# Probabilistic Crisscross Error Correction

RON M. ROTH*

October 20, 1995

**Abstract**

The crisscross error model in data arrays is considered, where the corrupted symbols are confined to a prescribed number of rows or columns (or both). Under the additional assumption that the corrupted entries are uniformly distributed over the channel alphabet, a probabilistic coding scheme is presented where the redundancy can get close to one half the redundancy required in minimum-distance decoding of crisscross errors.

**Keywords:** Crisscross errors, Rank metric, Probabilistic decoding.

*Hewlett-Packard Israel Science Center, Technion City, Haifa 32000, ISRAEL

# 1 Introduction

Consider an application where information symbols (such as bits or bytes) are stored in $m \times n$ arrays, with the possibility of some of the symbols recorded erroneously. The error patterns are such that all corrupted symbols are confined to a prescribed number of rows or columns (or both). We refer to such an error model as *crisscross errors*. A crisscross error pattern that is confined to two rows and three columns is shown in Figure 1.

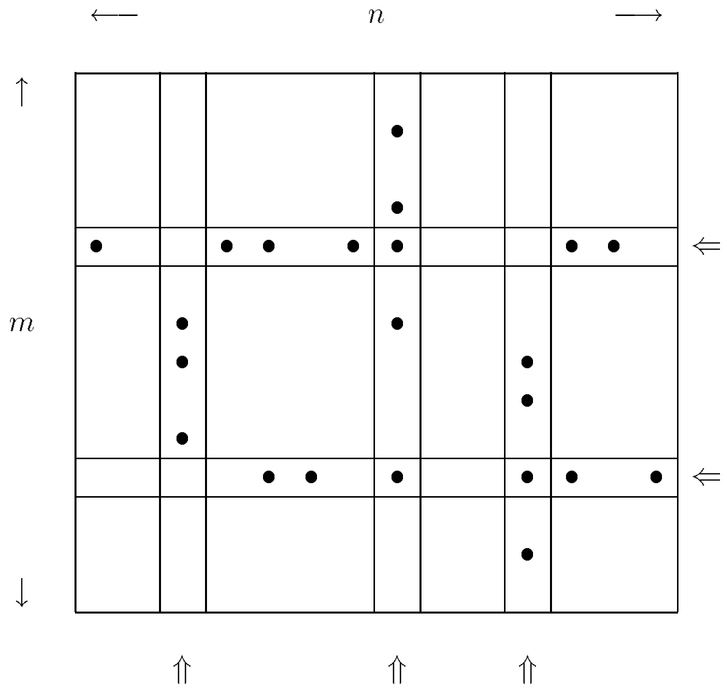


Figure 1: Typical crisscross error pattern.

Crisscross errors can be found in various data storage applications; see, for instance, [2], [4], [5], [8], [10], [11], [12], [13]. Such errors may occur in memory chip arrays, where row or column failures occur due to the malfunctioning of row drivers or column amplifiers. Crisscross errors can also be found in helical tapes, where the tracks are recorded in a direction which is (conceptually) perpendicular to the direction of the movement of the tape; misalignment of the reading head causes burst errors to occur along the track (and across the tape), whereas scratches on the tape usually occur along the tape (and across the tracks). Crisscross error-correcting codes can also be applied in linear magnetic tapes, where the tracks are written along the direction of the movement of the tape and, therefore, scratches cause bursts to occur along the tracks; still, the information and check symbols are usually recorded across the tracks. Computation of check symbols is equivalent to decoding of erasures at the check locations, and in this case these erasures are perpendicular to the erroneous tracks.

Crisscross errors can be analyzed through the following cover metric. A *cover* of an $m \times n$ array $\Gamma$ over a field $F$ is a set of rows or columns that contain all the nonzero entries in $\Gamma$. The *cover weight* of $\Gamma$ is the size of the smallest cover of $\Gamma$. The *cover distance* between two $m \times n$ arrays over $F$ is the cover weight of their difference. An $[m \times n, k, d_{\rm cov}]$ array code over $F$ is a $k$-dimensional linear subspace $\mathcal{C}$ of the vector space of all $m \times n$ matrices over $F$ such that $d_{\rm cov}$ is the smallest cover distance between any two distinct elements of $\mathcal{C}$ or, equivalently, the smallest cover weight of any nonzero element of $\mathcal{C}$. The parameter $d_{\rm cov}$ is referred to as the *minimum cover distance* of $\mathcal{C}$ and the term $mn-k$ stands for the *redundancy* of $\mathcal{C}$.

The Singleton bound on the minimum cover distance states that the minimum cover distance and the redundancy of any $[m \times n, k, d_{\rm cov}]$ array code over a field $F$ satisfy the relation

$$mn - k \;\geq\; (d_{\rm cov} - 1)\, n \;, \tag{1}$$

where we assume that $m \leq n$ (see [7] and [13]).

Let $\Gamma$ be the "transmitted" array and $\Gamma + E$ be the "received" array, where $E$ is the error array. The number of crisscross errors is bounded from below by the cover weight of $E$. Since cover distance is a metric, then by using an $[m \times n, k, d_{\rm cov}]$ array code, we can recover any pattern of up to $(d_{\rm cov}-1)/2$ crisscross errors. On the other hand, if we wish to be able to recover *any* pattern of up to $t$ crisscross errors, then we *must* use an array code with minimum cover distance which is at least $2t+1$. The Singleton bound on the minimum cover distance implies that the number of redundancy symbols must be at least $2tn$, namely, at least twice as large as the maximum number of erroneous symbols that need to be corrected.

In [7] and [13], it was shown how crisscross errors can be handled by applying array codes for the rank metric. A $\mu$-$[m \times n, k]$ array code $\mathcal{C}$ over a field $F$ is a $k$-dimensional linear subspace of the vector space of all $m \times n$ matrices over $F$ such that $\mu$ is the smallest rank of any nonzero matrix in $\mathcal{C}$. The parameter $\mu$ is referred to as the *minimum rank* of $\mathcal{C}$.

The Singleton bound on the minimum rank takes the form

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

where we assume that $m \leq n$. This bound was stated by Delsarte in [3]; see also Gabidulin [6] and Roth [13]. Furthermore, those references contain a construction of $\mu$-$[n \times n, k]$ array codes over the field $\mathsf{F}_q = GF(q)$ that attains this bound for every $\mu \leq n$. We describe next this optimal construction, which we denote by $\mathcal{C}(n \times n, t; q)$, where $\mu = t+1$. Let $\boldsymbol{\alpha} = [\,\alpha_i\,]_{i=1}^n$ be a row vector over $\mathsf{F}_{q^n} = GF(q^n)$ and $\boldsymbol{\omega} = [\,\omega_j\,]_{j=1}^n$ be a column vector over $\mathsf{F}_{q^n}$, each vector having entries that are linearly independent over $\mathsf{F}_q$. The array code $\mathcal{C}(n \times n, t; q)$ consists of all $n \times n$ matrices $\Gamma = [\,\Gamma_{i,j}\,]_{i,j=1}^n$ over $\mathsf{F}_q$ such that

$$\sum_{i,j=1}^{n} \Gamma_{i,j}\, \alpha_i^{q^\ell} \omega_j \;=\; 0 \;, \qquad 0 \leq \ell < t \;. \tag{3}$$

Two efficient decoding algorithms for $\mathcal{C}(n \times n, t; q)$ are presented in [6] and [13] for recovering any error pattern of rank $\leq t/2$. The construction $\mathcal{C}(n \times n, t; q)$ can be generalized to obtain optimal $(t+1)$-$[m \times n, k]$ array codes by means of code shortening. Namely, to form a $(t+1)$-$[m \times n, (m-t)n]$ array code for $m \leq n$, we take the $m \times n$ upper blocks of all the elements $\Gamma$ in a $(t+1)$-$[n \times n, (n-t)n]$ array code such that the last $n-m$ rows in each $\Gamma$ are zero.

The application of $\mu$-$[m \times n, k]$ array codes to crisscross error correction is based upon the observation that matrix rank is a metric and that the cover weight of an array is bounded from below by its rank. By using the elements of a $\mu$-$[m \times n, k]$ array code for transmission (or recording), we can recover any error array of rank $\leq (\mu-1)/2$ and, therefore, we can correct any pattern of up to $(\mu-1)/2$ crisscross errors. Thus, every $\mu$-$[m \times n, k]$ array code is also an $[m \times n, k, \mu]$ array code. The array codes defined by (3) are optimal with respect to the bound (2) and, as such, they are optimal with respect to (1).

Still, such an optimality criterion is based upon a worst-case decoding strategy where we are interested in being able to decode *any* pattern of up to $t$ crisscross errors, thus requiring to have at least $2tn$ redundancy symbols. The purpose of this work is to show that, by assuming a uniform distribution on the error values in each error location, and by allowing an acceptable probability of miscorrection, significant savings in the redundancy can be obtained. More specifically, we assume that the noisy channel acts on the transmitted array $\Gamma$ over $\mathsf{F}_q = GF(q)$ as follows:

**(P1)** The noisy channel selects a set $\mathsf{X}_{\mathrm{r}}$ of row indexes and a set $\mathsf{X}_{\mathrm{c}}$ of column indexes such that $|\mathsf{X}_{\mathrm{r}}| + |\mathsf{X}_{\mathrm{c}}| \leq t$. Note that no assumption is made on the selection of $\mathsf{X}_{\mathrm{r}}$ and $\mathsf{X}_{\mathrm{c}}$ other than limiting the sum of their sizes to be at most $t$. In particular, we do not assume any a priori probabilistic behavior on such a selection.

**(P2)** The channel marks entries in $\Gamma$ within the rows and columns that were selected in (P1). Again, except to their confinement to the rows and columns indexed by $\mathsf{X}_{\mathrm{r}}$ and $\mathsf{X}_{\mathrm{c}}$, no a priori assumption is made on the location of the entries that are marked.

**(P3)** Each marked entry in $\Gamma$ is set to a value which is uniformly distributed over $\mathsf{F}_q$ independently of the values chosen for the other marked entries in $\Gamma$. (In particular, a marked entry may still maintain the correct value with probability $1/q$.)

This probabilistic model of errors seems to approximate rather well the situation in reality, where crisscross errors are caused mainly by bursts. Bursts tend to overwrite the original data in $\Gamma$ independently of that data. In some cases, however, the retrieved bursty stream, which appears as rows or columns in $\Gamma + E$, might have a relatively small number of typical patterns (e.g., tendency of the patterns to contain runs of the same symbol). In such cases, we make the array $\Gamma$ appear random by the use of scramblers, thus forcing the error array $E$ to look random.

# 2 Code construction

In this section, we describe a construction of an $[n \times n, k]$ array code $\mathsf{C}(n \times n, t, \mathsf{p}; q)$ over $\mathsf{F}_q$ that can correct patterns of up to $t$ crisscross errors with a decoding failure probability which is bounded from above by $\mathsf{p}$. By this probability of failure we mean that for any selection of marked entries according to conditions (P1) and (P2), the probability that the values that were chosen according to (P3) resulted in an uncorrectable error array is at most $\mathsf{p}$. Due to the allowable miscorrection probability, the redundancy of these codes can get close to $tn$, namely one half the redundancy of $\mathcal{C}(n \times n, 2t; q)$. Note that the model allows to have up to $tn$ erroneous symbols in the worst case. Therefore, we must have at least as many as $tn$ redundancy symbols when $\mathsf{p} < 1 - (1/q)$. For the sake of simplicity we deal here with constructions of $[m \times n, k]$ array codes where $m = n$. The general case can be handled by code shortening.

Let $C_1$ and $C_2$ be two $[n, n-r, d]$ codes over $\mathsf{F}_q$ with $r \times n$ parity-check matrices $H_1$ and $H_2$, respectively. The specific value of $d$ (and $r$) will be set later on.

We define $\mathsf{C}(n \times n, t, \mathsf{p}; q)$ as an $[n \times n, k]$ array code over $\mathsf{F}_q$ consisting of all $n \times n$ arrays $\Gamma$ such that

$$\Gamma \in \mathcal{C}(n \times n, t; q) \qquad \text{and} \qquad \hat{\Gamma} = H_1 \Gamma H_2^T \in \mathcal{C}(r \times r, 2t; q) \ .$$

The redundancy $n^2 - k$ of $\mathsf{C}(n \times n, t, \mathsf{p}; q)$ is bounded from above by $tn + 2tr$.

We turn now to analyzing the correction capabilities of $\mathsf{C}(n \times n, t, \mathsf{p}; q)$. Throughout the sequel, we will make use of the following notations. For a matrix $A$ over $F$, we denote by $\mathsf{span}_{\mathrm{c}}(A)$ the linear space over $F$ which is spanned by the columns of $A$. Similarly, we denote by $\mathsf{span}_{\mathrm{r}}(A)$ the linear space which is spanned by the rows of $A$. The rank of $A$ will be denoted by $\mathsf{rank}(A)$. Clearly, $\mathsf{rank}(A) = \dim \mathsf{span}_{\mathrm{c}}(A) = \dim \mathsf{span}_{\mathrm{r}}(A)$.

Suppose that $\Gamma \in \mathsf{C}(n \times n, t, \mathsf{p}; q)$ is the transmitted array and that $\Gamma + E$ is the received array where $E$ is an $n \times n$ error array over $\mathsf{F}_q$ which was generated by the channel according to conditions (P1)–(P3). Let $\rho = \mathsf{rank}(E)$. We have $\rho \leq t$ and we can write

$$E \ = \ UAD \ , \tag{4}$$

where $U$ is an $n \times \rho$ matrix whose columns form a basis of $\mathsf{span}_{\mathrm{c}}(E)$, $D$ is a $\rho \times n$ matrix whose rows form a basis of $\mathsf{span}_{\mathrm{r}}(E)$, and $A$ is a $\rho \times \rho$ nonsingular matrix which will be referred to as the *intermediate matrix*. Note that the decomposition (4) of $E$ is not unique since $U$ and $D$ could be any bases of $\mathsf{span}_{\mathrm{c}}(E)$ and $\mathsf{span}_{\mathrm{r}}(E)$, respectively, and for any choice of bases we would have a unique intermediate matrix $A$ so that (4) holds.

Let $\hat{E}$ denote the $r \times r$ matrix $H_1 E H_2^T$. Clearly, $\mathsf{rank}(\hat{E}) \leq \mathsf{rank}(E) = \rho \leq t$. Using the decoding algorithm for $\mathcal{C}(r \times r, 2t; q)$ [6],[13] we can recover the array $\hat{E}$. Our goal now is to show how the parameters can be tuned so that we can recover $E$ from $\hat{E}$ and the syndrome values of $E$ computed for $\mathcal{C}(n \times n, t; q)$, with probability of success $\geq 1 - \mathsf{p}$, given our assumption (P3) on the distribution of the entries of $E$.

Denote by $E_c^+$ the $n \times |\mathsf{X}_c|$ submatrix of $E$ consisting of columns which are indexed by $\mathsf{X}_c$, and by $E_c^-$ the $n \times (n - |\mathsf{X}_c|)$ submatrix whose columns are indexed by $\{1, 2, \ldots, n\} - \mathsf{X}_c$. Similarly, the notations $E_r^+$ and $E_r^-$ will stand for submatrices of $E$ consisting of rows which are indexed by $\mathsf{X}_r$ and $\{1, 2, \ldots, n\} - \mathsf{X}_r$, respectively.

## 2.1 Guaranteeing partial bases of the spans of the error array

Our first step is to show how to set up the parameters $d$ and $r$ so that, with an acceptable probability of at least $1 - \mathsf{p}$, we can find an $n \times \sigma$ matrix $U_1$ such that

$$\mathsf{span}_c(E_c^-) \subseteq \mathsf{span}_c(U_1) \subseteq \mathsf{span}_c(E) . \tag{5}$$

Similarly, we find a $\tau \times n$ matrix $D_1$ such that

$$\mathsf{span}_r(E_r^-) \subseteq \mathsf{span}_c(D_1) \subseteq \mathsf{span}_r(E) . \tag{6}$$

Let $\nu$ be a positive integer which we set later on and choose $C_1$ and $C_2$ so that their minimum Hamming distance $d$ is at least $t + \nu$.

**Lemma 1.**

$$\mathsf{Prob}\left\{ \mathsf{rank}(E) = \mathsf{rank}(H_1 E) = \mathsf{rank}(E H_2^T) \right\} \geq 1 - 2q^{2t-\nu} .$$

**Proof.** Let $V$ be a submatrix of $E_c^-$ whose columns form a basis of $\mathsf{span}_c(E_c^-)$. Clearly, the number of columns in $V$ is bounded from above by $|\mathsf{X}_r|$. We denote by $W$ the matrix $[\, E_c^+ \; V \,]$ and by $h$ the number of columns in $W$. We have $h \leq t$ and $\mathsf{span}_c(E) = \mathsf{span}_c(W)$.

Fix $\mathbf{a} = [\, a_j \,]_{j=1}^h$ to be a column vector in $\mathsf{F}_q^h$. Clearly, $W\mathbf{a} = \mathbf{0}$ implies $H_1 W \mathbf{a} = \mathbf{0}$. We bound from above the probability of the event "$W\mathbf{a} \neq \mathbf{0}$ and $H_1 W \mathbf{a} = \mathbf{0}$." Write the $i$th entry of $W\mathbf{a}$ explicitly as $\sum_{j=1}^h W_{i,j} a_j$. We say that such an entry is *marked* if the entry $W_{i,j}$ was marked in (P2) for at least one index $j$ for which $a_j \neq 0$. Obviously, all unmarked entries in $W\mathbf{a}$ are zero. Furthermore, all marked entries in $W\mathbf{a}$ which are not indexed by $\mathsf{X}_r$ are uniformly distributed over $\mathsf{F}_q$. (The entries indexed by $\mathsf{X}_r$ might carry some dependency since $V$ is taken so that its columns are linearly independent; however, the nonzero entries of $V$ are confined to the rows indexed by $\mathsf{X}_r$.)

We distinguish between the following two cases:

*Case 1:* There are less than $\nu$ entries in $W\mathbf{a}$ which are marked. In particular, the Hamming weight of $W\mathbf{a}$ is less than $d$ and, therefore, $W\mathbf{a}$, if nonzero, cannot be a codeword of $C_1$.

*Case 2:* There are at least $\nu$ entries in $W\mathbf{a}$ which are marked. In particular, we have at least $\nu - |\mathsf{X}_\mathrm{r}| \geq \nu - t$ entries in $W\mathbf{a}$ that are uniformly distributed over $\mathsf{F}_q$. Now, $H_1$ is a parity-check matrix of a linear code with a minimum Hamming distance $d$ and, so, every $d-1$ columns in $H_1$ are linearly independent. In particular, assuming $t > 0$, every $\nu - t$ columns in $H_1$ are linearly independent. Therefore, for every $\mathbf{a} \neq \mathbf{0}$,

$$\mathsf{Prob}\Big\{ W\mathbf{a} \neq \mathbf{0} \quad \text{and} \quad H_1 W\mathbf{a} = \mathbf{0} \Big\} \leq \mathsf{Prob}\Big\{ H_1 W\mathbf{a} = \mathbf{0} \Big\} \leq q^{t-\nu} .$$

We now bound from below the probability of the event "for every $\mathbf{a} \in \mathsf{F}_q^h$, $W\mathbf{a} = \mathbf{0}$ if and only if $H_1 W\mathbf{a} = \mathbf{0}$," by bounding from above the probability of the complement event as follows:

$$\begin{aligned} 1 - \mathsf{Prob}\Big\{ \bigcap_{\mathbf{a}\in\mathsf{F}_q^h} \{ W\mathbf{a} = \mathbf{0} \Longleftrightarrow H_1 W\mathbf{a} = \mathbf{0} \} \Big\} &= \mathsf{Prob}\Big\{ \bigcup_{\mathbf{a}\in\mathsf{F}_q^h} \{ W\mathbf{a} \neq \mathbf{0} \Longleftrightarrow H_1 W\mathbf{a} = \mathbf{0} \} \Big\} \\ &\leq \sum_{\mathbf{a}\in\mathsf{F}_q^h} \mathsf{Prob}\Big\{ W\mathbf{a} \neq \mathbf{0} \Longleftrightarrow H_1 W\mathbf{a} = \mathbf{0} \Big\} \\ &= \sum_{\mathbf{a}\in\mathsf{F}_q^h} \mathsf{Prob}\Big\{ W\mathbf{a} \neq \mathbf{0} \quad \text{and} \quad H_1 W\mathbf{a} = \mathbf{0} \Big\} \\ &\leq q^h q^{t-\nu} \leq q^{2t-\nu} . \end{aligned}$$

It thus follows that $\mathsf{span}_\mathrm{c}(W)$ and $\mathsf{span}_\mathrm{c}(H_1 W)$ have the same dimension with probability at least $1 - q^{2t-\nu}$. Hence, this is also a lower bound on the probability that $\mathsf{rank}(E) = \mathsf{rank}(H_1 E)$. Iterating the proof for the row span of $E$, we obtain the desired result. □

**Lemma 2.**

$$\mathsf{Prob}\Big\{ \mathsf{rank}(\hat{E}) = \mathsf{rank}(E) \Big\} \geq 1 - 2q^{2t-\nu} .$$

**Proof.** Recalling that $\hat{E} = H_1 E H_2^T$, we obtain from (4) the equality

$$\hat{E} = \hat{U} A \hat{D} , \tag{7}$$

where

$$\hat{U} = H_1 U \quad \text{and} \quad \hat{D} = D H_2^T . \tag{8}$$

Now, by Lemma 1, both matrices $\hat{U}$ and $\hat{D}$ have full rank (equaling $\rho$) with probability $\geq 1 - 2q^{2t-\nu}$. Therefore, with such probability we have, $\mathsf{rank}(\hat{E}) = \mathsf{rank}(\hat{U}) = \mathsf{rank}(\hat{D}) = \mathsf{rank}(U) = \mathsf{rank}(D) = \mathsf{rank}(E)$. □

We now proceed as follows. Denote by $\mathcal{S}(n,t)$ the set of all column vectors $\mathbf{u} \in \mathsf{F}_q^n$ with Hamming weight $\leq t$. Start by picking a column vector $\mathbf{u}_1 \in \mathcal{S}(n,t)$ such that $H_1\mathbf{u}_1 \in$

$\mathsf{span}_c(\hat{E})$. Iterate this by picking column vectors $\mathbf{u}_\ell \in \mathcal{S}(n,t) - \mathsf{span}_c(\mathbf{u}_1\mathbf{u}_2\ldots\mathbf{u}_{\ell-1})$ such that $H_1\mathbf{u}_\ell \in \mathsf{span}_c(\hat{E})$. Continue in this manner until no more such vectors can be found. The vectors $\mathbf{u}_\ell$ will form the columns of the $n \times \sigma$ matrix $U_1$.

Now, the matrix $E_c^-$ has at most $|\mathsf{X}_r|$ nonzero rows. Therefore, any basis of $\mathsf{span}_c(E_c^-)$ must be entirely contained in $\mathcal{S}(n,t)$. On the other hand, by Lemma 2 it follows that, with probability $\geq 1 - 2q^{2t-\nu}$, any vector $\mathbf{u} \in \mathsf{span}_c(E_c^-)$ satisfies $H_1\mathbf{u} \in \mathsf{span}_c(\hat{E})$. Therefore, with probability $\geq 1-2q^{2t-\nu}$, we will have the first inclusion in (5). It remains to show how the parameter $\nu$ can be set so that, with probability $\geq 1-\mathsf{p}$, we will have $\mathbf{u}_\ell \in \mathsf{span}_c(E)$ for $\ell = 1,2,\ldots,\sigma$. More specifically, we find an upper bound in terms of $\nu$ for the probability of the event "$H_1\mathbf{u}_\ell \in \mathsf{span}_c(\hat{E})$ and $\mathbf{u}_\ell \notin \mathsf{span}_c(E)$."

**Lemma 3.** *For $\ell = 1,2,\ldots,\sigma$,*

$$\mathsf{Prob}\Big\{ H_1\mathbf{u}_\ell \in \mathsf{span}_c(\hat{E}) \quad \textit{and} \quad \mathbf{u}_\ell \notin \mathsf{span}_c(E) \Big\} \;\leq\; q^{2t-\nu} \; .$$

**Proof.** The proof is similar to that of Lemma 1. Suppose that $\mathbf{u}_\ell \in \mathcal{S}(n,t)$ is such that $\mathbf{u}_\ell \notin \mathsf{span}_c(E)$. Let $V$ and $W$ be as in the proof of Lemma 1 and fix $\mathbf{a}$ to be a vector in $\mathsf{F}_q^h$. We distinguish between the following two cases:

*Case 1:* There are less than $\nu$ entries in $W\mathbf{a}$ which are marked. Since the minimum Hamming distance of $C_1$ is $d \geq \nu + t$, we cannot have both $\mathbf{u}_\ell - W\mathbf{a} \in C_1$ and $\mathbf{u}_\ell \neq W\mathbf{a}$.

*Case 2:* There are at least $\nu$ entries in $W\mathbf{a}$ which are marked. Following similar arguments to those given in Case 2 of the proof of Lemma 1, it can be easily shown that the probability of having $H_1\mathbf{u}_\ell = H_1W\mathbf{a}$ is at most $q^{t-\nu}$.

Enumerating over all $\mathbf{a} \in \mathsf{F}_q^h$ yields

$$\begin{aligned} &\mathsf{Prob}\Big\{ H_1\mathbf{u}_\ell \in \mathsf{span}_c(\hat{E}) \quad \text{and} \quad \mathbf{u}_\ell \notin \mathsf{span}_c(E) \Big\} \\ &\quad \leq \; \mathsf{Prob}\Big\{ H_1\mathbf{u}_\ell \in \mathsf{span}_c(H_1W) \quad \text{and} \quad \mathbf{u}_\ell \notin \mathsf{span}_c(W) \Big\} \;\leq\; q^{2t-\nu} \; , \end{aligned}$$

as desired. □

Summing up those probabilities for all $\ell$, we conclude that with probability $\geq 1 - tq^{2t-\tau}$, our iterative procedure will come up with a matrix $U_1$ that satisfies the second inclusion in (5). Applying a similar algorithm for finding a matrix $D_1$, and taking into account that we require the equality $\mathsf{rank}(E) = \mathsf{rank}(\hat{E})$ to hold, we choose $\nu$ so that the upper bound $2(t{+}1)q^{2t-\nu}$ on the overall probability of failure is at most $\mathsf{p}$. Namely, we choose $\nu$ to be

$$\nu \;=\; 2t - \lfloor \log_q(2(t{+}1)/\mathsf{p}) \rfloor \; .$$

The value of $d$ is therefore set to $\nu + t = 3t - \lfloor \log_q(2(t{+}1)/\mathsf{p}) \rfloor$. If $C_1$ and $C_2$ are taken to be (extended) BCH codes over $\mathsf{F}_q$, then the value of $r$ will not exceed $(d{-}1)\lceil \log_q(n{-}1) \rceil$.

## 2.2 Guaranteeing uniqueness of the error array

Our next step is to show that, once $\hat{E}$, $U_1$, and $D_1$ are determined, the error array $E$ is completely specified. Recall that in the proof of Lemma 2, we map a decomposition $UAD$ of $E$ as in (4), into a decomposition $\hat{U}A\hat{D}$ of $\hat{E}$ as in (7) where $\hat{U} = H_1U$ and $\hat{D} = DH_2^T$ (see (8)). Furthermore, we have shown in Lemma 2 that, with probability $\geq 1 - \mathsf{p}$, the decomposition (7)–(8) is a proper one in the sense that the columns of $\hat{U}$, as well as the rows of $\hat{D}$, are linearly independent. It thus follows that, with probability $\geq 1 - \mathsf{p}$, the proper decompositions (4) of an error array $E$ obtained by (P1)–(P3) map *onto* the proper decompositions (7)–(8); furthermore, this mapping preserves the intermediate matrix $A$.

Given $U_1$ and $D_1$, we say that the decomposition (4) is *systematic* if $U_1$ occupies the first $\sigma$ columns of $U$ and $D_1$ occupies the first $\tau$ rows of $D$; namely, we have

$$U = \begin{bmatrix} U_1 \; U_2 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} .$$

We call the decomposition (7) systematic if it is the image of a systematic decomposition of $E$ under the mapping (8), namely, we have

$$\hat{U} = \begin{bmatrix} H_1U_1 & \hat{U}_2 \end{bmatrix} \quad \text{and} \quad \hat{D} = \begin{bmatrix} D_1H_2^T \\ \hat{D}_2 \end{bmatrix} .$$

Once the decoder for $\mathcal{C}(r \times r, 2t; q)$ has recovered $\hat{E}$, we compute $U_1$ and $D_1$ and then find a systematic decomposition (7) of $\hat{E}$. Such a decomposition, in turn, corresponds to a systematic decomposition of $E$ of the form

$$E = \begin{bmatrix} U_1 \; U_2 \end{bmatrix} A \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} , \tag{9}$$

where the matrices $U_1$, $D_1$, and $A$ are known. Now, with probability $\geq 1 - \mathsf{p}$, the $\sigma$ columns of $U_1$ span all but up to $\mathsf{X}_c$ columns of $E$. Therefore, the number of columns in $U_2$ is at most $|\mathsf{X}_c|$. Similarly, the number of rows in $D_2$ is at most $|\mathsf{X}_r|$. It remains to show that every choice of $U_2$ and $D_2$ in (9) yields, with probability $\geq 1 - \mathsf{p}$, the same error array $E$, given that $E$ is consistent with the syndrome values of the received array that are computed for the array code $\mathcal{C}(n \times n, t; q)$.

Indeed, suppose that the same syndrome values can be obtained for two error arrays,

$$E = \begin{bmatrix} U_1 \; U_2 \end{bmatrix} A \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} \quad \text{and} \quad \tilde{E} = \begin{bmatrix} U_1 \; \tilde{U}_2 \end{bmatrix} A \begin{bmatrix} D_1 \\ \tilde{D}_2 \end{bmatrix} .$$

Write

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} ,$$

where $A_{1,1}$ occupies the upper-left $\sigma \times \tau$ block of $A$. We then have

$$E = U_1 A_{1,1} D_1 + U_1 A_{1,2} D_2 + U_2 A_{2,1} D_1 + U_2 A_{2,2} D_2$$

and

$$\tilde{E} = U_1 A_{1,1} D_1 + U_1 A_{1,2} \tilde{D}_2 + \tilde{U}_2 A_{2,1} D_1 + \tilde{U}_2 A_{2,2} \tilde{D}_2 .$$

Subtracting the last two equations, we obtain

$$E - \tilde{E} = \underbrace{(U_1 A_{1,2} + \tilde{U}_2 A_{2,2})(D_2 - \tilde{D}_2)}_{B_1} + \underbrace{(U_2 - \tilde{U}_2)(A_{2,1} D_1 + A_{2,2} D_2)}_{B_2} .$$

Now, the rank of $B_1$ is bounded from above by the number of rows in $D_2 - \tilde{D}_2$ which, in turn, is at most $|\mathsf{X}_\mathrm{r}|$. Similarly, $\mathsf{rank}(B_2)$ is bounded from above by $|\mathsf{X}_\mathrm{c}|$. It follows that the rank of $E - \tilde{E}$ is at most $|\mathsf{X}_\mathrm{r}| + |\mathsf{X}_\mathrm{c}| \le t$. On the other hand, since the syndrome values of $E$ and $\tilde{E}$ for $\mathcal{C}(n \times n, t; q)$ are the same, then $E - \tilde{E}$ is in $\mathcal{C}(n \times n, t; q)$. Hence, we must have $E = \tilde{E}$.

# 3 Decoding

The decoding algorithm consists of the following stages:

**Step 1** Decoding the matrix $\hat{E} = H_1 E H_2^T$ from the syndrome values of $\hat{E}$ which are computed for the code $\mathcal{C}(r \times r; 2t; q)$.

**Step 2** Computing matrices $U_1$ and $D_1$ such that the inclusions (5) and (6) hold.

**Step 3** Computing the matrix $A$ out of a systematic decomposition (7) of $\hat{E}$ and recovering $E$ out of $U_1$, $D_1$, and $A$.

Step 1 can be carried out by one of the decoding algorithms for $\mathcal{C}(r \times r; 2t; q)$ [6], [13]. As for Step 2, we do not know yet of an algorithm better than enumerating over the elements of $\mathcal{S}(n, t)$ to find vectors $\mathbf{u}$ such that $H_1 \mathbf{u} \in \mathsf{span}_\mathrm{c}(\hat{E})$ (and similarly for the rows). Nevertheless, when the crisscross errors are not too short, then there is high probability to have bases of $\mathsf{span}_\mathrm{c}(E_\mathrm{c}^-)$ and $\mathsf{span}_\mathrm{r}(E_\mathrm{r}^-)$ that consist wholly of unit vectors. In such a case, the required search is linear in $n$. We bound from below the probability of having this in the next lemma.

**Lemma 4.** *Suppose that the channel marks in (P2) at least $\eta$ entries in each one of the rows and columns that were selected in (P1). Then,*

$$\mathsf{Prob}\left\{ \mathsf{rank}(E_\mathrm{c}^-) = |\mathsf{X}_\mathrm{r}| \quad \textit{and} \quad \mathsf{rank}(E_\mathrm{r}^-) = |\mathsf{X}_\mathrm{c}| \right\} > 1 - 2q^{t-\eta} .$$

**Proof.** For $i = 1, 2, \ldots, |\mathsf{X}_{\mathrm{r}}|$ we denote by $M_i$ the $i \times (n-|\mathsf{X}_{\mathrm{c}}|)$ submatrix of $E_{\mathrm{c}}^-$ whose rows are the portions within $E_{\mathrm{c}}^-$ of the first $i$ selected rows in (P1). Defining $\mathsf{rank}(M_0) = 0$, we show by induction on $i = 1, 2, \ldots, |\mathsf{X}_{\mathrm{r}}|$ that $\mathsf{rank}(M_i) > \mathsf{rank}(M_{i-1})$ with probability $\geq 1 - q^{(i-1)+|\mathsf{X}_{\mathrm{c}}|-\eta}$.

The first row selected in (P1) contains at least $\eta - |\mathsf{X}_{\mathrm{c}}|$ marked entries within $E_{\mathrm{c}}^-$. Therefore, $M_1$ is nonzero with probability $\geq 1 - q^{|\mathsf{X}_{\mathrm{c}}|-\eta}$.

As for the induction step, suppose without loss of generality that the last $i-1$ columns of $M_{i-1}$ contain a basis of $\mathsf{span}_{\mathrm{c}}(M_{i-1})$. Per our assumption, there are at least $\eta - |\mathsf{X}_{\mathrm{c}}| - (i-1)$ entries that were marked in (P3) within the first $n - |\mathsf{X}_{\mathrm{c}}| - (i-1)$ coordinates of the $i$th row of $M_i$. Let $\mathbf{u}$ denote a column of $M_i$ that contains one of those marked entries as its $i$th coordinate. The probability that $\mathbf{u}$ belongs to the linear span of the last $i-1$ columns of $M_i$ is at most $1/q$. Therefore, $\mathsf{rank}(M_i) = \mathsf{rank}(M_{i-1})$ with probability $\leq q^{(i-1)+|\mathsf{X}_{\mathrm{c}}|-\eta}$.

Hence,

$$\begin{aligned}
\mathsf{Prob}\left\{ \mathsf{rank}(E_{\mathrm{c}}^-) = |\mathsf{X}_{\mathrm{r}}| \right\} &= \mathsf{Prob}\left\{ \mathsf{rank}(M_{|\mathsf{X}_{\mathrm{r}}|}) = |\mathsf{X}_{\mathrm{r}}| \right\} \\
&= \mathsf{Prob}\left\{ \bigcap_{i=1}^{|\mathsf{X}_{\mathrm{r}}|} \{ \mathsf{rank}(M_i) > \mathsf{rank}(M_{i-1}) \} \right\} \\
&\geq 1 - \sum_{i=1}^{|\mathsf{X}_{\mathrm{r}}|} \mathsf{Prob}\left\{ \mathsf{rank}(M_i) = \mathsf{rank}(M_{i-1}) \right\} \\
&\geq 1 - \sum_{i=1}^{|\mathsf{X}_{\mathrm{r}}|} q^{(i-1)+|\mathsf{X}_{\mathrm{c}}|-\eta} \;>\; 1 - q^{t-\eta} .
\end{aligned}$$

Iterating the proof for $E_{\mathrm{r}}^-$, we obtain the desired result. $\square$

We turn now to Step 3. We assume that the error array $E$ can be written as

$$E = \begin{bmatrix} U_1 \; U_2 \end{bmatrix} A \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} , \tag{10}$$

where $A$ is a known nonsingular $\rho \times \rho$ matrix, $U_1$ is a known $n \times \sigma$ matrix with rank $\sigma$, and $D_1$ is a known $\tau \times n$ matrix with rank $\tau$. Our task is to compute the unknown $n \times (\rho - \sigma)$ matrix $U_2$ and the $(\rho - \tau) \times n$ matrix $D_2$ in (10), under the assumption that $t + \tau + \sigma \geq 2\rho$. Indeed, this inequality holds in our case since $\rho - \sigma \leq |\mathsf{X}_{\mathrm{c}}|$, $\rho - \tau \leq |\mathsf{X}_{\mathrm{r}}|$, and $|\mathsf{X}_{\mathrm{c}}| + |\mathsf{X}_{\mathrm{r}}| \leq t$.

Write $A$ in the form

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} , \tag{11}$$

where $A_{1,1}$ occupies the upper-left $\sigma \times \tau$ block of $A$. Let $\gamma$ denote the rank of $A_{2,2}$. We perform elementary operations on the rows and columns of $A$ — and respectively on the columns of $U$ and the rows of $D$ so that (10) still holds — as follows. Applying elementary operations on the last $\rho-\tau$ columns and the last $\rho-\sigma$ rows of $A$, we can get a new matrix $A$ with a $(\rho-\sigma)\times(\rho-\tau)$ submatrix $A_{2,2}$ of the form

$$A_{2,2} = \begin{bmatrix} 0 & 0 \\ 0 & I_\gamma \end{bmatrix} , \tag{12}$$

where $I_\gamma$ stands for a $\gamma\times\gamma$ identity matrix. Next, by row operations that change only the first $\sigma$ rows of $A$, followed by column operations that change only the first $\tau$ columns of $A$, we obtain a $\sigma\times(\rho-\tau)$ submatrix $A_{1,2}$ and a $(\rho-\sigma)\times\tau$ submatrix $A_{2,1}$ of the form

$$A_{1,2} = \begin{bmatrix} 0 & 0 \\ I_{\rho-\tau-\gamma} & 0 \end{bmatrix} \qquad \text{and} \qquad A_{2,1} = \begin{bmatrix} 0 & I_{\rho-\sigma-\gamma} \\ 0 & 0 \end{bmatrix} . \tag{13}$$

A final iteration of row and column operations that affect only the upper-left $\sigma\times\tau$ block of $A$ yields a $\sigma\times\tau$ submatrix $A_{1,1}$ of the form

$$A_{1,1} = \begin{bmatrix} I_{\tau+\sigma-\rho+\gamma} & 0 \\ 0 & 0 \end{bmatrix} . \tag{14}$$

Note that in all those operations, we never add any of the first $\sigma$ rows of $A$ to any of its last $\rho-\sigma$ rows. Therefore, the resulting matrix $U=[\,U_1\ U_2\,]$ is such that we still know the $n\times\sigma$ submatrix $U_1$. A similar rule applies also to the rows of $D$.

Write

$$U_1 = \begin{bmatrix} U_{1,1}\ U_{1,2} \end{bmatrix} \quad , \quad U_2 = \begin{bmatrix} U_{2,1}\ U_{2,2} \end{bmatrix} \quad , \quad D_1 = \begin{bmatrix} D_{1,1} \\ D_{1,2} \end{bmatrix} , \quad \text{and} \quad D_2 = \begin{bmatrix} D_{2,1} \\ D_{2,2} \end{bmatrix} ,$$

where $U_{1,1}$ occupies the first $\tau+\sigma-\rho+\gamma$ columns of $U_1$, $U_{2,1}$ occupies the first $\rho-\sigma-\gamma$ columns of $U_2$, $D_{1,1}$ occupies the first $\tau+\sigma-\rho+\gamma$ rows of $D_1$, and $D_{2,1}$ occupies the first $\rho-\tau-\gamma$ rows of $D_2$. By (10) and the choice of $A$ we thus have

$$E = E_{1,1} + E_{1,2} + E_{2,1} + E_{2,2} ,$$

where

$$E_{a,b} = U_{a,b}D_{b,a} , \qquad a,b\in\{1,2\} .$$

Let $\rho_{a,b}$ denote the rank of $E_{a,b}$. We have

$$\rho_{1,1}=\tau+\sigma-\rho+\gamma , \quad \rho_{1,2}=\rho-\tau-\gamma , \quad \rho_{2,1}=\rho-\sigma-\gamma , \quad \text{and} \quad \rho_{2,2}=\gamma ,$$

and those ranks sum to $\rho$. The matrix $E_{1,1}$, as well as the column span of $E_{1,2}$ and the row span of $E_{2,1}$, are known to the decoder.

We now present a decoding algorithm for recovering the matrix $E'=E-E_{1,1}$, whose rank equals $\rho-\rho_{1,1}=2\rho-\tau-\sigma-\gamma$. The algorithm is a generalization of the decoding procedure described in [13]. In analogy to conventional error-correcting codes, the partial information that we have on the column and row spans of $E_{1,2}$ and $E_{2,1}$ can be viewed as "erasure information," whereas $E_{2,2}$ needs full correction. The ranks of those three unknown matrices satisfy

$$\rho_{1,2} + \rho_{2,1} + 2\rho_{2,2} \;=\; 2\rho - \sigma - \tau \;\le\; t\;.$$

Pursuing the analogy with conventional error correction, we would expect such an inequality to suffice for recovering the unknown matrices while using the array code $\mathcal{C}(n \times n, t; q)$. Indeed, the results of Section 2 imply that this is indeed the case. The algorithm presented in [13] handles the special case where $\sigma = \tau = 0$ (and $t \ge 2\rho$).

Let $E''$ denote the sum $E_{1,2} + E_{2,2}$, which is of rank $\rho_{1,2} + \rho_{2,2} = \rho - \tau$. We first show how we can compute a matrix $U_{2,2}$, consisting of $\gamma$ linearly independent vectors that form together with the columns of $U_{1,2}$ a basis of $\mathsf{span}_c(E'') = \mathsf{span}_c(E_{1,2}\; E_{2,2})$. Such a basis will then allow us to recover the matrices $E''$ and $E_{2,1}$ by solving a set of linear equations. We point out that even though the matrix $E''$ will be uniquely determined, the specific decomposition of $E''$ into a sum $E_{1,2} + E_{2,2}$ will depend upon the particular matrix $U_{2,2}$ selected. Indeed, if the columns of $[\, U_{1,2}\; U_{2,2}\,]$ form a basis of $\mathsf{span}_c(E'')$, then we can always find matrices $D_{2,1}$ and $D_{2,2}$ such that $E'' = U_{1,2}D_{2,1} + U_{2,2}D_{2,2}$. Throughout the sequel we use a terminology which is similar to that in [13].

Since the matrix $E_{1,1}$ is known, the decoder can compute the syndrome vector $\mathsf{s} \in \mathsf{F}_{q^n}^t$ of $E' = E - E_{1,1} = [\, e'_{i,j}\,]_{i,j=1}^n$ for $\mathcal{C}(n \times n, r; q)$ with respect to the equations (3) as follows:

$$s_\ell \;=\; \sum_{i,j=1}^{n} e'_{i,j}\, \alpha_i^{q^\ell} \omega_j \;, \qquad 0 \le \ell < t\;. \tag{15}$$

Denote by $\mathsf{Y}_{a,b}$ the set of indexes of the $\rho_{a,b}$ columns in $U$ that belong to the submatrix $U_{a,b}$; e.g., if indexes start at 1, then $\mathsf{Y}_{2,1} = \{j : \sigma < j \le \rho - \gamma\}$. Note that $\mathsf{Y}_{a,b}$ also points at the rows of $D_{b,a}$ in $D$. We denote by $\mathsf{Y}$ the set $\{j : 0 < j \le \rho\} = \cup_{a,b\in\{1,2\}} \mathsf{Y}_{a,b}$ and by $\mathsf{Y}'$ the set $\mathsf{Y} - \mathsf{Y}_{1,1} = \mathsf{Y}_{1,2} \cup \mathsf{Y}_{2,1} \cup \mathsf{Y}_{2,2}$. Clearly, $|\mathsf{Y}'| = \rho - \rho_{1,1} = 2\rho - \sigma - \tau - \gamma$.

Write $U = [\,(U)_{i,k}\,]_{i,k}$ and $D = [\,(D)_{k,j}\,]_{k,j}$. By (10) and (15) we have

$$s_\ell = \sum_{k\in\mathsf{Y}'} \sum_{i=1}^{n} \alpha_i^{q^\ell} (U)_{i,k} \sum_{j=1}^{n} (D)_{k,j}\omega_j \;, \qquad 0 \le \ell < t\;. \tag{16}$$

Defining the row vector $\boldsymbol{\beta} = [\,\beta_k\,]_{k\in\mathsf{Y}} = \boldsymbol{\alpha} U$ and the column vector $\boldsymbol{\delta} = [\,\delta_k\,]_{k\in\mathsf{Y}} = D\boldsymbol{\omega}$, both in $\mathsf{F}_{q^n}^\rho$, we can rewrite (16) as

$$s_\ell = \sum_{k\in\mathsf{Y}'} \beta_k^{q^\ell} \delta_k \;, \qquad 0 \le \ell < t\;. \tag{17}$$

We now define the polynomials $\Phi(x)$, $\Psi(x)$, and $\Lambda(x)$ over $\mathsf{F}_{q^n}$ as

$$\Phi(x) \;=\; \sum_{m=0}^{\rho_{1,2}} \phi_m x^{q^m} \;=\; \prod_{\mathbf{y}\in\mathsf{span}_c(E_{1,2})} (x - \boldsymbol{\alpha}\mathbf{y})\;,$$

$$\Psi(x) = \sum_{m=0}^{\rho_{2,1}} \psi_m x^{q^m} = \prod_{\mathbf{z} \in \mathsf{span}_{\mathrm{r}}(E_{2,1})} (x - \mathbf{z}\boldsymbol{\omega}) ,$$

and

$$\Lambda(x) = \sum_{m=0}^{\rho-\tau} \lambda_m x^{q^m} = \prod_{\mathbf{u} \in \mathsf{span}_{\mathrm{c}}(E'')} (x - \boldsymbol{\alpha}\mathbf{u}) .$$

Those three polynomials are *linearized* polynomials over $\mathsf{F}_{q^n}$, namely, they have the form $\sum_m a_m x^{q^m}$ where $a_m \in \mathsf{F}_{q^n}$. Several properties of linearized polynomials are summarized in [9, Section 4.9]. In particular, if $a(x) = \sum_m a_m x^{q^m}$ is a linearized polynomial over $\mathsf{F}_{q^n}$, then the mapping $x \mapsto a(x)$ over the domain $\mathsf{F}_{q^n}$ is a linear transformation over $\mathsf{F}_q$. Thus, the set of roots of $a(x)$ in $\mathsf{F}_{q^n}$ forms a linear vector space over $\mathsf{F}_q$, as this set is a null space of a linear transformation.

The polynomial $\Phi(x)$ can be easily computed by the decoder by solving the following set of $\rho_{1,2} = \rho-\tau-\gamma$ linear equations over $\mathsf{F}_{q^n}$ for the coefficients $\phi_m$:

$$\Phi(\beta_k) = \sum_{m=0}^{\rho_{1,2}} \phi_m \beta_k^{q^m} = 0 , \qquad k \in \mathsf{Y}_{1,2} , \tag{18}$$

where $\phi_{\rho_{1,2}} = 1$. Similarly, the polynomial $\Psi(x)$ can be computed by solving the equations

$$\Psi(\delta_k) = \sum_{m=0}^{\rho_{2,1}} \psi_m \delta_k^{q^m} = 0 , \qquad k \in \mathsf{Y}_{2,1} , \tag{19}$$

with $\psi_{\rho_{2,1}} = 1$.

The coefficients of the polynomial $\Lambda(x)$ satisfy the following set of $\rho-\tau-\gamma$ linear equations over $\mathsf{F}_{q^n}$:

$$\Lambda(\beta_k) = \sum_{m=0}^{\rho-\tau} \lambda_m \beta_k^{q^m} = 0 , \qquad k \in \mathsf{Y}_{1,2} .$$

Those equations are linearly independent [9, p. 117]. Yet, when $\gamma > 0$, we will need additional constraints in order to determine $\Lambda(x)$ uniquely.

Let $S$ be the $(t-\rho+\tau) \times (\rho-\tau+1)$ matrix over $\mathsf{F}_{q^n}$ which is defined by

$$S = \begin{bmatrix} s_0^{q^n} & s_1^{q^n} & \cdots & s_{\rho-\tau}^{q^n} \\ s_1^{q^{n-1}} & s_2^{q^{n-1}} & \cdots & s_{\rho-\tau+1}^{q^{n-1}} \\ s_2^{q^{n-2}} & s_3^{q^{n-2}} & \cdots & s_{\rho-\tau+2}^{q^{n-2}} \\ \vdots & \vdots & \vdots & \vdots \\ s_{t-\rho+\tau-1}^{q^{n-(t-\rho+\tau-1)}} & s_{t-\rho+\tau}^{q^{n-(t-\rho+\tau-1)}} & \cdots & s_{t-1}^{q^{n-(t-\rho+\tau-1)}} \end{bmatrix} \tag{20}$$

and let $R$ denote the $(t+\sigma+\tau-2\rho+\gamma) \times (t-\rho+\tau)$ matrix over $\mathsf{F}_{q^n}$ which is given by

$$R = \left[ \begin{array}{ccccccccc} \psi_g^{q^{n-g}} & \psi_{g-1}^{q^{n-g}} & \cdots & \psi_1^{q^{n-g}} & \psi_0^{q^{n-g}} & 0 & \cdots & & \\ 0 & \psi_g^{q^{n-g-1}} & \psi_{g-1}^{q^{n-g-1}} & \cdots & \psi_1^{q^{n-g-1}} & \psi_0^{q^{n-g-1}} & 0 & \cdots & \\ \cdots & 0 & \psi_g^{q^{n-g-2}} & \psi_{g-1}^{q^{n-g-2}} & \cdots & \psi_1^{q^{n-g-2}} & \psi_0^{q^{n-g-2}} & 0 & \cdots \\ & \cdots & 0 & \ddots & & & & \ddots & 0 \\ & & \cdots & 0 & \psi_g^{q^{n-g'}} & \psi_{g-1}^{q^{n-g'}} & \cdots & \psi_1^{q^{n-g'}} & \psi_0^{q^{n-g'}} \end{array} \right], \tag{21}$$

where $g = \rho_{2,1} = \rho-\sigma-\gamma$ and $g' = g+t+\sigma+\tau-2\rho+\gamma-1 = t-\rho+\tau-1$. We denote by $R_L$ the submatrix consisting of the first $L$ rows of $R$.

**Lemma 5.** *Let* $\boldsymbol{\lambda} = [\,\lambda_m\,]_{m=0}^{\rho-\tau}$ *be the column vector of coefficients of the polynomial* $\Lambda(x) = \sum_{m=0}^{\rho-\tau} \lambda_m x^{q^m}$. *Then,*

$$RS\boldsymbol{\lambda} \;=\; \mathbf{0}\,.$$

**Proof.** Let $(S\boldsymbol{\lambda})_\ell$ denote the $\ell$th entry of $S\boldsymbol{\lambda}$. By (17) we have

$$\begin{aligned} (S\boldsymbol{\lambda})_\ell &= \sum_{m=0}^{\rho-\tau} \lambda_m s_{\ell+m}^{q^{n-\ell}} = \sum_{m=0}^{\rho-\tau} \lambda_m \Big( \sum_{k\in\mathsf{Y}'} \beta_k^{q^{\ell+m}} \delta_k \Big)^{q^{n-\ell}} \\ &= \sum_{k\in\mathsf{Y}'} \delta_k^{q^{n-\ell}} \sum_{m=0}^{\rho-\tau} \lambda_m \beta_k^{q^m} = \sum_{k\in\mathsf{Y}'} \delta_k^{q^{n-\ell}} \Lambda(\beta_k) \\ &= \sum_{k\in\mathsf{Y}_{2,1}} \delta_k^{q^{n-\ell}} \Lambda(\beta_k)\,, \qquad 0 \le \ell < t-\rho+\tau\,, \end{aligned} \tag{22}$$

where the last equality follows from having $\Lambda(\beta_k) = 0$ for $k \in \mathsf{Y}_{1,2} \cup \mathsf{Y}_{2,2}$. We thus obtain,

$$\begin{aligned} (RS\boldsymbol{\lambda})_m &= \sum_{\ell=m}^{m+g} \psi_{g+m-\ell}^{q^{n-g-m}} (S\boldsymbol{\lambda})_\ell = \sum_{\ell=0}^{g} \psi_{g-\ell}^{q^{n-g-m}} (S\boldsymbol{\lambda})_{m+\ell} \\ &= \sum_{\ell=0}^{g} \psi_{g-\ell}^{q^{n-g-m}} \sum_{k\in\mathsf{Y}_{2,1}} \delta_k^{q^{n-m-\ell}} \Lambda(\beta_k) \\ &= \sum_{k\in\mathsf{Y}_{2,1}} \Lambda(\beta_k) \Big( \sum_{\ell=0}^{g} \psi_{g-\ell} \delta_k^{q^{g-\ell}} \Big)^{q^{n-g-m}} \\ &= \sum_{k\in\mathsf{Y}_{2,1}} \Lambda(\beta_k) \left(\Psi(\delta_k)\right)^{q^{n-g-m}}\,, \qquad 0 \le m < t-\rho+\tau-g\,. \end{aligned} \tag{23}$$

The lemma now follows by observing that $\Psi(\delta_k) = 0$ for $k \in \mathsf{Y}_{2,1}$. $\square$

**Lemma 6.** *Let* $\mathbf{v} = [\,v_m\,]_{m=0}^{\rho-\tau}$ *be the column vector of coefficients of a linearized monic polynomial* $v(x) = \sum_{m=0}^{\rho-\tau} v_m x^{q^m}$ *such that* $v(\beta_k) = 0$ *for* $k \in \mathsf{Y}_{1,2}$ *and*

$$R_\gamma S\mathbf{v} \;=\; \mathbf{0}\;.$$

*Then* $v(x) = \Lambda(x)$.

**Proof.** Since $v(x)$ is a linearized polynomial, it suffices to show that $v(\beta_k) = 0$ for $k \in \mathsf{Y}_{2,2}$. The equality $R_\gamma S\mathbf{v} = \mathbf{0}$ implies

$$\sum_{\ell=0}^{g} \psi_{g-\ell}^{q^{n-g-m}} (S\mathbf{v})_{m+\ell} = 0\;, \qquad 0 \le m < \gamma$$

(compare with (23)). On the other hand, following (22), we have

$$(S\mathbf{v})_\ell \;=\; \sum_{k\in\mathsf{Y}'} \delta_k^{q^{n-\ell}} v(\beta_k)\;, \qquad 0 \le \ell < t{-}\rho{+}\tau\;.$$

Noting that $g + \gamma \le t - \rho + \tau$, we can combine the last two equations to obtain

$$\sum_{k\in\mathsf{Y}'} v(\beta_k) \Big(\sum_{\ell=0}^{g} \psi_{g-\ell}\delta_k^{q^{g-\ell}}\Big)^{q^{n-g-m}} \;=\; 0\;, \qquad 0 \le m < \gamma\;,$$

namely,

$$\sum_{k\in\mathsf{Y}'} v(\beta_k)\,(\Psi(\delta_k))^{q^{n-g-m}} \;=\; 0\;, \qquad 0 \le m < \gamma\;.$$

By assumption we have $v(\beta_k) = 0$ for $k \in \mathsf{Y}_{1,2}$. Since we also have $\Psi(\delta_k) = 0$ for $k \in \mathsf{Y}_{2,1}$, we end up with

$$\sum_{k\in\mathsf{Y}_{2,2}} v(\beta_k)\,(\Psi(\delta_k))^{q^{n-g-m}} \;=\; 0\;, \qquad 0 \le m < \gamma\;.$$

Now, the elements $\{\Psi(\delta_k)\}_{k\in\mathsf{Y}_{2,2}}$ are linearly independent over $\mathsf{F}_q$, or else $\Psi(x)$ would vanish at a nontrivial linear combination of the elements $\{\delta_k\}_{k\in\mathsf{Y}_{2,2}}$, which is impossible. By [9, p. 117] we conclude that the elements $v(\beta_k)$ must be zero for every $k \in \mathsf{Y}_{2,2}$. $\square$

It follows from Lemmas 5 and 6 that for any integer $L$ in the range $\gamma \le L \le t{+}\sigma{+}\tau{-}2\rho{+}\gamma$, there is a unique monic linearized polynomial $v(x)$ that vanishes at $\beta_k$ for all $k \in \mathsf{Y}_{1,2}$ and satisfies the set of equations

$$R_L S\mathbf{v} \;=\; \mathbf{0}\;.$$

**Lemma 7.** *Let* $v(x) = \sum_{m=0}^{\rho-\tau} v_m x^{q^m}$ *be a linearized polynomial over* $\mathsf{F}_{q^n}$*. Then* $v(\beta_k) = 0$ *for* $k \in \mathsf{Y}_{1,2}$ *if and only if there is a polynomial* $\Theta(x) = \sum_{m=0}^{\gamma} \theta_m x^{q^m}$ *over* $\mathsf{F}_{q^n}$ *such that*

$$v(x) \;=\; \Theta(\Phi(x))\;. \tag{24}$$

**Proof.** Clearly, the polynomial $\Theta(\Phi(x))$ vanishes at $\beta_k$ for all $k \in \mathsf{Y}_{1,2}$. The "only if" part can be proved by using an analog of a long division. Namely, it can be easily verified that for every linearized polynomial $v(x)$ we can find linearized polynomials $\Theta(x)$ and $\Delta(x)$ such that $v(x) = \Theta(\Phi(x)) + \Delta(x)$ and $\deg \Delta < \deg \Phi = q^{\rho-\tau-\gamma}$. Now, if $v(x)$ vanishes at $\beta_k$ for all $k \in \mathsf{Y}_{1,2}$, then so does $\Delta(x)$. This means that $\Delta(x)$ has at least $q^{\rho-\tau-\gamma}$ roots in $\mathsf{F}_{q^n}$, which implies that $\Delta(x) \equiv 0$. $\square$

Let $Q$ denote the $(\gamma+1) \times (\rho-\tau+1)$ matrix over $\mathsf{F}_{q^n}$ which is given by

$$Q = \begin{bmatrix} \phi_0 & \phi_1 & \cdots & \phi_{f-1} & \phi_f & 0 & \cdots & & & & \\ 0 & \phi_0^q & \phi_1^q & \cdots & \phi_{f-1}^q & \phi_f^q & 0 & \cdots & & & \\ \cdots & 0 & \phi_0^{q^2} & \phi_1^{q^2} & \cdots & \phi_{f-1}^{q^2} & \phi_f^{q^2} & 0 & \cdots & & \\ & \cdots & 0 & \ddots & & & & \ddots & 0 & & \\ & & \cdots & 0 & \phi_0^{q^\gamma} & \phi_1^{q^\gamma} & \cdots & \phi_{f-1}^{q^\gamma} & \phi_f^{q^\gamma} & & \end{bmatrix}, \qquad (25)$$

where $f = \rho_{1,2} = \rho-\tau-\gamma$. Then (24) is equivalent to having

$$\mathbf{v} = Q^T \boldsymbol{\theta} ,$$

where $\mathbf{v} = [\, v_m \,]_{m=0}^{\rho-\tau}$ and $\boldsymbol{\theta} = [\, \theta_m \,]_{m=0}^{\gamma}$.

Lemmas 5, 6, and 7 provide the means by which we can compute the matrix $U_{2,2}$. We first find the coefficients of $\Theta(x)$ by solving the set of equations

$$R_\gamma S\, Q^T \boldsymbol{\theta} = \mathbf{0} \qquad (26)$$

for $\boldsymbol{\theta} = [\, \theta_m \,]_{m=0}^{\gamma}$ and $\theta_\gamma = 1$. By those lemmas we have $Q^T \boldsymbol{\theta} = \boldsymbol{\lambda}$. Now, the mapping $x \mapsto \Lambda(x)$ is linear over $\mathsf{F}_q$. Hence, finding a full basis of the linear space (over $\mathsf{F}_q$) of roots of $\Lambda(x)$ in $\mathsf{F}_{q^n}$ is equivalent to finding a basis of the null space of an $n \times n$ matrix over $\mathsf{F}_q$ which represents the mapping $x \mapsto \Lambda(x)$ [1]. We choose $U_{2,2}$ so that the entries of $\boldsymbol{\alpha} U_{2,2}$ extend $\{\beta_k\}_{k \in \mathsf{Y}_{1,2}}$ to span the null space of such a matrix representation. The entries of $\boldsymbol{\alpha} U_{2,2}$ then form the set $\{\beta_k\}_{k \in \mathsf{Y}_{2,2}}$.

At this point, the set of equations (17) have become linear in the remaining unknown variables, which consist of $\{\beta_k\}_{k \in \mathsf{Y}_{2,1}}$, and $\{\delta_k\}_{k \in \mathsf{Y}_{1,2} \cup \mathsf{Y}_{2,2}}$. The number of those variables is $2\rho-\sigma-\tau-\gamma \le t$ and, in view of what we have shown in Section 2, the solution of (17) for these variables yields a unique error array $E$.

Following is a summary of Step 3 of the outlined decoding algorithm. We assume that the $\rho \times \rho$ matrix $A$ has the form (11)–(14) and that the matrices $U_{1,1}$, $U_{1,2}$, $D_{1,1}$, and $D_{1,2}$ are known. The decoder therefore knows the entries $\{\beta_k\}_{k \in \mathsf{Y}_{1,1} \cup \mathsf{Y}_{1,2}}$ in the row vector $\boldsymbol{\beta} = \boldsymbol{\alpha} U$, as well as the entries $\{\delta_k\}_{k \in \mathsf{Y}_{1,1} \cup \mathsf{Y}_{2,1}}$ in the column vector $\boldsymbol{\delta} = D\boldsymbol{\omega}$.

1. Compute the syndrome vector $\mathsf{s}$ of $E - U_{1,1} D_{1,1}$ as in (15).

2. Compute the coefficients of $\Phi(x)$ and $\Psi(x)$ by solving (18) and (19).

3. Compute the matrix $R_\gamma SQ^T$ using (20), (21), and (25).

4. Compute the coefficients of $\Theta(x)$ by solving (26).

5. Find a matrix $U_{2,2}$ so that the entries of $\boldsymbol{\alpha}U_{2,2}$ extend $\{\beta_k\}_{k\in\mathsf{Y}_{1,2}}$ to form the null space of the mapping $x \mapsto \Lambda(x) = \Theta(\Phi(x))$ over $\mathsf{F}_q$.

6. Find the remaining unknown components of $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ by solving the set of (already linear) equations (17).

7. The error array $E$ equals $UD$, where $\boldsymbol{\beta} = \boldsymbol{\alpha}U$ and $\boldsymbol{\delta} = D\boldsymbol{\omega}$.

# References


[1] E.R. Berlekamp, H. Rumsey, G. Solomon, *On the solution of algebraic equations over finite fields, Inform. Control,* 10 (1967), 553–564.

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

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

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

[6] 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).

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

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

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

[10] 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.

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

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

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