Tensor Codes for the Rank Metric*

Ron M. Roth†

Computer Science Department
Technion — Israel Institute of Technology
Haifa 32000, Israel.
e-mail: ronny@cs.technion.ac.il

Abstract

Linear spaces of $n \times n \times n$ tensors over finite fields are investigated where the rank of every nonzero tensor in the space is bounded from below by a prescribed number μ. Such linear paces can recover any $n \times n \times n$ error tensor of rank $\leq (\mu-1)/2$, and, as such, they can be used to correct three-way crisscross errors. Bounds on the dimensions of such spaces are given for $\mu \leq 2n+1$, and constructions are provided for $\mu \leq 2n-1$ with redundancy which is linear in n. These constructions can be generalized to spaces of $n \times n \times \cdots \times n$ hyper-arrays.

Keywords: Algebraic computation, Crisscross errors, Tensor rank.

*This work was presented in part at the *IEEE International Symposium on Information Theory,* Whistler, BC, Canada, September 1995.

†This research was supported by the Fund for the Promotion of Research at the Technion and by the Technion V.P.R Steiner Research Fund.

1 Introduction

An $n \times n \times n$ tensor over a field F is an $n \times n \times n$ array $\Gamma = [\,\Gamma_{i,j,\ell}\,]_{i,j,\ell=1}^{n}$ whose entries $\Gamma_{i,j,\ell}$ are in F. A tensor $\Gamma = [\,\Gamma_{i,j,\ell}\,]_{i,j,\ell=1}^{n}$ over F is called a *rank-one* tensor if there exist three nonzero vectors $\mathbf{a} = [\,a_1\, a_2\, \ldots\, a_n\,]$, $\mathbf{b} = [\,b_1\, b_2\, \ldots\, b_n\,]$, and $\mathbf{c} = [\,c_1\, c_2\, \ldots\, c_n\,]$ over F such that

$$\Gamma_{i,j,\ell} \;=\; a_i b_j c_\ell \quad \text{for} \quad i,j,\ell = 1,2,\ldots,n\;,$$

or, in shorthand notation,

$$\Gamma \;=\; \mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}\;.$$

The *rank* of an $n \times n \times n$ tensor Γ, denoted $\mathrm{rank}(\Gamma)$, is the smallest number ρ of rank-one tensors $\Gamma_1, \Gamma_2, \ldots, \Gamma_\rho$ such that $\Gamma = \sum_{m=1}^{\rho} \Gamma_m$. The rank of the all-zero tensor is zero. The definitions of tensor and tensor rank extend easily to $n \times m \times s$ tensors, where n, m, and s are not necessarily equal. However, for the sake of simplicity, we will assume throughout this paper that tensors are cubic, namely, $n = m = s$.

Tensor rank is a generalization of the well-known notion of matrix rank. Indeed, every $n \times n$ matrix over F of rank 1 has the form $[\,a_i b_j\,]_{i,j=1}^{n} = \mathbf{a} \otimes \mathbf{b}$, where $\mathbf{a} = [\,a_1\, a_2\, \ldots\, a_n\,]$ and $\mathbf{b} = [\,b_1\, b_2\, \ldots\, b_n\,]$ are nonzero vectors over F. Similarly, the rank of a matrix is the smallest number of rank-one matrices that sum to that matrix. It also follows from this definition that matrix rank (respectively, tensor rank) satisfies the triangle inequality: for any three matrices (respectively, tensors) Γ_1, Γ_2, and Γ_3 we have

$$\mathrm{rank}(\Gamma_1 - \Gamma_2) \;\le\; \mathrm{rank}(\Gamma_1 - \Gamma_3) + \mathrm{rank}(\Gamma_3 - \Gamma_2)\;.$$

Hence, the mapping $(\Gamma_1, \Gamma_2) \mapsto \mathrm{rank}(\Gamma_1 - \Gamma_2)$ is a *metric*.

Nevertheless, unlike the matrix case, deciding upon the value of the rank of a tensor is known to be NP-hard [13]. It is also known that there are no "nonsingular" tensors for $n > 1$, since the rank of any $n \times n \times n$ tensor is strictly smaller than n^2 for every such n [1], [15], [19].

The definition of rank extends to $\overbrace{n \times n \times \cdots \times n}^{\Delta \text{ times}}$ (in short, $n^{\times\Delta}$) tensors (or hyper-arrays) over F, where a rank-one $n^{\times\Delta}$ tensor $\Gamma = [\,\Gamma_{i_1,i_2,\ldots,i_\Delta}\,]_{i_1,i_2,\ldots,i_\Delta=1}^{n}$ is now of the form

$$\Gamma \;=\; \mathbf{a}_1 \otimes \mathbf{a}_2 \otimes \cdots \otimes \mathbf{a}_\Delta \;=\; \bigotimes_{m=1}^{\Delta} \mathbf{a}_m \tag{1}$$

for some Δ nonzero vectors $\mathbf{a}_m = [\,a_{m,1}\, a_{m,2}\, \ldots\, a_{m,n}\,]$, $m = 1,2,\ldots,\Delta$, over F; namely,

$$\Gamma_{i_1,i_2,\ldots,i_\Delta} \;=\; a_{1,i_1} a_{2,i_2} \cdots a_{\Delta,i_\Delta} \quad \text{for} \quad i_1, i_2, \ldots, i_m = 1,2,\ldots,n\;.$$

A μ-$[n^{\times\Delta}, k]$ tensor code $\mathcal{C}$ over a field F is a k-dimensional linear subspace of the vector space of all $n^{\times\Delta}$ tensors over F such that the rank of any nonzero tensor in $\mathcal{C}$ is at least

equal to μ, with equality holding for at least one tensor in $\mathcal{C}$. We call $n^\Delta - k$ the *redundancy* of $\mathcal{C}$ and μ the *minimum rank* of $\mathcal{C}$. We will use the term *array codes* for the case $\Delta = 2$.

A Singleton-type bound on the minimum rank states that the minimum rank and the redundancy of any μ-$[n^{\times\Delta}, k]$ tensor code over a field F satisfy the relation

$$n^\Delta - k \;\geq\; (\mu - 1)\, n \;. \tag{2}$$

This bound was stated by Delsarte in [7] for the case $\Delta = 2$ (see also [10] and the generalization for larger Δ in [26]). Furthermore, Delsarte obtained a construction of μ-$[n \times n, k]$ array codes over $GF(q)$ that attains this bound for every $\mu \leq n$ (see also [10] and [26]). We describe next this optimal construction, which we denote by $\mathcal{C}(n, \mu, 2; q)$ (the parameter 2 stands for $\Delta = 2$). Let $\boldsymbol{\beta} = [\,\beta_j\,]_{j=1}^n$ and $\boldsymbol{\omega} = [\,\omega_\ell\,]_{\ell=1}^n$ be two vectors in $GF(q^n)$, each with entries that are linearly independent over $GF(q)$. The array code $\mathcal{C}(n, \mu, 2; q)$ consists of all $n \times n$ matrices $\Gamma = [\,\Gamma_{j,\ell}\,]_{j,\ell=1}^n$ over $GF(q)$ such that

$$\sum_{j,\ell=1}^{n} \Gamma_{j,\ell}\, \beta_j^{q^s} \omega_\ell \;=\; 0 \;, \quad s = 0, 1, \ldots, \mu{-}2 \;. \tag{3}$$

The set $\{\, \Gamma\boldsymbol{\omega} \;:\; \Gamma \in \mathcal{C}(n, \mu, 2; q) \,\}$ forms a linear code of length n over $GF(q^n)$. It is shown in [10], [26] that such a code is generated by a matrix of the form $[\,\delta_j^{q^s}\,]_{s=0,j=1}^{n-\mu,n}$, where the δ_j's are linearly independent elements of $GF(q^n)$ over $GF(q)$. It follows that Γ is in $\mathcal{C}(n, \mu, 2; q)$ if and only if there exist $\eta_0, \eta_1, \ldots, \eta_{n-\mu} \in GF(q^n)$ such that

$$(\Gamma\boldsymbol{\omega})_j \;=\; \sum_{\ell=1}^{n} \Gamma_{j,\ell}\omega_\ell \;=\; \sum_{s=0}^{n-\mu} \eta_s \delta_j^{q^s} \;, \quad j = 1, 2, \ldots, n \;; \tag{4}$$

namely, Γ is a *matrix representation* of the linear transformation $\eta : GF(q^n) \to GF(q^n)$ over $GF(q)$ which is given by $x \mapsto \eta(x) = \sum_{s=0}^{n-\mu} \eta_s x^{q^s}$. Any nonzero η, being both a linear transformation and a polynomial of degree $\leq q^{n-\mu}$, has a null space of dimension $\leq n{-}\mu$. Hence, the rank of any nonzero Γ is at least μ. This is essentially the result obtained by Delsarte in [7, Section 6]. Using a different approach, this lower bound on the rank of any nonzero $\Gamma \in \mathcal{C}(n, \mu, 2; q)$ was also obtained in [10] and [26].

In [11] and [26], it was shown how a certain model of errors — so-called crisscross errors — can be handled optimally by using such array codes. A discussion was given in [26] also for larger Δ. In the crisscross model, an error corresponds to a corrupted line (i.e., a row or a column when $\Delta = 2$). If we let Γ be the "transmitted" tensor and $\Gamma + E$ be the "received" tensor, then it is not difficult to show that the number of crisscross errors is bounded from below by the rank of E. Since the mapping $(\Gamma_1, \Gamma_2) \mapsto \mathrm{rank}(\Gamma_1 - \Gamma_2)$ is a metric, then by using the elements of a μ-$[n^{\times\Delta}, k]$ tensor code for transmission, we can recover any error tensor of rank $\leq (\mu{-}1)/2$, and, therefore, we can correct any pattern of up to $(\mu{-}1)/2$ crisscross errors. There are various applications of the crisscross error model; see, for instance, [2], [8], [9], [18], [22], [23], [25], [26]. In particular, the three-way crisscross

model of errors in tensors (i.e., the case $\Delta = 3$) can be found in practice in certain memory chips [27].

Another important application of tensor rank is found also in the area of algebraic complexity, and, in particular, in the study of the computational complexity of sets of bilinear forms. See [16, Section 4.6.4], [28], [30]. By a set of bilinear forms we mean a set of expressions $Z = \{z_\ell\}_\ell$ over a field F, each expression z_ℓ having the form $\sum_{i,j} x_i \Gamma_{i,j,\ell} y_j$, where $\Gamma = [\,\Gamma_{i,j,\ell}\,]$ is a tensor over F and x_i and y_j are indeterminates. Examples of sets of bilinear forms include polynomial multiplication and matrix multiplication. It is known that the rank of Γ equals the *bilinear complexity* of the set Z, namely, the minimum number of noncommutative nonscalar multiplications in any so-called normal computation of Z by a straight-line algorithm (see [16, Section 4.6.4]). When the field F is sufficiently large, the bilinear complexity of Z coincides with the minimum number of noncommutative nonscalar multiplications in any straight-line algorithm that computes Z [30, Chapter III].

There is a very close relationship between the question of determining the bilinear complexity of multiplying two polynomials over $GF(q)$ and the problem of constructing good asymptotic families of error-correcting codes. See [4], [17], and the remarkable result by Chudnovsky and Chudnovsky [6].

The purpose of this work is to continue the work of [7], [10], and [26] and present constructions of linear spaces of $n^{\times\Delta}$ tensors for $\Delta \geq 3$ and to obtain bounds on the dimensions of such spaces. The apparent difficulty in handling the general tensor case (as opposed to the special case $\Delta = 2$) is due, in part, to the fact that although the typical rank of $n^{\times 3}$ tensors is quadratic in n [19], we do not know yet of any explicit construction of an infinite sequence of $n^{\times 3}$ tensors for increasing values of n with ranks that are at least super-linear in n. See also [3], [12], [15].

We will mainly concentrate here on μ-$[n^{\times\Delta}, k]$ tensor codes over finite fields with $\mu = O(n)$ (throughout this paper, $O(x)$ stands for an expression which is bounded from above by cx for some absolute constant c). We first present in Section 2 a sphere-packing bound for $\mu \leq 2n+1$. Then, in Section 3, we present a construction of μ-$[n^{\times\Delta}, k]$ tensor codes with redundancy at most $\binom{\mu+\Delta-3}{\Delta-1} n$. When we fix μ, this redundancy becomes *linear* in n, which is smaller than a redundancy proportional to $n \log_q n$ that would be needed in the simpler skewing crisscross coding method (Section 4). For $\mu \leq 2$, the construction attains the Singleton-type bound (2) on the minimum rank and for $\mu = 3$ the construction approaches the sphere-packing bound. We also point out an interesting connection between our construction and the set of bilinear forms that corresponds to modular polynomial multiplication (Section 3.2). Decoding algorithms for correcting one error and two errors are given in Section 5 and, finally, some remarks on the infinite-field case are given in Section 6.

2 Upper bounds

As mentioned before, the Singleton-type bound (2) on the minimum rank can be attained for every $\mu \leq n$ over every finite field when $\Delta = 2$. On the other hand, this bound on the minimum rank was sharpened in [26] for $\Delta = 3$ to

$$n^3 - k \;\geq\; (\mu - 1 + \sigma^2)\, n \;, \quad \text{where} \quad \sigma = n - \left\lceil \sqrt{n^2 - \mu + 1} \right\rceil \;. \tag{5}$$

Indeed, the value of σ is greater than zero when $\mu \geq 2n$, in which case the bound (5) is strictly stronger than (2). In this section, we improve on (2) for $\mu \leq 2n{+}1$ by using sphere-packing arguments.

Denote by $\mathcal{R}(n, \rho, \Delta; q)$ the number of $n^{\times\Delta}$ tensors of rank $\leq \rho$ over $GF(q)$.

Lemma 1. *The redundancy $n^\Delta{-}k$ of every $(2\rho{+}1)$-$[n^{\times\Delta}, k]$ tensor code over $GF(q)$ must be at least $\log_q \mathcal{R}(n, \rho, \Delta; q)$.*

Proof. Let $\mathcal{C}$ be a $(2\rho{+}1)$-$[n^{\times\Delta}, k]$ tensor code over $GF(q)$ and consider the space of $n^{\times\Delta}$ tensors over $GF(q)$. Since tensor rank defines a metric, the spheres of radius ρ in that space that are centered at the q^k elements of $\mathcal{C}$ must be disjoint. Hence, $q^k \cdot \mathcal{R}(n, \rho, \Delta; q) \leq q^{n^\Delta}$.□

Theorem 1. (Sphere-packing bound for $\mu = 3$). *The redundancy $n^\Delta{-}k$ of every 3-$[n^{\times\Delta}, k]$ tensor code satisfies*

$$n^\Delta - k \;\geq\; \Delta n \;-\; (\Delta - 1)\log_q(q-1) \;-\; O(\Delta/(q^n \ln q)) \;,$$

In particular, for $q = 2$,

$$n^\Delta - k \;\geq\; \Delta n \;-\; O(\Delta/2^n) \;.$$

Proof. We first calculate $\mathcal{R}(n, 1, \Delta; q)$. Every tensor Γ of rank 1 can be written as in (1) for some Δ nonzero vectors $\mathbf{a}_m$, $m = 1, 2, \ldots, \Delta$, over $GF(q)$. Furthermore, if the $\Delta{-}1$ vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{\Delta-1}$ are normalized so that the leading nonzero component in each of those vectors is 1, then the decomposition (1) is *unique*. Hence,

$$\mathcal{R}(n, 1, \Delta; q) \;=\; 1 \;+\; \frac{(q^n - 1)^\Delta}{(q-1)^{\Delta-1}} \;.$$

Now, by Lemma 1 we have, for 3-$[n^{\times\Delta}, k]$ tensor codes,

$$q^{n^\Delta - k} \;-\; 1 \;\geq\; \frac{(q^n - 1)^\Delta}{(q-1)^{\Delta-1}} \;,$$

which implies that the redundancy $n^{\Delta}-k$ satisfies the inequality

$$\begin{aligned} n^{\Delta}-k &> \Delta\log_q(q^n-1) \;-\; (\Delta-1)\log_q(q-1) \\ &= \Delta n \;+\; \Delta\log_q(1-(1/q^n)) \;-\; (\Delta-1)\log_q(q-1) \\ &= \Delta n \;-\; (\Delta-1)\log_q(q-1) \;-\; O(\Delta/(q^n\ln q))\;, \end{aligned}$$

as claimed. □

We turn to stating a sphere-packing bound for μ-$[n^{\times\Delta},k]$ tensor codes where $\mu=2\rho+1$ and $1\le\rho\le n$. A bound for even values of μ will be implied by a bound for $\mu-1$. For the sake of clarity we deal first with the case $\Delta=3$.

Lemma 2. *For $1\le\rho\le n$,*

$$\mathcal{R}(n,\rho,3;q) \;\ge\; 1 \;+\; \sum_{s=1}^{\rho}\binom{(q^n-1)/(q-1)}{s}\;\cdot\;\Big(\prod_{i=0}^{s-1}(q^n-q^i)\Big)^2\;\cdot\;\frac{1}{(q-1)^s}\;.$$

Proof. For $s\in\{1,2,\ldots,\rho\}$, let $\Gamma=[\,\Gamma_{i,j,\ell}\,]_{i,j,\ell=1}^{n}$ be given by

$$\Gamma \;=\; \sum_{t=1}^{s}(\mathbf{a}_t\otimes\mathbf{b}_t\otimes\mathbf{c}_t)\;, \tag{6}$$

where $A=\{\,\mathbf{a}_t\,\}_{t=1}^{s}$, $B=\{\,\mathbf{b}_t\,\}_{t=1}^{s}$, and $C=\{\,\mathbf{c}_t\,\}_{t=1}^{s}$ are three ordered sets, each consisting of s vectors in $(GF(q))^n$ such that —

(i) the vectors in A and B are normalized to have a leading nonzero component 1;

(ii) the vectors in B are linearly independent, and so are the vectors in C; and —

(iii) assuming some fixed ordering on the elements of $(GF(q))^n$, the vectors $\mathbf{a}_t$ are nonzero, distinct, and $\mathbf{a}_t<\mathbf{a}_{t'}$ for $t<t'$.

Consider the s rank-one $n\times n$ matrices $X_t=\mathbf{a}_t\otimes\mathbf{b}_t$, $t=1,2,\ldots,s$. We first claim that these matrices are linearly independent over $GF(q)$. Indeed, suppose that $\sum_{t=1}^{s}z_tX_t=0$ for some $z_t\in GF(q)$, that is

$$\sum_{t=1}^{s}z_t\,(\mathbf{a}_t\otimes\mathbf{b}_t) \;=\; \sum_{t=1}^{s}(z_t\,\mathbf{a}_t)\otimes\mathbf{b}_t \;=\; 0\;.$$

Since the s vectors $\mathbf{b}_t$ are linearly independent over $GF(q)$, we must have

$$z_t\,\mathbf{a}_t \;=\; \mathbf{0}\;,\quad t=1,2,\ldots,s\;.$$

However, the vectors $\mathbf{a}_t$ are nonzero, and, so, $z_t=0$ for $t=1,2,\ldots,s$.

For $\ell = 1, 2, \ldots, n$, define the $n \times n$ matrix $Y_\ell = [\,\Gamma_{i,j,\ell}\,]_{i,j=1}^{n}$ (i.e., Y_ℓ is a "slice" of Γ). By (6) we can write Y_ℓ as

$$Y_\ell \;=\; \sum_{t=1}^{s} c_{t,\ell} X_t \;, \tag{7}$$

where $c_{t,\ell}$ is the ℓth entry of $\mathbf{c}_t$. Furthermore, since the s vectors $\mathbf{c}_t$ are linearly independent, the sets $\{X_t\}_{t=1}^{s}$ and $\{Y_\ell\}_{\ell=1}^{n}$ span the same subspace of dimension s of $n \times n$ matrices over $GF(q)$.

Next we claim that the triple (A, B, C) for a given Γ, under conditions (i)–(iii), is uniquely defined. Indeed, suppose that there exists another triple of ordered sets, $A' = \{\,\mathbf{a}'_t\,\}_{t=1}^{s'}$, $B' = \{\,\mathbf{b}'_t\,\}_{t=1}^{s'}$, and $C' = \{\,\mathbf{c}'_t\,\}_{t=1}^{s'}$, satisfying (i)–(iii), such that

$$\Gamma \;=\; \sum_{t=1}^{s} (\mathbf{a}_t \otimes \mathbf{b}_t \otimes \mathbf{c}_t) \;=\; \sum_{t=1}^{s'} (\mathbf{a}'_t \otimes \mathbf{b}'_t \otimes \mathbf{c}'_t) \;.$$

The respective s' matrices $X'_t = \mathbf{a}'_t \otimes \mathbf{b}'_t$ are linearly independent over $GF(q)$ and, following Equation (7), we also have $Y_\ell = \sum_{t=1}^{s'} c'_{t,\ell} X'_t$. Hence, the sets $\{X_t\}_{t=1}^{s}$ and $\{X'_t\}_{t=1}^{s'}$ span the same linear subspace of $n \times n$ matrices over $GF(q)$, implying that $s' = s$ and that there exists a nonsingular $s \times s$ matrix $[\,u_{t,m}\,]_{t,m=1}^{s}$ such that

$$X'_t \;=\; \sum_{m=1}^{s} u_{t,m} X_m \;, \quad t = 1, 2, \ldots, s \;,$$

namely,

$$\mathbf{a}'_t \otimes \mathbf{b}'_t \;=\; \sum_{m=1}^{s} (u_{t,m} \mathbf{a}_m) \otimes \mathbf{b}_m \;, \quad t = 1, 2, \ldots, s \;. \tag{8}$$

Since each vector $\mathbf{a}'_t$ is nonzero, it follows by (8) that there exist coefficients $v_{t,m}$ such that

$$\mathbf{b}'_t \;=\; \sum_{m=1}^{s} v_{t,m} \mathbf{b}_m \;, \quad t = 1, 2, \ldots, s \;. \tag{9}$$

Plugging (9) back into (8) we obtain,

$$\mathbf{a}'_t \otimes \sum_{m=1}^{s} v_{t,m} \mathbf{b}_m \;=\; \sum_{m=1}^{s} (u_{t,m} \mathbf{a}_m) \otimes \mathbf{b}_m \;, \quad t = 1, 2, \ldots, s \;,$$

or

$$\sum_{m=1}^{s} (v_{t,m} \mathbf{a}'_t - u_{t,m} \mathbf{a}_m) \otimes \mathbf{b}_m \;=\; 0 \;, \quad t = 1, 2, \ldots, s \;.$$

By the linear independence of the vectors $\mathbf{b}_t$ we thus have

$$v_{t,m} \mathbf{a}'_t \;=\; u_{t,m} \mathbf{a}_m \;, \quad t = 1, 2, \ldots, s \;, \quad m = 1, 2, \ldots, s \;.$$

However, the vectors $\mathbf{a}_t$ are normalized, distinct, and ordered within A, and the same holds for the vectors $\mathbf{a}'_t$ within A'. Hence, we must have $\mathbf{a}_t = \mathbf{a}'_t$ and $v_{t,t} = u_{t,t}$; furthermore,

$u_{t,m} = v_{t,m} = 0$ if $t \neq m$. Returning to (9) and recalling that the vectors $\mathbf{b}_t$ and $\mathbf{b}'_t$ are normalized, we conclude that $\mathbf{b}_t = \mathbf{b}'_t$ and therefore $X_t = X'_t$. The uniqueness of $\mathbf{c}_t$ now follows from (7) and from the linear independence of the matrices X_t.

For distinct values of s we obtain disjoint sets of $n^{\times 3}$ tensors in (6). The lemma is obtained by counting the triples (A, B, C) that satisfy (i)–(iii) for $s = 1, 2, \ldots, \rho$. □

For $s \leq n$ we can write

$$\prod_{i=0}^{s-1}(q^n - q^i) \;=\; q^{ns}\prod_{i=0}^{s-1}(1 - q^{i-n}) \;\geq\; \lambda_q \cdot q^{ns}\;,$$

where $\lambda_q = \prod_{j=1}^{\infty}(1 - q^{-j})$ and $\lambda_q \geq \lambda_2 \approx 0.3$ (see [24, p. 755]). Therefore,

$$\mathcal{R}(n, \rho, 3; q) \;\geq\; 1 \;+\; \sum_{s=1}^{\rho}\frac{\lambda_q^3 q^{3ns}}{s!(q-1)^{2s}} \;>\; \frac{\lambda_2^3 q^{3n\rho}}{\rho!(q-1)^{2\rho}}\;.$$

By Lemma 1 it follows that

$$n^3 - k \;\geq\; 3n\rho \;-\; 2\rho\log_q(q-1) \;-\; \log_q(\rho!) \;-\; O(1/\ln q)\;.$$

Writing $\rho = \lfloor(\mu{-}1)/2\rfloor$ we thus obtain the following.

Theorem 2. *Let $\mu \leq 2n{+}1$. Then, for every μ-$[n^{\times 3}, k]$ tensor code,*

$$n^3 - k \;\geq\; \lfloor(\mu{-}1)/2\rfloor\Big(3n \;-\; 2\log_q(q-1) \;-\; \log_q\lfloor(\mu{-}1)/2\rfloor\Big) \;-\; O(1/\ln q)\;,$$

and, therefore,

$$n^3 \;-\; k \;\geq\; 3\lfloor(\mu{-}1)/2\rfloor\, n\,(1 - \epsilon(n))\;,$$

where $\lim_{n\to\infty}\epsilon(n) = 0$.

Theorem 2 can be generalized to any $\Delta \geq 3$ as follows. The vectors $\mathbf{a}_t$ in the proof of Lemma 2 are replaced by $s \leq \rho$ nonzero distinct normalized $n^{\times(\Delta-2)}$ rank-one tensors, in which case we have

$$\mathcal{R}(n, \rho, \Delta; q) \;\geq\; 1 \;+\; \sum_{s=1}^{\rho}\binom{((q^n-1)/(q-1))^{\Delta-2}}{s} \;\cdot\; \Big(\prod_{i=0}^{s-1}(q^n - q^i)\Big)^2 \;\cdot\; \frac{1}{(q-1)^s}\;.$$

This, in turn, implies

$$n^{\Delta} - k \;\geq\; \Delta\lfloor(\mu{-}1)/2\rfloor\, n\,(1 - \epsilon_{\Delta}(n))\;, \tag{10}$$

where $\lim_{n\to\infty}\epsilon_{\Delta}(n) = 0$.

In contrast, by a Gilbert-Varshamov-type bound obtained in [26], the inequality

$$n^{\Delta} - k \;\geq\; \Delta(\mu{-}1) \tag{11}$$

is a sufficient condition for having a μ-$[n^{\times\Delta}, k]$ tensor code over $GF(q)$. Thus, we have a factor close to 2 between the sphere-packing bound (10) and the Gilbert-Varshamov-type bound (11) on the redundancy of tensor codes. A similar gap is known to appear also in the respective bounds for high-rate conventional codes in the Hamming metric.

3 Construction of tensor codes

In this section, we provide a generalization of the construction $\mathcal{C}(n, \mu, 2; q)$ for larger Δ, starting with the $n^{\times 3}$ tensor case.

3.1 The $n^{\times 3}$ tensor case

Let $\mathcal{S}(n, \mu, 3; q)$ be the set of all pairs (r, s), where r and s range over all integers that satisfy the following two conditions:

(a) $0 \leq r, s < n$ and —

(b) there exists a (conventional) linear code over $GF(q)$ of length $\mu-1$, dimension $r+1$, and minimum Hamming distance at least equal to $s+1$.

Let $\boldsymbol{\alpha} = [\,\alpha_i\,]_{i=1}^{n}$, $\boldsymbol{\beta} = [\,\beta_j\,]_{j=1}^{n}$, and $\boldsymbol{\omega} = [\,\omega_\ell\,]_{\ell=1}^{n}$ be three vectors in $GF(q^n)$, each with entries that are linearly independent over $GF(q)$. The tensor code $\mathcal{C}(n, \mu, 3; q)$ is defined as the set of all tensors $\Gamma = [\,\Gamma_{i,j,\ell}\,]_{i,j,\ell=1}^{n}$ over $GF(q)$ such that

$$\sum_{i,j,\ell=1}^{n} \Gamma_{i,j,\ell}\, \alpha_i^{q^r}\, \beta_j^{q^s}\, \omega_\ell \;=\; 0 \quad \text{for every} \quad (r, s) \in \mathcal{S}(n, \mu, 3; q)\ . \tag{12}$$

We will hereafter use the notation $\boldsymbol{\alpha}^i$ for $[\,\alpha_1^i\ \alpha_2^i\ \ldots\ \alpha_n^i\,]$. The inner product of two vectors $\mathbf{x}$ and $\mathbf{x}'$ will be denoted by $\langle \mathbf{x}, \mathbf{x}' \rangle$, and we will extend this notation to tensors as follows: for two $n^{\times 3}$ tensors $\Gamma = [\,\Gamma_{i,j,\ell}\,]_{i,j,\ell=1}^{n}$ and $\Gamma' = [\,\Gamma'_{i,j,\ell}\,]_{i,j,\ell=1}^{n}$, we denote by $\langle \Gamma, \Gamma' \rangle$ the sum $\sum_{i,j,\ell=1}^{n} \Gamma_{i,j,\ell}\Gamma'_{i,j,\ell}$. Using these definitions, Equation (12) can be re-written as

$$\left\langle \Gamma\ ,\ \boldsymbol{\alpha}^{q^r} \otimes \boldsymbol{\beta}^{q^s} \otimes \boldsymbol{\omega} \right\rangle \;=\; 0 \quad \text{for every} \quad (r, s) \in \mathcal{S}(n, \mu, 3; q)\ . \tag{13}$$

We can represent the tensor code $\mathcal{C}(n, \mu, 3; q)$ as a conventional linear code of length n^2 over $GF(q^n)$ as follows. Let $\mathcal{H} = [\,\mathcal{H}_{(r,s),(i,j)}\,]_{(r,s),(i,j)}$ be the $|\mathcal{S}(n, \mu, 3; q)| \times n^2$ matrix over $GF(q^n)$ whose columns are indexed by pairs (i, j) such that $1 \leq i, j \leq n$, whose rows are indexed by pairs $(r, s) \in \mathcal{S}(n, \mu, 3; q)$, and whose entries are given by

$$\mathcal{H}_{(r,s),(i,j)} = \alpha_i^{q^r}\, \beta_j^{q^s}\ . \tag{14}$$

For each tensor $\Gamma = [\,\Gamma_{i,j,\ell}\,]_{i,j,\ell=1}^{n}$, we associate a vector $\boldsymbol{\gamma} = [\,\gamma_{(i,j)}\,]_{(i,j)}$ of length n^2 over $GF(q^n)$ whose entries are given by $\gamma_{(i,j)} = \sum_{\ell=1}^{n} \Gamma_{i,j,\ell}\omega_\ell$, $1 \leq i, j \leq n$. By (12) it follows that Γ is in $\mathcal{C}(n, \mu, 3; q)$ if and only if $\mathcal{H}\boldsymbol{\gamma} = \mathbf{0}$.

Next we obtain bounds on the redundancy and minimum rank of $\mathcal{C}(n, \mu, 3; q)$, making use of the following two lemmas.

Lemma 3. *Let $\delta_1, \delta_2, \ldots, \delta_h$ be elements of $GF(q^n)$ that span a linear space of dimension ρ over $GF(q)$. Then, the rows of the matrix*

$$\mathcal{D} \;=\; \begin{bmatrix} \delta_1 & \delta_2 & \ldots & \delta_h \\ \delta_1^q & \delta_2^q & \ldots & \delta_h^q \\ \delta_1^{q^2} & \delta_2^{q^2} & \ldots & \delta_h^{q^2} \\ \vdots & \vdots & \vdots & \vdots \\ \delta_1^{q^{\rho-1}} & \delta_2^{q^{\rho-1}} & \ldots & \delta_h^{q^{\rho-1}} \end{bmatrix}$$

are spanned by ρ linearly independent vectors whose components belong to $GF(q)$.

Proof. It is known that $\mathrm{rank}(\mathcal{D}) = \rho$ and that the right null space of $\mathcal{D}$ over $GF(q^n)$ is spanned by the columns of a full-rank $h \times (h-\rho)$ matrix U over $GF(q)$ (see [20, p. 109] and [26]). The left null space of U over $GF(q)$, in turn, is spanned by the rows of a full-rank $\rho \times h$ matrix G. It follows that the rows of $\mathcal{D}$ and G span the same linear space over $GF(q^n)$. □

Lemma 4. *Let $\boldsymbol{\alpha} = [\,\alpha_i\,]_{i=1}^{n}$ and $\boldsymbol{\beta} = [\,\beta_j\,]_{j=1}^{n}$ be vectors in $GF(q^n)$, each with entries that are linearly independent over $GF(q)$. Define the $n^2 \times n^2$ matrix $\mathcal{M} = [\,\mathcal{M}_{(r,s),(i,j)}\,]_{(r,s),(i,j)}$ over $GF(q^n)$ by*

$$\mathcal{M}_{(r,s),(i,j)} \;=\; \alpha_i^{q^r}\,\beta_j^{q^s}\;,$$

where the row index (r,s) ranges over pairs such that $0 \le r, s < n$ and the column index (i,j) ranges over pairs such that $1 \le i, j \le n$. Then the matrix $\mathcal{M}$ is nonsingular.

Proof. The matrix $\mathcal{M}$ is a Kronecker (direct) product of the two $n \times n$ matrices $[\,\alpha_i^{q^r}\,]_{r=0,i=1}^{n-1,n}$ and $[\,\beta_j^{q^s}\,]_{s=0,j=1}^{n-1,n}$. By Lemma 3, these two matrices are nonsingular. Hence, so is $\mathcal{M}$ [14, p. 244, Corollary 4.2.11]. □

Theorem 3. *The redundancy of $\mathcal{C}(n,\mu,3;q)$ equals $|\mathcal{S}(n,\mu,3;q)|\,n$.*

Proof. Lemma 4 implies that the rows of $\mathcal{H}$ as defined by (14) are linearly independent over $GF(q^n)$. □

Let $\mathcal{K}(\nu,d;q)$ denote the largest dimension of any linear code over $GF(q)$ of length ν and minimum Hamming distance at least d. Then condition (b) in the definition of $\mathcal{S}(n,\mu,3;q)$ amounts to the inequality

$$r \;<\; \mathcal{K}(\mu{-}1, s{+}1; q)\;.$$

Therefore, the set $\mathcal{S}(n,\mu,3;q)$ can be written as follows:

$$\mathcal{S}(n,\mu,3;q) \;=\; \Big\{\,(r,s)\;:\;0 \le s < n \quad\text{and}\quad 0 \le r < \min\{n, \mathcal{K}(\mu{-}1, s{+}1; q\}\,\Big\}\,. \qquad (15)$$

In particular, by the Singleton bound on the minimum Hamming distance of conventional linear codes, we have

$$\mathcal{K}(\mu{-}1, s{+}1; q) \;\leq\; \mu{-}s{-}1 \tag{16}$$

(and this inequality can be attained when $\mu \leq q{+}2$ by extended Reed-Solomon codes [21, Ch. 1]). Combining (15) and (16), we have that $(r,s) \in \mathcal{S}(n,\mu,3;q)$ implies the inequalities $0 \leq r, s < n$ and $r + s \leq \mu{-}2$. Hence, by Theorem 3 we obtain the following upper bound on the redundancy of $\mathcal{C}(n,\mu,3;q)$:

$$n^3 - k \;\leq\; \begin{cases} \binom{\mu}{2}\, n & \text{for } \mu = 1, 2, \ldots, n \\ n^3 - \binom{2n-\mu+1}{2}\, n & \text{for } \mu = n{+}1, n{+}2, \ldots, 2n{-}1 \end{cases} .$$

Theorem 4. *The minimum rank of $\mathcal{C}(n,\mu,3;q)$ is at least μ.*

Proof. The proof is carried out by induction on μ, where the induction base $\mu = 1$ is immediate. Given $\mu > 1$, we assume that the statement holds for every $\mu' < \mu$.

Let Γ be a tensor in $\mathcal{C}(n,\mu,3;q)$ and suppose that $\mathrm{rank}(\Gamma) < \mu$. Then there exist vectors $\mathbf{a}_t$, $\mathbf{b}_t$, $\mathbf{c}_t$ for $t = 1, 2, \ldots, \mu{-}1$ over $GF(q)$ such that

$$\Gamma \;=\; \sum_{t=1}^{\mu-1} (\mathbf{a}_t \otimes \mathbf{b}_t \otimes \mathbf{c}_t) \; .$$

Define

$$A_t \;=\; \langle \mathbf{a}_t, \boldsymbol{\alpha} \rangle \; , \quad B_t \;=\; \langle \mathbf{b}_t, \boldsymbol{\beta} \rangle \; , \quad \text{and} \quad C_t \;=\; \langle \mathbf{c}_t, \boldsymbol{\omega} \rangle \quad \text{for} \quad t = 1, 2, \ldots, \mu{-}1 \; .$$

The following chain of equalities can be easily verified for every $r, s \geq 0$:

$$\begin{aligned} \Big\langle \Gamma \;,\; \boldsymbol{\alpha}^{q^r} \otimes \boldsymbol{\beta}^{q^s} \otimes \boldsymbol{\omega} \Big\rangle \;&=\; \Big\langle \sum_{t=1}^{\mu-1} \Big(\mathbf{a}_t \otimes \mathbf{b}_t \otimes \mathbf{c}_t \Big) \;,\; \boldsymbol{\alpha}^{q^r} \otimes \boldsymbol{\beta}^{q^s} \otimes \boldsymbol{\omega} \Big\rangle \\ &=\; \sum_{t=1}^{\mu-1} \Big\langle \mathbf{a}_t \otimes \mathbf{b}_t \otimes \mathbf{c}_t \;,\; \boldsymbol{\alpha}^{q^r} \otimes \boldsymbol{\beta}^{q^s} \otimes \boldsymbol{\omega} \Big\rangle \\ &=\; \sum_{t=1}^{\mu-1} \langle \mathbf{a}_t, \boldsymbol{\alpha}^{q^r} \rangle \cdot \langle \mathbf{b}_t, \boldsymbol{\beta}^{q^s} \rangle \cdot \langle \mathbf{c}_t, \boldsymbol{\omega} \rangle \\ &=\; \sum_{t=1}^{\mu-1} A_t^{q^r} B_t^{q^s} C_t \end{aligned}$$

(the third equality follows from the mixed-product property; see [14, p. 244, Lemma 4.2.10]). Equation (13) is therefore equivalent to

$$\sum_{t=1}^{\mu-1} A_t^{q^r} B_t^{q^s} C_t \;=\; 0 \quad \text{for every} \quad (r,s) \in \mathcal{S}(n,\mu,3;q) \; . \tag{17}$$

We will show that (17) implies $\Gamma = \sum_{t=1}^{\mu-1}(\mathbf{a}_t \otimes \mathbf{b}_t \otimes \mathbf{c}_t) = 0$.

Let ρ denote the dimension of the linear space over $GF(q)$ which is spanned by the elements $\{A_t\}_{t=1}^{\mu-1}$. If $\rho = 0$ then $A_t = 0$ for all t and we are done. Therefore we assume that $\rho > 0$. Let $\mathcal{A}$ be the following $\rho \times (\mu-1)$ matrix over $GF(q^n)$:

$$\mathcal{A} \;=\; \begin{bmatrix} A_1 & A_2 & \dots & A_{\mu-1} \\ A_1^q & A_2^q & \dots & A_{\mu-1}^q \\ A_1^{q^2} & A_2^{q^2} & \dots & A_{\mu-1}^{q^2} \\ \vdots & \vdots & \vdots & \vdots \\ A_1^{q^{\rho-1}} & A_2^{q^{\rho-1}} & \dots & A_{\mu-1}^{q^{\rho-1}} \end{bmatrix} .$$

By Lemma 3, the rows of $\mathcal{A}$ are spanned by the rows of a $\rho \times (\mu-1)$ matrix G over $GF(q)$. Let d be the minimum Hamming distance of the conventional linear code over $GF(q)$ which is spanned by the rows of G and let $\mathbf{u} = [\, u_1\, u_2\, \dots\, u_{\mu-1}\,]$ be a codeword in that code with Hamming weight d. For $s = 0, 1, \dots, d-1$ we define the vectors

$$\mathbf{x}_s \;=\; \left[\, B_1^{q^s} C_1 \;\; B_2^{q^s} C_2 \;\; \dots \;\; B_{\mu-1}^{q^s} C_{\mu-1} \,\right] .$$

By (17) we have $\mathcal{A}\,\mathbf{x}_s = \mathbf{0}$ for $s = 0, 1, \dots, d-1$, and, so,

$$\sum_{t=1}^{\mu-1} u_t B_t^{q^s} C_t \;=\; \langle \mathbf{u}, \mathbf{x}_s \rangle \;=\; 0\,, \quad s = 0, 1, \dots, d-1\,. \tag{18}$$

Consider the $n \times n$ matrix

$$\Gamma' \;=\; [\,\Gamma'_{j,\ell}\,]_{j,\ell=1}^{n} \;=\; \sum_{t=1}^{\mu-1} u_t\,(\mathbf{b}_t \otimes \mathbf{c}_t)$$

over $GF(q)$. The matrix Γ' is a sum of d rank-one matrices and, therefore, $\mathrm{rank}(\Gamma') \le d$. On the other hand, for every $s \ge 0$ we have

$$\left\langle \Gamma'\,,\; \boldsymbol{\beta}^{q^s} \otimes \boldsymbol{\omega} \right\rangle \;=\; \left\langle \sum_{t=1}^{\mu-1} u_t\,(\mathbf{b}_t \otimes \mathbf{c}_t)\,,\; \boldsymbol{\beta}^{q^s} \otimes \boldsymbol{\omega} \right\rangle \;=\; \sum_{t=1}^{\mu-1} u_t B_t^{q^s} C_t\,.$$

Hence, by (18) we obtain

$$\sum_{j,\ell=1}^{n} \Gamma'_{j,\ell}\, \beta_j^{q^s} \omega_\ell \;=\; \left\langle \Gamma'\,,\; \boldsymbol{\beta}^{q^s} \otimes \boldsymbol{\omega} \right\rangle \;=\; 0\,, \quad s = 0, 1, \dots, d-1\,.$$

Comparing with (3), we conclude that Γ' is in $\mathcal{C}(n, d+1, 2; q)$. Hence, by the result of [7] we must have $\Gamma' = 0$.

Finally, let τ be such that $u_\tau \neq 0$. The tensor $u_\tau\Gamma$ is in $\mathcal{C}(n,\mu,3;q)$ and, as such, it is also a tensor in $\mathcal{C}(n,\mu{-}1,3;q)$. Now, the matrix $\Gamma' = \sum_{t=1}^{\mu-1} u_t\,(\mathbf{b}_t \otimes \mathbf{c}_t)$ is identically zero, and so is the tensor $\Gamma'' = \mathbf{a}_\tau \otimes \sum_{t=1}^{\mu-1} u_t\,(\mathbf{b}_t \otimes \mathbf{c}_t)$. Hence,

$$u_\tau\Gamma \;=\; u_\tau\Gamma \;-\; \Gamma'' \;=\; \sum_{t=1}^{\mu-1}(u_\tau\mathbf{a}_t - u_t\mathbf{a}_\tau)\otimes\mathbf{b}_t\otimes\mathbf{c}_t\ ,$$

which implies that $\mathrm{rank}(\Gamma) = \mathrm{rank}(u_\tau\Gamma) < \mu{-}1$. Applying the induction hypothesis on $\mathcal{C}(n,\mu{-}1,3;q)$, we must have $\Gamma = 0$. □

It follows from Theorems 3 and 4 that the tensor code $\mathcal{C}(n,\mu,3;q)$ attains the Singleton-type bound (2) on the minimum rank when $\mu = 2$. For $\mu = 3$ we get redundancy $3n$, which, in view of Theorem 1, is optimal over $GF(2)$ for sufficiently large n. (Over larger fields we still have an additive gap of $2\log_q(q-1)$.)

The tensor code $\mathcal{C}(n,\mu,3;q)$ becomes vacuous (namely, containing the zero vector only) when $\mathcal{S}(n,\mu,3;q)$ consists of all n^2 pairs (r,s) such that $0 \le r,s < n$. By (15), this happens whenever μ is large enough so that $\mathcal{K}(\mu{-}1,n;q) \ge n$. For $q = 2$ this occurs only if $\mu{-}1 \ge 3.52n - o(n)$ [4]. By (16), the tensor code $\mathcal{C}(n,\mu,3;q)$ is nonvacuous for every q if $\mu \le 2n{-}1$.

3.2 Dual representation and polynomial multiplication

We obtain next a dual representation of $\mathcal{C}(n,\mu,3;q)$ through a generator matrix over $GF(q^n)$. We will then use such a representation to establish a connection between $\mathcal{C}(n,\mu,3;q)$ and modular polynomial multiplication.

For a basis $\boldsymbol{\beta} = [\,\beta_j\,]_{j=1}^n$ of $GF(q^n)$ over $GF(q)$, we denote by $\boldsymbol{\beta}^\perp = [\,\beta_j^\perp\,]_{j=1}^n$ the *dual basis* of $\boldsymbol{\beta}$, namely, a basis of $GF(q^n)$ over $GF(q)$ such that

$$\sum_{\ell=0}^{n-1}(\beta_i\beta_j^\perp)^{q^\ell} \;=\; \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i\neq j \end{cases}\ . \tag{19}$$

A dual basis always exists [21, p. 118]. By (19) it follows that the matrices $\mathcal{B} = [\,\beta_i^{q^\ell}\,]_{i=1,\ell=0}^{n\ ,n-1}$ and $\mathcal{B}^\perp = [\,(\beta_j^\perp)^{q^\ell}\,]_{j=1,\ell=0}^{n\ ,n-1}$ satisfy $\mathcal{B}^\perp\mathcal{B}^T = I$. Therefore, we also have $\mathcal{B}^T\mathcal{B}^\perp = I$, and, so

$$\langle\,\boldsymbol{\beta}^{q^\ell}, (\boldsymbol{\beta}^\perp)^{q^m}\rangle \;=\; \begin{cases} 1 & \text{if } \ell=m \\ 0 & \text{if } \ell\neq m \end{cases}\ . \tag{20}$$

Denote by $\overline{\mathcal{S}}(n,\mu,3;q)$ the set of all pairs of integers (r,s) such that $0 \le r,s < n$ and $(r,s) \notin \mathcal{S}(n,\mu,3;q)$. Let $\mathcal{G}$ be the $(n^2 - |\mathcal{S}(n,\mu,3;q)|)\times n^2$ matrix over $GF(q^n)$ whose entries are given by

$$\mathcal{G}_{(r,s),(i,j)} \;=\; (\alpha_i^\perp)^{q^r}(\beta_j^\perp)^{q^s}\ , \tag{21}$$

where the row index (r,s) ranges over all elements of $\overline{\mathcal{S}}(n,\mu,3;q)$ and the column index (i,j) ranges over all integers $1 \le i,j \le n$. By Lemma 4 it follows that the matrix $\mathcal{G}$ has full rank; furthermore, by (20) we have that the columns of $\mathcal{G}^T$ span the right kernel of the matrix $\mathcal{H}$ given in (14). Hence the matrix $\mathcal{G}$ can be regarded as a generator matrix of $\mathcal{C}(n,\mu,3;q)$ in the following sense: a tensor $\Gamma = [\,\Gamma_{i,j,\ell}\,]_{i,j,\ell=1}^{n}$ is in $\mathcal{C}(n,\mu,3;q)$ if and only if there exist elements $\eta_{(r,s)} \in GF(q^n)$, indexed by $(r,s) \in \overline{\mathcal{S}}(n,\mu,3;q)$, such that

$$\sum_{\ell=1}^{n} \Gamma_{i,j,\ell}\omega_\ell \;=\; \sum_{(r,s)\in\overline{\mathcal{S}}(n,\mu,3;q)} \eta_{(r,s)}\mathcal{G}_{(r,s),(i,j)} \;=\; \sum_{(r,s)\in\overline{\mathcal{S}}(n,\mu,3;q)} \eta_{(r,s)}(\alpha_i^\perp)^{q^r}(\beta_j^\perp)^{q^s} \;, \quad 1 \le i,j \le n \;.$$

(This equality is, in fact, a generalization of (4); it is shown in [26] that the elements δ_j in (4) can be taken as $(\beta_j^\perp)^{q^\mu}$.)

It follows from the previous discussion that every nonvacuous tensor code $\mathcal{C}(n,\mu,3;q)$ contains the tensor $\Lambda = [\,\Lambda_{i,j,\ell}\,]_{i,j,\ell=1}^{n}$ which is defined by

$$\sum_{\ell=1}^{n} \Lambda_{i,j,\ell}\omega_\ell \;=\; \mathcal{G}_{(0,0),(i,j)} \;=\; \alpha_i^\perp\beta_j^\perp \;, \quad 1 \le i,j \le n \;.$$

This tensor represents the set of bilinear forms that corresponds to multiplication of elements in $GF(q^n)$ over $GF(q)$, which is equivalent to multiplying two polynomials of degree $< n$ modulo an irreducible polynomial of degree n over $GF(q)$. More specifically, let x and y be elements of $GF(q^n)$ and let $x = \sum_{i=1}^{n} x_i\alpha_i^\perp$ and $y = \sum_{j=1}^{n} y_j\beta_j^\perp$ be their representations with respect to the bases $\boldsymbol{\alpha}^\perp$ and $\boldsymbol{\beta}^\perp$, respectively, where $x_i, y_j \in GF(q)$. Then the representation of their product $z = xy$ with respect to the basis $\boldsymbol{\omega}$ satisfies

$$\sum_{\ell=1}^{n} z_\ell\omega_\ell \;=\; z \;=\; xy \;=\; \sum_{i,j=1}^{n} x_iy_j\alpha_i^\perp\beta_j^\perp \;=\; \sum_{i,j=1}^{n} x_iy_j \sum_{\ell=1}^{n} \Lambda_{i,j,\ell}\omega_\ell \;=\; \sum_{\ell=1}^{n} \omega_\ell\Big(\sum_{i,j=1}^{n} x_i\Lambda_{i,j,\ell}y_j\Big) \;,$$

i.e., the coefficients z_ℓ of z are obtained by the following bilinear forms:

$$z_\ell \;=\; \sum_{i,j=1}^{n} x_i\Lambda_{i,j,\ell}y_j \;, \quad \ell = 1,2,\ldots,n \;.$$

Hence, if $\mathcal{C}(n,\mu,3;q)$ is nonvacuous, then μ is a lower bound on the rank of Λ. For $q=2$, this yields the known lower bound of $3.52n - o(n)$ on this rank [4]. On the other hand, it is known that the rank of Λ is linear in n [6], so there is no hope that the true minimum rank of $\mathcal{C}(n,\mu,3;q)$ be quadratic. Yet, with some generalization of the construction, we can obtain a family of $n^{\times 3}$ tensor codes that attains a Gilbert-Varshamov-type bound, as we show next.

Let $\boldsymbol{\theta} = [\,\theta_{(i,j)}\,]_{(i,j)}$ be a vector of length n^2 over $GF(q^n)$ such that none of its entries is zero. Let $K = n^2 - |\mathcal{S}(n,\mu,3;q)|$, and define the $K \times n^2$ matrix $\mathcal{G}(\boldsymbol{\theta})$ over $GF(q^n)$ by

$$(\mathcal{G}(\boldsymbol{\theta}))_{(r,s),(i,j)} \;=\; \theta_{(i,j)}(\alpha_i^\perp)^{q^r}(\beta_j^\perp)^{q^s} \;,$$

where the indexes range as in (21). We fix $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, as well as K coordinates in $\boldsymbol{\theta}$ that correspond to independent columns in $\mathcal{G}$, and let each of the rest of the entries in $\boldsymbol{\theta}$ range over the nonzero values of $GF(q^n)$, thus forming an ensemble of $(q^n-1)^{n^2-K}$ tensor codes $\{\mathcal{C}(\boldsymbol{\theta})\}$, each code having dimension $k = nK$ over $GF(q)$. Now, every $n^{\times 3}$ nonzero tensor Γ (represented by a nonzero vector $\boldsymbol{\gamma} \in (GF(q^n))^{n^2}$) is contained in no more than one of the codes $\mathcal{C}(\boldsymbol{\theta})$. Suppose that μ' is such that $(q^n-1)^{n^2-K} > \mathcal{R}(n, \mu'-1, 3; q) - 1$, where $\mathcal{R}(n, \rho, 3; q)$ is the number of tensors over $GF(q)$ (including the zero tensor) with rank $\leq \rho$. Then there exists an $[n^{\times 3}, k = nK]$ tensor code $\mathcal{C}(\boldsymbol{\theta})$ in the ensemble with minimum rank $\geq \mu'$. Observing that $\mathcal{R}(n, \mu'-1, 3; q) \leq (q^n-1)^{3(\mu'-1)}$ for every $\mu' \geq 3$, it suffices to have

$$n^3 - k \;=\; (n^2 - K)\, n \;\geq\; 3\,(\mu' - 1)\, n \tag{22}$$

in order to guarantee that one of the tensor codes $\mathcal{C}(\boldsymbol{\theta})$ will have minimum rank $\geq \mu'$. This Gilbert-Varshamov-type bound coincides with the bound (11), except that here we take an ensemble which is much smaller than the set of all linear tensor codes.

3.3 The general tensor case

We now extend the constructions $\mathcal{C}(n, \mu, 2; q)$ and $\mathcal{C}(n, \mu, 3; q)$ to any $\Delta \geq 2$. Define the sets $\mathcal{S}(n, \mu, \Delta; q)$ inductively as follows. Let $\mathcal{S}(n, \mu, 2; q)$ consist of all integers r in the range $0 \leq r < \min\{n, \mu-1\}$, and, for $\Delta \geq 3$, let $\mathcal{S}(n, \mu, \Delta; q)$ consist of all integer $(\Delta-1)$-tuples $(r_1, r_2, \ldots, r_{\Delta-1})$ such that —

(a) $0 \leq r_m < n$ for all $m = 1, 2, \ldots, \Delta-1$, and —

(b) $(r_2, r_3, \ldots, r_{\Delta-1}) \in \mathcal{S}(n, d+1, \Delta-1; q)$, where d is such that there exists a conventional linear code over $GF(q)$ of length $\mu-1$, dimension r_1+1, and minimum Hamming distance at least equal to d.

Using the Singleton bound on the minimum Hamming distance of conventional linear codes, it can be easily shown by induction on Δ that the sum of entries of each $(\Delta-1)$-tuple in $\mathcal{S}(n, \mu, \Delta; q)$ is bounded from above by $\mu-2$. Hence, we have

$$|\mathcal{S}(n, \mu, \Delta; q)| \;\leq\; \binom{\mu + \Delta - 3}{\Delta - 1}.$$

This bound is tight if $\mu \leq \min\{n+1, q+2\}$.

For $m = 1, 2, \ldots, \Delta$, let $\boldsymbol{\alpha}_m$ be a vector of length n over $GF(q^n)$ whose entries form a basis of $GF(q^n)$ over $GF(q)$. We define the tensor code $\mathcal{C}(n, \mu, \Delta; q)$ as the set of all $n^{\times \Delta}$ tensors Γ such that

$$\Big\langle \Gamma \,,\, \bigotimes_{m=1}^{\Delta} \boldsymbol{\alpha}_m^{q^{r_m}} \Big\rangle \;=\; 0\,, \quad \text{where} \quad (r_1, r_2, \ldots, r_{\Delta-1}) \in \mathcal{S}(n, \mu, \Delta; q) \quad \text{and} \quad r_\Delta = 0\,. \tag{23}$$

Generalizing Theorem 3, it can be easily shown that the tensor code $\mathcal{C}(n,\mu,\Delta;q)$ has redundancy

$$n^{\Delta} - k \;=\; |\mathcal{S}(n,\mu,\Delta;q)|\, n \;.$$

Therefore, $\mathcal{C}(n,\mu,\Delta;q)$ is nonvacuous if $\mu \leq (\Delta{-}1)(n{-}1)+1$.

Theorem 5. *The minimum rank of $\mathcal{C}(n,\mu,\Delta;q)$ is at least μ.*

Proof. The proof is very similar to that of Theorem 4 and is carried out by double induction on Δ and μ. Let $\Gamma \in \mathcal{C}(n,\mu,\Delta;q)$ and suppose that $\mathrm{rank}(\Gamma) < \mu$. Then we can write

$$\Gamma \;=\; \sum_{t=1}^{\mu-1} \Big(\bigotimes_{m=1}^{\Delta} \mathbf{a}_{t,m} \Big) \;,$$

where $\mathbf{a}_{t,m}$ are vectors over $GF(q)$. Following the proof of Theorem 4, we define $A_{t,m} = \langle \mathbf{a}_{t,m}, \boldsymbol{\alpha}_m \rangle$ and we have

$$\Big\langle \Gamma \;,\; \bigotimes_{m=1}^{\Delta} \boldsymbol{\alpha}_m^{q^{r_m}} \Big\rangle \;=\; \sum_{t=1}^{\mu-1} \prod_{m=1}^{\Delta} A_{t,m}^{q^{r_m}} \;.$$

Hence, Equation (23) is equivalent to

$$\sum_{t=1}^{\mu-1} \prod_{m=1}^{\Delta} A_{t,m}^{q^{r_m}} \;=\; 0 \;, \tag{24}$$

where $(r_1, r_2, \ldots, r_{\Delta-1})$ ranges over all elements in $\mathcal{S}(n,\mu,\Delta;q)$. It remains to show that (24) implies $\Gamma = \sum_{t=1}^{\mu-1} (\otimes_{m=1}^{\Delta} \mathbf{a}_{t,m}) = 0$.

Let ρ denote the dimension of the linear space over $GF(q)$ which is spanned by the elements $\{A_{t,1}\}_{t=1}^{\mu-1}$ and let $\mathbf{u} = [\, u_1\, u_2\, \ldots\, u_{\mu-1}\,]$ be a nonzero vector over $GF(q)$ of minimum Hamming weight in the linear span of the rows of $\mathcal{A} = \Big[\, A_{t,1}^{q^r} \,\Big]_{r=0,t=1}^{\rho-1,\mu-1}$. By (24) and the inductive definition of $\mathcal{S}(n,\mu,\Delta;q)$ we obtain

$$\sum_{t=1}^{\mu-1} u_t \prod_{m=2}^{\Delta} A_{t,m}^{q^{r_m}} \;=\; 0 \quad \text{for every} \quad (r_2, r_3, \ldots, r_{\Delta-1}) \in \mathcal{S}(n, d{+}1, \Delta{-}1; q) \;,$$

where d is the Hamming weight of $\mathbf{u}$. It thus follows that the tensor $\Gamma' = \sum_{t=1}^{\mu-1} u_t (\bigotimes_{m=2}^{\Delta} \mathbf{a}_{t,m})$ is in $\mathcal{C}(n, d{+}1, \Delta{-}1; q)$, while its rank is bounded from above by d. Hence, by the induction hypothesis we must have $\Gamma' = 0$. We now continue as in the proof of Theorem 4. $\square$

As in the case $\Delta = 3$, the codes $\mathcal{C}(n,\mu,\Delta;q)$ attain the Singleton-type bound on the minimum rank when $\mu = 2$. For $\mu = 3$ we get redundancy which is close to the bound of Theorem 1 up to an additive gap of $(\Delta-1)\log_q(q-1)$. The Gilbert-Varshamov-type bound (22), when generalized to larger Δ, yields the sufficient condition $n^{\Delta} - k \;\geq\; \Delta(\mu'{-}1)n$ for the existence of a μ'-$[n^{\times\Delta}, k]$ tensor code (compare with (11)).

4 Application to crisscross error correction

The application of matrix rank to crisscross error correction is described in detail in [11] and [26]. In this section, we point out the advantage of applying the rank metric for crisscross error correction in tensors, even though the Singleton bound on the minimum rank cannot usually be attained when $\Delta \geq 3$. We will concentrate on the case $\Delta = 3$.

Let $\Gamma = [\Gamma_{i,j,\ell}]_{i,j,\ell=1}^{n}$ be an $n^{\times 3}$ tensor over $GF(q)$. A *line* in Γ is a set of n entries in Γ which are indexed by triples (i, j, ℓ), in which two out of the three indexes i, j, and ℓ are fixed, whereas the third index ranges over the integers between 1 and n. In other words, lines in tensors are generalizations of rows and columns in matrices [5, Ch. 1]. By one crisscross error we mean a line in Γ that got corrupted.

A *(line) cover* of a tensor Γ is a set of lines in Γ that contain all its nonzero entries. A cover weight of a tensor Γ is the size of a smallest cover of Γ. The cover distance between two tensors is the cover weight of their difference. The minimum cover distance of a tensor code is the smallest among the cover distances between any two distinct tensors in the code. Since we deal here with linear tensor codes, the minimum cover distance is the minimum cover weight of any nonzero tensor in the code. An $[n^{\times\Delta}, k]$ tensor code with minimum cover distance d will be called an $[n^{\times\Delta}, k, d]$ code. It is easy to check that cover distance is a metric. Therefore, an $[n^{\times\Delta}, k, d]$ tensor code can correct any pattern of up to $(d-1)/2$ crisscross errors. Furthermore, the cover weight of a tensor is bounded from below by its rank. Hence, every μ-$[n^{\times\Delta}, k]$ tensor code is also an $[n^{\times\Delta}, k, \mu]$ code.

We also mention here a generalization of the notion of term-rank for tensors. We say that entries in a tensor Γ are colinear if they lie on the same line in Γ. The term-rank of Γ is the largest number of nonzero entries in Γ that exist such that no two of them are colinear. Clearly, the cover weight of Γ is at least its term-rank, and, in case of matrices, these two numbers are actually equal [5, p. 6]. Such equality, however, does not always hold in the tensor case, as illustrated by the $2 \times 2 \times 3$ tensor Γ shown in Figure 1 (one can extend Γ by zero entries to form a cubic $3 \times 3 \times 3$ tensor). The entries a_i, $i = 1, 2, \ldots, 7$, are nonzero and no three of them are colinear. Therefore, we must have (at least) four lines to cover all the nonzero entries in Γ. On the other hand, consider the cycle $a_1 \to a_2 \to \ldots \to a_7 \to a_1$ that runs through all the nonzero entries in Γ: any two adjacent entries on the cycle are colinear, and, so, there can be found at most three nonzero entries in Γ such that no two of them are colinear.

$$\Gamma \;=\; \begin{array}{|ccc|}\hline 0 & a_1 & a_2 \\ \hline 0 & 0 & a_3 \\ a_6 & a_7 & 0 \\ \hline a_5 & 0 & a_4 \\ \hline\end{array}$$

Figure 1: Tensor of term-rank 3 and cover weight 4.

The Singleton bound on the cover distance takes the form

$$n^3 - k \;\geq\; (d-1)\, n \;,$$

and sphere-packing arguments show that this bound cannot be attained for certain values of q and n, e.g., $q = 2$ and $n < 8$ [26]. The latter reference also contains a Gilbert-Varshamov-type bound that guarantees the existence of an $[n^{\times 3}, k, d]$ tensor code over $GF(q)$ whenever

$$n^3 - k \;\geq\; (d-1)\, n\, (1 + \epsilon(n)) \;, \tag{25}$$

where $\lim_{n\to\infty} \epsilon(n) = 0$. Yet, the proof is nonconstructive.

Probably the simplest constructive technique to combat crisscross errors is the *skewing method*, by which we assign codewords of a conventional linear $[n^2, K, d]$ code over $GF(q)$ to n *wrapped-around hyper-diagonals* in Γ, where the sth hyper-diagonal consists of all entries $\Gamma_{i,j,\ell}$ that are indexed by $\{(i,j,\ell) \;:\; i+j+\ell \equiv s \pmod n\}$. It is easy to see that any line hits each such hyper-diagonal in exactly one entry. Hence, this scheme yields an $[n^{\times 3}, k = nK, d]$ tensor code. If $n^2 \leq q+1$, then by assigning codewords of extended Reed-Solomon codes to each hyper-diagonal, we attain the Singleton bound on the cover distance. Similarly, this bound can be attained when $d = 2$ or $K = 1$. As an example, when $q = n = 2$, we attain this bound for $d \in \{1, 2, 4\}$, whereas an exhaustive search has shown that there is no $[2{\times}2{\times}2, k, 3]$ tensor code over $GF(2)$ with $k > 2$. It is also worth noting that an exhaustive search has shown that there are $[2{\times}2{\times}2, 4, 3]$ tensor codes over $GF(3)$ which are inequivalent to any code obtained by the skewing method (note that by Lemma 1, we cannot have a 3-$[2{\times}2{\times}2, 4]$ tensor code over $GF(3)$).

Consider now an arbitrary q. If we fix d and let n grow, then by the sphere-packing bound for conventional linear codes, the redundancy of a tensor code obtained by the skewing method is bounded from below by an expression which is proportional to $n \log_q n$; namely, the redundancy must be *super-linear* in n. By using BCH codes over $GF(q)$, we can indeed attain redundancy of $2\frac{q-1}{q}(d-1)n \log_q n + O(n)$.

On the other hand, using the tensor code $\mathcal{C}(n, d, 3; q)$, we obtain a coding scheme where the redundancy is $|\mathcal{S}(n, d, 3; q)|\, n \;\leq\; \binom{d}{2} n$. Therefore, even though we do not get linear dependency in d, we do get linear dependency in n of the redundancy, as in (25). A similar savings is obtained also for larger Δ.

5 Decoding

In [10] and [26], efficient decoding algorithms are presented for $\mathcal{C}(n, \mu, 2; q)$ that can recover any error array of rank up to $(\mu-1)/2$. Those algorithms involve a computation of a so-called error-span polynomial whose roots form a linear space which is spanned by the columns of the error array. The degree of that polynomial is q^ρ, where ρ is the rank of the error array.

Hence, once we compute the error-span polynomial, we easily obtain the rank of the error array.

Since computing tensor rank is an intractable problem, it is unlikely that we will have an efficient decoding algorithm which computes an analog of the error-span polynomial for the error tensor with such a simple relationship between the degree and the rank of that tensor: otherwise, we could use the decoder to compute the rank of any tensor. Hence, if there is any efficient decoding algorithm for $\mathcal{C}(n, \mu, 3; q)$, then we expect such an algorithm to recover the error tensor without necessarily obtaining its rank. Such an algorithm that can handle any prescribed number of errors is not yet known.

Nevertheless, efficient decoding procedures can be obtained for the special cases of one-error and two-error correction. We describe such algorithms in Sections 5.1 and 5.2 below.

5.1 Decoding rank-one error tensors

We describe next an easy procedure for decoding a rank-one error tensor while using the code $\mathcal{C}(n, 3, \Delta; q)$. We consider here the case $\Delta = 3$; the case of larger Δ follows along the same lines.

Let Γ be the tensor that has been "transmitted" and let $Y = \Gamma + E$ be the tensor that has been "received", where $E = \mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}$. Since Γ satisfies (12) for $(r, s) \in \{(0,0), (0,1), (1,0)\}$, we can compute syndrome values $S_{(0,0)}$, $S_{(0,1)}$, and $S_{(1,0)}$ for E as follows:

$$S_{(0,0)} = \sum_{i,j,\ell=1}^{n} E_{i,j,\ell} \alpha_i \beta_j \omega_\ell \ , \quad S_{(0,1)} = \sum_{i,j,\ell=1}^{n} E_{i,j,\ell} \alpha_i \beta_j^q \omega_\ell \ , \quad \text{and} \quad S_{(1,0)} = \sum_{i,j,\ell=1}^{n} E_{i,j,\ell} \alpha_i^q \beta_j \omega_\ell \ .$$

Define

$$A \ = \ \langle \mathbf{a}, \boldsymbol{\alpha} \rangle \ , \quad B \ = \ \langle \mathbf{b}, \boldsymbol{\beta} \rangle \ , \quad \text{and} \quad C \ = \ \langle \mathbf{c}, \boldsymbol{\omega} \rangle \ . \tag{26}$$

Then,

$$S_{(0,0)} \ = \ ABC \ , \quad S_{(0,1)} \ = \ AB^q C \ , \quad \text{and} \quad S_{(1,0)} \ = \ A^q BC \ . \tag{27}$$

The tensor E is nonzero if and only if all three syndrome values are nonzero. If E is nonzero, we can use (27) and compute the vectors $\mathbf{a}$ and $\mathbf{b}$ (up to scaling by a nonzero element of $GF(q)$) by solving the following homogeneous linear equations in the entries of $\mathbf{a}$ and $\mathbf{b}$:

$$S_{(0,0)} A^q \ = \ S_{(1,0)} A \quad \text{and} \quad S_{(0,0)} B^q \ = \ S_{(0,1)} B \ .$$

The entries of $\mathbf{c}$ are then recovered by the equation $S_{(0,0)} = ABC$.

5.2 Decoding two crisscross errors

We now describe how one can decode a rank-two error tensor while using the code $\mathcal{C}(n, 5, 3; q)$. We then show how the decoding procedure can be simplified for the special case of double

crisscross error correction.

Let Γ be the transmitted tensor and let $Y = \Gamma + E$ be the received tensor, where

$$E \;=\; (\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}) \;+\; (\mathbf{x} \otimes \mathbf{y} \otimes \mathbf{z}) \;. \tag{28}$$

We first compute the following syndrome values

$$S_{(r,s)} \;=\; \sum_{i,j,\ell=1}^{n} E_{i,j,\ell} \alpha_i \beta_j \omega_\ell \;, \quad r, s \geq 0 \;, \quad r + s \leq 3 \;,$$

with the exception of $q = 2$: since $\mathcal{K}(4, 3; 2) = 1$ (see (15)), the syndrome value $S_{(1,2)}$ will not be available in the binary case.

Define A, B, and C as in (26), and let X, Y, and Z be given by

$$X \;=\; \langle \mathbf{x}, \boldsymbol{\alpha} \rangle \;, \quad Y \;=\; \langle \mathbf{y}, \boldsymbol{\beta} \rangle \;, \quad \text{and} \quad Z \;=\; \langle \mathbf{z}, \boldsymbol{\omega} \rangle \;.$$

Then,

$$S_{(r,s)} \;=\; A^{q^r} B^{q^s} C + X^{q^r} Y^{q^s} Z \;, \quad r, s \geq 0 \;, \quad r + s \leq 3 \;. \tag{29}$$

In particular, we have

$$S_{(r,0)} \;=\; A^{q^r} (BC) \;+\; X^{q^r} (YZ) \;, \quad r = 0, 1, 2, 3 \;,$$

and

$$S_{(0,s)} \;=\; B^{q^s} (AC) \;+\; Y^{q^s} (XZ) \;, \quad s = 0, 1, 2, 3 \;.$$

Applying one of the known decoding algorithms for $\mathcal{C}(n, 5, 2; q)$ with the syndrome values $S_{(r,0)}$, $r = 0, 1, 2, 3$, we obtain a basis of the linear span of $\{A, X\}$ over $GF(q)$. Similarly, the syndrome values $S_{(0,s)}$, $s = 0, 1, 2, 3$, yield a basis of the linear span of $\{B, Y\}$. If any of those linear spans is trivial, then $E = 0$ and we are done. Otherwise, if any of those linear spans has dimension 1, then the decoding problem reduces to that of decoding $\mathcal{C}(n, 5, 2; q)$. For example, if A and X are linearly dependent, then we can assume that $A = X$ and we have $S_{(0,s)} = A(B^{q^s} C + Y^{q^s} Z)$ for $s = 0, 1, 2, 3$. Hence, we can continue solving for B, C, Y, and Z by applying the decoding algorithm for $\mathcal{C}(n, 5, 2; q)$ with the syndrome values $A^{-1} S_{(0,s)}$, $s = 0, 1, 2, 3$.

Assume from now on that the linear span of $\{A, X\}$ has dimension 2 and let $\{U, V\}$ be a basis of this linear span as found by the decoding algorithm for $\mathcal{C}(n, 5, 2; q)$. Without loss of generality we can write

$$A = U + aV \quad \text{and} \quad X = V + xU \;, \quad \text{where} \quad a, x \in GF(q) \;. \tag{30}$$

Plugging (30) into (29) and re-arranging terms, we obtain

$$S_{(r,s)} \;=\; U^{q^r} \left(B^{q^s} C + x Y^{q^s} Z \right) \;+\; V^{q^r} \left(a B^{q^s} C + Y^{q^s} Z \right) \;, \quad r, s \geq 0 \;, \quad r + s \leq 3 \;. \tag{31}$$

For $s = 0, 1$, let Q_s and R_s denote the unique solutions of the following equations

$$S_{(0,s)} \;=\; UQ_s \;+\; VR_s \quad \text{and} \quad S_{(1,s)} \;=\; U^qQ_s \;+\; V^qR_s \;.$$

Indeed, by Lemma 3, Q_s and R_s always exist and they are uniquely determined. From (31) we have

$$Q_s \;=\; B^{q^s}C \;+\; xY^{q^s}Z \;, \quad s = 0, 1 \;, \tag{32}$$

and

$$R_s \;=\; aB^{q^s}C \;+\; Y^{q^s}Z \;, \quad s = 0, 1 \;. \tag{33}$$

Now, suppose that the value of a (and therefore A) is known. Multiplying both sides of (32) by a and subtracting from (33), we have

$$R_s - aQ_s \;=\; (1 - ax)Y^{q^s}Z \;, \quad s = 0, 1 \tag{34}$$

(note that $1 - ax \neq 0$ as we assume that A and X are linearly independent). If $R_s - aQ_s = 0$, then $YZ = 0$ and the problem reduces to that of rank-one error correction. Otherwise, we can obtain from (34) the equality

$$Y^q(R_0 - aQ_0) \;=\; Y(R_1 - aQ_1) \;,$$

which can be solved uniquely for Y, up to scaling by a nonzero element of $GF(q)$. Having found Y, we obtain from (34) the value of $Z' = (1 - ax)Z$.

Let W be an element of $GF(q^n)$ such that $\{W, Y\}$ is a basis of the (already known) linear span of $\{B, Y\}$. We can write $B = W + bY$, where b is an element of $GF(q)$ (which is yet to be found). Substituting $B = W + bY$ into (32) yields

$$Q_s \;=\; W^{q^s}C \;+\; Y^{q^s}(bC + xZ) \;, \quad s = 0, 1 \;. \tag{35}$$

Define $Z'' = bC + xZ = bC + x'Z'$, where $x' = x/(1 - ax)$. By Lemma 3, we can solve (35) uniquely for the values of C and Z''. Now, if the linear span of $\{C, Z'\}$ has dimension less than 2, then the decoding problem reduces again to that of the code $\mathcal{C}(n, 5, 2; q)$. Otherwise, we compute the coefficients b and x' in the decomposition of Z'' into a linear combination of C and Z'. This, in turn, allows us to find the values of B, x, X, and Z.

The decoding procedure we have just outlined assumes that the value of a is known. Therefore, for full decoding, we need to enumerate over all $a \in GF(q)$ to find a solution which is consistent with all ten (nine if $q = 2$) syndrome values. The decoding complexity thus amounts to $O(qn^3)$ arithmetic operations over $GF(q)$, on top of the syndrome computation, which requires $O(n^2)$ operations over $GF(q^n)$. We remark that the outlined algorithm can be extended to handle larger ranks of error tensors (while using the appropriate codes); however, the required number of operations will be proportional to a power of q which will become prohibitively large as the rank of the error tensor grows.

The linear dependency on q of the decoding complexity can be eliminated in the special case of crisscross error correction. In this case, there must be a decomposition of E as in (28) such that $\{\mathbf{a}, \mathbf{b}\}$ contains at least one unit vector, and so does $\{\mathbf{x}, \mathbf{y}\}$. Hence, we proceed as follows. First, we iterate the decoding algorithm for all possible assignments of the basis element $U = \langle \mathbf{u}, \boldsymbol{\alpha} \rangle$ such that $\mathbf{u}$ is a unit vector; there are at most two such assignments, since the linear span of $\{A, X\}$ has dimension 2 or less. In each such iteration, we set $a = 0$, thus forcing the vector $\mathbf{a}$ (i.e., the value A) to be the unit vector $\mathbf{u}$. In case we end up with a consistent error tensor for at least one of those assignments of U, then we are done. Otherwise, we conclude that neither of the vectors $\mathbf{a}$ and $\mathbf{x}$ is a unit vector, in which case there must be an assignment for both $\mathbf{b}$ and $\mathbf{y}$ as unit vectors. Hence, we switch between the roles of (A, X) and (B, Y) and iterate the decoding algorithm a third time, now forcing both $\mathbf{b}$ and $\mathbf{y}$ to be unit vectors that can be computed out of the known linear span of $\{B, Y\}$. Therefore, we will need no more than three iterations of the decoding algorithm, and no enumeration on the value of a will be required. Also, the decoding steps in the algorithm can be simplified when a is zero; e.g., we can solve for Y and Z directly from (33). For the special case of $q = 2$, we can do even better: in the binary case we have either $a = 0$ or $x = 0$, which calls for only two simplified iterations of the original algorithm.

6 The infinite-field case

The dependency of the bounds and constructions of array codes on the structure of the underlying field has already been pointed out in [26]. Therefore, it is not too surprising that such dependency exists for larger Δ as well. As this work is motivated by applications where the underlying field is finite, we will not pursue the discussion on infinite fields here beyond some comments and examples.

The construction $\mathcal{C}(n, \mu, \Delta; q)$ makes use of the fact that the field $GF(q)$ has an (algebraic) extension field of degree n, namely, the field $GF(q^n)$. Therefore, we can try to look at other fields that have such field extensions.

We demonstrate this for the $n^{\times 3}$ tensor case, starting with $\mu = 2$ (the case $\mu = 1$ is, of course, trivial for every field). Let F be a field and let Φ be a field extension of degree n of F. We take three vectors, $[\,\alpha_i\,]_{i=1}^n$, $[\,\beta_j\,]_{j=1}^n$, and $[\,\omega_\ell\,]_{\ell=1}^n$, over Φ, each with entries that are linearly independent over F. A 2-$[n^{\times 3}, k = n^3 - n]$ tensor code $\mathcal{C}$ is obtained by taking the set of all tensors $\Gamma = [\,\Gamma_{i,j,\ell}\,]_{i,j,\ell=1}^n$ over F that satisfy the equation

$$\sum_{i,j,\ell=1}^{n} \Gamma_{i,j,\ell} \alpha_i \beta_j \omega_\ell = 0 \; .$$

Indeed, by a simplified version of the proof of Theorem 4 for $\mu = 2$, it can be shown that every nonzero tensor in $\mathcal{C}$ has rank 2 or more (in fact, the proof of the theorem for $\mu = 2$ does

not depend on the characteristic of the field). Taking F and Φ to be the real and complex fields, respectively, we thus obtain a 2-$[2\times2\times2, 6]$ tensor code over the reals.

For larger values of μ, we make use of *conjugates* of basis elements, the same way we incorporated the powers $\alpha_i^{q^r}$ and $\beta_j^{q^s}$ in the definition of $\mathcal{C}(n, \mu, 3; q)$. More specifically, let Φ be an extension field of F of degree n and let $\mathrm{Aut}_F\Phi$ be the group of automorphisms over Φ which are linear over F. Further, assume that $\mathrm{Aut}_F\Phi$ is a cyclic group of size n with a generator $\varphi : \Phi \to \Phi$. We thus have $\mathrm{Aut}_F\Phi = \{\varphi^r\}_{r=0}^{n-1}$, where φ^r stands for r applications of φ and where φ^0 is the identity mapping (see [31, Chapters 1–3]). The conjugate class of an element $x \in \Phi$ is given by $\{\varphi^r(x)\}_{r=0}^{n-1}$.

We now generalize the construction $\mathcal{C}(n, \mu, 3; q)$ to F by defining a tensor code which consists of all tensors $\Gamma = [\,\Gamma_{i,j,\ell}\,]_{i,j,\ell=1}^{n}$ over F such that

$$\sum_{i,j,\ell=1}^{n} \Gamma_{i,j,\ell}\, \varphi^r(\alpha_i)\, \varphi^s(\beta_j)\, \omega_\ell \;=\; 0\;, \tag{36}$$

where $[\,\alpha_i\,]_{i=1}^{n}$, $[\,\beta_j\,]_{j=1}^{n}$, and $[\,\omega_\ell\,]_{\ell=1}^{n}$ are bases over Φ and r and s range over all integers such that $0 \le r, s < n$ and $r+s \le \mu-2$. Making use of the known properties of the automorphism φ, we can adapt the proof of Theorem 4 to show that every nonzero tensor that satisfies (36) must have rank $\ge \mu$. The construction $\mathcal{C}(n, \mu, 3; q)$ becomes a special case where $F = GF(q)$, $\Phi = GF(q^n)$, and $\varphi(x) = x^q$.

When F and Φ are the real and complex fields, respectively, we have $n = 2$ and $\varphi(x)$ is the conventional complex conjugate x^* of x. We thus obtain a 3-$[2\times2\times2, 2]$ tensor code over the reals which consists of all $2\times2\times2$ tensors $[\,\Gamma_{i,j,\ell}\,]_{i,j,\ell}$ that satisfy the following equations over the complex field:

$$\sum_{i,j,\ell=1}^{2} \Gamma_{i,j,\ell}\alpha_i\beta_j\omega_\ell = 0\;, \quad \sum_{i,j,\ell=1}^{2} \Gamma_{i,j,\ell}\alpha_i\beta_j^*\omega_\ell = 0\;, \quad \text{and} \quad \sum_{i,j,\ell=1}^{2} \Gamma_{i,j,\ell}\alpha_i^*\beta_j\omega_\ell = 0\;.$$

Every nonzero tensor that satisfies those equations must have rank 3 or more. In fact, by the upper bound on tensor rank in [15], it follows that the rank of every such nonzero tensor is exactly 3.

As another example, we construct μ-$[n^{\times3}, k]$ tensor codes over the rationals with $n^3 - k \le \binom{\mu}{2} n$ for every integer n such that $n+1$ is a prime p. Let $M_p(\xi)$ denote the irreducible rational polynomial $\sum_{i=0}^{p-1} \xi^i$ in the indeterminate ξ and define the cyclotomic extension field Φ of degree $n = p-1$ as the set of all rational polynomials $a(\xi)$ of degree $< p-1$, where the arithmetic is taken modulo $M_p(\xi)$. Fix a primitive element g in $GF(p)$. The automorphism φ maps an element $a(\xi) = \sum_{i=0}^{p-2} a_i\xi^i$ to the element $\sum_{i=0}^{p-2} a_i\xi^{i\cdot g} \bmod M_p(\xi)$. Setting $\alpha_i = \beta_i = \omega_i = \xi^i$, Equation (36) becomes

$$\sum_{i,j,\ell=1}^{n} \Gamma_{i,j,\ell}\, \xi^{i\cdot g^r + j\cdot g^s + \ell} \;\equiv\; 0 \mod M_p(\xi)\;.$$

The tensor code thus obtained is nonvacuous if and only if $\mu \leq 2n-1$.

Clearly, such techniques do not apply to tensor codes over the real field when $n > 2$ or over algebraically closed fields. Still, for $\mu = 2$ we can obtain $[n^{\times 3}, k = 3n-2]$ tensor codes over such fields (or over any field) by a construction which resembles the skewing method of Section 4, except that we do not wrap around the hyper-diagonals (see [26] for the case $\Delta = 2$). More specifically, for $s = 3, 4, \ldots, 3n$, let Q_s denote the set of index triples (i, j, ℓ) such that $1 \leq i, j, \ell \leq n$ and $i + j + \ell = s$ (unlike the skewing method, the index equality here is not modulo n). The tensor code consists of all tensors $\Gamma = [\Gamma_{i,j,\ell}]_{i,j,\ell=1}^{n}$ such that $\sum_{(i,j,\ell)\in Q_s} \Gamma_{i,j,\ell} = 0$ for every s. There are $3n-2$ values of s for which the sets Q_s are nonempty, and each such set contributes 1 to the redundancy of the code. Hence, the resulting overall redundancy is $3n-2$. It is an easy exercise to verify that every nonzero tensor in the resulting tensor code indeed has rank at least 2.

For a treatment of typical rank of tensors over algebraically closed fields, see [19] and [28].

References

[1] M.D. Atkinson, N.M. Stephens, *On the maximal multiplicative complexity of a family of bilinear forms, Linear Algebra Appl.,* 27 (1979), 1–8.

[2] M. Blaum, R.J. McEliece, *Coding protection for magnetic tapes: a generalization of the Patel-Hong code, IEEE Trans. Inform. Theory,* IT-31 (1985), 690–693.

[3] R.W. Brockett, *On the generic degree of a 3-tensor,* manuscript, Harvard University, 1976.

[4] R.W. Brockett, D. Dobkin, *On the optimal evaluation of a set of bilinear forms, Proc. Fifth ACM Symp. on Theory of Comput.* (1973), 88–95.

[5] R.A. Brualdi, H.J. Ryser, *Combinatorial Matrix Theory,* Cambridge University Press, Cambridge, UK, 1991.

[6] D.V. Chudnovsky, G.V. Chudnovsky, *Algebraic complexity and algebraic curves over finite fields, J. Complexity,* 4 (1988), 285–316.

[7] Ph. Delsarte, *Bilinear forms over a finite field, with applications to coding theory, J. Comb. Th. A,* 25 (1978), 226–241.

[8] S.A. Elkind, D.P. Siewiorek, *Reliability and performance of error-correcting memory and register codes, IEEE Trans. Computers,* C-29 (1980), 920–927.

[9] P.G. Farrell, *A survey of array error control codes, Europ. Trans. Telecomm. Rel. Technol.,* 3 (1992) 441–454.

[10] E.M. Gabidulin, *Theory of codes with maximum rank distance, Probl. Peredach. Inform.,* 21 (1985), 3–16 (in Russian; pp. 1–12 in the English translation).

[11] E.M. Gabidulin, *Optimal array error-correcting codes, Probl. Peredach. Inform.,* 21 (1985), 102–106 (in Russian).

[12] D.Yu. Grigoryev, *Some new bounds on tensor rank,* LOMI preprint E-2-1978, Leningrad, 1978.

[13] J. Håstad, *Tensor rank is NP-complete, J. Algorithms,* 11 (1990), 644-654.

[14] R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis,* Cambridge University Press, Cambridge, UK, 1991.

[15] T.D. Howell, *Global properties of tensor rank, Linear Algebra Appl.,* 22 (1978), 9–23.

[16] D.E. Knuth, *The Art of Computer Programming, Vol. 2: Seminumerical Algorithms,* Second Edition, Addison-Wesley, Reading, Massachusetts, 1981.

[17] A. Lempel, S.S. Winograd, *A new approach to error-correcting codes, IEEE Trans. Inform. Theory,* IT-23 (1977), 503–508.

[18] L. Levine, W. Meyers, *Semiconductor memory reliability with error detecting and correcting codes, Computer,* 9 (Oct. 1976), 43–50.

[19] T. Lickteig, *Typical tensorial rank, Linear Algebra Appl.,* 69 (1985), 95–120.

[20] R. Lidl, H. Niederreiter, *Finite Fields,* Cambridge University Press, Cambridge, UK, 1984.

[21] F.J. MacWilliams, N.J.A. Sloane, *The Theory of Error-Correcting Codes,* North-Holland, Amsterdam, 1977.

[22] W.F. Mikhail, R.W. Bartoldus, R.A. Rutledge, *The reliability of memory with single-error correction, IEEE Trans. Computers,* C-31 (1983), 560–564.

[23] A.M. Patel, S.J. Hong, *Optimal rectangular code for high density magnetic tapes, IBM J. Res. Dev.,* 18 (1974), 579–588.

[24] A.P. Prudnikov, Yu.A. Brychkov, O.I. Marichev, *Integrals and Series, Vol. 1: Elementary Functions,* Gordon and Breach Science Publishers, New York, 1986.

[25] P. Prunsinkiewicz, S. Budkowski, *A double track error-correction code for magnetic tape, IEEE Trans. Computers,* C-25 (1976), 642–645.

[26] R.M. Roth, *Maximum-rank array codes and their application to crisscross error correction, IEEE Trans. Inform. Theory,* IT-37 (1991), 328–336.

[27] C.H. Stapper, H.S. Lee, *Synergistic fault-tolerance for memory chips, IBM Burlington Technical Bulletin,* TR 19.90510 (1990).

[28] V. Strassen, *Vermeidung von Divisionen, J. Reine, Andew. Math.,* 264 (1973), 184–202.

[29] R. Westwick, *Spaces of linear transformations of equal rank, Linear Algebra Appl.,* 5 (1972), 49–64.

[30] S.S. Winograd, *Arithmetic Complexity of Computations*, SIAM, Philadelphia, Pennsylvania, 1980.

[31] D.J. Winter, *The Structure of Fields,* Springer, New York, 1974.

FIGURES

$$\Gamma \;=\; \begin{array}{|ccccc|}\hline & 0 & & a_1 & \multicolumn{1}{c|}{a_2} \\ 0 & & 0 & & a_3 \\ & a_6 & & a_7 & \multicolumn{1}{c|}{0} \\ \hline a_5 & & 0 & & a_4 \\ \hline \end{array}$$

Figure 1: Tensor of term-rank 3 and cover weight 4.

CAPTIONS

Figure 1: Tensor of term-rank 3 and cover weight 4.

RON M. ROTH (M'89) was born in Ramat Gan, Israel, in 1958. He received the B.Sc. degree in computer engineering, the M.Sc. in electrical engineering and the D.Sc. in computer science from Technion — Israel Institute of Technology, Haifa, Israel, in 1980, 1984 and 1988, respectively. Since 1988 he has been with the Computer Science Department at the Technion. During the academic years 1989–91 he was a Visiting Scientist at IBM Research Division, Almaden Research Center, San Jose, California, and, since 1994, he has been a consultant for Hewlett-Packard Company — Israel Science Center, Haifa, Israel. He is currently on a sabbatical leave from Technion, visiting Hewlett-Packard Laboratories, Palo Alto, California.

His research interests include coding theory, information theory, and their application to the theory of complexity.