Chapter 4: Orthogonal Frequency Division Multiplexing (OFDM)

Jeffrey G. Andrews

March 24, 2006

Orthogonal Frequency Division Multiplexing (OFDM) is a multicarrier modulation technique that has recently found wide adoption in a wide variety of high data rate communication systems, including Digital Subscriber Lines (DSL), Wireless LANs (802.11a/g/n), Digital Video Broadcasting, and now, WiMAX and other emerging wireless broadband systems like Flarion's proprietary Flash-OFDM and 4G/"Super 3G" cellular systems. OFDM's popularity for high-data rate applications stems primarily from its efficient and flexible management of intersymbol interference (ISI) in highly dispersive channels.

As emphasized in chapter 3, as the channel delay spread τ becomes an increasingly large multiple of the symbol time T_s, the ISI becomes very severe. By definition, a high data rate system will generally have $\tau \gg T_s$ since the number of symbols sent per second is high. In a non-line of sight (NLOS) system such as WiMAX that must transmit over moderate to long distances, the delay spread will also frequently be large. In short, wireless broadband systems of all types will suffer from severe ISI, and hence will require transmitter and/or receiver techniques that overcome the ISI. Although the 802.16 standards include single-carrier modulation techniques, the vast majority (if not all) 802.16-compliant systems will use the OFDM modes, which have also been selected as the preferred modes by the WiMAX forum.

This chapter has three main goals to help the WiMAX engineer or student understand how to use OFDM in a wireless broadband system:

1. Explain the elegance of multicarrier modulation, and how it works in theory.

2. Emphasize a practical understanding of OFDM system design, covering key concepts such as the cyclic prefix, frequency equalization, synchronization, and channel estimation.

3. Discuss implementation issues for WiMAX systems such as the peak-to-average ratio, and provide illustrative examples related to WiMAX.

4.1 The Multicarrier Concept

The basic idea of multicarrier modulation is quite simple, and follows naturally from the competing desires for high data rates and intersymbol interference (ISI) free channels. In order to have a channel that does not have ISI, the symbol time T_s has to be larger – often significantly larger – than the channel delay spread τ. Digital communication systems simply cannot function if ISI is present – an error floor quickly develops and as T_s approaches or falls below τ, the bit error rate becomes intolerable. As we have noted previously, for wideband channels that provide the high data rates needed by today's applications, the desired symbol time is usually much smaller than the delay spread, so intersymbol interference is severe.

In order to overcome this, multicarrier modulation divides the high-rate transmit bitstream into L lower-rate substreams, *each* of which has $T_s/L >> \tau$, and is hence ISI-free. These individual substreams can then be sent over L parallel subchannels, maintaining the total desired data rate. Typically the subchannels are orthogonal under ideal propagation conditions, in which case multicarrier modulation is often referred to as orthogonal frequency division multiplexing (OFDM). The data rate on each of the subchannels is much less than the total data rate, and so the corresponding subchannel bandwidth is much less than the total system bandwidth. The number of substreams is chosen to insure that each subchannel has a bandwidth less than

the coherence bandwidth of the channel, so the subchannels experience relatively flat fading. Thus, the ISI on each subchannel is small. Moreover, in the digital implementation of OFDM, the ISI can be completely eliminated through the use of a cyclic prefix.

Example 4.1 *A certain wideband wireless channel has a delay spread of* 1μ*sec. We assume that in order to overcome ISI, that* $T_s \geq 10\tau$.

1. *What is the maximum bandwidth allowable in this system?*
2. *If multicarrier modulation is used, and we desire a 5 MHz bandwidth, what is the required number of subcarriers?*

For part (1), if it is assumed that $T_s = 10\tau$ in order to satisfy the ISI-free condition, the maximum bandwidth would be $1/T_s = .1/\tau = 100$ KHz, far below the intended bandwidths for WiMAX systems.

In part (2), if multicarrier modulation is employed, the symbol time goes to $T = LT_s$. The delay spread criterion mandates that the new symbol time is still bounded to 10% of the delay spread, i.e. $(LT_s)^{-1} = 100KHz$. But the 5 MHz bandwidth requirement gives $(T_s)^{-1} = 5$ MHz. Hence, $L \geq 50$ allows the full 5 MHz bandwidth to be used with negligible ISI.

4.1.1 An elegant approach to intersymbol interference

Multicarrier modulation in its simplest form divides the wideband incoming data stream into L narrowband substreams, each of which is then transmitted over a different orthogonal frequency subchannel. As in example 4.1, the number of substreams L is chosen to make the symbol time on each substream much greater than the delay spread of the channel or, equivalently, to make the substream bandwidth less then the channel coherence bandwidth. This insures that the substreams will not experience significant ISI.

A simple illustration of a multicarrier transmitter and receiver is given in Figures 4.1 and 4.2. Essentially, a high-rate data signal of rate R bps and with a passband bandwidth B is broken into L parallel substreams each with rate R/L and passband bandwidth B/L. After passing through the channel $H(f)$, the received signal would appear as shown in 4.3, where we have assumed for simplicity that the pulse-shaping allows a perfect spectral shaping so that there is no subcarrier overlap[1]. As long as the number of subcarriers is sufficiently large to allow for the subcarrier bandwidth to be much less than the coherence bandwidth, i.e. $B/L \ll B_c$, then it can be insured that each subcarrier experiences approximately flat fading. The mutually orthogonal signals can then be individually detected, as shown in Figure 4.2.

[1] In practice, there would be some rolloff factor of β, so the actual consumed bandwidth of such a system would be $(1+\beta)B$. However, OFDM as we will see avoids this inefficiency by using a cyclic prefix.

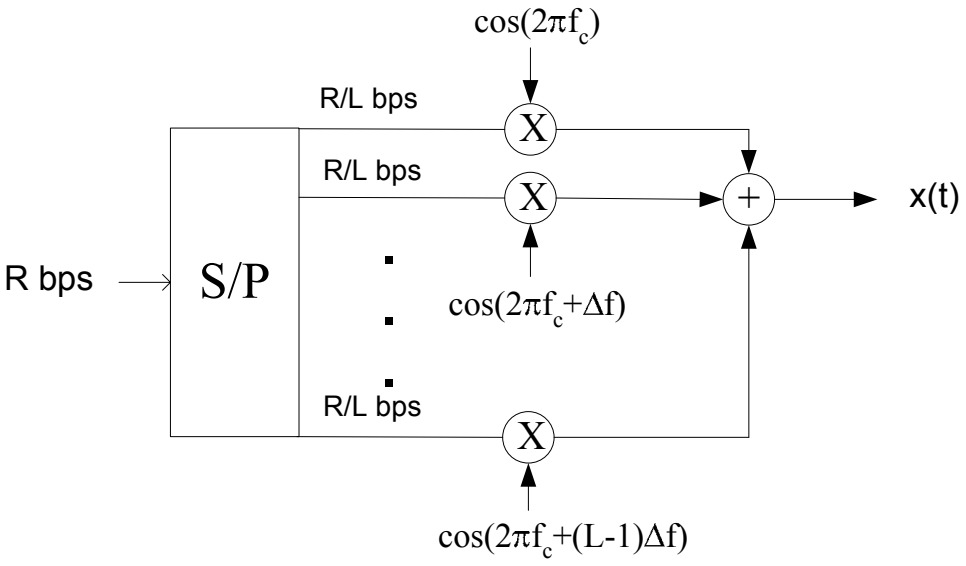


Figure 4.1: A Basic multicarrier transmitter: a high rate stream of R bps is broken into L parallel streams each with rate R/L and then multiplied by a different carrier frequency.

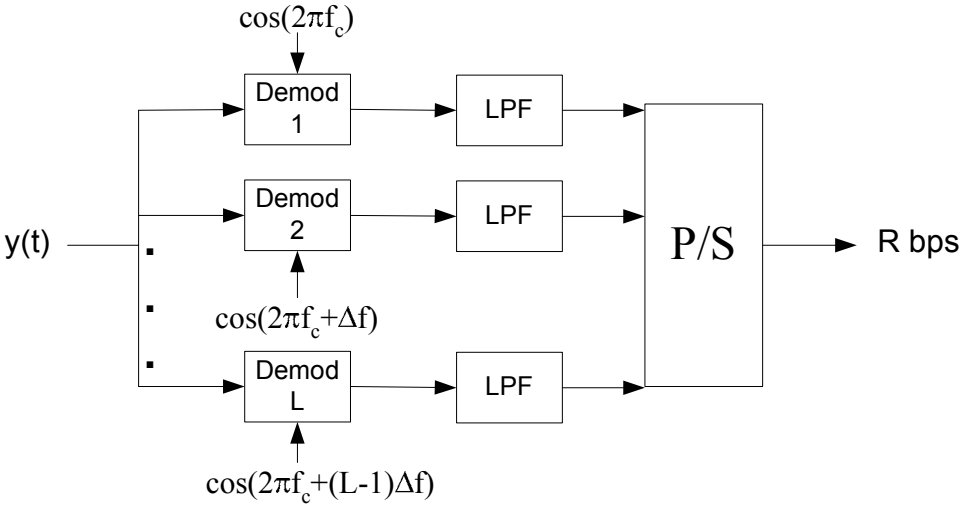


Figure 4.2: A Basic multicarrier receiver: each subcarrier is decoded separately, requiring L independent receivers.

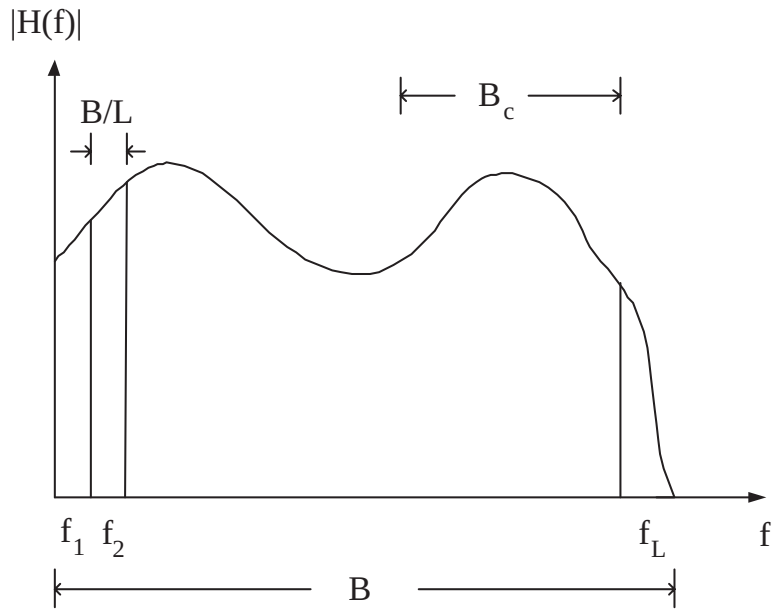


Figure 4.3: The transmitted multicarrier signal experiences approximately flat fading on each subchannel since $B/L \ll B_c$, even though the overall channel experiences frequency selective fading, i.e. $B > B_c$.

Hence, the multicarrier technique has an interesting interpretation in both the time and frequency domains. In the time domain, the symbol duration on each subcarrier has increased to $T = LT_s$, so by letting L grow larger, it can be assured that the symbol duration exceeds the channel delay spread, i.e. $T \gg \tau$, which is a requirement for ISI-free communication. In the frequency domain, the subcarriers have bandwidth $B/L \ll B_c$, which assures "flat fading", the frequency domain equivalent to ISI-free communication.

Although this simple type of multicarrier modulation is easy to understand, it has several crucial shortcomings. First, as noted above, in a realistic implementation, a large bandwidth penalty will be inflicted since the subcarriers can't have perfectly rectangular pulse shapes and still be time-limited. Additionally, very high quality (and hence, expensive) low pass filters will be required to maintain the orthogonality of the the subcarriers at the receiver. Most importantly, this scheme requires L independent RF units and demodulation paths. In Section 4.2, we show how OFDM overcomes these shortcomings.

4.2 OFDM Basics

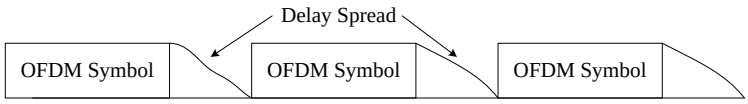
In order to overcome the daunting requirement for L RF radios in both the transmitter and receiver, OFDM employs an efficient computational technique known as the Discrete Fourier Transform (DFT), more commonly known as the Fast Fourier Transform (FFT)[2]. In this section, we will learn how the FFT (and its inverse, the IFFT) are able to create a multitude of orthogonal subcarriers using just a single radio.

[2]The FFT is a highly efficient implementation of the DFT.

4.2.1 Block transmission with guard intervals.

We begin by grouping L data symbols into a block known as an *OFDM symbol*. An OFDM symbol lasts for a duration of T seconds, where $T = LT_s$. In order to keep each OFDM symbol independent of the others after going through a wireless channel, it is necessary to introduce a guard time in between each OFDM symbol, as shown here:

This way, after receiving a series of OFDM symbols, as long as the guard time T_g is larger than the delay spread of the channel τ, each OFDM symbol will only interfere with itself.



Put simply, OFDM transmissions allow ISI *within* an OFDM symbol, but by including a sufficiently large guard band, it is possible to guarantee that there is no interference *between* subsequent OFDM symbols.

4.2.2 Circular Convolution and the DFT.

Now that subsequent OFDM symbols have been rendered orthogonal with a guard interval, the next task is to attempt to remove the ISI *within* each OFDM symbol. As described in Chapter 3, when an input data stream $x[n]$ is sent through a linear time-invariant FIR channel $h[n]$, the output is the linear convolution of the input and the channel, i.e. $y[n] = x[n] * h[n]$. However, let's imagine for a moment that it was possible to compute $y[n]$ in terms of a *circular* convolution, i.e.

$$y[n] = x[n] \circledast h[n] = h[n] \circledast x[n], \tag{4.1}$$

where

$$x[n] \circledast h[n] = h[n] \circledast x[n] \triangleq \sum_{k=0}^{L-1} h[k]x[n-k]_L \tag{4.2}$$

and the circular function $x[n]_L = x[n \bmod L]$ is a *periodic* version of $x[n]$ with period L. In other words, each value of $y[n] = h[n] \circledast x[n]$ is the sum of the product of L terms[3].

In this case of circular convolution, it would then be possible to take the DFT of the channel output $y[n]$ to get:

$$\mathrm{DFT}\{y[n]\} = \mathrm{DFT}\{h[n] \circledast x[n]\} \tag{4.3}$$

which yields in the frequency domain

$$Y[m] = H[m]X[m]. \tag{4.4}$$

[3]For a more thorough tutorial on circular convolution, the reader is referred to [34] or the "Connexions" web resource `http://cnx.rice.edu/`

Note that the duality between circular convolution in the time domain and simple multiplication in the frequency domain is a property unique to the DFT. The L point DFT is defined as

$$\mathrm{DFT}\{x[n]\} = X[m] \triangleq \frac{1}{\sqrt{L}} \sum_{n=0}^{L-1} x[n] e^{-j\frac{2\pi nm}{L}}, \tag{4.5}$$

while its inverse, the IDFT is defined as

$$\mathrm{IDFT}\{X[m]\} = x[n] \triangleq \frac{1}{\sqrt{L}} \sum_{m=0}^{L-1} X[m] e^{j\frac{2\pi nm}{L}}. \tag{4.6}$$

Referring to (4.4), this innocent formula actually describes an ISI-free channel in the frequency domain, where each input symbol $X[m]$ is simply scaled by a complex-value $H[m]$. So, given knowledge of the channel frequency response $H[m]$ at the receiver, it is trivial to recover the input symbol by simply computing

$$\hat{X}[m] = \frac{Y[m]}{H[m]}, \tag{4.7}$$

where the estimate $\hat{X}[m]$ will generally be imperfect due to additive noise, co-channel interference, imperfect channel estimation, and other imperfections that will be discussed later. Nevertheless, in principle, the ISI – which is the most serious form of interference in a wideband channel – has been mitigated.

A natural question to ask at this point is: "where does this circular convolution come from?" After all, nature provides a linear convolution when a signal is transmitted through a linear channel. The answer is that this circular convolution can be "faked" by adding a specific prefix called the *cyclic prefix* onto the transmitted vector.

4.2.3 The Cyclic Prefix.

The key to making OFDM realizable in practice is the utilization of the FFT algorithm, which has low complexity. In order for the IFFT/FFT to create an ISI-free channel, the channel must appear to provide a circular convolution, as seen in (4.4). If a cyclic prefix is added to the transmitted signal, as shown in Figure 4.4, then this creates a signal that appears to be $x[n]_L$, and so $y[n] = x[n] \circledast h[n]$.

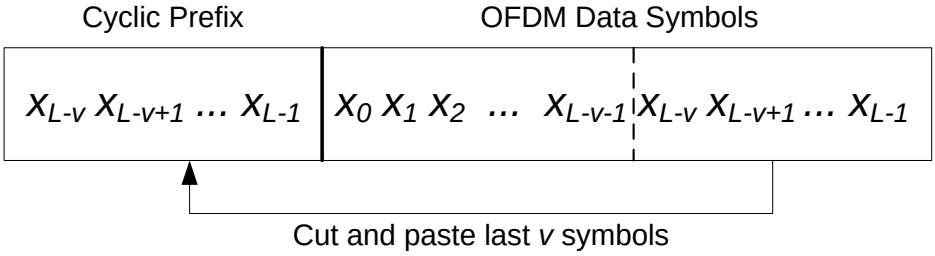


Figure 4.4: The OFDM cyclic prefix

Let's see how this works. If the maximum channel delay spread has a duration of $v + 1$ samples, then by adding a guard band of at least v samples between OFDM symbols, each OFDM symbol is made

independent of those coming before and after it, and so just a single OFDM symbol can be considered. Representing such an OFDM symbol in the time domain as a length L vector gives

$$\mathbf{x} = [x_1 \; x_2 \; \ldots x_L]. \tag{4.8}$$

After applying a cyclic prefix of length v, the actual transmitted signal is

$$\mathbf{x}_{cp} = [\underbrace{x_{L-v} \; x_{L-v+1} \; \ldots \; x_{L-1}}_{\text{Cyclic Prefix}} \underbrace{x_0 \; x_1 \; \ldots \; x_{L-1}}_{\text{Original data}}]. \tag{4.9}$$

The output of the channel is by definition $\mathbf{y}_{cp} = \mathbf{h} * \mathbf{x}_{cp}$, where $\mathbf{h}$ is a length $v+1$ vector describing the impulse response of the channel during the OFDM symbol[4]. The output $\mathbf{y}_{cp}$ has $(L+v)+(v+1)-1 = L+2v$ samples. The first v samples of $\mathbf{y}_{cp}$ contain interference from the preceding OFDM symbol, and so are discarded. The last v samples disperse into the subsequent OFDM symbol, so also are discarded. This leaves exactly L samples for the desired output $\mathbf{y}$, which is precisely what is required to recover the L data symbols embedded in $\mathbf{x}$.

Our claim is that these L samples of $\mathbf{y}$ will be equivalent to $\mathbf{y} = \mathbf{h} \circledast \mathbf{x}$. Various proofs are possible, the most intuitive is a simple inductive argument. Consider for the moment just y_0, i.e. the first element in $\mathbf{y}$. As shown in Figure 4.5, due to the cyclic prefix, y_0 depends on x_0 and the circularly wrapped values $x_{L-v} \ldots x_{L-1}$. That is:

$$\begin{aligned} y_0 &= h_0 x_0 + h_1 x_{L-1} + \ldots + h_v x_{L-v} \\ y_1 &= h_0 x_1 + h_1 x_0 + \ldots + h_v x_{L-v+1} \\ &\vdots \\ y_{L-1} &= h_0 x_{L-1} + h_1 x_{L-2} + \ldots + h_v x_{L-v-1} \end{aligned} \tag{4.10}$$

From inspecting (4.2), it can be seen that this is exactly the value of $y_0, y_1, \ldots, y_{L-1}$ resulting from $\mathbf{y} = \mathbf{x} \circledast \mathbf{h}$. Thus, by mimicking a circular convolution, a cyclic prefix that is at least as long as the channel duration allows the channel output $\mathbf{y}$ to be decomposed into a simple multiplication of the channel frequency response $\mathbf{H} = \mathrm{DFT}\{\mathbf{h}\}$ and the channel frequency domain input, $\mathbf{X} = \mathrm{DFT}\{\mathbf{x}\}$.

The cyclic prefix, although elegant and simple, is not entirely free. It comes with both a bandwidth and power penalty. Since v redundant symbols are sent, the required bandwith for OFDM increases from B to $\frac{L+v}{L}B$. Similarly, an additional v symbols must be counted against the transmit power budget. Hence, the cyclic prefix carries a power penalty of $10\log_{10}\frac{L+v}{L}$ dB in addition to the bandwidth penalty. In summary, the use of cyclic prefix entails data rate and power losses that are both

$$\text{Rate Loss} = \text{Power Loss} = \frac{L}{L+v}.$$

The "wasted" power has increased importance in an interference-limited wireless system, since it causes interference to neighboring users. One way to reduce the transmit power penalty is noted in Sidebar 4.1.

[4]It can generally be reasonably assumed that the channel remains constant over an OFDM symbol, since the OFDM symbol time T is usually much less that the channel coherence time, T_c.

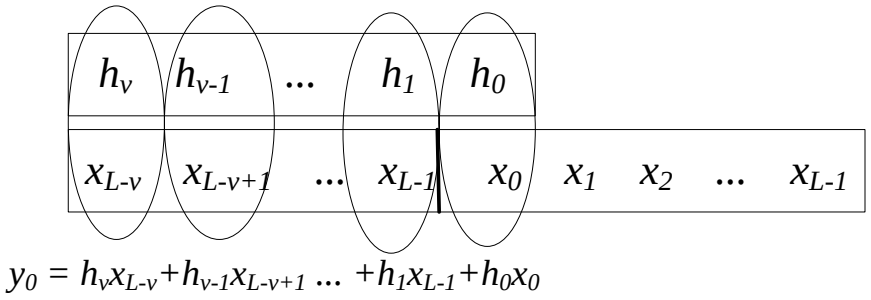


Figure 4.5: The OFDM cyclic prefix creates a circular convolution

It can be noted that for $L \gg v$, the inefficiency due to the cyclic prefix can be made arbitrarily small by increasing the number of subcarriers. However, as the latter parts of this chapter will explain, there are numerous other important sacrifices that must be made as L grows large. As with most system design problems, desirable properties such as efficiency must be traded off against cost and required tolerances.

Example 4.2 *Referring to Table 9.3, find the minimum and maximum data rate loss due to the cyclic prefix in WiMAX for a 10 MHz channel with a maximum delay spread $\tau = 5\mu$sec.*

At a symbol rate of 10 MHz, a delay spread of 5μsec affects 50 symbols, so we require a CP length of at least $v = 50$.

The minimum overhead will be for the largest number of subcarriers, referring to Table 9.3 this yields $L = 2048$. In this case, $v/L = 50/2048 = 1/40.96$ so the minimum guard band of $1/32$ will suffice. Hence, the data rate loss is only 1/32 in this case.

The maximum overhead occurs when the number of subcarriers is small. If $L = 128$, then $v/L = 50/128$, so even an overhead of $1/4$ won't be sufficient to preserve subcarrier orthogonality. More subcarriers are required. For $L = 256$, $v/L < 1/4$, so in this case ISI-free operation is possible, but at a data rate loss of 1/4.

Sidebar 4.1 *An Alternative Prefix*

One alternative to the cyclic prefix is to instead use a "zero prefix", which as in vector coding constitutes a null guard band. One commercial system that proposes this is the Intel/Texas Instruments "Multiband OFDM" system that has been proposed for ultrawideband (UWB) standardization under the context of the IEEE 802.15.3 subcommittee. As shown in Fig. 4.6 the Multiband OFDM transmitter just sends a prefix of null data so that there is no transmitter power penalty. At the receiver, the "tail" can be added back in, which recreates the effect of a cyclic prefix, so the rest of the OFDM system can function as usual.

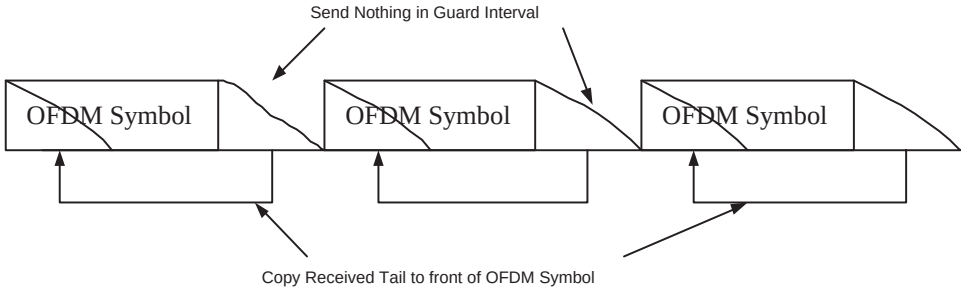


Figure 4.6: The OFDM Zero Prefix allows the circular channel to be recreated at the receiver.

Why wouldn't every OFDM system use a zero prefix then, since it reduces the transmit power by $10\log_{10}\left(\frac{L+v}{L}\right)$ dB? There are two reasons. First, the zero prefix generally increases the receiver power by $10\log_{10}\left(\frac{L+v}{L}\right)$ dB, since the tail now needs to be received (whereas with a cyclic prefix, it can be ignored). Second, additional noise from the received tail symbols is added back into the signal, causing a higher noise power $\sigma^2 \rightarrow \frac{L+v}{L}\sigma^2$. The designer must weigh these tradeoffs to determine whether a zero or cyclic prefix is preferable. WiMAX systems employ a cyclic prefix.

4.2.4 Frequency Equalization

In order for the received symbols to be estimated, the complex channel gains for each subcarrier must be known, which corresponds to knowing the amplitude and phase of the subcarrier. For simple modulation techniques like QPSK that don't use the amplitude to transmit information, just the phase information is sufficient.

After the FFT is performed, the data symbols are estimated using a one-tap frequency domain equalizer, or FEQ, as

$$\hat{X}_l = \frac{Y_l}{H_l} \tag{4.11}$$

where H_l is the *complex* response of the channel at the frequency $f_c + (l-1)\Delta f$, and therefore it both corrects the phase and equalizes the amplitude before the decision device. Note that although the FEQ inverts the channel, there is no problematic noise enhancement or coloring since both the signal and the noise will have their powers directly scaled by $|\frac{1}{H_l}|^2$.

4.2.5 An OFDM Block Diagram

Let us now briefly review the key steps in an OFDM communication system, each of which can be observed in Figure 4.7.

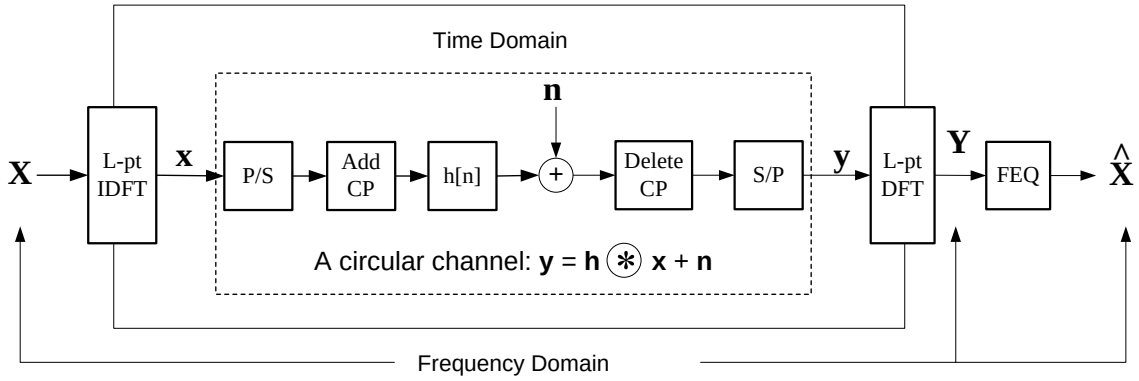


Figure 4.7: An OFDM system in vector notation. In OFDM, the encoding and decoding is done in the frequency domain, where $\mathbf{X}$, $\mathbf{Y}$, and $\hat{\mathbf{X}}$ contain the L transmitted, received, and estimated data symbols.

1. The first step in OFDM is to break a wideband signal of bandwidth B into L narrowband signals (subcarriers) each of bandwidth B/L. This way, the aggregate symbol rate is maintained, but each subcarrier experiences flat-fading, or ISI-free communication. The L subcarriers for a given OFDM symbol are represented by a vector $\mathbf{X}$, which contains the L current symbols.

2. In order to use a single wideband radio instead of L independent narrow band radios, the subcarriers are modulated using an IFFT operation.

3. In order for the IFFT/FFT to decompose the ISI channel into orthogonal subcarriers, a cyclic prefix of length v must be appended after the IFFT operation. The resulting $L + v$ symbols are then sent in serial through the wideband channel.

4. At the receiver, the cyclic prefix is discarded, and the L received symbols are demodulated using an FFT operation, which results in L data symbols, each of the form $Y_l = H_l X_l + N_l$ for subcarrier l.

5. Each subcarrier can then be equalized via an FEQ by simply dividing by the complex channel gain $H[i]$ for that subcarrier. This results in $\hat{\mathbf{X}}_l = X_l + N_l/H_l$.

We have neglected a number of important practical issues thus far. For example, we have assumed that the transmitter and receiver are perfectly synchronized, and that the receiver perfectly knows the channel (in order to perform the FEQ). In the next section, we overview the implementation issues for OFDM in WiMAX.

4.3 An Example: OFDM in WiMAX

To gain an appreciation for the time and frequency domain interpretations of OFDM, WiMAX systems can be used as an example. Although simple in concept, the subtleties of OFDM can be confusing if each signal processing step is not understood. To ground the discussion, we will consider a passband OFDM system, and then give specific values for the important system parameters.

Figure 4.8 shows an up close view of a passband OFDM modulation engine. The inputs to this figure are L independent QAM symbols (the vector $\mathbf{X}$), and these L symbols are treated as separate subcarriers. These L data-bearing symbols can be created from a bit stream by a symbol mapper and serial-to-parallel convertor (S/P). The L-point IFFT then creates a time domain L-vector $\mathbf{x}$ that is cyclic extended to have length $L(1+G)$, where G is the fractional overhead. This longer vector is then parallel-to-serial (P/S) converted into a wideband digital signal that can be amplitude modulated with a single radio at a carrier frequency of $f_c = \omega_c/2\pi$.

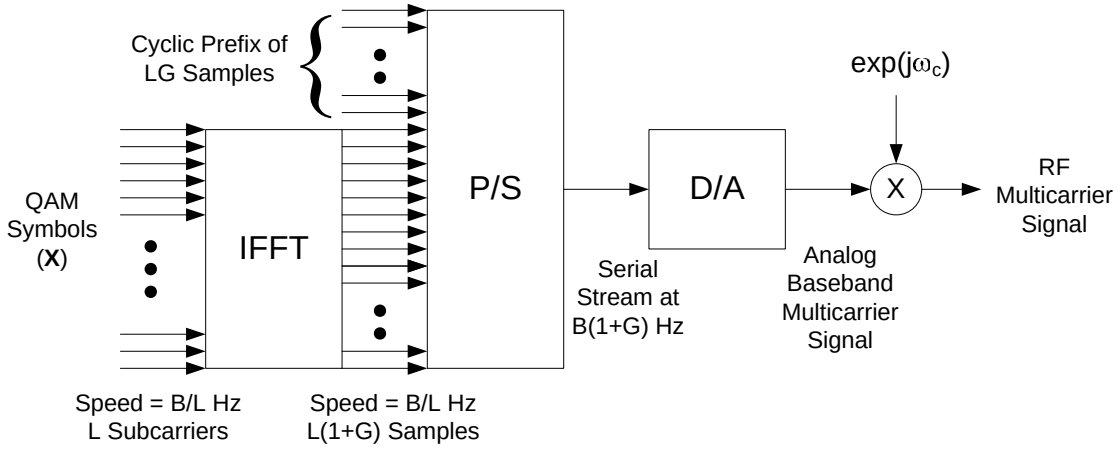


Figure 4.8: A closeup of the OFDM baseband transmitter.

This procedure appears to be relatively straightforward, but in order to be a bit less abstract, we will now use some plausible values for the different parameters. Chapter 9 enumerates all the legal values for the OFDM parameters B, L, L_g, and G. The key OFDM parameters are summarized in Table 4.1, along with some potential numerical values for these parameters. As an example, if 16QAM modulation was used ($M = 16$), the raw (neglecting coding) data rate of this WiMAX system would be:

$$R = \frac{B}{L}\frac{L_d \log_2(M)}{1+G} \tag{4.12}$$

$$= \frac{10^7 \text{MHz}}{1024}\frac{768 \log_2(16)}{1.125} = 24\text{Mbps}. \tag{4.13}$$

In words, there are L_d data-carrying subcarriers of bandwidth B/L, each carrying $\log_2(M)$ bits of data. An additional overhead penalty of $(1+G)$ must be paid for the cyclic prefix, since it consists of redundant information and sacrifices the transmission of actual data symbols.

Table 4.1: Summary of OFDM Parameters. WiMAX specified parameters are denoted by an asterisk ($*$): the other OFDM parameters can all be derived from these values.

Symbol	Description	Relation	Example WiMAX value
$B*$	Nominal bandwidth	$B = 1/T_s$	10 MHz
$L*$	No. of subcarriers	size of IFFT/FFT	1024
$G*$	Guard fraction	% of L for CP	1/8
L_g*	Data subcarriers	$L-$ pilot/null subcarriers	768
T_s	Sample time	$T_s = 1/B$	1 μsec
N_g	Guard symbols	$N_g = GL$	128
T_g	Guard time	$T_g = T_s N_g$	12.8 μsec
T	OFDM symbol time	$T = T_s(L + N_g)$	115.2 μsec
B_{sc}	Subcarrier Bandwidth	$B_{sc} = B/L$	9.76 KHz

4.4 Timing and Frequency Synchronization

In order to demodulate an OFDM signal, there are two important synchronization tasks that need to be performed by the receiver. First, the timing offset of the symbol and the optimal timing instants need to be determined. This is referred to as *timing synchronization*. Second, the receiver must align its carrier frequency as closely as possible with the transmitted carrier frequency. This is referred to as *frequency synchronization*. Compared to single-carrier systems, the timing synchronization requirements for OFDM are in fact somewhat relaxed, since the OFDM symbol structure naturally accommodates a reasonable degree of synchronization error. On the other hand, frequency synchronization requirements are significantly more stringent, since the orthogonality of the data symbols is reliant on their being individually discernible in the frequency domain.

Figure 4.9 shows a representation of an OFDM symbol in time (left) and frequency (right). In the time domain, the IFFT effectively modulates each data symbol onto a unique carrier frequency. In Figure 4.9 only two of the carriers are shown – the actual transmitted signal is the superposition of all the individual carriers. Since the time window is $T = 1\mu$sec and a rectangular window is used, the frequency response of each subcarrier becomes a "sinc" function with zero crossings every $1/T = 1$ MHz. This can be confirmed using the Fourier Transform $\mathcal{F}\{\cdot\}$ since

$$\begin{aligned} \mathcal{F}\{\cos(2\pi f_c) \cdot \mathrm{rect}(t/T)\} &= \mathcal{F}\{\cos(2\pi f_c)\} * \mathcal{F}\{\mathrm{rect}(2t/T)\} && (4.14) \\ &= \mathrm{sinc}\left(T(f - f_c)\right), && (4.15) \end{aligned}$$

where $\mathrm{rect}(x) = 1,\ x \in (-0.5, 0.5)$, and zero elsewhere. This frequency response is shown for $L = 8$ subcarriers in the right of Figure 4.9.

The challenge of timing and frequency synchronization can be appreciated by inspecting these two figures. If the timing window is slid to the left or right, a unique phase change will be introduced to each of the subcarriers. In the frequency domain, if the carrier frequency synchronization is perfect, the receiver samples at the peak of each subcarrier, where the desired subcarrier amplitude is maximized and the inter-

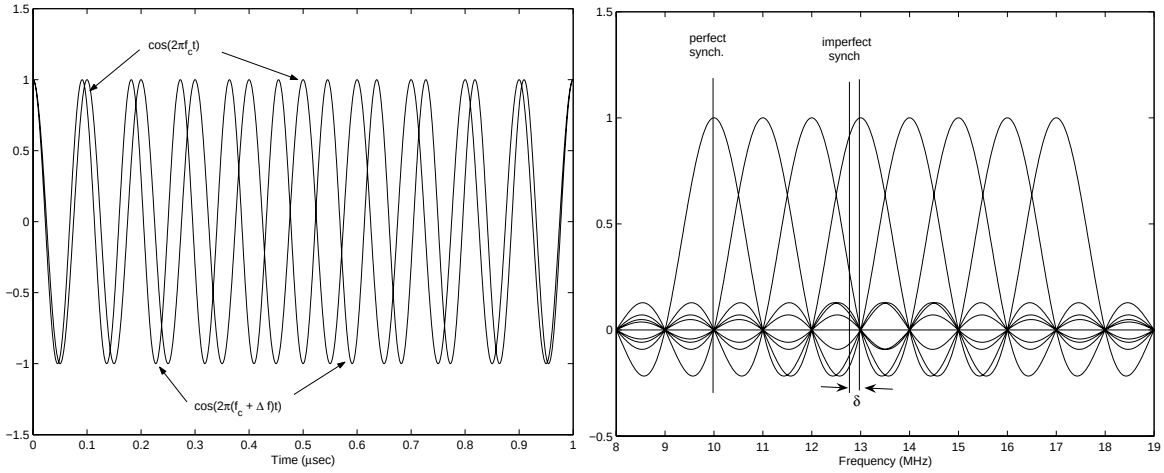


Figure 4.9: OFDM synchronization in time (left) and frequency (right). Here, two subcarriers in the time domain and eight subcarriers in the frequency domain are shown, where $f_c = 10$ MHz and the subcarrier spacing $\Delta f = 1$ Hz.

carrier interference (ICI) is zero. However, if the carrier frequency is misaligned by some amount δ, then some of the desired energy is lost, and more significantly, inter-carrier interference is introduced.

The following two subsections will more carefully examine timing and frequency synchronization. Although the development of good timing and frequency synchronization algorithms for WiMAX systems is the responsibility of each equipment manufacturer, we will give some general guidelines on what is required of a synchronization algorithm, and discuss the penalty for imperfect synchronization. It should be noted that synchronization is one of the most challenging problems in OFDM implementation, and the development of efficient and accurate synchronization algorithms presents an opportunity for technical differentiation and intellectual property.

4.4.1 Timing Synchronization

The effect of timing errors in symbol synchronization is somewhat relaxed in OFDM due to the presence of a cyclic prefix. In Section 4.2.3, we assumed that only the L time domain samples after the cyclic prefix were utilized by the receiver. Indeed, this corresponds to "perfect" timing synchronization, and in this case even if the cyclic prefix length N_g is equivalent to the length of the channel impulse response v, successive OFDM symbols can be decoded ISI free.

In the case that perfect synchronization is not maintained, it is still possible to tolerate a timing offset of τ seconds without any degradation in performance as long as $0 \leq \tau \leq T_m - T_g$, where as usual T_g is the guard time (cyclic prefix duration) and T_m is the maximum channel delay spread. Here, $\tau < 0$ corresponds to sampling earlier than at the ideal instant, whereas $\tau > 0$ is later than the ideal instant. As long as $0 \leq \tau \leq T_m - T_g$, the timing offset can be included by the channel estimator in the complex gain estimate for each subchannel and the appropriate phase shift can be applied by the FEQ without any loss in performance – at least in theory. This acceptable range of τ is referred to as the timing synchronization

margin, and is shown in Figure 4.10.

On the other hand, if the timing offset τ is not within this window $0 \leq \tau \leq T_m - T_g$, inter-symbol interference (ISI) occurs regardless of whether the phase shift is appropriately accounted for. This can be confirmed intuitively for the scenario that $\tau > 0$ and for $\tau < T_m - T_g$. For the case $\tau > 0$, the receiver loses some of the desired energy (since only delayed version of the early samples $x_0, x_1, ...$ are received), and also incorporates undesired energy from the subsequent symbol. Similarly for $\tau < T_m - T_g$, desired energy is lost while interference from the preceding symbol is included in the receive window. For both of these scenarios, the SNR loss can be approximated by

$$\Delta SNR(\tau) \approx -2\left(\frac{\tau}{LT_s}\right)^2, \tag{4.16}$$

which makes intuitive sense given the above arguments, and has been shown more rigorously in the literature on synchronization for OFDM. Important observations from this expression are that:

1. SNR decreases quadratically with the timing offset
2. Longer OFDM symbols are increasingly immune from timing offset, i.e. more subcarriers helps.
3. Since in general $\tau \ll LT_s$, timing synchronization errors are not that critical as long as the induced phase change is corrected.

In summary, to minimize SNR loss due to imperfect timing synchronization, the timing errors should be kept small compared to the guard interval, and a small margin in the cyclic prefix length is helpful.

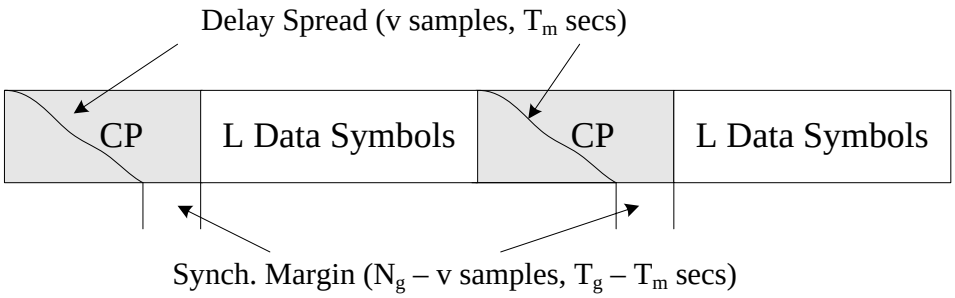


Figure 4.10:

4.4.2 Frequency Synchronization

OFDM achieves a high degree of bandwidth efficiency compared to other wideband systems. The subcarrier packing is extremely tight compared to conventional modulation techniques which require a guard band on the order of 50% or more, in addition to special transmitter architectures such as the Weaver architecture or Single-Sideband modulation that suppress the redundant negative-frequency portion of the passband signal. The price to be paid for this bandwidth efficiency is that the multicarrier signal shown in Figure 4.9 is very sensitive to frequency offsets due to the fact that the subcarriers overlap, rather than having each subcarrier truly spectrally isolated.

Since the zero crossings of the frequency domain sinc pulses all line up as seen in Figure 4.9, as long as the frequency offset $\delta = 0$, there is no interference between the subcarriers. One intuitive interpretation for this is that since the FFT is essentially a frequency sampling operation, if the frequency offset is negligible the receiver simply samples $\mathbf{y}$ at the peak points of the sinc functions, where the ICI is zero from all the neighboring subcarriers.

In practice, of course, the frequency offset is not always zero. The major causes for this are mismatched oscillators at the transmitter and receiver and Doppler frequency shifts due to mobility. Since precise crystal oscillators are expensive, tolerating some degree of frequency offset is essential in a consumer OFDM system like WiMAX. For example, if an oscillator is accurate to 0.1 parts per million (ppm), $f_{\text{offset}} \approx (f_c)(0.1ppm)$. If $f_c = 3$ GHz and the Doppler is 100 Hz, then $f_{\text{offset}} = 300 + 100$ Hz, which will degrade the orthogonality of the received signal, since now the received samples of the FFT will contain interference from the adjacent subcarriers. We'll now analyze this intercarrier interference (ICI) in order to better understand its effect on OFDM performance.

The matched filter receiver corresponding to subcarrier l can be simply expressed for the case of rectangular windows (neglecting the carrier frequency) as

$$x_l(t) = X_l e^{j\frac{2\pi lt}{LT_s}}, \tag{4.17}$$

where $1/LT_s = \Delta f$, and again LT_s is the duration of the data portion of the OFDM symbol, i.e. $T = T_g + LT_s$. An interfering subcarrier m can be written as

$$x_{l+m}(t) = X_m e^{j\frac{2\pi(l+m)t}{LT_s}}. \tag{4.18}$$

If the signal is demodulated with a fractional frequency offset of δ, $|\delta| \leq \frac{1}{2}$

$$\hat{x}_{l+m}(t) = X_m e^{j\frac{2\pi(l+m+\delta)t}{LT_s}}. \tag{4.19}$$

The ICI between subcarriers l and $l+m$ using a matched filter (i.e. the FFT) is simply the inner product between them:

$$I_m = \int_0^{LT_s} x_l(t)\hat{x}_{l+m}(t)dt = \frac{LT_s X_m \left(1 - e^{-j2\pi(\delta+m)}\right)}{j2\pi(m+\delta)}. \tag{4.20}$$

It can be seen that in the above expression, $\delta = 0 \Rightarrow I_m = 0$, and $m = 0 \Rightarrow I_m = 0$, as expected. The total average ICI energy per symbol on subcarrier l is then

$$ICI_l = E[\sum_{m \neq l} |I_m|^2] \approx C_0 (LT_s\delta)^2 \mathcal{E}_x, \tag{4.21}$$

where C_0 is a constant that depends of various assumptions and $\mathcal{E}_x$ is the average symbol energy [36, 27]. The approximation sign is due to the fact that this expression assumes that there are an infinite number of interfering subcarriers. Since the interference falls off quickly with m, this assumption is very accurate for subcarriers near the middle of the band, and is pessimistic by a factor of 2 for at either end of the band.

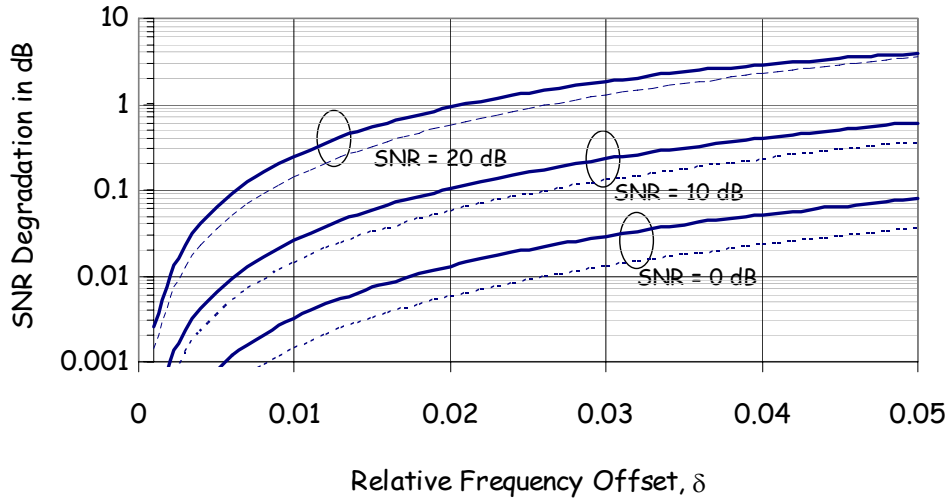


Figure 4.11: SNR loss as a function of the relative frequency offset δ. The solid lines are for a fading channel and the dotted lines are for an AWGN channel.

The SNR loss induced by frequency offset is given by

$$\Delta SNR = \frac{\mathcal{E}_x/N_o}{Ex/(N_o + C_0(LT_s\delta)^2\mathcal{E}_x)} \quad (4.22)$$

$$= 1 + C_0(LT_s\delta)^2 SNR \quad (4.23)$$

Important observations from the ICI expression (4.23) and Figure 4.11 are that:

1. SNR decreases quadratically with the frequency offset
2. SNR decreases quadratically with the number of subcarriers.
3. The loss in SNR is also proportional to the SNR itself.
4. In order to keep the loss negligible, say less than 0.1 dB, the frequency offset needs to be in the 1-2% range, or even lower to preserve high SNRs.
5. Therefore, this is a case where reducing the CP overhead by increasing the number of subcarriers causes an offsetting penalty, introducing a tradeoff.

In order to further reduce the ICI for a given choice of L, non-rectangular windows can also be used [30, 37].

4.4.3 Obtaining Synchronization in WiMAX

The preceding two sections discussed the consequences of imperfect time and frequency synchronization. There have been many synchronization algorithms developed in the literature, a by no means exhaustive list includes [39, 6, 42, 28, 16]. Generally, the proposed can be broken into the following two categories:

Pilot symbol based. In this class, known pilot symbols are transmitted. Since the receiver knows what was transmitted, attaining quick and accurate time and frequency synchronization is the easiest in this case, at the cost of surrendering some throughput. In the WiMAX downlink, the preamble consists of a known OFDM symbol that can be used to attain initial synchronization. In the WiMAX uplink, the periodic ranging (described in Chapter 9) can be used to synchronize. Since WiMAX is an OFDMA system with many competing users, each user implements the synchronization at the mobile. This requires the base station to communicate the frequency offset to the SS.

Blind, or Cyclic prefix based. "Blind" means that pilot symbols are not available to the receiver, so the receiver must do the best it can without explicitly being able to determine the effect of the channel. In the absence of pilot symbols, the cyclic prefix, which contains redundancy, can also be used to attain time and frequency synchronization [39]. This technique is effective when the number of subcarriers is large, or when the offsets are estimated over a number of consecutive symbols. The principle benefit of CP-based methods are that pilot symbols are not needed, so the data rate can nominally be increased. In WiMAX, accurate synchronization and especially channel estimation are considered important enough to warrant the use of pilot symbols, so the blind techniques are not usually used for synchronization. They could be used to track the channel in between preambles or ranging signals, but the frequency and timing offsets generally vary slowly enough that this is not required.

4.5 Channel Estimation in OFDM

TBD. This may just get covered in the Chapter 6 section on MIMO-OFDM channel estimation, since OFDM is just a special case of that.

4.6 The Peak to Average Ratio

OFDM signals have a higher Peak-to-Average Ratio (PAR) – often called a Peak-to-Average Power Ratio (PAPR) – than single carrier signals. The reason for this is that in the time domain, a multicarrier signal is the sum of many narrowband signals. At some time instances, this sum is large, at other times it is small, which means that the peak value of the signal is substantially larger than the average value. This high PAR is one of the most important implementation challenges that faces OFDM because it reduces the efficiency and hence increases the cost of the RF power amplifier, which is one of the most expensive components in the radio. In this section, we will quantify the PAR problem, explain its severity in WiMAX, and briefly overview some strategies for reducing the PAR.

4.6.1 The PAR Problem

When a high peak signal is transmitted through a nonlinear device such as a high power amplifier (HPA) or digital-to-analog converter (DAC), it generates out-of-band energy (spectral regrowth) and in-band distortion (constellation tilting and scattering). These degradations may affect the system performance severely. The nonlinear behavior of HPA can be characterized by amplitude modulation-amplitude modulation (AM/AM)

and amplitude modulation-phase modulation (AM/PM) responses. Figure 4.12 shows a typical AM/AM response for a HPA, with the associated input and output back off regions; IBO and OBO, respectively.

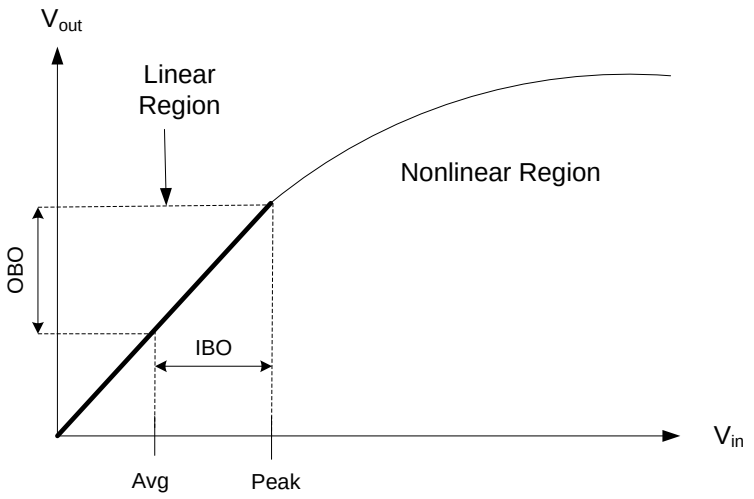


Figure 4.12: A typical power amplifier response. Operation in the linear region is required in order to avoid distortion, so the peak value must be constrained to be in this region, which means that on average, the power amplifier is underutilized by a "backoff" amount.

To avoid the undesirable nonlinear effects just mentioned, a waveform with high peak power must be transmitted in the linear region of the HPA by decreasing the average power of the input signal. This is called (input) *backoff* (IBO) and results in a proportional output backoff (OBO). High backoff reduces the power efficiency of the HPA, and may limit the battery life for mobile applications. In addition to inefficiency in terms of power, the coverage range is reduced and the cost of the HPA is higher than would be mandated by the average power requirements.

The input backoff is defined as

$$IBO = 10 \log_{10} \frac{P_{inSat}}{\bar{P}_{in}}, \tag{4.24}$$

where P_{inSat} is the saturation power (above which is the nonlinear region) and $\bar{P}_{in}$ is the average input power. The amount of backoff is usually greater than or equal to the PAR of the signal.

The power efficiency of a HPA can be increased by reducing the PAR of the transmitted signal. For example, the efficiency of a class A amplifier is halved when the input PAR is doubled or the operating point (average power) is halved [13, 5]. The theoretical efficiency limits for two classes of HPAs are shown in Fig. 4.13. Clearly, it would be desirable to have the average and peak values be as close together as possible in order to maximize the efficiency of the power amplifier.

In addition to the large burden placed on the HPA, a high PAR requires high resolution for both the transmitter's digital-to-analog convertor (DAC) and the receiver's analog-to-digital convertor (ADC), since the dynamic range of the signal is proportional to the PAR. High resolution D/A and A/D conversion places an additional complexity, cost, and power burden on the system.

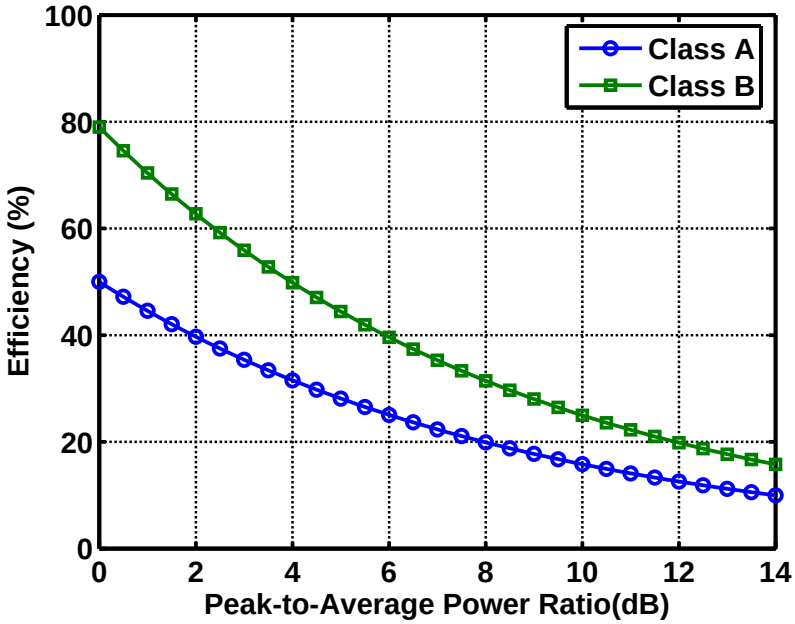


Figure 4.13: Theoretical efficiency limits of linear amplifiers [26]. A typical OFDM PAR is in the 10 dB range, so the PA efficiency is 50-75% lower than in a single carrier system.

4.6.2 Quantifying the PAR

Since multicarrier systems transmit data over a number of parallel frequency channels, the resulting waveform is the superposition of L narrowband signals. In particular, each of the L output samples from a L-pt IFFT operation involves the sum of L complex numbers, as can be seen from (4.6). Due to the Central Limit Theorem, the resulting output values $\{x_1, x_2, \ldots, x_L\}$ can be accurately modelled (particularly for large L) as complex Gaussian random variables with zero mean and variance $\sigma^2 = \mathcal{E}_x/2$, i.e. the real and imaginary parts both have zero mean and variance $\sigma^2 = \mathcal{E}_x/2$. The amplitude of the output signal is

$$|x[n]| = \sqrt{(Re\{x[n]\})^2 + (Im\{x[n]\})^2}, \tag{4.25}$$

which is Rayleigh distributed with parameter σ^2. The output power is therefore

$$|x[n]|^2 = (Re\{x[n]\})^2 + (Im\{x[n]\})^2, \tag{4.26}$$

which is exponentially distributed with mean $2\sigma^2$. The important thing to note is that the output amplitude and hence power are *random*, so the PAR is not a deterministic quantity either.

The PAR of the transmitted analog signal can be defined as

$$PAR \triangleq \frac{\max\limits_{t} |x(t)|^2}{E[|x(t)|^2]}, \tag{4.27}$$

where naturally the range of time to be considered has to be bounded over some interval. Generally, the PAR is considered for a single OFDM symbol, which consists of $L + N_g$ samples, or a time duration of T,

as this chapter has explained. Similarly, the discrete-time PAR can be defined for the IFFT output as

$$PAR \triangleq \frac{\max\limits_{l\in(0,L+N_g)} |x_l|^2}{E[|x_l|^2]} = \frac{\mathcal{E}_{\max}}{\mathcal{E}_x}. \tag{4.28}$$

It is important to recognize, however, that although the average energy of IFFT outputs $x[n]$ is the same as the average energy of the inputs $X[m]$ and equal to $\mathcal{E}_x$, the analog PAR is *not* generally the same as the PAR of the IFFT samples, due to the interpolation performed by the D/A convertor. Usually, the analog PAR is higher than the digital (Nyquist sampled) PAR. Since the PA is by definition analog, the analog PAR is what determines the PA performance. Similarly, DSP techniques developed to reduce the digital PAR may not always have the anticipated effect on the analog PAR, which is what matters. In order to bring the analog PAR expression in (4.27) and the digital PAR expression in (4.28) closer together, oversampling can be considered for the digital signal. That is, a factor M additional samples can be used to interpolate the digital signal in order to better approximate its analog PAR.

It can be proven that the maximum possible value of the PAR is L, which occurs when all the subcarriers add up constructively at a single point. However, although it is possible to choose an input sequence that results in this very high PAR, such an expression for PAR is misleading. For independent binary inputs, for example, the probability of this maximum peak value occurring is on the order of 2^{-L}.

Since the theoretical maximum (or similar) PAR value seldom occurs, a statistical description of the PAR is commonly used. The complementary cumulative distribution function (CCDF = 1 - CDF) of the PAR is the most commonly used measure. The distribution of the OFDM PAR has been studied by many researchers [43, 32, 31, 3]. Among these, van Nee and de Wild introduced a simple and accurate approximation of the CCDF for large $L(\geq 64)$:

$$\mathrm{CCDF(L}, \mathcal{E}_{\max}) = 1 - G(L, \mathcal{E}_{\max}) = 1 - F(L, \mathcal{E}_{\max})^{\beta L} = 1 - \left(1 - \exp(\frac{\mathcal{E}_{\max}}{2\sigma^2})\right)^{\beta L}, \tag{4.29}$$

where $\mathcal{E}_{\max}$ is the peak power level and β is a pseudo-approximation of the oversampling factor which is given empirically as $\beta = 2.8$ [43]. Note that the PAR is $\mathcal{E}_{\max}/2\sigma^2$ and $F(L, \mathcal{E}_{\max})$ is the CDF of a single Rayleigh distributed subcarrier with parameter σ^2. The basic idea behind this approximation is that unlike a Nyquist[5] sampled signal, the samples of an oversampled OFDM signal are correlated, making it hard to derive an exact peak distribution. The CDF of the Nyquist sampled signal power can be obtained by

$$G(L, \mathcal{E}_{\max}) = P(\max ||x(t)|| \leq \mathcal{E}_{\max}) = F(L, \mathcal{E}_{\max})^L, \tag{4.30}$$

Using this result as a baseline, the oversampled case can be approximated in a heuristic way by regarding the oversampled signal as generated by βL Nyquist sampled subcarriers. Note however that β is not equal to the oversampling factor M. This simple expression is quite effective for generating accurate PAR statistics for various scenarios, and sample results are displayed in Figure 4.14. As expected, the approximation is accurate for large L and the PAR of OFDM system increases with L but not nearly linearly.

[5]Nyquist sampling means the minimum allowable sampling frequency without irreversible information loss, i.e. no oversampling is performed.

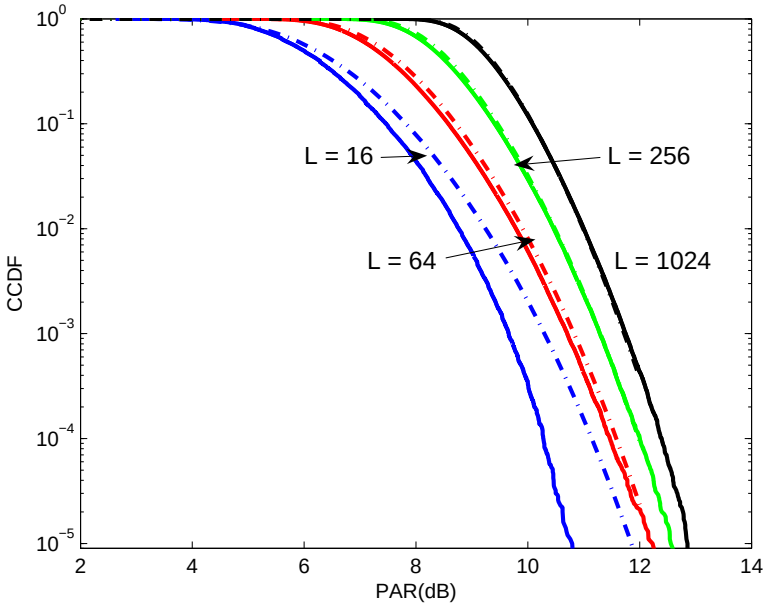


Figure 4.14: CCDF of PAR for QPSK OFDM system: L = 16, 64, 256, 1024. Solid line: simulation results, dotted line: approximation using $\beta = 2.8$.

4.6.3 Clipping: Living with a High PAR

In order to avoid operating the PA in the nonlinear region, the input power can be reduced by an amount about equal to the PAR. However, two important facts related to this IBO amount can be observed from Figure 4.14. First, since the highest PAR values are uncommon, it might be possible to simply "clip" off the highest peaks in order to reduce the IBO and effective PAR, at the cost of some hopefully minimal distortion of the signal. Second and conversely, it can be seen that even for a conservative choice of IBO, say 10dB, there is still a distinct possibility that a given OFDM symbol will have a PAR that exceeds the IBO and causes clipping.

Clipping, sometimes called "soft limiting", truncates the amplitude of signals that exceed the clipping level as

$$\tilde{x}_L[n] = \begin{cases} Ae^{j\angle x[n]}, & \text{if } |x[n]| > A \\ x[n], & \text{if } |x[n]| \leq A, \end{cases} \tag{4.31}$$

where $x[n]$ is the original signal, and $\tilde{x}[n]$ is the output after clipping. The soft limiter can be equivalently thought of as a peak cancellation technique, like that shown in Figure 4.15. The soft limiter output can be written in terms of the original signal and a cancelling, or "clipping signal" as

$$\tilde{x}[n] = x[n] + c[n], \quad \text{for } n = 0, \ldots, L-1 \tag{4.32}$$

where $c[n]$ is the clipping signal defined by

$$c[n] = \begin{cases} |A - |x[n]||\mathrm{e}^{j\theta[n]}, & \text{if } |x[n]| > A \\ 0, & \text{if } |x[n]| \leq A \end{cases} \qquad \text{for } n = 0, ..., NL-1 \tag{4.33}$$

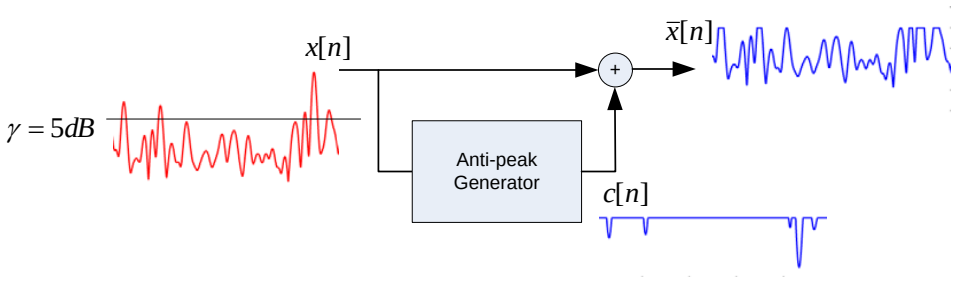


Figure 4.15: A peak cancellation as a model of soft limiter when $\gamma = 5dB$

where $\theta[n] = \arg(-x[n])$; i.e, the phase of $c[n]$ is out of phase with $x[n]$ by 180^o, and A is the *clipping level* which is defined as

$$\gamma \triangleq \frac{A}{\sqrt{E\{|x[n]|^2\}}} = \frac{A}{\sqrt{\mathcal{E}_x}} \tag{4.34}$$

In such a peak cancellation strategy, an anti-peak generator estimates peaks greater than clipping level. The clipped signal $\tilde{x}[n]$ can be obtained by adding a time-shifted and scaled signal $c[n]$ to the original signal $x[n]$. The exact clipping signal $c[n]$ can be generated to reduce PAR using a variety of techniques that are introduced shortly.

Obviously, clipping reduces the PAR at the expense of distorting the signal by the additive signal $c[n]$. The two primary drawbacks from clipping are (1) spectral regrowth (frequency domain leakage), which causes unacceptable interference to users in neighboring RF channels, and (2) distortion of the desired signal. We'll now consider these two effects separately.

Spectral Regrowth

. The clipping noise can be expressed in the frequency domain through the use of the DFT. The resulting clipped frequency domain signal $\tilde{X}$ is

$$\tilde{X}_k = X_k + C_k, \quad k = 0, \ldots, L-1, \tag{4.35}$$

where C_k represents the clipped off signal in the frequency domain. In Figure 4.16, the power spectral density of the original (X), clipped ($\tilde{X}$) and clipped-off (C) signals are plotted for different clipping ratios γ of 3, 5 and 7dB. The following deleterious effects are observed. First, the clipped-off signal C_k is strikingly increased as the clipping ratio is lowered from 7dB to 3dB. This increase shows the correlation between X_k and C_k inside the desired band at low clipping ratios, and causes the in-band signal to be attenuated as the clipping ratio is lowered. Second, it can be seen that the out-of-band interference caused by the clipped signal $\tilde{X}$ is determined by the shape of clipped-off signal C_k. Even the seemingly conservative clipping ratio of 7dB violates the specification for the transmit spectral mask of IEEE 802.16e, albeit barely.

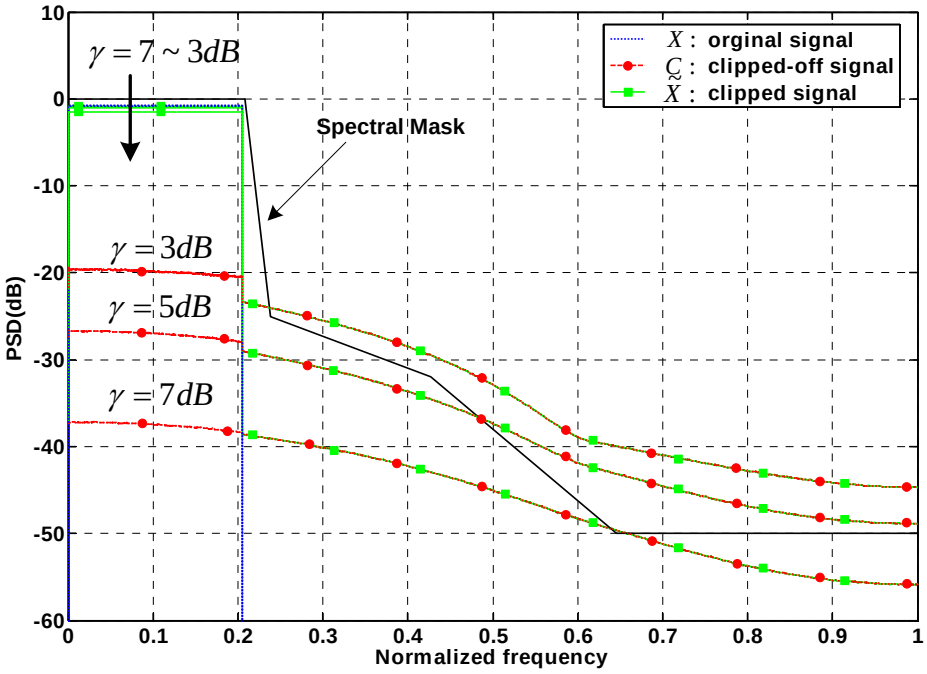


Figure 4.16: Power Spectral Density of the unclipped (original) and clipped (nonlinearly distorted) OFDM signals with 2048 block size and 64 QAM when clipping ratio (γ) is 3, 5, and 7dB in soft limiter.

In-band Distortion

. Although the desired signal and the clipping signal are clearly correlated, it is possible based on the Bussgang Theorem to model the in-band distortion due to the clipping process as the combination of uncorrelated additive noise and an attenuation of the desired signal [15, 14, 33]:

$$\tilde{x}[n] = \alpha x[n] + d[n], \quad \text{for } n = 0, 1, \ldots, L-1. \tag{4.36}$$

Now, $d[n]$ is uncorrelated with the signal $x[n]$ and the attenuation factor α is obtained by

$$\alpha = 1 - e^{-\gamma^2} + \frac{\sqrt{\pi}\gamma}{2}\text{erfc}(\gamma). \tag{4.37}$$

The attenuation factor α is plotted in Figure 4.17 as a function of the clipping ratio γ. The attenuation factor α is negligible when the clipping ratio γ is greater than 8dB, so high clipping ratios, the correlated clipped-off signal $c[n]$ in (4.33) can be approximated by uncorrelated noise $d[n]$. That is, $c[n] \approx d[n]$ as $\gamma \uparrow$. The variance of the uncorrelated clipping noise can be expressed assuming a stationary Gaussian input $x[n]$ as

$$\sigma_d^2 = \mathcal{E}_x(1 - \exp(-\gamma^2) - \alpha^2). \tag{4.38}$$

In WiMAX, the error vector magnitude (EVM) is used as a means to estimate the accuracy of the transmit filter and D/A converter as well as the PA non-linearity. The EVM is essentially the average error

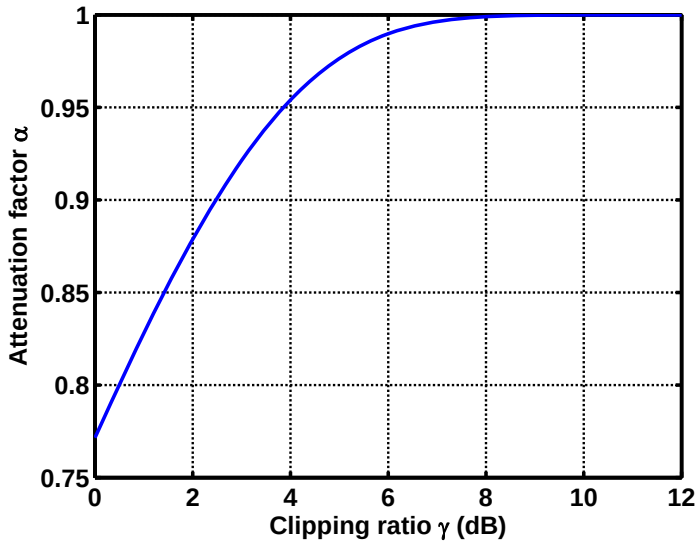


Figure 4.17: Attenuation factor α as a function of the clipping ratio γ

vector relative to the desired constellation point, and can be caused by an degradation in the system. The EVM over an OFDM symbol is defined as

$$\text{EVM} = \sqrt{\frac{\frac{1}{N}\sum_{k=1}^{L}(\Delta I_k^2 + \Delta Q_k^2)}{S_{\max}^2}} = \sqrt{\frac{\sigma_d^2}{S_{\max}^2}} \tag{4.39}$$

where $S_{\max}$ the maximum constellation amplitude. The concept of EVM is visually illustrated in Figure 4.18. In the case of clipping, a given EVM specification can be easily translated into an SNR requirement using the variance σ_d^2 of the clipping noise.

It is possible to define the signal-to-noise-plus-distortion ratio (SNDR) of one OFDM symbol in order to estimate the impact of clipped OFDM signals over an AWGN channel under the assumption that the distortion $d[n]$ is Gaussian and uncorrelated with the input and channel noise (that has variance $N_0/2$):

$$\text{SNDR} = \frac{\alpha^2 \mathcal{E}_x}{\sigma_d^2 + N_0/2}. \tag{4.40}$$

The bit-error probability (BEP) can be evaluated for different modulation types using the SNDR [14]. In the case of M-QAM and average power $\mathcal{E}_x$, the BEP can then be approximated as

$$P_b \approx \frac{4}{\log_2 M}\left(1 - \frac{1}{\sqrt{M}}\right) Q\left(\sqrt{\frac{3\mathcal{E}_x\alpha^2}{(\sigma_d^2 + N_0/2)(M-1)}}\right). \tag{4.41}$$

Figure 4.19 shows the BER for an OFDM system with L = 2048 subcarriers and 64-QAM modulation. As the SNR increases, the clipping error dominates the additive noise and an error floor is observed. The error floor can be inferred from (4.41) by letting the noise variance $N_0/2 \to 0$.

A number of additional studies of clipping in OFDM systems have been completed in recent years, e.g. [38, 14, 4, 3, 31]. In some cases clipping may be acceptable, but in WiMAX systems the margin for error

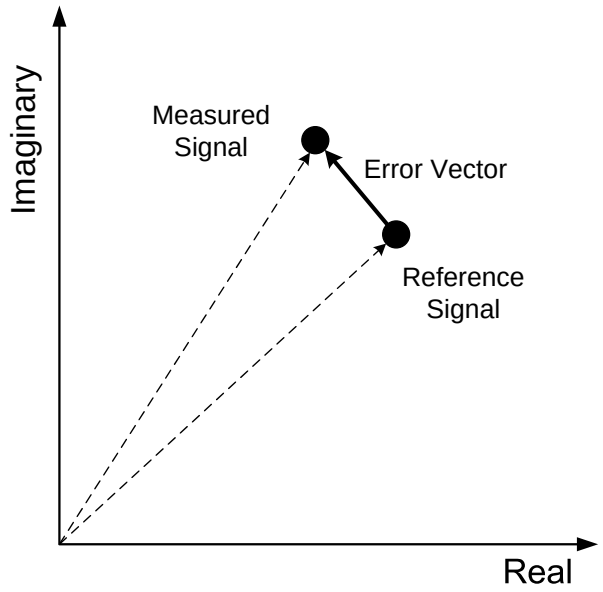


Figure 4.18: Illustrative example of EVM

is quite tight, as much of this book has emphasized. Hence, more aggressive and innovative techniques for reducing the PAR are being actively pursued by the WiMAX community in order to bring down the component cost and to reduce the degradation due to the nonlinear effects of the PA.

4.6.4 PAR Reduction Strategies

To alleviate the nonlinear effects mentioned above, numerous approaches have been pursued. The first plan of attach is to reduce PAR at the transmitter, either through peak cancellation or signal mapping [20]. Another set of techniques focuses on OFDM signal reconstruction at the receiver in spite of the introduced nonlinearities, e.g. [23, 41]. A further approach is to attempt to predistort the analog signal so that it will appear to have been linearly amplified [8]. In this section, we will overview techniques that attempt PAR reduction at the transmitter.

Peak Cancellation

This class of PAR reduction techniques applies an anti-peak signal to the desired signal. Although clipping is the most obvious such technique, and was discussed in detail in the last section, other important peak cancelling techniques include tone reservation (TR) and active constellation extension (ACE).

Clipping can be improved by an iterative clipping and filtering process, since the filtering can be used to subdue the spectral regrowth and in band distortion [2]. After several iterations of clipping and filtering, the residual in-band distortion can be restored by iterative estimation and cancellation of the clipping noise [9].

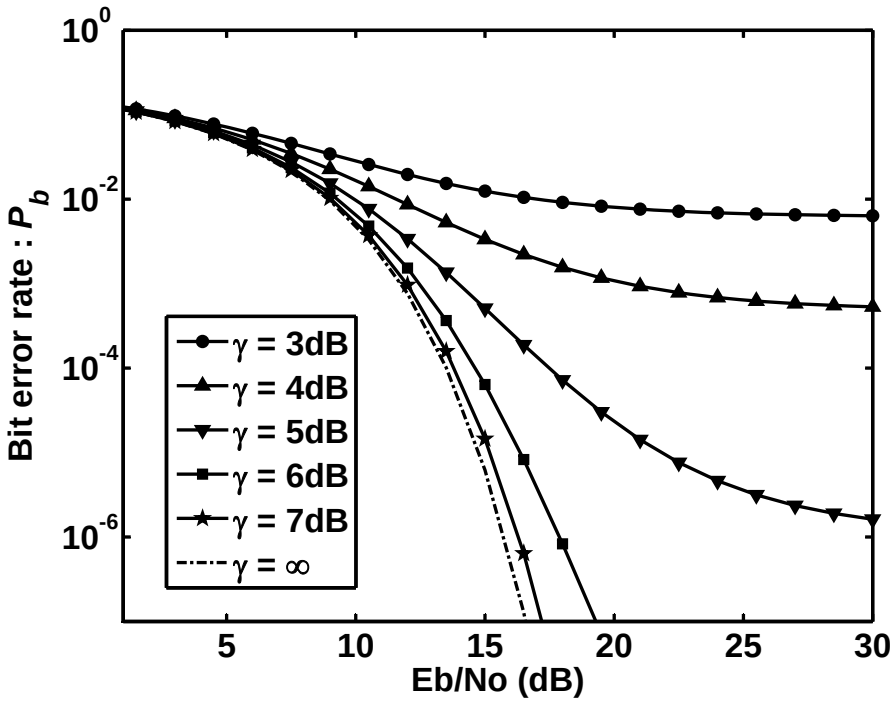


Figure 4.19: Bit error rate probability for a clipped OFDM signal in AWGN with different clipping ratios.

Tone reservation reduces PAR by intelligently adding power to unused carriers, such as the null subcarriers specified by WiMAX. The reduced signal can be expressed as

$$\tilde{x}[n] = x[n] + c[n] = Q_L(\mathbf{X} + \mathbf{C}) \tag{4.42}$$

where Q_L is the IDFT matrix of size $L \times L$, $\mathbf{X} = \{X_k, k \in \mathcal{R}^c\}$ is complex symbol vector, $\mathbf{C} = \{C_k, k \in \mathcal{R}\}$ is a complex tone reservation vector, and $\mathcal{R}$ is the reserved tones set. There is no distortion from TR because the reserved tone carriers and the data carriers are orthogonal. To reduce PAR using TR, a simple gradient algorithm and a Fourier projection algorithm have been proposed in [40] and [18], respectively. However, these technique converge slowly. For faster convergence, an active set approach is used in [25].

Another peak cancelling technique is active constellation extension (ACE) [24]. Essentially, the corner points of an M-QAM constellation can be extended without any loss of SNR, and this property can be used to decrease the PAR without negatively effecting the performance as long as ACE be allowed only in the case that the minimum distance is guaranteed. Unfortunately, the gain in PAR reduction is inversely proportional to the constellation size in M-QAM.

Signal Mapping

Signal mapping techniques share in common that some redundant information is added to the transmitted signal in a manner which reduces the PAR. This class includes coding techniques, selected mapping (SLM), and partial transmit sequence (PTS).

The main idea behind the various coding schemes is to select a low PAR codeword based on the desired transmit symbols [22, 35]. However, most of the decoding techniques for these codes require an exhaustive search so are feasible only for a small number of subcarriers. Moreover, it is difficult to maintain a reasonable coding rate in OFDM when the number of subcarriers grows large. The implementation prospects for the coding-based techniques appear dim.

In Selected mapping, one OFDM symbol is used to generate multiple representations that have the same information as the original symbol [29]. The basic objective is to select the one with minimum PAR, and the gain in PAR reduction is proportional to the number of the candidate symbols, but so is the complexity.

PTS is similar to SLM; however, the symbol in the frequency domain is partitioned into smaller disjoint subblocks. The objective is to design an optimal phase for the subblock set that minimizes the PAR. The phase can then be corrected at the receiver. The PAR reduction gain depends on the number of subblocks and the partitioning method. However, PTS has exponential search complexity with the number of subblocks.

SLM and PTS are quite flexible and effective, but their principle drawbacks are that the receiver structure must be changed and transmit overhead (power and symbols) are required to send the needed information for decoding. Hence, these techniques, in contrast to peak cancellation techniques, would require explicit support by the WiMAX standard.

4.7 The Computational Complexity Advantage of OFDM

One of the principal advantages of OFDM relative to single carrier modulation with equalization is that it requires much lower computational complexity for high data rate communication. In this section we compare the computational complexity of an equalizer with that of a standard IFFT/FFT implementation of OFDM.

An equalizer operation consists of a series of multiplications with several delayed versions of the signal. The number of delay taps in an equalizer depends on the symbol rate of the system and the delay spread in the channel. To be more precise, the number of equalizer taps is proportional to the bandwidth-delay spread product $T_m/T_s \approx BT_m$. We have been calling this quantity v, or the number of ISI channel taps. An equalizer with v taps performs v complex multiply and accumulate (CMAC) operations per received symbol. Therefore, the complexity of an equalizer is of the order

$$O(v \cdot B) = O(B^2 T_m). \tag{4.43}$$

In an OFDM system the IFFT and FFT are the main computational operations. It is well known that the IFFT and FFT each have a complexity of $O(L \log_2 L)$, where L is the FFT block size. In the case of OFDM, L is the number of subcarriers. As this chapter has shown, for a fixed cyclic prefix overhead, the number of subcarriers L must grow linearly with the bandwidth-delay spread product $v = BT_m$. Therefore, the computational complexity for each OFDM symbol is of the order $O(BT_m \log_2 BT_m)$. There are B/L OFDM symbols sent each second. Since $L \propto BT_m$, this means there are order $O(1/T_m)$ OFDM symbols per second, so the computational complexity in terms of CMACs for OFDM is

$$O(BT_m \log_2 BT_m) O(1/T_m) = O(B \log_2 BT_m). \tag{4.44}$$

Clearly, the complexity of an equalizer grows as the square of the data rate since both the symbol rate and the number of taps increases linearly with the data rate. For an OFDM system, the increase in complexity grows with the data rate only slightly faster than linearly. This difference is dramatic for very large data rates, as shown in Figure 4.20.

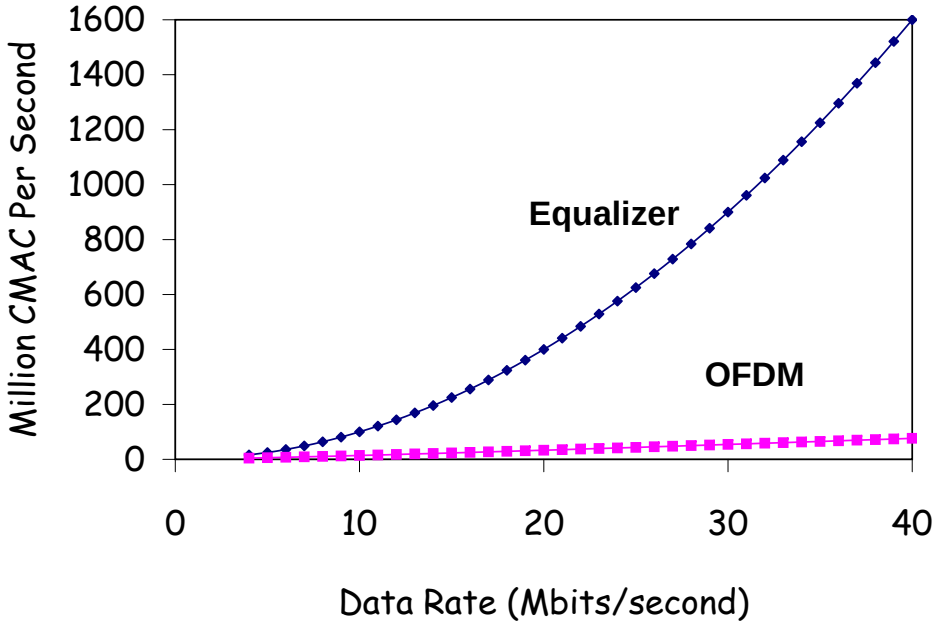


Figure 4.20: OFDM has an enormous complexity advantage over equalization for broadband data rates. The delay spread is $T_m = 2\mu$sec, the OFDM symbol period is $T = 20\mu$sec, and the considered equalizer is a DFE.

4.8 Simulating OFDM Systems

In this section we provide some resources to help engineers, students, and others get started on simulating an OFDM system. The popular LabVIEW simulation package from National Instruments can be used to develop VI's that implement OFDM in a graphical user interface. One of the author's research group at UT Austin has developed such a resource on the web [1] at:
http://www.ece.utexas.edu/~jandrews/molabview.html.html
Other communication system building blocks and multicarrier modulation tools are available in the National Instruments Modulation Toolkit.

OFDM functions can also be developed in Matlab[6]. Here, we provide Matlab code for a baseband OFDM transmitter and receiver for QPSK symbols. These functions can be modified to transmit and receive M-QAM symbols or passband (i.e. complex baseband) signals.

[6]These functions also can be used in MathScript for LabVIEW

```
function x=OFDMTx(N, bits, num, nu)
%=====================================================
% x=OFDMTx( N, bits, num, nu,zero_tones)
%
% QPSK transmitter for OFDM
%
% N is the FFT size
% bits is a {1,-1} stream of length (N/2-1)*2*num (baseband, zero DC)
% num is the number of OFDM symbols to produce
% nu is the cyclic prefix length
%=====================================================

if length(bits) ~= (N/2-1)*2*num
  error('bits vector of improper length -- Aborting');
end x=[];
real_index = -1;   % initial values
imag_index = 0;

for a=1:num
  real_index = max(real_index)+2:2:max(real_index)+N-1;
  imag_index = max(imag_index)+2:2:max(imag_index)+N-1;
  X = (bits(real_index) + j*bits(imag_index))/2;
  X=[0 X 0];     % zero nyquist and DC
  x_hold=sqrt(N)*ifft([X,conj(fliplr(X(2:length(X)-1)))]); %Baseband so Symmetric
  x_hold=[x_hold(length(x_hold)-nu+1:length(x_hold)),x_hold]; % Add CP
  x=[x,x_hold];
end
x=real(x);     % Baseband so take only the real part


function bits_out = OFDMRx(N,y,h,num_symbols,nu)
%=====================================================
% bits_out = OFDMRx(M,N,y,num_symbols,nu,zero_tones)
%
% N is the FFT size
% y is received channel output for all symbols (excess delay spread removed)
% h is the (estimated) channel impulse response
% num_symbols is the # of symbols (each of N*sqrt(M) bits and N+nu samples)
% nu is the cyclic prefix length
```

```
%=====================================================
if length(y) ~= (N + nu)*num_symbols
  error('received vector of improper length -- Aborting');
end bits_out=[]; bits_out_cur = zeros(1,N-2);

for a=1:num_symbols
  y_cur = y((a-1)*(N+nu)+1+nu:a*(N+nu));  % Get current OFDM symbol, strip CP
  X_hat = 1/sqrt(N)*fft(y_cur)./fft(h,N);  % FEQ
  X_hat = X_hat(1:N/2);
  real_index = 1:2:N-2-1;  % Don't inc. X_hat(1) because is DC, zeroed
  imag_index = 2:2:N-2;
  bits_out_cur(real_index) = sign(real(X_hat(2:length(X_hat))));
  bits_out_cur(imag_index) = sign(imag(X_hat(2:length(X_hat))));
  bits_out=[bits_out,bits_out_cur];
end
```

Sidebar 4.2 *A brief history of OFDM*

Although OFDM has become widely used only recently, the concept dates back some 40 years. In order to provide context for the reader, we give a brief history here of some landmark dates in OFDM.

1966: Chang shows in the Bell Labs technical journal that multicarrier modulation can solve the multipath problem without reducing data rate [7]. This is generally considered the first official publication on multicarrier modulation, although there were naturally precursory studies, including Holsinger's 1964 MIT dissertation [21] and some of Gallager's early work on waterfilling [17].

1971: Weinstein and Ebert show that multicarrier modulation can be accomplished using a "Discrete Fourier Transform" (DFT) [44].

1985: Cimini at Bell Labs identifies many of the key issues in OFDM transmission and does a proof of concept design [10].

1993: DSL adopts OFDM, also called "Discrete Multitone", following successful field trials/competitions at Bellcore vs. equalizer-based systems.

1999: IEEE 802.11 committee on wireless LANs releases 802.11a standard for OFDM operation in 5GHz UNI band.

2002: IEEE 802.16 committee releases OFDM based standard for wireless broadband access for metropolitan area networks under revision 802.16a.

2003: The "multi-band OFDM" standard for ultrawideband is developed, showing OFDM's usefulness in low-SNR systems.

4.9 Chapter Summary

This chapter has covered the theory of OFDM, as well as the important design and implementation related issues. The following were the key points:

- OFDM overcomes even severe inter-symbol interference through the use of the IFFT and a Cyclic Prefix.
- OFDM is the core modulation strategy used in WiMAX systems. Its multiple-access cousin OFDMA is the subject of Chapter 7.
- Key details on OFDM implementation such as synchronization and managing the peak-to-average ratio were covered in depth.
- In order to aid WiMAX engineers and students, examples relating OFDM to WiMAX, including simulation code, were provided.

Some references used in compiling this chapter not cited in the text: [19, 11, 12].

Bibliography

[1] J. G. Andrews and T. Kim. MIMO-OFDM Design using LabVIEW. `http://www.ece.utexas.edu/~jandrews/molabview.html.html`.

[2] J. Armstrong. Peak-to-average power reduction for OFDM by repeated clipping and frequency domain filtering. *Electronics Letters*, 38(8):246–247, Feb. 2002.

[3] A. Bahai, M. Singh, A. Goldsmith, and B. Saltzberg. A new approach for evaluating clipping distortion in multicarrier systems. *IEEE Journal on Selected Areas in Communications*, 20(5):1037–1046, 2002.

[4] P. Banelli and S. Cacopardi. Theoretical analysis and performance of OFDM signals in nonlinear AWGN channels. *IEEE Trans. on Communications*, 48(3):430–441, Mar. 2000.

[5] R. Baxley and G. Zhou. Power savings analysis of peak-to-average power ratio in OFDM. *IEEE Transactions on Consumer Electronics*, 50(3):792–798, 2004.

[6] H. Bolcskei. Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems. *IEEE Trans. on Communications*, 49:988–99, June 2001.

[7] R. W. Chang. Synthesis of band-limited orthogonal signals for multichannel data transmission. *Bell Systems Technical Journal*, 45:1775–96, Dec. 1966.

[8] S. Chang and E. J. Powers. A simplified predistorter for compensation of nonlinear distortion in OFDM systems. pages 3080–84, San Antonio, TX, Nov. 2001.

[9] H. Chen and A. Haimovich. Iterative estimation and cancellation of clipping noise for OFDM signals. *IEEE Communications Letters*, 7(7):305– 307, 2003.

[10] L. Cimini. Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing. *IEEE Trans. on Info. Theory*, 33(7):665–75, July 1985.

[11] J. M. Cioffi. *Digital Communications, Chapter 4: Multichannel Modulation*. Unpublished course notes, available at http://www.stanford.edu/class/ee379c/.

[12] J. M. Cioffi. A multicarrier primer. Stanford University/Amati T1E1 contribution, I1E1.4/91-157, Nov. 1991.

[13] S. C. Cripps. *RF Power Amplifiers for Wireless Communications*. Artech House, Norwood, MA, 1999.

[14] D. Dardari, V. Tralli, and A. Vaccari. A theoretical characterization of nonlinear distortion effects in OFDM systems. *IEEE Trans. on Communications*, 48(10):1755– 1764, Oct. 2000.

[15] M. Friese. On the degradation of OFDM-signals due to peak-clipping in optimally predistorted power amplifiers. In *Proc., IEEE Globecom*, pages 939–944, Nov. 1998.

[16] T. Fusco. *Synchronization techniques for OFDM systems*. PhD thesis, Universita di Napoli Federico II, 2005.

[17] R. G. Gallager. *Information Theory and Reliable Communications*. Wiley, 1968.

[18] A. Gatherer and M. Polley. Controlling clipping probability in DMT transmission. In *Proceedings of the Asilomar Conference on Signals, Systems and Computers*, pages 578–584, Nov. 1997.

[19] A. J. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[20] S. H. Han and J. H. Lee. An overview of peak-to-average power ratio reduction techniques for multicarrier transmission. *IEEE Wireless Communications*, 12(2):56–65, 2005.

[21] J. L. Holsinger. *Digital communication over fixed time-continuous channels with memory, with special application to telephone channels*. PhD thesis, Massachusetts Institute of Technology, 1964.

[22] A. E. Jones, T. A. Wilkinson, and S. K. Barton. Block coding scheme for reduction of peak to mean envelope power ratio of multicarrier transmission schemes. *Electronics Letters*, 30(25):2098–2099, Dec. 1994.

[23] D. Kim and G. Stuber. Clipping noise mitigation for OFDM by decision-aided reconstruction. *IEEE Communications Letters*, 3(1):4–6, Jan. 1999.

[24] B. Krongold and D. Jones. PAR reduction in OFDM via active constellation extension. *IEEE Transactions on Broadcasting*, 49(3):258– 268, 2003.

[25] B. Krongold and D. Jones. An active-set approach for OFDM PAR reduction via tone reservation. *IEEE Trans. on Signal Processing*, 52(2):495– 509, Feb. 2004.

[26] S. Miller and R. O'Dea. Peak power and bandwidth efficient linear modulation. *IEEE Trans. on Communications*, 46(12):1639–1648, Dec. 1998.

[27] P. Moose. A technique for orthogonal frequency division multiplexing frequency offset correction. *IEEE Trans. on Communications*, 42(10):2908–14, Oct. 1994.

[28] M. Morelli and U. Mengali. An improved frequency offset estimator for OFDM applications. *IEEE Communications Letters*, 3(3), Mar. 1999.

[29] S. Müller and J. Huber. A comparison of peak power reduction schemes for OFDM. In *Proc., IEEE Globecom*, pages 1–5, Nov. 1997.

[30] C. Muschallik. Improving an OFDM reception using an adaptive nyquist windowing. *IEEE Trans. Consumer Electron.*, 42(3):259–69, Aug. 1996.

[31] H. Nikopour and S. Jamali. On the performance of OFDM systems over a Cartesian clipping channel: a theoretical approach. *IEEE Transactions on Wireless Communications*, 3(6):2083–2096, 2004.

[32] H. Ochiai and H. Imai. On the distribution of the peak-to-average power ratio in OFDM signals. *IEEE Transactions on Communications*, 49(2):282–289, 2001.

[33] H. Ochiai and H. Imai. Performance analysis of deliberately clipped OFDM signals. *IEEE Trans. on Communications*, 50(1):89–101, Jan. 2002.

[34] A. V. Oppenheim and R. W. Schafer. *Discrete-Time Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1989.

[35] K. G. Paterson and V. Tarokh. On the existence and construction of good codes with low peak-to-average power ratios. *IEEE Trans. on Info. Theory*, 46(6):1974–87, Sept. 2000.

[36] T. Pollet, M. V. Bladel, and M. Moeneclaey. BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise. *IEEE Trans. on Communications*, 43(234):191–3, Feb/Mar/Apr 1995.

[37] A. Redfern. Receiver window design for multicarrier communication systems. *IEEE Journal on Sel. Areas in Communications*, 20(5):1029–36, June 2002.

[38] G. Santella and F. Mazzenga. A hybrid analytical-simulation procedure for performance evaluation in M-QAM-OFDM schemes in presence of nonlinear distortions. *IEEE Trans. on Veh. Technology*, 47(1):142–151, Feb. 1998.

[39] T. M. Schmidl and D. C. Cox. Robust frequency and timing synchronization for OFDM. *IEEE Trans. on Communications*, 45(12):1613 – 21, Dec. 1997.

[40] J. Tellado. *Multicarrier Modulation with low PAR: Applications to DSL and wireless*. Kluwer Academic Publishers, Boston, 2000.

[41] J. Tellado, L. Hoo, and J. Cioffi. Maximum-likelihood detection of nonlinearly distorted multicarrier symbols by iterative decoding. *IEEE Trans. on Communications*, 51(2):218– 228, Feb. 2003.

[42] J. van de Beek, M. Sandell, and P. Borjesson. ML estimation of time and frequency offset in OFDM systems. *IEEE Trans. on Signal Processing*, 45:1800–05, July 1997.

[43] R. van Nee and A. de Wild. Reducing the peak-to-average power ratio of OFDM. In *Vehicular Technology Conference, 1998. VTC 98. 48th IEEE*, volume 3, pages 2072–2076, 1998.

[44] S. Weinstein and P. Ebert. Data transmission by frequency-division multiplexing using the discrete fourier transform. *IEEE Trans. on Communications*, 19(5):628–34, Oct. 1971.