High Fidelity OFDM Simulation using MATLAB

Paul F. Roysdon, Ph.D.

Abstract— **Many applications demand multi-carrier wireless transmission in various environments. To evaluate this problem we consider an emerging technology for high data rates, Orthogonal Frequency Division Multiplexing (OFDM). This paper applies OFDM to a robust, high-fidelity, MATLAB simulation with a user interface for BPSK, QPSK, 16-PSK and 256-PSK modulation schemes with variable IFFT size, Number of Carriers, SNR, and Amplitude clipping by the communication channel. The simulation emulates a modern embedded Digital Signal Processor (DSP) and is applied to a 8-bit grey-scale image. An analysis of the transmitted results are presented, and a demonstration of the trade-off of Bit Error Rate vs. data-rate is provided.**

I. Introduction

Traditional single carrier transmission, the symbol period must be much greater than the delay time to avoid inter-symbol interference (ISI). Since data rate is inversely proportional to the symbol period, long symbol periods result in low data rate and communication inefficiency. Multi-carrier systems such as Frequency Division Multiplexing (FDM) divide the available bandwidth into several sub-carriers transmitted in parallel, resulting in higher data rates. However to avoid ISI, guard bands must be placed between adjacent carriers, resulting in a lower overall bandwidth. Orthogonal Frequency Division Multiplexing (OFDM) is a particularly clever multi-carrier transmission well suited for frequency selective channels and high data-rates. This technique transforms a frequency-selective wide-band channel into a group of non-selective narrow-band channels, which makes it robust against large delay spreads by preserving orthogonality in the frequency domain. Moreover, the addition of a cyclic redundancy at the transmitted signal reduces the complexity to only FFT processing, with one tap scalar equalization at the receiver.

The contributions of this paper are:

1) A novel way to use MATLAB to simulate wireless transmitter and receiver hardware.
2) A complete tutorial of OFDM with theory and application.

The document is organized as follows: First we review the available literature in Section II, then introduce the background and theory of OFDM in Section III, the MATLAB implementation is described in Section IV, the experimental results are provided and analyzed in Section V, and the conclusions and future work are discussed in Section VI.

II. Literature Review

More than 30 years ago, multi-carrier modulation was proposed, yet at the time only an analog design based on orthogonal waveforms was possible [1], [2]. The use of Discrete Fourier Transform (DFT) for modulation and demodulation (via the Inverse DFT) was first proposed in [3]. It was not until recently that OFDM emerged into commercial use, as recent developments in digital signal processing (DSP) and micro-processor technology has lowered the cost of the hardware and software necessary to implement OFDM systems [4], [5], [6].

III. Theory

The objective of this research paper is to familiarize the author, and demonstrate the concept of an OFDM system, through an investigation of its structure and how its performance is changed by varying some of its major parameters. This objective is met by developing a MATLAB simulation to synthetically model the transmit and receive signals of a basic OFDM system. Thus, the mechanism of an OFDM system can be studied, and the characteristics can be explored.

A. Overview

In digital communication systems, information or data is expressed in the form of bits. In OFDM, a symbol refers to a collection of various lengths of bits. Data is generated by applying symbols in the spectral space using various modulation schemes such as M-PSK, QAM, etc., and converting the spectra to the time domain by taking the Inverse Discrete Fourier Transform (IDFT). However in DSP's, is it more common to use the Inverse Fast Fourier Transform (IFFT) as it greatly reduces the computational load. Once the data is in the time domain, all carriers transmit in parallel to fully occupy the available frequency bandwidth.

During modulation, symbols are typically divided into frames such that the receiver will maintain sync with the received signal. In OFDM, symbols typically have long symbol periods which reduce the probability of ISI. To further reduce ISI, a guard interval, known as cyclic extension, or prefix, is added to each symbol. The cyclic prefix (CP) length is roughly 25% of the symbol, copied from the end, and added to the beginning of each symbol. The CP allows a signal uncertainty of up to the length of the CP and still enable the FFT to obtain the information of the entire symbol. Figure 1 illustrates the concept of CP.

B. Parameters and Characteristics

The number of carriers in an OFDM system is not only limited by the available spectral bandwidth, but also by the IFFT size based on the following relationship:

$$number\ of\ carriers \leq \frac{IFFT\ size}{2} - 2 \tag{1}$$

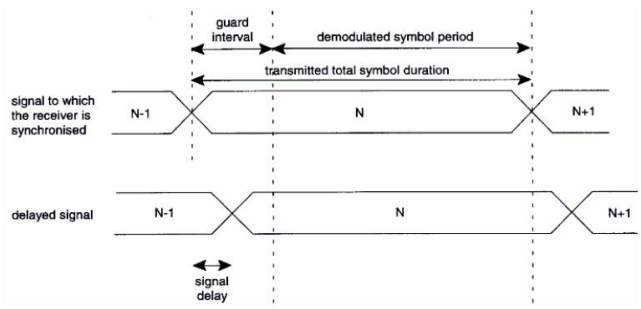

Fig. 1. Cyclic extension with and without signal delay.

which is determined by the complexity of the system. Increasing the complexity, increases the IFFT size (and computational cost); thus an increased number of carriers can be used, and higher data transmission rate achieved. Through schemes such as Phase Shift Keying (PSK), the modulation varies the data rate and Bit Error Rate (BER). Increased *M*-order (*M*-PSK) leads to larger symbol size, thus fewer symbols need to be transmitted and higher data rate is achieved. This comes at the cost of higher BER, due to reduced spacing of phases, as the sub-regions are divided on the range of 0-360 degrees. As the regions decrease in size (or spacing), the spacing (or margin) between the phases decreases, which increases the probability for incorrect decoding. OFDM signals also have high peak-to-average ratio, therefore it has a relatively high tolerance of peak power clipping due to transmission limitations.

C. Baseband vs. Passband

OFDM signals are *complex* baseband signals which must be converted to *real* passband signals for communication transmission. The complex baseband signal, $s(t)$ can be written as:

$$s(t) = s_I(t) + js_Q(t) \tag{2}$$

Where the real part, $s_I(t)$ is called the *in-phase component*, and the imaginary part $js_Q(t)$, is called the *quadrature component*. Thus the OFDM baseband signal becomes:

$$\begin{aligned} s_p(t) &= \sum_{k=0}^{N-1} (\Re\{s_k\}\cos(2\pi f_k t) - \Im\{s_k\}\sin(2\pi f_k t)) \\ &+ j\sum_{k=0}^{N-1} (\Re\{s_k\}\cos(2\pi f_k t) + \Im\{s_k\}\sin(2\pi f_k t)) \end{aligned} \tag{3}$$

where $\Re\{s_k\}$ and $\Im\{s_k\}$ denote the real and imaginary parts of the complex symbol s_k, and can be interpreted graphically as shown in figure 2.

Conversion from baseband to passband is as follows:

$$\begin{aligned} s_p(t) &= \Re\{s(t)\exp^{j2\pi f_c t}\} \\ &= s_I(t)\cos(2\pi f_c t) - s_Q(t)\sin(2\pi f_c t) \end{aligned} \tag{4}$$

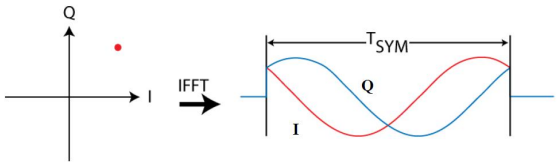

Fig. 2. Traditional in-phase and quadrature components.

where f_c is the carrier frequency of the system. We can further simplify the OFDM passband signal:

$$\begin{aligned} s_p(t) &= s_I(t)\cos(2\pi f_c t) - s_Q(t)\sin(2\pi f_c t) \\ &= \sum_{k=0}^{N-1} (\Re\{s_k\}\cos(2\pi(f_c + f_k)t) - \Im\{s_k\}\sin(2\pi(f_c + f_k)t)) \end{aligned} \tag{5}$$

The magnitude d_k and phase θ_k of the baseband complex symbol, $s_k = d_k \exp^{j\theta_k}$ can be represented in passband as:

$$s_p(t) \sum_{k=0}^{N-1} d_k \cos(2\pi(f_c + f_k)t + \theta_k) \tag{6}$$

D. Orthogonality

The clever aspect of OFDM is maintaining orthogonality in the frequency domain. If the integral of the product of two signals is zero over an arbitrary time period, then these two signals are said to be orthogonal to each other. Two sinusoids with frequencies that are integer multiples of a common frequency satisfy this criterion. Therefore, orthogonality is defined by:

$$\int_0^T \cos(2\pi n f_c t)\cos(2\pi m f_c t)dt = 0 \ (n \neq m) \tag{7}$$

where n and m are two unequal integers; f_c is the carrier frequency; T is the period over which the integration is taken. For OFDM, T or T_{SYM} is one symbol period and f_c is set to $\frac{1}{T}$ for optimal effectiveness.

E. Fading Channels

In wireless communication systems, the channel is subject to destructive interference due to multi-path and obstacles, both of which distort the signal as it traverses from the source to the receiver. This introduces the concept of fading, which is independent of frequency and location, and can vary with time.

Fading based on Time-Variations result in what is referred to as either Fast or Slow fading. Fading based on Time-Spreading results in either Flat fading, or Frequency Selective fading. Strong destructive interference is commonly referred to as *Deep Fading*, which is due to a severe drop in the channel SNR, and results in the temporary failure of the channel.

1) Fading based on Time-Variations: Time varying fading is defined as the rate at which the magnitude and phase change imposed by the channel on the signal changes. The coherence time T_{coh} is the minimum time required for the magnitude or phase change of the channel to become uncorrelated from its previous value.

$$T_{coh} \ll D_S \approxeq 1/4D_S \tag{8}$$

Slow fading occurs when the coherence time of the channel is large relative to the delay constraint or symbol period of the channel, T_S. It is roughly equal to the inverse of the Doppler Spread, D_S, when the Doppler Spread is low, and is the result of obstructions (known as shadowing). Both amplitude and phase change imposed by the channel can be considered constant over the period.

$$1/D_S \approxeq T_{coh} > T_S \tag{9}$$

where the Doppler Spread is defined as:

$$D_S = \max_{i,j} f_c(v_i - v_j) \tag{10}$$

Fast fading occurs when the coherence time of the channel is small relative to symbol period. It is roughly equal to the inverse of the Doppler Spread, when the Doppler Spread is high. Both amplitude and phase change imposed by the channel can vary considerably over the period.

$$1/D_S \approxeq T_{coh} < T_S \tag{11}$$

Doppler shift is the change in frequency due to a velocity difference between the source and the receiver, as well as influences from the environment. Doppler Spread is the result of different Doppler shifts, resulting in different rates of change in phase, contributing to a single fading channel.

2) Fading based on Time-Spreading: Time-Spreading fading is defined as a signal which arrives at the receiver by two different paths, one of which is changing, resulting in an interference with itself. Furthermore, as the carrier frequency of a signal is varied, the magnitude of the change in amplitude will vary.

Flat fading occurs when the coherence bandwidth of the channel, B_{coh}, is larger than the bandwidth of the signal, B_S. Therefore, all frequency components of the signal will experience the same magnitude of fading, and the spectral characteristics of transmitted signal are preserved. This is typically modeled as Rayleigh, or Ricean distribution.

$$B_S < B_{coh} \tag{12}$$

Frequency Selective fading occurs when the coherence bandwidth of the channel is smaller than the bandwidth of the signal. Different frequency components of the signal therefore experience uncorrelated fading, and the spectral characteristics of the transmitted signal are not preserved, however multi-path components are resolved.

$$B_S > B_{coh} \tag{13}$$

Frequency selectivity of the channel, as we will see in the following section, has desirable characteristics for OFDM because the frequency components of the signal are affected independently, and it is highly unlikely that all parts of the signal will be simultaneously affected by deep fade.

Fading can be partially mitigated through methods of diversity in time, frequency, or space, and commonly employ techniques such as: Diversity reception and transmission, MIMO, OFDM, Rake receivers, and Space−time codes.

F. Frequency Selective Channels

To better understand OFDM, it is necessary to first review frequency selectivity. Let $r(t)$ be the low-pass received signal:

$$r(t) = \int_{-\infty}^{\infty} c(\tau)x(t-\tau)d\tau + n(t). \tag{14}$$

From [7], [8], frequency selectivity occurs whenever the transmitted signal $x(t)$ occupies an interval bandwidth $[-\frac{W}{2}, \frac{W}{2}]$ greater than the coherence bandwidth B_{coh} of the channel $c(t)$. $c(t)$ is defined as the inverse of the delay spread T_d. Thus the frequency components of $x(t)$ with frequency separation greater than B_{coh} are subject to different channel gains (this is a key element of OFDM).

Consider a typical time impulse response $c(t)$ of a channel. For common high data-rates schemes, the symbol rate T is small compared to T_d, thus the signals are subject to frequency selectivity (this is also referred to as broadband signals).

Now consider the multi-path channel, modeled as an impulse response given by:

$$c(t) = \sum_{l=0}^{M-1} \lambda_l g(t - \tau_l) \tag{15}$$

where $g(t)$ is transmitting filter and T_d is the duration of the multi-path or delay spread T_d. The multi-path complex gains are $(\lambda_l)_{l=0,\ldots,M-1}$ and $(\tau_l)_{l=0,\ldots,M-1}$ are the corresponding time delays. The variance of each gain as well as the time delays can be determined from propagation measurements. Let $T = \frac{1}{W}$ be the sampling rate, and W denote the signal bandwidth. To simplify the equations we will assume an ideal transmitting filter ($G(f) = 1$ for $f \in [-\frac{W}{2}, \frac{W}{2}]$ and 0 outside, where $G(f)$ is the Fourier transform of the transmitting filter $g(t)$).

To recover the transmitted signal $x(t)$, a method called equalization is required, and is complicated by the frequency selectivity behavior of the channel. Furthermore, as complexity increases, so does the equalizer with channel memory. Thus, with increasing data-rates, the cost of an equalizer becomes prohibitively expensive.

This problem is easily solved with OFDM, as the channel convolution is changed to a multiplicative problem via FFT's, greatly simplifying the task of equalization. Additionally,

OFDM adds redundancy, the CP, which circularizes the channel effect. From [9] we see that circular convolution can be diagonalized in an FFT basis, the multi-path time domain channel is transformed into a set of parallel frequency flat fading channels. Furthermore, OFDM exploits newer DSP technology of digital FFT modulators. From each of these advantages, we can see the distinct advantage of OFDM over other multi-carrier technologies. Figure 3, illustrates the OFDM flow diagram.

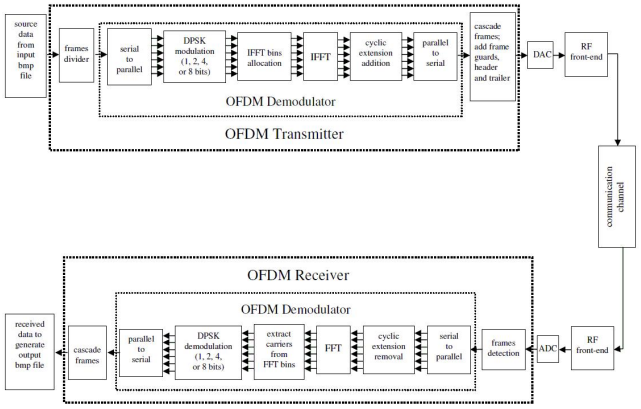

Fig. 3. OFDM model.

G. Conventional OFDM

Consider the conventional OFDM; for simplicity we will ignore noise for now. Incoming high-rate data is first split into N-rate sub-carriers, and transmitted in blocks of size N : $s(k) = [s_1(k), ..., s_i(k), ..., s_N(k)]$ where k is the block OFDM symbol number, and i is the carrier index. Each block symbol is precoded by an IFFT matrix $\mathbf{F_N^H} = \mathbf{F_N^{-\infty}}$ to yield a time domain block vector $x(k) = [x_1(k), ..., x_i(k), ..., x_N(k)]$. After the IFFT, a guard interval of D samples is inserted at the beginning of each symbol $[x_{N-D+1}(k), ..., x_N(k), x_1(k), ..., x_i(k), ..., x_N(k)]$, consisting of a cyclic extension (roughly the last 25%) of the time domain symbol of size larger than the channel impulse response ($D > L - 1$). The CP is appended to each symbol to transform the multi-path linear convolution into a circular convolution. Finally a Parallel to Serial (P/S) and Digital to Analog Conversion (DAC) is performed before the signal is sent through a frequency-selective channel. The figures 4 & 5 illustrate the time versus frequency domain for multiple carriers. Figure 6 illustrates the cyclic prefix used in OFDM.

Applying the discrete time model to the channel, it can be represented by a linear Finite Impulse Response (FIR) filter with channel impulse response $c_N = [c_1, ..., c_{L-1}, 0, ..., 0]$. Generally we can assume that D is smaller than N, where $D = \frac{N}{4}$, and greater than $(L-1)$. Notice that the redundancy factor is equal to $\frac{N}{N+D}$. To avoid spectrally inefficient transmissions, N must be greater than D, based on:

$$\lim_{N\to\infty} \frac{N}{N+D} = 1 \tag{16}$$

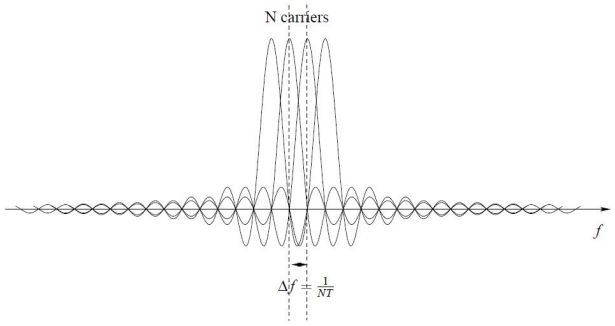

Fig. 4. Frequency representation of OFDM.

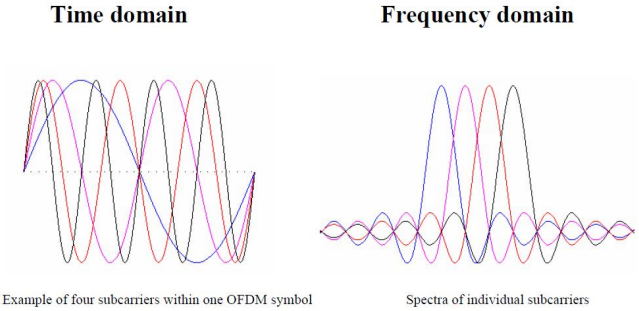

Fig. 5. Time vs. Frequency Domain.

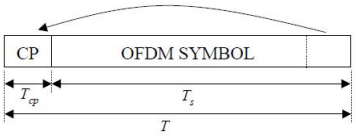

Fig. 6. Time representation of OFDM.

Where D is fixed, and the redundancy factor approaches unity as the number of carriers increases. However, the FFT complexity per carrier grows with the size of N. Notice also that to maintain circular convolution, the channel should not change inside one OFDM symbol. Finally, the carrier spacing is related to the factor $\frac{1}{NT}$ and reduces as N increases. This implies that there is no diversity gain in a fixed channel by increasing N. Therefore the choice of N depends on the user and and type of channel (slow varying, fast fading, high diversity channel, impulse response length, etc.) and complexity desired.

Symmetric operations are performed at the receiver: down conversion, Analog to Digital Conversion (ADC), etc.. The discrete time received signal with cyclic prefix, r^{CP}, has the following expression:

$$\mathbf{r^{CP}} = \begin{bmatrix} r_1^{CP}(k) \\ \vdots \\ r_{N+D}^{CP}(k) \end{bmatrix}_{\mathbf{(N+D)x1}}$$

which is:

$$\mathbf{r^{CP}} = \mathbf{H_{ISI}} \begin{bmatrix} x_{N-D+1}(k) \\ \vdots \\ x_N(k) \\ x_1(k) \\ \vdots \\ x_N(k) \end{bmatrix}_{\mathbf{(N+D)x1}} + \mathbf{H_{IBI}} \begin{bmatrix} x_{N-D+1}(k-1) \\ \vdots \\ x_N(k-1) \\ x_1(k-1) \\ \vdots \\ x_N(k-1) \end{bmatrix}_{\mathbf{(N+D)x1}}$$

where H_{ISI} is:

$$\mathbf{H_{ISI}} = \begin{bmatrix} c_0 & 0 & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & & & \vdots \\ c_{L-1} & & \ddots & \ddots & & \vdots \\ 0 & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \dots & 0 & c_{L-1} & \dots & c_0 \end{bmatrix}_{\mathbf{(N+D)x(N+D)}}$$

where H_{IBI} is:

$$\mathbf{H_{IBI}} = \begin{bmatrix} 0 & \dots & 0 & c_{L-1} & \dots & c_1 \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & & \ddots & & \ddots & c_{L-1} \\ \vdots & & & \ddots & & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & \dots & \dots & \dots & \dots & 0 \end{bmatrix}_{\mathbf{(N+D)x(N+D)}}$$

and:

$$\begin{bmatrix} x_1(k) \\ \vdots \\ x_N(k) \end{bmatrix}_{Nx1} = \mathbf{F_N^H} \begin{bmatrix} s_1(k) \\ \vdots \\ s_N(k) \end{bmatrix}_{\mathbf{Nx1}}$$

$\mathbf{H_{ISI}}$ represents ISI generated by the frequency selective channel inside an OFDM block at time $k[x_{N-D+1}(k), ..., x_N(k), x_1(k), ..., x_i(k), ..., x_N(k)]$ while $\mathbf{H_{IBI}}$ corresponds to inter-block interference (IBI) between two consecutive OFDM block transmissions at time $k[x_{N-D+1}(k), ..., x_N(k), x_1(k), ..., x_i(k), ..., x_N(k)]$ and at time $(k-1)[x_{N-D+1}(k-1), ..., x_N(k-1), x_1(k-1), ..., x_i(k-1), ..., x_N(k-1)]$.

Here we denote channel transfer function as:

$$H(z) = \sum_{k=0}^{L-1} c_k z^{-k} \tag{17}$$

and the Fourier Transform as:

$$\begin{aligned} \mathbf{H} &= \mathbf{F_{N-1}} \mathbf{c_N} \\ &= [H(0), H(e^{j2\pi/N}), ..., H(e^{(j2\pi N-1)/N})]^T \\ &= [h_1, ..., h_N]^T \end{aligned} \tag{18}$$

The first D samples of the received signal r^{CP} are then discarded to reduce inter-block interference.

$$\begin{bmatrix} r_{D+1}^{CP}(k) \\ \vdots \\ r_{N+D}^{CP}(k) \end{bmatrix}_{Nx1} = \begin{bmatrix} r_1(k) \\ \vdots \\ r_N(k) \end{bmatrix}$$

which is:

$$\begin{bmatrix} & c_{L-1} & \dots & c_0 & & 0 \\ & & \ddots & & \ddots & \\ 0_{D-(L-1)} & & & \ddots & & \\ \vdots & & & & \ddots & \\ & 0 & & c_{L-1} & \dots & c_0 \end{bmatrix} \times \begin{bmatrix} x_{N-D+1}(k) \\ \vdots \\ x_N(k) \\ x_1(k) \\ \vdots \\ x_N(k) \end{bmatrix}_{(N+D)x1}$$

which can be written as:

$$\begin{bmatrix} r_1(k) \\ \vdots \\ r_N(k) \end{bmatrix}_{Nx1} = \begin{bmatrix} c_0 & 0 & \dots & c_{L-1} & & \dots & c_1 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ c_{L-1} & & \ddots & \ddots & & & c_{L-1} \\ 0 & \ddots & & \ddots & & \ddots & \vdots \\ \vdots & \ddots & \ddots & & c_{L-1} & \ddots & 0 \\ 0 & \dots & 0 & c_{L-1} & c_{L-1} & \dots & c_0 \end{bmatrix}_{NxN} \times \mathbf{F_N^H} \begin{bmatrix} s_1(k) \\ \vdots \\ \vdots \\ \vdots \\ s_N(k) \end{bmatrix}_{\mathbf{Nx1}}$$

We can see from the previous equations that cyclic redundancy has enabled us to turn the linear convolution into a circular convolution. Recall from the properties of the DFT, provided $x(n)$ and $h(n)$ are finite length sequences, their linear convolution is:

$$y(n) = x(n) * h(n) \tag{19}$$

The N-point circular convolution of $x(n)$ with $h(n)$ is:

$$h(n) \circledast x(n) = [\sum_{k=-\infty}^{\infty} y(n+kN)] \Re_N(n) \tag{20}$$

Thus, the circular convolution of two sequences is found by performing the linear convolution and aliasing the result. It follows, if $y(n)$ is of length N or less, $y(n-kN)\Re_N(n) = 0$ for $k \neq 0$ and, $h(n) \circledast x(n) = h(n) * x(n)$.

From [10], any circulant matrix is diagonal in the Fourier basis, therefore with an FFT we can diagonalize the channel effects:

$$\begin{bmatrix} y_1(k) \\ \vdots \\ y_N(k) \end{bmatrix}_{Nx1} = \mathbf{F_N} \begin{bmatrix} r_1(k) \\ \vdots \\ r_N(k) \end{bmatrix}_{\mathbf{Nx1}} \tag{21}$$

Which can be expanded to:

$$\begin{bmatrix} y_1(k) \\ \vdots \\ y_N(k) \end{bmatrix}_{Nx1}$$

$$= \mathbf{F_N} \begin{bmatrix} c_0 & 0 & \dots & c_{L-1} & & \dots & c_1 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ c_{L-1} & & \ddots & \ddots & & & c_{L-1} \\ 0 & \ddots & & \ddots & & \ddots & \vdots \\ \vdots & \ddots & \ddots & & c_{L-1} & \ddots & 0 \\ 0 & \dots & 0 & c_{L-1} & c_{L-1} & \dots & c_0 \end{bmatrix}_{\mathbf{NxN}}$$

$$\times \mathbf{F_N^H} \begin{bmatrix} s_1(k) \\ \vdots \\ s_N(k) \end{bmatrix}_{\mathbf{Nx1}}$$

Because circular convolution yields a multiplication in the frequency domain, $[s_1(k), ..., s_N(k)]$ is transmitted over N parallel flat fading channels, with complex frequency attenuation h_i

$$\begin{bmatrix} y_1(k) \\ \vdots \\ y_N(k) \end{bmatrix}_{Nx1}$$

$$= \begin{bmatrix} h_1 & 0 & \dots & \dots & \dots & \dots \\ 0 & h_2 & 0 & & & \vdots \\ \vdots & 0 & h_3 & 0 & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & h_N \end{bmatrix}_{NxN} \times \begin{bmatrix} s_1(k) \\ \vdots \\ s_N(k) \end{bmatrix}_{Nx1}$$

In the case of a transmission with Additive White Gaussian Noise (AWGN), the time Gaussian noise vector $[b_1(k), ..., b_N(k)]^T$ is multiplied at the receiver by the FFT demodulator:

$$\begin{bmatrix} n_1(k) \\ \vdots \\ n_N(k) \end{bmatrix}_{Nx1} = \mathbf{F_N} \begin{bmatrix} b_1(k) \\ \vdots \\ b_N(k) \end{bmatrix}_{\mathbf{Nx1}} \tag{22}$$

Based on [9] the statistics of a Gaussian vector does not change by orthogonal transform, and $[n_1(k), ..., n_N(k)]^T$ is an AWGN vector with the same variance.

H. The Pro's and Con's of OFDM

1) Complexity: For a given delay spread, ODFM complexity does not grow as fast as similar systems which use only a single carrier with an equalizer. Take for example the single carrier case when the sampling rate is reduced by a factor of two, the resulting equalizer must be doubled in both length and speed, thus complexity grows quadratically with the inverse of the sampling rate and a matrix inversion is required. Alternatively, multi-carrier OFDM grows only slightly faster than linearly, primarily due to the fact that scalar equalization is performed at the receiver. This remains true only if the impulse response of the channel is shorter than the cyclic prefix, there is no ISI, and each constellation is multiplied by the channel frequency coefficient.

However, the channel still has to be compensated by a multiplication of each FFT output by a single coefficient:

$$\hat{s}_i(k) = g_i h_i s_i(k) + g_i(k) n_i(k) \tag{23}$$

The matrix equivalent equation is: $\hat{s}_i(k) = \mathbf{Gy(k)}$ with $\mathbf{G}$ as a diagonal matrix.

We find in the literature, other equalization schemes such as zero forcing (ZF) or minimum mean-square error (MMSE) equalization which accounts for noise.

ZF equalization: $g_i = \frac{h_i^*}{|h_i|^2} = \frac{1}{h_i}$

MMSE equalization: $gi = \frac{h_i^*}{|h_i|^2 + \sigma_i^2}$

where σ_i^2 is the noise variance on the carrier i. And the coefficients $(h_i(k))_{i=1,...,N}$ can either be known or estimated.

2) Attenuation: In multi-channel systems, the noise variance changes for each carrier depending on the channel frequency response. Fortunately in the frequency domain, the channel attenuation can be determined by a learning sequence [11], or blind estimation [12]. A method called denoising estimation exploits the time structure of the channel. Turbo estimation [13], uses the time and frequency autocorrelation function of the channel. For example, IEEE802.11a systems apply two OFDM blocks at the start and end of each frame for synchronization.

3) Efficiency: Efficiency, in the spectral sense, is increased in OFDM due to frequency overlapping with different carriers.

4) Diversity: Unlike single-carrier systems, OFDM is limited in channel diversity which limits its use in fading environments. Due to flat fading, the information on one OFDM sub-channel can be lost entirely if deep fade occurs, [5]. We also see that OFDM diversity will be less than a single-carrier system employing the same error correcting code with high diversity. Moreover, the performance can be greatly impacted if we consider the Rayleigh behavior the same environment without the use of coded schemes such as: convolution codes, block codes, multidimensional constellations, turbo-codes [14], [15], [16], [17], [18], [19], [20] with interleaving.

Interleaving applies replicas of the same information to different sub-channels such that each sub-channel is subject to different fading, resulting in a higher probability that the

receiver will be able to recover the information. Interleaving has three primary forms:

- *Time Interleaving* sends coded bits at different time intervals longer than the coherence time of the channel, which works well in fast-fading environments, but results in longer delays in the received information.
- *Frequency Interleaving* coded bits are sent over different frequencies separated by the coherence bandwidth, and is suitable for scattering environments.
- *Space Interleaving* coded bits are sent over different antennas, and commonly employ the simple space-time Alamouti scheme, [21].

IV. Implementation

Modeling of the OFDM system in Matlab, required each block shown in figure 3 to be implemented as a function with defined input and output variables. To make the simulation more meaningful, an 8-bit gray-scale (256 gray levels) bitmap image file (*.bmp) was used as the source data instead of a greatly simplified signal. To transmit, the image data is converted to symbols ($size = bits/symbol$) with size based on a user choice of the *M*-PSK modulation, which is then separated into blocks by the OFDM transmitter function. Before the blocks are transmitted, the CP's are inserted. The channel is modeled with AWGN and amplitude clipping. The receiver detects the start and end of each block, then demodulation, and upon completion of the transmission, the data is assembled into the output image file. Simulation diagnostics are performed at the end of each run to present the user with valuable transmission statistics and figures.

A. I/O

The program input requires a $n \times m$ image file, where n is the height of the image (in pixels) and m is the width. Note that the source file is an 8-bit gray-scale bitmap, therefore its word size is 8 bits/word. When converted, the word size will depend on the modulation symbol selected. For example, if QPSK is chosen (4 bits/word) the data stream of 8-bits/word is [36, 182, 7] will go through the following process:

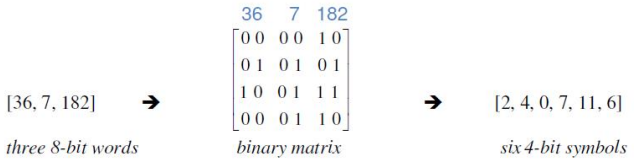

Fig. 7. QPSK example conversion of 8-bit words to 4-bit words.

If the receiver does not recover the entire symbol after demodulation, i.e. has missing bits, then the remaining demodulated bits are used and converted for the output image. This can result in incorrect decoding of pixels (gray scale value), missing pixels, or incorrect length of rows and columns. In the case of the latter, the receiver will pad the missing information with its nearest estimate of the missing bits. Similarly, if the data stream is longer than than expected due to incorrect processing caused by channel noise, the receiver will trim the data to the correct length. During the simulation, these occurrences are logged as errors, which are realistic and useful for evaluation of the channel.

B. Transmitter

1) Modulator: The modulator takes input data and divides it into frames based on the symbol size which is defined as $symbols\ per\ frame = ceil(2^{13}/carrier\ count)$ which limits the total number of symbols which can fill a frame within the interval $[2^{13}, 2 \cdot (2^{13-1})]$, or $[8192, 16382]$. These bounds were chosen as limits for the simulation, and do reflect commonly available DSP's. In the event that the total number of symbols are fewer than the total per frame, then the symbols are modulated at once.

2) Headers & Frame Guards: To assist the receiver, each frame contains a header and frame guard at the beginning and end of the modulated signal, as illustrated in figure 8. This assists the receiver in recovering the data. The headers are scaled to the RMS level of the modulated time signal.

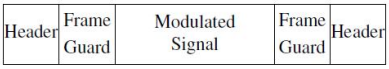

Fig. 8. Header, Frame guard, and modulated signal.

3) Modulation: During modulation, the data is converted from serial to parallel paths. If the case exists where the data stream does not entirely fill the carriers, zeros are used to pad the data stream. Take for example a data stream with 400 carriers and capacity of 30 symbols/carrier, any data set fewer than 12,000 total symbols will be padded with zeros. Figure 9 illustrates this example with a 2D matrix representing a carrier, with rows as one symbol period.

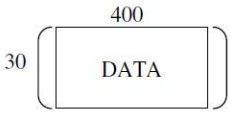

Fig. 9. Initial data matrix.

Continuing our example, we apply QPSK modulation. Consider the example illustrated in figure 10, where every symbol under QPSK modulation is translated to its corresponding phase value from 0-360 degrees and unitary magnitude, resulting in complex valued phase and magnitude. Also modeled in simulation were, BPSK which has two phases located at $\frac{\pi}{2}$ and $\frac{3\pi}{2}$, 16-PSK which has 16 equidistant phases, and 256-PSK.

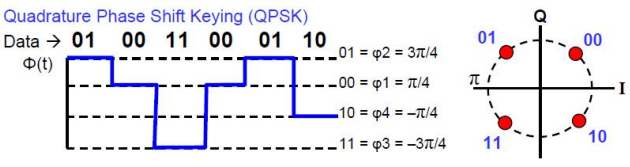

Fig. 10. QPSK example.

4) DPSK: Differential Phase Shift Keying (D-PSK) Modulation appends each carrier, or column of the matrix, with uniform random numbers defined by the symbol size. This reference data is then used by the receiver for demodulation of the carrier. This is illustrated in figure 11, resulting in a 400x31 matrix.

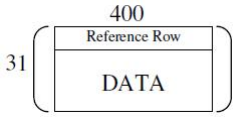

Fig. 11. Differential Phase Shift Keyed matrix.

5) IFFT: Prior to performing the IFFT, the matrix is expanded to contain the carriers (and their values stored as columns), conjugated carriers, and padded with zeros based on the size of the IFFT. The illustration in figure 12 applies an IFFT of size 1024.

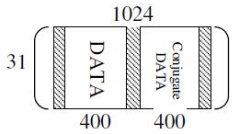

Fig. 12. Pre-IFFT matrix.

6) Periodic Time Guard: Just prior to transmission, the matrix is further expanded to include a Periodic Time Guard or cyclic prefix, to assist the receiver in synchronization and error recovery of the data stream. We illustrate this with roughly 25% of the symbol copied and added to the beginning of the symbol period, see figure 13, which concludes the modulation of the data stream.

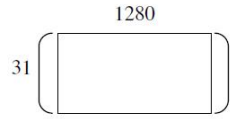

Fig. 13. Modulated matrix with cyclic prefix.

C. Channel

For simulation purposes, both AWGN and peak power clipping were implemented and are variable input values at the program runtime. This allows the user to simulate various environments with each of the modulation schemes modeled (BPSK, QPSK, 16-PSK, 256-PSK).

It is important to note for the simulation that the *original* modulated data is saved then cleared from memory to replicate the real system, wherein the receiver only has knowledge of the modulated data *influenced by the channel*. The saved data is later used for system diagnostics and error analysis.

D. Receiver

1) Frame detector: To detect the start of the signal frame, the frame detector employs a moving sum over the first frame of the sampled signal to detect the index bounds of the header and frame guard. Once detected, the frame guard is discarded from the search algorithm (not the date frame, which happens later), resulting in the starting index of the useful *i*-th data frame. Figure 14 illustrates this, notice too the IBI-free part discussed in sub-section III-G.

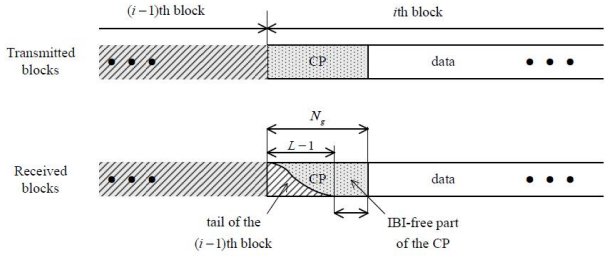

Fig. 14. CP of the *i*-th data frame.

2) Periodic Cyclic Prefix Removal: The reverse action is performed on the demodulated frames after converting a frame of discrete time signal from serial to parallel, such that the cyclic prefix is discarded as illustrated in figure 15

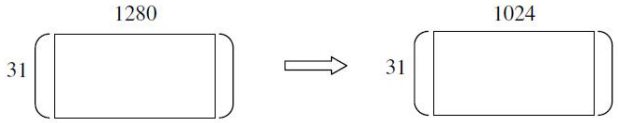

Fig. 15. Cyclic prefix removal.

3) FFT: Once the CP is removed, the time signal must be returned to the frequency domain, or spectral space, and the locations of the carriers are extracted to recover the complex valued matrix, as illustrated in figure 16.

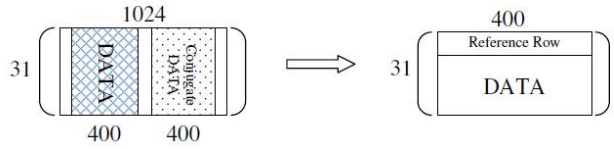

Fig. 16. Received data extracted from FFT bins.

4) DPSK Demodulation: Demodulation in the receiver is the symmetric operation to modulation in the transmitter, whereby the demodulator processes each frame sequentially after each carrier is extracted from the FFT bins.

First, the reference row is used to synchronize the receiver, then discarded.

Then each element of the complex matrix is returned to a phase between 0-360, and values translated to the symbol size. This is illustrated in figure 17. The data steam is then reshaped from parallel to serial to obtain the demodulated frame.

Finally, if the frame is padded with zeros, the zeros are detected and removed by calculating the remainder of the total number of data symbols over the number of carriers,

otherwise the frame demodulation is complete. For simulation purposes, the assembly of the transmission is performed at the end of the transmission, though it is possible to update the image by sequential rows, however this dramatically increases the program execution time.

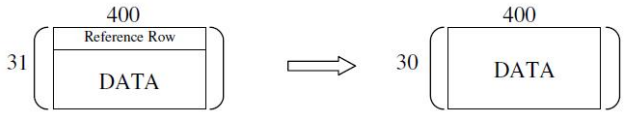

Fig. 17. DPSK Demodulation.

E. Error Calculations

Four error metrics are used for evaluation of the OFDM simulation.

1) Phase Error: Phase error is defined by the difference between the received phase and the translated (initial) phase for the corresponding symbol, before transmission.

2) Bit Error Rate (BER): BER is simply calculated by dividing the total number of errors by total number of demodulated symbols.

3) Data loss: In the event that one or more full rows of pixels may be missing at the output of the receiver, the post simulation diagnostics calculate the number of missing data samples and the total data transmitted, as well as the percentage of data loss.

4) Percent Error of Pixels in the Received Image: Percent Error of Pixels is a more meaningful metric for the end-user, which is a comparison of the received image and original image, pixel by pixel.

V. Experimental Results

This section is presented in two sub-sections: results of the simulation, and analysis of the results.

A. Simulation Results

Several figures were generated to understand the OFDM signals and overall performance of the different PSK modulations schemes. The following figures 19 - 22, illustrate comparative results of run time versus modulation scheme, BER vs. modulation scheme, BER vs. SNR, and Pixel Error vs. SNR. These plots are helpful in understanding the effect of SNR on each modulation and the resultant errors.

The remaining figures (located in the Appendix) illustrate the transmitted photo of the widly-used ”Lenna” image (the original is shown in figure 18), and the received transmission based on variations of modulation scheme, IFFT length, and SNR.

B. Anaysis of Results

Figure 19, confirms that increased number of symbols per frame increases the throughput of the system as more symbols are modulated at once, resulting in faster transmit and receive operations.

In contrast, the Bit Error Rate increases dramatically as the order of the modulation is increased, as illustrated in figure 20. Thus it can be viewed as a trade-off for decreasing the run-time.

Fig. 18. Original Lenna photo.

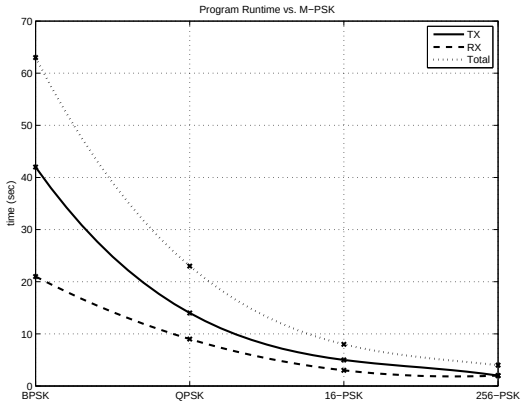

Fig. 19. Program Runtime vs. modulation scheme.

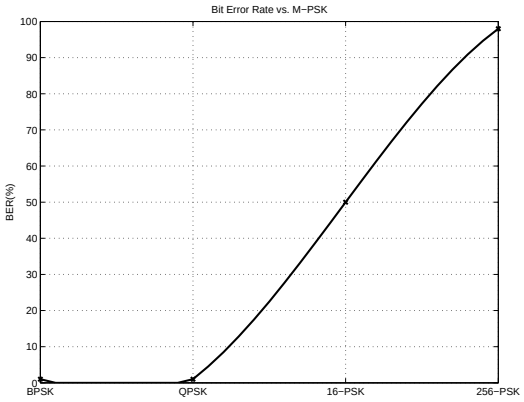

Fig. 20. BER vs. modulation scheme.

To illustrate that SNR is inversely proportional to error rates, figure 21 shows the relationship for all four methods used. Thus we can deduce that higher order PSK requires larger SNR to minimize BER.

Similarly, higher order PSK requires larger SNR to reduce the data (pixel) error, as illustrated in figure 22, and evident in figures 29 - 36.

The first six figures in the Appendix illustrate a single

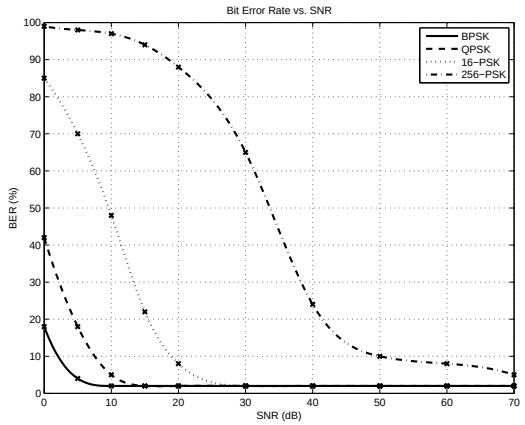

Fig. 21. BER vs. SNR.

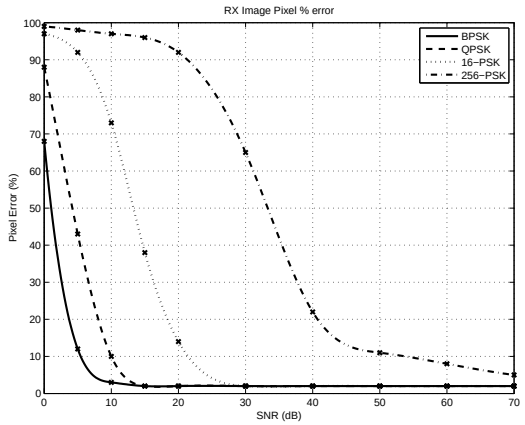

Fig. 22. Pixel Error vs. SNR.

symbol of QPSK modulation with its' respective IFFT bins, modulation, phase and magnitude. Because the first and last portion of the received modulated data have higher probability of errors due to incorrect synchronization, the middle of the frame is used instead.

Figure 23 shows unity magnitude of the modulated carriers, also identifying that the IFFT bins are nearly fully utilized by the carriers.

Figure 24 shows the constellation of four phase distributions of the QPSK symbols at the transmitter.

Figure 25 shows the time domain signals for each carrier.

Figure 26 correctly shows the channel influenced magnitude (no longer unity as in figure 23) of the modulated carriers at the receiver.

Figure 27 correctly shows the channel influenced scatter of the phase (no longer equidistant as in figure 23) of the modulated carriers at the receiver.

The polar plot in figure 28 shows the phase and magnitude grouping and distribution after the channel. Demonstrating the unlikelihood of ISI with QPSK modulation.

Based on the results contained in figure 28, the average phase error is 11.9 degrees, which is well below the phase tolerance of 45 degrees.

The remaining figures (29 - 36) in the Appendix illustrate the variation of PSK Modulation and SNR. In the case of the previous example using QPSK, results for which are illustrated in figures 23-28, as SNR is increased above 15dB, the BER converges to roughly 1.5% and pixel error to 2%. It is obvious from figures 31 - 32 that the resulting error is acceptable, and in fact nearly unnoticeable. This result is expected when the symbol size is not the same as the word size of the source data. This occurs when one or more symbols of the 4-bit word are decoded incorrectly, resulting in mistranslation of the 8-bit word to which they are mapped. Conversely, when SNR is relatively small yielding a BER and pixel error near 90%, the image is still observable because pixel errors in the gray-scale image are only off by a small number of gray levels. In fact, in this case nearly all of the pixels are off by a few gray levels, yet the image overall is intact. As a result we see the trade-off between BER and data rate, the requirements of which will vary upon application.

VI. Conclusions and Future Work

This research paper introduced and demonstrated the application of OFDM using MATLAB. While the use of MATLAB is certainly common for DSP analysis, this simulation tool is unique in that four different modulation schemes were evaluated for overall system performance, with the ability to vary the IFFT, number of sub-carriers, and SNR for each of the four selected modulation types. In the analysis of results, we demonstrated the trade-off between BER and data-rate using an 8-bit gray scale image. Further work may be performed on the current MATLAB simulation, and apply QAM (Quadrature Amplitude Modulation) which is common in telecommunications, or CDMA (Code Division Multiple Access) which is used in Global Positioning Systems (GPS) and mobile phone handsets.

VII. Acknowledgments

This research builds on the technical conversations with Prof. Yingbo Hua. These technical collaborations are greatly appreciated.

References

[1] B. Saltzberg, "Performance of an Efficient Parallel Data Transmission System," *IEEE Trans. on Communications*, vol. 15, no. 6, 1967.

[2] Z. Wang and G.B., "Wireless Multicarrier Communications: Where Fourier Meets Shannon," *IEEE Signal Processing Magazine*, vol. 17, pp. 29–48, 2000.

[3] S. B. Weinstein and P. M. Ebert, "Data Transmission by Frequency-Division Multiplexing using the Discrete Fourier Transform," *IEEE Trans. on Communications*, vol. 19, no. 5, 1971.

[4] R. van Nee et al., "New high-rate wireless LAN standards," *IEEE Communications Magazine*, vol. 40, no. 5, pp. 140–147, 2002.

[5] R. van Nee and R. Prasad, *OFDM for Wireless Multilmedia Communications*. Artech House Publishers, 2000.

[6] R. van Nee, "A new OFDM standard for high rate wireless LAN in the 5 GHz band," in *Proc. of the IEEE Vehicular Technology Conference*, 1999, p. 258262.

[7] J. P. E. Biglieri and S. Shamai(Shitz), "Fading channels: Information-Theoretic and Communications Aspects," *IEEE Trans. on Information Theory*, vol. 44, no. 6, p. 26192692, 1998.

[8] J. G. Proakis, *Digital Communications*. Mc Graw Hill, 1995.

[9] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Prentice-Hall, 2010.

[10] G. H. Golub and C. F. V. Loan, *Matrix Computations*. John Hopkins University Press, 1996.

[11] A. D. M. Bossert and V. Zyablov, "Improved channel estimation with decision feedback for OFDM systems," *Electronics Letters*, vol. 34, no. 11, p. 10641065, 1998.

[12] U. Tureli and H. Liu, "Blind carrier synchronization and channel identification for OFDM communications," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 6, May 1998, pp. 3509–3512.

[13] E. Jaffrot and M. Siala, "Turbo channel estimation for OFDM systems on highly time and frequency selective channels," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, June 2000.

[14] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo codes," *IEEE Trans. on Communications*, vol. 44, no. 10, p. 12611271, 1996.

[15] A. Viterbi, "Error Bounds for Convolutional Coders and Asymptoticaly Optimum Decoding Algorithm," *IEEE Trans. on Information*, vol. 13, pp. 55–67, 1967.

[16] J. Boutros, "Lattice codes for Rayleigh Fading Channels," Ph.D. thesis, cole Nationale Suprieure desTlcommunications, Tech. Rep., 1996.

[17] ——, "Les Turbo Codes Paralleles et Series," *Polycopi du cours de Telecom Paris (translated to English)*, 1998.

[18] E. O. J. Hagenauer and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. on Information Theory*, vol. 42, no. 2, p. 429445, 1996.

[19] G. Ungerboeck, "Treillis Coded Modulation with Redundant Signals Sets, Part. I: Introduction. Part. II: State of the Art." *IEEE Trans. on Communications*, vol. 25, pp. 5–21, 1987.

[20] E. Zehavi, "8-PSK Trellis Codes for a Rayleigh Channel," *IEEE Trans. on Communications*, vol. 40, pp. 873–884, 1992.

[21] S. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, p. 14511458, 1998.

APPENDIX

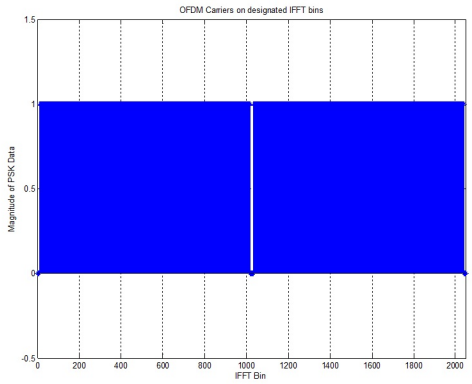

Fig. 23. QPSK magnitude vs. IFFT Bins.

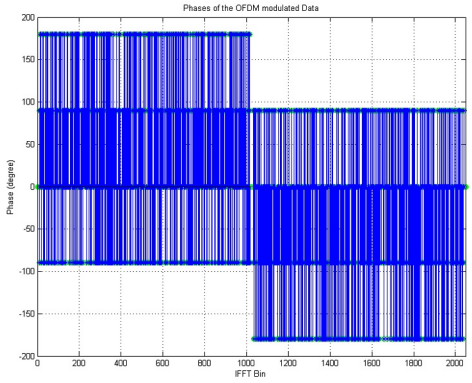

Fig. 24. QPSK phase vs. IFFT Bins.

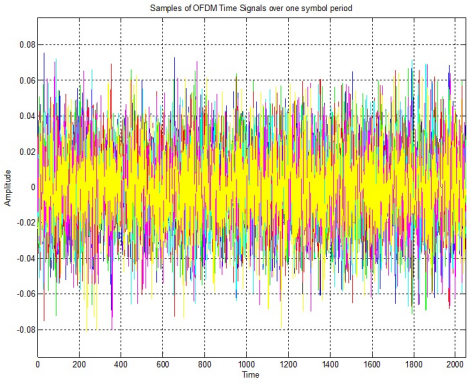

Fig. 25. QPSK time signals for one symbol period.

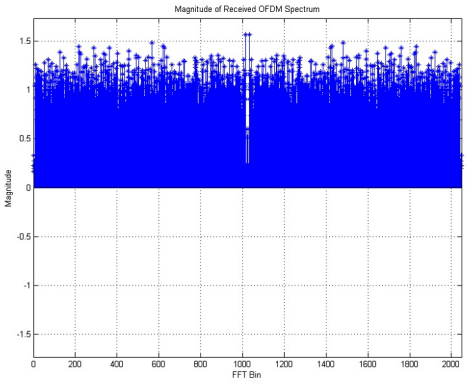

Fig. 26. QPSK received spectrum magnitude.

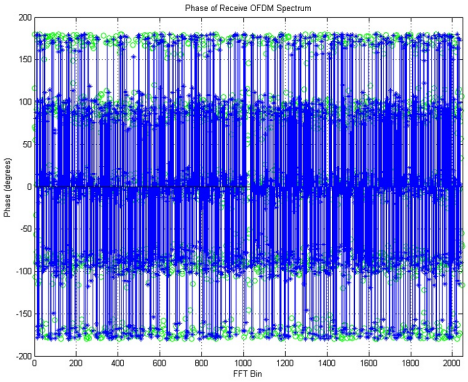

Fig. 27. QPSK received spectrum phase.

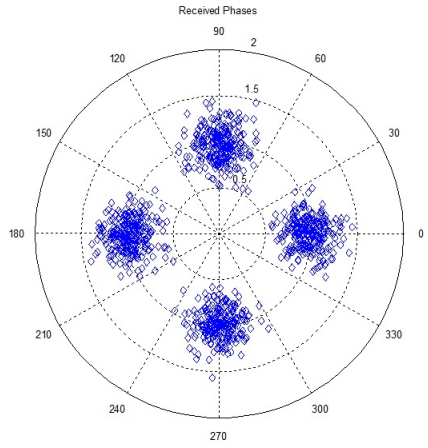

Fig. 28. QPSK received phase and magnitude distributions.

(a) (b)

Fig. 29. BPSK: SNR=0dB (a), BPSK: SNR=5dB (b).

(a) (b)

Fig. 30. BPSK: SNR=10dB (a), BPSK: SNR=15dB (b).

(a) (b)

Fig. 31. QPSK: SNR=0dB (a), QPSK: SNR=5dB (b).

(a) (b)

Fig. 32. QPSK: SNR=10dB (a), QPSK: SNR=15dB (b).

(a) (b)

Fig. 33. 16-PSK: SNR=0dB (a), 16-PSK: SNR=5dB (b).

(a) (b)

Fig. 34. 16-PSK: SNR=15dB (a), 16-PSK: SNR=30dB (b).

(a) (b)

Fig. 35. 256-PSK: SNR=0dB (a), 256-PSK: SNR=10dB (b).

(a) (b)

Fig. 36. 256-PSK: SNR=30dB (a), 256-PSK: SNR=70dB (b).