Abstracts

Optimal additive and linear b-symbol codes for large distances

SASCHA KURZ

For a finite *alphabet* $\mathcal{A}$ a *code* C of *length* n and *minimum distance* d is a subset of $\mathcal{A}^n$ such that any two elements differ in at least d positions. E.g. $C = \{cabbdb, bcabbd, abcdbb, cdadcc, acdcdc, dacccd, dcbdaa, bdcada, cbdaad\}$ is a code with length 6 and minimum distance $d = 5$ over the alphabet $\mathcal{A} = \{a, b, c, d\}$. Given parameters n, d, and $\#\mathcal{A}$, the aim is to maximize the code size $\#C$. In our example size 9 is indeed maximal [3]. Alternatively, one can minimize n given d, $\#\mathcal{A}$, and $\#C$. For alphabet $\mathcal{A} = \mathbb{F}_q$ we say that C is *linear* if it is linearly closed. The parameters of a linear code are related by the so-called *Griesmer bound* [6, 12]

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

where $k = \log_q \#C$. Interestingly enough, this bound can always be attained with equality if the minimum distance d is sufficiently large [12]. If C is only additively closed we call it *additive*. Each additive code is linear over some subfield so that we set $\mathcal{A} = \mathbb{F}_{q^h}$ and assume that C is linear over $\mathbb{F}_q$, so that $\#C = q^k$.

Let $G \in \mathbb{F}_q^{k\times n}$ be a *generator matrix* of a linear code C, i.e. a matrix whose rows form a basis of C. The columns of G span 1-dimensional subspaces which form a multiset of n points in the projective geometry $\mathrm{PG}(k-1, q)$ such that each hyperplanes contains at most $n - d$ points [5]. Similarly, an additive code over $\mathcal{A} = \mathbb{F}_{q^h}$ is given as the $\mathbb{F}_q$-row span of a full-rank matrix $G \in \mathbb{F}_{q^h}^{k\times n}$. Choosing an $\mathbb{F}_q$-basis $\mathcal{B}$ of $\mathbb{F}_{q^h}$ we can rewrite G as $\widetilde{G} \in \mathbb{F}_q^{kh\times n}$. Writing $\mathbb{F}_4 \simeq \mathbb{F}_2[\omega]/\left(\omega^2 + \omega + 1\right)$, we can start with the generator matrix of a linear code, interprete it as the generator matrix of an additive code and use the basis $\mathcal{B}$ to obtain the example

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & \omega & \omega^2 \end{pmatrix} \to \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & \omega & \omega & \omega & \omega \\ 1 & 0 & 1 & \omega & \omega^2 \\ \omega & 0 & \omega & \omega^2 & 1 \end{pmatrix} \to \begin{pmatrix} 00 & 10 & 10 & 10 & 10 \\ 00 & 01 & 01 & 01 & 01 \\ 10 & 00 & 10 & 01 & 11 \\ 01 & 00 & 01 & 11 & 10 \end{pmatrix}.$$

The blocks of h subsequent columns span subspaces of dimension at most h, which are elements in $\mathrm{PG}(k-1, q)$ of geometric dimension at most $h-1$. Indeed, additive codes over $\mathcal{A} = \mathbb{F}_{q^h}$ with length n, minimum distance d, and size q^k are in one-to-one correspondence to multisets of n subspaces of geometric dimension at most $h-1$ in $\mathrm{PG}(k-1, q)$ such that each hyperplane contains at most $n-d$ of those subspaces [1]. Replacing those subspaces by their contained points we obtain a multiset of points that corresponds to a linear code over $\mathbb{F}_q$ with length $n \cdot (q^h - 1)/(q - 1)$ and minimum distance $q^{h-1} \cdot d$ (assuming some non-degeneracy) [9]. Applying the

Griesmer bound (1) to the obtained linear code yields, see [1, Theorem 12],

$$n \ge \left\lceil \frac{g_q\big(r, d\cdot q^{h-1}\big)\cdot (q-1)}{q^h-1} \right\rceil = \left\lceil \frac{(q-1)\cdot \sum\limits_{i=0}^{r-1} \left\lceil d\cdot q^{h-1-i}\right\rceil}{q^h-1} \right\rceil . \tag{2}$$

One main conclusion of our work is that this bound can always be attained with equality if the minimum distance d is sufficiently large. Reverting the above chain of changes between coding theory and geometry boils down to the problem of partitioning a multiset of points into subspaces of geometric dimension $h-1$, which is rather hard in general, so that we consider a special subclass.

Each minimum distance $d \in \mathbb{N}$ can be uniquely written as $d = \sigma q^{k-1} - \sum_{i=1}^{k-1} \varepsilon_i \cdot q^{i-1}$, where $\sigma \in \mathbb{N}$ and $\varepsilon_i \in \{0,1,\dots,q-1\}$ for all $1 \le i \le k-1$. With this, the Griesmer bound from (1) is attained with equality, i.e. $n = g_q(k,d)$, iff

$$n = \sigma \cdot \frac{q^k-1}{q-1} - \sum_{i=1}^{k-1} \varepsilon_i \cdot \frac{q^i-1}{q-1}. \tag{3}$$

In order to describe a special variant of the Solomon–Stiffler construction, see [12], we assume a chain $S_1 \le S_2 \le \dots \le S_k$ of subspaces S_i with algebraic dimension i in $\mathrm{PG}(k-1,q)$. For each subspace T we denote its *characteristic function* by χ_T, i.e. $\chi_T(P) = 1$ if point P is contained in T and $\chi_T(P) = 0$ otherwise. With this, $\mathcal{M} = \sigma\chi_{S_k} - \sum_{i=1}^{k-1} \varepsilon_i \chi_{S_i}$ is a multiset of points whose corresponding linear code attains the Griesmer bound (1) if $\sigma \ge \sum_{i=1}^{k-1} \varepsilon_i$ and $0 \le \varepsilon_i \le q-1$, i.e. if d is sufficiently large. More generally, we say that a multiset of points $\mathcal{M}$ in $\mathrm{PG}(k-1,q)$ has type $\sigma[k] - \sum_{i=1}^{k-1} \varepsilon_i[i]$, where $\sigma \in \mathbb{N}$ and $\varepsilon_i \in \mathbb{Z}$ for all $1 \le i \le k-1$, if $\mathcal{M} = \sigma\chi_{S_k} - \sum_{i=1}^{k-1} \varepsilon_i \chi_{S_i}$. Note that σ needs to be sufficiently large. We say that a multiset of points is *h-partitionable* if it can be written as the sum of characteristic functions of subspaces of algebraic dimension h.

Theorem ([9]) Let q be a prime power, $k > h \ge 1$, $g := \gcd(k,h)$, and $\varepsilon_1, \dots, \varepsilon_{k-1} \in \mathbb{Z}$ such that q^{h-i} divides ε_i for all $1 \le i < h$ and $\sum_{i=1}^{k-1} \varepsilon_i \cdot \frac{q^i-1}{q-1} \equiv 0 \pmod{\frac{q^g-1}{q-1}}$. Then, there exists a $\sigma \in \mathbb{N}$ such that there exists an h-partionable multiset of points in $\mathrm{PG}(k-1,q)$ with type $\left(\sigma + t\cdot \frac{q^h-1}{q^g-1}\right)[k] - \sum\limits_{i=1}^{k-1} \varepsilon_i[i]$ for all $t \in \mathbb{N}$.

We remark that the stated conditions are also necessary for the assumed chain $S_1 \le \dots \le S_k$ and that the corresponding proof is constructive. As an example we mention the existence of a 2-partionable multiset of points in $\mathrm{PG}(7,2)$ with type $t[8] - [7] - [5] - [3]$ for each $t \ge 3$. Those $85t - 55$ lines in $\mathrm{PG}(7,2)$ have the property that each hyperplane contains at most $21t - 13$ lines, i.e. the corresponding additive code C_t over $\mathbb{F}_4$ has length $n_t = 85t - 55$, minimum distance $d_t = 64t - 42$, and cardinality $\#C_t = 2^8$. We remark that $\mathbb{F}_4$-linear codes with the same minimum distance and cardinality require lengths at least $85t - 53$, i.e. additive codes outperform linear codes for the stated parameters of d, $\#\mathcal{A}$, and $\#C$.

Corollary For given q, k, h, and sufficiently large d the Griesmer bound for additive codes (2) can always be attained. Moreover, additive codes outperform linear codes if h does not divide k or if the the difference of (2) and (1) is positive.

Open problem Find constructions for additive codes outperforming linear codes for relatively small minimum distances, see e.g. [2, 7, 8, 9], or decrease the necessary σ in the Solomon–Stiffler type construction. Find improved upper bounds.

In storage applications the reading device is sometimes insufficient to isolate adjacent symbols, which makes it necessary to adjust the standard coding-theoretic error model. Cassuto and Blaum studied a model where pairs of adjacent symbols are read in every step and introduced the so-called symbol-pair metric for codes [4]. This notion was generalized to the b-symbol metric where b-tuples of adjacent symbols are read at every step [13]. For linear codes each b subsequent columns of a generator matrix span a subspace with dimension at most b, so that the Griesmer bound for additive codes (2) applies [11]. Again, the Griesmer bound can always be attained if the minimum distance is sufficiently large [10].

Open problem Find Griesmer type bound bounds and attaining Solomon–Stiffler type constructions for other metric spaces.

References

[1] S. Ball, M. Lavrauw, and T. Popatia. Griesmer type bounds for additive codes over finite fields, integral and fractional MDS codes. *Des. Codes Cryptogr.*, 93(1):175–196, 2025.

[2] S. Ball and T. Popatia. Additive codes from linear codes. *arXiv preprint 2506.03805*, 9 pages, 2025.

[3] G. T. Bogdanova, A. E. Brouwer, S. N. Kapralov, and P. R. J. Östergård. Error-correcting codes over an alphabet of four elements. *Des. Codes Cryptogr.*, 23:333–342, 2001.

[4] Y. Cassuto and M. Blaum. Codes for symbol-pair read channels. *IEEE Trans. Inf. Theory*, 57(12), 8011–8020, 2011.

[5] S. M. Dodunekov and J. Simonis. Codes and projective multisets. *Electron. J. Combin.*, 5:1–23, 1998.

[6] J. H. Griesmer. A bound for error-correcting codes. *IBM J. Res. Dev.*, 4(5):532–542, 1960.

[7] C. Guan, R. Li, Y. Liu, and Z. Ma. Some quaternary additive codes outperform linear counterparts. *IEEE Trans. Inf. Theory*, 69(11):7122–7131, 2023.

[8] C. Guan, J. Lv, G. Luo, and Z. Ma. Combinatorial constructions of optimal quaternary additive codes. *IEEE Trans. Inf. Theory*, 70(11):7690–7700, 2024.

[9] S. Kurz. Additive codes attaining the Griesmer bound. *arXiv preprint 2412.14615*, 100 pages, 2024.

[10] S. Kurz. Linear codes for b-symbol read channels attaining the Griesmer bound. *arXiv preprint 2507.07728*, 27 pages, 2025.

[11] G. Luo, M. F. Ezerman, C. Güneri, S. Ling, and F. Özbudak. Griesmer bound and constructions of linear codes in b-symbol metric. *IEEE Trans. Inf. Theory*, 70(11):7840–7847, 2024.

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

[13] E. Yaakobi, J. Bruck, and P. H. Siegel. Constructions and decoding of cyclic codes over b-symbol read channels. *IEEE Trans. Inf. Theory*, 62(4), 1541–1551, 2016.

Reporter: H. Liu