Full Frequency Diversity Codes for Single Input Single Output Systems

Enis Akay and Ender Ayanoglu

Center for Pervasive Communications and Computing
Department of Electrical Engineering and Computer Science
The Henry Samueli School of Engineering
University of California, Irvine
Irvine, California 92697-2625
Email: eakay@uci.edu ayanoglu@uci.edu

***Abstract*— Due to the severe conditions of wireless channels, it is crucial for wireless systems to accommodate some sort of diversity to achieve reliable communication. Different types of diversity techniques such as temporal, frequency, code, and spatial have been developed in the literature. In addition to the destructive multipath nature of wireless channels, frequency selective channels pose intersymbol interference (ISI) while offering frequency diversity for successfully designed systems. Orthogonal frequency division multiplexing (OFDM) has been shown to fight ISI very well by converting the frequency selective channel into parallel flat fading channels. On the other hand, bit interleaved coded modulation (BICM) was shown by Zehavi and later by Caire *et al* to have high performance for flat fading Rayleigh channels. It is natural to combine BICM and OFDM to exploit the common ground of both techniques to improve overall system performance. In this paper we show both analytically and via simulations that for L tap frequency selective fading channels, BICM-OFDM can achieve a diversity order of $\min(d_{free}, L)$, where d_{free} is the minimum Hamming distance of the convolutional code used for BICM.**

I. Introduction

Wireless communication channels suffer from severe attenuation due to the destructive addition of multiple paths in the propagation media and due to interference from other users. In some cases it is impossible for the receiver to make a correct decision on the transmitted signal unless some form of diversity (less attenuated replica of the signal) is used. In order to fight the severe conditions of wireless channels, different diversity techniques have been developed. Some forms of diversity can be listed as temporal, frequency, spatial and code diversity.

Zehavi [1] showed that the code diversity could be improved by bit-wise interleaving. Using an appropriate soft-decision bit metric at a Viterbi decoder, Zehavi achieved a code diversity equal to the smallest number of distinct bits, rather than channel symbols, along any error event. On the other hand, the diversity order of systems based on Ungerboeck's trellis coded modulation (TCM, [2]) with a symbol interleaver is the minimum number of distinct symbols between codewords. Thus, diversity can only be increased by preventing parallel transitions and increasing the constraint length of the code. As a result, BICM shows performance improvement when compared to TCM [1]. Following Zehavi's paper, Caire *et al* [3] presented the theory behind BICM. Their work illustrated tools to evaluate the performance of BICM with tight error probability bounds, and design guidelines. A brief overview of BICM is given in Section II for reader's convenience.

In recent years deploying multiple transmit antennas has become an important tool to improve diversity. The use of multiple transmit antennas allowed significant diversity gains for wireless communications. In general, spatial diversity systems are called space-time (ST) codes and some important results can be listed as [4], [5], [6], [7]. In these papers the multi input multi output (MIMO) wireless channel is assumed to be flat fading. However, when there is frequency selectivity in the channel, the design of appropriate space-time codes becomes a more complicated problem due to the existence of intersymbol interference (ISI). On the other hand, frequency selective channels offer additional frequency diversity [8], [9]. OFDM can be used to combat ISI and therefore can simplify the code design problem for frequency selective channels. Using a cyclic prefix (CP), OFDM converts a frequency selective channel into parallel flat fading channels. Some space-time-frequency coded systems have been proposed to exploit the diversity order in space, time and frequency, [10], [11], [12], [13], [14], [15]. Out of these papers [14] combines space time block codes (STBC) of [5] and [6] with BICM-OFDM to achieve diversity in space time and frequency. Reference [15] uses BICM-OFDM directly with multiple antennas and without external STBC to achieve higher data rate in the cost of lower diversity.

In this paper we limit ourselves to single input single output (SISO) wireless channels. The reader is urged to note that unlike [14] and [15], we use a SISO system instead of MIMO and through Section III and IV we will *formally prove* that BICM-OFDM systems achieve a diversity order of $\min(d_{free}, L)$, where d_{free} is the minimum Hamming distance of the convolutional code.

In Section V simulation results supporting our analysis are presented. Finally, we end the paper with a brief conclusion in Section VI where we restate the important results of this paper.

0-7803-8521-7/04/$20.00 © 2004 IEEE

II. Bit-Interleaved Coded Modulation (BICM)

BICM can be obtained by using a bit interleaver, π, between an encoder for a binary code $\mathcal{C}$ and a memoryless modulator over a signal set $\chi \subseteq \mathbb{C}$ of size $|\chi| = M = 2^m$ with a binary labeling map $\mu : \{0,1\}^m \to \chi$. Gray encoding is used to map the bits onto symbols and plays an important role in BICM's performance [3]. Gray labeling allows parallel independent decoding for each bit [16]. If set partition labeling or mixed labeling is used, then an iterative decoding approach should be used to achieve high performance [17]. Note that, due to the ability of independent parallel decoding of Gray labeling, iterative decoding does not introduce any performance improvement [17]. Therefore, non-iterative maximum likelihood (ML) decoding is considered in this paper.

During transmission, the code sequence $\underline{c}$ is interleaved by π, and then mapped onto the signal sequence $\underline{x} \in \chi$. The signal sequence $\underline{x}$ is then transmitted over the channel.

The bit interleaver can be modeled as $\pi : k' \to (k, i)$ where k' denotes the original ordering of the coded bits $c_{k'}$, k denotes the time ordering of the signals x_k transmitted, and i indicates the position of the bit $c_{k'}$ in the symbol x_k.

Let χ_b^i denote the subset of all signals $x \in \chi$ whose label has the value $b \in \{0,1\}$ in position i. Then, the maximum likelihood (ML) bit metrics are given by [3]

$$\lambda^i(y_k, c_{k'}) = \begin{cases} \max\limits_{x \in \chi^i_{c_{k'}}} \log p_{\theta_k}(y_k|x), & \text{perfect CSI} \\ \max\limits_{x \in \chi^i_{c_{k'}}} \log p(y_k|x), & \text{no CSI} \end{cases} \tag{1}$$

where θ_k denotes the channel state information (CSI) for the time order k.

Following (1), the bit metrics for the flat fading Rayleigh channels can be calculated using the ML criterion with CSI as [1]

$$\lambda^i(y_k, c_{k'}) = \min_{x \in \chi^i_{c_{k'}}} \|y_k - \rho x\|^2 \tag{2}$$

where ρ denotes the Rayleigh coefficient and $\|(\cdot)\|^2$ represents the squared Euclidean norm of $(\cdot)$.

The ML decoder at the receiver can make decisions according to the rule

$$\hat{\underline{c}} = \arg\min_{\underline{c} \in \mathcal{C}} \sum_{k'} \lambda^i(y_k, c_{k'}). \tag{3}$$

III. System Model of BICM-OFDM

The system deploys only one transmit and one receive antenna (SISO). One OFDM symbol has K subcarriers where each subcarrier corresponds to a symbol from a constellation map χ. As given in Section II, constellation size $|\chi| = 2^m$. It is assumed that the interleaver's depth is Km so that bits are interleaved within one OFDM symbol. By doing so, the decoder does not need to wait for the arrival of multiple OFDM symbols to start decoding.

A convolutional encoder is used to generate the binary code at the transmitter. For k_0/n_0 rate convolutional code with given number of states, the one with the highest minimum Hamming distance, d_{free}, is picked from tables, e.g., [18]. The output bit $c_{k'}$ of a convolutional encoder is interleaved and mapped onto the subcarrier $x(k)$ at the ith location.

Consider a frequency selective channel $\underline{h} = [h_0\ h_1\ \cdots\ h_{(L-1)}]^T$ with L taps. Each tap is assumed to be statistically independent and modeled as a zero mean complex Gaussian random variable with variance $1/2L$ per complex dimension. The fading model is assumed to be quasi-static, i.e., the fading coefficients are constant over the transmission of one packet, but independent from one packet transmission to the next. It is assumed that the taps are spaced at integer multiples of the symbol duration, which is the worst case scenario in terms of designing full diversity codes [19].

A cyclic prefix (CP) of appropriate length is added to each OFDM symbol. Adding CP converts the linear convolution of the transmitted signal and the L-tap channel into a circular convolution. When CP is removed and FFT is taken at the receiver, the received signal is given by

$$y(k) = x(k)H(k) + n(k), \quad 0 \le k \le K-1 \tag{4}$$

where $x(k)$ is the transmitted signal at the kth subcarrier, $n(k)$ is complex additive white Gaussian noise with zero mean and variance $1/(2 \cdot SNR)$ per complex dimension making the variance of $n(k)$ $N_0 = 1/SNR$, and $H(k)$ is given by

$$H(k) = \sum_{l=0}^{L-1} h(l) W_K^{lk}, \text{ where } W_K \triangleq e^{-j\frac{2\pi}{K}}. \tag{5}$$

$H(k)$ can also be written as

$$H(k) = \underline{W}^H(k)\underline{h} \tag{6}$$

where $\underline{W}(k) = [1 \quad W_K^k \quad W_K^{2k} \quad \ldots \quad W_K^{(L-1)k}]^H$ is an $L \times 1$ vector. Note that the transmitted symbols are assumed to have average energy of 1 so that with the channel and AWGN models described here, the received signal to noise ratio is SNR.

IV. Diversity order of BICM-OFDM

In this section we will show that for an L-tap frequency selective channel, BICM-OFDM can achieve a diversity order of $\min(d_{free}, L)$ without the use of multiple antennas. Since d_{free} of convolutional codes can be high, this is a significant result.

Assume the code sequence $\underline{c}$ is transmitted and $\hat{\underline{c}}$ is detected. Then, the PEP of $\underline{c}$ and $\hat{\underline{c}}$ given CSI can be written as, using (2) and (3),

$$P(\underline{c} \to \hat{\underline{c}}|\mathbf{H}) = P\left(\begin{array}{l} \sum\limits_{k'} \min\limits_{x \in \chi^i_{c_{k'}}} \|y(k) - xH(k)\|^2 \ge \\ \sum\limits_{k'} \min\limits_{x \in \chi^i_{\hat{c}_{k'}}} \|y(k) - xH(k)\|^2 \end{array} \right) \tag{7}$$

For a convolutional code with rate k_0/n_0, and minimum Hamming distance d_{free}, the Hamming distance between $\underline{c}$

and $\underline{\hat{c}}$, $d(\underline{c}-\underline{\hat{c}})$, is at least d_{free}. Assume $d(\underline{c}-\underline{\hat{c}}) = d_{free}$ for $\underline{c}$ and $\underline{\hat{c}}$ under consideration for PEP analysis, which is the worst case scenario between any two codewords. Then, $\chi^i_{c_{k'}}$ and $\chi^i_{\hat{c}_{k'}}$ are equal to one another for all k' except for d_{free} distinct values of k'. Therefore, inequality on the right hand side of (7) shares the same terms on all but d_{free} summation points, and the summations can be simplified to only d_{free} terms for PEP analysis.

$$P(\underline{c} \to \underline{\hat{c}}|\mathbf{H}) = P\left(\begin{array}{c} \sum\limits_{k',d_{free}} \min\limits_{x\in\chi^i_{c_{k'}}} \|y(k) - xH(k)\|^2 \geq \\ \sum\limits_{k',d_{free}} \min\limits_{x\in\chi^i_{\hat{c}_{k'}}} \|y(k) - xH(k)\|^2 \end{array}\right) \tag{8}$$

where $\sum_{k',d_{free}}$ means that the summation is taken with index k' over d_{free} different values of k'.

Note that for binary codes and for the d_{free} points at hand, $\hat{c}_{k'} = \bar{c}_{k'}$, where $\bar{(\cdot)}$ denotes the binary complement of $(\cdot)$. For the d_{free} bits let's denote

$$\begin{aligned} \tilde{x}(k) =& \arg \min_{x\in\chi^i_{c_{k'}}} \|y(k) - xH(k)\|^2 \\ \hat{x}(k) =& \arg \min_{x\in\chi^i_{\bar{c}_{k'}}} \|y(k) - xH(k)\|^2 \end{aligned} \tag{9}$$

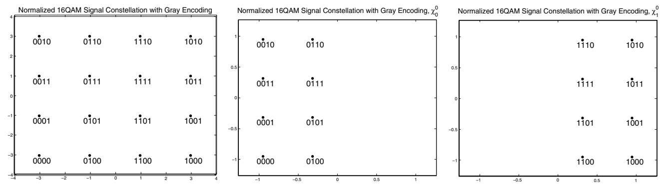
It is easy to see that $\tilde{x}(k) \neq \hat{x}(k)$ since $\tilde{x}(k) \in \chi^i_{c_{k'}}$ and $\hat{x}(k) \in \chi^i_{\bar{c}_{k'}}$ where $\chi^i_{c_{k'}}$ and $\chi^i_{\bar{c}_{k'}}$ are complement sets of constellation points within the signal constellation set χ (see Figure 1 for 16 QAM example). Also, $\|y(k)-x(k)H(k)\|^2 \geq \|y(k) - \tilde{x}(k)H(k)\|^2$ and $x(k) \in \chi^i_{c_{k'}}$.

For convolutional codes, due to their trellis structure, d_{free} distinct bits between any two codewords occur in consecutive trellis branches. The bit interleaver can be designed such that consecutive $\lceil d_{free}/n_0 \rceil n_0$ bits are mapped onto $\lceil d_{free}/n_0 \rceil n_0$ different symbols. This guarantees that there exists d_{free} distinct pairs of $(\tilde{x}(k), \hat{x}(k))$, and d_{free} distinct pairs of $(x(k), \hat{x}(k))$. Note that, if there is no bit interleaver following the encoder, there are only $\lceil d_{free}/m \rceil$ distinct pairs. The PEP of BICM-OFDM can be rewritten as

$$\begin{aligned} P(\underline{c} \to \underline{\hat{c}}|\mathbf{H}) =& P\left(\sum_{k,d_{free}} \begin{array}{c} \|y(k) - \tilde{x}(k)H(k)\|^2 - \\ \|y(k) - \hat{x}(k)H(k)\|^2 \end{array} \geq 0\right) \\ \leq& P\left(\sum_{k,d_{free}} \begin{array}{c} \|y(k) - x(k)H(k)\|^2 - \\ \|y(k) - \hat{x}(k)H(k)\|^2 \end{array} \geq 0\right) \end{aligned} \tag{10}$$

$$= P\left(\left[\sum_{k,d_{free}} \|\left(x(k) - \hat{x}(k)\right)H(k)\|^2\right] - \beta \leq 0\right) \tag{11}$$

where $\beta = \sum_{k,d_{free}} \beta(k)$ and $\beta(k) = (\hat{x}(k) - x(k))^* H^*(k) n(k) + (\hat{x}(k) - x(k)) H(k) n^*(k)$. For given $\mathbf{H}$, $\beta(k)$s are independent zero-mean complex Gaussian random variables with variance $2N_0\|(\hat{x}(k) - x(k))H(k)\|^2$. Consequently, β is a complex Gaussian random variable with zero mean and variance $2N_0 \sum_{k,d_{free}} \|(\hat{x}(k) - x(k))H(k)\|^2$. Note that, the upper bound in (10) is tight, since for high SNR values $\tilde{x}(k) = x(k)$. Following (11), PEP can be found as



(a) χ (b) χ_0^0 (c) χ_0^1

Fig. 1. Gray encoded 16 QAM constellation

$$\begin{aligned} P(\underline{c} \to \underline{\hat{c}}|\mathbf{H}) \leq& P\left(\beta \geq \sum_{k,d_{free}} \|\left(x(k) - \hat{x}(k)\right)H(k)\|^2\right) \\ \leq& Q\left(\sqrt{\frac{\sum\limits_{k,d_{free}} \|\left(x(k) - \hat{x}(k)\right)H(k)\|^2}{2N_0}}\right) \end{aligned} \tag{12}$$

where $Q(\cdot)$ is the well-known Q-function. Let's denote $d(k) = x(k) - \hat{x}(k)$, which are non-zero for the d_{free} points considered for PEP analysis. In order to find PEP, $\sum_{k,d_{free}} \|d(k)H(k)\|^2$ has to be calculated. Using (6),

$$\begin{aligned} \sum_{k,d_{free}} \|d(k)H(k)\|^2 =& \sum_{k,d_{free}} \underline{h}^H \underline{W}_K(k) d^*(k) d(k) \underline{W}_K^H(k) \underline{h} \\ =& \underline{h}^H \left[\sum_{k,d_{free}} A_k\right] \underline{h} = \underline{h}^H \mathbf{A} \underline{h} \end{aligned} \tag{13}$$

where $\mathbf{A}$ and A_k's are $L \times L$ matrices and A_k is given by

$$A_k = |d(k)|^2 \begin{bmatrix} 1 & W_K^k & \cdots & W_K^{(L-1)k} \\ W_K^{-k} & 1 & \cdots & W_K^{(L-2)k} \\ \vdots & \vdots & \ddots & \vdots \\ W_K^{-(L-1)k} & W_K^{-(L-2)k} & \cdots & 1 \end{bmatrix}_{L\times L} \tag{14}$$

As can be seen from (14), the rank of each A_k matrix is one. However, due to the special form of the A_k matrices, the rank of the matrix $\mathbf{A} = \sum_{k,d_{free}} A_k$ is $r = rank(\mathbf{A}) = \min(d_{free}, L)$ (see Appendix for the proof). Note that each A_k is also Hermitian with a square root $\underline{W}(k)d^*(k)$ such that $A_k = \underline{W}(k)d^*(k)(\underline{W}(k)d^*(k))^H$. From linear algebra, it is known that any matrix with a square root is positive semidefinite [4], [20]. Also, any non-negative linear combination of positive semidefinite matrices is positive semidefinite. It follows that A_ks and $\mathbf{A}$ are positive semidefinite, and the singular value decomposition of $\mathbf{A}$ can be written as [20]

$$\mathbf{A} = V^H \Lambda V \tag{15}$$

where V is a $L \times L$ unitary matrix and Λ is a $L \times L$ diagonal matrix with eigenvalues of $\mathbf{A}$, $\{\lambda_i\}_{i=1}^{L}$, on the diagonal. Note that, eigenvalues of any positive semidefinite matrix are real and non-negative. Let's write $V\underline{\mathrm{h}} = [v_1 \quad v_2 \quad \cdots \quad v_L]^T$, then

$$\sum_{k,d_{free}} \|d(k)H(k)\|^2 = \underline{\mathrm{h}}^H \mathbf{A} \underline{\mathrm{h}} = \underline{\mathrm{h}}^H V^H \Lambda V \underline{\mathrm{h}} = \sum_{l=1}^{L} \lambda_l |v_l|^2 \tag{16}$$

Since each tap of the channel is modeled as an independent complex Gaussian random variable with zero mean and equal variance, v_l's are also complex Gaussian random variables. Then $|v_l|$'s are Rayleigh distributed with pdf $2|v_l|e^{-|v_l|^2}$. Using an upper bound for the Q function $Q(x) \leq (1/2)e^{-x^2/2}$, PEP can be upper bounded as

$$\begin{aligned} P(\underline{\mathrm{c}} \to \hat{\underline{\mathrm{c}}}) =& E\left[P(\underline{\mathrm{c}} \to \hat{\underline{\mathrm{c}}}|\mathbf{H})\right] \\ \leq& E\left[\frac{1}{2}\exp\left(-\frac{\sum_{l=1}^{L}\lambda_l|v_l|^2}{4N_0}\right)\right] = \frac{1}{\prod_{l=1}^{L}\left(1+\frac{\lambda_l}{4N_0}\right)} \end{aligned} \tag{17}$$

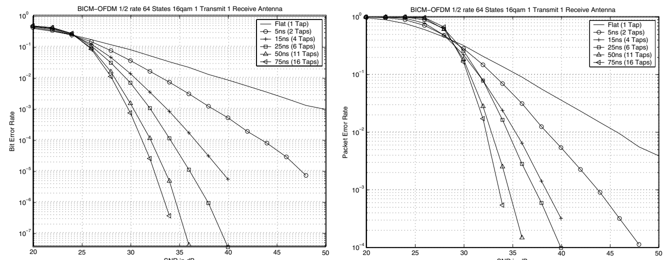
For the rank of $\mathbf{A}$, $r = \min(d_{free}, L)$, without loss of generality we can order the λ_l's such that, $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_r$ and $\lambda_{r+1} = \ldots = \lambda_L = 0$. Using $N_0 = 1/SNR$ from Section III, for high SNR values, PEP becomes upper bounded by

$$P(\underline{\mathrm{c}} \to \hat{\underline{\mathrm{c}}}) \leq \frac{1}{\prod_{l=1}^{r}\left(1+\frac{\lambda_l SNR}{4}\right)} \simeq \left(\prod_{l=1}^{r}\lambda_l\right)^{-1}\left(\frac{SNR}{4}\right)^{-r} \tag{18}$$

It can be easily seen from (18) that the diversity order of BICM-OFDM system is $r = \min(d_{free}, L)$. For example, the industry standard $1/2$ rate 64 state (133,171) convolutional encoder has $d_{free} = 10$. Therefore, a BICM-OFDM system with this convolution code can achieve a diversity order of 10 without implementing any additional antennas, or using any other diversity technique. In order to even further increase the diversity order of the system, one can in addition add multiple antennas using STBC to multiply the diversity order of BICM-OFDM with the number of transmit and receive antennas (see [21]). Or, multiple antennas can be used to increase the throughput of the system, while BICM-OFDM is used to provide the necessary diversity order. Also, a low complexity Viterbi decoder can be implemented for BICM-OFDM systems without any performance degradation [22] and [23]. Thus, a low complexity, easy to implement, and a high diversity order system can be easily generated by BICM-OFDM.

V. Simulation Results

In the simulations below we used 64 subcarriers for each OFDM symbol. One symbol has a duration of 4 μs of which 0.8 μs is CP. 250 bytes are sent with each packet and the channel is assumed to be the same through the transmission of one packet. The maximum delay spread of the channel is set to be ten times the root mean square (rms) delay spread.



(a) BER vs SNR

(b) PER vs SNR

Fig. 2. BICM-OFDM results using 1/2 rate 64 states $d_{free} = 10$ code

Figure 2 shows the simulation results for different rms delay spread values of the channel. (133,171) 1/2 rate 64 states convolutional code used in BICM-OFDM and interleaved bits are modulated using 16 QAM. As can be seen from the figure as the number of taps for the channel increases, its diversity increases as well. Another interesting observation is that while diversity for 50 ns and 75 ns channels reach the maximum diversity, 75 ns channel shows slightly better coding gain. This is due to the fact that the product of the eigenvalues in (18) is higher for higher delay spreads once the maximum diversity is reached.

VI. Conclusion

BICM and OFDM are used widely in many wireless communication systems. In this paper we showed the two can be combined to achieve a high diversity order. We illustrated both analytically and via simulations that the maximum diversity that is inherited in frequency selective channels can be fully and successfully achieved. If the convolutional code being used has a minimum Hamming distance of d_{free}, we showed that the diversity order of BICM-OFDM is $\min(d_{free}, L)$ for an L tap frequency selective fading channel. Simulations also showed that, when $L \geq d_{free}$, as the delay spread increases the coding gain increases, improving the system performance.

Appendix
Proof of rank $\min(d_{free}, L)$

Note that in general the number of subcarriers $K \geq d_{free}$ and $K \geq L$, and these are assumed to be the case in this paper. In order to have a clearer presentation we will denote $D = d_{free}$ and without loss of generality, we can reorder the D different A_k matrices so that $\mathbf{A} = \sum_{k=1}^{D} A_k$. Assume for now, $D \leq L$. Then, it is known that [20] $rank(\mathbf{A}) = r \leq \sum_{k=1}^{D} rank(A_k) = D$. Let's denote $a_k \triangleq W_K^k$. Note that $a_k^{-1} = a_k^*$, and a_ks lie on the unit circle on the complex plane and $a_i \neq a_j$ for $i \neq j$, $1 \leq i, j \leq K$. Then, A_ks become

$$A_k = |d(k)|^2 \begin{bmatrix} 1 & a_k & \cdots & a_k^{(L-1)} \\ a_k^{-1} & 1 & \cdots & a_k^{(L-2)} \\ \vdots & \vdots & \ddots & \vdots \\ a_k^{-(L-1)} & a_k^{-(L-2)} & \cdots & 1 \end{bmatrix}$$

$$\mathbf{A} = \sum_{k=1}^{D} A_k$$

$$= \begin{bmatrix} \sum_{k=1}^{D} |d_k|^2 & \sum_{k=1}^{D} |d_k|^2 a_k & \cdots & \sum_{k=1}^{D} |d_k|^2 a_k^{L-1} \\ \sum_{k=1}^{D} |d_k|^2 a_k^{-1} & \sum_{k=1}^{D} |d_k|^2 & \cdots & \sum_{k=1}^{D} |d_k|^2 a_k^{L-2} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{D} |d_k|^2 a_k^{-(L-1)} & \cdots & \cdots & \sum_{k=1}^{D} |d_k|^2 \end{bmatrix} \tag{A.1}$$

Clearly, if the rank of $\mathbf{A}$ is r, then there exists a sub-matrix within $\mathbf{A}$ of size $r \times r$ such that the determinant of the sub-matrix is nonzero [20]. Consider the sub-matrix A_D of size $D \times D$ of $\mathbf{A}$.

$$A_D = \begin{bmatrix} \sum_{k=1}^{D} |d_k|^2 & \sum_{k=1}^{D} |d_k|^2 a_k & \cdots & \sum_{k=1}^{D} |d_k|^2 a_k^{D-1} \\ \sum_{k=1}^{D} |d_k|^2 a_k^{-1} & \sum_{k=1}^{D} |d_k|^2 & \cdots & \sum_{k=1}^{D} |d_k|^2 a_k^{D-2} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{D} |d_k|^2 a_k^{-(D-1)} & \cdots & \cdots & \sum_{k=1}^{D} |d_k|^2 \end{bmatrix}$$

$A_D = B_D C_D$ where

$$B_D = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_1^{-1} & a_2^{-1} & \cdots & a_D^{-1} \\ a_1^{-2} & a_2^{-2} & \cdots & a_D^{-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{-(D-1)} & a_2^{-(D-1)} & \cdots & a_D^{-(D-1)} \end{bmatrix}$$

$$C_D = \begin{bmatrix} |d_1|^2 & |d_1|^2 a_1 & \cdots & |d_1|^2 a_1^{(D-1)} \\ |d_2|^2 & |d_2|^2 a_2 & \cdots & |d_2|^2 a_2^{(D-1)} \\ \vdots & \vdots & & \vdots \\ |d_D|^2 & |d_D|^2 a_D & \cdots & |d_D|^2 a_D^{(D-1)} \end{bmatrix} \tag{A.2}$$

It is evident that, since $a_k^{-1} = a_k^*$, B_D^H is a Vandermonde matrix of size D. The determinant of a Vandermonde matrix can be easily calculated by [20]

$$\det(B_D^H) = \prod_{\substack{i,j \\ i>j}}^{D} (a_i - a_j) \tag{A.3}$$

which is non-zero, since $a_i \neq a_j$ for $i \neq j, 1 \leq i, j \leq D \leq K$. Therefore $rank(B_D^H) = D = rank(B_D)$ and B_D is full rank. Note that, C_D is equal to B_D^H with each row multiplied by a positive scalar. Since multiplying rows of a matrix with nonzero scalars does not change the rank of a matrix, C_D is also full rank with rank D. This shows $\det(A_D) = \det(B_D)\det(C_D)$ is nonzero, confirming A_D is a full rank matrix with rank D. Since A_D is a sub-matrix of $\mathbf{A}$, then $rank(\mathbf{A}) \geq D = d_{free}$, concluding $rank(\mathbf{A}) = D \leq L$.

If $L < D$, then $\mathbf{A}$ is a sub-matrix of A_D. Again from (A.2) and (A.3), A_D is a full rank matrix with rank D due to the fact that $a_i \neq a_j$ for $i \neq j$, $1 \leq i, j \leq D \leq K$. Since any sub-matrix of a full rank matrix is also full rank, then $\mathbf{A}$ is full rank with $rank(\mathbf{A}) = L$. Consequently, $rank(\mathbf{A}) = \min(D, L) = \min(d_{free}, L)$.

References

[1] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, no. 5, pp. 873–884, May 1992.

[2] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 1, pp. 55–67, January 1982.

[3] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, May 1998.

[4] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, March 1998.

[5] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, October 1998.

[6] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[7] ——, "Space-time block coding for wireless communications: Performance results," *IEEE J. Select. Areas Commun.*, vol. 17, no. 3, pp. 451–460, March 1999.

[8] H. Bolcskei and A. J. Paulraj, "Space-frequency coded broadband OFDM systems," in *Proc. WCNC*, vol. 1, September 2000, pp. 1–6.

[9] B. Lu and X. Wang, "Space-time code design in OFDM systems," in *Proc. IEEE GLOBECOM*, vol. 2, 27 Nov - 1 Dec 2000, pp. 1000–1004.

[10] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri, "Space-time coded OFDM for high data-rate wireless communication over wideband channels," in *Proc. VTC*, vol. 3, May 1998, pp. 2232–2236.

[11] H. Bolcskei and A. J. Paulraj, "Space-frequency codes for broadband fading systems," in *Proc. ISIT*, June 2001, p. 219.

[12] H. E. Gamal, A. R. H. Jr., Y. Liu, M. P. Fitz, and O. Y. Takeshita, "On the design of space-time and space-frequency codes for MIMO frequency-selective fading channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 9, pp. 2277–2292, September 2003.

[13] Z. Liu and G. B. Giannakis, "Space-time-frequency coded OFDM over frequency-selective fading channels," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2465–2476, October 2002.

[14] Z. Hong and B. L. Hughes, "Robust space-time codes for broadband OFDM systems," in *Proc. IEEE WCNC'02*, vol. 4, March 2002, pp. 105–108.

[15] D. Rende and T. F. Wong, "Bit interleaved space-frequency coded modulation for OFDM systems," in *Proc. IEEE ICC'03*, vol. 4, May 2003, pp. 2827–2831.

[16] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel codes: Theoretical concepts and practical design rules," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, July 1999.

[17] X. Li and J. A. Ritcey, "Trellis-coded modulation with bit interleaving and iterative decoding," vol. 17, no. 4, pp. 715–724, April 1999.

[18] S. Lin and D. J. Costello, *Error Control Coding: Fundamentals and Applications*. Prentice-hall, 1983.

[19] Y. Liu, M. P. Fitz, and O. Y. Takeshita, "Outage probability and space-time code design criteria for frequency selective fading channels with fractional delay," in *Proc. IEEE ISIT*, June 2001.

[20] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge Unversity Press, 1990.

[21] E. Akay and E. Ayanoglu, "Bit interleaved coded modulation with space time block codes for OFDM systems," in *IEEE VTC Fall '04*, Los Angeles, USA, September 2004.

[22] ——, "Low complexity decoding of bit interleaved coded modulation," in *Proc. IEEE ICC '04*, Paris, France, June 2004.

[23] ——, "High performance viterbi decoder for OFDM systems," in *Proc. IEEE VTC Spring*, Milan, Italy, May 2004.