On the Shannon Theorem and noise spheres in n-dimensional space.

Akhila Raman

1. Introduction

It is well known that Shannon's Capacity Limit for a communication channel in the presence of Additive White Gaussian Noise(AWGN) is given by $C = W\log_2(1+\frac{P}{N})$ where C is the Channel Capacity in bits per second, W is the channel bandwidth in Hz, P is the average signal power and N is the average noise power.[1] [Shannon's 1949 paper is reproduced here.]

• Shannon used the idea of n-dimensional sphere(**n-sphere**) of radius $Rs = \sqrt{2WTN}$ for noise, where $n = 2WT$ refers to the total number of signal samples in duration T and T is the duration of the signal in seconds. Let us call this radius $Rs = \sqrt{2WTN}$ as **Shannon Radius** of noise n-sphere.[We can either use a n-sphere is a (n+1) dimensional Euclidean space or (n-1) sphere in a n-dimensional Euclidean space.] [In this paper, we use n for number of dimensions $N_d = n = 2WT$. We use the notation (n', k') for error correcting code. $n' = n$ for binary signalling $M = 2$ which assigns bits 0 and 1 to symbols -1 and $+1$. We scale average noise power $N = 1$.]

• **Shannon's Volume Argument**

He used a larger n-sphere of radius $R = \sqrt{2WT(P+N)}$ for received signal and asked this question: "How many different transmitted signals can be found which will be distinguishable?". He stated the answer "Certainly not more than the volume of the sphere of radius $R = \sqrt{2WT(P+N)}$ divided by the volume of a sphere of radius $Rs = \sqrt{2WTN}$, since **overlap of the noise spheres** results in **confusion** as to the message at the receiving point". [Page 7 of Shannon's 1949 paper is reproduced here.] Thus he derived the maximum number of distinguishable transmitted signals as $M \leq (\frac{R}{Rs})^n = (\frac{\sqrt{2WT(P+N)}}{\sqrt{2WTN}})^{2WT} = (1+\frac{P}{N})^{WT}$ and hence he derived the channel capacity as $C = \frac{\log_2(M)}{T} = W\log_2(1+\frac{P}{N})$. [2]

• **Noise confinement argument**

Shannon implied that, as $n = 2WT \to \infty$, noise samples will become fully **confined to** Shannon sphere of radius $Rs = \sqrt{2WTN}$. [He said that, "for large T, perturbation(noise) will almost certainly be to some point **near the surface of a sphere** of radius $\sqrt{2WTN}$ centered at the original signal point. More precisely, by taking sufficiently large T we can insure (with probability as near to one as we wish) that the perturbation will lie **within** a sphere of radius $\sqrt{2WT(N+\epsilon)}$ where ϵ is arbitrarily small. The noise regions can therefore be thought of roughly as **sharply**

defined billiard balls, when $2WT$ is very large."][2] [Page 7 of Shannon's 1949 paper is reproduced here.]

Let us follow these lines of inquiry:

Section A: We will examine the Packing Density of spheres in n-dimensional space and the gaps between the spheres and examine if this can reduce the as $n = 2WT \to \infty$. We will show that practically realizable Limit is less than that specified by Shannon's Limit, which remains the upper bound.

Section B: For the **specific** case of Orthogonal Frequency Division Multiplexing(**OFDM**) with spectral efficiency $\frac{C}{W} = 2$, let us examine noise spheres in n-dimensional Euclidean space and show that they **do indeed** overlap, as $n = 2WT \to \infty$. Let us add Gallager's Low Density Parity Check (LDPC) codes (n', k') and show that we can have $2^{k'}$ transmitted signals with non-overlapping noise spheres but code rate $R = \frac{k'}{n'}$ is less than 1 and we need to supply more power than required by Shannon's expression.

Section C: Modern OFDM systems with LDPC codes **do not** typically compare received signal with $2^{k'}$ transmitted signals(soft decision decoding) and hence do not use the concept of non-overlapping noise spheres, due to computational complexity. Instead, they perform symbol by symbol detection, which involves decoding the I-Q symbol in each complex DFT coefficient to obtain the bits, followed by LDPC decoding performed on hard decision bits.

We will derive the **Modified Limit** for OFDM system with **symbol by symbol** detection which uses Gallager's LDPC codes and show that it performs **better** than the system in Section B (which performs soft decision detection in n-dimensional space and needs non-overlapping noise spheres). Hence it will be argued that Shannon's expression derived using noise spheres is **different** from the expression derived using symbol by symbol detection, in the specific case of this OFDM system with LDPC.

Section D: We will examine the **Modified Limit** for other LDPC implementations.

Note that we have chosen Gallager's LDPC codes because they have the desirable property that, as $n' \to \infty$, keeping $\frac{k'}{n'}$ constant, $\frac{d_{min}}{n'}$ and hence error correction rate also remain constant, where d_{min} is the minimum distance of that code.

2. Section A

• In Shannon's theorem, each transmitted signal point is located on the surface of a (n-1) dimensional sphere of radius $R_p = \sqrt{2WT * P}$, in a n-dimensional Euclidean space, which is surrounded by Shannon's noise sphere of radius $R_s = \sqrt{2WT * N}$ in n-dimensional space and the received signal point is located on the surface of a (n-1) dimensional sphere of radius $R = \sqrt{2WT * (P + N)}$. The total number of samples is given by $n = 2WT$ and is a common scale factor in the radius of both signal and noise spheres.

• **Packing Density in n-spheres**

The problem of packing density in (n-1) dimensional hyperspheres is well studied. For example, in **3-dimensional** space, for **equal radius** spheres, the densest packing uses approximately $\pi\frac{\sqrt{2}}{6} = 74\%$ of the volume.[Reference in Wikipedia]. This Packing Density Reduction Factor(**PDRF**)of 74% is due to the **gaps** between the equal radius spheres. In our case, Shannon's noise spheres are equal radius spheres which surround each transmitted signal point and reside within the larger volume sphere of radius R and we **cannot fill the gaps** between the noise spheres and we must expect reduction in due to the gaps.[Shannon's Volume Argument discussed before, assumes that the entire volume is packed tight with equal radius noise spheres, **without accounting for gaps** between the noise spheres.]

As the number of dimensions n increases, Packing Density Reduction Factor(**PDRF**) also **reduces further**, as per table below. [Hypersphere Packing].

Number of Dimensions n	Packing density δ_n
2	$\pi\frac{\sqrt{3}}{6} = 0.9069$
3	$\pi\frac{\sqrt{2}}{6} = 0.74048$
4	$\frac{\pi^2}{16} = 0.61685$
5	$\pi^2\frac{\sqrt{2}}{30} = 0.46526$
6	$\pi^3\frac{\sqrt{3}}{144} = 0.37295$
7	$\frac{\pi^3}{105} = 0.2953$
8	$\frac{\pi^4}{384} = 0.25367$

• It is clear that Packing Density reduces, as number of dimensions n increases. This is also supported by a related problem known as "**Kissing Number Problem**", which is defined as the number of non-overlapping unit spheres((n-1) dimensional spheres) that can be arranged such that they each touch another given unit sphere (in n- dimensional Euclidean space) .[Kissing Number Problem].

In 2-dimensions, the kissing number is 6, as shown in figure below.

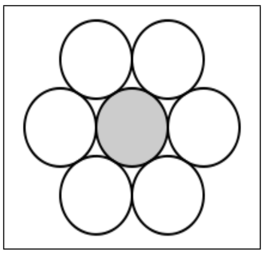

Figure 1:

• In the case of our Shannon's noise spheres of radius $R_s = \sqrt{2WT * N}$ packed inside the larger received signal sphere of radius $R = \sqrt{2WT * (P + N)}$, for any given choice of $n = 2WT$, the **common scale factor** $2WT$ between the smaller and larger spheres can be **omitted** and for our choice of $N = 1$, Shannon's noise spheres become unit radius spheres($R_s = 1$). The larger sphere has radius $R = 3$ which means $P = 8$. This corresponds to channel spectral efficiency $\frac{C}{W} = log_2(1 + \frac{P}{N}) = 3.1699$.[We choose $R = 3$ and $R_s = 1$ to make an analogy to the kissing number problem]

• Packing density is given by the Ratio of total area of the $6 + 1 = 7$ smaller spheres of radius R_s to the area of larger sphere of radius R (for our Shannon's spheres example) is given by

$$\frac{7\pi R_s^2}{\pi R^2} = \frac{7N}{(P + N)} = 0.77778 \tag{1}$$

• As the number of dimensions n increases, it can be shown that Packing density reduces further. Packing density in n-dimensions is given by the Ratio of total n-volume of the $(N_k + 1)$ smaller spheres of radius R_s to the n-volume of larger sphere of radius R where N_k is the kissing number in n-dimensions.

Packing density for our Shannon's spheres example is given by the following equation.

$$\delta_n = (N_k + 1) * \frac{V_n(R_s)}{V_n(R)} = (N_k + 1) * \frac{(\frac{\pi^{\frac{n}{2}} R_s^n}{\Gamma(\frac{n}{2}+1)})}{(\frac{\pi^{\frac{n}{2}} R^n}{\Gamma(\frac{n}{2}+1)})} = (N_k + 1) * (\frac{R_s}{R})^n \tag{2}$$

For our choice of $N = 1, P = 8$, we get $\delta_n = (N_k + 1) * (\sqrt{\frac{N}{(P+N)}})^n = (N_k + 1) * (\frac{1}{3})^n$.

This is calculated in the table below and it is clear that, as the number of dimensions n increases, Packing density reduces further, of order of $(\frac{1}{K})^n$ where $K >= 1.5$. [Kissing Number in n-dimensions]

Number of Dimensions n	Kissing Number Upper Bound N_k	Packing density δ_n
2	6	0.77778
3	12	0.48148
4	24	0.30864
5	44	0.18519
6	78	0.10837
7	134	0.061728
8	240	0.036732
16	7355	$1.7088e - 4$
24	196560	$6.9596e - 7$

• **Packing Density Reduction in the asymptotic case $n \to \infty$**

For the general case of n, as the number of dimensions $n \to \infty$, **Grigorii Kabatiansky and Vladimir Levenshtein(1978)** have shown that the **upper bounds for packing density is given**

by $2^{-0.599*n}$. [Kabatiansky, G. A.; Levenshtein, V. I. (1978), "Bounds for packings on a sphere and in space", Problemy Peredachi Informatsii 14: 3–25.] ["Sphere packing bounds via spherical codes" Henry Cohn, Yufei Zhao, Duke Math. J. 163, no. 10 (2014), 1965-2002. Page 2, Eq. 1.1. Link to Paper]. This result applies for all general sphere packing problems including lattice packing.

This implies that Shannon's $C = Wlog_2(1 + \frac{P}{N})$ will be reduced by the packing density factor as follows.

$$\begin{aligned} C = Wlog_2(1+\frac{P}{N}) &= \lim_{T\to\infty} \frac{2WT}{T} log_2(\sqrt{(1+\frac{P}{N})}) \\ &= \lim_{T\to\infty} \frac{1}{T} log_2(\sqrt{(1+\frac{P}{N})})^n \\ C_{reduced} &= \lim_{T\to\infty} \frac{1}{T} log_2[(\sqrt{(1+\frac{P}{N})})^n * 2^{-0.599*n}] \\ C_{reduced} &= C - 0.599 * 2W = C - 1.2 * W \\ \frac{C_{reduced}}{W} &= \frac{C}{W} - 1.2 \end{aligned} \tag{3}$$

We should expect a reduction in by 1.2, due to packing density limitations, provided Kabatiansky and Levenshtein's result holds.

3. Section B

Let us examine the specific case of Orthogonal Frequency Division Multiplexing(OFDM) and analyze the noise sphere overlapping. OFDM signal is expressed as a $N_f = n$-point DFT given by

$$s[m] = \frac{1}{n} \sum_{k=0}^{n-1} S[k] e^{i\frac{2\pi}{n}k*m} \tag{4}$$

where $S[k] = I[k] + iQ[k]$ are the complex DFT/FFT coefficients which carry information symbols.[Appendix A]

Example 1:

• Let the spectral efficiency factor $\frac{C}{W} = 2$ and let us assume unit noise power $N = 1$. Then $(1 + \frac{P}{N}) = 2^2 = 4$. Then we get $P = 3$. This means we need to use a scale factor of $\sqrt{3}$ in the FFT domain for OFDM constellation points.

• In OFDM block of duration T, given a signal bandwidth W, sampling rate(number of samples/second) is $2W$ and number of samples per T seconds is given by the number of FFT samples in a single OFDM block $N_f = 2WT = n$ where n is the number of dimensions. In OFDM, information carrying symbols are in the frequency domain. Both signal and noise samples are scaled by a factor of $\sqrt{\frac{n}{2}} = \sqrt{WT}$ in the frequency domain, as shown in Appendix A [5].

• This spectral efficiency $\frac{C}{W} = 2$ can be achieved by **antipodal signalling** by loading one bit represented by a symbol ± 1 on each frequency domain(FFT) sample, with scaling factor $\sqrt{3}$. [Each FFT sample carries 1 bit,$M = 2$, total number of bits in an OFDM block of duration T seconds equals total number of FFT samples in the OFDM block and is given by $n = 2WT$, hence total number of bits/second is given by $C = 2W$ which means $\frac{C}{W} = 2$.]

• We can represent transmitted signal points in a single OFDM block of length $N_f = n = 2WT$ samples by geometric representation in a n-dimensional space. Each transmitted signal point in FFT domain is represented by a n-tuple. For example, if the size of OFDM block is $n = 4$ samples, if we use binary signalling on each FFT sample with $M = 2$, each of $2^n = 16$ transmitted signal points is given by $(1, 1, 1, 1), (1, 1, 1, -1), ..., (-1, -1, -1, -1)$ multiplied by scale factor of $\sqrt{3}$, for a given spectral efficiency factor $\frac{C}{W} = 2$ and each of these 16 transmitted signal points are located on the $2^4 = 16$ vertices of a 4-dimensional cube which sits insides a hypersphere in 4-dimensional space with its 16 vertices touching the hypersphere.

• Let us examine Shannon's noise sphere of radius $R_s = \sqrt{2WTN} = \sqrt{nN}$ around each transmitted signal point and see if they overlap. Note that only when $n \to \infty$, all the noise samples will be confined to the surface of this sphere of radius $R_s = \sqrt{nN}$. Which is why, we are examining if noise spheres of this radius R_s begin to overlap for finite n dimensions. If Shannon's noise spheres overlap for finite n, as $n \to \infty$, all noise samples will be confined to the surface of this sphere, while the overlap continues to exist and hence Shannon's full cannot be realized.

In The Figure below,on the left, OFDM constellation for $n = 2WT = 3$ is plotted and we can see that Shannon's noise spheres of radius $\sqrt{2WTN} = \sqrt{nN}$ around each transmitted signal point, do not overlap and just barely touch each other.[Click here to view the plot in full scale]

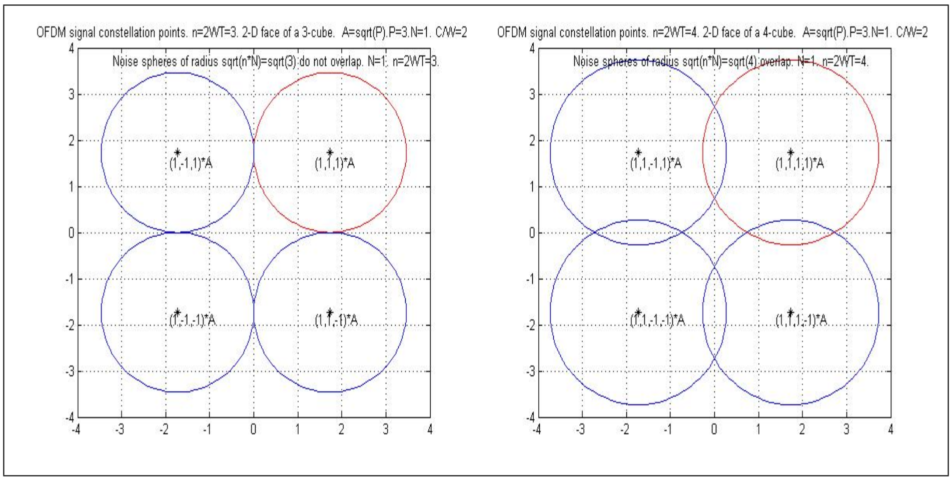

Figure 2:

In The Figure above, on the right, OFDM constellation for $n = 2WT = 4$ is plotted and we can see that Shannon's noise spheres of radius $\sqrt{2WTN} = \sqrt{nN}$ around each transmitted signal point, **do overlap** . [Click here to view the plot in full scale]

For $n >= 4$, Shannon's noise spheres of radius $\sqrt{nN}$ **grow larger and larger and overlap** with increasing number of adjacent transmitted signal points. It is clear that we cannot use all the 2^n transmitted signal points and get the full . Instead we need to use a (n', k') error correcting code with $n' = n$ and use $2^{k'}$ out of $2^{n'}$ transmitted codewords and figure out what reduced can we achieve, if we use the idea of Shannon's n dimensional noise spheres and decode in n dimensional space and find the nearest neighbor for n dimensional received signal. This is explained in Section B.1.

Section B.1:

• Given the fact that Shannon's noise sphere of radius $R_s = \sqrt{nN}$ surrounding each transmitted signal point, starts overlapping with more and more adjacent signal points, as $n \to \infty$, we can calculate the signal power required to keep the $2^{k'}$ spheres non-overlapping, when we use a (n', k') error correcting code with certain $\frac{d_{min}}{n'}$ where $n' = n.(M = 2)$.

If we use an **additional** scale factor B for each transmitted signal point, for example, $(1, 1, ...1) * \sqrt{3} * B$, noise spheres surrounding each transmitted signal point **will not** overlap, if the radius of Shannon's Noise sphere is less than half the Euclidean distance($d_{eucl_{min}}$) between transmitted signal points corresponding to 2 valid codewords separated by the minimum distance d_{min} of that (n', k') error correcting code.[If we use a (n', k') error correcting code on the bits loaded in FFT domain, with certain $\frac{d_{min}}{n'}$, where d_{min} is the minimum distance of that (n', k') code, then the Euclidean distance between any 2 valid codewords has a minimum value of $d_{eucl_{min}} = 2\sqrt{d_{min}}$.]

$$R_s = \sqrt{nN} \leq \frac{d_{eucl_{min}}}{2}$$
$$\sqrt{nN} \leq B\sqrt{3}\sqrt{d_{min}} \tag{5}$$

For Gallager's parity-check codes, $\frac{d_{min}}{n'} = 0.11 = K$ for $\frac{k'}{n'} = \frac{1}{2}$. We assumed average noise power $N = 1$ by convention. For binary signalling $M = 2$ and $n' = n$. Hence

$$\sqrt{n} \leq B\sqrt{3}\sqrt{nK}$$
$$\sqrt{n} \leq \sqrt{3B^2 nK} \tag{6}$$

This implies that $3B^2 = \frac{1}{K} = \frac{1}{0.11}$ and hence $B = 1.74$. Hence we see that

• To achieve $\frac{C}{W} = 2$, Shannon's theorem requires average signal power $P = 3$ for average noise power $N = 1$. This required a scale factor of $\sqrt{3}$ to be applied to the OFDM constellation $(1, 1, ..1)$. We showed that this scale factor of $\sqrt{3}$ was not sufficient to keep the noise spheres non-overlapping.

• We showed that we require an **additional** scale factor of $B = 1.74$ to be applied to the constellation to keep $2^{k'}$ noise spheres non-overlapping when we use (n', k') error correcting code with $\frac{k'}{n'} = \frac{1}{2}$. Effective **source** was reduced by the factor $\frac{k'}{n'} = \frac{1}{2}$. This implies an **additional** average signal power of $B^2 = 3.03$ is required, compared to Shannon's expression, to achieve a **channel** $\frac{C}{W} = 2$ and **source** $\frac{C}{W} = 1$.

• Using the same method used in Section C Eq. 12, we can show that the modified channel capacity for an OFDM system with (n', k') error correcting code, is given by

$$C = W * log_2(1 + 3P_0) = W * log_2(1 + \frac{3P}{A^2}) = W * log_2(1 + \frac{3PK}{N})$$
$$(\frac{C}{W})_{OFDM_{SPHERE}(source)} = log_2(1 + \frac{3PK}{N}) * \frac{k'}{n'} \tag{7}$$

where $P = P_0 A^2$ and A is the overall scale factor applied to the transmitted signal points. In the previous example, this would be $A = B\sqrt{3}$ and $A^2 = \frac{N}{K}$.

• **Note 1: Shannon's theorem uses all of the M^n spheres**

Shannon's Limit $C = Wlog_2(1 + \frac{P}{N})$ can be rewritten as $C = \lim_{T\to\infty} \frac{2WT}{T} log_2 \sqrt{(1 + \frac{P}{N})} = \lim_{T\to\infty} \frac{2WT}{T} log_2 M = \lim_{T\to\infty} \frac{1}{T} log_2[M^n]$ where $M = \sqrt{(1 + \frac{P}{N})}$ is the number of symbols per dimension in a $n = 2WT$ dimensional space. It is clear that M^n is the total number of transmitted signal points in a n dimensional space and Shannon's noise sphere of radius $\sqrt{nN}$ is around each of these transmitted signal points and as $n \to \infty$, AWGN noise samples are confined to the surface of this sphere. Shannon's theorem requires that these M^n noise spheres **do not overlap** (and cause confusion) as $n \to \infty$. This implies that **all** the M^n signal points are transmitted, to realize the full , for the example of $M = 2$ binary signalling in each FFT sample, this means all 2^n signal points or codewords are transmitted and the received signal points can be perfectly decoded given that the noise spheres do not overlap and all the 2^n signal points contribute to the channel capacity and channel bit rate equals the source bit rate.[Or at least $M^n * f(n)$ signal points are transmitted, where $\lim_{T\to\infty} \frac{1}{T} log_2[f(n)] \to 0$].

This is in contrast with modern OFDM systems with (n', k') error correction codes where we transmit only $2^{k'}$ out of $2^{n'}$ codewords and the receiver performs symbol by symbol detection on each FFT sample and to produce the bits and then does error correction decoding and hence the source bit rate is reduced by the factor $\frac{k'}{n'}$.

- **Note 2**

In the OFDM constellation with $M = 2$(binary signalling on each FFT sample), in n dimensional space,with 2^n transmitted signal points represented by
$(1, 1, 1, ...1), 1, 1, ..., -1), ...(-1, -1, ... - 1)$, it is easy to show that
each transmitted signal point has $\binom{n}{r}$ number of signal points separated by Euclidean distance of $d_{eucl} = 2\sqrt{r}$ for $r = 1, 2, ...n$.

- **Note 3**

If we use Gray coding such that any 2 adjacent transmitted signal points differ by only 1 bit [use +1 to represent bit 1 and use −1 to represent bit 0], we can show that Euclidean distance $d_{eucl} = 2\sqrt{d_{bits}}$ [**Appendix B**] where d_{bits} is the number of bits by which any 2 transmitted signal points differ.

- **Note 4: M=4 case** . $M = 4$. see **Appendix B**

4. Section C

Modern OFDM systems with LDPC codes **do not** typically compare received signal with $2^{k'}$ transmitted signals(soft decision decoding) and hence do not use the concept of non-overlapping noise spheres. Instead, they perform symbol by symbol detection, which involves decoding the I-Q symbol in each complex DFT coefficient to obtain the bits, followed by LDPC decoding performed on hard decision bits. We will derive the **Modified Limit** for this system with **symbol by symbol** detection.[In this section, we will decode I and Q separately to obtain bits.]

Let us start with a **Reference System** with transmitted symbols per dimension as follows: $\pm1, \pm3, \pm5, ... \pm (M - 1)$ where M is the number of symbols and $\log_2(M)$ is the number of bits per symbol. Number of dimensions in OFDM is $n = 2WT$ where W is the bandwidth and T is the time duration of OFDM symbol. The symbols in the Reference System are spaced by 2 units. We will apply scale factor A later to these symbols, for a given noise power N, to achieve a specified Signal to Noise Ratio(SNR).

This Reference System is plotted in the **Figure**.[Click here for Figure]

The power P_0 of this Reference System is given by

$$P_0 = \frac{2}{M} \sum_{r=0}^{(\frac{M}{2}-1)} (2r+1)^2 \tag{8}$$

• This **Reference system** has a constellation with M symbols.
The number of bits/symbol = $\log_2(M)$.
Each symbol is loaded on a FFT sample in OFDM and the total number of samples per OFDM block of duration T seconds = $n = 2WT$ samples.
The total number of bits per OFDM block = $n \log_2(M)$.
The total number of bits per second $C = \frac{n}{T} \log_2(M)$.
Spectral Efficiency of this OFDM system is given by $(\frac{C}{W})_{OFDM} = \frac{n}{WT} \log_2(M) = 2\log_2(M)$.

• We can show that $(\frac{C}{W})_{OFDM} = 2\log_2(M) = \log_2(1+3P_0)$. First we show that $M^2 = 1+3P_0$.

$$\begin{aligned}
P_0 = \frac{2}{M}[1^2+3^2+...(M-1)^2] = \frac{2}{M}[(1^2+2^2+3^2+...(M)^2)-(2^2+4^2+...(M)^2)] \\
= \frac{2}{M}[(1^2+2^2+3^2+...(M)^2)-2^2(1^2+2^2+...(\frac{M}{2})^2)] \\
= \frac{2}{M}[(\frac{M*(M+1)*(2M+1)}{6})-4(\frac{\frac{M}{2}*(\frac{M}{2}+1)*(M+1)}{6})] \\
= \frac{2}{M}[(\frac{M*(M+1)}{6})*(2M+1-2*(\frac{M}{2}+1))] \\
= \frac{2}{M}[(\frac{M*(M+1)}{6})*(M-1)] = \frac{2}{M}(\frac{M*(M^2-1)}{6}) \\
1+3P_0 = 1+\frac{6}{M}(\frac{M*(M^2-1)}{6}) \\
1+3P_0 = M^2 \\
(\frac{C}{W})_{OFDM} = 2\log_2(M) = \log_2(1+3P_0)
\end{aligned} \tag{9}$$

For example,

[1] **M=2** : Binary antipodal signalling with symbols ± 1.

$P_0 = \frac{(1^2+(-1)^2)}{2} = 1$.
$2\log_2(M) = 2$.

$\log_2(1+3P_0) = \log_2(4) = 2$.

[2] **M=4** : 4 symbols $\pm 1, \pm 3$.

$P_0 = \frac{(1^2+(3)^2)}{4} + \frac{((-1)^2+(-3)^2)}{4} = 5$.
$2\log_2(M) = 4$.
$\log_2(1+3P_0) = \log_2(16) = 4$.

Number of Symbols M	Spectral Efficiency $(\frac{C}{W})_{OFDM} = 2\log_2(M)$	$P_0 = \frac{2}{M}\sum_{r=0}^{(\frac{M}{2}-1)}(2r+1)^2$	$\log_2(1+3P_0)$
2	2	1	2
4	4	5	4
8	6	21	6
16	8	85	8

From the above table, we can see that columns 2 and 4 match. We can rewrite the channel spectral efficiency of this Reference system as follows.

$$(\frac{C}{W})_{OFDM} = \log_2(1+3P_0) \tag{10}$$

Given that both the signal power and noise power are proportional to signal bandwidth W, given N_0 is the noise spectral density, we can write the total signal power P and total noise power N as follows. A is the scale factor applied to the reference system, to achieve specified SNR, which will scale proportional to the bandwidth.

$$\begin{aligned} P &= P_0 A^2 \\ N &= N_0 W \end{aligned} \tag{11}$$

We can rewrite the channel spectral efficiency of the actual system which uses symbol by symbol detection as follows.

$$(\frac{C}{W})_{OFDM} = \log_2(1+3P_0) = \log_2(1+\frac{3P}{A^2}) \tag{12}$$

We wish to find an expression for constellation scale factor A in terms of noise power N such that, as number of dimensions(samples) $n \to \infty$, probability of bit error(BER) $\to 0$ and above can be achieved. This is done in Section C.1

Section C.1

Let us consider the above OFDM system and use an error-correction code. For example, let us use Gallager's Low Density Parity Check(LDPC) codes[3] (n', j', k') with minimum distance ratio $\delta_0 = \frac{d_{min}}{n'}$. $H(\lambda) = -\lambda \ln(\lambda) - (1-\lambda)\ln(1-\lambda)$ and $H(\delta_0) = (1-R)\ln(2); R = \frac{k'}{n'}$ and δ_0 corresponds to the equiprobable ensemble of parity check codes and code rate R Vs δ_0 is reproduced in the figure below for convenience. [Click here to view the plot in full scale]

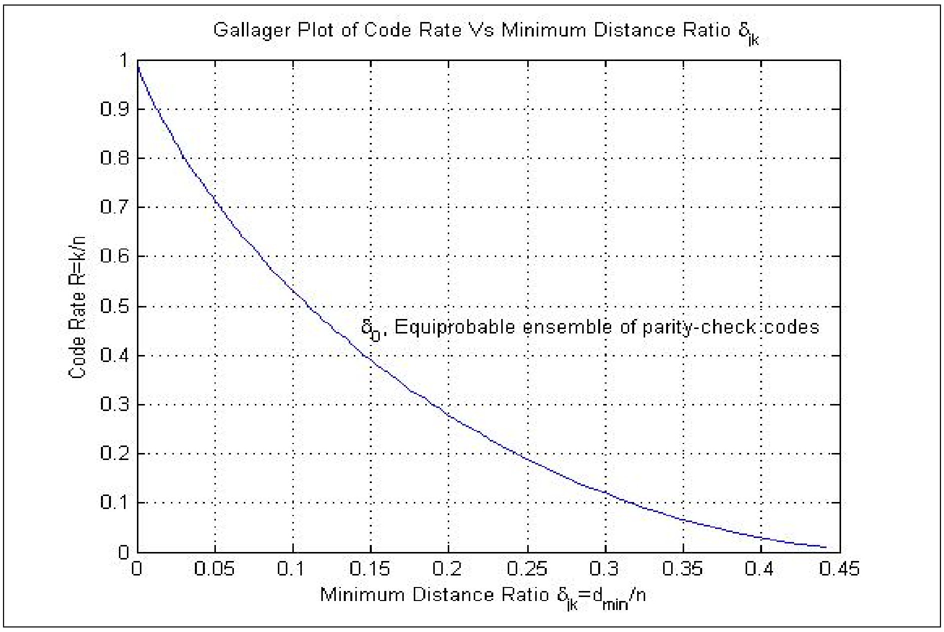

Figure 3:

For example, from the figure above, we can see that, for an example code rate $R = \frac{k'}{n'} = \frac{1}{2}$, minimum distance ratio $\delta_0 = \frac{d_{min}}{n'} = 0.11$ as $n' \to \infty$. Given that $d_{min} = 2t + 1$, where t is the number of corrected bit errors, error correction rate ECR is given by $\lim_{n' \to \infty} \frac{t}{n'} = \frac{d_{min}}{2n'} - \frac{1}{n'} = \lim_{n' \to \infty} \frac{d_{min}}{2n'} = 0.055$.

We require the uncorrected **raw BER** of our OFDM system, which uses this Gallager LDPC codes, to be **less than** error correction rate **ECR**. That would ensure that as $n' \to \infty$, $BER \to 0$.

Example 1:

- **M=2**: Each FFT sample in OFDM block has 2 symbols ± 1.

The Raw BER of this binary signalling system is given by the well known expression
$P_e = \frac{1}{2} erfc(\sqrt{\frac{E_b}{N_0}}) = \frac{1}{2} erfc(\sqrt{\frac{E_{b(BPSK)}}{N_0}})$
where $E_b = E_{b(BPSK)}$ is the Energy Per Bit for a Binary Signalling System and N_0 is the Noise Spectral Density.

Given that the average signal power is given by $P = \frac{E_b}{T_b} = \frac{E_{b(BPSK)}}{T_b} = P_{BPSK}$ and average noise power is given by $N = N_0 W$, we can rewrite the BER as
$P_e = \frac{1}{2}erfc(\sqrt{\frac{P_{BPSK}}{N}WT_b})$.

For this OFDM system of block duration T with binary signalling, $T_b = \frac{T}{n} = T_s = \frac{1}{f_s}$ where $f_s = 2W$ is the sampling frequency. Hence we can write the BER of this OFDM system with Binary Signalling as follows.

$$P_e = \frac{1}{2}erfc(\sqrt{\frac{P_{BPSK}}{2N}}) \tag{13}$$

For Binary signalling $P_{BPSK} = P_0 A^2 = A^2$ given that $P_0 = 1$. Hence $P_e = \frac{1}{2}erfc(\sqrt{\frac{A^2}{2N}})$

We require this Raw BER to be less than Error Correction Rate(ECR) of the associated Gallager's LDPC code.

$$P_e = \frac{1}{2}erfc(\sqrt{\frac{A^2}{2N}}) \leq \lim_{n' \to \infty} \frac{d_{min}}{2n'} \tag{14}$$

The minimum value of $\frac{E_b}{N_0}$ which satisfied above inequality is given by $(\frac{E_b}{N_0})_{min} = \frac{A^2}{2N}$.

Now we can rewrite Eq. 12 as follows.

$$(\frac{C}{W})_{OFDM} = \log_2(1 + \frac{3P}{A^2}) = \log_2(1 + \frac{3P}{2N(\frac{E_b}{N_0})_{min}}) \tag{15}$$

Note that $(\frac{E_b}{N_0})_{min}$ is a function of M and $\frac{d_{min}}{n'}$.

For $\frac{d_{min}}{2n'} = 0.055$ and code rate $\frac{k'}{n'} = 0.5$, $(\frac{E_b}{N_0})_{min} = 1.28$ and $M = 2$

$$(\frac{C}{W})_{OFDM} = \log_2(1 + \frac{3P}{2N * 1.28}) = \log_2(1 + \frac{P * 1.171875}{N}) \tag{16}$$

This seems to perform better than Shannon's Limit which uses n-dimensional noise spheres, $\frac{C}{W} = log_2(1+\frac{P}{N})$, but we should remember that $\frac{C}{W}$ is the **source** spectral efficiency while $(\frac{C}{W})_{OFDM}$ is the **channel** spectral efficiency and the **source spectral efficiency** is given by $(\frac{C}{W})_{OFDM(source)} = \frac{k}{n}\log_2(1 + \frac{P*1.171875}{N})$.

$$(\frac{C}{W})_{OFDM_{SYM}(source)} = \frac{k'}{n'}\log_2(1+\frac{P*1.171875}{N})$$

(17)

• **Example 2**:

M=4: Each FFT sample in OFDM block has 4 symbols $\pm1, \pm3$. See Appendix C.1. The expression for BER for general case of M is derived in Appendix C. For each case of M, we can deduce the value of A which satisfies the expression $BER = 0.055$(If we use Gallager's LDPC codes with $\frac{k'}{n'} = \frac{1}{2}$).

• **Section C.3**

We can compare the expressions for **source** spectral efficiency for OFDM (adjusting for $\frac{k}{n} = \frac{1}{2}$) in Eq. 17 with the spectral efficiency computed by Shannon's original expression.

• Table below lists $\frac{E_b}{N_0}$ Vs $\frac{C}{W}$. For Shannon's expression, source $\frac{C}{W}$ and channel $\frac{C}{W}$ are the same. For our OFDM expression, they are different and this explains the first 2 columns.

• we can see that, in order to achieve the same **source** capacity per unit bandwidth, we need **higher** $\frac{E_b}{N_0}$ with our expression for OFDM, compared to Shannon's expression, for the specific case of Gallager's LDPC codes with $\frac{d_{min}}{n} = 0.11$ for $\frac{k}{n} = \frac{1}{2}$.

• The fourth column corresponds to $\frac{E_b}{N_0}$ obtained for OFDM with Gallager's parity-check codes, with symbol by symbol detection as in Eq. 15 [$(\frac{C}{W})_{OFDM_{SYM}(source)} = \frac{k}{n}\log_2(1+\frac{3P}{2N*(\frac{E_b}{N_0})_{min}})$] and adjusts $\frac{E_b}{N_0} = \frac{P}{N}\frac{1}{\frac{C}{W}\frac{k}{n}}$ for code rate $\frac{k}{n} = \frac{1}{2}$ and corresponds to source bit rate per unit bandwidth.

• The last column corresponds to $\frac{E_b}{N_0}$ obtained for OFDM with Gallager's parity-check codes, with detection based on noise spheres, as per Eq. 7 [For $M = 2$, $(\frac{C}{W})_{OFDM_{SPHERE}(source)} = log_2(1+\frac{3P*0.11}{N}) * \frac{k'}{n'}$].

$Channel\frac{C}{W}$	$Source\frac{C}{W}$	$\frac{E_b}{N_0}_{(shannon)}$ in dB	$\frac{E_b}{N_0}_{(OFDM_{SYM})}$ (source)	$\frac{E_b}{N_0}_{(OFDM_{SPHERE})}$ (source)
2	1	0	4.08	9.5861
4	2	1.77	7.19	13.576

• It is clear from above table that Symbol by Symbol detection of OFDM system with binary signalling $M = 2$ and Gallager's parity-check codes $\frac{d_{min}}{n'} = 0.11$, **performs better** than a corresponding system which uses soft decision detection based on noise spheres in n-dimensions(choose the nearest neighbour to a received signal point in n-dimensional Euclidean space). The two detection systems are **NOT** equivalent.

• In Section D, we will show that, with modern LDPC codes with higher $\frac{d_{min}}{n}$ or $\frac{d_{min_{average}}}{n}$ where $d_{min_{average}}$ is the average minimum distance between codewords, we can get higher $\frac{C}{W}$ with lower $\frac{E_b}{N_0}$ than the fourth column.

5. Section D

• Modern LDPC codes may have some codewords which are closer in terms of number of bits they differ, but a much higher average minimum distance $d_{min_{average}}$ between codewords, hence we will consider $\frac{d_{min_{average}}}{n'}$. Average minimum distance is given by $d_{min_{average}} = \frac{1}{2^{k'}} \sum_{m=1}^{2^{k'}} d_{min}(m)$ where $d_{min}(m)$ is the minimum distance between codeword m and all other codewords, minimum number of bits by which codeword m differs from other codewords.

For example, consider a code with $n' = 4, k' = 2$ which has $2^{k'} = 4$ codewords. Let us choose the codewords as $(1,1,1,1),(1,1,-1,1),(-1,-1,-1,-1),(-1,-1,1,1))$. The first 2 codewords have a $d_{min_1} = 1$ but the rest of 2 codewords have a $d_{min_2} = 2$, hence the average inter-codeword minimum distance $d_{min_{average}}$ is closer to 1.5.

$$d_{min_{average}} = \frac{1}{4} * 1 + \frac{1}{4} * 1 + \frac{1}{4} * 2 + \frac{1}{4} * 2 = 1.5 \tag{18}$$

• Radford Neal's (10000,5000) LDPC implementation in C language[4] produces $BER = 0$ for $\frac{E_b}{N_0} = 1.41$ dB (signal amplitude $A = 1$, noise standard deviation $\sqrt{N} = 0.85$) while it produces $BER = 2.438e-02$ for $\frac{E_b}{N_0} = 0.92$ dB (($A = 1$, $\sqrt{N} = 0.9$)). [LDPC implementation].

We can use this code with OFDM with $M = 2$ signalling and $\frac{C}{W} = 2$. BER is given by $BER = 0.5 * erfc(\sqrt{\frac{A^2}{2N}}) = 0.1197$ [Note D.1]. We require this $BER < \frac{t}{n'}$ of the code where $\frac{t}{n'}$ is the **average** error correction rate.

This implies that this code has a $\frac{d_{min_{average}}}{n'} = \frac{(2t+1)}{n'} = 0.24$, when $n' = 10000; k' = 5000$, where $d_{min_{average}} = 2t + 1$ is the average minimum distance between codewords.

• We can compute the $\frac{d_{min_{average}}}{n'}$ required for a future error correcting code which will produce the **same** $(\frac{C}{W})_{OFDM_{SYM}(source)} = \frac{k}{n} \log_2(1 + \frac{3P}{2N(\frac{E_b}{N_0})_{min}})$ for $\frac{k}{n} = \frac{1}{2}$, as that of $(\frac{C}{W})_{shannon} = \log_2(1 + \frac{P}{N})$, for the same source $\frac{C}{W} = 1$ and same $P = 1$. We require $(\frac{E_b}{N_0})_{min} = 0.5$ and BER is given by $BER = 0.5 * erfc(\sqrt{(\frac{E_b}{N_0})_{min}}) = 0.15866$.

We require this BER to be less than error correction rate $\frac{t}{n'} = \frac{(d_{min_{average}} - 1)}{2n'}$ of the future error correcting code as $n \rightarrow \infty$. We require $\frac{d_{min_{average}}}{n'} \geq 0.15866 * 2 = 0.31731$ for the future error correcting code with $\frac{k}{n} = \frac{1}{2}$, which is combined with an OFDM system which uses symbol by symbol detection [System 1], which produces the same source $\frac{C}{W} = 1$ and same average signal power $P = 1$ as compared to Shannon's $\frac{C}{W}$ expression which uses non-overlapping noise spheres in n-dimensional space for detection. [System 2].

• We have already shown that the 2 systems are different and **can produce different results** because they use **different methods of detection**. It may be possible for our System 1 which uses symbol by symbol detection, when combined with a future error correcting code with $0.31731 < \frac{d_{min_{average}}}{n'} < 1$ and with $\frac{k}{n} = \frac{1}{2}$, to exceed the performance of Shannon's expression for $\frac{C}{W} = 1$. This **will not violate** Shannon's Capacity expression because Shannon's Capacity expression is valid for a receiver which uses soft decision detection in n-dimensions using non-overlapping noise spheres and is not valid for our System which uses symbol by symbol detection, followed by error correction decoding on hard bits. This is represented in the Figure below.

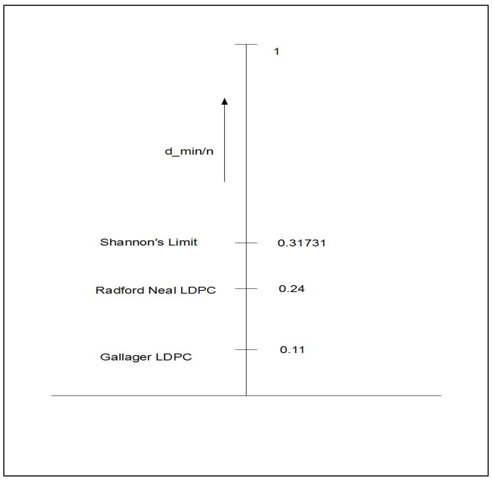

Figure 4:

- **Note D.1**

In numerical simulations in Matlab or C, when we generate a Gaussian Random Variable(RV) of standard deviation σ to simulate zero mean noise[Matlab: w_m=randn(1,N_t)*σ], the average power of noise $w[m]$ is given by $N = \frac{1}{N_t}\sum_{m=1}^{N_t} w[m]^2$ and the standard deviation of noise $w[m]$ is given by $\sigma = \sqrt{\frac{1}{N_t}\sum_{m=1}^{N_t} w[m]^2}$. It is clear that average noise power N always equals square of standard deviation of noise $N = \sigma^2$ in numerical simulations. We know that $N = N_0W; \sigma^2 = \frac{N_0}{2}$. This implies that $N_0W = \frac{N_0}{2}$, which implies that $W = \frac{f_s}{2} = \frac{1}{2}$ and $f_s = 1$ in numerical simulations, where W is the signal bandwidth and $f_s = 2W$ is the sampling frequency.

6. Conclusion

In this paper, we have examined the capacity expression for OFDM system combined with Gallager's LDPC codes, which uses symbol by symbol detection.

7. References

[1] Claude E. Shannon. Communication in the Presence of Noise. Proceedings of the IRE, vol. 37, no. 1, pp. 10–21, Jan. 1949. Reprinted in PROCEEDINGS OF THE IEEE, VOL. 86, NO. 2, FEBRUARY 1998. pp.447-457

`http://web.stanford.edu/class/ee104/shannonpaper.pdf` Click here for Shannon 1949 Paper

[2] Claude E. Shannon. Communication in the Presence of Noise. p.453 [Reprinted in PROCEEDINGS OF THE IEEE, VOL. 86, NO. 2, FEBRUARY 1998. pp.447-457]

[3] Robert Gallager. Low Density Parity Check Codes. 1963. Click here for Gallager's 1963 book

[4] Radford Neal's LDPC implementation

NOISE STANDARD DEVIATION 0.85, Eb/N0 = 1.41 dB
Bit error rate (on message bits only): 0.000e+00

NOISE STANDARD DEVIATION 0.90, Eb/N0 = 0.92 dB
Bit error rate (on message bits only): 2.438e-02

[5] THE HISTORY OF ORTHOGONAL FREQUENCY-DIVISION MULTIPLEXING, STEPHEN B. WEINSTEIN, IEEE Communications Magazine, November 2009 History of OFDM .

Chang, R. W. (1966). "Synthesis of band-limited orthogonal signals for multi-channel data transmission". Bell System Technical Journal 45 (10): 1775–1796.

8. Appendix A

Let us consider an Orthogonal Frequency Division Multiplexing(OFDM) signal expressed as a $N_{fft} = N_f$ -point DFT given by

$$s[m] = \frac{1}{N_f} \sum_{k=0}^{N_f-1} S[k] e^{i\frac{2\pi}{N_f} k*m} \tag{19}$$

where $S[k] = I[k] + iQ[k]$ is complex in general. Let us consider specific case of $s[m]$ as a real signal. This implies that the real part of $S[k]$ has even symmetry and the imaginary part of $S[k]$ has odd symmetry. For real $s[m]$, $Q[0] = 0$ and $Q[\frac{N_f}{2}] = 0$ are not used. Hence we can write

$$s[m] = \frac{2}{N_f} \sum_{k=1}^{\frac{N_f}{2}-1} [I[k] \cos(\frac{2\pi}{N_f} k * m) - Q[k] \sin(\frac{2\pi}{N_f} k * m)] + \frac{1}{N_f} I[0] + \frac{1}{N_f} I[\frac{N_f}{2}]$$

(20)

In the context of Shannon's capacity theorem, Shannon used $n = 2WT$ dimensional space. This corresponds to $n = N_f = 2WT$ in our example of N_f -point DFT. The signal $s[m]$ has a duration of T seconds and bandwidth W and is sampled at the Nyquist rate of $2W$ samples per second. Hence the total number of samples N_f in duration T equals $n = N_f = \frac{T}{\frac{1}{2W}} = 2WT$. [As an example, let us consider a speech signal $s[m]$ of duration $T = 1$ msec and bandwidth $W = 4$ KHz sampled at the rate of $2W = 8$ KHz. Hence the total number of samples N_{fft} in duration T equals $N_f = \frac{T}{\frac{1}{2W}} = 2WT = 8000 * 10^{-3} = 8$ which gives a 8-point DFT.] Replacing $N_{(f)}$ by n in Eq. (20) to compare with Shannon's notations,

$$s[m] = \frac{2}{n}\sum_{k=1}^{\frac{n}{2}-1}[I[k]\cos(\frac{2\pi}{n}k*m) - Q[k]\sin(\frac{2\pi}{n}k*m)] + \frac{1}{n}I[0] + \frac{1}{n}I[\frac{n}{2}]$$

(21)

Bit Loading

Now we can take n bits and load $(\frac{n}{2}+1)$ bits in the frequency domain, onto $I[k]$ and $(\frac{n}{2}-1)$ bits onto $Q[k]$, using binary antipodal signalling[Bit 1 is coded as $+A$ and Bit 0 is coded as $-A$. $M = 2$.].

If we wish to achieve a $\frac{C}{W} = 2$ in Shannon's Capacity equation $C = W\log_2(1+\frac{P}{N})$, for average noise power $N = 1$(AWGN noise standard deviation of 1), this gives average signal power $P = 3$.

It is easy to show that the average signal power of $s[m]$ is $P = \frac{1}{n}\sum_{m=0}^{n-1}s^2[m]$ and $P = 3$ if $A = \sqrt{\frac{3n}{2}}$. Given that the summation of the cosines and sines over the OFDM symbol period T is zero, we have

$$P = \frac{1}{n}\sum_{m=0}^{n-1}s^2[m] = \frac{1}{n}\sum_{m=0}^{n-1}\frac{4}{n^2}(\sum_{k=1}^{\frac{n}{2}-1}[I[k]\cos(\frac{2\pi}{n}k*m) + Q[k]\sin(\frac{2\pi}{n}k*m)])^2 + \frac{1}{n}\sum_{m=0}^{n-1}[(\frac{1}{n}I[0] + \frac{1}{n}I[\frac{n}{2}])^2]$$
$$P = \frac{4}{n^3}\sum_{m=0}^{n-1}\sum_{k=1}^{\frac{n}{2}-1}[\frac{I^2[k]}{2} + \frac{Q^2[k]}{2}] + \frac{1}{n}\sum_{m=0}^{n-1}[(\frac{1}{n}I[0] + \frac{1}{n}I[\frac{n}{2}])^2]$$
$$= \frac{4}{n^3}\sum_{m=0}^{n-1}(\frac{n}{2}-1)A^2 + \frac{1}{n^2}2A^2 = \frac{4}{n^3}\frac{n^2}{2}A^2 - A^2\frac{4}{n^2} + \frac{2A^2}{n^2}$$
$$P = \frac{2}{n}A^2 - \frac{2A^2}{n^2} = \frac{2A^2}{n}[1-\frac{1}{n}] \approx \frac{2A^2}{n}$$

(22)

Note that the cross product terms in $(\frac{1}{n}I[0] + \frac{1}{n}I[\frac{n}{2}])^2$ vanish over time, given that $I[\frac{n}{2}]$ and $I[0]$ can take values of $+A$ or $-A$ across time, with equal probability.

As $n \to \infty$, average signal power $P = \frac{2A^2}{n}[1 - \frac{1}{n}] \approx \frac{2A^2}{n}$. Hence $A = \sqrt{\frac{Pn}{2}}$. If $A = \sqrt{\frac{3n}{2}}$, we have $P = 3$.

For higher order signalling, $M > 2$, average signal power is given by $P = P_0 B$ where $P_0 = \frac{2}{M}\sum_{r=0}^{M-1}(2r+1)^2$. In this case, DFT samples are scaled by a factor $A = \sqrt{\frac{Bn}{2}}$.

Appendix A.1

AWGN (Additive White Gaussian Noise) $w[m]$ is added to $s[m]$ to yield the received signal $r[m] = s[m] + w[m]$. Similar to the signal $s[m]$, we can express $w[m]$ as a DFT as follows.

$$w[m] = \frac{2}{n}\sum_{k=1}^{\frac{n}{2}-1}[I_w[k]\cos(\frac{2\pi}{n}k * m) + Q_w[k]\sin(\frac{2\pi}{n}k * m)] + \frac{1}{n}I_w[0] + \frac{1}{n}I_w[\frac{n}{2}] \tag{23}$$

where $I_w[k]$ and $Q_w[k]$ are Gaussian Random Variables(RV) of standard deviation $\sigma = \sqrt{\frac{Nn}{2}}$ for noise average power N. For $N = 1$, we have $\sigma = \sqrt{\frac{n}{2}}$. [$w[n]$ is an AWGN time sequence whose average power in time domain $N_t = \frac{1}{n}\sum_{m=0}^{n-1} w^2[m]$ equals corresponding average power in frequency domain $N_{fr} = \frac{1}{n^2}\sum_{k=0}^{n-1} W^2[k]$ as per Parseval's relation, where $W[k] = W_k I_w[k] + iQ_w[k]$. Given that w[n] is an ergodic random process, time averages(or frequency domain averages) equal ensemble averages and hence the average power of each RV $W[k]$ is given by $\frac{1}{E}\sum_{e=0}^{E-1}(|W_k[e]|)^2$ and equals $N_{fr} * n$ where E is the number of random experiments. The average power of each RV $I_w[k]$ and $Q_w[k]$ is $N_{fr} * \frac{n}{2}$ and standard deviation of each RV is $\sqrt{\frac{N_{fr}*n}{2}}$ which is $\sigma = \sqrt{\frac{n}{2}}$ for $N = 1$.]

Hence, in the Euclidean constellation space, for $P = 3$ and $N = 1$, each DFT coefficient of signal $I[k]$ and $Q[k]$ is represented by $+\sqrt{\frac{3n}{2}}$ or $-\sqrt{\frac{3n}{2}}$. Each DFT coefficient of noise $I_w[k]$ and $Q_w[k]$ has a standard deviation represented by $+\sqrt{\frac{n}{2}}$ or $-\sqrt{\frac{n}{2}}$.

Normalizing the Euclidean constellation points by $\sqrt{\frac{n}{2}}$, we have each normalized DFT coefficient of signal $I[k]$ and $Q[k]$ is represented by $+\sqrt{3}$ or $-\sqrt{3}$. Each normalized DFT coefficient of noise $I_w[k]$ and $Q_w[k]$ has a normalized standard deviation represented by $+1$ or -1 for the case of $\frac{C}{W} = 2, P = 3, N = 1$.

9. Appendix B :

Appendix B.1

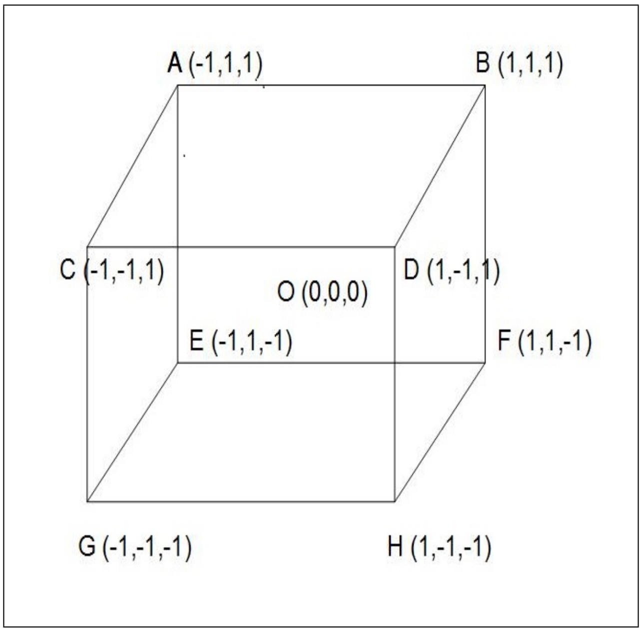

Figure 5:

Gray coded OFDM constellation for $M = 2$ is plotted in the figure above upto 3 dimensions. For $M = 2$, bit 1 is coded as $+1$ and bit 0 is coded as -1 and it is automatically graycoded and adjacent constellation points differ by only 1 bit. Bits (111) are assigned to Constellation point $B(1, 1, 1)$ and Bits (000) are assigned to Constellation point $G(-1, -1, -1)$. A block of $k' = k$ source bits are mapped to $n' = n$ bits using (n', k') error correction code and n' bits are gray coded into 2^n symbols using $M = 2$-ary signalling and are represented in n-dimensional Euclidean space. We transmit $2^{k'}$ codewords out of $2^{n'}$ possible sets of codewords.

One half the Euclidean distance between any 2 constellation points d_{eucl} for this $M = 2$ constellation is given by $\frac{d_{eucl}}{2} = \sqrt{d_{bits}}$ where d_{bits} is the number of bits by which the 2 constellation points differ.

For example, if $d_{min} = 3$ for this code, as in the Figure above, and we transmit constellation points B, G corresponding to 2 codewords, the Euclidean distance $d_{eucl_{min}} = d_{BG} = 2 * \sqrt{3}$ and $\frac{d_{BG}}{2} = \frac{d_{eucl_{min}}}{2} = \sqrt{d_{min}}$ where $d_{eucl_{min}}$ is the Euclidean distance between 2 valid codewords separated by the minimum distance for that code.

We require the radius of the noise sphere in n-dimensions, namely $\sqrt{nN} < \sqrt{P}\frac{d_{eucl_{min}}}{2}$; $\sqrt{nN} < \sqrt{Pd_{min}}$. If we use Gallager's LDPC codes with $\frac{d_{min}}{n} = 0.11 = K$ for $\frac{k}{n} = \frac{1}{2}$ and for $N = 1$, this implies we require $\sqrt{n} < \sqrt{PnK}$ or average signal power $P = \frac{1}{K} = 9$.

Appendix B.2

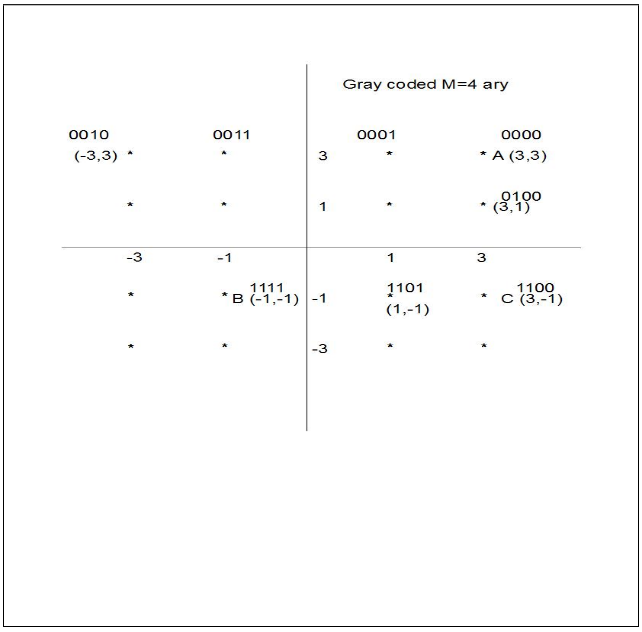

Figure 6:

Gray coded OFDM constellation for $M = 4$ is plotted in the figure above upto 2 dimensions. A block of $k' = 2k$ source bits are mapped to $n' = 2n$ bits using (n', k') error correction code and $n' = 2n$ bits are gray coded into 4^n symbols using $M = 4$-ary signalling and are represented in n-dimensional Euclidean space. We transmit 2^{2k} codewords out of 2^{2n} possible sets of codewords.

One half the Euclidean distance between any 2 constellation points d_{eucl} for this $M = 4$ constellation is given by $\frac{d_{eucl}}{2} \geq \sqrt{d_{bits}}$ where d_{bits} is the number of bits by which the 2 constellation points differ.

We require the radius of the noise sphere in n-dimensions, namely $\sqrt{nN} < \sqrt{P}\frac{d_{eucl_{min}}}{2}$ where $d_{eucl_{min}}$ is the Euclidean distance between 2 valid codewords separated by the minimum distance for that code. We require $\sqrt{nN} < \sqrt{Pd_{min}}$. If we use Gallager's LDPC codes with $\frac{d_{min}}{n'} = 0.11 = K$ for $\frac{k'}{n'} = \frac{1}{2}$ and for $N = 1$, $n' = 2n$, this implies we require $\sqrt{n} < \sqrt{2PnK}$ or average signal power $P = \frac{1}{2*K} = 4.5$.

Appendix B.3

We can extend this analogy to $M = 8$ case. We have 8 symbols per dimension, $\pm 1, \pm 3, \pm 5, \pm 7$. A block of $k' = 3k$ source bits are mapped to $n' = 3n$ bits using (n', k') error correction code and $n' = 3n$ bits are gray coded into M^n symbols using $M = 8$-ary signalling and are represented in n-dimensional Euclidean space. We transmit 2^{3k} codewords out of 2^{3n} possible sets of codewords.

One half the Euclidean distance between any 2 constellation points d_{eucl} for this $M = 8$ constellation is given by $\frac{d_{eucl}}{2} \geq \sqrt{d_{bits}}$ where d_{bits} is the number of bits by which the 2 constellation points differ.

We require the radius of the noise sphere in n-dimensions, namely $\sqrt{nN} < \sqrt{P}\frac{d_{eucl_{min}}}{2}$ where $d_{eucl_{min}}$ is the Euclidean distance between 2 valid codewords separated by the minimum distance for that code. We require $\sqrt{nN} < \sqrt{Pd_{min}}$. If we use Gallager's LDPC codes with $\frac{d_{min}}{n'} = 0.11 = K$ for $\frac{k'}{n'} = \frac{1}{2}$ and for $N = 1$, $n' = 3n$, this implies we require $\sqrt{n} < \sqrt{3PnK}$ or average signal power $P = \frac{1}{3*K} = 3$.

$M > 8$

In general, for $M > 8$, we require $\sqrt{nN} < \sqrt{Pd_{min}}$. If we use Gallager's LDPC codes with $\frac{d_{min}}{n'} = 0.11 = K$ for $\frac{k'}{n'} = \frac{1}{2}$ and for $N = 1$, $n' = n\log_2(M)$, this implies we require $\sqrt{n} < \sqrt{\log_2(M)PnK}$ or average signal power $P = \frac{1}{\log_2(M)*K} < 3$.

Alternatively, we can hold the average signal power constant at $P = 3$(similar to Shannon's expression) and instead opt for higher $\frac{k'}{n'} > \frac{1}{2}$ which corresponds to smaller $\frac{d_{min}}{n'} = K < 0.11$ so that $\sqrt{n} < \sqrt{\log_2(M)PnK}$, so we need $K < \frac{1}{\log_2(M)P}$ which means $K < \frac{1}{3\log_2(M)}$. As $\frac{C}{W} = 2\log_2(M)$ increases, K decreases and $\frac{k'}{n'}$ increases towards 1.

10. Appendix C

Let us derive the expression for Bit Error Rate(BER) for M-ary PAM system used in each FFT sample in OFDM. It is plotted in the Figure below.

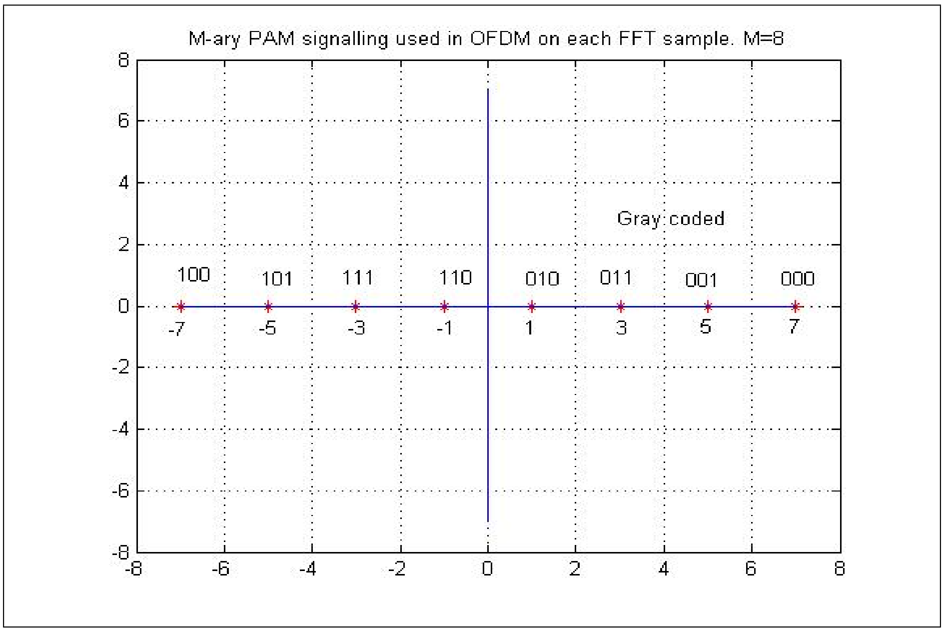

Figure 7:

This reference constellation is scaled by a factor A such that average signal power $P = P_0 A^2$ where P_0 is the average power of the reference constellation. A is chosen to achieve a desired Signal to Noise Ratio(SNR) which yields a desired BER.

Let us start with the outermost symbol numbered 1 to the right, which has a value of $+7A$. Given that the symbol error rate for this symbol is given by the probability that noise takes a value larger than half the distance between adjacent symbols, we can write $P_{ser} = \frac{1}{2}erfc(\sqrt{\frac{E_b}{N_0}}) = \frac{1}{2}erfc(\sqrt{\frac{E_{bpsk}}{N_0}}) = \frac{1}{2}erfc(\sqrt{\frac{P_{bpsk}}{N}WT_b}) = \frac{1}{2}erfc(\sqrt{\frac{P_{bpsk}}{2N}}) = \frac{1}{2}erfc(\sqrt{\frac{A^2}{2N}})$. Note that the distance between adjacent symbols is $2A$ and E_b in above expression corresponds to the Energy Per Bit for BPSK constellation and NOT the overall $E_b$for this M-ary signalling system. We can write the BER contributed by this **outermost symbol** as follows.

$$BER_1 = \frac{1}{M}[\sum_{m=0}^{M-3}(\frac{1}{2}erfc((2m+1)\sqrt{\frac{A^2}{2N}}) - \frac{1}{2}erfc((2m+3)\sqrt{\frac{A^2}{2N}})] * d_{bits}(s_1, s_{(m+2)})) \\ + \frac{1}{2}erfc((2M-3)\sqrt{\frac{A^2}{2N}}) * d_{bits}(s_1, s_{(M)})] \tag{24}$$

where $d_{bits}(s_1, s_{(m+2)})$ is the number of bits which differ in Symbol 1 and Symbol $m+2$.

We can write the BER contributed by this **inner symbols** for symbol number $i = 2, 3, ..M/2$ as follows in terms of 2 terms, first term accounts for symbols to the left of this symbol and second term accounts for symbols to the right of this symbol.

$$BER_i = \frac{1}{M}[\sum_{m=0}^{M-1-i}(\frac{1}{2}erfc((2m+1)\sqrt{\frac{A^2}{2N}}) - \frac{1}{2}erfc((2m+3)\sqrt{\frac{A^2}{2N}})] * d_{bits}(s_i, s_{(m+1+i)})) \\ + \frac{1}{2}erfc((2M+1-2i)\sqrt{\frac{A^2}{2N}}) * d_{bits}(s_i, s_{(M)})] \\ +[\sum_{m=0}^{i-3}(\frac{1}{2}erfc((2m+1)\sqrt{\frac{A^2}{2N}}) - \frac{1}{2}erfc((2m+3)\sqrt{\frac{A^2}{2N}})] * d_{bits}(s_i, s_{(i-m-1)})) \\ + \frac{1}{2}erfc((2i-3)\sqrt{\frac{A^2}{2N}}) * d_{bits}(s_i, s_{(1)})] \tag{25}$$

For symbols numbered $\frac{M}{2}+1, ...M$, the BER is the same as the corresponding symbols $\frac{M}{2}, \frac{M}{2}-1, ..1$. Hence the total BER of this M-ary signalling system is given by

$$BER = \frac{1}{log_2(M)}\sum_{i=1}^{M} BER_i \tag{26}$$

- **Appendix C.1**

Let us use Gray Coding for this constellation. +3 corresponds to bits 00, +1 corresponds to bits 01, −1 corresponds to bits 11, −3 corresponds to bits 10.

The Raw BER of this 4-ary signalling system can be expressed as a weighted sum of BER expression for binary signalling case $M = 2$ as follows[as per Appendix C].

$$\begin{aligned} P_{e_{(M=4)}} &= \frac{1}{2}[0.5 * [P_0 + (P_0 - P_1) + 2 * P_1] + 0.5 * [(P_0 - P_1) + 2(P_1 - P_2) + P_2]] \\ P_0 &= \frac{1}{2} erfc(\sqrt{\frac{E_{b_{(BPSK)}}}{N_0}}) = \frac{1}{2} erfc(\sqrt{\frac{P_{BPSK}}{2N}}) = \frac{1}{2} erfc(\sqrt{\frac{A^2}{2N}}) \\ P_1 &= \frac{1}{2} erfc(3\sqrt{\frac{E_{b_{(BPSK)}}}{N_0}}) = \frac{1}{2} erfc(3\sqrt{\frac{P_{BPSK}}{2N}}) = \frac{1}{2} erfc(\sqrt{\frac{A^2}{2N}}) \\ P_2 &= \frac{1}{2} erfc(5\sqrt{\frac{E_{b_{(BPSK)}}}{N_0}}) = \frac{1}{2} erfc(5\sqrt{\frac{P_{BPSK}}{2N}}) = \frac{1}{2} erfc(\sqrt{\frac{A^2}{2N}}) \\ P_3 &= \frac{1}{2} erfc(7\sqrt{\frac{E_{b_{(BPSK)}}}{N_0}}) = \frac{1}{2} erfc(7\sqrt{\frac{P_{BPSK}}{2N}}) = \frac{1}{2} erfc(\sqrt{\frac{A^2}{2N}}) \end{aligned} \tag{27}$$

where $E_{b_{(BPSK)}}$ is the Energy Per Bit for a Binary Signalling System and N_0 is the Noise Spectral Density and the average signal power for Binary signalling system is given by $P_{BPSK} = A^2$ and average noise power is given by $N = N_0 W$ and $WT_b = \frac{1}{2}$.

We require this Raw BER $P_{e_{(M=4)}}$ to be less than Error Correction Rate(ECR) of the associated Gallager's LDPC code.
$P_{e_{(M=4)}} \leq \lim_{n' \to \infty} \frac{d_{min}}{2n'}$.

The minimum value of $\frac{E_b}{N_0}$ which satisfied above equality is given by $(\frac{E_b}{N_0})_{min} = \frac{A^2}{2N}$.

For $\frac{d_{min}}{2n'} = 0.055$ and code rate $\frac{k}{n} = 0.5$, $(\frac{E_b}{N_0})_{min}$ is around 1.05 for $M = 4$.

$(\frac{C}{W})_{OFDM} = \log_2(1 + \frac{3P}{2N*1.05}) = \log_2(1 + \frac{P*1.4285}{N})$.

This seems to perform better than Shannon's Capacity Limit which uses n dimensional noise spheres, $\frac{C}{W} = log_2(1 + \frac{P}{N})$, but we should remember that $\frac{C}{W}$ is the source spectral efficiency while $(\frac{C}{W})_{OFDM}$ is the channel spectral efficiency and the **source spectral efficiency** is given by $(\frac{C}{W})_{OFDM(source)} = \frac{k}{n} \log_2(1 + \frac{P*1.4285}{N})$.