Embedded Rank Distance Codes for ISI channels

S. Dusad* S. N. Diggavi* A. R. Calderbank†

Abstract

Designs for transmit alphabet constrained space-time codes naturally lead to questions about the design of rank distance codes. Recently, diversity embedded multi-level space-time codes for flat fading channels have been designed from sets of binary matrices with rank distance guarantees over the binary field by mapping them onto QAM and PSK constellations. In this paper we demonstrate that diversity embedded space-time codes for fading Inter-Symbol Interference (ISI) channels can be designed with provable rank distance guarantees. As a corollary we obtain an asymptotic characterization of the fixed transmit alphabet rate-diversity trade-off for multiple antenna fading ISI channels. The key idea is to construct and analyze properties of binary matrices with a particular structure (Toeplitz structure) induced by ISI channels.

1 Introduction

Over the past decade significant progress has been made in constructing space-time codes that achieve the optimal rate-diversity trade-off for *flat-fading* channels when there are transmit alphabet constraints [19, 16]. Far less attention has been given to space-time code design and analysis for fading channels with memory, *i.e.,* Inter-Symbol Interference (ISI) channels which are encountered in broadband multiple antenna communications. There have been several constructions of space-time codes for fading ISI channels using multi-carrier techniques (see for example [18] and references therein). However, since these inherently increase the transmit alphabet size, and the right framework to study such constructions is through the *diversity-multiplexing* (D-M) trade-off [20]. We have examined diversity embedded codes for ISI channels in [7, 8], by considering the diversity-multiplexing trade-off[1].

As in space-time code design for flat-fading channels, it is natural to ask for a characterization of the rate-diversity trade-off for ISI channels with transmit alphabet constraints[2]. The problem of constructing space-time codes with fixed transmit alphabet constraints is partially motivated by the need to control the transmit spectrum as well as the peak-to-average (PAR) ratio of the transmitted signal. For example,

*EPFL, Lausanne, Switzerland, Email: {sanket.dusad,suhas.diggavi}@epfl.ch. S. Dusad was supported in part by SNSF Grant # 200021-105640/1. S N. Diggavi is part of the SNF NCCR-MICS center on wireless sensor networks.

†Princeton University, Email: calderbank@math.princeton.edu. A R. Calderbank was supported in part by NSF grant # 1096066.

[1] In [7, 8] we show that for MISO/SIMO/SISO ISI fading channels, the D-M trade-off is successively refinable. An interesting aspect of these results is that the correlations of the frequency domain channel response is crucially used in establishing the result.

[2] Throughout this paper we restrict our attention to a transmit alphabet constraint, *i.e.,* the transmit alphabet is restricted to be from the set $\mathcal{A}$. Therefore this imposes a maximal rate of $M_t \log |\mathcal{A}|$ bits and we normalize the rate by $\log |\mathcal{A}|$ and state the rate in terms of a number in $[0, M_t]$ symbols per transmission.

if we restrict transmission to PSK alphabet, it is clear that we have a unit peak-to-average ratio (PAR) making it possible to use efficient non-linear amplifiers (requiring small PAR), which are more efficient and hence suitable for mobile devices. Another important reason to consider this problem is a fundamental theoretical question, which is motivated by the origins of space-time codes for flat-fading channels in [19] where the constructions were for fixed transmit alphabet. For this constraint, there exists a trade-off between rate and diversity, for the flat-fading case. In this paper we ask the corresponding question for fading ISI channels. Since space-time code design with diversity order guarantees requires control over the rank distance of the codewords [19], the main topic of this paper is to design codes with rank distance guarantees for ISI channels.

Diversity embedded codes were introduced in [2] which allowed different parts of a message to have different diversity order guarantees. These codes allowed diversity to be viewed as a systems resource that can be allocated judiciously to achieve a target rate-diversity trade-off in wireless communications. A class of such multi-level diversity embedded codes suitable for flat-fading channels was constructed in [5, 1, 6]. In this paper we extend these constructions to ISI channels.

The corresponding question, of what is studied in this paper, can be also be posed in the context of the trade-off between diversity and multiplexing rate. Such an information-theoretic question, for the flat-fading case, has been posed and partially answered in [3, 4]. For scalar ISI channels, we have studied code designs for rate-growth (multiplexing rate) codes and the diversity embedding properties in [7]. There we have shown that the diversity multiplexing trade-off for the scalar ISI channel is actually successively refinable. However, the code designs and criteria for the rate-growth codes are quite different from those needed for the fixed rate, transmit alphabet constrained codes, which are the focus of this paper.

For the case of a scalar ISI channel with $\nu + 1$ taps and a single transmit antenna, it can be shown by a simple argument (see for example [20]) that an uncoded transmission scheme can achieve a diversity order of $(\nu + 1)$. The best case scenario for the rate-diversity trade-off for ISI channels with multiple transmit antennas would be similar to the flat-fading case, but with a $(\nu + 1)-$fold increase in the diversity order. However, in the multiple transmit antenna case, it is not obvious that a space-time code designed for a flat-fading channel can achieve such a $(\nu+1)-$fold increase in the diversity order. All that can be guaranteed is that a space-time code that achieves diversity order d over a flat-fading channel will still achieve diversity order d over a fading ISI channel [19]. In particular in Example 1 of Section 7 we provide an example of a code which achieves particular points on rate-diversity trade-off for flat-fading channels and fails to do so in the case of ISI channels. Therefore, the design of codes for fading ISI channels cannot be immediately done by using the codes for flat-fading channels. However, in this paper we see that codes designed for the fading ISI channel can be used successfully to achieve the rate-diversity trade-off for the flat-fading case as well.

A finite alphabet construction to exploit the potential diversity gain from ISI channels with M_t multiple transmit antennas was proposed in [10] for the maximal diversity case. But the rate of the code for this construction was $1/M_t$ as opposed to the maximal potential rate of 1. In this paper we show that as the transmission block size increases we can construct codes that have rate 1 and achieve the maximal diversity order of $(\nu + 1)M_t$. We characterize the rate diversity tradeoff for the ISI channels and construct codes which achieve this tradeoff (asymptotically in block size). We build on the construction technique introduced in [6] to design diversity embedded codes for ISI channels that guarantee multiple reliability (diversity) levels. Given that we can achieve a $(\nu+1)-$fold increase in the diversity order for ISI channels, this flexibility could be quite important.

The main contribution of this paper is the construction of multi-level diversity embedded codes with finite alphabet constraints for the fading ISI channel. In doing so, we also establish several other related re-

sults. Analogous to the (diversity-embedded) space-time code design criteria for flat-fading channels, we establish (diversity-embedded) space-time code design criteria for transmission over fading ISI channels. These design criteria require the construction of rank-distance codes, underlying which is the construction of binary codes with Toeplitz constraints having rank-distance guarantees. We believe that the maximal construction of such codes established in this paper, is one of the main technical ideas. These constructions allow a $(\nu+1)$ multiplicative gain in diversity order when transmitted over an ISI channel with $(\nu+1)$ taps (as compared to the performance over flat-fading channels). In constructing the multi-level diversity embedded codes for ISI channels, we also specialize it for classical space-time codes for ISI channels, and therefore establish the rate diversity tradeoff for fading ISI channels.

The outline of the paper is as follows. We extend the rate-diversity trade-off bound from [19] and develop the diversity embedded code design criteria for fading ISI channels in Section 2. The basic multi-level construction of diversity-embedded space-time code for fading ISI channels is given in Section 3. We also show that this construction can be specialized to asymptotically achieve the diversity-rate trade-off for ISI channels. The key ingredient is the construction of binary codes for ISI channels with rank-distance guarantees, and this is done in Section 5 and Section 6. This is perhaps the most important technical contribution of this paper. We also construct of convolutional codes suitable for transmission over the ISI channel in Section 4. In Section 7 we give examples of codes constructed by the method given in the paper along with their numerical performance.

2 Problem Statement and code design criteria

In Section 2.1, we define the ISI channel model . Section 2.2 recalls the code design criteria for diversity embedded codes for flat-fading channels given in [6] and extends it to the fading ISI case. These criteria give the connection between embedded rank-distance codes and diversity-embedded space-time codes. The rate-diversity trade-off for flat-fading channels is reviewed in Section 2.3, and a simple upper bound for the corresponding trade-off for the fading ISI case is established. The subsections 2.4 and 2.5 are based on [6] and reproduced here for completeness. In Section 2.4, we review the principle of set-partitioning and give algebraic properties of such partitions in Section 2.5. These properties would be useful in *lifting* rank properties of binary matrices over binary fields to the complex domain, thereby giving provable diversity embedded code constructions.

2.1 Channel Model

Our focus in this paper is on the quasi-static frequency selective (ISI) channel with $(\nu+1)$ taps where we transmit information coded over M_t transmit antennas and have M_r antennas at the receiver. Furthermore, we make the standard assumption that the transmitter has no channel state information, whereas the receiver is able to perfectly track the channel. The code is designed over a large enough block size $T \geq T_{thr}$ transmission symbols, where T_{thr} is specified in the constructions given in Section 3. The received vector at time n after demodulation and sampling can be written as,

$$\mathbf{y}[n] = \mathbf{H}_0\mathbf{x}[n] + \mathbf{H}_1\mathbf{x}[n-1] + \ldots + \mathbf{H}_\nu\mathbf{x}[n-\nu] + \mathbf{z}[n] \tag{1}$$

where, $\mathbf{y} \in \mathbb{C}^{M_r \times 1}$, $\mathbf{H}_l \in \mathbb{C}^{M_r \times M_t}$ represents the matrix ISI channel, $\mathbf{x}[n] \in \mathbb{C}^{M_t \times 1}$ is the space-time coded transmission sequence at time n with transmit power constraint P and $\mathbf{z} \in \mathbb{C}^{M_t \times 1}$ is assumed to be additive white (temporally and spatially) Gaussian noise with variance σ^2. The matrix $\mathbf{H}_l$ consists of fading coefficients h_{ij} which are i.i.d. $\mathcal{CN}(0,1)$ and fixed for the duration of the block length (T).

Consider a transmission scheme in which we transmit over a period $T-\nu$ and send (fixed) known symbols[3] for the last ν transmissions. For the period of communication we can equivalently write the received data as,

$$\underbrace{\begin{bmatrix} \mathbf{y}[0] \\ \vdots \\ \mathbf{y}[T-1] \end{bmatrix}^t}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} \mathbf{H}_0 \\ \vdots \\ \mathbf{H}_\nu \end{bmatrix}^t}_{\mathbf{H}} \underbrace{\begin{bmatrix} \mathbf{x}[0] & \mathbf{x}[1] & \dots & \mathbf{x}[T-\nu-1] & 0 & \dots & 0 \\ 0 & \mathbf{x}[0] & \mathbf{x}[1] & \dots & \mathbf{x}[T-\nu-1] & 0 & 0 \\ \dots & & \vdots & \ddots & & . & \vdots \\ 0 & \dots & 0 & \mathbf{x}[0] & \mathbf{x}[1] & \dots & \mathbf{x}[T-\nu-1] \end{bmatrix}}_{\mathbf{X}} + \mathbf{Z} \qquad (2)$$

i.e.,

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z}, \qquad (3)$$

where $\mathbf{Y} \in \mathbb{C}^{M_r \times T}$, $\mathbf{H} \in \mathbb{C}^{M_r \times (\nu+1)M_t}$, $\mathbf{X} \in \mathbb{C}^{(\nu+1)M_t \times T}$, $\mathbf{Z} \in \mathbb{C}^{M_r \times T}$. *Notice that the structure in (2) is different from the flat-fading case, since the channel imposes a Toeplitz structure on the equivalent space-time codewords* $\mathbf{X}$ *given in (2)-(3).* This structure makes the design of space-time codes different than in the flat-fading case. For reference, the space-time codeword is completely determined by the matrix $\mathbf{X}^{(1)}$ given by,

$$\mathbf{X}^{(1)} = \begin{bmatrix} \mathbf{x}[0] & \mathbf{x}[1] & \dots & \mathbf{x}[T-\nu-1] & 0 & \dots & 0 \end{bmatrix}. \qquad (4)$$

2.2 Diversity-embedded code design criteria

A scheme with diversity order d has an error probability at high SNR behaving as $\bar{P}_e(\text{SNR}) \approx \text{SNR}^{-D}$ [19]. More formally,

Definition 1 *A coding scheme which has an average error probability $\bar{P}_e(SNR)$ as a function of SNR that behaves as*

$$\lim_{SNR \to \infty} \frac{\log(\bar{P}_e(SNR))}{\log(SNR)} = -D \qquad (5)$$

is said to have a diversity order of D.

The fact that the diversity order of a space-time code is determined by the rank of the codeword difference matrix is well known [19, 11]. Therefore, for flat-fading channels, it has been shown that the diversity order achieved by a space-time code is given by [19]

$$D = M_r \min_{\mathbf{C}_1 \neq \mathbf{C}_2} \text{rank}(\mathbf{C}_1 - \mathbf{C}_2)\ , \qquad (6)$$

where $\mathbf{C}_1, \mathbf{C}_2 \in \mathbb{C}^{M_t \times T}$ are the space-time codewords. Clearly the analysis in [19, 11] can be easily extended to fading ISI channels, and we can write

$$D = M_r \min_{\mathbf{X}_1 \neq \mathbf{X}_2} \text{rank}(\mathbf{X}_1 - \mathbf{X}_2)\ , \qquad (7)$$

where $\mathbf{X}_1, \mathbf{X}_2 \in \mathbb{C}^{(\nu+1)M_t \times T}$ are matrices with structure given in (2).

[3] Taken without loss of generality to be 0.

It is easy to see from the structure of $\mathbf{X}$ in (2) that the rank of the matrix $\mathbf{X}$ is *at most* $(\nu+1)$ times the rank of the matrix $\mathbf{X}^{(1)}$ (see (4)), *i.e.*,

$$\text{rank}(\mathbf{X}) \leq (\nu+1)\text{rank}(\mathbf{X}^{(1)}). \tag{8}$$

The codebook structure proposed in [6] takes two information streams and outputs the transmitted sequence $\{\mathbf{x}(k)\}$. The objective is to ensure that each information stream gets the designed rate and diversity levels. Let $\mathcal{E}$ denote the message set from the first information stream and $\mathcal{F}$ denote that from the second information stream. Then analogous to Definition 1, we can write the diversity order for the messages as,

$$D_H = \lim_{\text{SNR}\to\infty} \frac{\log \bar{P}_e^E(SNR)}{\log(\text{SNR})},\ D_L = \lim_{\text{SNR}\to\infty} \frac{\log \bar{P}_e^F(SNR)}{\log(\text{SNR})}. \tag{9}$$

Design criteria for fading ISI channels: The space-time codeword for fading ISI channels have the structure given in (2). To translate this to the diversity embedded case, we annotate it with given messages $\mathbf{a} \in \mathcal{E}, \mathbf{b} \in \mathcal{F}$, as $\mathbf{X}_{\mathbf{a},\mathbf{b}}$. Clearly we can then translate the code design criterion from (7) to diversity embedded codes for ISI channels as,

$$\min_{\mathbf{a}_1\neq\mathbf{a}_2\in\mathcal{E}} \min_{\mathbf{b}_1,\mathbf{b}_2\in\mathcal{F}} \text{rank}(\mathbf{X}_{\mathbf{a}_1,\mathbf{b}_1} - \mathbf{X}_{\mathbf{a}_2,\mathbf{b}_2})) \geq D_E/M_r. \tag{10}$$

In an identical manner, we can show for the message set $\mathcal{F}$, we need the following to hold.

$$\min_{\mathbf{b}_1\neq\mathbf{b}_2\in\mathcal{F}} \min_{\mathbf{a}_1,\mathbf{a}_2\in\mathcal{E}} \text{rank}(\mathbf{X}_{\mathbf{a}_1,\mathbf{b}_1} - \mathbf{X}_{\mathbf{a}_2,\mathbf{b}_2})) \geq D_F/M_r. \tag{11}$$

As one can easily see, these are simple generalizations of the diversity-embedded code design criteria developed in [2] to the fading ISI case.

2.3 Rate-Diversity Trade-off for Flat Fading Channels

For a given diversity order, it is natural to ask for upper bounds on achievable rate. For a flat Rayleigh fading channel, this has been examined in [19] where the following result was obtained.

Theorem 2 *([19, 15]) Given a constellation of size $|\mathcal{A}|$ and a system with diversity order $D = dM_r$, then the rate R that can be achieved is given by*

$$R \leq (M_t - d + 1) \tag{12}$$

in symbols per transmission, i.e., *the rate is $R\log_2|\mathcal{A}|$ bits per transmission.*

Just as Theorem 2 shows the trade-off between achieving high-rate and high-diversity given a fixed transmit alphabet constraint for a flat fading channel, there also exists a trade-off between achievable rate and diversity for frequency selective channels, and we aim to characterize this trade-off[4]. A corollary will be an upper bound on the performance of diversity embedded codes for ISI channels. This can be seen by observing that we can easily extend the proof in [19, 15] to the case where we have the Toeplitz structure as given in (2). Note that the diversity order of the codes for fading ISI channel is given by the rank of the corresponding (Toeplitz) codeword difference matrix. Since this rank is upper bounded as seen in (8), we see that we immediately obtain a trivial upper bound for the rate-diversity trade-off for the fading ISI case as follows.

[4]It is tempting to guess that the trade-off for the fading ISI case is just a $(\nu+1)-$fold increase in the diversity order.

Lemma 3 *If we use a constellation of size $|\mathcal{A}|$ and the diversity order of the system is $D_{isi} = d(\nu+1)M_r$, then the rate R in symbols per transmission that can be achieved is upper bounded as*

$$R \leq (M_t - d + 1). \tag{13}$$

Note that in Theorem 8, we establish a corresponding lower bound that asymptotically (in block size) matches this upper bound. Note that due to the zero padding structure for ISI channels, the effective rate R^{eff} is going to be smaller than the rate of space-time code. Since we do not utilize ν transmissions over a block of T transmissions for each of the antennas we can only hope for a rate R asymptotically in transmission block size T.

2.4 Set Partitioning of QAM and QPSK Constellations

Let $\Gamma_1, \ldots, \Gamma_L$ be a L-level partition where Γ_i is a refinement of partition Γ_{i-1}. We view this as a rooted tree, where the root is the entire signal constellation and the vertices at level i are the subsets that constitute the partition Γ_i. In this paper we consider only binary partitions, and therefore subsets of partition Γ_i can be labeled by binary strings $a_1, \ldots, a_i$, which specify the path from the root to the specified vertex.

Signal points in QAM constellations are drawn from some realization of the integer lattice $\mathbb{Z}^2$. We focus on the particular realization shown in Figure 1, where the integer lattice has been scaled by $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ to give the lattice $D_2 = \{(a,b)|a,b \in \mathbb{Z}, a+b \equiv 0(\mathrm{mod}2)\}$, and then translated by $(1,0)$. The constellation is formed by taking all the points from Λ that fall within a bounding region $\mathcal{R}$. The size of the constellation is proportional to the area of the bounding region, and in Figure 1, the bounding region encloses 16 points.

Binary partitions of QAM constellations are typically based on the following chain of lattices

$$D_2 \supset 2\mathbb{Z}^2 \supset 2D_2 \supset 4\mathbb{Z}^2 \supset \ldots 2^{i-1}D_2 \supset 2^i\mathbb{Z}^2 \supset 2^iD_2 \supset \ldots$$

In Figure 1, the subsets at level 1 are, to within translation, cosets of $2\mathbb{Z}^2$ in D_2 and the subsets at level 2 are cosets of $2D_2$. In general the subsets at level $2i$ are pairs of cosets of 2^iD_2 where the union is a coset of $2^i\mathbb{Z}^2$, and the subsets at level $2i+1$ are pairs of cosets of $2^{i+1}\mathbb{Z}^2$ where the union is a coset of 2^iD_2. Note that implicit in Figure 1 is a binary partition of QPSK, where the points $1, -1, i, -i$ are labeled $00, 01, 11, 10$ respectively. Binary partitions of PSK constellations are described in Section 2.5.

2.5 Algebraic properties of binary partitions

The QAM constellations can be represented through a lattice chain $\Lambda|\Lambda_1|\Lambda_2|\ldots$, where $\Lambda = \mathbb{Z}^2$ is the integer lattice. The lattices in the chain are produced with the generator matrix $\mathbf{G}^k$ where $\mathbf{G} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Given this, we can represent the 2^k-QAM constellation as $\Lambda|\Lambda_k$, *i.e.,* the coset representatives of Λ in Λ_k. The lattice Λ can also be written as the set of Gaussian integers $Z[i] = \{a + bi : a, b \in \mathbb{Z}\}$. Similarly we can write the lattice Λ_k as $\{(a+bi)(1-i)^k : a, b \in \mathbb{Z}\}$. This decomposition of the QAM constellation is illustrated in Figure 2. Therefore, using this we can represent any point s in a 2^L-QAM constellation using a L-length bit string as

$$s - c(L) \equiv \sum_{l=0}^{L-1} b_l(1-i)^l \mod (1-i)^L, \tag{14}$$

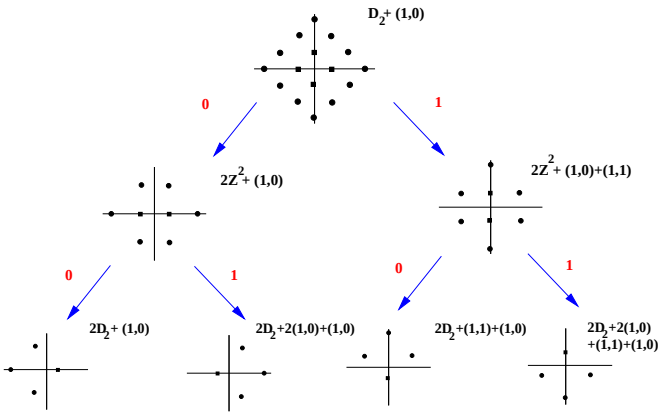

Figure 1: A binary partition of a QAM constellation

where we define $f \equiv g \mod (1-i)^L$ if there exist $c, d \in \mathbb{Z}$ such that $f = (c+di)(1-i)^L + g$. Also in (14) the constant $c(L) = \frac{1}{2}$ for odd L and $\frac{1}{2}(1+i)$ for even L.

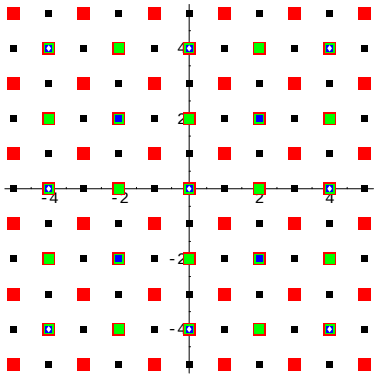

Figure 2: Decomposition for QAM constellations

Binary partitions of PSK constellations are based on a chain of subfields of the cyclotomic field $\mathbb{Q}(\xi_{2^L})$ obtained by adjoining $\xi_{2^L} = \exp(2\pi i/2^L)$ to the rational field $\mathbb{Q}$. Analogous to (14), points in the 2^L-PSK constellation can be represented as

$$s = \prod_{l=0}^{L-1} (\xi^{2^l})^{b_l}, \tag{15}$$

where $\xi = \xi_{2^L} = \exp(2\pi i/2^L)$ and ξ^{2^L} is a primitive element for $\mathbb{Q}(\xi_{2^L})$. Note that $1-\xi$ is prime in $\mathbb{Z}[\xi]$ and the quotient $\mathbb{Z}[\xi]/(1-\xi)$ is the field $\mathbb{Z}_2$.

The field $\mathbb{Q}(\xi_{2^L})$ is a degree 2^{L-1} extension of $\mathbb{Q}$. Every rational number is a quotient a/b, where $a, b \in \mathbb{Z}$, and every complex number in $\mathbb{Q}(i)$ is a quotient a/b, where a, b are Gaussian integers. In general every complex number in $\mathbb{Q}(\xi_{2^L})$ is a quotient a/b, where a, b are integer linear combinations of $1, \xi_{2^L}, \ldots, \xi_{2^L}^{2^{L-1}-1}$ and $b \neq 0$. For more details about cyclotomic fields see [21]. Note that $\xi_{2^L}^{2^{L-1}} = -1$, so that $\xi_{2^L}^{j} = -\xi_{2^L}^{2^{L-1}+j}$, for $j = 0, 1, \ldots, 2^{L-1}-1$.

We have a chain of fields

$$\mathbb{Q} = \mathbb{Q}(\xi_2) \subset \mathbb{Q}(i) = \mathbb{Q}(\xi_4) \subset \mathbb{Q}(\xi_8) \ldots \subset \mathbb{Q}(\xi_{2^L}).$$

These observations are used to establish the performance of the multi-level diversity embedded codes in the proof of Theorem 7, using similar techniques as in [5, 6].

3 Diversity embedded codes for ISI channels

In this section we will first recall the construction of multi-level (non-linear) space-time codes for transmission over *flat fading* channels that are matched to a binary partition of a QAM or PSK constellation (see [6]). We will give the construction and refer the reader to [6] for proofs of code performance for the flat-fading case. Following this we will use the structure imposed by the ISI on the space time code as in (2) to construct multilevel codes *suitable for ISI channels* using binary matrices which are constructed in Section 6.

3.1 Multi-Level Constructions for Flat Fading Channels

Given an L-level binary partition of a QAM or PSK signal constellation, a space-time codeword is an array $\mathbf{K} = \{\mathbf{K}_1, \mathbf{K}_2, \ldots, \mathbf{K}_L\}$ determined by a sequence of binary matrices, where matrix, $\mathbf{K}_i$ specifies the space-time array at level i. A *multi-level space-time code* is defined by the choice of the constituent sets of binary matrices $\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_L$. These sets of binary matrices provide rank guarantees necessary to achieve the diversity orders required for each message set. For $i = 1, \ldots, L$ the binary matrix $\mathbf{K}_i$ is required to be in the set $\mathcal{K}_i$.

Given message sets $\{\mathcal{E}_i\}_{i=1}^{L}$, they are mapped to the space-time codeword $\mathbf{X}$ as shown below

$$\{\mathcal{E}_i\}_{i=1}^{L} \xrightarrow{f_1} \mathbf{K} = \begin{bmatrix} K(1,1) & \ldots & K(1,T) \\ \vdots & \vdots & \vdots \\ K(M_t,1) & \ldots & K(M_t,T) \end{bmatrix} \xrightarrow{f_2} \mathbf{X} = \begin{bmatrix} x(1,1) & \ldots & x(1,T) \\ \vdots & \vdots & \vdots \\ x(M_t,1) & \ldots & x(M_t,T) \end{bmatrix}, \qquad (16)$$

where the matrix $\mathbf{K}$ is specified by $K(m,n) \in \{0,1\}^{\log(|\mathcal{A}|)}$ *i.e.,* binary string and $x(m,n) \in \mathcal{A}$. This construction is illustrated in Figure 3 for a constellation size of L bits.

In summary, given the message set, we first choose the matrices $\mathbf{K}_1, \ldots . \mathbf{K}_L$. The first mapping f_1 is obtained by taking matrices and constructing the matrix $\mathbf{K} \in \mathbb{C}^{M_t \times T}$ each of whose entries is constructed by concatenating the bits from the corresponding entries in the matrices $\mathbf{K}_1, \ldots . \mathbf{K}_L$ into L-length bit-string. This matrix is then mapped to the space-time codeword through a constellation mapper f_2, for example the set-partition mapping given in Section 2.4. Using this sequence of L matrices, we obtain the space-time codeword as seen in Figure 3.

For flat fading channels the sets $\mathcal{K}_l, l = 1, \ldots, L$ are binary $M_t \times T$ matrices such that for any distinct pair of matrices $\mathbf{A}, \mathbf{B} \in \mathcal{K}$ the rank of $\mathbf{A} - \mathbf{B}$ is at least d. The size of the codebook is therefore at most $2^{(M_t - d + 1)T}$ since at least one of the first $(M_t - d + 1)$ rows of $\mathbf{A}$ and $\mathbf{B}$ must be distinct. There is a classical example [9] that achieves the bound (this construction was also given in [15, 16]).

With the rate achieved on the l^{th} layer defined as $R_l = \frac{1}{T} \log |\mathcal{K}_l|$ it can be shown [6] that this construction for QAM constellations achieves the rate-diversity tuple $(R_1, M_r d_1, \ldots, R_L, M_r d_L)$, with the overall equivalent single layer code achieving rate-diversity point, $(\sum_l R_l, M_r d_L)$. Optimal decoding employs a

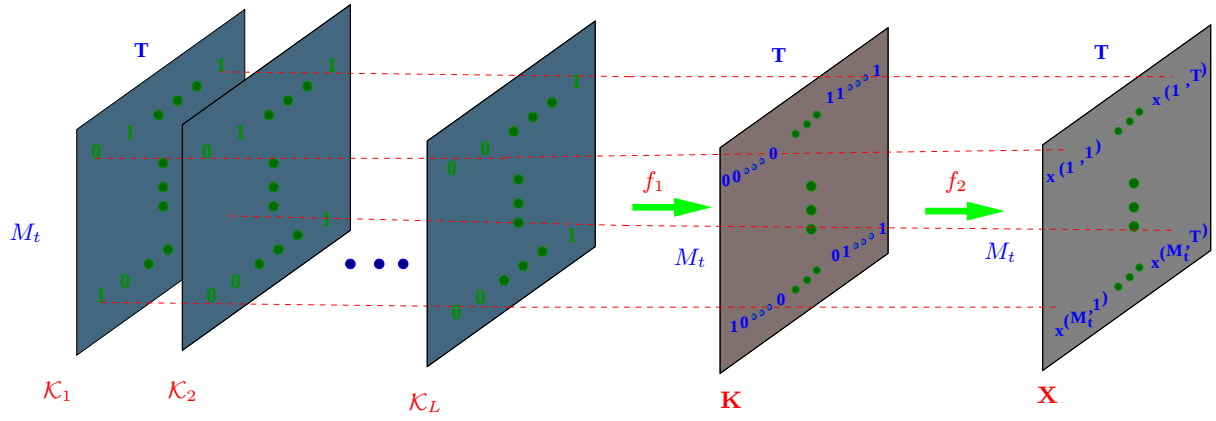

Figure 3: Schematic representation of the multi-level code construction.

maximum-likelihood decoder which jointly decodes the message sets. This is the decoder for which the performance is summarized in Theorem 4.

Theorem 4 *[6] Let $\mathcal{C}$ be a multi-level space-time code for a QAM or M-PSK constellation of size 2^L with M_t transmit antennas that is determined by constituent sets of binary matrices $\mathcal{K}_l, l = 1, \ldots, L$ with binary rank guarantees $d_1 \geq d_2 \ldots \geq d_L$. For joint maximum-likelihood decoding, the input bits that select the codeword from the ith matrix $\mathcal{K}_i$ are guaranteed diversity $D_i = d_i M_r$ in the complex domain when transmitted over a flat fading channel.*

3.2 Multi-Level Construction for ISI Channels

In this section we use the idea of multi-level diversity embedded codes for flat fading channels as in Section 3.1 and the structure imposed by the ISI on the space time code as in (2) to motivate construction and analysis of a class of binary matrices as follows.

We apply the idea suggested by the constructions of multi-level codes for flat-fading channels to the fading ISI case. We do this by applying a zero padding as seen in (4) along with mappings of binary matrices to the transmit signal alphabet. That is, we use the mapping given in (16) for a block size of T with the constraint that the last ν entries of the mapping lead to *given* alphabets (taken to be zero without loss of generality). This is combined with binary sets $\mathcal{K}_{\nu,d}$, which we specify in Definition 5. This means that over a time period T, we transmit a sequence $\mathbf{x}[0], \mathbf{x}[1], \ldots, \mathbf{x}[T-\nu-1]$ which are mapped from the inputs bits using a structure given in (16). Therefore, given that we transmit the sequence shown in (17),

$$\mathbf{X}^{(1)} = \left[\begin{array}{cccccccc} \mathbf{x}[0] & \mathbf{x}[1] & \ldots & \mathbf{x}[T-\nu-1] & 0 & \ldots & 0 \end{array} \right], \tag{17}$$

we need a mapping from a binary matrix as in (16). For a constellation of size 2^L, we do this by taking message sets $\{\mathcal{E}_i\}_{i=1}^{L}$ and mapping them to a codeword with the structure given in (17) as follows,

$$\{\mathcal{E}_i\}_{i=1}^{L} \xrightarrow{f_1} \mathbf{K}^{(1)} = \left[\begin{array}{ccc} K(1,1) & \ldots & K(1,T) \\ \vdots & \vdots & \vdots \\ K(M_t,1) & \ldots & K(M_t,T) \end{array} \right] \xrightarrow{f_2} \mathbf{X}^{(1)} = \left[\begin{array}{cccccccc} \mathbf{x}[0] & \mathbf{x}[1] & \ldots & \mathbf{x}[T-\nu-1] & 0 & \ldots & 0 \end{array} \right] \tag{18}$$

where the $(m,n)^{th}$ entry of $\mathbf{K}^{(1)}$ is given by $K(m,n) \in \{0,1\}^{\log(|\mathcal{A}|)}$ *i.e.,* binary string. Since the mapping f_2 is just the set-partitioning mapping specified in Section 2.4, we need the last ν columns of $\mathbf{K}^{(1)}$ to be

given constants for *all* choices of the message sets $\{\mathcal{E}_i\}_{i=1}^{L}$. That is, we need the following structure for the matrix $\mathbf{K}^{(1)}$,

$$\mathbf{K}^{(1)} = \left[\begin{array}{ccccccc} \mathbf{k}[0] & \mathbf{k}[1] & \dots & \mathbf{k}[T-\nu-1] & \mathbf{0} & \dots & \mathbf{0} \end{array} \right], \tag{19}$$

where, as before, $\{\mathbf{k}[i]\}$ are columns of binary strings of length L, and with no loss of generality, we have specified the last ν columns of $\mathbf{K}^{(1)}$ to be the zero strings.

Given that we have an ISI channel, the transmitted codeword with the structure given in (17) gives an equivalent codeword matrix with a Toeplitz structure, as specified in (2). This Toeplitz structure is equivalent to mapping a Toeplitz matrix $\mathbf{K}$ of binary strings with the structure:

$$\mathbf{K} = \left[\begin{array}{ccccccc} \mathbf{k}[0] & \mathbf{k}[1] & \dots & \mathbf{k}[T-\nu-1] & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{k}[0] & \mathbf{k}[1] & \dots & \mathbf{k}[T-\nu-1] & \mathbf{0} & \mathbf{0} \\ \dots & & \vdots & \ddots & . & . & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{k}[\mathbf{0}] & \mathbf{k}[1] & \dots & \mathbf{k}[T-\nu-1] \end{array} \right], \tag{20}$$

to $\mathbf{X}$ using the constellation mapping f_2. Therefore, as in the flat fading case, given the message set, we first choose the binary matrices $\mathbf{K}_1^{(1)}, \dots . \mathbf{K}_L^{(1)}$, each of which have the structure specified below in (21). These put together give us the matrix of binary strings $\mathbf{K}^{(1)}$. This in turn, due to the ISI channel, is related to $\mathbf{K}$, the Toeplitz matrix of binary strings, given above in (20). Therefore, the choice of matrices $\mathbf{K}_1^{(1)}, \dots . \mathbf{K}_L^{(1)}$, for the ISI case, naturally is equivalent to a choice of Toeplitz binary matrices, $\mathbf{K}_1, \dots . \mathbf{K}_L$, as specified in (22) below.

Therefore, for the multi-level coding structure we have used, analogous to the flat fading case studied in [6], we need to study the rank properties of sets of binary Toeplitz matrices as specified below. Consider the matrix $\mathbf{B} \in \mathbb{F}_2^{M_t \times T}$, with the following structure,

$$\mathbf{B} = \left[\begin{array}{ccccccc} \mathbf{c}[0] & \mathbf{c}[1] & \dots & \mathbf{c}[T-\nu-1] & \mathbf{0} & \dots & \mathbf{0} \end{array} \right], \tag{21}$$

where $\mathbf{c}[n] \in \mathbb{F}_2^{M_t \times 1},\ n = 0, \dots, T-\nu-1$. We define a mapping $\Theta_\nu : \mathbb{F}_2^{M_t \times T} \rightarrow \mathbb{F}_2^{(\nu+1)M_t \times T}$ by,

$$\Theta_\nu\left(\mathbf{B}\right) = \left[\begin{array}{ccccccc} \mathbf{c}[0] & \mathbf{c}[1] & \dots & \mathbf{c}[T-\nu-1] & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{c}[0] & \mathbf{c}[1] & \dots & \mathbf{c}[T-\nu-1] & \mathbf{0} & \mathbf{0} \\ \dots & & \vdots & \ddots & . & . & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{c}[\mathbf{0}] & \mathbf{c}[1] & \dots & \mathbf{c}[T-\nu-1] \end{array} \right]. \tag{22}$$

Definition 5 *Define $\mathcal{K}_{\nu,d} \subset \{\mathbf{B} : \mathbf{B} \in \mathbb{F}_2^{M_t \times T}\}$ to be the set of $M_t \times T$ binary matrices such that for $T \geq T_{thr}$, they satisfy the following properties:*

- *The last ν entries of all the M_t rows are zero.*
- *For any distinct pair of matrices $\mathbf{A}, \mathbf{B} \in \mathcal{K}_{\nu,d}$ the rank of $[\Theta_\nu(\mathbf{A}) - \Theta(\mathbf{B})]$ is at least $d(\nu+1)$.*
- $|\mathcal{K}_{\nu,d}| \geq 2^{T(M_t-d+1)-\nu M_t}$.

Note that in Section 3.1 the first step in code construction was constructing the sets $\mathcal{K}_l, l = 1, \dots, L$ from which the matrices $\mathbf{K}_1, \dots . \mathbf{K}_L$ were chosen. In the case of flat fading channels there are constructions by [9] but these do not satisfy the rank guarantee properties in Definition 5. We will postpone the construction of such sets of binary matrices $\{\mathcal{K}_{\nu,d}\}$ to Section 6, where we show that we can set $T_{thr} = R\nu + (M_t - 1)(\nu+1)(2^R - 1)$. More formally, in Section 6, we show that,

Lemma 6 *For block size $T \geq T_{thr} = R\nu + (M_t - 1)(\nu + 1)(2^R - 1)$, there exist sets of binary matrices $\mathcal{K}_{\nu,d}$ which satisfy the requirements of Definition 5.*

Adapted easily from [6] we can state the formal construction guarantee for the diversity embedded code for transmission over the ISI channel as follows.

Theorem 7 *Let $\mathcal{C}$ be a multi-level space-time code for a QAM or PSK constellation of size 2^L with M_t transmit antennas that is determined by constituent sets of binary matrices $\mathcal{K}_l = \mathcal{K}_{\nu,d_l}, l = 1, \ldots, L$, such that $d_1 \geq d_2 \ldots \geq d_L$. For joint maximum-likelihood decoding, the input bits that select the codeword from the lth set $\mathcal{K}_l$ are guaranteed diversity $D_l = d_l(\nu + 1)M_r$ in the complex domain when transmitted over an ISI channel with $\nu + 1$ taps.*

The proof of the Theorem 7 follows from the same techniques as in [6] by mapping binary matrices with desired rank guarantees to rank guarantees in complex domain and *crucially utilizes the set partitioning concepts introduced in Section 2.4 and the algebraic properties of the binary partitions in Section 2.5*. In particular, given sets of (Toeplitz) binary matrices $\mathcal{K}_l = \mathcal{K}_{\nu,d_l}, l = 1, \ldots, L$, which have rank guarantees of $\{d_l\}$, given the set-partitioning mapping f_2, we can lift the binary rank properties to the complex domain. Therefore, the main challenge, addressed in this paper, is the construction of such sets of binary matrices with rank guarantees.

Therefore the codewords from lth layer achieve a rate $R_l = \frac{1}{T}\log|\mathcal{K}_{\nu,d_l}|$ and diversity order $d_l(\nu + 1)M_r$. From Definition 5 it follows that the size of $\mathcal{K}_{\nu,d_l}$ can be made at least as large as $2^{T(M_t - d_l + 1) - \nu M_t}$. Similar to [6] this construction for QAM constellations achieves the rate-diversity tuple $(R_1, M_r d_1(\nu + 1), \ldots, R_L, M_r d_L(\nu + 1))$, with the overall equivalent single layer code achieving rate-diversity point, $(\sum_l R_l, M_r d_L(\nu + 1))$.

In particular, we can construct a space-time code by choosing identical diversity requirements for all the layers, *i.e.*, $d_1 = d_2 = \ldots = d_L$. From this we conclude that the rate diversity tradeoff for the ISI channel can be characterized as follows:

Theorem 8 (Rate Diversity Tradeoff for ISI Channels) *Consider transmission over a $\nu + 1$ tap ISI channel with M_t transmit antennas from a QAM or PSK signal constellation $\mathcal{A}$ with $|\mathcal{A}| = 2^L$ and communication over a time period T such that $T \geq T_{thr}$. For diversity order $D_{isi} = d(\nu + 1)M_r$, the rate diversity tradeoff is given by,*

$$(M_t - d + 1) - \frac{\nu}{T}M_t \leq R^{eff} \leq (M_t - d + 1)\,,$$

where R_{eff} is the effective rate of transmission which includes the overhead due to the zero padding.

The lower bound follows directly from Theorem 7 and the upper bound follows from Lemma 3. Note that the bounds in the above theorem are tight as $T \to \infty$.

4 Diversity Embedded Trellis Codes

The construction of diversity embedded trellis codes for ISI channels is quite similar to the construction of block codes. Again the idea is to construct binary convolutional codes with the following properties.

Definition 9 *Define $\mathcal{P}_{\nu,d} \subset \{\mathbf{B} : \mathbf{B} \in \mathbb{F}_2^{M_t \times T}\}$ to be the set of $M_t \times T$ binary matrices such that for $T \geq T_{thr}$ they satisfy the following properties:*

- *The last ν entries of all the M_t rows are equal to zero.*
- *For any distinct pair of matrices $\mathbf{A}, \mathbf{B} \in \mathcal{P}_{\nu,d}$ the rank of $\Theta_\nu(\mathbf{A}) - \Theta_\nu(\mathbf{B})$ is at least $d(\nu+1)$.*
- $\log|\mathcal{P}_{\nu,d}| \geq R(T - \nu - (\nu+1)(M_t-1)(2^R-1)(2^{R-1}))$, *where* $R = M_t - d + 1$.

Using these sets of matrices obtained by appropriately choosing the underlying convolutional codes the diversity embedding properties are ensured.

We will postpone the construction of such sets of binary matrices to Section 4.1 where using Lemma 6 along with particular choices of convolutional codes we show the following result.

Lemma 10 *For block size $T \geq T_{thr} = (2^R-1)\nu + (2^R-1)(\nu+1)\big((M_t-2)(2^R-1)+R\big)$, where $R = M_t - d + 1$, there exist sets of binary matrices $\mathcal{P}_{\nu,d}$ which satisfy the requirements of Definition 9.*

As in the case of block codes in Section 3, given a L-level binary partition of a QAM or PSK signal constellation, a diversity embedded convolutional space-time codeword is defined by an array $\mathbf{P} = \{\mathbf{P}^{(1)}, \mathbf{P}^{(2)}, \ldots, \mathbf{P}^{(L)}\}$ determined by a sequence of binary matrices where, matrix $\mathbf{P}^{(i)}$ specifies the space-time array at level i. Adapted easily from [6] we can state the formal construction guarantee for the diversity embedded trellis code for ISI channels as follows:

Theorem 11 *Let $\mathcal{C}$ be a multi-level space-time code for a QAM or PSK constellation of size 2^L with M_t transmit antennas that is determined by constituent sets of binary matrices $\mathcal{P}_l = \mathcal{P}_{\nu,d_l}, l = 1, \ldots, L$, such that $d_1 \geq d_2 \ldots \geq d_L$. For joint maximum-likelihood decoding, the input bits that select the codeword from the lth set $\mathcal{P}_l$ are guaranteed diversity $D_l = d_l(\nu+1)M_r$ in the complex domain when transmitted over an ISI channel with $\nu+1$ taps.*

The proof of Theorem 11 follows from the same techniques as in [6] by mapping binary matrices with desired rank guarantees to rank guarantees in complex domain. As in the proof of Theorem 7, the main difficulty is in constructing these sets of binary matrices with thee given rank guarantees, using convolutional codes. We give such a construction in Section 4.1. Therefore the codewords from lth layer achieve a rate $R_l = \frac{1}{T}\log|\mathcal{P}_{\nu,d_l}|$ and diversity order $d_l(\nu+1)M_r$. From Definition 9 it follows that the size of $\mathcal{P}_{\nu,d_l}$ can be made at least as large as $2^{R(T-\nu-(\nu+1)(M_t-1)(2^R-1)(2^{R-1}))}$, which in the limit as $T \to \infty$ tends to 2^R. Similar to [6] this construction for QAM constellations achieves the rate-diversity tuple $(R_1, M_r d_1(\nu+1), \ldots, R_L, M_r d_L(\nu+1))$, with the overall equivalent single layer code achieving rate-diversity point, $(\sum_l R_l, M_r d_L(\nu+1))$.

We illustrate the idea by examining the construction for each of the layers. The construction is shown in Figure 4. Given the input stream for each layer i, the first block in the figure maps these inputs to the coefficients of R_i polynomials $u_{i,j}(D), j = 1, \ldots, R_i$ in $\mathbb{F}_2[D]$. The second block multiplies the input vector $\{u_{i,j}(D)\}_{j=1}^{R_i}$ by the generator matrix $\mathbf{G}_i(D)$, with special structure which we define in the following subsection, and generates a vector $\mathbf{u}_i(D)$ of polynomials. The final block Ω then maps this vector $\mathbf{p}_i(D)$ to a binary matrix $\mathbf{P}_i \in \mathbb{F}_2^{M_t \times T}$.

We define the set $\mathcal{P}_i = \mathcal{P}_{\nu,i}$ to be the set of all output matrices $\mathbf{P}_i$ for all possible inputs on the stream i. Note that these sets satisfy the properties in Definition 9.

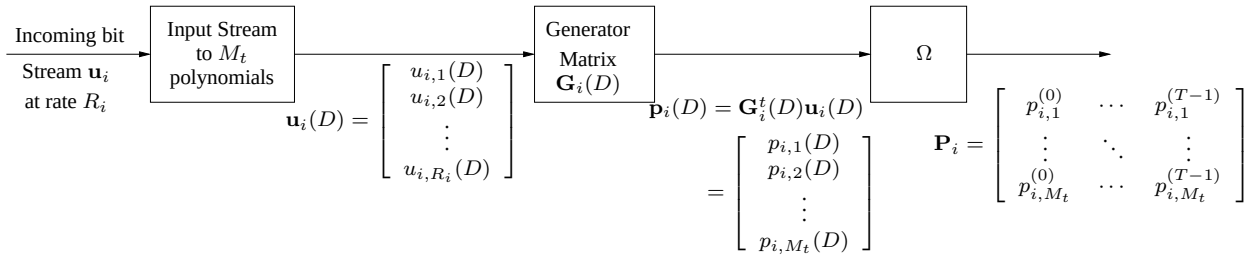

Figure 4: Binary matrices for each layer

4.1 Binary Convolutional Codes

Explicit construction of full diversity maximum rate binary convolutional codes was first shown in [12]. This was extended for general points on the rate diversity tradeoff for flat fading channels in [16]. We will give constructions for such sets of binary matrices for ISI channels in this section.

Consider the construction for a particular layer above. We will see the construction of rate R symbols per transmission, and rank distance of $(\nu+1)(M_t - R + 1)$ binary codes for transmission over the ISI channel. Represent the generator matrix or transfer function matrix $\mathbf{G}$ for this code by an $R \times M_t$ generator matrix given by,

$$\mathbf{G} = \begin{bmatrix} g_1^{(1)}(D) & g_1^{(2)}(D) & \cdots & g_1^{(M_t)}(D) \\ g_2^{(1)}(D) & g_2^{(2)}(D) & \cdots & g_2^{(M_t)}(D) \\ \vdots & \ddots & \ddots & \vdots \\ g_R^{(1)}(D) & g_R^{(2)}(D) & \cdots & g_R^{(M_t)}(D) \end{bmatrix}. \tag{23}$$

Denoting $\xi = D^{(\nu+1)(2^R-1)}$ we choose

$$g_l^{(q)}(D) = \xi^{(q-1)2^{(l-1)}}. \tag{24}$$

The input message polynomial is represented by the vector of message polynomial

$$\mathbf{u}(D) = \begin{bmatrix} u_1(D) & u_2(D) & \cdots & u_R(D) \end{bmatrix}^t, \tag{25}$$

where $u_i(D) \in \mathbb{F}_2[D]$. The code polynomial vector is given by

$$\begin{aligned} \mathbf{p}(D) &= \mathbf{G}^t(D)\mathbf{u}(D) \\ &= \begin{bmatrix} p_1(D) & p_2(D) & \cdots & p_{M_t}(D) \end{bmatrix}^t. \end{aligned} \tag{26}$$

The $M_t \times T$ code matrix which is actually transmitted on the antenna is given by

$$\mathbf{P} = \begin{bmatrix} p_1^0 & \cdots & p_1^{T-1} \\ \vdots & \ddots & \vdots \\ p_{M_t}^0 & \cdots & p_{M_t}^{T-1} \end{bmatrix} \tag{27}$$

where p_i^j is the j^{th} coefficient of the polynomial p_i in (26). We make a distinction between $\mathbf{p}(D)$ which is a vector of polynomials in D, and $\mathbf{P}$ which is a binary matrix. This mapping is denoted by Ω *i.e.* $\Omega(\mathbf{p}(D)) = \mathbf{P}$

Note that in order that the matrix $\mathbf{P}$ satisfies the structure in (21) we require the ν largest coefficients of each $p_i(D)$ in (26) to be zero, *i.e.*,

$$p_i^j = 0 \qquad \forall\, i \in \{1, \ldots, M_t\} \text{ and } \forall\, j \in \{T-\nu-1, \ldots, T-1\}. \tag{28}$$

With this constraint we get that,

$$\begin{aligned} deg(u_i(D)) &\leq T - 1 - \nu - \max_{u,v} deg(g_u^{(v)}) \\ &= T - 1 - \nu - (\nu+1)(M_t - 1)(2^R - 1)(2^{R-1}), \end{aligned}$$

where the last equality follows from our particular choice of $g_u^{(v)}$ given in (24). Note that this convolutional code corresponds to an effective rate of

$$\begin{aligned} R^{eff} &= \frac{\log\left(2^{T-1-\nu-(\nu+1)(M_t-1)(2^R-1)(2^{R-1})+1} \cdot 2^R\right)}{T} \\ &= \frac{R\left(T-\nu-(\nu+1)(M_t-1)(2^R-1)(2^{R-1})\right)}{T} \quad bits/Tx, \end{aligned} \tag{29}$$

which asymptotically tends to R as $T \to \infty$.Also observe that,

$$\Theta(\mathbf{P}) = \begin{bmatrix} \Omega(\mathbf{p}(D)) \\ \Omega(D\mathbf{p}(D)) \\ \vdots \\ \Omega(D^\nu \mathbf{p}(D)) \end{bmatrix}. \tag{30}$$

From this we can conclude that,

$$\Theta(\mathbf{P})) = \Omega\left(\tilde{\mathbf{G}}^t(D)\mathbf{u}(D)\right)$$

where $\tilde{\mathbf{G}} \in \mathbb{F}_2^{R\times(\nu+1)M_t}$ is given by,

$$\tilde{\mathbf{G}} = \begin{bmatrix} g_1^{(1)}(D) & g_1^{(2)}(D) & \cdots & g_1^{(M_t)}(D) & Dg_1^{(1)}(D) & \cdots & Dg_1^{(M_t)}(D) & \cdots & D^\nu g_1^{(1)}(D) & \cdots & D^\nu g_1^{(M_t)}(D) \\ g_2^{(1)}(D) & g_2^{(2)}(D) & \cdots & g_2^{(M_t)}(D) & Dg_2^{(1)}(D) & \cdots & Dg_2^{(M_t)}(D) & \cdots & D^\nu g_2^{(1)}(D) & \cdots & D^\nu g_2^{(M_t)}(D) \\ \vdots & \ddots & \ddots & \vdots & & & & & & & \\ g_R^{(1)}(D) & g_R^{(2)}(D) & \cdots & g_R^{(M_t)}(D) & Dg_R^{(1)}(D) & \cdots & Dg_R^{(M_t)}(D) & \cdots & D^\nu g_R^{(1)}(D) & \cdots & D^\nu g_R^{(M_t)}(D) \end{bmatrix}. \tag{31}$$

With our particular choice of $g_l^{(q)}(D)$, given in (24), we can write this as,

$$\tilde{\mathbf{G}}^t = \begin{bmatrix} 1 & \dots & 1 & 1 \\ \xi & \xi^2 & \dots & \xi^{2^{R-1}} \\ \xi^2 & (\xi^2)^2 & \dots & (\xi^2)^{2^{R-1}} \\ \vdots & & \vdots & \\ \xi^{(M_t-1)} & (\xi^{(M_t-1)})^2 & \dots & (\xi^{(M_t-1)})^{2^{R-1}} \\ D & \dots & D & D \\ & \vdots & & \vdots \\ D\xi^{(M_t-1)} & D(\xi^{(M_t-1)})^2 & \dots & D(\xi^{(M_t-1)})^{2^{R-1}} \\ & \vdots & & \vdots \\ & \vdots & & \vdots \\ D^\nu\xi^{(M_t-1)} & D^\nu(\xi^{(M_t-1)})^2 & \dots & D^\nu(\xi^{(M_t-1)})^{2^{R-1}} \end{bmatrix}. \tag{32}$$

Define the polynomial

$$f(x) = \sum_{l=0}^{R-1} u_l(D)x^{2^l}, \tag{33}$$

where $\{u_l(D)\}_{l=0}^{R-1} \in \mathbb{F}_2[D]$. Then from (32) with $\xi = D^{(\nu+1)(2^R-1)}$ we have,

$$\tilde{\mathbf{G}}^t(D)\mathbf{u}(D) = \begin{bmatrix} f(1) \\ f(\xi) \\ \vdots \\ f(\xi^{(M_t-1)}) \\ Df(1) \\ \vdots \\ D^\nu f(\xi^{(M_t-1)}) \end{bmatrix}.$$

The proof now that the left null space of $\Omega\left(\tilde{\mathbf{G}}^t(D)\mathbf{u}(D)\right)$ over $\mathbb{F}_2$ is of dimension at most $d(\nu+1)$ is the same as the proof of Theorem 21 by choosing T such that,

$$T \geq (2^R-1)\nu + (2^R-1)(\nu+1)\left((M_t-2)(2^R-1)+R\right).$$

Therefore, given the result of Theorem 21, which is proved in Section 6.3, we can prove the rank guarantees of the convolutional codes.

5 Rate Guarantees

In this section we will give background needed for construction of binary codes $\mathcal{K}_{\nu,d}$ with properties given in Definition 5. We start in Section 5.1 with a representation of $\mathcal{K}_{\nu,d}$ in terms of polynomials over $\mathbb{F}_{2^T}$ which will be useful in proving the construction of binary codes $\mathcal{K}_{\nu,d}$. In Section 5.2 we will list some definitions which we will use in proving rank guarantees in Section 6. Note that these definitions are not required for constructing $\mathcal{K}_{\nu,M_t}$, *i.e.*, maximal rank sets, for which the proof is much simpler as seen in Section 6.1. Finally in Section 5.3 we will show that $|\mathcal{K}_{\nu,d}| \geq 2^{RT-\nu M_t}$, where $R = M_t - d + 1$. The rank properties of $\mathcal{K}_{\nu,d}$ are given in Section 6.

5.1 Polynomial representation

Given a rate R, we define the linearized polynomial

$$f(x) = \sum_{l=0}^{R-1} f_l x^{2^l}, \tag{34}$$

where $\{f_l\}_{l=0}^{R-1} \in \mathbb{F}_{2^T}$. To develop the binary matrices with structure given in (21), we define $\mathbf{c}_f \in \mathbb{F}_{2^T}^{M_t}$ as

$$\mathbf{c}_f = \begin{bmatrix} f(1) & f(\xi) & \dots & f(\xi^{(M_t-1)}) \end{bmatrix}^t, \tag{35}$$

where $\xi = \alpha^{(2^R-1)(\nu+1)}$, and α is a primitive element of $\mathbb{F}_{2^T}$. Let $\mathbf{f}^{(0)}(\xi^i)$ and $\mathbf{f}^{(k)}(\xi^i)$ be the representations of $f(\xi^i)$ and $\alpha^k f(\xi^i)$ in the basis $\{\alpha^0, \alpha^1, \dots, \alpha^{T-1}\}$ respectively *i.e.*, $\mathbf{f}^{(0)}(\xi^i), \mathbf{f}^{(k)}(\xi^i) \in \mathbb{F}_2^{1\times T}$. We obtain a matrix representation $\mathbf{C}_f \in \mathbb{F}_2^{M_t\times T}$ of $\mathbf{c}_f$ as,

$$\mathbf{C}_f = \begin{bmatrix} \mathbf{f}^{(0)t}(1) & \mathbf{f}^{(0)t}(\xi) & \dots & \mathbf{f}^{(0)t}(\xi^{(M_t-1)}) \end{bmatrix}^t. \tag{36}$$

Now, in order to get the structure required in (21), we need to study the requirements of f so that the last ν elements in $\mathbf{C}_f$ are 0 for all the M_t rows. Note that the j^{th} row of $\mathbf{C}_f$ is given by the binary expansion of $f(\xi^{j-1}) \in \mathbb{F}_{2^T}$ in terms of the basis $\{\alpha^0, \alpha, \dots, \alpha^{T-1}\}$, where α is a primitive element of $\mathbb{F}_{2^T}$. The coefficients in this basis expansion can be obtained using the trace operator described below for completeness[5].

Consider an extension field $\mathbb{F}_{2^T}$ of the base field $\mathbb{F}_2$. If $\alpha \in \mathbb{F}_{2^T}$ is a primitive element of $\mathbb{F}_{2^T}$ then $(\alpha^0, \alpha^1, \dots, \alpha^{T-1})$ form a basis of $\mathbb{F}_{2^T}$ over $\mathbb{F}_2$ and any element $\beta \in \mathbb{F}_{2^T}$ can be uniquely represented in the form,

$$\beta = \beta_0\alpha^0 + \beta_1\alpha^1 + \dots \beta_{T-1}\alpha^{T-1} \quad \text{with } \beta_i \in \mathbb{F}_2, \text{ for } 0 \le i \le (T-1).$$

To solve for the coefficients β_i we will use the trace function and trace dual bases. Note that for any element $\beta \in \mathbb{F}_{2^T}$ the trace of the element β relative to the base field $\mathbb{F}_2$ is defined as,

$$\mathrm{Tr}_{2^T/2}(\beta) = \beta + \beta^2 + \beta^{2^2} + \dots + \beta^{2^{T-1}}.$$

Given that $\beta, \tilde{\beta} \in \mathbb{F}_{2^T}$ the trace function satisfies the following properties,

- $\mathrm{Tr}_{2^T/2}(\beta) \in \mathbb{F}_2$.
- $\mathrm{Tr}_{2^T/2}(\beta + \tilde{\beta}) = \mathrm{Tr}_{2^T/2}(\beta) + \mathrm{Tr}_{2^T/2}(\tilde{\beta})$.
- $\mathrm{Tr}_{2^T/2}(\lambda\beta) = \lambda\mathrm{Tr}_{2^T/2}(\beta)$, if $\lambda \in \mathbb{F}_2$.

Given the basis $(\alpha^0, \alpha^1, \dots, \alpha^{T-1})$ the corresponding trace dual basis $(\theta_0, \theta_1, \dots, \theta_{T-1})$ is defined to be the unique set of elements which satisfy the following relation for $0 \le i, j \le (T-1)$,

$$\mathrm{Tr}_{2^T/2}\left(\theta_i\alpha^j\right) = \begin{cases} 0 & \text{for } i \ne j \\ 1 & \text{for } i = j. \end{cases}$$

[5]More background can be found in standard textbooks on finite fields [14, 17].

The fact that the trace dual basis exists and is unique can be found in standard references such as [14, 17]. Therefore given $\beta \in \mathbb{F}_{2^T}$, we can find β_i by using the properties of the trace function and noting that,

$$\mathrm{Tr}_{2^T/2}\left(\theta_i \beta\right) = \mathrm{Tr}_{2^T/2}\left(\theta_i \sum_{j=0}^{T-1} \beta_j \alpha^j\right) = \sum_{j=0}^{T-1} \mathrm{Tr}_{2^T/2}\left(\theta_i \beta_j \alpha^j\right) = \sum_{j=0}^{T-1} \beta_j \mathrm{Tr}_{2^T/2}\left(\theta_i \alpha^j\right) = \beta_i,$$

where the last equality follows from the definition of the trace dual basis. Therefore, binary matrix $\mathbf{B}$ given in (21) can be represented in terms of the set $\mathcal{S}$ defined as

$$\mathcal{S} = \left\{ f : f \in \mathbb{F}_{2^T}^R, \mathrm{Tr}_{2^T/2}\left(\theta_i f(\xi^j)\right) = 0 \; \forall i \in \{T-\nu, \ldots, T-1\} \text{ and } \forall j \in \{1, \ldots, M_t\} \right\}. \tag{37}$$

Associate to $f \in \mathcal{S}$ the codeword vector $\mathbf{u}_f \in \mathbb{F}_{2^T}^{(\nu+1)M_t \times 1}$ given by,

$$\mathbf{u}_f = \left[\begin{array}{cccc} \mathbf{c}_f^t & \alpha \mathbf{c}_f^t & \ldots & \alpha^\nu \mathbf{c}_f^t \end{array} \right]^t. \tag{38}$$

Associate with every such codeword $\mathbf{u}_f$ the codeword matrix $\mathbf{U}_f \in \mathbb{F}_2^{(\nu+1)M_t \times T}$ given by the representation of each element of $\mathbf{u}_f$ in the basis $\{\alpha^0, \alpha^1, \ldots, \alpha^{T-1}\}$.

Since $f \in \mathcal{S}$ we know that the last ν elements in $\mathbf{C}_f$ are 0 for all the M_t rows. Therefore we can see that $\mathbf{f}^{(k)}(\xi^i)$ is a cyclic shift by k positions of $\mathbf{f}(\xi^i)$. Hence, for $i \in \{0, 1, \ldots, \nu\}$ we can write,

$$\mathbf{C}_f^{(i)} = \left[\begin{array}{cccc} \mathbf{f}^{(i)t}(1) & \mathbf{f}^{(i)t}(\xi) & \ldots & \mathbf{f}^{(i)t}(\xi^{(M_t-1)}) \end{array} \right]^t, \tag{39}$$

where $\mathbf{C}_f^{(i)}$ represents the matrix obtained by a cyclic shift of all the rows of the matrix $\mathbf{C}_f$ by i positions.

For transmission over an ISI channel, as seen in Section 3.2, it can be shown from equation (2) that the effective binary transmitted codeword matrix for a particular f is of the form

$$\mathbf{U}_f = \left[\begin{array}{cccc} \mathbf{C}_f^t & \mathbf{C}_f^{(1)t} & \ldots & \mathbf{C}_f^{(\nu)t} \end{array} \right]^t. \tag{40}$$

We will show in Section 6.3 that indeed for $R = M_t - d + 1$, $\mathcal{K}_{\nu,d} = \{\mathbf{C}_f : f \in \mathcal{S}\}$, *i.e.*, $\mathrm{rank}(\mathbf{U}_f) \geq d(\nu+1), \forall f \in \mathcal{S}$.

5.2 Notation and Definitions

We will need the following definitions in the construction of the basis vectors of the null space of $\mathbf{U}_f$.

1. We define a set $\Gamma \subseteq \mathbb{F}_{2^T}$ which will be used extensively in the proof in Section 6 as,

$$\Gamma = \{\gamma \in \mathbb{F}_{2^T} : \gamma = \sum_{t=0}^{\nu} \delta_t \alpha^t, \delta_t \in \mathbb{F}_2\}. \tag{41}$$

2. Given a binary vector $\mathbf{b} \in \mathbb{F}_2^{(\nu+1)M_t \times 1}$ define $\Psi : \mathbb{F}_2^{(\nu+1)M_t \times 1} \rightarrow \Gamma^{1 \times M_t}$ as,

$$\Psi(\mathbf{b}) = \underbrace{\left[\begin{array}{ccc} \sum_{i=0}^{\nu} b_{iM_t+1}\alpha^i & \ldots & \sum_{i=0}^{\nu} b_{iM_t+M_t}\alpha^i \end{array} \right]}_{\mathbf{g}}. \tag{42}$$

Note that the mapping Ψ is a one-to-one mapping between $\mathbf{b}$ and $\mathbf{g}$, due to the linear independence of $\{\alpha^0, \alpha, \ldots, \alpha^\nu\}$.

3. For a given fixed $\mathbf{c}_f \in \mathbb{F}_{2^T}^{M_t \times 1}$ define $\mathcal{G}_f \subseteq \Gamma^{1\times M_t}$ such that,

$$\mathcal{G}_f = \{\mathbf{g} \in \Gamma^{1\times M_t} : \mathbf{g}\mathbf{c}_f \quad = 0\} \tag{43}$$

4. Motivated by the mapping in (42), for each $\mathbf{g}^{(i)} \in \mathcal{G}_f$ we will use the following representation:

$$\begin{aligned} \mathbf{g}^{(i)} &= \begin{bmatrix} g_0^{(i)} & \dots & g_{M_t-1}^{(i)} \end{bmatrix} \\ g_k^{(i)} &= \sum_{j=0}^{\nu} \delta_{k,j}^{(i)} \alpha^j \quad \text{where } \delta_{k,j}^{(i)} \in \mathbb{F}_2. \end{aligned} \tag{44}$$

5. For an element $\gamma \in \Gamma$ given by $\gamma = \sum_{j=0}^{\nu} \delta_j \alpha^j$, define

$$deg(\gamma) = \max_j \{j : \delta_j \neq 0\}. \tag{45}$$

6. For each $\mathbf{g} \in \mathcal{G}_f$ define,

$$deg(\mathbf{g}) = \max_k \{j : \delta_{k,j} \neq 0\}. \tag{46}$$

7. For each $\mathbf{g}^{(i)} \in \mathcal{G}_f$ define a function $\Phi : \Gamma^{1\times M_t} \to \mathbb{F}_2^{1\times M_t}$ by,

$$\Phi(\mathbf{g}^{(i)}) \quad = \begin{bmatrix} \delta_{0,0}^{(i)} & \delta_{1,0}^{(i)} & \dots & \delta_{M_t-1,0}^{(i)} \end{bmatrix}. \tag{47}$$

8. Given a set of elements $\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \dots, \mathbf{g}^{(d)} \in \Gamma^{1\times M_t}$ define,

$$\mathcal{D}\left(\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \dots, \mathbf{g}^{(d)}\right) \quad = \left\{\mathbf{g} : \mathbf{g} = \textstyle\sum_{i=1}^{d} \gamma_i \mathbf{g}^{(i)}, \text{ where for all } i,\ \gamma_i \in \Gamma,\ \gamma_i \mathbf{g}^{(i)} \in \Gamma^{1\times M_t}\right\}. \tag{48}$$

Note that it then directly follows that,

$$|\mathcal{D}\left(\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \dots, \mathbf{g}^{(d)}\right)| \quad \leq 2^{d(\nu+1)}. \tag{49}$$

5.3 Set cardinality

Using the polynomial representation given in Section 5.1, we can give a lower bound on the rate as follows:

Theorem 12 *If $T > (\nu+1)M_t$ then a lower bound to the cardinality of the set $\mathcal{S}$ is given by $|\mathcal{S}| \geq 2^{RT-\nu M_t}$ or lower bound to effective rate is given by,*

$$R_{eff} = \frac{1}{T}\log|\mathcal{S}| R_{eff} \geq R - \frac{\nu M_t}{T}. \tag{50}$$

Proof. Let $\lambda_{\theta_i,\beta_j}$ be the mapping,

$$\lambda_{\theta_i,\beta_j} : \begin{bmatrix} f_{R-1} & \dots & f_1 & f_0 \end{bmatrix}^t \mapsto \mathrm{Tr}_{2^T/2}(\theta_i f(\beta_j)),$$

for some $\beta_j \in \mathbb{F}_{2^T}, j = 1, \ldots, M_t$. This is homomorphism of the $\mathbb{F}_2$-vector space $\mathbb{F}_{2^T}^R$ into $\mathbb{F}_2$. The cardinality of the set $\mathcal{S}$ is given by,

$$|\mathcal{S}| = \left| \bigcap_{i,j} \ker(\lambda_{\theta_i,\beta_j}) \right| \; i \in \{T-\nu, \ldots, T-1\} \; \& \; j \in \{1, \ldots, M_t\}.$$

Note that the range space of $\lambda_{\theta_i,\beta_j}$ is the range of the trace function, *i.e.*, $\{0, 1\}$. Since $T > (\nu+1)M_t$ and the rank of the equivalent matrix transformation of $[\lambda_{\theta_{T-\nu},\beta_0}, \ldots, \lambda_{\theta_{T-1},\beta_{M_t-1}}]^t$ at most νM_t, by the rank nullity theorem it follows that the null space is of dimension at least $RT - \nu M_t$. Therefore the cardinality of the set $\mathcal{S}$ is lower bounded by,

$$|\mathcal{S}| \geq 2^{RT-\nu M_t}.$$

■

An intuition for the above result can be obtained as follows. Note that of the total possible 2^{RT} polynomials, it can be easily shown that for any entry of the matrix $\mathbf{C}_f$, for exactly half of the 2^{RT} polynomials the value will be equal to 0. Therefore, requiring ν rows of any particular row of $\mathbf{C}_f$ to be zero reduces the admissible f to $2^{RT-\nu}$. Equivalently it reduces the dimension of admissible f from RT to $RT-\nu$. If these polynomials evaluate to 0 for all the remaining M_t rows as well, then it is possible to conclude that the size of the set $\mathcal{S}$ is equal to $2^{RT-\nu}$. Since this is not necessarily true, taking intersection of M_t subspaces of the admissible polynomials, each of dimension $(RT-\nu)$, it follows that $|\mathcal{S}| \geq 2^{RT-\nu M_t}$.

The Theorem 12 implies that we do not lose too much, in terms of rate, by the zero padding at the end of the transmission block. In particular it is a constant factor which does not depend on T and therefore can be made small by taking large enough T. Note that this lower bound could be loose, and we may not lose as much rate as $\frac{\nu M_t}{T}$.

We still need to show that this set satisfies the rank guarantees, which we will do next in Section 6.

6 Rank Guarantees

In Section 5, see (37), we have already constructed codes (binary sets) $\mathcal{S}$ which satisfy the structure in (21) and that $|\mathcal{S}| \geq 2^{T(M_t-d+1)-\nu M_t}$. Therefore, this set $\mathcal{S}$ is a good candidate for the construction of $\mathcal{K}_{\nu,d}$, needed for the multilevel construction of Section 3.2. In this section we will prove that the set $\mathcal{S}$ in (37) also satisfies the rank guarantees given in Definition 5 and hence proving Lemma 6. To illustrate the proof techniques, we will first prove the rank guarantees for the maximal rank binary codes *i.e.*, $\mathcal{K}_{\nu,M_t}$ in Section 6.1. However, the argument for arbitrary rank needs a more sophisticated argument. We will explore the structure of the null space of $\mathbf{U}_f$ and find a basis for it in Section 6.2. Using the structure of the basis we will finally bound the cardinality and dimension of the null space giving the required rank guarantees for $\mathcal{K}_{\nu,d}$ with $T_{thr} = R\nu + (M_t-1)(\nu+1)(2^R-1)$.

6.1 Maximal rank distance codes

In this section we will show that that if $R = 1$ then for all $f \in \mathcal{S}, \text{rank}(\mathbf{U}_f) \geq M_t(\nu+1)$. In fact for this case $T_{thr} = M_t(\nu+1)$ is enough.

Theorem 13 ((Maximal rank distance codes)) *With $R = 1$ let $f(x) = f_0 x$, as in (34), and $T \geq M_t(\nu + 1)$. Then for $\mathcal{S}$ defined in (37), $\frac{1}{T}\log|\mathcal{S}| \geq 1 - \frac{\nu M_t}{T}$ and $\forall f \in \mathcal{S}$, rank$(\mathbf{U}_f) \geq M_t(\nu+1)$ over the binary field.*

Proof. The rate lower bound is directly from Theorem 12. We prove the result by contradiction. Suppose that $\mathcal{O} = \{\mathbf{U}_f : f \in \mathcal{S}\}$ has rank distance less than $(\nu+1)M_t$. Then there exists a vector $\mathbf{u}_f \neq \mathbf{0}$ for some $f \in \mathcal{S}$ such that the corresponding binary matrix $\mathbf{U}_f$ has binary rank less than $(\nu+1)M_t$ (as the code is linear). So there exists a non-trivial binary vector space $\mathcal{B} \subseteq \mathbb{F}_2^{(\nu+1)M_t}$ such that for every $\mathbf{b} \in \mathcal{B}$,

$$\mathbf{b}^t\mathbf{U}_f \;=\; \mathbf{0} \iff \sum_{i=1}^{(\nu+1)M_t} b_i \mathbf{U}_f(i,j) = 0, \;\; j = 1, \ldots, T, \tag{51}$$

where $\mathbf{U}_f(i,j)$ is the $(i,j)^t h$ entry of $\mathbf{U}_f$ and we have used $(\cdot)^t$ to denote vector transpose. Since each row of $\mathbf{U}_f$ is an expansion of the rows of $\mathbf{u}_f$ in the basis $\{\alpha^0, \alpha, \ldots, \alpha^{T-1}\}$, using the basis expansion this can be written as operations over $\mathbb{F}_{2^T}$ as follows:

$$\mathbf{b}^t\mathbf{u}_f \;=\; \sum_{i=1}^{(\nu+1)M_t} b_i \mathbf{u}_f(i) = \sum_{i=1}^{(\nu+1)M_t} b_i \sum_{j=1}^{T} \mathbf{U}_f(i,j)\alpha^{j-1} = \sum_{j=1}^{T} \alpha^{j-1} \left[\sum_{i=1}^{(\nu+1)M_t} b_i \mathbf{U}_f(i,j) \right]. \tag{52}$$

Due to the linear independence of $\{\alpha^0, \alpha, \ldots, \alpha^{T-1}\}$ it is clear from (51) and (52) that,

$$\mathbf{b}^t\mathbf{U}_f = 0 \iff \mathbf{b}^t\mathbf{u}_f = 0. \tag{53}$$

Now, we suppose that for $\mathbf{b} \neq \mathbf{0}$,

$$\begin{aligned}\mathbf{b}^t\mathbf{u}_f &= \sum_{i=0}^{\nu}\sum_{k=0}^{M_t-1} b_{i+k(\nu+1)}\alpha^i f(\alpha^{k(\nu+1)}) = \sum_{i=0}^{\nu}\sum_{k=0}^{M_t-1} b_{i+k(\nu+1)}\alpha^i f_0 \alpha^{k(\nu+1)} \\ &= f_0\left(\sum_{i=0}^{\nu}\sum_{k=0}^{M_t-1} b_{i+k(\nu+1)}\alpha^{i+k(\nu+1)}\right) = 0.\end{aligned} \tag{54}$$

Thus, for every $\mathbf{b} \in \mathcal{B}$ the element $\left(\sum_{i=0}^{\nu}\sum_{k=0}^{M_t-1} b_{i+k(\nu+1)}\alpha^{i+k(\nu+1)}\right)$ is a zero of $f(x)$. But we know that $\{\alpha^{i+k(\nu+1)}\}$ are linearly independent for $k \in \{0,1,\ldots,M_t-1\}$ and $i \in \{0,1,\ldots,\nu\}$ as $T \geq (\nu+1)M_t$. Therefore there is only one trivial solution to the equation (54) *i.e.*, $b_{i+k(\nu+1)} = 0$ for $i = \{0,\ldots,\nu\}, k = \{1,\ldots,\nu\}$. This contradicts the fact that the null space is non-trivial since we cannot have $\mathbf{b} \neq \mathbf{0}$ and $\mathbf{b} \in \mathcal{B}$. Hence all matrices in $\mathcal{O}$ have rank equal to $M_t(\nu+1)$. ∎

6.2 Minimal Basis Vectors

To prove the rank distance properties in this subsection we will show the existence of elements which satisfy the following properties.

Definition 14 (Properties of Minimal Basis Vectors) *Given a fixed nonzero vector $\mathbf{c}_f \in \mathbb{F}_{2^T}^{M_t \times 1}$ define the associated $\mathcal{G}_f$ as in equation (43). Then the elements $\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \ldots, \mathbf{g}^{(d)} \in \mathcal{G}_f$ are called the minimal basis vectors if they satisfy the following properties:*

(i). For each $\mathbf{g}^{(i)}$*,* $\exists\ k$ *such that* $\delta_{k,0}^{(i)} = 1$*,* i.e.*,* $\Phi(\mathbf{g}^{(i)}) \neq \mathbf{0}$*.*

(ii). $\Phi(\mathbf{g}^{(1)}), \ldots, \Phi(\mathbf{g}^{(d)})$ *are linearly independent over* $\mathbb{F}_2$*.*

(iii). For every subset $S \subseteq \{1, \ldots, d\}$*, there do not exist* $\{\gamma_i\}$ *such that,*

$$\left\{\gamma_i : i \in S, \gamma_i \in \Gamma \text{ and } \gamma_i \mathbf{g}^{(i)} \in \Gamma^{1\times M_t}\right\},$$

and

$$deg(\sum_{i\in S} \gamma_i \mathbf{g}^{(i)}) < \max_{i\in S} deg(\gamma_i \mathbf{g}^{(i)}),$$

are simultaneously satisfied.

(iv). We have,

$$\mathcal{G}_f = \mathcal{D}\left(\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \ldots, \mathbf{g}^{(d)}\right), \tag{55}$$

where $\mathcal{D}(\cdot, \ldots, \cdot)$ *is defined as in (48) to be,*

$$\mathcal{D}\left(\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \ldots, \mathbf{g}^{(d)}\right) = \left\{\mathbf{g} : \mathbf{g} = \textstyle\sum_{i=1}^{d} \gamma_i \mathbf{g}^{(i)},\ \textit{where for all } i,\ \gamma_i \in \Gamma,\ \gamma_i \mathbf{g}^{(i)} \in \Gamma^{1\times M_t}\right\}. \tag{56}$$

To prove the existence of such minimal basis vectors, we need the following lemmas. We state the Lemma 15 required in the proofs and then prove it in the appendix.

Lemma 15 *Assume there exist* p *elements* $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)} \in \mathcal{G}_f$ *which do not satisfy property (iii)* i.e.*, for some subset* $S \subseteq \{1, \ldots, p\}$ *there exist* $\{\gamma_i\}$ *such that*

$$\left\{\gamma_i : i \in S, \gamma_i \in \Gamma \text{ and } \gamma_i \mathbf{g}^{(i)} \in \Gamma^{1\times M_t}\right\}$$

and,

$$deg(\sum_{i\in S} \gamma_i \mathbf{g}^{(i)}) < \max_{i\in S} deg(\gamma_i \mathbf{g}^{(i)})$$

are simultaneously satisfied. Then there exists a set $S' \subseteq S$ *and* $k \in S$*,* $k \notin S'$ *such that,*

$$deg\left(\mathbf{g}^{(k)} + \sum_{i\in S'} \gamma_i' \mathbf{g}^{(i)}\right) < deg\left(\mathbf{g}^{(k)}\right)$$

and

$$deg\left(\gamma_i' \mathbf{g}^{(i)}\right) \leq deg\left(\mathbf{g}^{(k)}\right) \quad \forall i \in S',$$

where $\gamma_i' \mathbf{g}^{(i)} \in \Gamma^{1\times M_t}$ *for all* $i \in S'$*.*

Lemma 16 *If there exist* p *elements* $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)} \in \mathcal{G}_f$ *satisfying (i), (ii) and (iii) in Definition 14 but not satisfying (iv) then it is possible to form* $\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p)}, \tilde{\mathbf{g}}^{(p+1)}$ *satisfying (i), (ii) and,*

$$\mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)}) \subset \mathcal{D}(\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p)}, \tilde{\mathbf{g}}^{(p+1)}). \tag{57}$$

Proof. ***Existence of element:*** Since $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)}$ satisfy (i), (ii) and (iii) but not (iv), there exists $\mathbf{g}^{(p+1)} \in \mathcal{G}_f$ such that $\mathbf{g}^{(p+1)} \notin \mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)})$. If $\Phi(\mathbf{g}^{(p+1)}) = \mathbf{0}$, then clearly $\mathbf{g}^{(p+1)} = \alpha^t \breve{\mathbf{g}}^{(p+1)}$, where $\Phi(\breve{\mathbf{g}}^{(p+1)}) \neq \mathbf{0}$, since only the common $\alpha^{(\cdot)}$ factor is taken out of $\mathbf{g}^{(p+1)}$. Note that clearly $\breve{\mathbf{g}}^{(p+1)} \in \mathcal{G}_f$ and since $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)}$ satisfy (iii), it can be shown that $\breve{\mathbf{g}}^{(p+1)} \notin \mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)})$.[6]

Construction of new elements satisfying (i): If $\Phi(\mathbf{g}^{(p+1)})$ is linearly independent of $\Phi(\mathbf{g}^{(1)}), \ldots, \Phi(\mathbf{g}^{(p)})$ then $(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p+1)})$ satisfy (i) and (ii) and (57) follows directly by choosing $\tilde{\mathbf{g}}^{(i)} = \mathbf{g}^{(i)}, i = 1, \ldots p+1$.

If $\Phi(\mathbf{g}^{(1)}), \ldots, \Phi(\mathbf{g}^{(p+1)})$ are not linearly independent then,

$$w_1 \Phi(\mathbf{g}^{(1)}) + w_2 \Phi(\mathbf{g}^{(2)}) + \ldots w_{p+1} \Phi(\mathbf{g}^{(p+1)}) = \mathbf{0} \tag{58}$$

for $w_1, \ldots, w_{p+1} \in \mathbb{F}_2$ and not all equal to zero. Let $\mathbf{g}^{(k)}$ be such that $w_k \neq 0$ and

$$deg(\mathbf{g}^{(k)}) \geq deg(\mathbf{g}^{(i)}) \quad \forall i, \text{ such that } w_i = 1. \tag{59}$$

Since $w_1, \ldots, w_{p+1} \in \mathbb{F}_2$, it can be seen that,

$$\begin{aligned} w_1 \Phi(\mathbf{g}^{(1)}) + w_2 \Phi(\mathbf{g}^{(2)}) + \ldots w_{p+1} \Phi(\mathbf{g}^{(p+1)}) &= \Phi(w_1 \mathbf{g}^{(1)} + \ldots + w_{p+1} \mathbf{g}^{(p+1)}) \\ &= \mathbf{0}. \end{aligned}$$

Therefore, there is a common $\alpha^{(\cdot)}$ factor in $w_1 \mathbf{g}^{(1)} + \ldots + w_{p+1} \mathbf{g}^{(p+1)}$ *i.e.,* there is $\tilde{\mathbf{g}}^{(k)}$ and t such that,

$$\left(w_1 \mathbf{g}^{(1)} + \ldots + w_k \mathbf{g}^{(k)} + \ldots + w_{p+1} \mathbf{g}^{(p+1)}\right) = \alpha^t \tilde{\mathbf{g}}^{(k)},$$

where t is chosen to be the minimum value such that $\Phi(\tilde{\mathbf{g}}^{(k)}) \neq \mathbf{0}$. Using this define,

$$\tilde{\mathbf{g}}^{(k)} = \alpha^{-t} \left(w_1 \mathbf{g}^{(1)} + \ldots + \mathbf{g}^{(k)} + \ldots + w_{p+1} \mathbf{g}^{(p+1)}\right) \tag{60}$$

$$\tilde{\mathbf{g}}^{(i)} = \mathbf{g}^{(i)} \quad \forall i \neq k, \tag{61}$$

where the fact that $w_k = 1$ has been used. Note that from (59) and (60)

$$deg(\tilde{\mathbf{g}}^{(k)}) \leq deg(\mathbf{g}^{(k)}) - t. \tag{62}$$

Clearly (i) is satisfied for $\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p+1)}$. Moreover, $\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p+1)} \in \mathcal{G}_f$ since $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p+1)} \in \mathcal{G}_f$.

Proof of containment (57): It will now be shown that,

$$\mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p+1)}) \subset \mathcal{D}(\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p+1)}). \tag{63}$$

Let $\mathbf{g} \in \mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p+1)})$, *i.e.,*

$$\mathbf{g} = \gamma_1 \mathbf{g}^{(1)} + \ldots + \gamma_{p+1} \mathbf{g}^{(p+1)} \tag{64}$$

[6] Assume that $\breve{\mathbf{g}}^{(p+1)} \in \mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)})$ but $\mathbf{g}^{(p+1)} \notin \mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)})$. Since $\mathbf{g}^{(p+1)} \in \mathcal{G}_f$ it follows that $\mathbf{g}^{(p+1)} \in \Gamma^{1 \times M_t}$. The set $\mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)})$ contains all combinations of $\gamma_i \mathbf{g}^{(i)}$ such that $\gamma_i \mathbf{g}^{(i)} \in \Gamma^{1 \times M_t}$. This is possible only if for some set of $\{\gamma_i\}$, $\sum \alpha^t (\gamma_i \mathbf{g}^{(i)}) \in \Gamma^{1 \times M_t}$ but $\alpha^t \gamma_k \mathbf{g}^{(k)} \notin \Gamma^{1 \times M_t}$ for some k. This implies that,

$$deg\left(\alpha^t \gamma_k \mathbf{g}^{(k)} + \sum \alpha^t (\gamma_i \mathbf{g}^{(i)})\right) < deg\left(\alpha^t \gamma_k \mathbf{g}^{(k)}\right).$$

Since $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)}$ satisfy (iii) this is not possible. Therefore, $\mathbf{g}^{(p+1)} \in \mathcal{G}_f$ can always be chosen such that $\mathbf{g}^{(p+1)} \notin \mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)})$ and $\Phi(\mathbf{g}^{(p+1)}) \neq \mathbf{0}$.

such that $\gamma_i \mathbf{g}^{(i)} \in \Gamma^{1\times M_t}$. Since $\gamma_i \mathbf{g}^{(i)} \in \Gamma^{1\times M_t}$ it follows that,

$$deg(\gamma_i) + deg(\mathbf{g}^{(i)}) \leq \nu, \tag{65}$$

where the definitions given in (45) and (46) have been used. Now consider,

$$\tilde{\gamma}_k = \gamma_k \alpha^t \tag{66}$$
$$\tilde{\gamma}_i = w_i \gamma_k + \gamma_i \quad \forall i \neq k. \tag{67}$$

Then,

$$\begin{aligned}
\tilde{\gamma}_1 \tilde{\mathbf{g}}^{(1)} + \ldots + \tilde{\gamma}_{p+1} \tilde{\mathbf{g}}^{(p+1)} &= \gamma_k \alpha^t \left[\alpha^{-t} (w_1 \mathbf{g}^{(1)} + \ldots + \mathbf{g}^{(k)} + \ldots + w_{p+1} \mathbf{g}^{(p+1)}) \right] \\
&\quad + \sum_{i \neq k} (w_i \gamma_k + \gamma_i) \mathbf{g}^{(i)} \\
&= \gamma_k \mathbf{g}^{(k)} + \sum_{i \neq k} \left(w_i \gamma_k \mathbf{g}^{(i)} \right) + \sum_{i \neq k} \left(\gamma_i \mathbf{g}^{(i)} + w_i \gamma_k \mathbf{g}^{(i)} \right) \\
&= \gamma_1 \mathbf{g}^{(1)} + \ldots + \gamma_{p+1} \mathbf{g}^{(p+1)},
\end{aligned}$$

where the last step follows as the field has characteristic 2. To verify that $\tilde{\gamma}_i \in \Gamma$ and $\tilde{\gamma}_i \tilde{\mathbf{g}}^{(i)} \in \Gamma^{1\times M_t}$ we have the following three cases:

- $i \neq k$, $w_i = 0$.
  From (67) $\tilde{\mathbf{g}}^{(i)} = \gamma_i$ as $w_i = 0$ and $\tilde{\mathbf{g}}^{(i)} = \mathbf{g}^{(i)}$ as $i \neq k$. Therefore, from (65) it follows that $\tilde{\gamma}_i \in \Gamma$ and $\tilde{\gamma}_i \tilde{\mathbf{g}}^{(i)} \in \Gamma^{1\times M_t}$.

- $i \neq k$, $w_i = 1$.
  For this case, an upper bound to the degree of $\tilde{\gamma}_i \tilde{\mathbf{g}}^{(i)}$ is given by

  $$\begin{aligned}
  deg\left(\tilde{\gamma}_i \tilde{\mathbf{g}}^{(i)}\right) &\leq \max\left(deg(\gamma_k) + deg(\tilde{\mathbf{g}}^{(i)}), deg(\gamma_i) + deg(\tilde{\mathbf{g}}^{(i)})\right) \\
  &\leq \max\left(deg(\gamma_k) + deg(\mathbf{g}^{(i)}), deg(\gamma_i) + deg(\mathbf{g}^{(i)})\right) \quad \text{from } (61) \\
  &\leq \max\left(deg(\gamma_k) + deg(\mathbf{g}^{(k)}), deg(\gamma_i) + deg(\mathbf{g}^{(i)})\right) \quad \text{from } (59) \\
  &\leq \nu,
  \end{aligned}$$

  where the last inequality follows from (65). Therefore, $\tilde{\gamma}_i \in \Gamma$ and $\tilde{\gamma}_i \tilde{\mathbf{g}}^{(i)} \in \Gamma^{1\times M_t}$.

- $i = k$.
  Note that,

  $$\begin{aligned}
  deg(\tilde{\gamma}_k) + deg(\tilde{\mathbf{g}}^{(k)}) &= deg(\gamma_k) + t + deg(\tilde{\mathbf{g}}^{(k)}) \quad \text{from } (66) \\
  &\leq deg(\gamma_k) + deg(\mathbf{g}^{(k)}) \quad \text{from } (62) \\
  &\leq \nu, \quad \text{from } (65).
  \end{aligned}$$

  Therefore, $\tilde{\gamma}_k \in \Gamma$ and $\tilde{\gamma}_k \tilde{\mathbf{g}}^{(k)} \in \Gamma^{1\times M_t}$.

Hence $\forall \mathbf{g} \in \mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p+1)})$, $\mathbf{g} \in \mathcal{D}(\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p+1)})$. Therefore,

$$\mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p+1)}) \subset \mathcal{D}(\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p+1)}). \tag{68}$$

Proof of elements satisfying (ii) and termination condition: Since $t \geq 1$,

$$deg(\tilde{\mathbf{g}}^{(k)}) < deg(\mathbf{g}^{(k)}) \tag{69}$$

and $deg(\tilde{\mathbf{g}}^{(i)}) = deg(\mathbf{g}^{(i)})\ \forall i \neq k$. Therefore for the new set $\{\tilde{\mathbf{g}}^{(i)}\}_{i=1}^{p+1}$, the degree is smaller than or equal to that of the previous set $\{\mathbf{g}^{(i)}\}_{i=1}^{p+1}$. Since the degree of atleast one element is being reduced and the maximal degree of the set is bounded above by ν, if this step is iterated, the process will terminate. If $\Phi(\tilde{\mathbf{g}}^{(1)}), \ldots, \Phi(\tilde{\mathbf{g}}^{(p+1)})$ are linearly independent then the process terminates and the elements $(\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p)}, \tilde{\mathbf{g}}^{(p+1)})$ satisfy the conditions in the lemma. If not, continue the process defined in (60) till $\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p+1)}$ are obtained such that $\Phi(\tilde{\mathbf{g}}^{(p)}), \ldots, \Phi(\tilde{\mathbf{g}}^{(p+1)})$ are linearly independent or $deg(\tilde{\mathbf{g}}^{(1)}) = \ldots = deg(\tilde{\mathbf{g}}^{(p+1)}) = 0$. If the former occurs, the required set $\{\tilde{\mathbf{g}}^{(i)}\}$ has been obtained. If the latter occurs, and if $\Phi(\tilde{\mathbf{g}}^{(p)}), \ldots, \Phi(\tilde{\mathbf{g}}^{(p+1)})$ are linearly independent, again the objective is achieved. Now, if the latter occurs, *i.e.*, $deg(\tilde{\mathbf{g}}^{(1)}) = \ldots = deg(\tilde{\mathbf{g}}^{(p+1)}) = 0$ and $\Phi(\tilde{\mathbf{g}}^{(1)}), \ldots, \Phi(\tilde{\mathbf{g}}^{(p+1)})$ are linearly dependent, then since the degrees are equal to zero just take the set of independent $\tilde{\mathbf{g}}^{(i)}$. It can be seen that using these sets of vectors it is possible to satisfy (i) and (ii). Note that $\mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p+1)})$ cannot be equal to the set $\mathcal{G}_f$ without the elements $(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p+1)})$ satisfying properties (i), (ii)[7]. Therefore, using this iterative process it is possible to construct the required set $\{\tilde{\mathbf{g}}^{(i)}\}$ since in (68) it has already been shown that the nesting property needed in (57) is satisfied. ∎

Note that in Lemma 16 $\mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)})$ is a proper subset of $\mathcal{D}(\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(t)})$ as the element $\mathbf{g}^{(p+1)}$ is not contained in $\mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p)})$.

Lemma 17 *If there exist $\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p+1)} \in \mathcal{G}_f$ satisfying (i) and (ii) but not satisfying (iii) in Definition 14 it is possible to construct $\hat{\mathbf{g}}^{(1)}, \ldots, \hat{\mathbf{g}}^{(p+1)}$ satisfying (i), (ii) and (iii) in Definition 14 and*

$$\mathcal{D}(\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p+1)}) \subset \mathcal{D}(\hat{\mathbf{g}}^{(1)}, \ldots, \hat{\mathbf{g}}^{(p+1)}). \tag{71}$$

Proof. ***Construction of new elements satisfying (i):*** Let $\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p+1)} \in \mathcal{G}_f$ represent the elements satisfying (i) and (ii) but not satisfying (iii) in Definition 14. From Lemma 15 it follows that there exists a set $S \subset \{1, \ldots, d\}$, $k \notin S$ and $\{\gamma_i\}_{i \in S}$, $\gamma_i \in \Gamma$, such that,

$$deg\left(\tilde{\mathbf{g}}^{(k)} + \sum_{i \in S} \gamma_i \tilde{\mathbf{g}}^{(i)}\right) \ < \ deg(\tilde{\mathbf{g}}^{(k)}) \tag{72}$$

and also,

$$deg(\gamma_i \tilde{\mathbf{g}}^{(i)}) \ \leq \ deg(\tilde{\mathbf{g}}^{(k)}) \quad \forall i \in S. \tag{73}$$

[7] If property (ii) is not satisfied, $\Phi(\mathbf{g}^{(1)}), \ldots, \Phi(\mathbf{g}^{(p+1)})$ are not linearly independent, *i.e.*,

$$w_1 \Phi(\mathbf{g}^{(1)}) + w_2 \Phi(\mathbf{g}^{(2)}) + \ldots w_{p+1} \Phi(\mathbf{g}^{(p+1)}) = 0 \tag{70}$$

for $w_1, \ldots, w_{p+1} \in \mathbb{F}_2$ and not all equal to zero. Since $\sum_i w_i \Phi(\mathbf{g}^{(i)}) = \Phi\left(\sum_i w_i (\mathbf{g}^{(i)})\right) = 0$, there is a common $\alpha^{(\cdot)}$ factor in $w_1 \mathbf{g}^{(1)} + \ldots + w_{p+1} \mathbf{g}^{(p+1)}$ and the element $\mathbf{g} = \sum_i \underbrace{\alpha^{-1} w_i}_{\gamma_i} \mathbf{g}^{(i)}$ is contained in $\mathcal{G}_f$ but not in $\mathcal{D}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(p+1)})$, since $\gamma_i \notin \Gamma$.

There might be multiple such sets S and we choose the one which has highest degree of $\tilde{\mathbf{g}}^{(k)}$. Therefore it follows that $deg(\tilde{\mathbf{g}}^{(k)}) \geq deg(\tilde{\mathbf{g}}^{(i)})$ for all $i \in S$. Define,

$$\begin{aligned} \hat{\mathbf{g}}^{(k)} &= \tilde{\mathbf{g}}^{(k)} + \sum_{i \in S} \gamma_i \tilde{\mathbf{g}}^{(i)} \\ \hat{\mathbf{g}}^{(i)} &= \tilde{\mathbf{g}}^{(i)} \quad \forall i \neq k. \end{aligned} \tag{74}$$

To see that property (i) is satisfied by $\hat{\mathbf{g}}^{(1)}, \ldots, \hat{\mathbf{g}}^{(p+1)}$ first observe that for $i \neq k$,

$$\hat{\mathbf{g}}^{(i)} = \tilde{\mathbf{g}}^{(i)} \Rightarrow \Phi(\hat{\mathbf{g}}^{(i)}) \neq \mathbf{0} \quad \forall i \neq k. \tag{75}$$

To prove that $\Phi(\hat{\mathbf{g}}^{(k)}) \neq \mathbf{0}$ represent γ_i in (72) as $\gamma_i = \sum_{b=0}^{\nu} \delta_b^{(i)} \alpha^b$ where $\delta_b^{(i)} \in \mathbb{F}_2$. Then,

$$\begin{aligned} \Phi(\hat{\mathbf{g}}^{(k)}) &= \Phi\left(\tilde{\mathbf{g}}^{(k)} + \sum_{i \in S} \sum_{b=0}^{\nu} \delta_b^{(i)} \alpha^b \tilde{\mathbf{g}}^{(i)}\right) \\ &= \Phi(\tilde{\mathbf{g}}^{(k)}) + \sum_{i \in S} \left(\delta_0^{(i)} \Phi(\tilde{\mathbf{g}}^{(i)})\right) + \sum_{i \in S} \sum_{b=1}^{\nu} \delta_b^{(i)} \Phi(\alpha^b \tilde{\mathbf{g}}^{(i)}) \\ &= \Phi(\tilde{\mathbf{g}}^{(k)}) + \sum_{i \in S} \delta_0^{(i)} \Phi(\tilde{\mathbf{g}}^{(i)}). \end{aligned} \tag{76}$$

Note that since $\delta_0 \in \mathbb{F}_2$ and $\Phi(\tilde{\mathbf{g}}^{(1)}), \ldots, \Phi(\tilde{\mathbf{g}}^{(p+1)})$ were independent to begin with as $\{\tilde{\mathbf{g}}^{(i)}\}$ satisfy (i). Therefore it follows that $\Phi(\hat{\mathbf{g}}^{(k)}) \neq \mathbf{0}$ and hence $\hat{\mathbf{g}}^{(1)}, \ldots, \hat{\mathbf{g}}^{(p+1)}$ satisfy property (i) of Definition 14.

Proof of elements satisfying (ii): The proof of the linear independence of $\{\Phi(\hat{\mathbf{g}}^{(i)})\}_i$ follows by contradiction. $\{\Phi(\hat{\mathbf{g}}^{(i)})\}_{i \neq k}$ are linearly independent since $\{\tilde{\mathbf{g}}^{(i)}\}$ satisfy (i) and (ii) of Definition 14 and $\hat{\mathbf{g}}^{(i)} = \tilde{\mathbf{g}}^{(i)}, i \neq k$. Assuming that $\Phi(\hat{\mathbf{g}}^{(k)})$ is linearly dependent on $\{\Phi(\hat{\mathbf{g}}^{(i)})\}_{i \neq k}$ it follows that

$$\Phi(\hat{\mathbf{g}}^{(k)}) = \sum_{i \neq k} \theta_i \Phi(\hat{\mathbf{g}}^{(i)}) = \sum_{i \neq k} \theta_i \Phi(\tilde{\mathbf{g}}^{(i)})$$

for $\theta_i \in \mathbb{F}_2$. Since $\Phi(\hat{\mathbf{g}}^{(k)}) \neq \mathbf{0}$, $\{\theta_i\}$ are not all equal to zero. Due to (76) this implies that,

$$\Phi(\tilde{\mathbf{g}}^{(k)}) + \sum_{i \in S} \delta_0^{(i)} \Phi(\tilde{\mathbf{g}}^{(i)}) = \sum_{i \neq k} \theta_i \Phi(\tilde{\mathbf{g}}^{(i)}),$$

or equivalently

$$\Phi(\tilde{\mathbf{g}}^{(k)}) + \sum_{i \in S} \delta_0^{(i)} \Phi(\tilde{\mathbf{g}}^{(i)}) + \sum_{i \neq k} \theta_i \Phi(\hat{\mathbf{g}}^{(i)}) = \mathbf{0}.$$

This contradicts the linear independence of $\Phi(\tilde{\mathbf{g}}^{(1)}), \ldots, \Phi(\tilde{\mathbf{g}}^{(p+1)})$ as $k \notin S$. Therefore $\Phi(\hat{\mathbf{g}}^{(k)})$ is linearly independent of $\Phi(\hat{\mathbf{g}}^{(i)}) = \Phi(\tilde{\mathbf{g}}^{(i)})$ for all $i \neq k$ and hence $\hat{\mathbf{g}}^{(1)}, \ldots, \hat{\mathbf{g}}^{(p+1)}$ satisfy (i) and (ii) of Definition 14.

Proof of containment (71): To show (71) let $\tilde{\mathbf{g}} \in \mathcal{D}(\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p+1)})$, *i.e.*,

$$\tilde{\mathbf{g}} = \tilde{\gamma}_1 \tilde{\mathbf{g}}^{(1)} + \ldots + \tilde{\gamma}_{p+1} \tilde{\mathbf{g}}^{(p+1)}. \tag{77}$$

Choose $\hat{\gamma}_i = \tilde{\gamma}_i$ if $i \notin S$, and $\hat{\gamma}_i = \tilde{\gamma}_i + \gamma_i\tilde{\gamma}_k$ for all $i \in S$ where $\{\gamma_i\}_{i\in S}$ is defined in (72). Since $k \notin S$, it follows that,

$$\begin{aligned}\hat{\gamma}_1\hat{\mathbf{g}}^{(1)} + \ldots + \hat{\gamma}_{p+1}\hat{\mathbf{g}}^{(p+1)} &= \tilde{\gamma}_k\left(\tilde{\mathbf{g}}^{(k)} + \sum_{i\in S}\gamma_i\tilde{\mathbf{g}}^{(i)}\right) + \sum_{i\in S}\left(\tilde{\gamma}_i + \gamma_i\tilde{\gamma}_k\right)\tilde{\mathbf{g}}^{(i)} + \\ &\qquad \sum_{i\notin S, i\neq k}\tilde{\gamma}_i\tilde{\mathbf{g}}^{(i)} \\ &= \tilde{\gamma}_k\tilde{\mathbf{g}}^{(k)} + \left(\sum_{i\in S}\tilde{\gamma}_k\gamma_i\tilde{\mathbf{g}}^{(i)}\right) + \left(\sum_{i\in S}\tilde{\gamma}_i\tilde{\mathbf{g}}^{(i)}\right) + \left(\sum_{i\in S}\gamma_i\tilde{\gamma}_k\tilde{\mathbf{g}}^{(i)}\right) + \left(\sum_{i\notin S, i\neq k}\tilde{\gamma}_i\tilde{\mathbf{g}}^{(i)}\right) \\ &= \tilde{\gamma}_k\tilde{\mathbf{g}}^{(k)} + \left(\sum_{i\neq k}\tilde{\gamma}_i\tilde{\mathbf{g}}^{(i)}\right)\end{aligned}$$

where the last step follows as the characteristic of the field is 2. It still needs to be shown that $\hat{\gamma}_i \in \Gamma$ and $\hat{\gamma}_i\hat{\mathbf{g}}^{(i)} \in \Gamma^{1\times M_t}$. Consider the following three cases:

- $i \notin S, i \neq k$.
  Since $i \notin S$, $\hat{\gamma}_i = \tilde{\gamma}_i$, and the definition in (74) implies that $\hat{\mathbf{g}}^{(i)} = \tilde{\mathbf{g}}^{(i)}, i \notin S, i \neq k$. Since $\gamma_i \in \Gamma$ and $\gamma_i\tilde{\mathbf{g}}^{(i)} \in \Gamma^{1\times M_t}$, for all i in (77) it can be concluded that $\hat{\gamma}_i \in \Gamma$ and $\hat{\gamma}_i\hat{\mathbf{g}}^{(i)} \in \Gamma^{1\times M_t}$, for all $i \notin S, i \neq k$.

- $i \in S$.
  Now, for $i \in S$, by the definition above $\hat{\gamma}_i = \tilde{\gamma}_i + \gamma_i\tilde{\gamma}_k$ and $\hat{\gamma}_i\hat{\mathbf{g}}^{(i)} = \hat{\gamma}_i\tilde{\mathbf{g}}^{(i)}$. Therefore, for $i \in S$,

$$deg(\hat{\gamma}_i) \leq \max\{deg(\tilde{\gamma}_i), deg(\gamma_i) + deg(\tilde{\gamma}_k)\}, \tag{78}$$
$$deg(\hat{\gamma}_i\hat{\mathbf{g}}^{(i)}) \leq \max\{deg(\tilde{\gamma}_i\tilde{\mathbf{g}}^{(i)}), deg(\gamma_i) + deg(\tilde{\gamma}_k) + deg(\tilde{\mathbf{g}}^{(i)})\}. \tag{79}$$

  Note that due to (73),

$$deg(\tilde{\mathbf{g}}^{(i)}) + deg(\gamma_i) \leq deg(\tilde{\mathbf{g}}^{(k)}) \quad \forall i \in S. \tag{80}$$

  Also, from (77), by definition, $\tilde{\gamma}_i \in \Gamma$, $\forall i$ and $\tilde{\gamma}_i\tilde{\mathbf{g}}^{(i)} \in \Gamma^{1\times M_t}\forall i$, hence

$$deg(\tilde{\mathbf{g}}^{(i)}) + deg(\tilde{\gamma}_i) \leq \nu, \ \forall i, \tag{81}$$

  and in particular,

$$deg(\tilde{\mathbf{g}}^{(k)}) + deg(\tilde{\gamma}_k) \leq \nu. \tag{82}$$

  From (80) and (82), it follows that for $i \in S$,

$$deg(\gamma_i) + deg(\tilde{\gamma}_k) + deg(\tilde{\mathbf{g}}^{(i)}) \leq \nu, \tag{83}$$

  and by definition in (77) $deg(\tilde{\gamma}_i) \leq \nu$. This implies that $deg(\gamma_i) + deg(\tilde{\gamma}_k) \leq \nu$ and hence from (78), $deg(\hat{\gamma}_i) \leq \nu$, *i.e.*, $\hat{\gamma}_i \in \Gamma, i \in S$. Again, by definition in (77) it is known that $deg(\tilde{\gamma}_i\tilde{\mathbf{g}}^{(i)}) \leq \nu$. Combining this and (83), it follows from (79) that $deg(\hat{\gamma}_i\hat{\mathbf{g}}^{(i)}) \leq \nu$ and hence $\hat{\gamma}_i\hat{\mathbf{g}}^{(i)} \in \Gamma^{1\times M_t}$.

- $i = k$.
  For $i = k$, since $\hat{\gamma}_k = \tilde{\gamma}_k$ it is clear that $\hat{\gamma}_k \in \Gamma$. It needs to be shown that $\hat{\gamma}_k \hat{\mathbf{g}}^{(k)} \in \Gamma^{1\times M_t}$ *i.e.*,

$$deg\left(\tilde{\gamma}_k\left(\tilde{\mathbf{g}}^{(k)} + \sum_{i\in S}\gamma_i\tilde{\mathbf{g}}^{(i)}\right)\right) \leq \nu.$$

  Note that $deg(\tilde{\gamma}_k\tilde{\mathbf{g}}^{(k)}) \leq \nu$ follows directly from definition in (81). Also,

$$\begin{aligned} deg(\tilde{\gamma}_k\gamma_i\tilde{\mathbf{g}}^{(i)}) &\leq deg(\tilde{\gamma}_k) + deg(\gamma_i) + deg(\tilde{\mathbf{g}}^{(i)}) \\ &\leq deg(\tilde{\mathbf{g}}^{(k)}) + deg(\tilde{\gamma}_k) \qquad \text{from } (73) \\ &\leq \nu, \end{aligned}$$

  where the last inequality follows from (82).

Therefore it follows that,

$$\mathcal{D}(\tilde{\mathbf{g}}^{(1)}, \ldots, \tilde{\mathbf{g}}^{(p+1)}) \subset \mathcal{D}(\hat{\mathbf{g}}^{(1)}, \ldots, \hat{\mathbf{g}}^{(p+1)}).$$

Proof of elements satisfying (iii) and termination condition: Note that the degree of one of the elements of the set $\hat{\mathbf{g}}^{(1)}, \ldots, \hat{\mathbf{g}}^{(p+1)}$ (specifically $\hat{\mathbf{g}}^{(k)}$) is strictly less than the degree of $\tilde{\mathbf{g}}^{(k)}$ and the degree of all other elements is the same. If $\hat{\mathbf{g}}^{(1)}, \ldots, \hat{\mathbf{g}}^{(p+1)}$ satisfy (iii) then terminate, otherwise repeat the process. Note that at each iteration the degree of one of the elements is decreased by at least 1. Since at the beginning of the process the degree was finite, continue this process either until the property (iii) is satisfied or the maximum degree of elements not satisfying (iii) is equal to zero. At this point if property (iii) is not satisfied, from Lemma 15, for some $S' \subset \{1, \ldots, p+1\}$ if follows that[8]

$$deg\left(\tilde{\mathbf{g}}^{(k)} + \sum_{i\in S'}\tilde{\mathbf{g}}^{(i)}\right) < deg(\tilde{\mathbf{g}}^{(k)}) = 0.$$

This is possible only if,

$$\tilde{\mathbf{g}}^{(k)} + \sum_{i\in S'}\tilde{\mathbf{g}}^{(i)} = \mathbf{0}.$$

But since $\tilde{\mathbf{g}}^{(i)} = \Phi(\tilde{\mathbf{g}}^{(i)})$ for , $i \in S'$ and $i = k$, and as $\{\tilde{\mathbf{g}}^{(i)}\}$ satisfy property (ii) there is a contradiction. Therefore property (iii) will be satisfied when the degree of all the elements is 0. Note that $\mathcal{D}(\hat{\mathbf{g}}^{(1)}, \ldots, \hat{\mathbf{g}}^{(p+1)})$ cannot be equal to the set $\mathcal{G}_f$ without the elements $(\hat{\mathbf{g}}^{(1)}, \ldots, \hat{\mathbf{g}}^{(p+1)})$ satisfying property (iii) [9]. ∎

Given these two lemmas we will show that given a fixed nonzero $\mathbf{c}_f \in \mathbb{F}_{2^T}^{M_t\times 1}$, and the associated $\mathcal{G}_f$ defined as in equation (43), there exist minimal basis vectors satisfying the properties in Definition 14, reproduced in the following theorem for completeness.

[8] Since from Lemma 15, if property (iii) is not satisfied, then $deg\left(\gamma_i\mathbf{g}^{(i)}\right) \leq deg\left(\mathbf{g}^{(k)}\right) \quad \forall i \in S'$, and hence for this case, $\gamma_i = 1, i \in S'$.

[9] If property (iii) is not satisfied, then for some subset $S \subseteq \{1, \ldots, p\}$ there exist $\{\gamma_i\}_{i\in S}$ such that,

$$deg(\sum_{i\in S}\gamma_i\hat{\mathbf{g}}^{(i)}) < \max_{i\in S} deg(\gamma_i\hat{\mathbf{g}}^{(i)}). \tag{84}$$

Theorem 18 (Existence of Minimal Basis Vectors) *Given a fixed nonzero $\mathbf{c}_f \in \mathbb{F}_{2^T}^{M_t \times 1}$ define the associated $\mathcal{G}_f$ as in equation (43). Then there exist elements $\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \ldots, \mathbf{g}^{(d)} \in \mathcal{G}_f$ such that they satisfy the following properties:*

(i). For each $\mathbf{g}^{(i)}$, $\exists\ k$ such that $\delta_{k,0}^{(i)} = 1$, i.e., *$\Phi(\mathbf{g}^{(i)}) \neq \mathbf{0}$.*

(ii). $\Phi(\mathbf{g}^{(1)}), \ldots, \Phi(\mathbf{g}^{(d)})$ are linearly independent over $\mathbb{F}_2$.

(iii). For every subset $S \subseteq \{1, \ldots, d\}$, there do not exist $\{\gamma_i\}$ such that,

$$\left\{\gamma_i : i \in S, \gamma_i \in \Gamma \text{ and } \gamma_i \mathbf{g}^{(i)} \in \Gamma^{1 \times M_t}\right\},$$

and

$$deg(\sum_{i \in S} \gamma_i \mathbf{g}^{(i)}) < \max_{i \in S} deg(\gamma_i \mathbf{g}^{(i)}),$$

are simultaneously satisfied.

(iv). $\mathcal{G}_f = \mathcal{D}\left(\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \ldots, \mathbf{g}^{(d)}\right)$

Proof. Clearly let us assume $\mathcal{G}_f$ is not empty. Then $\exists$ a $\mathbf{g}^{(1)} \in \mathcal{G}_f$ such that $\delta_{k,0}^{(1)} = 1$ for some k, since otherwise in the $\mathbf{g}^{(1)}$ picked we can take out $\alpha^{(\cdot)}$ factor and still have it in $\mathcal{G}_f$. Clearly properties (ii) and (iii) of Definition 14 are satisfied trivially. If $\mathbf{g}^{(1)}$ satisfies property (iv) then we are done. If not, we proceed to build the set $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(d)}$. If $\mathbf{g}^{(1)}$ does not satisfy (iv) it means that $\exists\ \mathbf{g}^{(2)} \in \mathcal{G}_f$ such that $\mathbf{g}^{(2)} \neq \gamma_1 \mathbf{g}^{(1)}$ for any $\gamma_1 \in \Gamma$ and $\gamma_1 \mathbf{g}^{(1)} \in \Gamma^{1 \times M_t}$. From Lemma 16 we can construct either $\tilde{\mathbf{g}}^{(1)}, \tilde{\mathbf{g}}^{(2)}$ (or just $\tilde{\mathbf{g}}^{(1)}$) such that they satisfy (i) and (ii) and

$$\mathcal{D}(\mathbf{g}^{(1)}, \mathbf{g}^{(2)}) \subset \mathcal{D}(\tilde{\mathbf{g}}^{(1)}, \tilde{\mathbf{g}}^{(2)}). \tag{85}$$

If (iii), (iv) are also satisfied, then $d = 2$.

If (iii) is not satisfied by these vectors $\tilde{\mathbf{g}}^{(1)}, \tilde{\mathbf{g}}^{(2)}$ we can construct $\hat{\mathbf{g}}^{(1)}, \hat{\mathbf{g}}^{(2)}$ from Lemma 17 which satisfy (i), (ii) and (iii) [10].

Let $t = \max_{i \in S} deg(\gamma_i \mathbf{g}^{(i)})$ and define $k = \text{argmax}_{i \in S}\, deg(\gamma_i \mathbf{g}^{(i)})$. Note then that the element,

$$\mathbf{g} = \alpha^{\nu - t + 1} \left(\sum_{i \in S} \gamma_i \hat{\mathbf{g}}^{(i)} \right)$$

is contained in $\mathcal{G}_f$ as the elements $(\hat{\mathbf{g}}^{(1)}, \ldots, \hat{\mathbf{g}}^{(p+1)})$ satisfy property (iii) for the $\{\gamma_i\}$. But,

$$\mathbf{g} = \underbrace{\alpha^{\nu - t + 1} \gamma_k}_{\gamma} \hat{\mathbf{g}}^{(k)} + \sum_{i \in S, i \neq k} \alpha^{\nu - t + 1} \gamma_i \hat{\mathbf{g}}^{(i)}$$

is not contained in $\mathcal{D}(\hat{\mathbf{g}}^{(1)}, \ldots, \hat{\mathbf{g}}^{(p+1)})$ because $\gamma \notin \Gamma$.

[10] The reason we need the property (iii) is as follows. If we take any element $\mathbf{g} \in \mathcal{G}_f$ then if $\alpha \mathbf{g} \in \Gamma^{1 \times M_t}$, then $\alpha \mathbf{g}$ is also in $\mathcal{G}_f$. This may not be captured in our definition of $\mathcal{D}$ framework for the following reason. If $deg\left[\tilde{\mathbf{g}}^{(1)} + \gamma \tilde{\mathbf{g}}^{(2)}\right] < deg(\tilde{\mathbf{g}}^{(1)})$ and $deg(\tilde{\mathbf{g}}^{(1)}) \geq deg(\tilde{\mathbf{g}}^{(2)})$, then for some t, $\alpha^t(\tilde{\mathbf{g}}^{(1)} + \gamma \tilde{\mathbf{g}}^{(2)}) \in \mathcal{G}$ but $\alpha^t \tilde{\mathbf{g}}^{(1)} + \alpha^t \gamma \tilde{\mathbf{g}}^{(2)} \notin \mathcal{D}(\tilde{\mathbf{g}}^{(1)}, \tilde{\mathbf{g}}^{(2)})$ since $\alpha^t \tilde{\mathbf{g}}^{(1)}$ or $\alpha^t \gamma \tilde{\mathbf{g}}^{(2)} \notin \Gamma^{1 \times M_t}$.

Now if $\hat{\mathbf{g}}^{(1)}, \hat{\mathbf{g}}^{(2)}$ satisfy (iv) then we are done, otherwise we again use the Lemma 16 with $\hat{\mathbf{g}}^{(1)}, \hat{\mathbf{g}}^{(2)}$ as the input vectors. Repeat this process until $\hat{\mathbf{g}}^{(1)}, \ldots, \hat{\mathbf{g}}^{(d)}$ satisfy the properties (i), (ii), (iii) and (iv). It is assured that the process terminates since $|\mathcal{G}_f| \leq |\Gamma^{1\times M_t}| \leq 2^{(\nu+1)M_t}$, *i.e.*, $|\mathcal{G}_f|$ is finite. ∎

Note that from property (ii) it follows that the elements are such that $\Phi(\mathbf{g}^{(1)}), \ldots, \Phi(\mathbf{g}^{(d)})$ are linearly independent only over $\mathbb{F}_2$. The following lemma shows that as long as $T > (\nu+1)M_t$ this is sufficient to guarantee the independence of $\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \ldots, \mathbf{g}^{(d)}$ over $\mathbb{F}_{2^T}$ as well.

Lemma 19 *Consider elements $\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \ldots, \mathbf{g}^{(d)} \in \mathcal{G}_f$ such that $\Phi(\mathbf{g}^{(1)}), \ldots, \Phi(\mathbf{g}^{(d)})$ are linearly independent over $\mathbb{F}_2$. If the size of the extension field $\mathbb{F}_{2^T}$ is such that $T > (\nu+1)M_t$ then these vectors are linearly independent over $\mathbb{F}_{2^T}$ as well.*

Proof. Clearly $d \leq M_t$ otherwise the property (ii) in the theorem 18 will be violated. Define,

$$\mathbf{Q} \;= \begin{bmatrix} \mathbf{g}^{(1)} & \ldots & \mathbf{g}^{(d)} \end{bmatrix}^t$$

and

$$\mathbf{H} \;= \begin{bmatrix} \Phi(\mathbf{g}^{(1)}) & \ldots \Phi(\mathbf{g}^{(d)}) \end{bmatrix}^t.$$

By the linear independence of $\Phi(\mathbf{g}^{(1)}), \ldots, \Phi(\mathbf{g}^{(d)})$ we conclude that $\mathbf{H}$ has full rank over $\mathbb{F}_2$. Therefore, there exist d linearly independent columns over $\mathbb{F}_2$ in $\mathbf{H} \in \mathbb{F}_2^{d\times M_t}$. Select these d columns and form the matrix $\hat{\mathbf{H}} \in \mathbb{F}_2^{d\times d}$ which is of rank d. Therefore $det(\hat{\mathbf{H}}) = 1$ as $det(\hat{\mathbf{H}}) \in \mathbb{F}_2$. Select these same columns in the matrix $\mathbf{Q}$ and form the matrix $\hat{\mathbf{Q}} \in \Gamma^{d\times d}$. Note that since $\hat{\mathbf{Q}} \in \Gamma^{d\times d}$,

$$det(\hat{\mathbf{Q}}) \;= \textstyle\sum_{k=0}^{d\nu} \delta_k \alpha^k.$$

The linear independence of $1, \alpha, \ldots, \alpha^{T-1}$ follows as $T > (\nu+1)M_t \geq (\nu+1)d$. Moreover, note that since $\delta_0 = det(\hat{\mathbf{H}}) \neq 0$ it can be concluded that $det(\hat{\mathbf{Q}}) \neq 0$. Hence the vectors $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(d)}$ are linearly independent over $\mathbb{F}_{2^T}$. ∎

6.3 General Rank Distance Codes

In this section we will prove the required rank guarantees for $\mathcal{S}$ with $T_{thr} = R\nu + (M_t - 1)(\nu+1)(2^R - 1)$ and therefore show that $\mathcal{K}_{\nu,d}$ is given by this set. We state the following lemma required in the proof of the rank guarantees and prove it in the Appendix.

Lemma 20 *Consider a matrix $\mathbf{P} \in \mathbb{F}_{2^T}^{R\times R}$ defined as,*

$$\mathbf{P} = \begin{bmatrix} \mathbf{g}^{(1)} \\ \mathbf{g}^{(2)} \\ \vdots \\ \mathbf{g}^{(R)} \end{bmatrix} \begin{bmatrix} 1 & \ldots & 1 & 1 \\ \xi^{2^{R-1}} & \ldots & \xi^2 & \xi \\ (\xi^2)^{2^{R-1}} & \ldots & (\xi^2)^2 & \xi^2 \\ \vdots & & \vdots & \\ (\xi^{(M_t-1)})^{2^{R-1}} & \ldots & (\xi^{(M_t-1)})^2 & \xi^{(M_t-1)} \end{bmatrix}$$

where $\xi = \alpha^{(2^R-1)(\nu+1)}$ and the vectors $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(R)} \in \Gamma^{M_t\times 1}$ are linearly independent over $\mathbb{F}_{2^T}$. If,

$$T \geq (2^R - 1)\nu + (2^R - 1)(\nu+1)\left((M_t - 2)(2^R - 1) + R\right)$$

then $det(\mathbf{P}) \neq 0$.

Theorem 21 *Let $f(x) = \sum_{l=0}^{R-1} f_l x^{2^l}$ as in (34) and $T \geq T_{thr}$. Then for $\mathcal{S}$ defined in (37), $\frac{1}{T}\log|\mathcal{S}| \geq R - \frac{\nu M_t}{T}$ and $\forall f \in \mathcal{S}$, rank$(\mathbf{U}_f) \geq (M_t - R + 1)(\nu + 1)$ over the binary field.*

Proof. The rate bound is directly from Theorem 12. If $\mathcal{O} = \{\mathbf{U}_f : f \in \mathcal{S}\}$ has rank distance $(\nu + 1)M_t - \mathcal{D}$ then there exists a vector $\mathbf{u}_f \neq \mathbf{0}$ for some $f \in \mathcal{S}$ such that the corresponding binary matrix $\mathbf{U}_f$ has binary rank equal to $(\nu+1)M_t - \mathcal{D}$ (as the code is linear). Equivalently there exists some $f \in \mathcal{S}$ for which there exists a binary vector space $\mathcal{B}_f \subseteq \mathbb{F}_2^{(\nu+1)M_t}$ of dimension $\mathcal{D}$ such that for every $\mathbf{b} \in \mathcal{B}_f$, just as we saw in (53), we have

$$\mathbf{b}^t\mathbf{U}_f = \mathbf{0} \iff \mathbf{b}^t\mathbf{u}_f = 0. \tag{86}$$

Note that the size of $\mathcal{B}_f$ is $2^{\mathcal{D}}$. Rewriting the above we have that $\forall \mathbf{b} \in \mathbb{F}_2^{1\times M_t(\nu+1)}$ and $\mathbf{b} \in \mathcal{B}_f$,

$$\underbrace{\begin{bmatrix} b_1 & b_2 & \dots & b_{(\nu+1)M_t} \end{bmatrix}}_{\mathbf{b}} \begin{bmatrix} f(1) \\ f(\xi) \\ \vdots \\ f(\xi^{(M_t-1)}) \\ \alpha f(1) \\ \vdots \\ \alpha^\nu f(\xi^{(M_t-1)}) \end{bmatrix} = 0. \tag{87}$$

Let the function Ψ be as in (42) such that it maps $\mathcal{B}_f$ to $\mathcal{G}_f$. Since Ψ is a one-to-one mapping, as seen in (42) in Section 5.2, it immediately follows that

$$|\mathcal{B}_f| = |\mathcal{G}_f|.$$

With the representation $\mathbf{g} = \Psi(\mathbf{b})$, (87) can be rewritten as,

$$\underbrace{\begin{bmatrix} g_1 & g_2 & \dots & g_{M_t} \end{bmatrix}}_{\mathbf{g}} \underbrace{\begin{bmatrix} f(1) \\ f(\xi) \\ \vdots \\ f(\xi^{(M_t-1)}) \end{bmatrix}}_{\mathbf{c}_f} = 0$$

where $g_i \in \Gamma$, or equivalently as

$$\begin{bmatrix} g_1 & g_2 & \dots & g_{M_t} \end{bmatrix} \underbrace{\begin{bmatrix} 1 & \dots & 1 & 1 \\ \xi^{2^{R-1}} & \dots & \xi^2 & \xi \\ (\xi^2)^{2^{R-1}} & \dots & (\xi^2)^2 & \xi^2 \\ \vdots & & \vdots & \\ (\xi^{(M_t-1)})^{2^{R-1}} & \dots & (\xi^{(M_t-1)})^2 & \xi^{(M_t-1)} \end{bmatrix}}_{\mathbf{W}\in\mathbb{F}_{2^T}^{M_t\times R}} \begin{bmatrix} f_{R-1} \\ f_{R-2} \\ \vdots \\ f_0 \end{bmatrix} = 0.$$

If the only element in $\mathcal{G}_f$ is the all zero vector then, $\mathcal{D} = 0$, $\mathbf{U}_f$ has full binary rank, and the result has already been shown in Theorem 13. If not, by Theorem 18 there exists a set of minimal vectors, $\mathcal{M} = \{\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \dots, \mathbf{g}^{(d)}\}$ for $\mathcal{G}_f$.

If $d \leq R-1$ it implies that $|\mathcal{G}_f| \leq 2^{(R-1)(\nu+1)}$, and therefore $\mathcal{D} = dim(\mathcal{B}_f) \leq (R-1)(\nu+1)$ which in turn would imply that all matrices in $\mathcal{O}$ have rank at least $(M_t - R + 1)(\nu+1)$. We will prove that $d \leq R-1$ by contradiction. Let us assume that there are more than $R-1$ such minimal vectors *i.e.*, $d > R-1$. Taking any R of the minimal vectors of the solution space $\mathcal{G}_f$ we conclude that,

$$\underbrace{\begin{bmatrix} \mathbf{g}^{(1)} \\ \mathbf{g}^{(2)} \\ \vdots \\ \mathbf{g}^{(R)} \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 & 1 \\ \xi^{2^{R-1}} & \cdots & \xi^2 & \xi \\ (\xi^2)^{2^{R-1}} & \cdots & (\xi^2)^2 & \xi^2 \\ \vdots & & \vdots & \\ (\xi^{(M_t-1)})^{2^{R-1}} & \cdots & (\xi^{(M_t-1)})^2 & \xi^{(M_t-1)} \end{bmatrix}}_{\mathbf{P}} \begin{bmatrix} f_{R-1} \\ \vdots \\ f_1 \\ f_0 \end{bmatrix} = \mathbf{0}_{R\times 1}, \tag{88}$$

where $\mathbf{P} \in \mathbb{F}_{2^T}^{R\times R}$. This is possible iff,

$$det(\mathbf{P}) = 0.$$

As shown in Lemma 20, by the linear independence of $\{\alpha^0, \alpha^1, \ldots, \alpha^{T-1}\}$ it follows that the determinant cannnot be equal to zero. Therefore there can be at most $R-1$ basis vectors. From (49) and property (iii) of theorem 18, since

$$\begin{aligned} |\mathcal{B}_f| &= 2^{\mathcal{D}} \\ |\mathcal{G}_f| &\leq 2^{(R-1)(\nu+1)} \\ |\mathcal{B}_f| &= |\mathcal{G}_f|, \end{aligned}$$

it follows that $\mathcal{D} \leq (R-1)(\nu+1)$. Therefore all matrices in $\mathcal{O}$ have rank at least $(M_t - R+1)(\nu+1)$. ■

The consequence of Theorem 21 is that $\mathcal{K}_{\nu,d} = \{\mathbf{C}_f : f \in \mathcal{S}\}$ satisfies the requirements of Definition 5 and therefore can be used to construct diversity embedded codes for fading ISI channels as done in Theorem 7.

7 Examples and Discussion

We will start off by giving an example of a code which has full diversity equal to M_t when transmitted over the flat fading channel but does not have the maximum possible diversity of $(\nu+1)M_t$ when transmitted over an ISI channel with $\nu+1$ taps.

Example 1: Consider construction of a code for $M_t = 2$, $T = 5$ with rate $R = 1$ and BPSK signaling using code constructions given in [9, 15]. To design these codes, use the field extension $\mathbb{F}_{2^5}$ with the primitive polynomial given by $x^5 + x^4 + x^2 + x + 1$ and the primitive element α. Define,

$$f(x) = f_0 x$$

where $f_0 \in \mathbb{F}_{2^5}$ depends on the input message. The space time codeword is obtained as,

$$\mathbf{C}_{f_0} = \begin{bmatrix} \mathbf{f}^t(1) & \mathbf{f}^t(\alpha) \end{bmatrix}^t \tag{89}$$

where $\mathbf{f}(\alpha^i)$ is the representation of $f(\alpha^i)$ as a binary 1×5 row vector and$\mathbf{C}_f \in \mathbb{F}_2^{2\times 5}$. As was shown in [9, 15] this code achieves full diversity $M_t = 2$ *i.e.*, $\mathbf{C}_{f_0}$ has rank 2 for all nonzero $f_0 \in \mathbf{F}_{2^5}$.

Now assume that we use this code for transmission over an ISI channel with $\nu = 1$. Since this is a linear code, the rank distance of the code is the minimum rank of a nonzero codeword. Therefore the space time codeword corresponding to $f_0 = 1$ is given by,

$$\mathbf{C}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}. \tag{90}$$

When transmitted over the ISI channel we see that the equivalent space time codeword is given by,

$$\Theta\left(\mathbf{C}_1\right) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \tag{91}$$

Since

$$\text{rank}(\Theta\left(\mathbf{C}_1\right)) = 3 < 4,$$

we conclude that the space time codeword which achieves full diversity $M_t = 2$ over the flat fading channel does not achieve the maximum possible diversity of $(\nu+1)M_t = 4$ over the ISI channel.

Example 2: Similarly this can be shown to hold true for any diversity point. Consider for example the case of $M_t = 3$, $T = 7$, $R = 2$ and BPSK signaling using code constructions given in [9, 15]. Use the field extension $\mathbb{F}_{2^7}$ with the primitive element α. Define,

$$f(x) = f_1 x^2 + f_0 x$$

where $f_0 \in \mathbb{F}_{2^7}$ depends on the input message as before. The space time codeword is obtained as,

$$\mathbf{C}_{f_0} = \begin{bmatrix} \mathbf{f}^t(1) & \mathbf{f}^t(\alpha) & \mathbf{f}^t(\alpha^2) \end{bmatrix}^t \tag{92}$$

where $\mathbf{f}(\alpha^i)$ is the representation of $f(\alpha^i)$ as a binary 1×5 row vector and $\mathbf{C}_f \in \mathbb{F}_2^{2\times 5}$. As was shown in [9, 15] this code achieves diversity $d = 2$ *i.e.,* $\mathbf{C}_{f_0}$ has rank 2 for all nonzero $f_0 \in \mathbf{F}_{2^7}$. But it can be seen as before that the space time codeword corresponding to $(f_1, f_0) = (0, 1)$ does not achieve the maximum possible diversity of $(\nu+1)M_t$ when transmitting over the ISI channel with $\nu+1$ taps.

Example 3: Consider construction of a BPSK code for $M_t = 2$, $\nu = 1$, $T = 5$ with rate $R = 1$ and hence $R^{eff} = \frac{3}{5}$. To design these codes, use the field extension $\mathbb{F}_{2^5}$ with the primitive polynomial given by $x^5 + x^4 + x^2 + x + 1$ and the primitive element α. The set of codeword polynomials which satisfy the constraints in (37) are given by,

$$\mathcal{S} = \left\{0, \alpha, \alpha^{17}, \alpha^{19}, \alpha^{21}, \alpha^{24}, \alpha^{26}, \alpha^{31}\right\}. \tag{93}$$

This set is of cardinality

$$|\mathcal{S}| = 2^{RT-\nu M_t} = 2^{5-2} = 8.$$

Corresponding to every element f in $\mathcal{S}$ consider the codeword vector,

$$\mathbf{c}_f = \begin{bmatrix} f(1) & f(\alpha^2) \end{bmatrix}^t \tag{94}$$

where $\mathbf{c}_f \in \mathbb{F}_{2^5}^{2\times 1}$. Let $\mathbf{C}_f \in \mathbb{F}_2^{2\times 5}$ be the representation of each element of $\mathbf{c}_f$ in the basis $\{\alpha^0, \alpha^1, \alpha^2, \alpha^3, \alpha^4\}$, *i.e.*,

$$\mathbf{C}_f = \left[\begin{array}{cc} \mathbf{f}^t(1) & \mathbf{f}^t(\alpha^2) \end{array} \right]^t, \tag{95}$$

where $\mathbf{f}(\alpha^i)$ is the representation of $f(\alpha^i)$ as a binary 1×5 row vector. Then the 2×5 space time code has rate,

$$R = 1 - \frac{\nu M_t}{T} = 1 - \frac{2}{5} = \frac{3}{5}$$

and gives diversity 4 when transmitted over the ISI channel with $\nu = 1$. The corresponding 8 codewords $\mathbf{X}^{(1)}$ as given in (4) are,

$$\begin{array}{cc} \left[\begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right], \\ \left[\begin{array}{ccccc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right], \left[\begin{array}{ccccc} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right] \\ \left[\begin{array}{ccccc} 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{array}\right], \left[\begin{array}{ccccc} 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{array}\right], \\ \left[\begin{array}{ccccc} 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{array}\right], \left[\begin{array}{ccccc} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{array}\right]. \end{array}$$

In figure 5 we give the performance of a full diversity code which is designed for $M_t = 2$, $M_r = 1$, $\nu = 1$ and 4-QAM signal constellation. We plot the logarithm of the error probability as a function of SNR (in dB). Note that the slope of the error probability curve is approximately equal to 4 which is expected since we are using full diversity codes on both the layers.

From the construction of these codes, one might be tempted to conclude that the analysis for these codes is quite similar to that of cyclic codes. But the peculiar structure of the solution space, *i.e.*, the fact that given a vector in the solution space $\mathcal{B}$ not all circular shifts of the vector remain in $\mathcal{B}$, makes it difficult to analyze. The main contribution of this work is the construction of binary matrices with a particular structure which consequently characterizes the rate diversity tradeoff for the ISI channel. Note that as seen in Example 1 and Example 2, codes which give guaranteed diversity orders for flat fading MIMO channel, when used for transmission over ISI channel do not necessarily give the multiplicative diversity gain of $(\nu + 1)$. The tools and techniques developed over here could also have independent interest in designing codes in various other wireless or distributed settings.

Acknowledgments

We would like to thank Amin Shokrollahi for interesting and helpful discussions about this work, and in particular for discussions on the proof of Theorem 12.

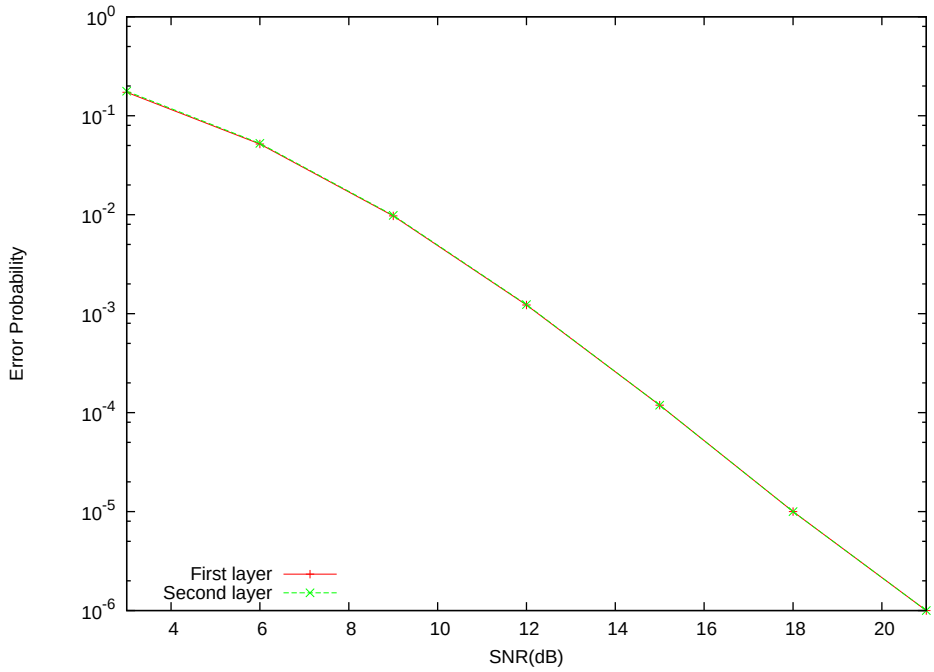

Figure 5: Error Performance of full diversity codes with $M_t = 2$, $\nu = 1$, $R_1 = R_2 = \frac{3}{5}$ and $D_1 = D_2 = 4$.

8 Appendix

Proof. [of Lemma 15] By assumption there exists a subset $S \subseteq \{1, \ldots, p\}$ for which there exist $\{\gamma_i\}_{i \in S}$ such that,

$$deg(\sum_{i \in S} \gamma_i \mathbf{g}^{(i)}) < \max_{i \in S} deg(\gamma_i \mathbf{g}^{(i)}). \tag{96}$$

Let $t = \max_{i \in S} deg(\gamma_i \mathbf{g}^{(i)})$ and define $\mathcal{T} = \left\{ i : deg(\gamma_i \mathbf{g}^{(i)}) = t \right\}$. Note that,

$$deg(\gamma_i \mathbf{g}^{(i)}) = deg(\alpha^{deg(\gamma_i)} \mathbf{g}^{(i)}). \tag{97}$$

This allows us to see that (96) implies that,

$$deg\left(\sum_{i \in \mathcal{T}} \alpha^{deg(\gamma_i)} \mathbf{g}^{(i)} \right) < t. \tag{98}$$

Denote $w = \min_{i \in \mathcal{T}}(deg(\gamma_i))$ to be the minimum degree of γ_i for $i \in \mathcal{T}$ and $k = \text{argmin}_{i \in \mathcal{T}}(deg(\gamma_i))$. Define $S' = \mathcal{T} \backslash \{k\}$ where $\backslash$ is the set difference operator. We have the following two cases depending on the value of w:

- If $w = 0$ then,

$$deg(\mathbf{g}^{(k)} + \sum_{i \in S'} \alpha^{deg(\gamma_i)} \mathbf{g}^{(i)}) < t = deg(\mathbf{g}^{(k)})$$

and,

$$t = deg\left(\alpha^{deg(\gamma_i)}\mathbf{g}^{(i)}\right) \leq deg\left(\mathbf{g}^{(k)}\right) = t \quad \forall i \in S'$$

which shows that if $w = 0$, then the claim is true.

- If $w \neq 0$ the common α^w factor can be taken out of $\sum_{i\in\mathcal{T}} \alpha^{deg(\gamma_i)}\mathbf{g}^{(i)}$. Then,

$$deg(\mathbf{g}^{(k)} + \sum_{i\in S'} \alpha^{deg(\gamma_i)-w}\mathbf{g}^{(i)}) < (t-w) = deg(\mathbf{g}^{(k)})$$

and,

$$t - w = deg\left(\alpha^{deg(\gamma_i)-w}\mathbf{g}^{(i)}\right) \leq deg\left(\mathbf{g}^{(k)}\right) = t - w \quad \forall i \in S'.$$

Hence the claim is proved. ∎

To prove the Lemma 20 we will make use of the Cauchy Binet formula reproduced here for completeness.

Definition 22 Cauchy Binet Formula *[13] Let $\mathbf{A}$ be a $m \times n$ matrix and $\mathbf{B}$ be a $n \times m$ matrix. If S is a subset of $\{1, \ldots, n\}$ with m elements, let $\mathbf{A}_S$ represent the $m \times m$ matrix whose columns are those columns of $\mathbf{A}$ that have indices from S. Similarly, let $\mathbf{B}_S$ represent the $m \times m$ matrix whose rows are those rows of $\mathbf{B}$ that have indices from S. The Cauchy-Binet formula then states that*

$$det(\mathbf{AB}) = \sum_S det(\mathbf{A}_S)\det(\mathbf{B}_S), \tag{99}$$

where the sum extends over all possible subsets S of $\{1, \ldots, n\}$ with m elements.

Note that the Cauchy Binet formula holds for matrices with entries from any commutative rings. Given this definition, the proof of Lemma 20 proceeds as follows.

Proof. [of Lemma 20] The matrix $\mathbf{P}$ is given by,

$$\mathbf{P} = \underbrace{\begin{bmatrix} \mathbf{g}^{(1)} \\ \mathbf{g}^{(2)} \\ \vdots \\ \mathbf{g}^{(R)} \end{bmatrix}}_{\mathbf{M}\in\Gamma^{R\times M_t}} \underbrace{\begin{bmatrix} 1 & \ldots & 1 & 1 \\ \xi^{2^{R-1}} & \ldots & \xi^2 & \xi \\ (\xi^2)^{2^{R-1}} & \ldots & (\xi^2)^2 & \xi^2 \\ \vdots & & \vdots & \\ (\xi^{(M_t-1)})^{2^{R-1}} & \ldots & (\xi^{(M_t-1)})^2 & \xi^{(M_t-1)} \end{bmatrix}}_{\mathbf{W}\in\mathbb{F}_{2^T}^{M_t\times R}}$$

Using Gaussian elimination (which can be applied over any finite field), we reduce the matrix $\mathbf{M}$ to its row echelon form,

$$\breve{\mathbf{Y}} = \begin{bmatrix} \breve{\mathbf{g}}^{(1)} \\ \breve{\mathbf{g}}^{(2)} \\ \vdots \\ \breve{\mathbf{g}}^{(R)} \end{bmatrix},$$

where

$$deg(\breve{\mathbf{g}}^{(k)}) \leq 2^{(R-k)}\nu.$$

Note that this pivoting and reduction to a row echelon form is a full rank operation and preserves the rank of $\mathbf{P}$. Therefore

$$det(\mathbf{P}) = K det\left(\breve{\mathbf{Y}}\mathbf{W}\right),$$

where $K \in \mathbb{F}_{2^T}$ and $K \neq 0$. Let the columns containing the pivots in $\breve{\mathbf{Y}}$ be denoted by $\tilde{S}$. Therefore, by the Cauchy Binet formula we have

$$K^{-1}det(\mathbf{P}) = det\left(\breve{\mathbf{Y}}_{\tilde{S}}\right) det\left(\mathbf{W}_{\tilde{S}}\right) + \sum_{S \neq \tilde{S}} det\left(\breve{\mathbf{Y}}_{S}\right) det\left(\mathbf{W}_{S}\right). \tag{100}$$

Note that for all S such that $det\left(\breve{\mathbf{Y}}_{S}\right) \neq 0$ the maximum coefficient of ξ in $det\left(\mathbf{W}_{S}\right)$ is less than the maximum coefficient of ξ in $det\left(\mathbf{W}_{\tilde{S}}\right)$ by at least 1. Therefore

$$deg(det\left(\mathbf{W}_{\tilde{S}}\right)) - deg(det\left(\mathbf{W}_{S}\right)) \geq (2^R - 1)(\nu + 1). \tag{101}$$

Also note that

$$deg(det(\breve{\mathbf{Y}}_{S})) - deg(det(\breve{\mathbf{Y}}_{\tilde{S}})) \leq \nu + 2\nu + 2^2\nu + \ldots + 2^{R-1}\nu = \nu(2^R - 1). \tag{102}$$

Combining (101) and (102),

$$deg(det\left(\mathbf{W}_{\tilde{S}}\right) det(\breve{\mathbf{Y}}_{\tilde{S}})) - deg(det\left(\mathbf{W}_{S}\right) det(\breve{\mathbf{Y}}_{S})) \geq (2^R - 1)(\nu + 1) - \nu(2^R - 1) > 0.$$

By the linear independence of $\{1, \alpha, \ldots, \alpha^{T-1}\}$ there exists a term in $det\left(\mathbf{W}_{\tilde{S}}\right) det(\breve{\mathbf{Y}}_{\tilde{S}})$ with a power of α which is not canceled by any other term in the equation (100). Therefore it follows that $K^{-1}det(\mathbf{P}) \neq 0$ implying $det(\mathbf{P}) \neq 0$. Hence proved. ∎

References

[1] A. R. Calderbank, S. N. Diggavi and N. Al-Dhahir. Space-Time Signaling based on Kerdock and Delsarte-Goethals Codes. *IEEE International Conference on Communications (ICC), pp 483 - 487, Paris, June 2004.*

[2] S. N. Diggavi, N. Al-Dhahir, and A. R. Calderbank, Diversity embedding in multiple antenna communications, advances in network information theory. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science, pages 285-301, 2004.*

[3] S. N. Diggavi and D. Tse, *Fundamental Limits of Diversity-Embedded Codes over Fading Channels*, IEEE International Symposium on Information Theory (ISIT), pp 510–514. September, 2005.

[4] S. N. Diggavi and D. Tse, *On opportunistic codes and broadcast codes with degraded message sets*, IEEE Information Theory Workshop (ITW), pp 227–231, March, 2006.

[5] S. N. Diggavi, S. Dusad, A. R. Calderbank, N. Al-Dhahir, On Diversity Embedded codes, *Proceedings of Allerton Conference on Communication, Control, and Computing*, Illinois, September 2005.

[6] S. N. Diggavi, A. R. Calderbank, S. Dusad and N. Al-Dhahir, Diversity embedded space-time codes, *IEEE Transactions on Information Theory, 2008: 54(1): 33-50, Jan 2008.*

[7] S. Dusad and S N. Diggavi, On successive refinement of diversity for fading ISI channels, *Proceedings of Allerton Conference on Communication, Control, and Computing*, Illinois, September 2006.

[8] S. Dusad and S. N. Diggavi, Successive refinement of diversity for fading ISI MISO channels, *IEEE International Symposium on Information Theory (ISIT), July 2008*.

[9] E. Gabidulin, Theory of codes with maximum rank distance. *Probl. Per. Inform., 21:3–16, Jan/March 1985.*

[10] H. E. Gamal, A. R. Hammons, Y. Liu, M. P. Fitz, O. Y. Takeshita, On the Design of Space-Time and Space-Frequency Codes for MIMO Frequency Selective Fading Channels, *IEEE Transactions on Information Theory, 49(9):2277–2291, September 2003.*

[11] J-C. Guey, M. P. Fitz, M. R. Bell, and W-Y. Kuo, Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels. *IEEE Transactions on Communications, 47(4):527–537, April 1999.*

[12] A. R. Hammons, Jr., H. El-Gamal, On the theory of space-time codes for PSK modulation, *IEEE Transactions on Information Theory*, Vol 46, No. 2, pp 524–542, March 2000.

[13] R. Horn and C. Johnson *Matrix Analysis*. Cambridge University Press, 1990

[14] R. Lidl and H. Niederreiter *Finite Fields*. Cambridge University Press, 1997

[15] H. F. Lu and P. V. Kumar, Rate-diversity trade-off of space-time codes with fixed alphabet and optimal constructions for PSK modulation, *IEEE Transactions on Information Theory, 49(10):2747–2752, October 2003.*

[16] H. F. Lu and P. V. Kumar, A unified construction of space-time codes with optimal rate-diversity tradeoff, *IEEE Transactions on Information Theory, 51(5):1709–1730, May 2005.*

[17] R. J. McEliece *Finite Fields for Computer Scientists and Engineers*. Kluwer Academic Publishers, 2003

[18] W. Su, Z. Safar, M. Olfat, R. Liu, Obtaining full-diversity space-frequency codes from space-time codes via mapping *IEEE Transactions on Signal Processing, 51(11):2905–2916, November 2003.*

[19] V. Tarokh, N. Seshadri, and A.R. Calderbank. Space-time codes for high data rate wireless communications: Performance criterion and code construction. *IEEE Transactions on Information Theory, 44(2):744–765, March 1998.*

[20] D. N. C. Tse and P. Viswanath, Fundamentals of Wireless Communication,*Cambridge University Press, 2005*

[21] L. Washington *Introduction to cyclotomic fields*, *Springer Verlag*, 2nd edition June 1997.