3F4: Data Transmission

Handout 8: Orthogonal Frequency Division Multiplexing (OFDM)

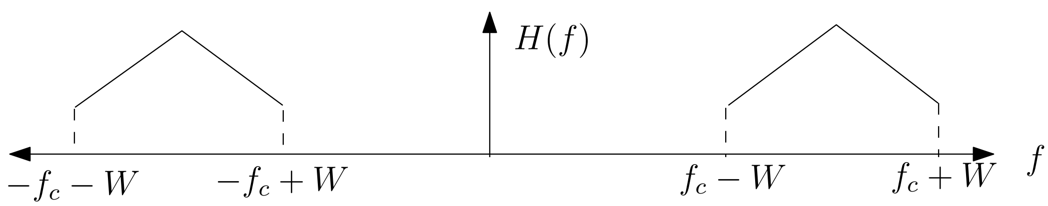
Ramji Venkataramanan

Signal Processing and Communications Lab, CUED

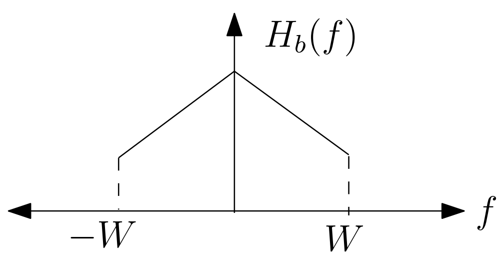
ramji.v@eng.cam.ac.uk

Lent Term 2019

As in the previous handout, the goal is to communicate over a dispersive (frequency-selective) channel.
For example,



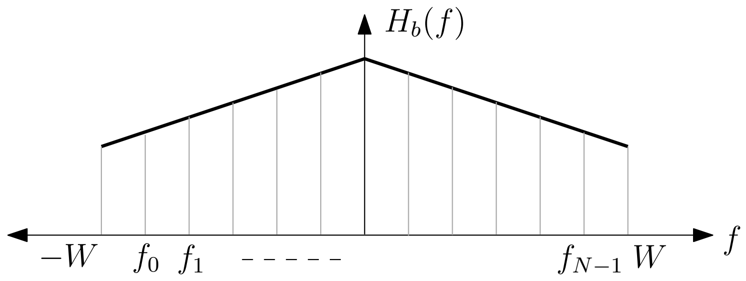
The *baseband equivalent* frequency response of the channel is



- We can focus on understanding how to communicate over this (possibly complex-valued) baseband equivalent channel
- Then if $x_b(t)$ is the (complex) transmitted waveform for the baseband equiv. channel, for the passband channel we transmit

$$x(t) = \text{Re}[x_b(t)e^{j2\pi f_c t}]$$

The main idea in OFDM



- Choose a set of frequencies $-W \le f_0, \ldots, f_{N-1} \le W$ that define an set of *orthogonal* sinusoids over a symbol period T.
- Each of these N orthogonal sinusoids is a *sub-carrier*, which carries an independent PAM/QAM symbol in the symbol time T.
- When passed through the dispersive channel, we ensure that each subcarrier f_n sees just a *scalar* channel gain $H_b(f_n)$, for $n = 0, \ldots, (N-1)$.
- Thus we effectively convert the wideband dispersive channel into N *non-interfering* narrowband channels.

- Let $X[n]$ be QAM symbol to be transmitted via the nth sub-carrier f_n, for $n = 0, \ldots, N-1$.
- The transmitted complex baseband waveform is

$$x_b(t) = \sum_{n=0}^{N-1} X[n]\, e^{j2\pi f_n t}\, \mathbf{1}_{\{t \in [0,T]\}} \tag{1}$$

Here $\mathbf{1}_{\{t \in [0,T]\}}$ denotes the indicator function that equals 1 for $t \in [0, T]$, and 0 otherwise.

- We want the sub-carriers to be orthogonal, i.e.,

$$\langle e^{j2\pi f_n t}\mathbf{1}_{\{t \in [0,T]\}},\ e^{j2\pi f_m t} \rangle = 0, \quad \text{for } n \neq m.$$

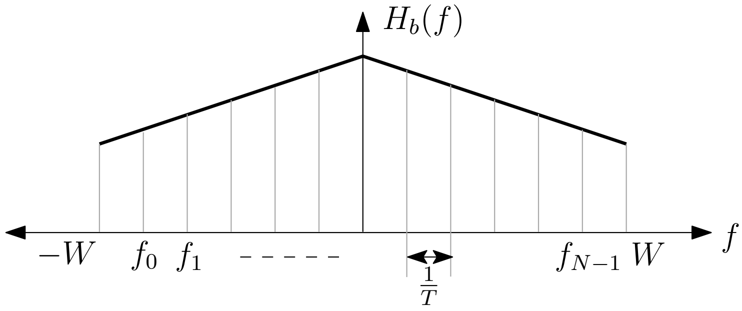
This will be true if $(f_n - f_m)T =$ non-zero integer. To see this, we compute the inner product:

$$\int_0^T e^{j2\pi f_n t}\, e^{-j2\pi f_m t}\, dt = \frac{e^{j2\pi(f_n - f_m)T} - 1}{j2\pi(f_n - f_m)} = 0,$$

where the last equality holds whenever $(f_n - f_m)T =$ non-zero integer.

We can therefore ensure orthogonality by choosing sub-carriers $f_0, \ldots, f_{N-1}$ such that adjacent frequencies are spaced $\frac{1}{T}$ apart:

$$f_n = f_0 + \frac{n}{T}, \quad n = 0, \ldots, (N-1).$$



To understand what happens when $x_b(t)$ is passed through the channel with impulse response $h_b(t)$, we look at the operation in frequency domain.

- Let us first examine the spectrum of the transmitted waveform in Eq. (1). Note that

$$P_n(f) = \mathcal{F}\left[e^{\mathrm{j}2\pi f_n t}\, \mathbf{1}_{\{t\in[0,T]\}}\right] = T\mathrm{sinc}\left(\pi(f-f_n)T\right) e^{-\mathrm{j}\pi f T} \tag{2}$$

- Therefore, for a fixed set of information symbols $\{X[0], \ldots, X[N-1]\}$, the spectrum of $x_b(t)$ in (1) is:

$$X_b(f) = \sum_{n=0}^{N-1} X[n]\, P_n(f)$$

- This is passed through the channel with freq. response $H_b(f)$.
- Therefore the output of the baseband equivalent channel (without noise) is

$$Y_b(f) = \sum_{n=0}^{N-1} X[n]\, P_n(f) H_b(f). \tag{3}$$

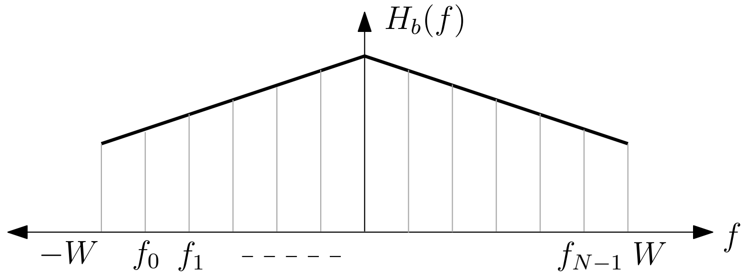
- From Eq. (2), notice that $|P_n(f)|$ is small, when $|f - f_n| \gg \frac{1}{T}$.
- If we take T large enough, so that $H_b(f)$ is roughly constant over a few multiples of $\frac{1}{T}$, then $P_n(f)H_b(f) \approx P_n(f)H_b(f_n)$.

- Hence, from Eq. (3) we have

$$Y_b(f) \approx \sum_{n=0}^{N-1} X[n]\, P_n(f) H_b(f_n),$$

or, using Eq. (2) to invert $P_n(f)$ we get

$$y_b(t) \approx \sum_{n=0}^{N-1} H_b(f_n)\, X[n]\, e^{j2\pi f_n t}\, \mathbf{1}_{\{t\in[0,T]\}}. \qquad (4)$$



Comparing Eqs. (1) and (4), we see that the information symbol in the nth sub-carrier is just multiplied by the scalar $H_b(f_n)$.

At the Rx, to extract the information symbol from each sub-carrier, we just compute inner products.
Assuming no additive noise, from Eq. (4), we have for $n = 0, \ldots, (N-1)$:

$$Y[n] := \frac{1}{T}\langle y_b(t),\, e^{j2\pi f_n t}\rangle = \frac{1}{T}\int_0^T y_b(t)\, e^{-j2\pi f_n t} \approx X[n] H_b(f_n)$$

With additive noise, we get

$$Y[n] \approx X[n] H_b(f_n) + N[n], \qquad n = 0, \ldots, (N-1).$$

- Thus, at the demod output, we have N *non-interfering* AWGN channels, one for each sub-carrier.
- The channel coefficient $H_b(f_n)$ determines the snr available to the nth AWGN channel.
- $N[n]$ is complex Gaussian noise with the same variance for each sub-carrier n.

Note that OFDM is similar to FSK with N orthogonal sinusoids, with the additional feature that we send information by modulating the amplitude of each sub-carrier with a QAM symbol.

- The ideas outlined in the scheme above have been known for decades.
- What's made OFDM attractive in recent years is that it can be efficiently implemented in digital hardware (DSP).
- Furthermore, the DSP implementation can restore *exact* orthogonality of the sub-carriers at the cost of a small overhead.

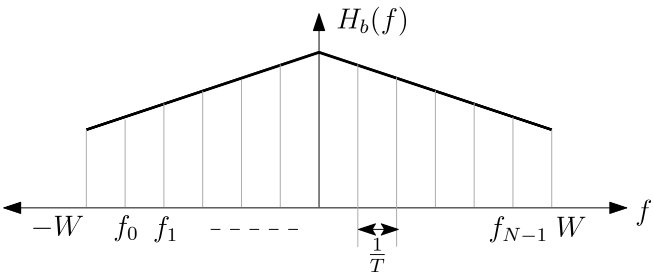
DSP implementation of OFDM

Notation: For a sequence $X[n]$, $n = 0, \ldots, (N-1)$, we denote its *inverse DFT* sequence by $x[k]$, $k = 0, \ldots, (N-1)$.

Recall that for $n, k \in \{0, \ldots, (N-1)\}$:

$$x[k] = \sum_{n=0}^{N-1} X[n] e^{j2\pi nk/N}, \tag{5}$$

$$X[n] = \frac{1}{N} \sum_{k=0}^{N-1} x[k] e^{-j2\pi nk/N}. \tag{6}$$



- Recall that the sub-carriers are spaced $\frac{1}{T}$ apart with frequencies $f_0 + \frac{n}{T}$, for $n = 0, 1 \dots, N$
- The baseband signal in Eq. (1) can then be written as

$$x_b(t) = e^{j2\pi f_0 t} \sum_{n=0}^{N-1} X[n]\, e^{j2\pi nt/T}\, \mathbf{1}_{\{t \in [0,T]\}}$$

- The passband bandwidth of up-converted $x(t)$ will be $\approx \frac{N}{T}$
- Hence we can accurately represent the $x_b(t)$ by sampling at rate $\frac{1}{T_s} \geq \frac{N}{T}$.
- T_s is the sampling interval for the baseband equivalent OFDM signal $x_b(t)$.

- For notational convenience, we will take $f_0 = 0$.
  The assumption $f_0 = 0$ just means that the baseband frequency content won't be centered around the origin, but this can be adjusted during implementation.
- With this assumption, the baseband signal is:

$$x_b(t) = \sum_{n=0}^{N-1} X[n]\, e^{j2\pi nt/T}\, \mathbf{1}_{\{t \in [0,T]\}} \tag{7}$$

- We will also take the sampling interval to be $T_s = \frac{T}{N}$.
- Then the T_s-spaced samples of $x_b(t)$ in the interval $[0, T]$ are:

$$x_b(kT_s) = \sum_{n=0}^{N-1} X[n]\, e^{j2\pi nkT_s/T} = \sum_{n=0}^{N-1} X[n]\, e^{j2\pi nk/N}, \quad 0 \leq k < N.$$

That is, the N samples of $x_b(t)$ for $0 \leq t < T$ are just the **inverse DFT** of the symbol sequence $\{X[n]\}$, $n = 0, \dots, (N-1)$.

OFDM transmitter

We make the inverse DFT (IDFT) explicit using the following notation:

$$x[k] = x_b(kT_s) = \sum_{n=0}^{N-1} X[n]\, e^{j2\pi nkT_s/T} = \sum_{n=0}^{N-1} X[n]\, e^{j2\pi nk/N}. \qquad (8)$$

Thus the transmitter can be implemented as follows:

- Compute the IDFT $\{x[k]\}_{0\leq k<N}$ of the source symbols $\{X[n]\}_{0\leq n<N}$. This can be efficiently implement in hardware using FFT algorithms.
- Use the samples $\{x[k] = x_b(kT_s)\}$ to generate the complex baseband waveform $x_b(t)$ using a digital-to-analog converter.
- The D/A converter is an *interpolating* filter that generates $x_b(t)$ from its samples. For example,

$$x_b(t) = \sum_{k=0}^{N-1} x[k]p(t-kT_s), \qquad (9)$$

with $p(t) = \mathrm{sinc}(\pi t/T_s)$. (Recall from 1B paper 6 that the ideal interpolating filter is a low-pass filter.)
- Upconvert to obtain passband waveform: $x(t) = \mathrm{Re}[x_b(t)e^{j2\pi f_c t}]$

At the Receiver

- The baseband equivalent of the transmitted OFDM signal (over a period T) is

$$x_b(t) = \sum_{k=0}^{N-1} x[k]p(t-kT_s),$$

where $T_s = T/N$.
- The (noiseless) received baseband equivalent signal is

$$y_b(t) = \sum_{k=0}^{N-1} x[k]\, g(t-kT_s),$$

where $g(t) = p(t) \star h(t) \star q(t)$ is the **overall filter**, which includes channel impulse response $h(t)$ and receive filter $q(t)$.
- Sampling the received signal at rate $1/T_s$, we get

$$y[k] = y_b(kT_s) = \sum_{m=0}^{N-1} x[m]g[k-m], \quad \text{for } k = 0, \dots, (N-1).$$

Here $g[\ell] = g(\ell T_s)$ is assumed to be non-zero only for $0 \leq \ell \leq L$.

Hence the convolution generating $y[k]$ can be written as

$$y[k] = \sum_{\ell=0}^{L} g[\ell]\, x[k-\ell], \qquad \text{for } k = 0, \ldots, N-1. \tag{10}$$

- Typically the number of sub-carriers N is larger than L.
- Eq. (10) represents a *linear* convolution between the length-N sequences $[x[0], \ldots, x[N-1]]$ and $[g[0], \ldots, g[L], 0, \ldots, 0]$.
- Can we take N-point DFT of (10) and claim that

$$Y[n] = G[n]\, X[n], \quad n = 0, \ldots, (N-1) \; ? \tag{11}$$

- No! Because term-by-term multiplication of DFT corresponds to **circular** convolution in the discrete-time domain.
- Hence, for Eq. (11) to be true, what we actually want is:

$$\tilde{y}[k] = \sum_{\ell=0}^{L} g[\ell]\, x[(k-\ell) \mod N], \qquad \text{for } k = 0, \ldots, N-1. \tag{12}$$

Linear vs. circular convolution

Let us examine the difference between linear convolution and the circular convolution, recalling that L is smaller than N.

- For $k = L, L+1, \ldots, (N-1)$, notice that $y[k]$ in Eq. in (10) and $\tilde{y}[k]$ in (12) are identical.
- But for $k = 0, 1, \ldots, (L-1)$:

  Linear convolution:

$$y[k] = g[0]x[k] + g[1]x[k-1] + \ldots + g[k]x[0] \tag{13}$$

  Circular convolution:

$$\begin{aligned}\tilde{y}[k] = {} & g[0]x[k] + g[1]x[k-1] + \ldots + g[k]x[0] \qquad (14)\\ & + \underbrace{g[k+1]x[N-1] + g[k+2]x[N-2] + \ldots g[L]x[N-L+k]}_{\text{these terms are not present in (13)}}.\end{aligned}$$

The difference is due to the fact that in the linear conv. in (13)

$$g[k+1]x[-1] + \ldots + g[L]x[-L+k)] = 0$$

as $x[-1] = \ldots = x[-(L-k)] = 0$, for $0 \leq k \leq (L-1)$.

The cyclic prefix

To make linear conv. mimic the circular one in (14), we define

$$x[-L] = x[N-L], \quad x[-L+1] = x[N-L+1], \\ \ldots, x[-1] = x[N-1]. \tag{15}$$

and transmit the symbols $x[-L], \ldots, x[-1]$ *before* the block $x[0], \ldots, x[N-1]$.

- $x[-L], \ldots, x[-1]$ is called the **cyclic prefix** for the block $x[0], \ldots, x[N-1]$.
- With the cyclic prefix, the channel output for $0 \le k \le N-1$ is

$$\begin{aligned} y[k] &= \sum_{\ell=0}^{L} g[\ell]\, x[k-\ell] \\ &= g[0]x[k] + g[1]x[k-1] + \ldots + g[k]x[0] \qquad (16) \\ &\quad + \underbrace{g[k+1]x[N-1] + g[k+2]x[N-2] + \ldots g[L]x[N-L+k]}_{\text{these terms are present in the linear conv. due to the cyclic prefix Eq. (15)}} \end{aligned}$$

- The receiver computes the N-point DFT of $y[0], \ldots, y[N-1]$ (discarding the outputs $y[-L], \ldots, y[-1]$) to get

$$Y[n] = G[n]X[n], \quad \ldots, k = 0, \ldots, (N-1).$$

- With additive noise, the channel output in the DFT domain is

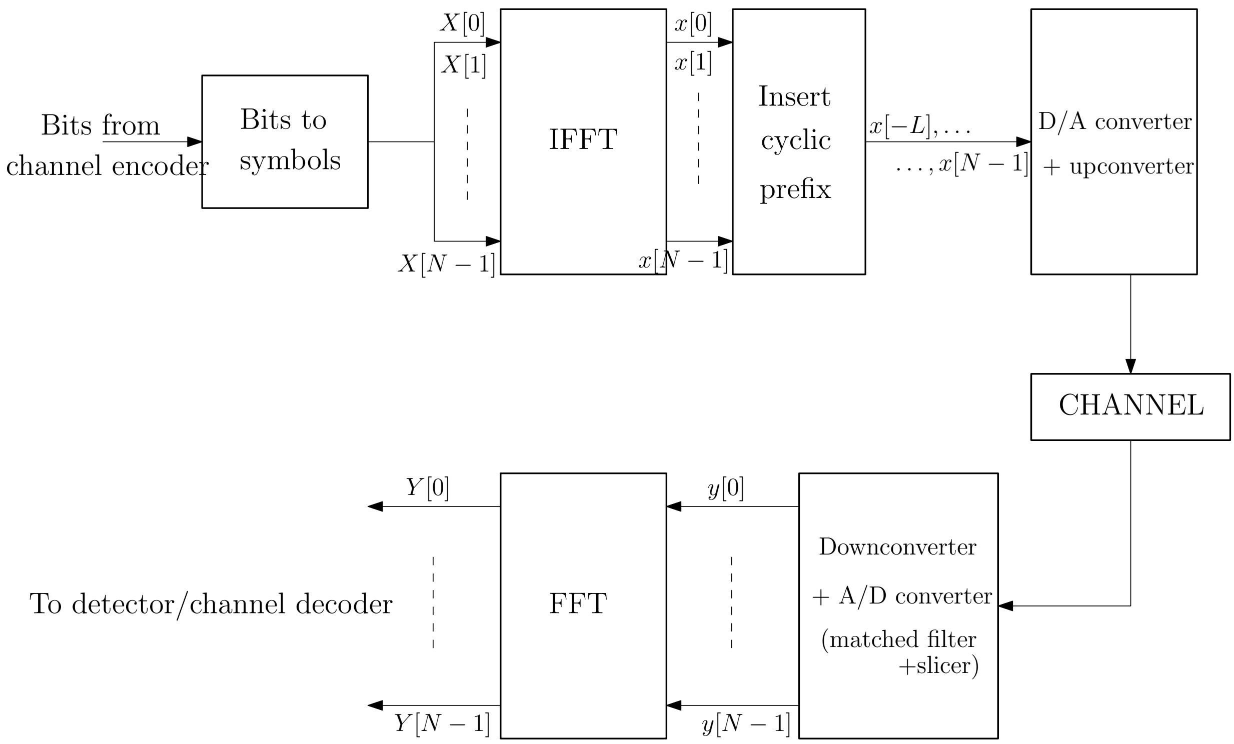
$$Y[n] = G[n]X[n] + N[n], \quad \ldots, k = 0, \ldots, (N-1).$$

That is, we have N *non-interfering* AWGN channels with no ISI.

Overhead due to cyclic prefix:

- For each block of N information symbols, we send a cyclic prefix of L symbols.
- Recall that N information symbols are to be transmitted in $t \in [0, T)$. Additionally, the cyclic prefix has be transmitted beforehand in $t \in [-LT_s, 0)$, where $T_s = T/N$.
- This prefix interval of length LT_s is called the "guard interval" or "guard period" and is necessary to ensure perfect orthogonality of N channels at the output.

Block diagram of OFDM Tx + Rx implementation



Summary

Two methods for communication over wide-band frequency selective channels:

1) Channel equalisation: Use standard *single-carrier* PAM/QAM, and deal with ISI at the receiver by post-processing the received sequence to extract the information symbols.

2) OFDM: A *multi-carrier* scheme which uses N orthogonal carriers. Orthogonality of the carriers can be maintained at the receiver via a cyclic prefix at the cost of a small overhead.

$$Y[n] = G[n]X[n] + N[n], \quad , n = 0, \ldots, (N-1).$$

- The snr for the channel corresponding to the nth carrier depends on the channel gain $G[n] \propto H_b(f_n)$.
- In coded OFDM, an outer channel code is used to correct symbol errors that may occur on the low snr subcarriers.
- We can also vary the symbol rate across sub-carriers: transmit higher rate over sub-carriers with larger channel gain.

OFDM very widely used: 4G, 5G wireless standards, DSL, Wi-Fi, digital video broadcast, ...

In Part 1 of 3F4, we covered:

- Signal space
- Optimal transmit/receive filters for PAM baseband modulation on AWGN channels
- Passband modulation: QAM and FSK
- Dealing with ISI in wideband frequency selective channels: Equalisation and OFDM

Some topics we did not cover: *synchronization* techniques to estimate delays, frequency/phase shifts introduced by the channel

Part 2 of the module (6 lectures by Prof. Ioannis Kontoyiannis) will discuss channel coding, convolutional codes, network algorithms

Exam information:

- Standard IIA format: 1.5 hours, you answer 3 out of 4 questions.
- 2-3 questions from Part 1 of the module, 1-2 from Part 2.
- List of relevant past Tripos questions available on Moodle.