathrm{PG}(k-1,q)$ and $H$ is a hyperplane. Each non-zero codeword $c$ corresponds to a hyperplane $H$ and we have $\mathrm{wt}(c) = n - \mathcal{M}(H)$. For the other direction, each hyperplane corresponds to the $q-1$ multiples of a non-zero codeword. A multiset of points $\mathcal{M}$ is called *$m$-arc* if $\mathcal{M}(H) \leq m$ for every hyperplane $H$. I.e. a full-length $[n,k,d]_q$ code corresponds to a spanning $(n-d)$-arc. Similarly, a multiset of points $\mathcal{M}$ is called *$m$-minihyper* (or *$m$-fold blocking set w.r.t. hyperplanes*) if $\mathcal{M}(H) \geq m$ for every hyperplane $H$. If $\mathcal{M}$ is a spanning $m$-arc and $\mathcal{M}'$ is an $m'$-minihyper in $\mathrm{PG}(k-1,q)$, then $\mathcal{M} - \mathcal{M}'$ is a spanning $(m-m')$-arc iff it is a spanning multiset of points. In coding theoretic terms the deletion of $\mathcal{M}'$ is called *puncturing*. For a multiset of points $\mathcal{M}$ we denote the maximum point multiplicity by $\gamma(\mathcal{M})$, so that e.g. $\gamma(\mathcal{M}) = 1$ iff the corresponding linear code $C$ is projective. If $K$ is some subspace not equal to the ambient space $V$, then the *projection $\mathcal{M}_K$ of $\mathcal{M}$ through $K$* is the multiset of points in the factor geometry $V/K$ with multiplicity $\mathcal{M}_K(S/K) = \mathcal{M}(S) - \mathcal{M}(K)$ for every subspace $K \subseteq S \subseteq V$. The dimension of the ambient space drops by $\dim(K)$ and the cardinality drops by $\mathcal{M}(K)$, while the $m$-arc property is preserved. In coding theory terms $\mathcal{M}_K$ corresponds to a codimension $\dim(K)$ subcode of $C$ when $\mathcal{M}$ corresponds to $C$ and projection is called *truncation* or *shortening*. If $\mathcal{M}, \mathcal{M}'$ are spanning $m$- and spanning $m'$-arcs in $\mathrm{PG}(k-1,q)$, respectively, then $\mathcal{M} + \mathcal{M}'$ is a spanning $(m+m')$-arc in $\mathrm{PG}(k-1,q)$. If $\mathcal{M}'$ spans a $(k-1)$-space we can choose a hyperplane $H$ for the embedding, so that $\mathcal{M}+\mathcal{M}'$ is a spanning $(\leq \max\{m+m', \mathcal{M}(H)+\#\mathcal{M}'\})$-arc in $\mathrm{PG}(k-1,q)$. In terms of corresponding generator matrices $G, G'$ this corresponds to the juxtaposition $\left(G|\tilde{G}'\right)$ where $\tilde{G}'$ is obtained from $G'$ by suitably padding one row. An example is given by the unique $[38,6,18]_2$ Griesmer code that can be obtained in this way from a $[32,6,16]_2$ and a $[6,5,2]_2$ code.

We call a multiset of points $\mathcal{M}$ $\Delta$-divisible iff its corresponding linear code is, i.e. if we have $\mathcal{M}(H) \equiv \#\mathcal{M} \pmod{\Delta}$ for every hyperplane $H$. Similarly, we call a multiset of points in $\mathrm{PG}(k-1,2)$ self-orthogonal (SO) if its corresponding linear code is. Clearly, projection preserves divisibility and self-orthogonality.

**Lemma 4.** *Let $\mathcal{M}$ be a spanning multiset of points in* $\mathrm{PG}(k-1,2)$*, where $k \geq 2$. Then, $\mathcal{M}$ is SO iff $\mathcal{M}(H) \equiv \mathcal{M}(K) \equiv \#\mathcal{M} \pmod 2$ for every hyperplane $H$ and every subspace $K$ of codimension* 2*.*

*Proof.* By $C$ denote the corresponding $[n,k]_2$ code. Let $c$ denote an arbitrary non-zero codeword and $H$ the corresponding hyperplane. The statements $\mathrm{wt}(c) \equiv 0 \pmod 2$, $\mathcal{M}(H) \equiv \#\mathcal{M} \pmod 2$, and $\langle c, c\rangle = 0$ are pairwise equivalent. Now let $c, c'$ be two different non-zero codewords and $H, H'$ be the corresponding hyperplanes. We have $\langle c, c'\rangle = 0$ iff $|\mathrm{supp}(c) \cap \mathrm{supp}(c')| \equiv 0 \pmod 2$. Setting $K := H \cap H'$ we observe

$$\begin{aligned} |\mathrm{supp}(c) \cap \mathrm{supp}(c')| &= \#\mathcal{M} - \mathcal{M}(H \cup H') = \#\mathcal{M} - \mathcal{M}(H) - \mathcal{M}(H') + \mathcal{M}(K) \\ &= (\#\mathcal{M} - \mathcal{M}(H)) + (\#\mathcal{M} - \mathcal{M}(H')) - (\#\mathcal{M} - \mathcal{M}(K)) \equiv \#\mathcal{M} - \mathcal{M}(K) \pmod 2. \end{aligned}$$

■

**Corollary 1.** *If $\mathcal{M}_1$ is an arbitrary multiset of points and $\mathcal{M}_2, \mathcal{M}_3$ are SO multisets of points in* $\mathrm{PG}(k-1,2)$*, then $2\lambda \cdot \mathcal{M}_1$, $\mathcal{M}_2 + \mathcal{M}_3$, and $\lambda \cdot \mathcal{M}_2$ are SO multisets of points for all $\lambda \in \mathbb{N}_{\geq 1}$. If $\mathcal{M}_2 - \mathcal{M}_3$ is a multiset of points, then it is also SO.*

So, $\mathcal{S}_2(k)$ and $\mathcal{RM}_2(1,k)$ are SO iff $k \geq 3$.

**Corollary 2.** *Let $\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2$ be the sum of two multisets of points in* $\mathrm{PG}(k-1,2)$*, where $\mathcal{M}_1$ is SO. Then, $\mathcal{M}$ is SO iff $\mathcal{M}_2$ is SO.*

In coding theory terms, Corollary 2 was shown in [68, Lemma 3.1] for the special case where $\mathcal{M}_1$ in $\mathrm{PG}(k-1,2)$, where $k \geq 3$, has constant point multiplicity. In [67, Lemma 3.1] the extra assumption on $\mathcal{M}_1$ was removed. Corollary 1 directly implies the following construction, cf. [67].

**Lemma 5.** *Let $\mathcal{M}$ be a spanning SO multiset of points in* $\mathrm{PG}(k-1,2)$ *such that $\mathcal{M}(H) \leq \#\mathcal{M} - d$ for every hyperplane $H$ and $P$ some point with $\mathcal{M}(P) \geq 3$. Then, $\mathcal{M} - 2\chi_P$ corresponds to an $[\#\mathcal{M} - 2, k, d-2]_2$ SO code.*[5]

With this, we can easily show that the Griesmer SO bound in Inequality (5) can always be attained with equality if $d$ is sufficiently large. For $d \equiv 0 \pmod 4$ any $[g_2(k,d), k, d]_2$ Griesmer code is 4-divisible by Theorem 1, which implies SO. For $d \equiv 2 \pmod 4$ we combine the same reasoning to $d' = d + 2 \equiv 0 \pmod 4$ with Lemma 5. Since sufficiently large $d$ implies sufficiently large $n$ and the existence of a point $P$ with multiplicity at least 3, it just remains to show $g_2(k,d') = g_2(k,d) + 3$, see e.g. [67] or Lemma 7 for details.

It is a nice exercise to show that a multiset of points $\mathcal{M}$ in $\mathrm{PG}(k-1,2)$, where $k \leq 4$, is SO iff it can be recursively obtained from Corollary 1 using just the characteristic functions of 3- and 4-spaces. For $k > 4$ there are more SO multisets, see e.g. [7, Appendix B].

[5] Note that $\mathcal{M}(P) \geq 2$ is sufficient if $d > 2$.

The first three MacWilliams identities can be shown by double-counting in the geometric setting, see e.g. [52, Section 2.2]. Let $C$ be a linear code, $\mathcal{M}$ its corresponding multiset of points, $c \in C$, and $H$ the hyperplane corresponding to $c$. The partition induced by the support of $c$ and its complement corresponds to the restriction of $\mathcal{M}$ to $H$ and its complement $\overline{H}$. Over $\mathbb{F}_2$, a dual codeword of weight three geometrically corresponds to a line $L$. The fact that $L$ intersects $H$ in either three or one point directly translates to a condition for the dual variables in the split weight enumerator for $c$.

## 2.3 Parameters of Griesmer codes and the Solomon–Stiffler construction

Here we want to briefly describe the Solomon–Stiffler construction [73] from 1965 that attains the Griesmer bound for all parameters if the minimum distance is sufficiently large.[6] Using $[i]_q = \frac{q^i-1}{q-1}$ for the number of points of an $i$-space in $\mathrm{PG}(k-1,q)$ we consider the following parameterization of the Griesmer bound.

**Lemma 6.** *Let $k$ and $d$ be positive integers. Write $d$ as*

$$d = \sigma q^{k-1} - \sum_{i=1}^{k-1} \varepsilon_i q^{i-1}, \tag{11}$$

*where $\sigma \in \mathbb{N}_0$ and the $0 \leq \varepsilon_i < q$ are integers for all $1 \leq i \leq k-1$. Then, Inequality (1) is satisfied with equality iff*

$$n = \sigma [k]_q - \sum_{i=1}^{k-1} \varepsilon_i [i]_q, \tag{12}$$

*which is equivalent to*

$$n - d = \sigma [k-1]_q - \sum_{i=1}^{k-1} \varepsilon_i [i-1]_q. \tag{13}$$

**Remark 1.** *Given $k$ and $d$ Equation (11) always determines $\sigma$ and the $\varepsilon_i$ uniquely. This is different for Equation (13) given $k$ and $n - d = s$. Here it may happen that no solution with $0 \leq \varepsilon_i \leq q-1$ exists. By relaxing to $0 \leq \varepsilon_i \leq q$ we can ensure existence and uniqueness is enforced by additionally requiring $\varepsilon_j = 0$ for all $j < i$ where $\varepsilon_i = q$ for some $i$, and by additionally specifying $\varepsilon_1$. The same is true for Equation (12) given $k$ and $n$. For more details we refer to [34, Chapter 2] which also gives pointers to Hamada's work on minihypers, see e.g. [37].*

Lemma 6 and Lemma 1 imply the following observations.

**Lemma 7.** *Let $k \geq 2$.*

[6] While it is widely known, there still appear papers claiming new constructions of linear codes with desired properties that fall into this class, cf. [81].

(1) *The residual code of an* $[n,k,d]_q$ *Griesmer code w.r.t. a codeword of weight* $d$ *is an* $\left[n-d,k-1,\left\lceil\frac{d}{q}\right\rceil\right]_q$ *Griesmer code.*

(2) *For each positive integer* $r$ *we have* $g_q\left(k,d+rq^{k-1}\right)=g_q(k,d)+r\cdot\frac{q^k-1}{q-1}$.

(3) *We have* $d\equiv 0 \pmod{q^{k-1}}$ *iff* $\varepsilon_i=0$ *for all* $1\le i\le k-1$. *Otherwise we have* $d\equiv 0 \pmod{q^{i-1}}$ *and* $d\not\equiv 0 \pmod{q^i}$ *for the smallest index* $1\le i\le k-1$ *with* $\varepsilon_i\neq 0$.

(4) *For* $d\equiv 0 \pmod 4$ *we have* $g_2(k,d)=g_2(k,d-2)+3$.

For Lemma 7.(1) see e.g. [8, Lemma 29], [24] and for Lemma 7.(2) see e.g. [8, Lemma 21], [27]. Clearly also Lemma 7.(4) can be generalized to arbitrary field sizes $q$ and different powers of $q$ for the step size.

Using the notation and restrictions from Lemma 6, fixing some proper subspaces $S_1,\dots,S_l$, and denoting the ambient space $\mathrm{PG}(k-1,q)$ by $V$, we can state a geometric variant of the Solomon–Stiffler construction. We say that $\mathcal{M}=\sigma\chi_V-\sum_{j=1}^{l}\chi_{S_j}$ is of *Solomon–Stiffler type* $[\sigma;\varepsilon_1,\dots,\varepsilon_{k-1}]$, where $l=\sum_{i=1}^{k-1}\varepsilon_i$, iff exactly $\varepsilon_i$ subspaces $S_j$ have dimension $1\le i\le k-1$. If $\mathcal{M}$ is a multiset of points, i.e. if $\mathcal{M}(P)\ge 0$ for all points $P$, then $\mathcal{M}$ corresponds to a Griesmer code. An easy sufficient condition is $\sigma\ge\sum_{i=1}^{k-1}\varepsilon_i$, where the right hand side is upper bounded by $(q-1)(k-1)$. For more refined sufficient conditions allowing a suitable selection of the $S_j$ we e.g. refer to [4]. Codes of Solomon–Stiffler type are not the only known class of Griesmer codes, see e.g. [4, 20, 38, 40, 41, 42]. For $q=2$ and $d\le 2^{k-1}$ Griesmer codes have to be either of Solomon–Stiffler or Belov type [4], see [39]. The latter case cannot occur if $d$ is doubly-even and a singly-even example is given by the unique $[38,6,18]_2$ code. So, due to the Griesmer SO bound in Inequality (5) a projective Griesmer SO code has to be of Solomon–Stiffler type cf. [67]. The multiset of points $\sigma\cdot\chi_V-\mathcal{M}$, i.e. the multiset that is subtracted, is an optimal minihyper (anticode in coding theory terms). For corresponding classification results we refer e.g. to [37, 74].

# 3 Optimal binary self-orthogonal codes with dimension at most seven

The determination of the minimum possible length $\tilde{n}_2(k,d)$ of a self-orthogonal $k$-dimensional binary linear code with minimum distance $d$ was completed for $k\le 7$ in [67], see e.g. [13, 47, 48, 61, 68] for previous results. In this section we briefly redo the entire computation and simplify some arguments. All of our lower bounds have computer-free proofs, while e.g. the non-existence of $[41,7,18]_2$, $[56,7,26]_2$, $[63,7,30]_2$, and $[103,7,50]_2$ SO codes was shown in [67] using exhaustive enumeration via the software package `Q-EXTENSION` [14]. We also state general constructions that cover cases that previously relied on computer searches.

First we summarize some easy improvements of $\tilde{n}_2(k,d)\ge\tilde{g}_2(k,d)$.

**Lemma 8.** *Let $d$ be even and $k \geq 2$.*

*(1) If $d \equiv 2 \pmod 4$ and $n < n_2(k-1, d+2)$, then no $[n,k,d]_2$ SO code exists.*
*(2) For integers $2 \leq t \leq k-1$ and $d = 2^t - 2$ we have $\tilde{n}_2(k,d) \geq 2^{t+1} + k - t - 3 = \tilde{g}_2(k,d) + 1$.*
*(3) For integers $2 \leq t \leq k-2$ and $d = 2^{k-1} - 2^t - 2$ we have $\tilde{n}_2(k,d) \geq 2^k - 2^{t+1} - 1 = \tilde{g}_2(k,d) + 1$.*
*(4) If $n_2(k,d+2) > n$ and $n - d < n_2\left(k-1, 2\left\lceil \frac{d}{4} \right\rceil\right)$, then no $[n,k,d]_2$ SO code exists.*

*Proof.* Proposition 1.(B) implies the existence of an $[\leq n, k-1, d+2]_2$ subcode, which yields (1). For an $[n', k-1, 2^t]_2$ code $C'$, we write $2^t = 2 \cdot 2^{k-2} - \sum_{i=t}^{k-2} 2^i$, so that

$$n' \geq g_2\left(k-1, 2^t\right) = 2 \cdot \left(2^{k-1} - 1\right) - \sum_{i=t}^{k-2} \left(2^{i+1} - 1\right) = 2^k - 2 - 2^{t+1} \cdot \left(2^{k-t-1} - 1\right) + (k-t-1) = 2^{t+1} + k - t - 3.$$

Similarly, we compute $g_2(k, 2^t - 2) = 2^{t+1} + k - t - 5$, so that $\tilde{n}_2(k, 2^t - 2) \geq n' \geq 2^{t+1} + k - t - 3 = g_2(k, 2^t - 2) + 2 = \tilde{g}_2(k, 2^t - 2) + 1$, i.e. statement (2). For an $\left[n'', k-1, 2^{k-1} - 2^t\right]_2$ code $C''$, we write $2^{k-1} - 2^t = 2 \cdot 2^{k-2} - 2^t$, so that

$$n'' \geq g_2\left(k-1, 2^{k-1} - 2^t\right) = 2 \cdot \left(2^{k-1} - 1\right) - \left(2^{t+1} - 1\right) = 2^k - 2^{t+1} - 1$$

and

$$g_2\left(k, 2^{k-1} - 2^t - 2\right) = \left(2^k - 1\right) - \left(2^{t+1} - 1\right) - 3 = 2^k - 2^{t+1} - 3,$$

which implies statement (3). For (4) we apply Lemma 2 with a codeword of minimum weight $d$ (using $n_2(k, d+2) > n$ and the non-existence of codewords of weight $d+1$). ∎

Applying Lemma 8.(2) with $t \in \{k-1, k-2\}$ gives [62, Theorem 6] and applying Lemma 8.(3) with $t = u - 1$ gives [62, Theorem 9]. Combining Lemma 8.(1) with known lower bounds for $n_2(k,d)$ [35] yields e.g. the following non-existence results.

**Corollary 3.** *There are no* $[13,4,6]_2$, $[14,5,6]_2$, $[22,5,10]_2$, $[29,5,14]_2$, $[15,6,6]_2$, $[23,6,10]_2$, $[30,6,14]_2$, $[39,6,18]_2$, $[46,6,22]_2$, $[54,6,26]_2$, $[61,6,30]_2$, $[17,7,6]_2$, $[25,7,10]_2$, $[41,7,18]_2$, $[47,7,22]_2$, $[62,7,30]_2$, $[72,7,34]_2$, $[79,7,38]_2$, $[87,7,42]_2$, $[94,7,46]_2$, $[103,7,50]_2$, $[110,7,54]_2$, $[118,7,58]_2$, $[125,7,62]_2$, $[18,8,6]_2$, $[26,8,10]_2$, $[34,8,14]_2$, $[42,8,18]_2$, $[49,8,22]_2$, $[58,8,26]_2$, $[74,8,34]_2$, $[81,8,38]_2$, $[137,8,66]_2$, $[144,8,70]_2$, $[152,8,74]_2$, $[159,8,78]_2$, $[168,8,82]_2$, $[175,8,86]_2$, $[183,8,90]_2$, $[190,8,94]_2$, $[200,8,98]_2$, $[207,8,102]_2$, $[215,8,106]_2$, $[222,8,110]_2$, $[231,8,114]_2$, $[238,8,118]_2$, $[246,8,122]_2$, *and* $[253,8,126]_2$ *SO codes.*

Combining Lemma 8.(4) with known lower bounds for $n_2(k,d)$ [35] yields e.g. the following non-existence results.

**Corollary 4.** *There are no* $[105,8,50]_2$, $[112,8,54]_2$, *and* $[129,9,62]_2$ *SO codes.*

For the constructive side we have:

**Lemma 9.** *Let $d$ be even and $k \geq 2$.*

*(1) If $d \equiv 0 \pmod 4$ and a $[g_2(k,d),k,d]_2$ Griesmer code exists, then it is self-orthogonal.*
*(2) If $d \equiv 2 \pmod 4$, $d > 2^{k-1}$ and a $[g_2(k,d+2),k,d+2]_2$ Griesmer code exists, then there is a $[g_2(k,d)+1,k,d]_2$ self-orthogonal code.*
*(3) If $d \equiv 2 \pmod 4$, $d > 2^{k-1}$ and a doubly-even $[n+2,k,d+2]_2$ code exists, then there is an $[n,k,d]_2$ self-orthogonal code.*

*Proof.*

(1) By Theorem 1 each such code is doubly-even and thus self-orthogonal.

(2) Let $C$ be the initial Griesmer code and $\mathcal{M}$ be the corresponding multiset of points. As observed in (1), $C$ and $\mathcal{M}$ are self-orthogonal. Since $d > 2^{k-1}$ we have $\#\mathcal{M} \geq 2^k$ and there exists a point $P$ with multiplicity $\mathcal{M}(P) \geq 2$. With this $\mathcal{M} - 2\chi_P$ is SO, has minimum distance at least $d$, and cardinality $g_2(k,d+2) - 2 = g_2(k,d) + 1$, see Subsection 2.3. Since $d > 2^{k-1} \geq 2$ we have that $\mathcal{M} - 2\chi_P$ is spanning.

(3) Let $C$ be the initial doubly-even code and $\mathcal{M}$ be the corresponding multiset of points. Similar to (2) we conclude that $C$ is self-orthogonal and $\mathcal{M} - 2\chi_P$ has the desired properties for some suitable point $P$.

■

To treat the case $d \equiv 2 \pmod 4$ for $d \leq 2^{k-1}$ we first construct special non-projective codes and then apply the idea of the Solomon–Stiffler construction, i.e. the removal of suitable minihypers.

**Lemma 10.** *For $k \geq r \geq 2$ let $V$ denote $\mathrm{PG}(k-1,2)$, $A \leq V$ an $r$-space, and $B \leq A$ an $(r-1)$-space. Then, $\chi_V + \chi_A - 2\chi_B$ corresponds to a $2^{r-1}$-divisible $\left[2^k,k,2^{k-1}\right]_2$ non-projective two-weight code with maximum weight $2^{k-1} + 2^{r-1}$.*

*Proof.* Consider the different possible intersections of hyperplanes with $A$ and $B$. ■

We remark that the six non-equivalent full-length $[64,6,32]_2$ codes are all of this type (allowing $r = 1$ as a degenerated case). Since the construction guarantees points of multiplicity 2 for $r \neq k$ and self-orthogonality for $r \geq 3$, we can obtain a $[62,6,30]_2$ SO code, that turns out to be optimal. Similar to the Solomon-Stiffler construction we can remove several suitable selected subspaces. To this end, we mention that for each $k/2 \leq r < k$ with $k \geq 3$ there exists a partition of $\mathrm{PG}(k-1,q)$ into one $r$-space and $q^r$ $(k-r)$-spaces [5, 16].

**Proposition 4.** *Let $k \geq 5$, $d \equiv 2 \pmod 4$ with $d \leq 2^{k-1}$, and $\varepsilon_1, \ldots, \varepsilon_{k-1}$ be uniquely defined as in Lemma 6. Denote the largest index of the $\varepsilon_i \neq 0$ by $j$ and the second largest by $h$, where we set $h = -\infty$ if there is no second largest index.*

*(1) If $j = 2$, then there exists a $[g_2(k,d)+2, k, d]_2$ SO code.*
*(2) If $j = k-2$ and $d = 2^{k-1} - 2^{j-1} - 2$, then there exists a $[g_2(k,d)+2, k, d]_2$ SO code.*
*(3) If $3 \leq j \leq k-3$ and $j + h \leq k$, then there exists a $[g_2(k,d)+2, k, d]_2$ SO code.*

*Proof.* We construct a corresponding multiset of points in $\mathrm{PG}(k-1, 2)$. Denote the ambient space by $V$.

(1) Here we have $d = 2^{k-1} - 2$ and a construction is given by Lemma 10 with $r = 3$.

(2) Let $K$ be a $(k-2)$-space, $B$ a disjoint line, and $A$ a plane that contains $B$. Let $Q$ be the point that equals $K \cap A$, and $P \neq Q$ be a point that is contained in $A$ but not in $B$. With this, the desired multiset of points is given by $\chi_V - \chi_K + \chi_A - 2\chi_B - 2\chi_P$.

(3) If $h = 2$, then choose a $j$-space $K$ and proceed as in (2). If $h > 2$, then set $j' := \max\{j, \lceil k/2 \rceil\}$, choose a $j'$-space $K'$, a $j$-space $K$ inside $K'$, and a (vector space) partition of $V$ into $K'$ and $l := 2^{j'}$ $(k-j')$-spaces $T_1, \ldots, T_l$. For each $3 \leq i \leq h$ choose an $i$-space $S_i$ inside the first $m := h - 2$ partition elements $T_r$ so that they are pairwise different. Choose a plane $A$ in $T_{m+1}$ and a line $B$ that is contained in $A$ as well as a point $P \notin K$ that is contained in $A$ but not in $B$. Since $h \leq k - j'$ and $l \geq 2^{k/2} \geq k/2 \geq m + 1$, this is always possible. With this, the desired multiset of points is given by $\chi_V - \chi_K - \sum_{i=3}^{h} \varepsilon_i \chi_{S_i} + \chi_A - 2\chi_B - 2\chi_P$.

■

We remark that the construction in Proposition 4 e.g. yields a $[95, 7, 46]_2$ SO code, cf. [67] where such a code was found by a computer search. We can construct another family of non-projective two-weight SO codes.

**Lemma 11.** *For $k \geq 4$ there exists a $\left[2^k - 4, k, 2^{k-1} - 4\right]_2$ two-weight SO code with second weight $2^{k-1}$.*

*Proof.* Let $V$ denote the ambient space $\mathrm{PG}(k-1, 2)$, $E$ a plane, and $P \in E$ a point. Then counting points for the possible hyperplane intersections yields that the multiset of points $\mathcal{M} = \chi_V - \chi_E + 4 \cdot \chi_P$ corresponds to a linear code with the stated properties. ■

If $C$ is an $[n, k, d]_2$ SO code with maximum weight $w$, then adding an all-one codeword of even length $n' \geq n$ (that is not already present) gives a $[n', k+1, \min\{d, n' - w\}]_2$ SO code. This easy construction is especially effective when applied to a, possibly non-projective, two-weight code. An optimal example is given by the $[124, 8, 60]_2$ SO code listed in Appendix A. See e.g. [18] for projective two-weight codes and [54] for non-projective two-weight codes. In Appendix A a non-projective $[183, 6, 92]_2$ two-weight code appears as a subcode of a $[186, 8, 92]_2$ SO code. Instead of adding an all-one codeword a 62-fold repetition of a $[3, 2, 2]_2$ simplex code is used.

**Lemma 12.** *(Cf. [67, Remark 3.2(2)]) Let $C$ be an $[n,k,d]_2$ SO code with a codeword of weight $w$, then we have $\tilde{n}_2(k,\tilde{d}) \leq n + \tilde{n}_2(k-1,d')$, where $\tilde{d} = \min\{d+d', w\}$ and $d'$ is arbitrary.*

*Proof.* Let the multiset of points $\mathcal{M}$ in $\mathrm{PG}(k-1,2)$ correspond to $C$, $\mathcal{M}'$ correspond to an $[n',k-1,d']_2$ SO code with $n' = \tilde{n}_2(k-1,d')$, and $H$ be a hyperplane with $\mathcal{M}(H) = n - w$. Then $\widetilde{\mathcal{M}} := \mathcal{M} + \mathcal{M}'$, where $\mathcal{M}'$ is embedded in $H$, is an $\widetilde{m}$-arc with $\widetilde{m} = \max\{n-d+n'-d', n-w+n'\} = \max\{n+n'-(d+d'), (n+n')-w\}$. ∎

**Corollary 5.** *There exist* $[47,6,22]_2$, $[55,6,26]_2$, $[80,7,38]_2$, $[82,7,40]_2$, $[88,7,42]_2$, $[90,7,44]_2$, $[147,8,72]_2$, $[155,8,76]_2$, $[171,8,84]_2$, $[176,8,86]_2$, $[178,8,88]_2$, $[201,8,98]_2$, $[203,8,100]_2$, $[208,8,102]_2$, $[210,8,104]_2$, $[216,8,106]_2$, $[223,8,110]_2$, $[232,8,114]_2$, $[239,8,118]_2$, $[247,8,122]_2$, *and* $[254,8,126]_2$ *SO codes.*

*Proof.* Apply Lemma 12 with $\mathcal{RM}_2(1,k-1)$ as $C$ and $w = 2^{k-1}$. ∎

Lemma 12 also yields $[\tilde{n}_2(7,d),7,d]_2$ SO codes for all even $48 \leq d \leq 64$ with $d \neq 52$. The question arises whether there exist other SO codes, with length near the Griesmer bound and a large maximum weight, besides first-order Reed–Muller codes. The generator matrices

$$\begin{pmatrix}
11111100000000011111111000000001111111111111110000000001100101101000000000\\
00000011100000000000000011111111111111111111111100011100000100111001000000100\\
00000000011111111111111111111111000000001111111100000000000011101000100000010\\
00011100000111000011110000111100001111000011111111110000101101010001000001\\
01100101101001100110011001100110011001100110011001101100111110100011100001000000\\
10011111110110101010001110101010101011101000101100011011100010011000001011\\
10101000100011010100101011001011010010101011010001100101110111001000001111
\end{pmatrix}$$

and

$$\begin{pmatrix}
11110000000011110000111111111110000000001111000011111111001001011100000000\\
00001111000000001111111111111000111000000000011111111111100010011101000000000\\
00000000111111111111000011110000001110001111111100001111000011101001000000\\
11111111111111111111111111111000000000011100000000000000000001111111000100100\\
00110011001100110011001100110110111110110011001100110011111110010000100011\\
01110100011111010001010011111011010111010101010101010101100100110000010111\\
10011001101101100110010101011101101000101001100100010111011010010100000001101
\end{pmatrix}$$

yield optimal self-orthogonal $[75,7,36]_2$ codes with weight distributions $36^{75}40^{51}60^1$ and $36^{76}40^{50}64^1$. The orders of their automorphism groups are given by 1440 and 96, respectively. However, using them as code $C$ in Lemma 12 does not yield any optimal codes.

**Lemma 13.** *No* $[10,5,4]_2$ *SO code exists.*

*Proof.* Let $(I_5|G')$ be a generator matrix of such a code $C$, where $I_5$ denotes the $5 \times 5$ unit matrix. Since $C$ is even and has minimum distance 4 the weights of the rows of $G'$ are contained in $\{3,5\}$. Orthogonality rules out weight 5 and implies that each pair of different rows has exactly two 1s in common. There cannot be three rows of $G'$ with the same two 1s in common, since choosing a fourth row would be impossible. However, for one row there are just 3 ways to choose two 1s, so that there are at most four rows. ∎

Note that there are numerous ways to show this result, e.g. using results for self-dual codes. We remark that $[10,5,4]_2$ codes with weight distributions $4^{10}5^{16}8^5$ and $4^{15}6^{15}10^1$ do

exist. However, the only doubly-even $[\leq 10, 4]_2$ codes have weight distributions $4^{14}8^1$ and $4^{10}8^5$, cf. [46].

**Lemma 14.** *We have $\tilde{n}_2(k, 2) = 2k$ for all $k \geq 1$.*

*Proof.* The lower bound is given by $n \geq 2k$ and a construction is given by doubling the multiset of points given by a unit matrix. ∎

**Proposition 5.** *For even $d \geq 2$ we have*

$$\tilde{n}_2(3, d) = \begin{cases} \tilde{g}_2(3, d) + 1 & : \quad d = 2, \\ \tilde{g}_2(3, d) & : \quad \text{otherwise}. \end{cases}$$

*Proof.* Assume $d \geq 4$, so that all lower bounds are given by the Griesmer SO bound. For doubly-even $d \geq 4$ and for singly-even $d \geq 6$ there exist Griesmer codes with minimum distance $d$ or $d + 2$, so that we can apply Lemma 9. ∎

Note that it suffices to provide constructions for $d \leq 2 + 2^2 = 6$. For even $d \geq 2$ we have $n_2(3, d) = g_2(3, d)$.

**Proposition 6.** *For even $d \geq 2$ we have*

$$\tilde{n}_2(4, d) = \begin{cases} \tilde{g}_2(4, d) + 2 & : \quad d = 2, \\ \tilde{g}_2(4, d) + 1 & : \quad d = 6, \\ \tilde{g}_2(4, d) & : \quad \text{otherwise}. \end{cases}$$

*Proof.* Assume $d \geq 4$. Besides the Griesmer SO bound, the lower bounds are given by Corollary 3 for $d = 6$. Doubling a $[7, 4, 3]_2$ code yields a construction for $d = 6$. For doubly-even $d \geq 4$ and for singly-even $d \geq 10$ there exist Griesmer codes, so that we can apply Lemma 9. ∎

Note that it suffices to provide constructions for $d \leq 6 + 2^3 = 14$. For even $d \geq 2$ we have $n_2(4, d) = g_2(4, d)$.

**Proposition 7.** *For even $d \geq 2$ we have*

$$\tilde{n}_2(5, d) = \begin{cases} \tilde{g}_2(5, d) + 3 & : \quad d = 2, \\ \tilde{g}_2(5, d) + 2 & : \quad d = 4, \\ \tilde{g}_2(5, d) + 1 & : \quad d \in \{6, 10, 14\}, \\ \tilde{g}_2(5, d) & : \quad \text{otherwise}. \end{cases}$$

*Proof.* Assume $d \geq 4$. Besides the Griesmer SO bound, the lower bounds are given by Lemma 13 for $d = 4$ and Corollary 3 for $d \in \{6, 10, 14\}$. Constructions for $d \in \{4, 6, 10, 14\}$ can be found in [13]. For doubly-even $d \geq 8$ and for singly-even $d \geq 18$ there exist Griesmer codes, so that we can apply Lemma 9. ∎

Note that it suffices to provide constructions for $d \leq 14 + 2^4 = 30$. For even $d \geq 2$ we have $n_2(5,d) = g_2(5,d) + 1$ for $d \in \{4,6\}$ and $n_2(5,d) = g_2(5,d)$ otherwise.

**Lemma 15.** *No* $[31,6,14]_2$ *SO code exists.*

*Proof.* Assume that $C$ is a $[31,6,14]_2$ SO code and $\mathcal{M}$ its corresponding multiset of points. From Lemma 8.(2) we conclude that $C$ is of full length. Since projection through a point for $\mathcal{M}$ corresponds to taking a subcode in $C$ and $\tilde{n}_2(5,14) = 30$, we have $\mathcal{M}(P) \in \{0,1\}$ for all points $P$. The doubly-even weight subcode of $C$ is a $[31,5,16]_2$ code, i.e. a 5-dimensional simplex code. Thus, the non-zero doubly-even codewords have weight 16. Using the residual SO code argument in Lemma 2 and Lemma 3 we conclude that the only possible singly-even weight is 14. Using the MacWilliams equations we obtain a contradiction. ∎

There exists an even $[30,6,14]_2$ code with weight distribution $14^{31}16^{31}30^1$.

**Proposition 8.** *For even* $d \geq 2$ *we have*

$$\tilde{n}_2(6,d) = \begin{cases} \tilde{g}_2(6,d) + 4 & : \quad d = 2, \\ \tilde{g}_2(6,d) + 2 & : \quad d \in \{4,14\}, \\ \tilde{g}_2(6,d) + 1 & : \quad d \in \{6,8,10,12,18,20,22,26,30\}, \\ \tilde{g}_2(6,d) & : \quad \textit{otherwise}. \end{cases}$$

*Proof.* Assume $d \geq 4$. Besides the Griesmer SO bound, the lower bounds are given by $n \geq 2k$ for $d = 4$, $\tilde{n}_2(6,d) \geq n_2(6,d)$ for $d \in \{8,12,20\}$, Lemma 15 for $d = 14$, and Corollary 3 for $d \in \{6,10,18,22,26,30\}$. Constructions for $d \leq 22$ can be found in [13] and [7, Appendix B].[7] For $d \in \{26,30\}$ Proposition 4 gives examples. For doubly-even $d \geq 24$ and for singly-even $d \geq 34$ there exist Griesmer codes, so that we can apply Lemma 9. ∎

Note that it suffices to provide constructions for $d \leq 30 + 2^5 = 62$. For even $d \geq 2$ we have $n_2(6,d) = g_2(6,d) + 1$ for $d \in \{4,\ldots,14\} \cup \{20\}$ and $n_2(6,d) = g_2(6,d)$ otherwise.

Similar to Lemma 15 we obtain the following non-existence results.

**Lemma 16.** *No* $[32,7,14]_2$, $[56,7,26]_2$, *and* $[63,7,30]_2$ *SO codes exist.*

*Proof.* Assume that $C$ is an $[n,k,d]_2$ SO code with parameters from the above list and $\mathcal{M}$ its corresponding multiset of points. From Lemma 8.(1) we conclude that $C$ is of full length. Projection through points and the known values for $\tilde{n}_2(6,d)$ yield $\gamma(\mathcal{M}) \leq 1$, i.e., $C$ is projective.

[7] For the unique self-orthogonal $[11,5,4]_2$ code [13] states its weight distribution $4^{10}6^{16}8^5$, its order 1920 of the automorphism group as well as its type, and refers to [66], where it is mentioned as $\overline{B}_{11}$ occurring as a truncation of the unique self-orthogonal $[12,6,4]_2$ code $B_{12}$. While both codes are projective, the parameters of the first one are omitted in the list of [7, Appendix B]. We remark that there are exactly three non-equivalent $[11,5,4]_2$ codes with weight distribution $4^{10}6^{16}8^5$ having different orders of the automorphism group and that the doubly-even subcode of the $[11,5,4]_2$ code is a 4-divisible $[10,4,4]_2$ code with weight distribution $4^{10}8^5$.

With this, Lemma 3 excludes weights $n-1$ and $n-2$. Denote the doubly-even weight subcode of $C$ by $C_4$. For our parameters it turns out that it is a doubly-even $[n, k-1, d+2]_2$ Griesmer code with $d+2 \leq 2^{k-2}$, i.e., it has to be of Solomon–Stiffler type $[1; \varepsilon_1, \ldots, \varepsilon_{k-2}]$, so that we can compute the possible weight distributions of $C_4$ from the $\varepsilon_i$. Since $C_4$ contains all doubly-even codewords from $C$, this restricts the possible weights of $C$ as does the residual SO code argument in Lemma 2.

- For $[32, 7, 14]_2$ the residual SO code argument in Lemma 2 excludes weights $\{18, 22, 26\}$ and $C_4$ has weight distribution $16^{62}32^1$, so that the non-zero weights of $C$ have to be contained in $\{14, 16, 32\}$. With this, the MacWilliams equations give a contradiction for the existence of $C$.[8]

- For $[56, 7, 26]_2$ the residual SO code argument excludes $\{34, 42, 46, 50\}$ and $C_4$ has weight distribution $28^{56}32^7$, so that the non-zero weights of $C$ have to be contained in $\{26, 28, 30, 32, 38\}$. With this, the MacWilliams equations give a contradiction for the existence of $C$.[9]

- For $[63, 7, 30]_2$ the residual SO code argument excludes $\{34, 38, 42, 46, 50, 54, 58\}$ and $C_4$ has weight distribution $32^{63}$, so that the non-zero weights of $C$ have to be contained in $\{30, 32\}$. With this, the MacWilliams equations give a contradiction for the existence of $C$.

■

We remark that the unique even $[30, 6, 14]_2$ code has weight distribution $14^{31}16^{31}30^1$. The non-existence of a $[62, 7, 30]_2$ SO code was also shown in [67, Proposition 6.4] and the non-existence of a $[63, 7, 30]_2$ SO code in [67, Remark 6.14] by an exhaustive computer search.

**Lemma 17.** *No* $[33, 7, 14]_2$ *SO code exists.*

*Proof.* Let $C$ be a $[33, 7, 14]_2$ SO code and $\mathcal{M}$ be its corresponding multiset of points. From Lemma 16 we conclude that $C$ has full length. The subcode $C_4$ of doubly-even weight codewords is either a $[32, 6, 16]_2$ or a $[33, 6, 16]_2$ code. Since the residual codes of $C$ have to be even, there is no codeword of weight 32, which excludes the $[32, 6, 16]_2$ Griesmer code with weight distribution $16^{62}32^1$ as candidate for $C_4$. Due to the residual code argument in Lemma 1 the $[33, 6, 16]_2$ code $C_4$ cannot contain a codeword with weight in $\{20, 24, 28\}$. However, $C_4$ cannot consist of non-zero codewords of weight 16 only. ■

We remark that there exist three non-equivalent full-length $[33, 6, 16]_2$ codes – none of them is doubly-even.

[8] Alternatively, $\tilde{n}_2(6, 14) = 32$ forces $\mathcal{M}(P) = 0$ under projection, which is impossible.

[9] More precisely, using $A_{28} = 56$, $A_{32} = 7$, and $B_2 = 0$ yields $A_{26} = \frac{136}{3} \notin \mathbb{N}$. Alternatively, one can deduce the existence of a codeword of weight 26 from the Griesmer bound and exploit one of the facts that the residual $[30, 6, 14]_2$ code contains a codeword of weight 30 and does not contain a codeword of weight 20. However, both statements would need a small separate proof.

**Proposition 9.** *For even $d \geq 2$ we have*

$$\tilde{n}_2(7,d) = \begin{cases} \tilde{g}_2(7,d)+5 & : \ d = 2, \\ \tilde{g}_2(7,d)+3 & : \ d \in \{4,14\}, \\ \tilde{g}_2(7,d)+2 & : \ d \in \{6,10,16,18,26,28,30\}, \\ \tilde{g}_2(7,d)+1 & : \ d \in \{8,12,20,22,24,34,\ldots,46,50,54,58,62\}, \\ \tilde{g}_2(7,d) & : \ \textit{otherwise}. \end{cases}$$

*Proof.* Assume $d \geq 4$. Besides the Griesmer SO bound, the lower bounds are given by $n \geq 2k$ for $d = 4$, $\tilde{n}_2(7,d) \geq n_2(7,d)$ for $d \in \{8,12,16,20,24,28,36,40,44\}$, Lemma 17 for $d = 14$, Lemma 16 for $d \in \{26,30\}$, and Corollary 3 for $d \in \{6,10,18,22,34,38,42,46,50,54,58,62\}$. Constructions for $d \leq 16$ can be found in [13] and [66]. For constructions with $18 \leq d \leq 36$ we refer to [67]. For $38 \leq d \leq 44$ Corollary 5 gives examples. For singly-even $46 \leq d \leq 62$ Proposition 4 gives examples. For doubly-even $d \geq 48$ and for singly-even $d \geq 66$ there exist Griesmer codes, so that we can apply Lemma 9. ∎

Note that it suffices to provide constructions for $d \leq 62 + 2^6 = 126$. For even $d \geq 2$ we have $n_2(7,d) = g_2(7,d) + 2$ for $d \in \{14,16,26,28\}$, $n_2(7,d) = g_2(7,d) + 1$ for $d \in \{4,\ldots,12\} \cup \{18,\ldots,24\} \cup \{30\} \cup \{36,\ldots,44\}$, and $n_2(7,d) = g_2(7,d)$ otherwise.

# 4 Optimal binary self-orthogonal codes with dimension 8

In this section we completely determine $\tilde{n}_2(8,\cdot)$, see Theorem 2. The existence of a linear code with a prescribed minimum distance and a prescribed subgroup of the automorphism group can be formulated as a system of Diophantine linear equations, see e.g. [15]. Here we use the fact that each doubly-even binary linear code is self-orthogonal and solve the corresponding systems with integer linear programming.

**Lemma 18.** *There exist* $[82,8,38]_2$, $[84,8,40]_2$, $[113,8,54]_2$, $[115,8,56]_2$, $[138,8,66]_2$, $[140,8,68]_2$, $[145,8,70]_2$, $[147,8,72]_2$, $[153,8,74]_2$, $[155,8,76]_2$, $[160,8,78]_2$, $[162,8,80]_2$, $[169,8,82]_2$, $[171,8,84]_2$, $[176,8,86]_2$, $[178,8,88]_2$, $[184,8,90]_2$, $[186,8,92]_2$, $[68,9,30]_2$, $[70,9,32]_2$, $[84,9,38]_2$, $[93,9,42]_2$, $[100,9,46]_2$, $[116,9,54]_2$, $[118,9,56]_2$, $[124,9,58]_2$, *and* $[131,9,62]_2$ *SO codes.*

*Proof.* In Appendix A we have listed generator matrices of non-projective doubly-even $[84,8,40]_2$, $[115,8,56]_2$, $[140,8,68]_2$, $[147,8,72]_2$, $[155,8,76]_2$, $[171,8,84]_2$, $[186,8,92]_2$, $[70,9,32]_2$, and $[118,9,56]_2$ codes. For the parameters $[162,8,80]_2$, $[178,8,88]_2$, $[86,9,40]_2$, $[95,9,44]_2$, $[102,9,48]_2$, $[126,9,60]_2$, and $[133,9,64]_2$ it turns out that the codes stored in the `BKLC` database [19] are doubly-even and non-projective. These $[n,k,d]_2$ codes are SO themselves and decreasing the point multiplicity of a suitable point by 2 gives an $[n-2,k,d-2]_2$ SO code. ∎

For singly-even self-orthogonal multisets of points we can fix a point $P$ and require $\#\mathcal{M} - \mathcal{M}(H) \equiv 0 \pmod 4$ for every hyperplane $H \supseteq P$ and $\#\mathcal{M} - \mathcal{M}(H) \equiv 2 \pmod 4$ otherwise, see Proposition 1.(B). Alternatively, we can start with suitable candidates for the subcode of the doubly-even codewords and use `LinCode` [9] to search for an extension to the full SO code. We present examples of $[82, 8, 38]_2$ and $[113, 8, 54]_2$ SO codes obtained this way in Appendix A. Next we provide the necessary non-existence results for $[n, 8, d]_2$ SO codes.

**Lemma 19.** *No* $[65, 8, 30]_2$ *SO code exists.*

*Proof.* Let $C$ be a $[65, 8, 30]_2$ SO code. From $n_2(8, 30) = 65$ we conclude that $C$ is of full length, so that Lemma 3 yields that there are no codewords of weight 64. Denote the doubly-even weight subcode of $C$ by $C_4$. Since a $[64, 7, 32]_2$ code is a Griesmer code with weight distribution $32^{126}64^1$, $C_4$ is a full-length 4-divisible $[65, 7, 32]_2$ code. Applying the residual code argument from Lemma 1 to $C_4$ yields that there are no codeword weights strictly between 32 and 64. However, since there is no codeword of weight 64 in $C \supset C_4$, the MacWilliams equations contradict the existence of $C_4$.[10] ∎

The non-existence of $[65, 8, 30]_2$ SO codes was previously shown via exhaustive enumeration using `Q-EXTENSION` [14], see [62, Remark 20].

**Lemma 20.** *No* $[96, 8, 46]_2$, $[120, 8, 58]_2$, *and* $[127, 8, 62]_2$ *SO codes exist.*

*Proof.* Assume that $C$ is an $[n, k, d]_2$ SO code with parameters from the above list and $\mathcal{M}$ its corresponding multiset of points. From Lemma 8.(1) we conclude that $C$ is of full length. Projection through points and the known values for $\tilde{n}_2(7, d)$ yield $\gamma(\mathcal{M}) \leq 1$, i.e., $C$ is projective. With this, Lemma 3 excludes weights $n - 1$ and $n - 2$. Denote the doubly-even weight subcode of $C$ by $C_4$. For our parameters it turns out that it is a doubly-even $[n, k - 1, d + 2]_2$ Griesmer code with $d + 2 \leq 2^{k-2}$, i.e., it has to be of Solomon–Stiffler type $[1; \varepsilon_1, \ldots, \varepsilon_{k-2}]$, so that we can compute the possible weight distributions of $C_4$ from the $\varepsilon_i$. Since $C_4$ contains all doubly-even codewords from $C$, this restricts the possible weights of $C$ as does the residual SO code argument in Lemma 2.

- For $[96, 8, 46]_2$ the residual SO code argument excludes all singly-even weights in $\{50, \ldots, 90\}$ and $C_4$ has weight distribution $48^{124}64^3$, so that the non-zero weights of $C$ have to be contained in $\{46, 48, 64\}$. With this, the MacWilliams equations give a contradiction for the existence of $C$.

- For $[120, 8, 58]_2$ the residual SO code argument in Lemma 2 and Lemma 3 excludes all singly-even weights in $\{62, 66\} \cup \{74, \ldots, 118\}$ and $C_4$ has weight distribution $60^{112}64^{15}$, so

[10] Alternatively, one can use the fact that one-weight codes have to be repeated simplex codes and that the parameters do not match.

that the non-zero weights of $C$ have to be contained in $\{58, 60, 64, 70\}$. Assume that $C'$ is the full-length projective even $[50, 7, 24]_2$ residual code of a codeword $c'$ of weight 70. From the residual code argument in Lemma 1 we conclude that the non-zero weights of $C'$ are contained in $\{24, 32, 48, 50\}$. Since weight 24 has to occur (due to the Griesmer bound) and weight 26 does not occur, there also cannot be a codeword of weight 50. With this, the MacWilliams equations for $C'$ yield a contradiction ($A_{48} < 0$). Thus, $C$ cannot contain a codeword of weight 70. With this, the MacWilliams equations give a contradiction for the existence of $C$.[11]

- For $[127, 8, 62]_2$ the residual SO code argument excludes all singly-even weights $66 \leq w \leq 122$ and $C_4$ has weight distribution $64^{127}$, so that the non-zero weights of $C$ have to be contained in $\{62, 64\}$. With this, the MacWilliams equations give a contradiction for the existence of $C$.

■

The non-existence of $[96, 8, 46]_2$ SO codes was previously shown via exhaustive enumeration using `Q-EXTENSION` [14], see [62, Remark 20]. There is a unique $[50, 7, 24]_2$ code [50]. However, its weight distribution is given by $24^{108}32^{19}$ and it is non-projective.

**Lemma 21.** *There exist no* $[90, 8, 42]_2$, $[97, 8, 46]_2$, *and* $[52, 9, 22]_2$ *SO codes.*

*Proof.* Using `LinCode` [9] we have classified all candidates for the subcode of the doubly-even codewords and checked that they cannot be extended to the full SO code. The parameters and the number of non-equivalent codes are stated in Table 1. ■

| $[n, k-1, d+2]_2$ | $[90, 7, 44]_2$ | $[96, 7, 48]_2$ | $[97, 7, 48]_2$ | $[120, 7, 60]_2$ | $[121, 7, 60]_2$ |
|---|---|---|---|---|---|
| # codes | 6 | 1 | | 1 | |
| # 4-div. codes | 3 | 1 | 1 | 1 | 6 |
| $[n, k-1, d+2]_2$ | $[51, 8, 24]_2$ | $[52, 8, 24]_2$ | | | |
| # codes | 1 | 8 | | | |
| # 4-div. codes | 1 | 2 | | | |

Table 1: Enumeration of (almost) optimal full-length codes.

The non-existence of $[52, 9, 22]_2$ SO codes was also shown in [70]. `LinCode` needs six hours for $[90, 8, 42]_2$ and five minutes for $[97, 8, 46]_2$. Checking possible extensions of the six doubly-even full-length $[121, 7, 60]_2$ codes using `LinCode` requires 88 hours of computation time. For the

[11] Alternatively, we may use the MacWilliams equations with $B_1 = B_2 = 0$, $A_{60} = 112$, $A_{64} = 15$, and $\{58, 70\}$ as possibilities for the singly-even weights to conclude $A_{70} \notin \mathbb{N}$.

unique $[120,7,60]_2$ code we have aborted the computation after 12 days and a memory usage of more than 30 GB. So we pursue a different approach based on the split weight enumerator.[12]

**Lemma 22.** *There exists no* $[121,8,58]_2$ *SO code.*

*Proof.* Let $C$ be a $[121,8,58]_2$ SO code and let $\mathcal{M}$ be its corresponding multiset of points. From Lemma 20 we conclude that $C$ has full length. Since $\tilde{n}_2(7,58) = 119$, we have $\gamma(\mathcal{M}) \leq 2$. If $c \in C$ is a codeword of weight 118, then the residual code is a full-length $[3,k',2]_2$ code, so that $k' = 2$. The subcode $C'$ of $C$ shortened onto $\mathrm{supp}(c)$ has dimension 6 and its non-zero codewords have weights contained in $\{58,60,118\}$. Note that $C'$ is self-orthogonal itself and the subcode spanned by the doubly-even codewords has dimension at least 5, which e.g. contradicts the classification of one-weight codes in [6]. Thus, there is no codeword of weight 118. From the residual code argument in Lemma 2 we conclude that the weights of the non-zero singly-even codewords are contained in

$$\{\underline{58},62,70,74,78,86,94,102\}.$$

Since no $[121,8,60]_2$ code exists weight 58 has to occur. By $C_4$ we denote the subcode consisting of the doubly-even codewords of $C$ and by $\mathcal{M}_4$ the corresponding multiset of points. Using `LinCode` we have exhaustively enumerated all doubly-even $[\leq 121,7,60]_2$ codes. Their weight distributions and some further data are given as follows:

| $n$ | w.e. | # Aut | $B_2$ |
|---|---|---|---|
| $120^{13}$ | $60^{112}64^{15}$ | 13872660480 | 0 |
| 121 | $60^{108}64^{15}76^4$ | 483840 | 15 |
| | $60^{104}64^{15}68^8$ | 258048 | 7 |
| | $60^{96}64^{31}$ | 884736 | 3 |
| | $60^{98}64^{27}68^2$ | 258048 | 4 |
| | $60^{105}64^{14}68^772^1$ | 225792 | 8 |
| | $60^{100}64^{23}68^4$ | 737280 | 5 |

Thus, the weights of the non-zero doubly-even codewords in $C$ are contained in

$$\{\underline{60},64,68,72,76\}.$$

Since no $[121,7,64]_2$ code exists weight 60 has to occur. Using `LinCode` we have exhaustively enumerated the following candidates for the residual codes $\mathrm{Res}(C;c)$:

[12] Alternatively, one may use any linear constraint solver that exploits the 13872660480 automorphisms of the $[120,7,60]_2$ code.

[13] A $[120,7,60]_2$ Griesmer code has to be of Solomon–Stiffler type $[1;0,0,1,0,0,0]$.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| wt($c$) | 58 | 62 | 70 | 74 | 78 | 86 | 94 |
| Res($C;c$) | $[63,7,30]_2$ | $[59,7,28]_2$ | $[51,7,24]_2$ | $[47,7,22]_2$ | $[43,7,20]_2$ | $[35,7,16]_2$ | $[27,7,12]_2$ |
| # (even) | 20 | 143 | 68 | 1 | 7 | 4 | 1 |

Having the possible weight distributions of these residual codes at hand we iteratively apply the linear programming method for the split weight enumerator w.r.t. the partition given by the support of a codeword $c$ and its complement. Each time when we say that we apply the split weight enumerator for wt($c$) we consider all pairs of weight distributions for $C_4$, Res($C;c$) and prescribe them. We then state the conclusions that we draw from the respective set of infeasible linear programs.

The split weight enumerator for wt($c$) = 58 excludes the weights $\{72, 102\}$ for $C$ and the cases $A_{60} \in \{100, 104, 105\}$ for $C$, $C_4$. Applying the split weight enumerator for weights 94, 86, 78, 62 excludes these weights. Applying the split weight enumerator for wt($c$) = 70 shows that there cannot be such a codeword. Applying the split weight enumerator for wt($c$) = 74 yields $A_{60} = 96$, which implies the unique weight distribution $58^{108}60^{96}64^{31}74^{20}$ with $B_2 = 5$. Since $C_4 \subseteq C$, we have $C^\perp \subseteq C_4^\perp$, so $5 = B_2(C) \leq B_2(C_4)$, which contradicts $B_2(C_4) = 3$ for weight distribution $60^{96}64^{31}$. Thus, there is no codeword of weight 74. Applying the split weight enumerator for wt($c$) = 58 again, yields a contradiction. ■

Note that we have not used the computer-aided classified residual and subcodes explicitly, but only used lists of possible weight distributions. It seems very likely that one can iteratively apply the linear programming method for the split weight enumeration with partitions of reasonable size to obtain a *computer-free* proof. This should also be the case for the two parameter sets in Lemma 21 with $k = 8$.[14]

We remark that there exist exactly six non-equivalent $[119, 7, 58]_2$ SO codes with maximum weight at most 76. Five of them violate $A_{66} = 0$ or $A_{62} \leq 20$. The generator matrix

$$\begin{pmatrix}
110000110011110001111000000001111000011111111000011111111000000000000000011111111000000001111111111111111000001011100000 \\
001100001111110000000111100000000111111111111000000000000011111111000000000000000011111111111111111111111100000111010000 \\
000011111100110000000000011111111111100001111000000000000000000000011111111111111111111111110000000011111111000011010010 00 \\
000000000000001000000000000000000000000000000011001111111111111111111111111111111111111111111111111111111111111111110000000100 \\
000000000000000101111111111111111111111111111111001100001111000011110000111100001111000011110000111100001111111100000000010 \\
111111111111110010011001100110011001100110011111100110011001100110011001100110011001100110011001100110011001100110000000001 \\
110101010101001010100010101010101110101010111010101010101110100010001110101000111010101010111010001010101010110100000000
\end{pmatrix}$$

yields a $[119, 7, 58]_2$ SO code with weight distribution $58^{52}60^{56}62^{8}64^{7}74^{4}$ and an automorphism group of order 32256. Interestingly enough, it contains 14 duplicated columns. In Table 2 we state the number of $[n, k, 58]_2$ SO codes with non-zero weights in $\{58, 60, 62, 64, 68, 70, 72, 74, 76, 78,$

[14]Proofs that do not use exhaustive computer enumerations of codes were found with AI tools e.g. for the cases $[90, 8, 42]_2$, $[97, 8, 46]_2$, and $[121, 8, 58]_2$. The proofs have been checked by independent AI reviews and exact computer checks, but not yet by a human expert, see the GitHub repository `https://github.com/yygotm/self-orthogonal-dim8` by Hirotaka Shimizu.

84, 86, 94} for moderate parameters. The `LinCode` computation took 22 hours for the stated values with $n \leq 118$.

| **k/n** | **105** | **106** | **107** | **108** | **109** | **110** | **111** | **112** | **113** | **114** | **115** | **116** | **117** | **118** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **3** | 5 | 10 | 18 | 29 | 49 | 73 | 110 | 157 | 220 | 294 | 393 | 495 | | |
| **4** | | | | | | 0 | 2 | 10 | 47 | 195 | 701 | 2322 | 6943 | |
| **5** | | | | | | | | | | 0 | 2 | 38 | 754 | 17397 |
| **6** | | | | | | | | | | | | 0 | 1 | 26 |
| **7** | | | | | | | | | | | | | | 0 |

Table 2: Number of non-equivalent full-length $[n, k, 58]_2$ SO codes with non-zero weights in $\{58, 60, 62, 64, 68, 70, 72, 74, 76, 78, 84, 86, 94\}$.

**Theorem 2.** *For even $d \geq 2$ we have*

$$\tilde{n}_2(8,d) = \begin{cases} \tilde{g}_2(8,d) + 6 & : \ d = 2, \\ \tilde{g}_2(8,d) + 4 & : \ d = 4, \\ \tilde{g}_2(8,d) + 3 & : \ d \in \{14, 26, 28, 30, 32, 42, 46, 58, 60\}, \\ \tilde{g}_2(8,d) + 2 & : \ d \in \{6, 10, 12, 16, \ldots, 22, 34, \ldots, 40, 44, 48, \ldots, 56, 62\}, \\ \tilde{g}_2(8,d) + 1 & : \ d \in \{8, 24, 66, \ldots, 94, 98, \ldots, 106, 110, 114, 118, 122, 126\}, \\ \tilde{g}_2(8,d) & : \ \textit{otherwise}. \end{cases}$$

*Proof.* Assume $d \geq 4$. Besides the Griesmer SO bound, the lower bounds are given by $n \geq 2k$ for $d = 4$, $\tilde{n}_2(8,d) \geq n_2(8,d)$ for doubly-even $8 \leq d \leq 60$, $68 \leq d \leq 92$, and $d \in \{100, 104\}$, Lemma 19 for $d = 30$, Lemma 21 for $d \in \{42, 46\}$, Corollary 4 for $d \in \{50, 54\}$, Lemma 22 for $d = 58$, Lemma 20 for $d = 62$, and Corollary 3 for singly-even $6 \leq d \leq 26$, $66 \leq d \leq 126$, and $d \in \{34, 38\}$. Constructions for $d \leq 16$ can be found in [13] and [66]. Constructions for $d \in \{38, 40, 54, 56\} \cup \{66, \ldots, 92\}$ in Lemma 18, see Appendix A for the corresponding generator matrices. For constructions with $d \in \{d \equiv 0 \pmod 2 : 18 \leq d \leq 60\} \backslash \{38, 42, 46, 54\}$ we refer to [62]. Generator matrices for $d \in \{42, 46, 94\}$ can be found in [1]. For $d \in \{72, 76, 84, 86, 88, 98, \ldots, 106, 110, 114, 118, 122, 126\}$ Corollary 5 gives examples. For doubly-even $d \geq 108$, $d \in \{64, 96\}$ and for singly-even $d \geq 130$ there exist Griesmer codes, so that we can apply Lemma 9. ∎

For even $d \geq 2$ we have $n_2(8,d) = g_2(8,d) + 3$ for $d \in \{26, \ldots, 32\} \cup \{60\}$, $n_2(8,d) = g_2(8,d) + 2$ for $d \in \{10, \ldots, 20\} \cup \{34, \ldots, 58\}$, $n_2(8,d) = g_2(8,d) + 1$ for $d \in \{4, 6, 8, 22, 24, 62\} \cup \{68, \ldots, 92\} \cup \{100, \ldots, 104\}$, and $n_2(8,d) = g_2(8,d)$ otherwise.

# 5 Optimal binary self-orthogonal codes with dimension 9 and minimum distance at most 64

In this section we partially determine $\tilde{n}_2(9,d)$ for $d \leq 64$, see Theorem 5 and Corollary 6. We remark that for even $d \leq 64$ the exact values $n_2(9,d)$ are known if $d \leq 28$ or $d \in \{32, 40, 50, 52, 64\}$. For the construction of self-orthogonal 9-dimensional binary codes the methods from Section 3 and Section 4 suffice for our purpose. This is only partially true for the presented non-existence results, so that we have to enhance our toolbox. As a by-product we improve the previously best known lower bound for $n_2(9,44)$ in Theorem 4 and show divisibility results for some putative optimal $[n,9,d]_2$ codes in Proposition 10.

For some set $\mathcal{F} \subset \mathbb{N}$ we say that $C$ is a linear code with forbidden weights $\mathcal{F}$ if $\mathrm{wt}(c) \notin \mathcal{F}$ for all $c \in C$. Let $y_r$ correspond to $y \in C$ in $\mathrm{Res}(C;x)$. From Equation (2) and Equation (4) we conclude

$$\mathrm{wt}(y_r) = \frac{\mathrm{wt}(y) + \mathrm{wt}(x+y) - \mathrm{wt}(x)}{2} = \frac{\mathrm{wt}(x) + \Big(\mathrm{wt}(y) - \mathrm{wt}(x)\Big) + \Big(\mathrm{wt}(x+y) - \mathrm{wt}(x)\Big)}{2}, \tag{14}$$

which directly implies:

**Lemma 23.** *Let $C$ be a $2^r$-divisible binary code and $y_r$ correspond to $y \in C$ in* $\mathrm{Res}(C;x)$. *If* $\mathrm{wt}(x) \equiv 2^r \pmod{2^{r+1}}$ *and* $\mathrm{wt}(y_r) \equiv 0 \pmod{2^r}$, *then either* $\mathrm{wt}(y)$ *or* $\mathrm{wt}(x+y)$ *is congruent to* $2^r$ *modulo* $2^{r+1}$.

A codeword that is not even is called odd. Applying Lemma 23 with $r = 0$, Equation (14) implies the following improvement of the residual code argument from Lemma 1.

**Lemma 24.** *Let $C$ be an $[n,k,d]_2$ code such that the weight of each odd codeword is at least $d'$. If $c \in C$ is a codeword with odd weight $w < 2d$, then* $\mathrm{Res}(C;c)$ *is an* $\left[n-w, k-1, \frac{d+d'-w}{2}\right]_2$ *code.*

Gaps in the weight spectrum can sometimes be exploited by a combination of Lemma 23 and Equation (14) to exclude further weights.

**Lemma 25.** *An even* $[93,9,44]_2$ *code does not contain a codeword of weight* 46.

*Proof.* Let $C$ be an even $[93,9,44]_2$ code. The residual code argument in Lemma 1 gives that the weights of the non-zero codewords are contained in

$$\{44, 46, 48, 56, 58, 60, 64, 72, 74, 76, 80, 84, 88, 90, 92\}.$$

Assume that $x$ is a codeword of weight 46. Using `LinCode` we obtain that $\mathrm{Res}(C;x)$ is a $[47,8,21]_2$ code with weight distribution $21^{66}22^{78}23^{30}24^{30}29^{30}30^{18}31^{2}32^{1}$.[15] By applying Lemma 23 with

[15] Up to equivalence there is a unique $[47,8,21]_2$ code.

$y_r \in \mathrm{Res}(C;x)$ of weight 32 we can assume that $\mathrm{wt}(y) \equiv 2 \pmod 4$. Using Equation (14) we compute

| $\mathrm{wt}(y)$ | 46 | 58 | 74 | 90 |
|---|---|---|---|---|
| $\mathrm{wt}(x+y)$ | 64 | 52 | 36 | 20 |

so that $\mathrm{wt}(x+y) = 64$. In $\mathrm{Res}(C;x+y)$ the codewords $x, y \in C$ correspond to a residual codeword of weight 14. However, using `LinCode` we verify that a $[29,8,12]_2$ residual code $\mathrm{Res}(C;x+y)$ is doubly-even.[16] ∎

Our next tool is a divisibility criterion for binary non-Griesmer codes, see Lemma 26.

**Theorem 3.** *([33, Theorem 2.6][17]) If $C$ is an $[n,k,d]_2$-code with $n = g_2(k,d) + t$, then $C$ admits a generator matrix $G$ where every row has a weight between $d$ and $d+t$.*

If $n = g_2(k,d) + t$ we speak of a *Griesmer plus $t$ code and call $t$ the* Griesmer defect*.*

**Lemma 26.** *Let $C$ be an $[n,k,d]_2$-code that admits a generator matrix $G$ where every row has a Hamming weight that is divisible by $2^r$ with $r \geq 1$. If each residual code w.r.t. a row of $G$ is $2^{r-1}$-divisible, then $C$ is $2^r$-divisible.*

*Proof.* For two codewords $x, y \in C$ and $y_r \in \mathrm{Res}(C;x)$ being the residual of $y$ we have

$$\mathrm{wt}(x+y) = \mathrm{wt}(x) + \mathrm{wt}(y) - 2 \cdot \left|\mathrm{supp}(x) \cap \mathrm{supp}(y)\right|$$

and

$$\mathrm{wt}(y_r) = \mathrm{wt}(y) - \left|\mathrm{supp}(x) \cap \mathrm{supp}(y)\right|,$$

see Equation (2) and Equation (4), respectively. For each row $x$ of $G$ the weight is divisible by $2^r$ and $\mathrm{wt}(y_r)$ is divisible by $2^{r-1}$. So, if the weight of $y$ is divisible by $2^r$, then also the weight of $x+y$ is divisible by $2^r$. Since each non-zero codeword can be written as a sum of pairwise different rows of $G$ we can prove the statement by induction on the number of summands. ∎

**Proposition 10.**

- *Each $[33,7,14]_2$ code with forbidden weight 15 is 2-divisible.*
- *Each $[49,8,22]_2$ code with forbidden weight 23 is 2-divisible.*
- *Each $[61,8,28]_2$ code with forbidden weights $\{29,30,31\}$ is 4-divisible.*
- *Each $[78,9,36]_2$ code is 4-divisible.*
- *Each even $[93,9,44]_2$ code is 4-divisible.*
- *Each $[117,9,56]_2$ code is 8-divisible.*

[16] There are 73 non-equivalent $[29,8,12]_2$ codes, which are all doubly-even and have a maximum weight of at most 28.

[17] In [31] the statement was generalized to arbitrary field sizes, cf. [79, Corollary 6].

- *Each* $[125, 9, 60]_2$ *code is* 4*-divisible.*

*Proof.* For each $[n, k, d]_2$ code $C$ of the above list we compute the Griesmer defect $t = n - g_2(k, d)$, use the residual code argument in Lemma 1 and other tailored arguments to exclude some small weights. If the remaining values in $\{d, d+1, \ldots, d+t\}$ are divisible by $2^r$ and the corresponding residual codes are $2^{r-1}$-divisible, then we can apply Theorem 3 and Lemma 26 to conclude that $C$ is $2^r$-divisible. For $r > 1$ the divisibility of the residual codes is obtained from preceding parameters in the list, if not stated otherwise.

- A $[33, 7, 14]_2$ code is a Griesmer plus 3 code. Since $C$ does not contain a codeword of weight 15 the non-existence of a $[16, 6, 7]_2$ code and Lemma 24 excludes weight 17. Thus, $C$ is spanned by codewords of weight 14 and 16.

- A $[49, 8, 22]_2$ code is a Griesmer plus 2 code. Since $C$ does not contain a codeword of weight 23 it is spanned by codewords of weight 22 and 24.

- A $[61, 8, 28]_2$ code is a Griesmer plus 3 code. The forbidden weights imply that $C$ is spanned by codewords of weight 28. The corresponding residual codes are $[33, 7, 14]_2$ codes and Equation (14) implies that there is no codeword of weight 15.

- A $[78, 9, 36]_2$ code is a Griesmer plus 2 code with forbidden weights $\{37, 38, 39\}$, so that it is spanned by codewords of weight 36. The corresponding residual codes are $[42, 8, 18]_2$ Griesmer plus 2 codes and Equation (14) excludes weight 19, i.e. they are spanned by even codewords.

- A $[93, 9, 44]_2$ code is a Griesmer plus 2 code. The assumption that it is even and Lemma 25 imply that it is spanned by codewords of weight 44. The corresponding residual codes are $[49, 8, 22]_2$ codes and Equation (14) implies that there is no codeword of weight 23.

- A $[117, 9, 56]_2$ code is a Griesmer plus 3 code with forbidden weights $\{57, \ldots, 63\}$, so that it is spanned by codewords of weight 56. The corresponding residual codes are $[61, 8, 28]_2$ codes and Equation (14) excludes the weights $\{29, 30, 31\}$.

- A $[125, 9, 60]_2$ code is a Griesmer plus 3 code with forbidden weights $\{61, 62, 63, 65, 66\}$, so that it is spanned by codewords of weight 60. The corresponding residual codes are $[65, 8, 30]_2$ Griesmer plus 3 codes and Equation (14) excludes the weights $\{31, 33\}$, i.e. they are spanned by even codewords.

■

The 4-divisibility of $[78, 9, 36]_2$ codes is also shown in [56, Lemma 3.6] and the 8-divisibility of $[117, 9, 56]_2$ codes in [56, Theorem 3.10]. In [56, Theorem 3.3] it was shown that there exists

a unique $[46, 9, 20]_2$ code. Interestingly enough, it has a trivial automorphism group and is doubly-even. We will use the 4-divisibility of even $[93, 9, 44]_2$ codes to show their non-existence in Theorem 4.

**Lemma 27.** *There is a unique* $[99, 8, 48]_2$ *code. It has weight distribution* $48^{210}56^{36}64^9$ *and an automorphism group of order* 144*. A generator matrix is given by*

$$\begin{pmatrix}
111000000001111111111101110111111111111000110000000011111100101111101101000000111111000000001110000\\
000111100001111111100001000111111111111100000011110000110000110111100101100000001111001111111110011\\
000000011110000000011111000011111110000010011110011000011111100111000010010000011110000111001111111\\
111111111111111111111110000000000000000001011110000110000000011111110100010000111111111110001111100\\
000111111110000111111110111000011100000001001111111111001111010001011000010000000110000000011000\\
011001100110011001100111011001100110011111101010101011101110100001011010000010000111111001110111111\\
001010101010101010101000111010101010110011001101001010101101101111100000010001100001111100011100\\
110010001101001011010110010000101111100110000011101110111001011011001100000000010101010110110010101
\end{pmatrix}.$$

*Proof.* Let $C$ be a $[99, 8, 48]_2$ code. Since no $[98, 8, 48]_2$ code exists it has full length. The Griesmer defect of $C$ is 2 and the residual code argument in Lemma 1 excludes all odd weights except $\{49, 97, 99\}$. The residual code of a codeword of weight 49 is a $[50, 7, 24]_2$ code. It is unique and has weight distribution $24^{108}32^{19}$, see e.g. [50] where a precursor of `LinCode` was used for the exhaustive enumeration.[18] With this we compute the unique weight distribution $48^{108}49^{128}64^{19}$. Using `LinCode` we have exhaustively enumerated all possibilities for the even-weight subcode. It is a unique $[96, 7, 48]_2$ code with weight distribution $48^{124}64^3$ – contradiction.[19] Thus, $C$ is spanned by codewords of weight 48 and 50, so that Lemma 26 implies that $C$ is even.

From the residual code argument in Lemma 1 we conclude that the non-zero weights are contained in $\{48, 50, 52, 56, 64, 72, 80, 96, 98\}$. Assume that $c_{50}$ is a codeword of weight 50. Its residual code is a $[49, 7, 23]_2$ code, which we exhaustively enumerate using `LinCode`. There are exactly two non-equivalent such codes and they have weight distributions $23^{52}24^{56}31^{12}32^7$ or $23^{48}24^{60}31^{16}32^3$, respectively. Thus, the non-zero weights of $C$ have to be contained in $\{48, 50, 56, 64\}$. Using `LinCode` we have enumerated those $[99, 8, 48]_2$ codes. The unique such code has weight distribution $48^{210}56^{36}64^9$, i.e. no codeword of weight 50 exists. Assume that $c_{98}$ is a codeword of weight 98. Since $\mathrm{Res}(C; c_{98})$ is a $[1, 1, 1]_2$ code, $C$ shortened onto $\mathrm{supp}(c_{98})$ is a $[98, 7, 48]_2$ code containing a codeword of weight 98, which is impossible (using the non-existence of a codeword of weight 50). Thus, $C$ is 4-divisible.

Assume that $c_{52}$ is a codeword of weight 52. Its residual code is a $[47, 7, 22]_2$ code, which we exhaustively enumerate using `LinCode`. There is just one such code and it has weight distribution $22^{78}24^{30}30^{18}32^1$. Thus, the non-zero weights of $C$ have to be contained in $\{48, 52, 56, 64\}$. Using `LinCode` we have enumerated those $[99, 8, 48]_2$ codes. Again, the unique such code has weight distribution $48^{210}56^{36}64^9$, i.e. no codeword of weight 52 exists. Thus, $C$ is 8-divisible.

Since the residual codes are 4-divisible, there cannot be a codeword of weight 96, whose residual code would have effective length 3. Using `LinCode` we have enumerated all 8-divisible

[18] Due to the existence of a unique $[51, 8, 24]_2$ projective two-weight code this classification result is not contained in [10].

[19] Alternatively, we can use the MacWilliams equations to conclude $B_{98} = -1 < 0$.

$[99, 8, 48]_2$ codes with maximum weight at most 80.[20] ∎

We remark that the $[99, 8, 48]_2$ code in Lemma 27 cannot be extended to a $[101, 9, 48]_2$ code. So, if a $[101, 9, 48]_2$ code exists, it has to be projective. For the multiset of points $\mathcal{M}$ corresponding to the unique $[99, 8, 48]_2$ code we have $\mathcal{M}(P) \leq 1$ for every point $P$ and $\mathcal{M}(E) \leq 6$ for every plane $E$, so that we cannot construct a $[92, 8, 44]_2$ code via $\mathcal{M} - \chi_E$.

**Lemma 28.** *There exists no* $[99, 9, 46]_2$ *SO code.*

*Proof.* Let $C$ be a $[99, 9, 46]_2$ SO code. In Lemma 27 we have classified the possible subcode of doubly-even codewords. Via a `LinCode` computation we verify that no extension to a $[99, 9, 46]_2$ code exists. ∎

Using similar methods as in Lemma 22 we obtain the following non-existence result.

**Lemma 29.** *There exists no* $[130, 9, 62]_2$ *SO code.*

*Proof.* By Corollary 4, $C$ has full length. The residual code argument in Lemma 2 excludes all weights in $\{66, 70, 74, 86, 90, 98, 102, 106, 114, 118, 122\}$. By $C_4$ we denote the subcode of the doubly-even weight codewords. A $[128, 8, 64]_2$ code is a Griesmer code with weight distribution $64^{254}128^1$. For a full-length doubly-even $[129, 8, 64]_2$ code the residual code argument in Lemma 1 excludes all non-zero weights except $\{64, 128\}$, which is impossible due to the MacWilliams identities.[21] Now assume that $C_4$ is a full-length doubly-even $[130, 8, 64]_2$ code. The residual code argument in Lemma 1 excludes all non-zero weights except $\{64, 68, 80, 128\}$. If $c \in C_4$ is a codeword of weight 128, then the residual code is a $[2, 1, 2]_2$ code and $C'$ arising from $C_4$ by restricting to $\mathrm{supp}(c)$ is a full length $[128, 8, 62]_2$ code with weight distribution $62^{64}64^{126}66^{64}128^{122}$, so that $C_4$ has weight distribution $64^{190}68^{64}128^1$. If there is no codeword of weight 128 in $C_4$, but a codeword of weight 80, then `LinCode` shows that the residual code is a unique $[50, 7, 24]_2$ code with weight distribution $24^{108}32^{19}$, cf. [50]. Counting yields the unique weight distribution $64^{235}80^{20}$ for $C_4$. If the maximum weight of $C_4$ is 68, then the residual code is a $[62, 7, 30]_2$ code. `LinCode` shows that there are exactly two such non-equivalent codes, both with a codeword of weight 62, which yields a contradiction in our setting. Thus, we have classified exactly three possible weight distributions for $C_4$.[23]

Since no $[130, 9, 64]_2$ code exists, there is a codeword $c \in C$ with weight 62. The residual code $\mathrm{Res}(C; c)$ is a $[68, 8, 32]_2$ code. Using `LinCode` we exhaustively enumerate all 153 such codes,

[20] We remark that there are unique residual codes for putative codewords of weight 72 and 80, which e.g. imply $A_{72} + A_{80} \leq 1$.

[21] More precisely, the MacWilliams identities give $A_{64} = 252$, $A_{128} = 3$, and $B_3 = -2603 < 0$.

[22] Such a code indeed exists and it has an automorphism group of order 1981808640.

[23] Clearly, $C_4$ cannot consist of non-zero codewords of weight 64 only, since e.g. the MacWilliams equations give a contradiction.

where all of them are even automatically. Having the possible weight distributions for the doubly-even codewords and the residual codes of a codeword of weight 62 at hand, we use the split weight enumerator with the condition $B_2(C) \leq B_2(C_4)$ to obtain a contradiction.[24] ∎

**Theorem 4.** *There is no* $[93, 9, 44]_2$ *code.*

*Proof.* Assume that $C$ is an even $[93, 9, 44]_2$ code[25] and $\mathcal{M}$ be its corresponding multiset of points. Since no $[91, 8, 44]_2$ code exists $C$ is projective. From Proposition 10 we conclude that $C$ is doubly-even. Thus, there cannot be a codeword of weight 92 whose residual would be 2-divisible having effective length 1. If $c_{88}$ is a codeword of weight 88, then $C$ shortened onto $\mathrm{supp}(c_{88})$ is a $[88, k', 44]_2$ code with non-zero weights in $\{44, 88\}$, so that $k' \leq 4$, see e.g. [6]. However, $\mathrm{Res}(C; c_{88})$ is an even $[5, k'', 2]_2$ code, where $k'' \leq 4$, which contradicts $k' + k'' = 9$. The residual code argument in Lemma 1 excludes the weights $\{52, 68\}$. Using `LinCode` we exhaustively enumerate all two non-equivalent $[33, 8, 14]_2$ and all 73 non-equivalent $[29, 8, 12]_2$ codes, which are the residual codes of codewords of weight 60 and 64, respectively. In both cases applying the split weight enumerator with respect to the support of such a codeword and its complement yields a contradiction for each possible weight distribution of the residual code. Thus, $C$ does not contain codewords of weight 60 or 64. Using these constraints, the integer programming method yields $B_3 > 0$,[26] i.e. there exists a line $L$ where all three points have multiplicity one in $\mathcal{M}$. With this, the projection $\mathcal{M}_L$ corresponds to a $[90, 7, 44]_2$ code $C'$. Using `LinCode` we exhaustively enumerate all six non-equivalent $[90, 7, 44]_2$ codes. Since each of them contains either a codeword of weight 60 or a codeword of weight 64, we obtain a contradiction. ∎

**Theorem 5.** *For even* $2 \leq d \leq 64$ *we have the following bounds for* $\tilde{n}_2(9, d)$*:*

| $d$ | 2 | 4 | 6 | 8 | 10 | 12 | 14 | 16 | 18 | 20 | 22 | 24 | 26 | 28 | 30 | 32 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\tilde{n}_2(9, d)$ | 18 | 18 | 20 | 21 | 29 | 30 | 36 | 38 | 45 | 46 | 53 | 54 | 61 | 62 | 68 | 70 |

| $d$ | 34 | 36 | 38 | 40 | 42 | 44 | 46 | 48 | 50 |
|---|---|---|---|---|---|---|---|---|---|
| $\tilde{n}_2(9, d)$ | 77 | 78–79 | 84 | 86 | 92–93 | 94–95 | 100 | 101–102 | 108–109 |

| $d$ | 52 | 54 | 56 | 58 | 60 | 62 | 64 |
|---|---|---|---|---|---|---|---|
| $\tilde{n}_2(9, d)$ | 110 | 115–116 | 117–118 | 124 | 125–126 | 131 | 133 |

*Proof.* Assume $d \geq 4$. Besides the Griesmer SO bound, the lower bounds are given by $n \geq 2k$ for $d = 4$, $\tilde{n}_2(9, d) \geq n_2(9, d)$[27] for doubly-even $8 \leq d \leq 64$, $\tilde{n}_2(9, d) \geq n_2(8, d + 2)$ for singly-even

[24] Alternatively, we may verify that for each possible weight distribution of $C_4$ the first three MacWilliams equations complemented by $B_2(C) \leq B_2(C_4)$ do not have a solution.

[25] If $C$ is a $[93, 9, 44]_2$ code that is not even, then puncturing and adding a parity bit gives an even $[93, 9, 44]_2$ code.

[26] The rational point $A_{44} = \frac{2224}{7}$, $A_{48} = 131$, $A_{56} = 59$, $A_{72} = \frac{23}{7}$ satisfies $B_1 = B_2 = B_3 = 0$, i.e. we need integer rounding cuts.

[27] Besides the data in [35] we use the non-existence of a $[85, 9, 40]_2$ code [56, Theorem 3.7] and the non-existence of a $[93, 9, 44]_2$ code shown in Theorem 4.

$d \in \{6, \ldots, 58\} \backslash \{46\}$, Lemma 28 for $d = 46$, Lemma 29 for $d = 62$, and Lemma 21 or [70] for $d = 22$. Constructions are given for doubly-even $8 \leq d \leq 64$ in [19], for $d \in \{2, 4, 6, 10, 14\}$ in [13], for $d \in \{22, 26, 50\}$ in [64], for $d \in \{18, 34\}$ in Appendix A, and for $d \in \{30, 38, 42, 46, 54, 56, 58, 62\}$ in Lemma 18. ∎

The values $\tilde{n}_2(9, d)$ were previously determined in [13] for all $d \leq 16$. A $[84, 9, 38]_2$ SO code was also found in [64]. To set the open cases of $\tilde{n}_2(9, d)$ into context, we remark that for even $d \leq 64$ the exact values $n_2(9, d)$ are known if $d \leq 28$ or $d \in \{32, 40, 50, 52, 64\}$. The $[92, 8, 44]_2$ code stored in the `BKLC` database cannot be extended to a $[92, 9, 42]_2$ SO code. Currently no other doubly-even $[92, 8, 44]_2$ code is known. The fact that each doubly-even code is self-orthogonal, the bounds in Theorem 5, and the divisibility results in Proposition 10 imply:

**Corollary 6.** *We have $\tilde{n}_2(9, d) = n_2(9, d)$ for $d \in \{36, 56, 60\}$.*

# Declaration on generative AI

In short, generative AI was used strictly for copy-editing, correcting grammatical errors, and identifying inconsistencies in the notation, without altering the mathematical content. Everything else, including programming and setting up the computations, was done by the author himself. The models used were the free versions of GPT-5.6 Luna (OpenAI) and Claude Sonnet 5.5 (Anthropic).

On October 8, 2026, Hirotaka Shimizu contacted me via email to inform me that he had found a computer-assisted proof suggesting that the maximum possible minimum distance of a self-orthogonal $[52, 9]_2$ code is 20. A written summary and a verification in Lean 4 were provided. However, an intermediate step relied on one of my computer enumerations [50, 51], where (a precursor of) `LinCode` [9] was used. The next day, I sent him a direct `LinCode` enumeration that verifies his result. Moreover, I mentioned that, although I had not previously worked on SO codes, I expected that the possible minimum distances of binary SO codes in dimension 8, as well as those in dimension 9 for *small* values of $d$, could be completely determined using standard techniques and some effort. Still slightly taken aback by OpenAI's math drop, I started asking myself what properties a useful research paper should have. The present paper arose in that context. During the process of writing the paper, I realized how cumbersome and predictable some of the steps involved are. Generative AI can be really useful for simplifying some of the arguments and extending the range of parameters that can be handled, cf. Footnote 14. I have decided not to incorporate any of these fascinating technical possibilities into this paper, but I intend to work with Hirotaka Shimizu to develop self-contained and simplified proofs of the stated results and extensions of these results that do not rely on extensive computer enumerations. To this end, we intend to make extensive use of generative AI.

# References


[1] J. An, N. Kaplan, J.-L. Kim, J. Luo, and G. Wang. Shortest self-orthogonal embeddings of binary linear codes. *arXiv preprint 2511.05440v2*, 2026.

[2] A. Ashikhmin and A. Barg. Minimal vectors in linear codes. *IEEE Transactions on Information Theory*, 44(5):2010–2017, 1998.

[3] E. F. Assmus and H. F. Mattson. New 5-designs. *Journal of Combinatorial Theory*, 6(2):122–151, 1969.

[4] B. I. Belov, V. N. Logachev, and V. P. Sandimirov. Construction of a class of linear binary codes achieving the Varshamov-Griesmer bound. *Problems of Information Transmission*, 10(3):211–217, 1974.

[5] A. Beutelspacher. Partitions of finite vector spaces: An application of the Frobenius number in geometry. *Archiv der Mathematik*, 31(1):202–208, 1978.

[6] A. Bonisoli. Every equidistant linear code is a sequence of dual Hamming codes. *Ars Combinatoria*, 18:181–186, 1984.

[7] I. Bouyukliev. On the binary projective codes with dimension 6. *Discrete Applied Mathematics*, 154(12):1693–1708, 2006.

[8] I. Bouyukliev. Classification of Griesmer codes and dual transform. *Discrete Mathematics*, 309(12):4049–4068, 2009.

[9] I. Bouyukliev, S. Bouyuklieva, and S. Kurz. Computer classification of linear codes. *IEEE Transactions on Information Theory*, 67(12):7807–7814, 2021.

[10] I. Bouyukliev and D. B. Jaffe. Optimal binary linear codes of dimension at most seven. *Discrete Mathematics*, 226(1-3):51–70, 2001.

[11] I. Bouyukliev, D. B. Jaffe, and V. Vavrek. The smallest length of eight-dimensional binary linear codes with prescribed minimum distance. *IEEE Transactions on Information Theory*, 46(4):1539–1544, 2000.

[12] I. Bouyukliev and P. R. J. Östergård. Classification of self-orthogonal codes over $\mathbb{F}_3$ and $\mathbb{F}_4$. *SIAM Journal on Discrete Mathematics*, 19(2):363–370, 2005.

[13] I. Bouyukliev, S. Bouyuklieva, T. A. Gulliver, and P. R. J. Östergård. Classification of optimal binary self-orthogonal codes. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 59:33–87, 2006.

[14] I. G. Bouyukliev. What is Q-extension? *Serdica Journal of Computing*, 1(1):115–130, 2007.

[15] M. Braun, A. Kohnert, and A. Wassermann. Optimal linear codes from matrix groups. *IEEE Transactions on Information Theory*, 51(12):4247–4251, 2005.

[16] T. Bu. Partitions of a vector space. *Discrete Mathematics*, 31(1):79–83, 1980.

[17] A. R. Calderbank and P. W. Shor. Good quantum error-correcting codes exist. *Physical Review A*, 54(2):1098–1105, 1996.

[18] R. Calderbank and W. M. Kantor. The geometry of two-weight codes. *Bulletin of the London Mathematical Society*, 18(2):97–122, 1986.

[19] J. Cannon, M. Grassl, and G. White. The Magma BKLC (Best Known Linear Codes) database. Integrated within the Magma Computational Algebra System, 2026. See also Markus Grassl's online database at `http://codetables.de`.

[20] H. Chen. Griesmer and optimal linear codes from the affine Solomon-Stiffler construction. *IEEE Transactions on Information Theory*, 71(9):6834–6843, 2025.

[21] N. Chigira, M. Harada, and M. Kitazume. Permutation groups and binary self-orthogonal codes. *Journal of Algebra*, 309(2):610–621, 2007.

[22] J. H. Conway and N. J. A. Sloane. A new upper bound on the minimal distance of self-dual codes. *IEEE Transactions on Information Theory*, 36(6):1319–1333, 1990.

[23] J. H. Conway and N. J. A. Sloane. *Sphere Packings, Lattices and Groups*, volume 290 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, New York, 3rd edition, 1999.

[24] H. Deng. The generalized Griesmer and antiGriesmer bounds. *arXiv preprint 2608.04495*, 2026.

[25] H. Deng, H. Huang, and Q. Xiang. Divisibility of Griesmer codes. *Journal of Combinatorial Theory, Series A*, 222:106181, 2026.

[26] H. Deng and W. Zhang. A proof of Ward's conjecture on Griesmer codes. *arXiv preprint 2610.29275*, 2026.

[27] S. Dodunekov. *Optimal Linear Codes*. Doctor of mathematical sciences thesis, Institute of Mathematics, Bulgarian Academy of Sciences, Sofia, Bulgaria, 1985.

[28] S. Dodunekov, S. Guritman, and J. Simonis. Some new results on the minimum length of binary linear codes of dimension nine. *IEEE Transactions on Information Theory*, 45(7):2543–2546, November 1999.

[29] S. Dodunekov and J. Simonis. Codes and projective multisets. *The Electronic Journal of Combinatorics*, 5(1):R37, 1998.

[30] S. M. Dodunekov. Minimal block length of a linear $q$-ary code with specified dimension and code distance. *Problems of Information Transmission*, 20(4):239–249, 1984. Translated from Problemy Peredachi Informatsii, 20(4): 9–20, 1984.

[31] S. M. Dodunekov. A comment on the weight structure of generator matrices of linear codes. *Problems of Information Transmission*, 26(2):173–176, 1990. Translated from Problemy Peredachi Informatsii, Vol. 26, No. 2, pp. 101–104.

[32] S. M. Dodunekov and N. L. Manev. Minimum possible block length of a linear binary code for some distances. *Problems of Information Transmission*, 20(1):8–14, 1984. Translated from Problemy Peredachi Informatsii, Vol. 20, No. 1, pp. 12–18, January–March, 1984.

[33] S. M. Dodunekov and N. L. Manev. An improvement of the Griesmer bound for some

small minimum distances. *Discrete Applied Mathematics*, 12(2):103–114, 1985.

[34] P. Govaerts. *Classifications of blocking set related structures in Galois geometries*. PhD thesis, Ghent University, 2003.

[35] M. Grassl. Bounds on the minimum distance of linear codes and quantum codes. Online available at `http://www.codetables.de`. Accessed on 2026-10-27.

[36] J. H. Griesmer. A bound for error-correcting codes. *IBM Journal of Research and Development*, 4(5):532–542, 1960.

[37] N. Hamada. A characterization of some $[n, k, d; q]$-codes meeting the Griesmer bound using a minihyper in a finite projective geometry. *Discrete Mathematics*, 116(1-3):229–268, 1993.

[38] N. Hamada, T. Helleseth, and Ø. Ytrehus. A new class of nonbinary codes meeting the Griesmer bound. *Discrete Applied Mathematics*, 47(3):219–226, 1993.

[39] T. Helleseth. A characterization of codes meeting the Griesmer bound. *Information and Control*, 50(2):128–159, 1981.

[40] T. Helleseth. New constructions of codes meeting the Griesmer bound. *IEEE Transactions on Information Theory*, 29(3):434–439, 1983.

[41] T. Helleseth. Projective codes meeting the Griesmer bound. *Discrete Mathematics*, 106:265–271, 1992.

[42] T. Helleseth and H. van Tilborg. A new class of codes meeting the Griesmer bound. *IEEE Transactions on Information Theory*, 27(5):548–555, 1981.

[43] W. C. Huffman and V. Pless. *Fundamentals of Error-Correcting Codes*. Cambridge University Press, Cambridge, UK, 2003.

[44] D. B. Jaffe. A brief tour of split linear programming. In *International Symposium on Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes*, pages 164–173. Springer, 1997.

[45] W. Jin, K. Li, and L. Qu. Several new classes of self-orthogonal minimal linear codes violating the Ashikhmin–Barg condition. *IEEE Transactions on Information Theory*, 71(10):7699–7714, 2025.

[46] M. Kiermaier and S. Kurz. Classification of $\delta$-divisible linear codes spanned by codewords of weight $\delta$. *IEEE Transactions on Information Theory*, 69(6):3544–3551, 2023.

[47] J.-L. Kim and W.-H. Choi. Self-orthogonality matrix and Reed–Muller codes. *IEEE Transactions on Information Theory*, 68(11):7159–7164, 2022.

[48] J.-L. Kim, Y.-H. Kim, and N. Lee. Embedding linear codes into self-orthogonal codes and their optimal minimum distances. *IEEE Transactions on Information Theory*, 67(6):3701–3707, 2021.

[49] S. Kurz. Additive codes attaining the Griesmer bound. *arXiv preprint 2412.14615*, 2025.

[50] S. Kurz. Classification of 8-divisible binary linear codes with minimum distance 24. *arXiv preprint 2012.06163*, 2020.

[51] S. Kurz. Codes of Nodal Sextics with Many Nodes. in *Varieties of Nodal Surfaces, Coding*

*Theory and Discriminants of Cubic Hypersurfaces*, F. Catanese, Ed., Lecture Notes of the Unione Matematica Italiana, vol. 31. Cham: Springer, 2026, chapter 5, 193–198.

[52] S. Kurz. Divisible codes. *arXiv preprint 2112.11763*, 2021.

[53] S. Kurz. Linear codes for $b$-symbol read channels attaining the Griesmer bound. *arXiv preprint 2507.07728*, 2025.

[54] S. Kurz. Non-projective two-weight codes. *Entropy*, 26(4), 289, 2024.

[55] S. Kurz. Optimal additive quaternary codes of dimension 3.5 and 4. *Serdica Journal of Computing*, 19(1):37–71, 2025.

[56] S. Kurz. The $[46, 9, 20]_2$ code is unique. *Advances in Mathematics of Communications*, 15(3):415–422, 2021.

[57] I. Landjev. The geometric approach to linear codes. In *Finite Geometries*, pages 247–256. Kluwer Academic Publishers, 2001.

[58] I. N. Landjev. Linear codes over finite fields and finite projective geometries. *Discrete Mathematics*, 213(1-3):211–244, 2000.

[59] K. Li, H. Chen, W. Jin, and L. Qu. Constructions of binary self-orthogonal singly-even wide minimal linear codes with few weights. *Designs, Codes and Cryptography*, 94(9):186, 2026.

[60] P. Li, Q. Cao, and F. Li. The construction of two classes of self-orthogonal cyclic codes. *Advances in Mathematics of Communications*, 25:44–59, 2026.

[61] R. Li, Z. Xu, and X. Zhao. On the classification of binary optimal self-orthogonal codes. *IEEE Transactions on Information Theory*, 54(8):3778–3782, 2008.

[62] S. Li, M. Shi, T. Helleseth, and S. Ling. Nonexistence of several infinite families of binary self-orthogonal codes. *IEEE Transactions on Information Theory*, 72(1):285–297, 2026.

[63] Z. Li and R. Li. On construction of ternary optimal self-orthogonal codes. *Computational and Applied Mathematics*, 43(3):134, 2024.

[64] H. Lim, J. An, and J.-L. Kim. Embedding Linear Codes into Optimal Self-Orthogonal Codes Based on Constraint-Preserving Genetic Algorithms. *arXiv preprint 2609.02135v2*, 2026.

[65] F. J. MacWilliams and N. J. A. Sloane. *The theory of error-correcting codes*, volume 16. Elsevier, 1977.

[66] V. Pless. A classification of self-orthogonal codes over GF(2). *Discrete Mathematics*, 3(1-3):209–246, 1972.

[67] M. Shi, S. Li, T. Helleseth, and J.-L. Kim. Binary self-orthogonal codes which meet the Griesmer bound or have optimal minimum distances. *Journal of Combinatorial Theory, Series A*, 214:106027, 2025.

[68] M. Shi, S. Li, and J.-L. Kim. Two conjectures on the largest minimum distances of binary self-orthogonal codes with dimension 5. *IEEE Transactions on Information Theory*, 69(7):4507–4512, 2023.

[69] M. Shi, N. Liu, and J.-L. Kim. Classification of binary self-orthogonal codes of lengths from

16 to 20 and its application. *Journal of Applied Mathematics and Computing*, 69(1):1173–1203, 2023.

[70] H. Shimizu. Proof candidate that $d_{SO}(52, 9) = 20$ (no self-orthogonal $[52, 9, 22]$ code). `https://github.com/yygotm/self-orthogonal-52-9`, 2026. Assembled 2026-10-08. AI-written proof with independent computer checks.

[71] J. Simonis. Restrictions on the weight distribution of binary linear codes imposed by the structure of Reed-Muller codes. *IEEE Transactions on Information Theory*, 40(1):194–196, January 1994.

[72] J. Simonis. MacWilliams identities and coordinate partitions. *Linear Algebra and its Applications*, 216:81–91, 1995.

[73] G. Solomon and J. J. Stiffler. Algebraically punctured cyclic codes. *Information and Control*, 8(2):170–179, 1965.

[74] L. Storme. Linear codes meeting the Griesmer bound, minihypers, and geometric applications. *Le Matematiche*, 59(1–2):367–392, 2004.

[75] V. D. Tonchev. Self-orthogonal designs and extremal doubly even codes. *Journal of Combinatorial Theory, Series A*, 52(2):197–205, 1989.

[76] M. A. Tsfasman and S. G. Vlăduţ. *Algebraic-Geometric Codes*, volume 58 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 1991.

[77] Z.-X. Wan. A characteristic property of self-orthogonal codes and its application to lattices. *Bulletin of the Belgian Mathematical Society-Simon Stevin*, 5(2-3):477–482, 1998.

[78] H. Ward. Divisible codes – a survey. *Serdica Mathematical Journal*, 27(4):263–278, 2001.

[79] H. N. Ward. Divisibility of codes meeting the Griesmer bound. *Journal of Combinatorial Theory, Series A*, 83(1):79–93, 1998.

[80] C. Xiang and C. Tang. Ternary self-orthogonal codes from square functions. *Designs, Codes and Cryptography*, 94(2):29, 2026.

[81] C. Xie, H. Chen, H. Zhou, Y. Li, and H. Lao. Optimal, almost optimal few-weight linear codes and related quantum codes. *IEEE Transactions on Information Theory*, 71(6):4250–4259, 2025.

[82] Z. Xie, H. Chen, X. Cao, G. Luo, and Y. Chen. Optimal, best known $q$-ary self-orthogonal cyclic codes and related constructions. *IEEE Transactions on Information Theory*, 2026. To appear.

[83] G. Zhang. A hyperplane bound and Ward's divisibility conjecture. *arXiv preprint 2610.26200*, 2026.

[84] Y. Zhang, D. Zheng, and X. Wang. A new class of self-orthogonal linear codes and their applications. *Discrete Mathematics*, 349(4):114935, 2026.

# A Generator matrices of self-orthogonal codes

In this section we list generator matrices for *SO* codes found by computer searches. The examples of 4-divisible non-projective SO codes are found by prescribing a subgroup of the automorphism group and solving an integer linear programming formulation. The singly-even SO codes were found by starting from some candidate for the doubly-even weight subcode and extension using `LinCode` [9]. Generator matrices with length at least 120 are stated in octal format after appending them with zero columns on the right to ensure that the length is a multiple of three. An $[n,k,d]_q$ code is called *minimal* if the support of every non-zero codeword is not strictly contained in the support of another codeword. The Ashikhmin–Barg condition $\frac{d}{w_{\max}} > \frac{q-1}{q}$ implies minimality [2], where $w_{\max}$ denotes the maximum weight. It is satisfied for several of the following examples and can be easily read off from the stated weight distribution. For examples of minimal self-orthogonal codes violating the Ashikhmin–Barg condition we refer e.g. to [45].

A $[82,8,38]_2$ SO code is given by

$$\begin{pmatrix}
11100000000011110000000001111111100000001111111111111111001001011000111000110000000\\
00011110000000000111110000000000001111111111111111111111100000000011111100010100000\\
00000001100000000000011111111111111111110000000111111111000100111111001001000100000\\
00000000011111111111111110000111100011110001111000011111000011101001010000000001000\\
01111111101100110111100110011001101100110110011001100001111111000011101111100000100\\
11100111111101110001110111011101000010010000010011001100000000001111101110000001 0\\
10101011110111011010101110111001011101001000110010101011110000001001100000100000001\\
10001110111010000110010010100011100111010111001001010010011111011011110001000000000
\end{pmatrix}.$$

It has weight distribution $38^{83}40^{108}46^{42}48^{17}54^{3}56^{2}$ and a trivial automorphism group.

A non-projective 4-divisible $[84,8,40]_2$ code is given by

$$\begin{pmatrix}
000000000000000000000000000000000000011111111111111111111111111111111111111111111\\
000000000000000000001111111111111111000000000000000000000000111111111111111111111\\
000000000000111111110000000011111111000000000000011111111111000000000000011111111111\\
000000111111000011110000001100001111000000111111000011111111000000111111000000001111\\
000111001111001100010011110101110011000011000111000100001111000001000011000011110111\\
001001010011010001101100111000011100001101111001011000110001001110011101011100010001\\
011110000101101110110011011010100101010100001010101100000010010011100110100101101010\\
110010000011000110100101110110100010101010011110111001000100110110001011100011100001
\end{pmatrix}.$$

It has weight distribution $40^{192}48^{57}56^{6}$ and an automorphism group of order 6.

A $[113,8,54]_2$ SO code is given by

$$\begin{pmatrix}
11000000110000111100001111111100111100000000000001111000000001111111100000000111111111111111100100110011111000000\\
00110000001100000011111111111110000001111000000000001111000000000000111111111111111111111111100010001111101000000\\
00001100000001111111111000011110000000000011110000000000000111111111111111111100000000111111110000111110011001000000\\
00000011111111001100110011001100000000000000011111111111111111000011110000111100001111000011100101010100100010000\\
00000000000000000000000000000001011111111111111111111111111111111111111111111111111111111111111111111111000000000000000001000\\
11111111111111111111111111111010011001100110011001100110011001100110011001100110011001100110011111000000000000000000100\\
01010101010101010101010101011101010101010101010101010101010101010101010101010101010101010101111101010010000001 0\\
01011100101100101111100001011000110101000101001100101101100101101100100011101010001110110100111010000110100000001
\end{pmatrix}.$$

It has weight distribution $54^{91}56^{120}62^{28}64^{7}70^{7}78^{2}$ and an automorphism group of order 21.

A non-projective 4-divisible $[115, 8, 56]_2$ code is given by

$$\begin{pmatrix}
10000100010000001111000011111111111011111111111111110100000011110001111001111111000011111111111111111111111110\\
01000010001111000000111111111111110000000000001111110010000000001000000000000000001111111111111111110000111111110\\
00100001000000011111111110000111111000000001111000011000100000000010000001111111111111100000000111111111100001111110\\
00010000111111110011001100110011111000011101110011000000100000000011111111000011100110000111100001100010011011110\\
00001111100011010101011101010101110011001100011100001000000000111110011101001100110100001100110011000111001111100110\\
11110010010101100101010110010111011100110110000011000001001100011110001011000011010100111100001100011011110011000\\
11111011110000111101111111100000111010101101101010000001000110000001100000000110000000000001111110000111100111111\\
01100000111001110011011000100110011011100110001101100000000101011110100100010101011001010101010101011011010101011
\end{pmatrix}.$$

It has weight distribution $56^{210}64^{36}72^{8}80^{1}$ and an automorphism group of order 42.

A non-projective 4-divisible $[124, 8, 60]_2$ code is given by

$$\begin{pmatrix}
77777777777777777777777777777777777777774\\
00000000000000000000077777777777777777774\\
00000000007777777777000000000077777777774\\
00000777770000077777000007777700000777774\\
00377003770017700177001770017700377001774\\
07417074170361703617036070360703417037607 4\\
31463114631466314623146311463314463140631 4\\
52525325252522522524525253252552512525725 24
\end{pmatrix}.$$

It has weight distribution $60^{127}64^{127}124^{1}$ and an automorphism group of order 332943851520. It can be obtained by adding the all-one codeword to the $[124, 7, 60]_2$ non-projective two-weight code obtained from Lemma 11 with $k = 7$.

A non-projective 4-divisible $[140, 8, 68]_2$ code is given by

$$\begin{pmatrix}
00000000000000000000037777777777777777777777776\\
00000000001777777777740000000000037777777777776\\
00000377776000007777740000037777740000037777776\\
00177400776003770017740017740017740017740017776\\
03607417036074170360741760740360740760340760076\\
14631463146314631463146163143463143163443161736\\
74007417740077607400740777055400777020744666706\\
25252525252525252525252265264525245265511322112
\end{pmatrix}.$$

It has weight distribution $68^{126}72^{126}76^{2}128^{1}$ and an automorphism group of order 1152.

A non-projective 4-divisible $[147, 8, 72]_2$ code is given by

$$\begin{pmatrix}
00000000000000000000037777777777777777777777777\\
00000000001777777777740000000000017777777777777\\
00000377776000007777740000007777776000000037777777\\
00177740037600177000774000777003777600003774000777\\
03601740741603637017034037077034017607740740370 17\\
37616141006235647021154747001144763610341043473 77\\
54222246342606731160664153117243164671547004730 63\\
44464613056254352263341255223611235222653311551 25
\end{pmatrix}.$$

It has weight distribution $72^{210}80^{36}88^{6}96^{3}$ and an automorphism group of order 18.

A non-projective 4-divisible $[155, 8, 76]_2$ code is given by

$$\begin{pmatrix}
00000000000000000000000077777777777777777777777776\\
00000000003777777777777770000000000017777777777777776\\
00000177774000000177777770000037777600000001777777776\\
00177603774003777600177770077403776000777760007777 6\\
03607614034034017607603771703741403601700176077007 76\\
14631634754354761671614372734356435636303631717036\\
64217617140065162172021570315427056746514656123610 46\\
35572245265576266614646637566572164253725366664526312
\end{pmatrix}.$$

It has weight distribution $76^{204}80^{31}92^{19}108^{1}$ and an automorphism group of order 6.

A non-projective 4-divisible $[171, 8, 84]_2$ code is given by

$$\begin{pmatrix}
000000000000000000000000077777777777777777777777777777777\\
000000000000077777777777700000000000000007777777777777777\\
000000077777700000077777700000000777777777000000077777777\\
000077700077700077700077700007777000077770000777700007777\\
017703701701700703701701701770077007701770017003701770077\\
060714306316307134316306316070707070716070363074316070707\\
363101713766411754466011766311133113366311477314466311133\\
525224520533123325522522532533311231122722525525522532215
\end{pmatrix}.$$

It has weight distribution $84^{186}88^{30}92^{37}96^{1}124^{1}$ and an automorphism group of order 240.

A non-projective 4-divisible $[186, 8, 92]_2$ code is given by

$$\begin{pmatrix}
000000000000000000000000000007777777777777777777777777777777\\
000000000000000377777777777777000000000000000007777777777777\\
000000017777777400000007777777000000017777777700000000377777\\
000077760001777400077770001777000177760000777700007777400037\\
007700760176037403700770076037007601760077007701770037403740\\
171717163636347414103030306043171636363717171706070141414103\\
630303660747541474614171416354630207460763630316300607460747\\
252525225252652525225252522545252452525125252532531252525325
\end{pmatrix}.$$

It has weight distribution $92^{189}96^{63}124^{3}$ and an automorphism group of order 16515072. The three codewords of weight 124 form a 62-fold repetition of the $[3, 2, 2]_2$ simplex code, i.e. they have full support. Removing these codewords, i.e. taking a suitable largest subcode not containing them, gives a non-projective two-weight code, see e.g. [54], with weight distribution $92^{48}96^{15}$ and an automorphism group of order 4608. A generator matrix is given by

$$\begin{pmatrix}
000000000000000000000000000007777777777777777777777777777777\\
000077760001777400077770001777000177760000777700007777400037\\
007700760176037403700770076037007601760077007701770037403740\\
171717163636347414103030306043171636363717171706070141414103\\
630303660747541474614171416354630207460763630316300607460747\\
252525225252652525225252522545252452525125252532531252525325
\end{pmatrix}$$

Let the multiset $\mathcal{M}$ correspond to the shortened full-length $[183, 6, 92]_2$ code. Writing $V$ for the ambient space PG(5, 2) and using two suitable disjoint sets of points $\mathcal{M}_1$, $\mathcal{M}_2$, we can write $\mathcal{M} = 3 \cdot \chi_V - \mathcal{M}_1 + \mathcal{M}_2$, so that we have $\mathcal{M}(P) = 2$ iff $\mathcal{M}_1(P) = 1$ and $\mathcal{M}(P) = 4$ iff $\mathcal{M}_2(P) = 1$. The set of points $\mathcal{M}_1 + \mathcal{M}_2$ corresponds to a projective $[12, 5, 4]_2$ SO code with weight distribution $4^{6}6^{16}8^{9}$ and an automorphism group of order 1152. A generator matrix is given by

$$\begin{pmatrix}
000000111111\\
011011000110\\
100111001001\\
111001011011\\
001100110011
\end{pmatrix}.$$

We remark $\tilde{n}_2(5, 4) = 11$ and that there are five non-equivalent $[12, 5, 4]_2$ SO codes [13].

A $[45, 9, 18]_2$ SO code is given by

$$\begin{pmatrix}
111111111000000111111000100011111110010000000\\
111111000111000111000111010011111001101000000\\
000000000000111111111111001000011111100100000\\
111111111111111111000000000100000111100010000\\
000111001001011000011011111100101010100001000\\
001011010010101011001001001111101011000000100\\
000101100110010101010000111001110111100000010\\
110110010011110001100101100111010010000000001\\
010011111011010111110110111110011011100000000
\end{pmatrix}.$$

It has weight distribution $18^{60}20^{135}22^{115}24^{79}26^{71}28^{37}30^{9}32^{4}34^{1}$ and a trivial automorphism group.

A non-projective 4-divisible $[70, 9, 32]_2$ code is given by

$$\begin{pmatrix}
0100000000110010001011100110100110001100000001001111101011010111011101\\
0010010000000100110011100111011110100000010101101100010101100011110011\\
1001010001110000000010010001111011111110000100001011010011000101110011\\
0000110001000110111110110000011010000101010010011101010010001101011101\\
0000000100010110010111000100010001010111101101001100101101011001100011\\
0000001011100110011000110111101110100111011100101100110011111110011\\
0000000000001110101011010011101110111011110000010111011011001000000000\\
1000001000010111101111100010011000000000101011010010010101010111101110\\
1000001000010110111000001110001011011011100100010110110000100011101110
\end{pmatrix}.$$

It has weight distribution $32^{318}40^{192}64^{1}$ and an automorphism group of order 774144.

A $[77, 9, 34]_2$ SO code is given by

$$\begin{pmatrix}
00000000000000000000000000000011111111111111111111111111111111111111110001\\
00000000000000001111111111111100000000000000000000001111111111111111111110000\\
00001111111111110000011111111100000011111111111111000000000011111111111110010\\
00010000001111110011100000111100011100000001111111100000011111000001111110000\\
00110000110001110100100011001100101100011110000011100001100011011110011110100\\
01110111110110010000000101000101100000100110000100101110101101101111100110000\\
11000011111110001011001001001001010101001010011111100011011001100100111110000\\
11101101010010110101000010010011101001110000101001000101110111000110101010000\\
11111011000001011000010110011100001101100100110101010110011011000000110000000
\end{pmatrix}.$$

It has weight distribution $34^{97}36^{147}38^{67}40^{61}42^{61}44^{37}46^{25}48^{6}50^{6}52^{4}$ and a trivial automorphism group. The doubly-even $[77, 8, 36]_2$ subcode has an automorphism group of order 6. We remark that there exist more than 62000 optimal $[77, 8, 36]_2$ codes, but we do not know an example with a larger automorphism group.

A non-projective 4-divisible $[118, 9, 56]_2$ code is given by

$$\begin{pmatrix}
10000000000000000000011111100000001100000001111111111111000000001111111111110000000011000000011111111\\
01100110000000000000001011101101001101100110111000010110111010011111001010101100110100000011100110101011110101011010000\\
00011110000000000000000111111011010010111100011100101010010110100010111010100111100001100010001101100110101100110011110\\
00000001010101000000000010110110011111011001100001001101001011010111011110111111010010010110010000101010111101101011010000\\
00000000110011000000001100100011111000011111011010111010101010101011000000110101101001011101011010110011011010101100110 0\\
00000000001111000000011000011001100011001100010010110011011001100001010111110101010011001011011000111111011001111110 01\\
00000000000000101010111101010110101011010010111110101000011001100011001101001100111111011000100101101010101010100110 10\\
00000000000000011001100000010101011010110100001111010110000111110110010101000011111011110111010101010001100111011101\\
00000000000000000111101110001100110010101011101101011010110011011100101111011001101001011110010011001110000111101000000
\end{pmatrix}$$

It has weight distribution $56^{348}64^{127}72^{36}$ and an automorphism group of order 384.