CHANNEL ESTIMATION IN DVB-T AND OFDM SYSTEMS

RETHNAKARAN PULIKKOONATTU

ABSTRACT. This document outlines the theoretical framework behind pilot based channel estimation in an OFDM system. Fundamental concepts of optimal frequency domain estimation of a finite length channel are derived and illustrated. Practical limitations and implementation complexity of optimal algorithms often prompt to design sub-optimal estimation schemes. Some such circuits are discussed. Specifics of DVB-T from a channel estimation perspective is considered while discussing the specifics.

This document is a highly condensed version of the internal technical report on channel estimation prepared by the same authors from ST (Genesis Microchip group). Since this document is intended to a wider readership, the IP and implementation details of algorithms and design aspects are not discussed. For internal ST employees, the detailed report discussing all aspects of channel estimation in IP are available for reference. The scope of this document hence is restricted to the discussion of general channel estimation of OFDM systems with specific system under study as DVB-T.

Date: 01 March 2007.
Key words and phrases. DVB.
ST Microelectronics (Genesis Microchip) India, www.st.com .

CONTENTS

1. Introduction 3
1.1. Data detection 5
1.2. Estimating the channel taps (Training) 5
2. Channel estimation 6
2.1. OFDM channel estimation: Deterministic noiseless channel 8
2.2. Deterministic channel with AWGN 9
3. Fading channel scenario 10
4. Optimum Channel estimation for a wireless channel 11
5. Channel estimation implementation 17
6. Pilot pattern in DVB-T 17
7. Appendix B: Linear Prediction 19
8. Simulink implementation 21
References 22

1. INTRODUCTION

A block diagram of OFDM system using cylcic prefix (or cyclic suffix) is shown in Figure 1 .

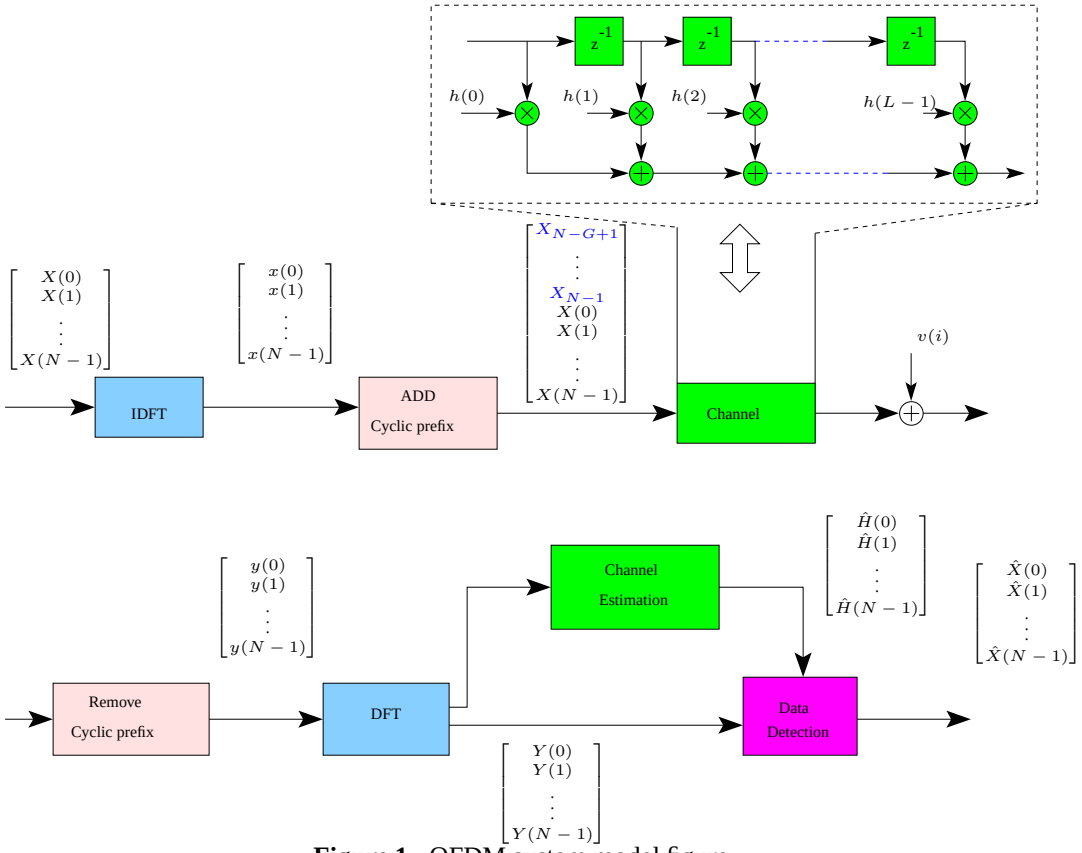

Figure 1. OFDM system model figure

To avoid inter symbol interference [1] the cyclic prefix length G must be atleast equal to the FIR channel length L. While a larger G helps in improved detection (in noisy channels), the effective transmission rate and bandwidth are sacrificed with it. Here for analysis, we nevertheless assume $G \leq L$. For $L > G$, channel output vector (receiver input) due to n-th OFDM transmission is given by

$$\begin{bmatrix} 0 & \dots & h_{L-1} & \dots & h_1 & h_0 & 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & \dots & h_2 & h_1 & h_0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & h_{L-1} & h_{L-2} & \dots & h_1 & h_0 & 0 \\ 0 & \dots & 0 & \dots & 0 & h_{L-1} & h_{L-2} & \dots & h_1 & h_0 \end{bmatrix} \begin{bmatrix} X_{N-G+1}(m) \\ \vdots \\ X_{N-2}(m) \\ X_{N-1}(m) \\ X_0(m) \\ X_1(m) \\ X_2(m) \\ X_3(m) \\ \vdots \\ X_{N-G+1}(m) \\ \vdots \\ X_{N-2}(m) \\ X_{N-1}(m) \end{bmatrix} + \mathbf{v}(n)$$

[1] Cylclic prefix also helps us to realise a simple equalizer structure to reduce ISI damage as illustrated in section xxx. Among the other (added) benefits of cyclic prefix includes data aided frame/time synchronization schemes

After the removal of cyclic prefix (chop the G columns of matrix $\mathbf{C}(n)$ and first G columns of receive input vector) we get the time domain signal (corruped by AWGN noise) vector $y(n)$ equal to,

$$\begin{bmatrix} h_0 & \dots & 0 & \dots & 0 & h_{L-1} & h_{L-2} & \dots & h_2 & h_1 \\ h_1 & h_0 & 0 & \dots & 0 & 0 & h_{L-1} & \dots & h_3 & h_2 \\ h_2 & h_1 & h_0 & \dots & 0 & 0 & 0 & \dots & h_4 & h_3 \\ h_3 & h_2 & h_1 & \dots & 0 & 0 & 0 & \dots & h_5 & h_4 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{L-2} & h_{L-3} & h_{L-4} & \dots & h_0 & 0 & 0 & \dots & 0 & h_{L-1} \\ h_{L-1} & h_{L-2} & h_{L-3} & \dots & h_1 & h_0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & h_{L-1} & h_{L-2} & h_{L-3} & h_{L-4} & \dots & h_0 & 0 \\ 0 & \dots & 0 & 0 & h_{L-1} & h_{L-2} & h_{L-3} & \dots & h_1 & h_0 \end{bmatrix} \begin{bmatrix} X_0(m) \\ X_1(m) \\ X_2(m) \\ X_3(m) \\ \vdots \\ X_{N-G+1}(m) \\ \vdots \\ X_{N-2}(m) \\ X_{N-1}(m) \end{bmatrix} + \begin{bmatrix} V_0(m) \\ V_1(m) \\ V_2(m) \\ V_3(m) \\ \vdots \\ V_{N-G+1}(m) \\ \vdots \\ V_{N-2}(m) \\ V_{N-1}(m) \end{bmatrix}$$

or simply,

$$\begin{aligned} \mathbf{y}(n) &= \mathbf{C}(n)\mathbf{x}(n) + v_N(n) \\ &= \mathbf{C}(n)\mathbf{F}_N^*\mathbf{X}(n) + \mathbf{v}(n) \end{aligned} \tag{1}$$

Since the time domain channel matrix $\mathbf{C}(n)$ is circulant, it can be decomposed using the well known Fourier diagonalization to,

$$\mathbf{C}(n) = \mathbf{F}_N^* \underbrace{\begin{bmatrix} H_1(n) & 0 & 0 & 0 & 0 \\ 0 & H_2(n) & 0 & \ddots & 0 \\ 0 & 0 & H_3(n) & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & 0 & H_{N-1}(n) \end{bmatrix}}_{\triangleq \mathbf{\Lambda}(n)} \mathbf{F}_N \tag{2}$$

Using Eq.2 on Eq.1, equivalent (and simple) frequency domain representation of the system can be arrived.

$$\mathbf{Y}(n) = \mathbf{\Lambda}(n)\mathbf{X}(n) + \mathbf{V}(n) \tag{3}$$

Figure 2 and Eq.4 shows the simplified model of OFDM .

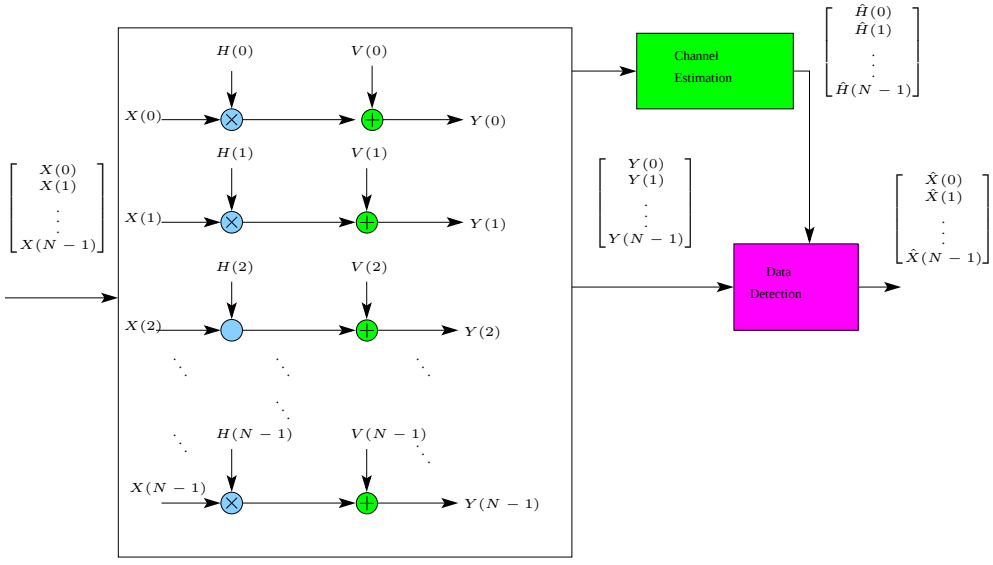

Figure 2. Frequency domain equivalent of OFDM system

1.1. **Data detection.** Ignoring the subscript n we could simply write,

$$\mathbf{Y}(n) = \mathbf{\Lambda}(n)\mathbf{X}(n) + \mathbf{V}(n) \tag{4}$$

Using results in Appendix (TBD) The LMMSE estimate vector $\hat{X}$ of the data vector X can be obtained.

$$\hat{X} = \mathbb{R}_{XX}\Lambda^* \left(R_{VV} + \Lambda\mathbb{R}_{XX}\Lambda^*\right)^{-1} Y \tag{5}$$

$$= \Lambda^* \left(\mathbb{D} + \Lambda\Lambda^*\right)^{-1} Y \tag{6}$$

D is again a diagonal matrix defined as,

$$\mathbb{D} \triangleq \begin{bmatrix} \frac{1}{SNR_0} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{SNR_1} & 0 & \vdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & \frac{1}{SNR_{N-1}} \end{bmatrix} \tag{7}$$

Note that, Eq 6 is derived from Eq. 5 for the AWGN case where,

$$R_{VV} = \Sigma = \begin{bmatrix} \sigma_0^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_1^2 & 0 & \vdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & \sigma_{N-1}^2 \end{bmatrix} \tag{8}$$

Eq 6 is the same as the following (for individual elements of an OFDM symbol),

$$\hat{X_k} = H_k^* \left(\frac{1}{SNR_k} + |H_k|^2\right)^{-1} Y_k, \quad k = 0, 1, 2, \dots, N-1 \tag{9}$$

1.2. **Estimating the channel taps (Training).** To recover the transmitted data, the channel taps (Λ) are needed. The most comon training method is to allocate some tones to known training data. Channel taps can be estimated using these pilots.

Let usassume that, M (where $1 \leq M \leq N$) out of the N tones in the OFDM symbol $Y(n)$ correspond to a known (training) data (see Figure 7 for such an arrangement of pilots, usually called Comb type pilots). Let us denote, the M pilot positions (tones) as $\{k_1, k_2, \dots, k_M\}$ and the training data containing on these tones as $\{S_1, S_2, \dots, S_M\}$. Using these we can write the linear relationship,

$$\underbrace{\begin{bmatrix} S_1(n) & 0 & 0 & \dots & 0 \\ 0 & S_2(n) & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & S_M(n) \end{bmatrix}}_{\mathbf{S}_p(n)} \underbrace{\begin{bmatrix} \mathbf{1}_{k_1}^* \\ \mathbf{1}_{k_2}^* \\ \mathbf{1}_{k_3}^* \\ \vdots \\ \mathbf{1}_{k_M}^* \end{bmatrix}}_{\triangleq\Gamma} \begin{bmatrix} H_0(n) \\ H_1(n) \\ H_2(n) \\ \vdots \\ H_{L-1}(n) \end{bmatrix} + \begin{bmatrix} V_{k_1}(n) \\ V_{k_2}(n) \\ V_{k_3}(n) \\ \vdots \\ V_{k_M}(n) \end{bmatrix} = \begin{bmatrix} Y_{k_1}(n) \\ Y_{k_2}(n) \\ Y_{k_3}(n) \\ \vdots \\ Y_{k_M}(n) \end{bmatrix} \tag{10}$$

Where $\mathbf{1}_\nu$ is a column vector of size $L \times 1$with all positions except at row number ν are zeros. The ν row contain the value 1 . For example $\mathbf{1}_2$ represents,

$$\mathbf{1}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \mathbf{1}_3^* = \begin{bmatrix} 0 & 0 & 1 & 0 & \dots & 0 \end{bmatrix} \tag{11}$$

$$\mathbf{S}_p(n)\Gamma(n)H(n) + V(n) = Y_p(n) \tag{12}$$

LS solution to this problem would provide the estimate,

$$\hat{H}(n) = \left(\Gamma^*\mathbf{S}_p^*(n)\mathbf{S}_p(n)\Gamma(n)\right)^{-1} \Gamma^*(n)\mathbf{S}_p^*(n)Y_p(n) \tag{13}$$

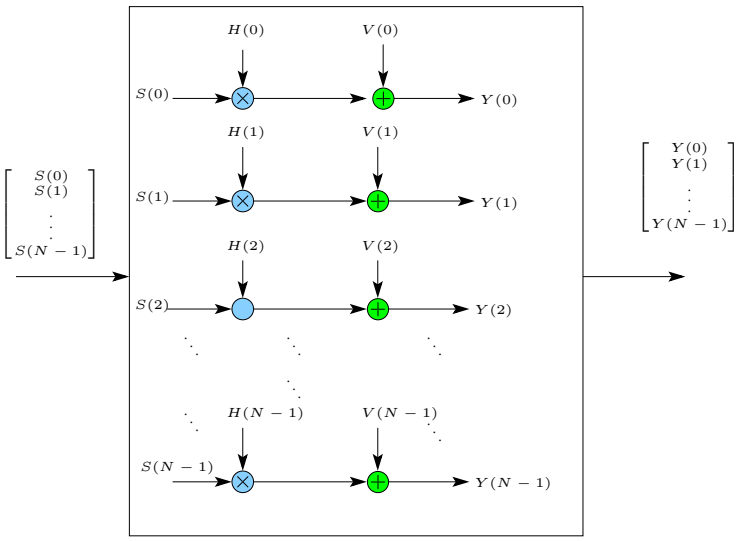

Figure 3. Frequency domain equivalent of OFDM system

From a statistical estimation point of view, the channel estimate is not optimum because, the statistical correlations (channel, input/output) are not considered here. Thus the least square algorithm for channel estimation provide only a sub optimal esimate. However, this is close to the best one can do, when the statistcs of the channel are unknown at the receiver.

Using this method, the channel taps [2] are estimated on every symbol, assuming that some subcarrier positions (one every symbol) are allocated for for pilots. In practice channel is not often estimated on a symbol by symbol scale[3], but instead on every few symbols or so. Similarly, the number of pilot subcarriers on a given (pilot) symbol) decides the closeness of the estimate to the actual channel. The higher the number M of pilot subcarriers, the better the estimate would be. The penalty again is wasting the bandwidth, whcih otherwise could have been used for data transmission.

2. Channel estimation

Most of the practical OFDM based systems have limited number of pilots in the time-frequency grid. It is from these pilots, the channel estimate at all time (symbol index) and frequency(sub carrier index) positions have to be computed. Thus the problem in prestine form resembles to a two dimensional estimation problem. The complex channel gain[4] varies across sub carriers (frequency axis) as well as from symbol to symbol (time axis).

It is much easier to visualise multicarrier systems in frequency domain. The task of channel estimation involves the computation of the channel gain $H(k, n)$ at certain subcarrier k (i.e., $f = k\Delta_f$) for OFDM symbol index n (i.e., n-th OFDM symbol in time).

While there are non-pilot based schemes studied in the literature [Ratna], in this document, we considered only pilot based channel estimation methods. Pilot based schemes uses known[5] (to both transmitter and receiver) data sent at known positions (fixed subcarriers and fixed symbols). At certain positions on the time frequency grid, the data sent is known. These known data is often called as pilot data and their positions as pilot positions.

Channel estimation techniques greatly works for the given pilot arrangement considered. There are different ways these pilots cane be arranged on a 2D grid, based on the channel variations in time and frequency. Some of the commonly used pilot arrangements are shown in Figure **??**, Figure **??** and Figure **??**.

[2]Strictly speaking, the frequency domain channel is estimated here. The time domain tap values are obtained by performing a DFT on them. That is $\hat{h}(n) = DFT(H(n)))$

[3]The channel may not vary (drastically) from one symbol to another symbol, due to which the channel need to be estimated only at intervals of few symbols. Performing channel estimation too often demands extensive computational complexity, besides consuming available bandwidth for pilots. Design of pilots and their positions are thus based on the channel charactersitcs and also the system complexity.

[4]Here channel gain refers to the channel attenuation. In general this value is a complex number, which means both amplitude and phase of the input are distorted by the channel

[5]known both to transmitter as well as receiver

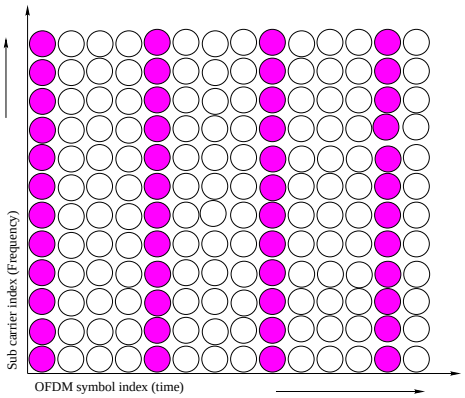

Figure 4. Block type pilot arrangement for OFDM

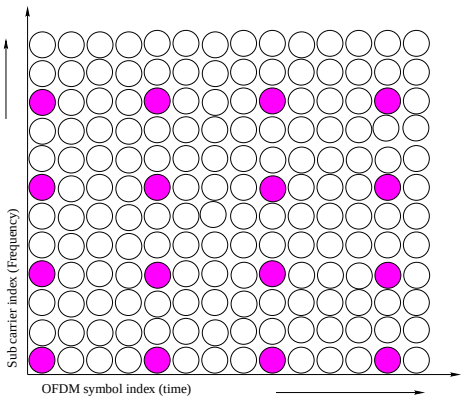

Figure 5. Rectangular arrangement of pilots in OFDM

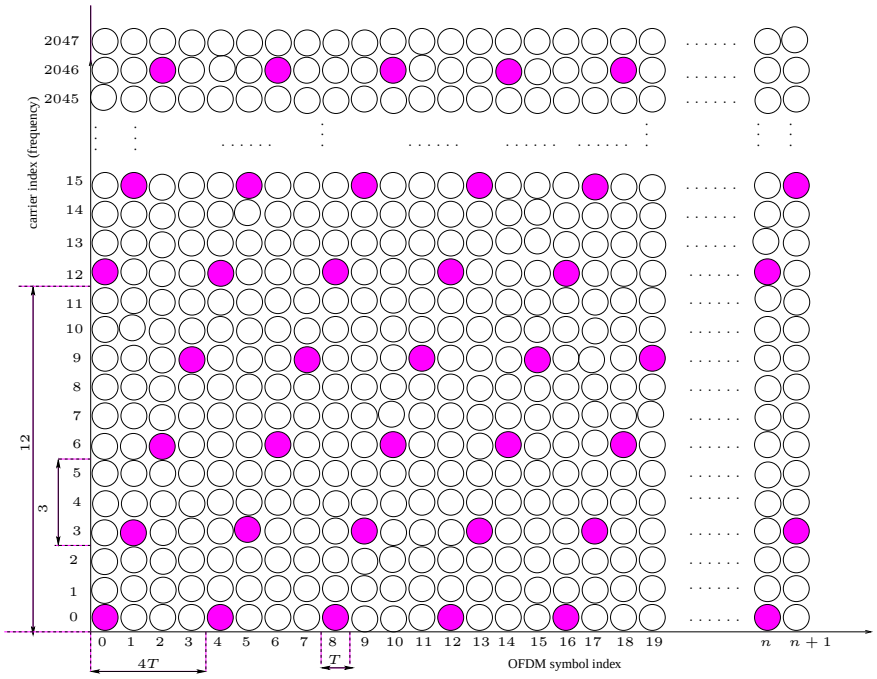

Figure 6. Diagonal Pilots: Used in DVB-T

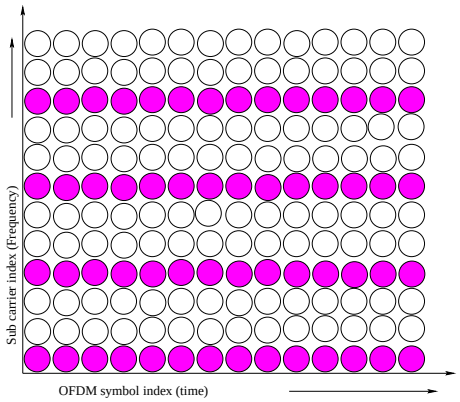

Figure 7. Comb type pilot arrangement for OFDM

These pilots are usually called scattered pilots (in DVB-T parlance for example) to distinguish them from the continuous pilots (also known as Comb type pilots).

At certain positions in time and frequency, the modulation symbols $X(k,n)$ will be replaced by known pilot symbols $S(k,l)$. At these positions, the channel can be measured. Figure **??** shows a rectangular grid with pilot symbols at every third frequency and every fourth time slot (OFDM symbol). The pilot density is thus $1/12$, that is, $1/12$ of the whole capacity is used for channel estimation. This lowers not only the data rate, but also the available energy E_b/N_0 . Both must be taken into account in the evaluation of the spectral and the power efficiency of such a system.

It is useful, when the power on the pilot symbols are higher than that on data. broadly speaking, this helps in better detection and estimation since it can provide higher noise immunity. But they are not good on rectangular pilot arrangements like that in Figure **??** because that would cause a higher average power for every fourth OFDM symbol, which is not desirable for reasons of transmitter implementation (burden on power amplifiers and other analog circuits). A diagonal like Figure **??** however can have this option availed, as indeed done in DVB-T.

In the build up to channel estimation in practical systems (say DVB-T) we will consider a series of idealistic scenarios to illustrate the concept of OFDM channel estimation.

Let us first consider a deterministic channel in noiseless scenario.

2.1. **OFDM channel estimation: Deterministic noiseless channel.** In this case the channel in time domain has an FIR structure with L taps. Stacking up the DFT output (frequency domain representation of channel output) of all the M pilots, we get the vector Y_{pilot}

$$Y_{pilot} = \begin{bmatrix} Y_{p_1} \\ Y_{p_2} \\ Y_{p_3} \\ \vdots \\ \vdots \\ Y_{p_{M-1}} \\ Y_{p_M} \end{bmatrix} = \begin{bmatrix} H_{p_1} S_p \\ H_{p_2} S_p \\ H_{p_3} S_p \\ \vdots \\ \vdots \\ H_{p_{M-1}} S_p \\ H_{p_M} S_p \end{bmatrix} \tag{14}$$

Dividing the DFT output vector Y_{pilot} by the known pilot data, we get the frequency domain channel coefficients $\{H_{p_i}\}$ at all the pilot positions. Once the channel coefficients at the M pilot positions in the two dimensional grid are obtained, the corresponding values at other points (non-pilot positions on the 2D grid) are obtained using 2D-interpolation. If any prior know-hows on the nature of the channel response, that could be used in the interpolation algorithm. When such informations are unavailable, then various interpolation techiques may be used to find the best fit. Simple hold based interpolation, polynomial,spline interpolation and sinc interpolation algorithms are some of the commonly used interpolation schemes. There is a trade off between complexity and closer match (to ideal channel) have a say in chosing a particular interpolation method.

Let us investigate an example channel response to see how different methods fares with respect to channel estimation. Figure 8 shows the exact channel response in time frequency plane. The channel variation in time (for certain fixed

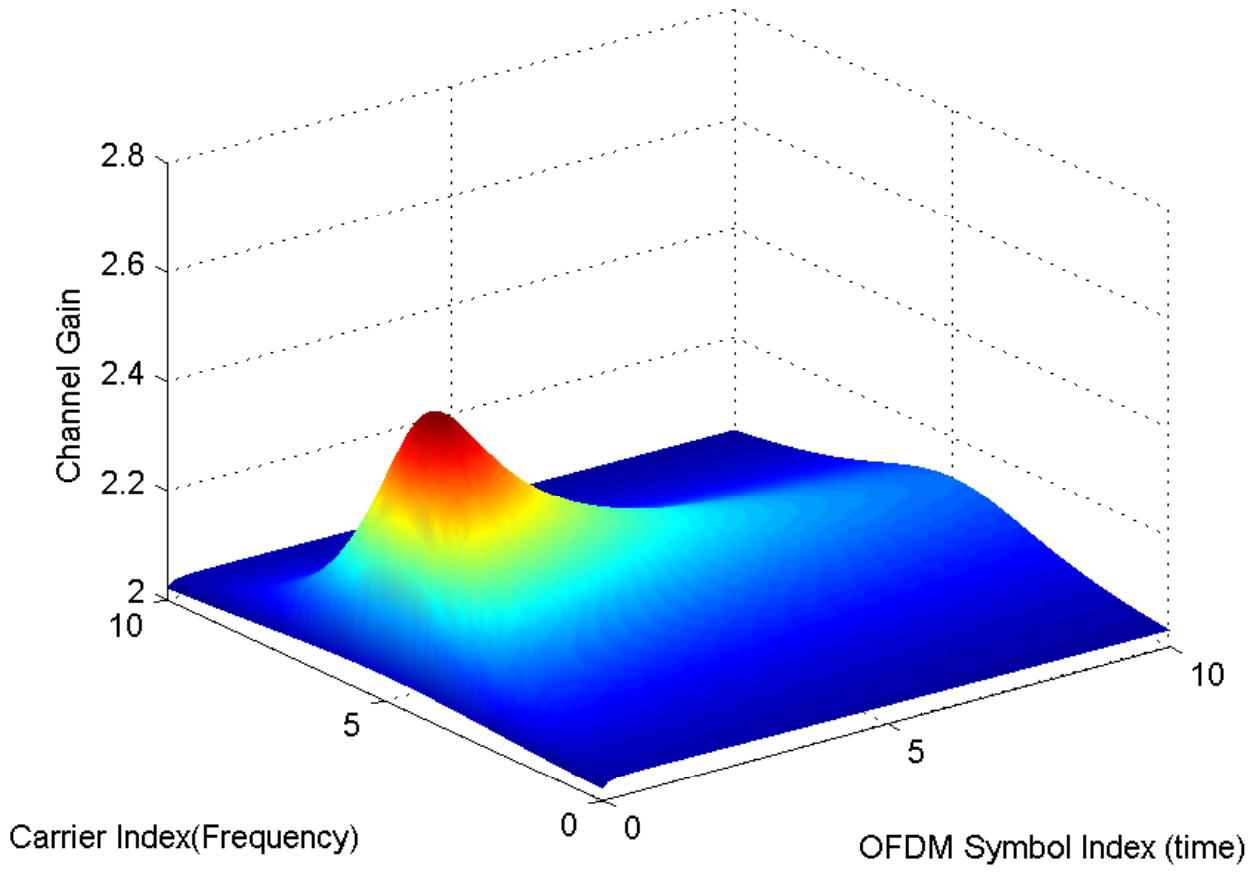

Figure 8. Ideal Channel response

frequency) is shown in Figure 9 and channel variation in frequency (for some fixed time instants) is depicted in Figure 10. Figure 11 illustrates the comparative performance of various 2D interpolation schemes on the channel estimation.

2.2. **Deterministic channel with AWGN.** The 2D-interpolation schemes mentioned in the noiseless case may not yield the best estimate of the channel, in the presence of (the omniprescent) AWGN in the system (channel model). In such cases, some optimality criteria is devised to compute the channel estimate. One commonly used optimality crieria is the minimum least square error criteria alias LS esimate. Here again, first a noisy (rough) estimate of the channel at pilot positions are computed. This is used by the LS estimator to compute the channel gain at all positions in the grid.

As done for the noiseless case, channel estimates at the M pilot positions are estimated first. The DFT output corresponding to these M pilot positions is stacked to form the vector Y_{pilot} given by,

$$Y_{pilot} = \begin{bmatrix} Y_{p_1} \\ Y_{p_2} \\ Y_{p_3} \\ \vdots \\ \vdots \\ Y_{p_{M-1}} \\ Y_{p_M} \end{bmatrix} = \begin{bmatrix} H_{p_1}S_p + N_{p_1} \\ H_{p_2}S_p + N_{p_1} \\ H_{p_3}S_p + N_{p_1} \\ \vdots \\ \vdots \\ H_{p_{M-1}}S_p + N_{p_1} \\ H_{p_M}S_p + N_{p_1} \end{bmatrix} \tag{15}$$

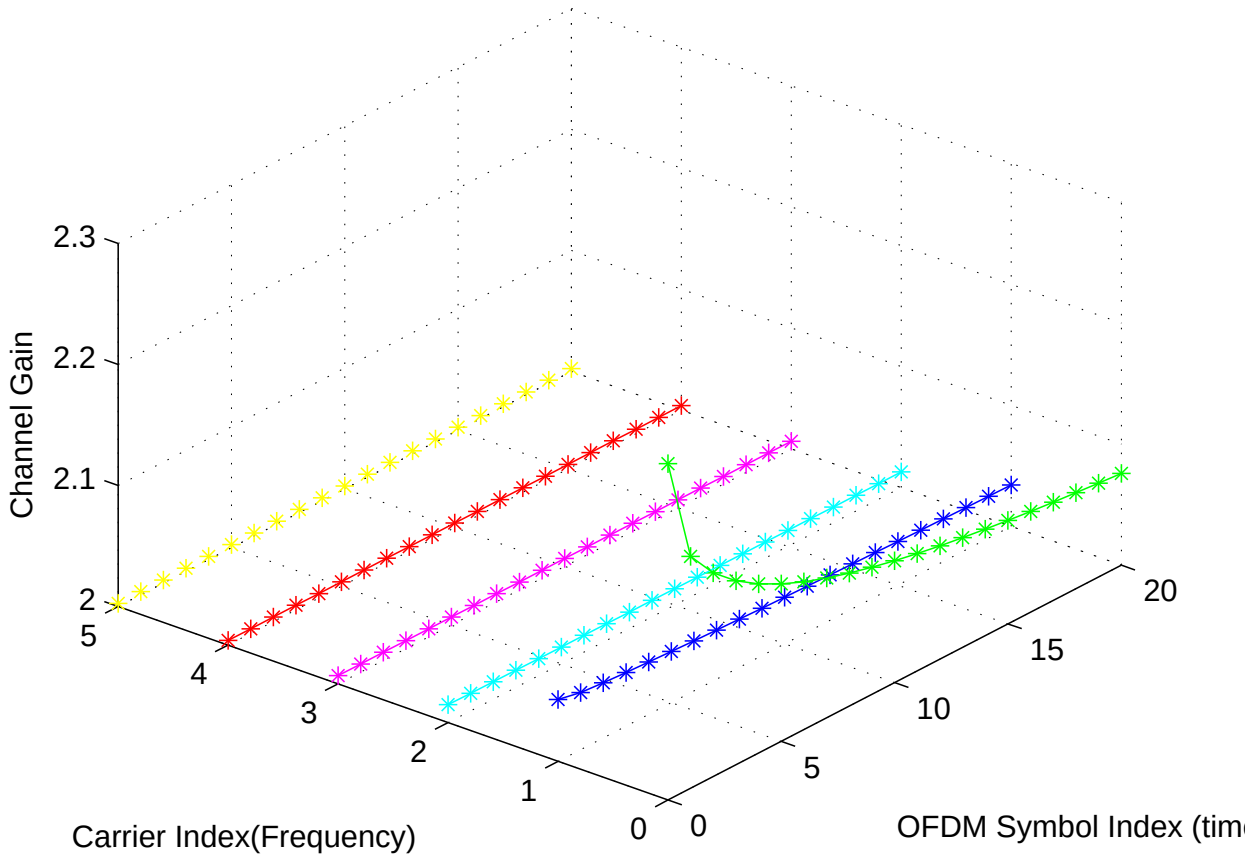

Figure 9. Channel response; Frequency view

On division by S_p, the exact channel response at the pilot cannot be obtained, due to the presence of the additive noise. However, a division is performed and the resultant value is chosen as the initial estimate (noisy nevertheess) at the pilot positions. A conventional 2D interpolation (as done in the noiseless case) may be performed to obtain the channel estimates at all points in the grid, but that is likely to be very crude or atleast sub optimal.A LS estimate result in a 2D filter.

3. Fading channel scenario

Fading is reminiscent of typical wireless communication channel. Such a channel can again be modeled as an FIR filter, with the tap coefficients non-deterministic (random). It is sometimes possible to model the probability distribution function (pdf) of these random coefficients. We will consider a simple fading model without additive noise [4] [5].

Consider the grid of Figure **??** for an OFDM system with carrier spacing $\Delta_f = 1/T = 1$kHz and symbol duration $T_s = 1250\mu s$ [4]. At every third frequency, the channel will be measured once in the time $4T_s = 5ms$, that is, the unknown signal (the time-variant channel) is sampled at the sampling frequency of $200Hz$. For a noise-free channel, we can conclude from the sampling theorem that the signal can be recovered from the samples if the maximum Doppler frequency v_{max} fulfills the condition $v_{max} < 100Hz$ More generally, for a pilot spacing of $4T_s$, the condition $v_{max}T_s < 1/8$ must be fulfilled. In frequency direction, the sample spacing is $3kHz$. From the (frequency domain) sampling theorem, we conclude that the delay power spectrum must be inside an interval of the length of $333\mu s$. Since the guard interval already has the length $250\mu s$, this condition is automatically fulfilled if we can assume that all the echoes lie within the guard interval. We can now start the interpolation (according to the sampling theorem) either in time or in frequency direction and then calculate the interpolated values for the other direction.

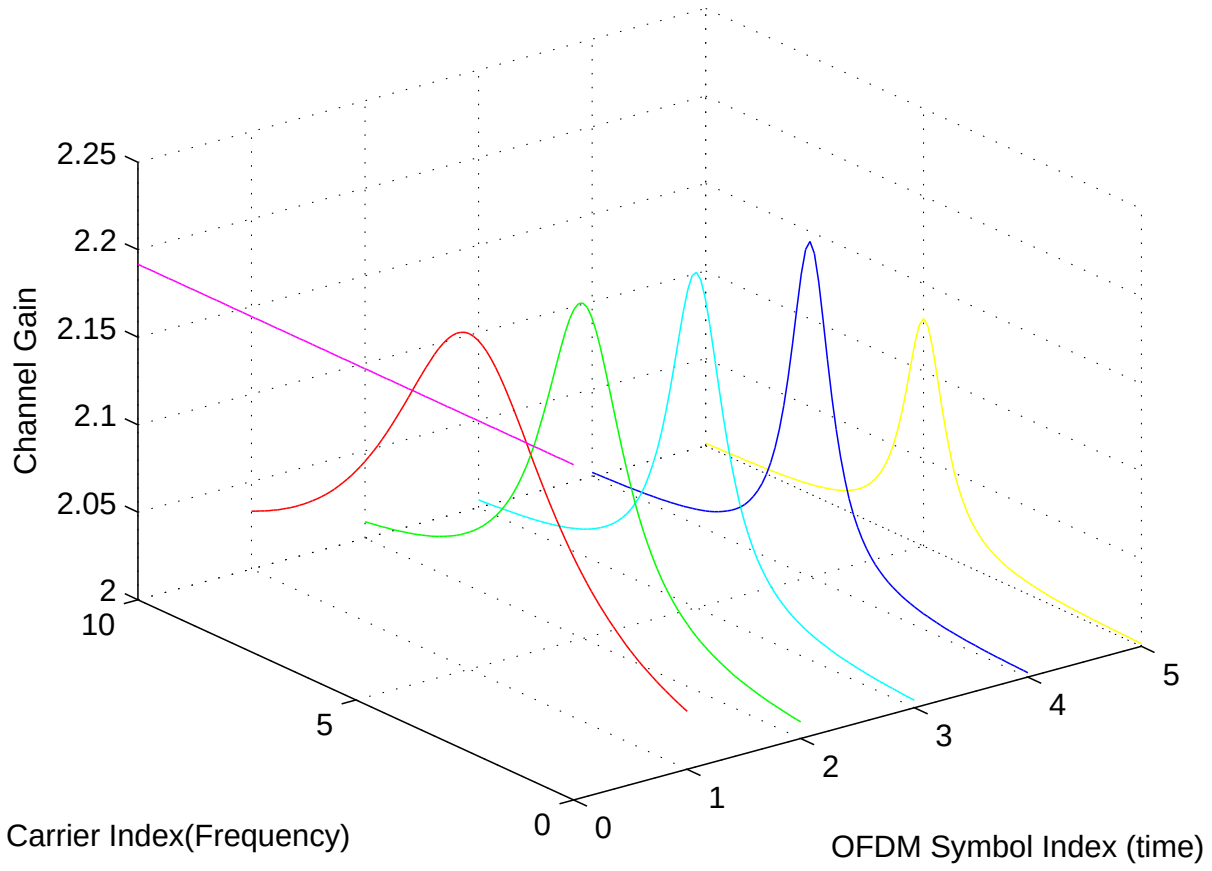

Figure 10. Channel response; Time view

Simpler interpolations are possible and may be used in practice for a very coherent channel, for example, linear interpolation or piecewise constant approximation. However, for a really time-variant and frequency-selective channel, these methods are not adequate.

For a noisy channel, even the interpolation given by the sampling theorem is not the best choice because the noise is not taken into account. The optimum linear estimator will be derived in the next subsection.

The density of the grid has to be matched to the incoherency of the channel, that is, to the time-frequency fluctuations described by the scattering function. To illustrate this by a

4. Optimum Channel estimation for a wireless channel

Practical wireless channels are hostile. It is not possible to unquley model a wireless channel that represent all practical communication links. Wireless channels are time varying and can be frequency selective. Generally speaking, such channels are modeled by a random process with a scattering function of time and frequency.

Optimum estimate in the minimum mean square error (MMSE) sense often leads to nonlinear solution (except in some special cases). However a linear MMSE (suboptimal, yet the best MMSE under linearity constraint) can be wroked out. It is found that such an estimate for a pilot based OFDM channel estimation leads to 2D Wiener filter. A block diagram of such an estimator is shown in Figure 12. The estimator for a general pilot arrangement is derived below.

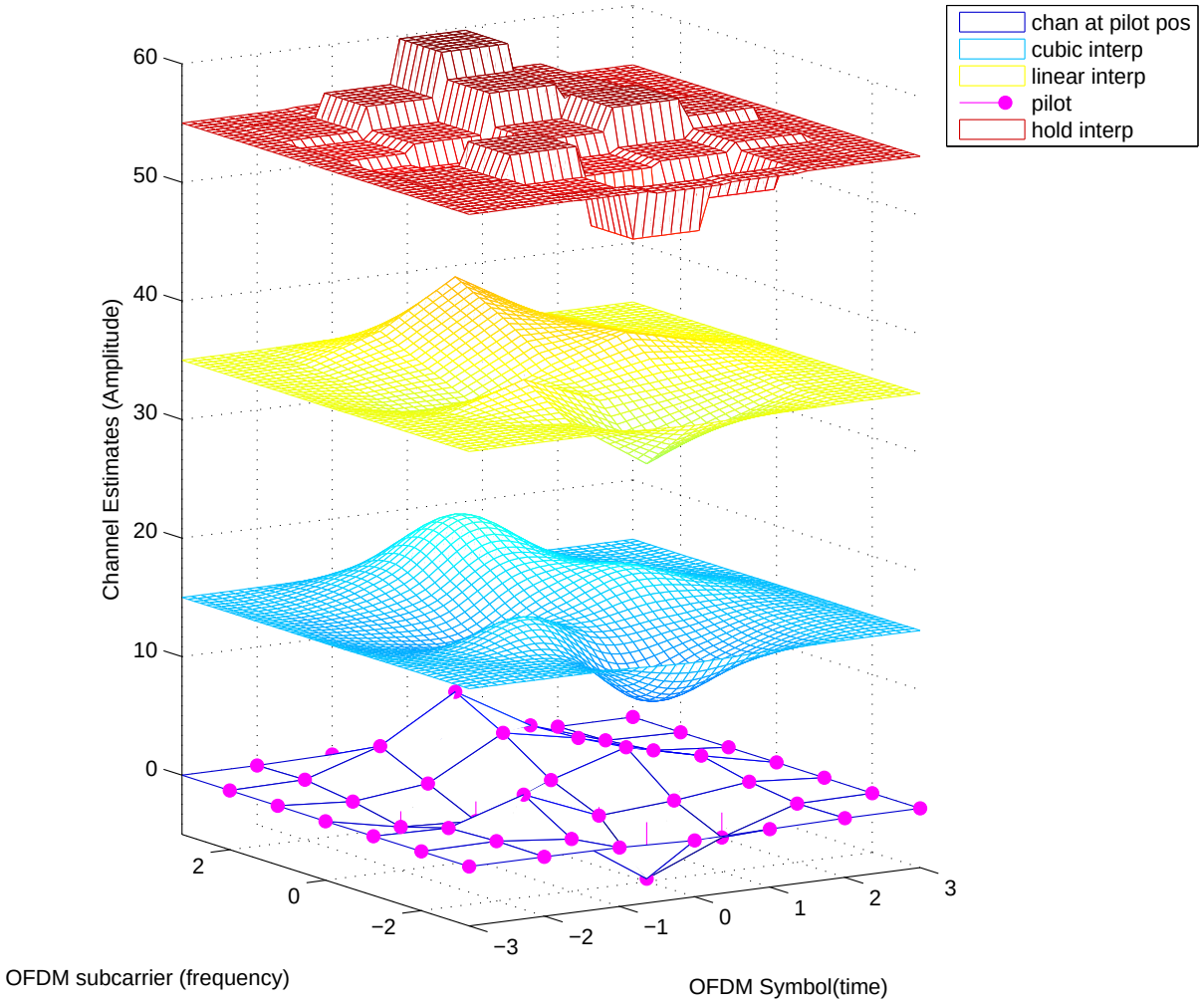

Figure 11. DVB-T Channel estimation using 2-D interpolation

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \dots & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \dots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \dots & \dots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \dots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & \ddots & 1 & 0 \end{bmatrix}}_{\Gamma} \begin{bmatrix} H(0\Delta_f, 0T_s) \\ H(0\Delta_f, T_s) \\ H(0\Delta_f, 2T_s) \\ H(0\Delta_f, 3T_s) \\ H(0, 4T_s) \\ H(0\Delta_f, 5T_s) \\ \cdots \\ \ddots \\ \vdots \\ H((K-1)\Delta_f, (N-3)T_s) \\ H((K-1)\Delta_f, (N-2)T_s) \\ H((K-1)\Delta_f, (N-1)T_s) \end{bmatrix} + \underbrace{\begin{bmatrix} V_1/S_1 \\ V_2/S_2 \\ V_3/S_3 \\ \vdots \\ \ddots \\ \vdots \\ V_{M-1}/S_{M-1} \\ V_M/S_M \end{bmatrix}}_{V_p} = \underbrace{\begin{bmatrix} Y_1/S_1 \\ Y_2/S_2 \\ Y_3/S_3 \\ \vdots \\ \ddots \\ \vdots \\ Y_{M-1}/S_{M-1} \\ Y_M/S_M \end{bmatrix}}_{Y_p}$$

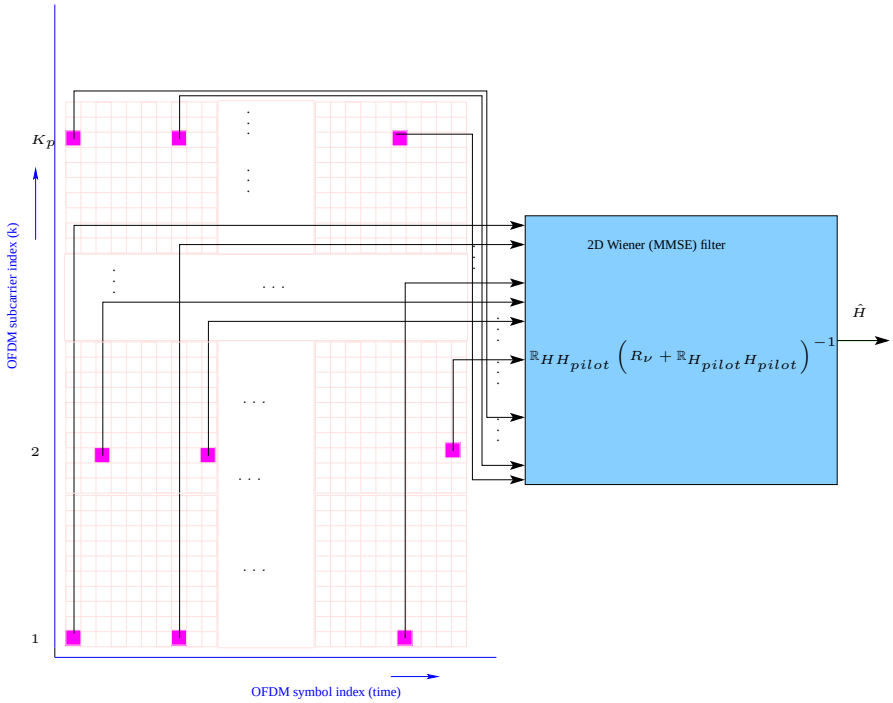

Figure 12. 2D MMSE channel estimator

or simply,

$$\Gamma H + \Psi_p = \tilde{H}_p \tag{16}$$

Here $\tilde{H}_p$ is the noisy channel estimate at the pilot positions, given by,

$$\tilde{H}_p = \begin{bmatrix} \tilde{H}_1 \\ \tilde{H}_2 \\ \tilde{H}_3 \\ \vdots \\ \ddots \\ \vdots \\ \tilde{H}_{M-1} \\ \tilde{H}_M \end{bmatrix} = \begin{bmatrix} Y_1/S_1 \\ Y_2/S_2 \\ Y_3/S_3 \\ \vdots \\ \ddots \\ \vdots \\ Y_{M-1}/S_{M-1} \\ Y_M/S_M \end{bmatrix} \tag{17}$$

The M pilot positions in the channel grid $\{H_1^p, H_2^p, \ldots\ldots H_M^p\}$ are shown in red colour.

$$\begin{bmatrix} H(0\Delta_f, 0T_s) \rightarrow H_1^p \\ H(0\Delta_f, T_s) \\ H(0\Delta_f, 2T_s) \\ H(0\Delta_f, 3T_s) \\ H(0, 4T_s) \rightarrow H_2^p \\ H(0\Delta_f, 5T_s) \\ \cdots \\ \ddots \\ \vdots \\ H((K-1)\Delta_f, (N-3)T_s) \\ H((K-1)\Delta_f, (N-2)T_s) \rightarrow H_M^p \\ H((K-1)\Delta_f, (N-1)T_s) \end{bmatrix} \tag{18}$$

From Eq 16 a linear MMSE estimate can be computed as,

The correlation matrix R_H is given by

Let us denote the M pilot positions using a vector pair, $(k_1^p, n_1^p), (k_2^p, n_2^p), (k_3^p, n_3^p), \ldots, \ldots, (k_M^p, n_M^p)$, with $n_j^p, k_j^p \in \mathbb{Z}$ and $0 \leq k_j^p \leq K-1, 0 \leq n_j^p \leq N-1$, If we were to think of the pilot positions in a one dimensional vector form, the pilot positions would be at $p_1, p_2, p_3, \ldots, \ldots, p_M$, where $p_j \in \mathbb{Z}, 1 \leq p_j \leq M$. It is easy to establish the relationship

$$p_j = 1 + \left(k_j^p - 1\right) K + k_j^p \tag{19}$$

We will denote the channel estimates at the pilot positions as a single array (vector) H_{pilot},

$$H_{\text{pilot}} = \begin{bmatrix} H(0\Delta_f, 0T_s) \rightarrow H_{p_1} \\ H(0\Delta_f, 4T_s) \rightarrow H_{p_2} \\ H(0\Delta_f, 8T_s) \rightarrow H_{p_3} \\ \cdots \\ \ddots \\ \vdots \\ H(\Delta_f, T_s) \\ H(\Delta_f, 5T_s) \\ H(\Delta_f, 9T_s) \\ \cdots \\ \ddots \\ \vdots \\ H((K-1)\Delta_f, (N-6)T_s) \\ H((K-1)\Delta_f, (N-2)T_s) \end{bmatrix} \tag{20}$$

In general for arbitrary pilot positioning, we can write this vector (using the pilot position notation p_i) as,

$$H_{\text{pilot}} = \begin{bmatrix} H(k_1^p\Delta_f, n_1^pT_s) \rightarrow H_{p_1} \\ H(k_2^p\Delta_f, n_2^pT_s) \rightarrow H_{p_2} \\ H(k_3^p\Delta_f, n_3^pT_s) \rightarrow H_{p_3} \\ \cdots \\ \ddots \\ \vdots \\ H(k_j^p\Delta_f, n_j^pT_s) \rightarrow H_{p_j} \\ \cdots \\ \ddots \\ \vdots \\ H(k_{M-1}^p\Delta_f, n_{M-1}^pT_s) \rightarrow H_{p_{M-1}} \\ H(k_M^p\Delta_f, n_M^pT_s) \rightarrow H_{p_M} \end{bmatrix} \tag{21}$$

For the DVB-T case, the first few pilots (scattered pilots) in the frequency-time grid are positioned at $(0,0), (0,4), (0,8), \ldots, \ldots, (3,1)$, and so on. That is $\rightarrow H_{p_1} = H(0\Delta_f, 0T_s)$, $\rightarrow H_{p_2} = H(0\Delta_f, 4T_s)$, $\rightarrow H_{p_3} = H(0\Delta_f, 8T_s)$ etc.,

More generally,

$$\mathbb{R}_H C^* = \begin{bmatrix} \text{Col}_{p_1}(\mathbb{R}_H) & \text{Col}_{p_2}(\mathbb{R}_H) & \text{Col}_{p_3}(\mathbb{R}_H) & \ldots & \ldots \text{Col}_{p_M}(\mathbb{R}_H) \end{bmatrix} \tag{22}$$

Where, $\text{Col}_{p_i}(\mathbb{R}_H)$ is the p_i-th column vector of the channel correlation matrix $\mathbb{R}_H$.

R

The matrix $\mathbb{R}_H C^*$ has dimension $L \times M$. Mathematically, $\mathbb{R}_H C^* \in \mathbb{C}^{L \times M}$.

We find that $\mathbb{R}_H C^*$ is nothing but the cross-correlation matrix formed by correlating H with H_{pilot}. That is.,

$$\mathbb{R}_H \Gamma^* = E\left[H H_{\text{pilot}}\right] \tag{23}$$

Similarly,

$$\Gamma \mathbb{R}_H \Gamma^* = \begin{bmatrix} \text{Row}_{p1}(\mathbb{R}_H \Gamma^*) \\ \text{Row}_{p2}(\mathbb{R}_H \Gamma^*) \\ \text{Row}_{p3}(\mathbb{R}_H \Gamma^*) \\ \ddots \\ \vdots \\ \ddots \\ \text{Row}_{p_M}(\mathbb{R}_H \Gamma^*) \end{bmatrix} \tag{24}$$

In other words,

$$\Gamma \mathbb{R}_H \Gamma^* = E\left[H_{\text{pilot}} H_{\text{pilot}}^*\right] \tag{25}$$

For our special case (DVB-T pilot positions),

$$\Gamma\mathbb{R}_H\Gamma^* = \begin{bmatrix} R_H(0\Delta_f, 0T_s) & R(0\Delta_f, -4T_s) & \ddots & R_H(-3\Delta_f, -5T_s) \\ R_H(0\Delta_f, 4T_s) & R(0\Delta_f, 0T_s) & \ddots & R_H(-3\Delta_f, -1T_s) \\ R_H(0\Delta_f, 8T_s) & R(0\Delta_f, 4T_s) & \ddots & R_H(-3\Delta_f, 3T_s) \\ \vdots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ R_H(3\Delta_f, T_s) & R(3\Delta_f, -3T_s) & \ddots & R_H(0\Delta_f, -4T_s) \\ R_H(3\Delta_f, 5T_s) & R(3\Delta_f, 1T_s) & \ddots & R_H(0\Delta_f, 0T_s) \\ \vdots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ R_H(k_M^p\Delta_f, n_M^pT_s) & R(k_M^p\Delta_f, (n_M^p-4)T_s) & \ddots & R_H((k_M^p-3)\Delta_f, (n_M^p-5)T_s) \end{bmatrix} \tag{26}$$

$$\hat{H}_{\text{LMMSE}} = \mathbb{R}_H\Gamma^* \left(R_\nu + \Gamma\mathbb{R}_H\Gamma^*\right)^{-1} \tilde{H} \tag{27}$$

In AWGN case, R_ν becomes a diagonal matrix with entries of noise to pilot power ratio. In this case, $\nu = V/S$ (awgn noise samples scaled down by the pilot values) Eq 28 becomes,

$$\hat{H}_{\text{LMMSE}} = \mathbb{R}_H\Gamma^* \left(D + \Gamma\mathbb{R}_H\Gamma^*\right)^{-1} \tilde{H} \tag{28}$$

where

$$D = diag(\sigma_1^2/E_s, \sigma_2^2/E_s, \ldots, \sigma_{M-1}^2/E_s) \tag{29}$$

$\tilde{H} = H_p + \nu$ is the vector with noisy estimates at pilot positions.

While the optimum channel estimate is mathematically pleasing, the heavy computational burden often force such techniques from implementation. Moreover, the estimate requires the time-frequency correlation (scattering functions) of the channel. While the time correlations and frequency correlations are some times individually known, the 2D scattering functions are difficult to model. If the 2D correlations can be decomposed to two 1D correlations (one in time and one in frequency), then the optimum 2D filtering can be simplified to two separate 1D estimation filters. To retain optimality, it is mandatory that the decomposition in Figure 13 hold good. This is very rarely met for arbitrary channel models.

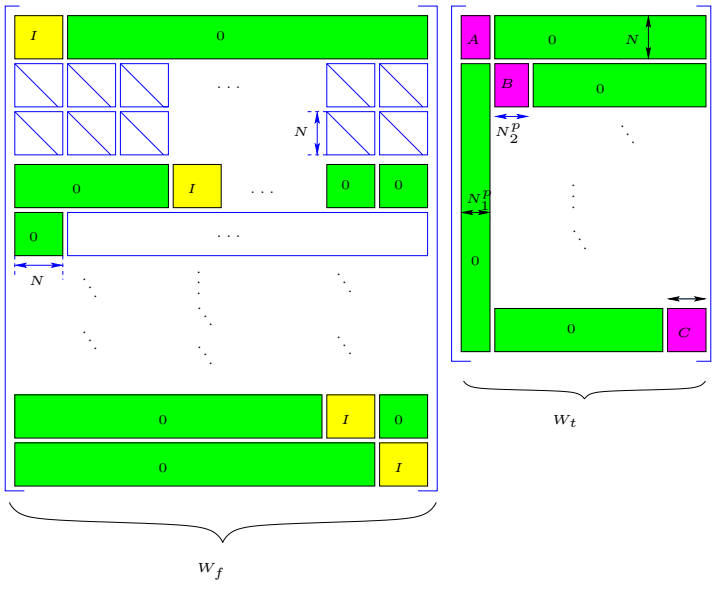

Figure 13. 2D to 1D matrix decomposition

Since the optimum 2D filter is prohibitively complex, simplified estimates are tried out. One obvious method is to have individual 1D filtering. In this method, first a filtering across time (for a fixed subcarrier pilot) is performed, followed by a frequency domain 1D filtering. The order of filtering (whether time interpolation to be done first or frequency domain interpolation) is decided by the pilot positioning. In rectangular, symmetric pilot arrangements, the order doesnt alter performance, but for scattered pilots (diagonal schemes for intance) the order does matter. Figure 14 illustrates the $2 \times 1D$ filtering concept.

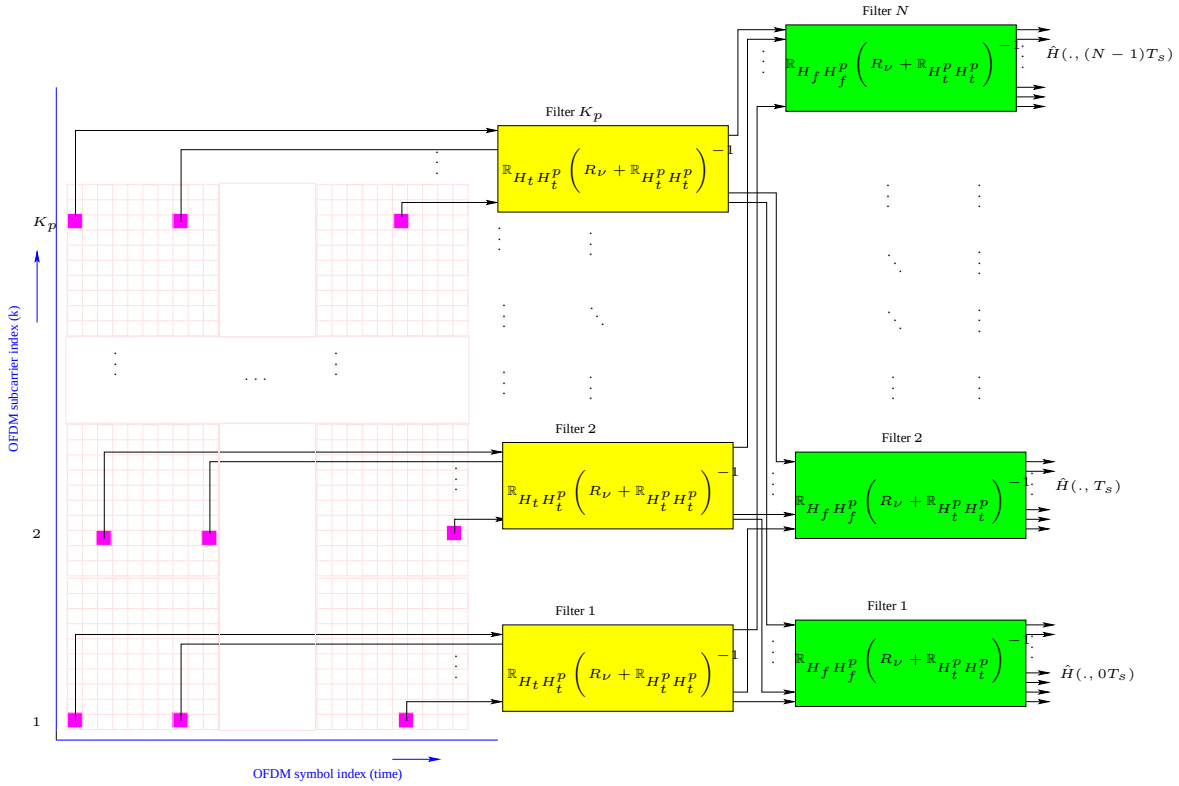

Figure 14. 2 by 1D Wiener

The advantage of doing individual 1D filtering besides, the computational simplification is that, the correlation matrices can be precomputed, for some channel models. The time autocorrelation function can be derived from the Doppler spectrum (inverse Fourier transform of Doppler spectrum gives the autocorrelation function in time), whereas the frequency correlation can be infered from the delay power spectrum.
(Ratna: Some examples to be given)

5. Channel estimation implementation

The section from here on wards are not for public use.

6. Pilot pattern in DVB-T

In DVB-T, two kind of pilots are used. Even though, there is no difference in signal properties (or power), they are distinguished by the respective names, due to their arrangements on the time-frequency (2D) OFDM grid.

scattered pilots are distributed on the 2D grid in a diagonal fashion. Their time (OFDM symbol index n) and frequency (sub carrier index k) form the following simple relationship.

$$k = k_{\text{min}} + 3(n \mod 4) + 12p, \quad n, k, p \in \mathbb{Z}, k_{\text{min}} \leq k \leq k_{\text{max}}, n \in [0, 67] \tag{30}$$

Where k_{min} and k_{max} are pre defined values for the mode of operation. They are different for 2K and 8K modes. k_{min} value is 171 and 687 for 2K and 8K modes respectively. The corresponding k_{max} values are 687 and 7503 in respective order.(Ratna:Values need to be added/verified). Since one frame composite of 68 symbols, the range of values n take is limited in the interval $[0, 67]$

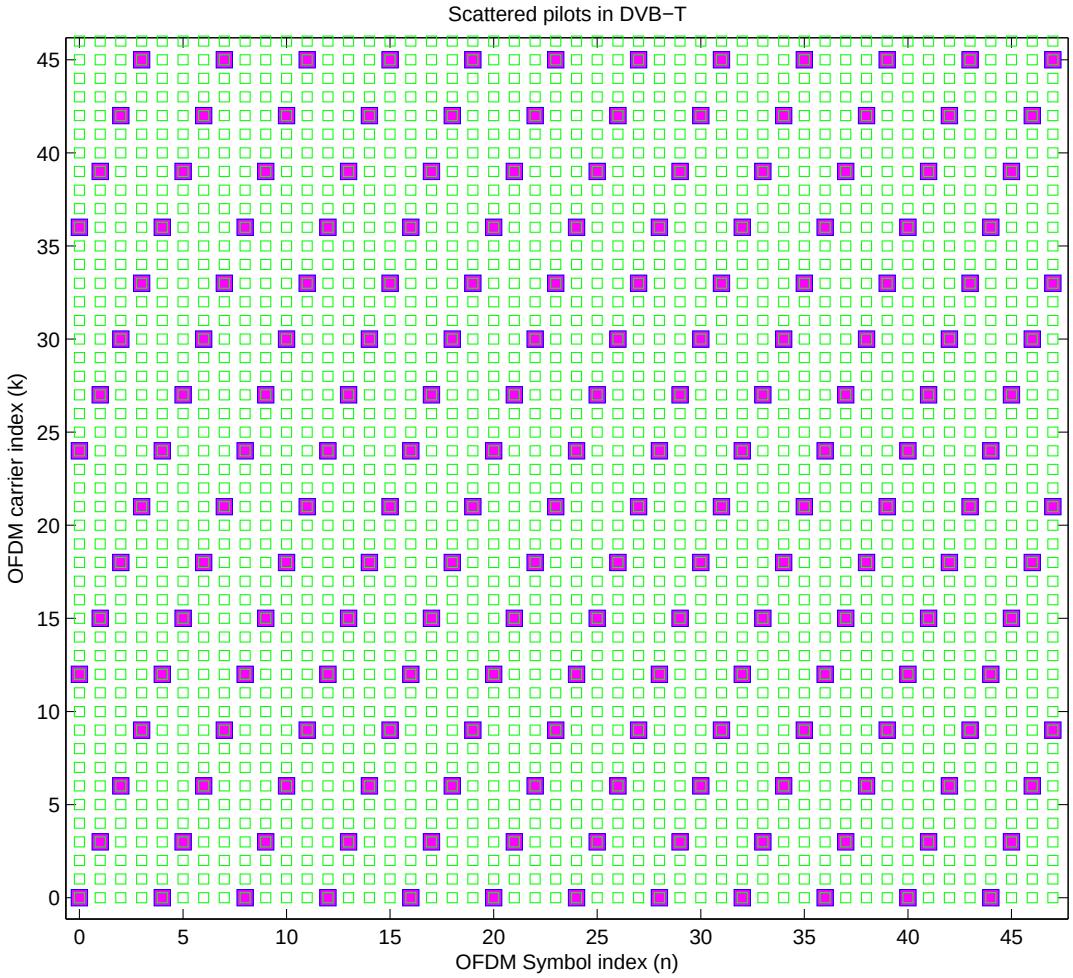

Figure 15. DVB-T scattered pilots: Scattered pilots shown in magenta marker. Data carriers are shown in green. Only a segment of the complete time frequency grid shown

Mode	Scattered Pilot Carrier index
2K	0 48 54 87 141 156 192 201 255 279 282 333 432 450 483 525 531 618 636 714 759 765 780 804 873 888 918 939 942 969 984 1050 1101 1107 1110 1137 1140 1146 1206 1269 1323 1377 1491 1683 1704
8K	0 48 54 87 141 156 192 201 201 255 279 279 282 282 333 333 432 432 450 450 483 483 525 525 531 531 618 618 636 636 714 714 759 759 765 765 780 780 804 804 873 873 888 888 918 918 939 939 942 942 969 969 984 984 1050 1050 1101 1101 1107 1107 1110 1110 1137 1137 1140 1140 1146 1146 1206 1206 1269 1269 1323 1323 1377 1377 1491 1491 1683 1683 1704 1704 1752 1758 1791 1845 1860 1896 1905 1959 1983 1986 2037 2136 2154 2187 2229 2235 2322 2340 2418 2463 2469 2484 2508 2577 2592 2622 2643 2646 2673 2688 2754 2805 2811 2814 2841 2844 2850 2910 2973 3027 3081 3195 3387 3408 3456 3462 3495 3549 3564 3600 3609 3663 3687 3690 3741 3840 3858 3891 3933 3939 4026 4044 4122 4167 4173 4188 4212 4281 4296 4326 4347 4350 4377 4392 4458 4509 4515 4518 4545 4548 4554 4614 4677 4731 4785 4899 5091 5112 5160 5166 5199 5253 5268 5304 5313 5367 5391 5394 5445 5544 5562 5595 5637 5643 5730 5748 5826 5871 5877 5892 5916 5985 6000 6030 6051 6054 6081 6096 6162 6213 6219 6222 6249 6252 6258 6318 6381 6435 6489 6603 6795 6816

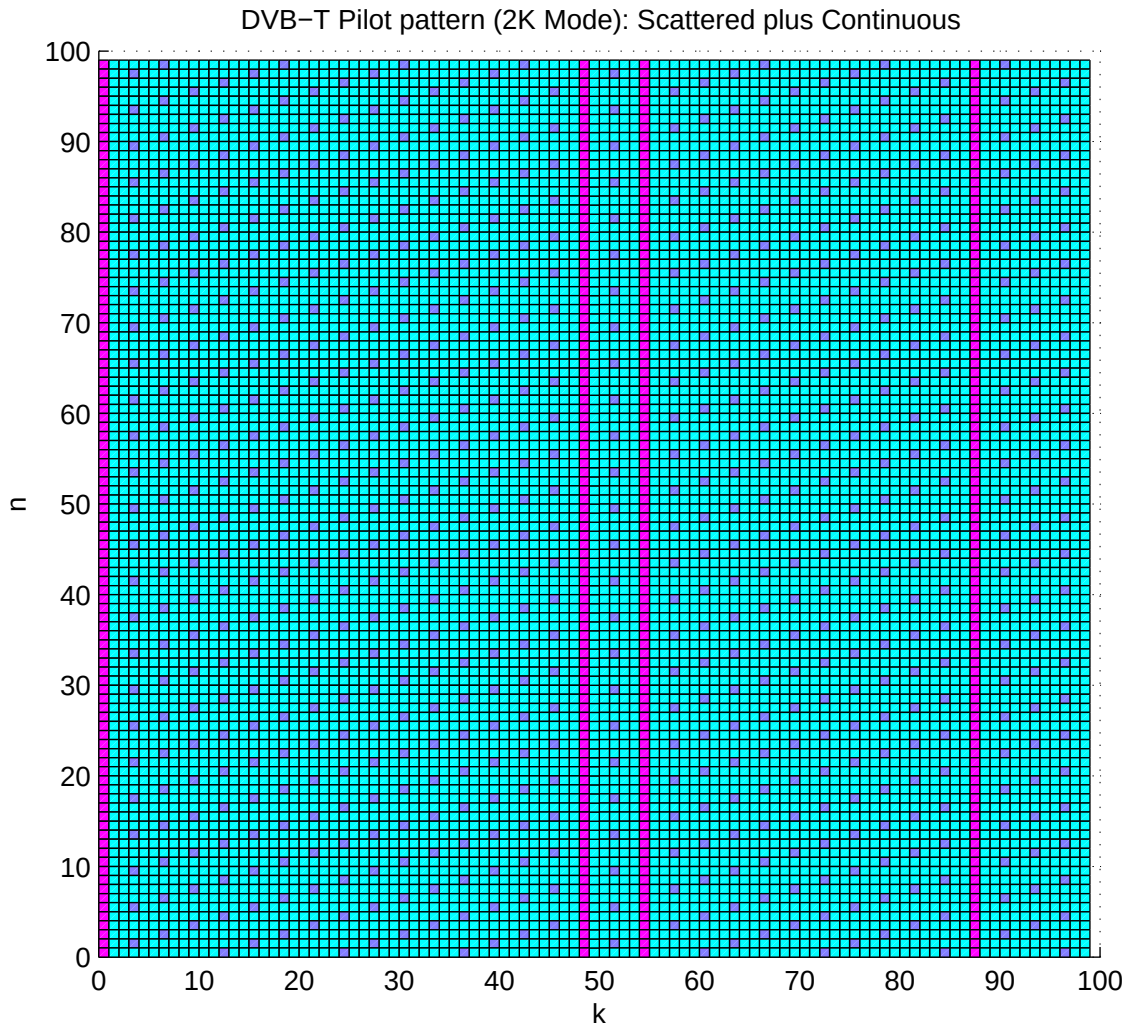

Figure 16. DVB-T pilots: Scattered pilots shown in magenta and data carriers in green. Continuous pilots shown in blue markers.It may be observed that continuous pilots may coincided with scattered pilots at the continuous pilot tones. Only a segment of the time frequency grid shown in figure.

7. Appendix B: Linear Prediction

$$\hat{u}(n+1) = \sum_{k=0}^{N-1} w(k)u(n-k) \tag{31}$$

where $w(k)$ for $k = 0, 1, 2, \ldots, N-1$ are the coefficients of the prediction filter. The linear predictor may be cast into the Wiener filtering problem, where the desired response $d(n)$ set equal to $x(n+1)$

The Wiener filtering setup for (linear) prediction has the Wiener-Hopf system of equations as,

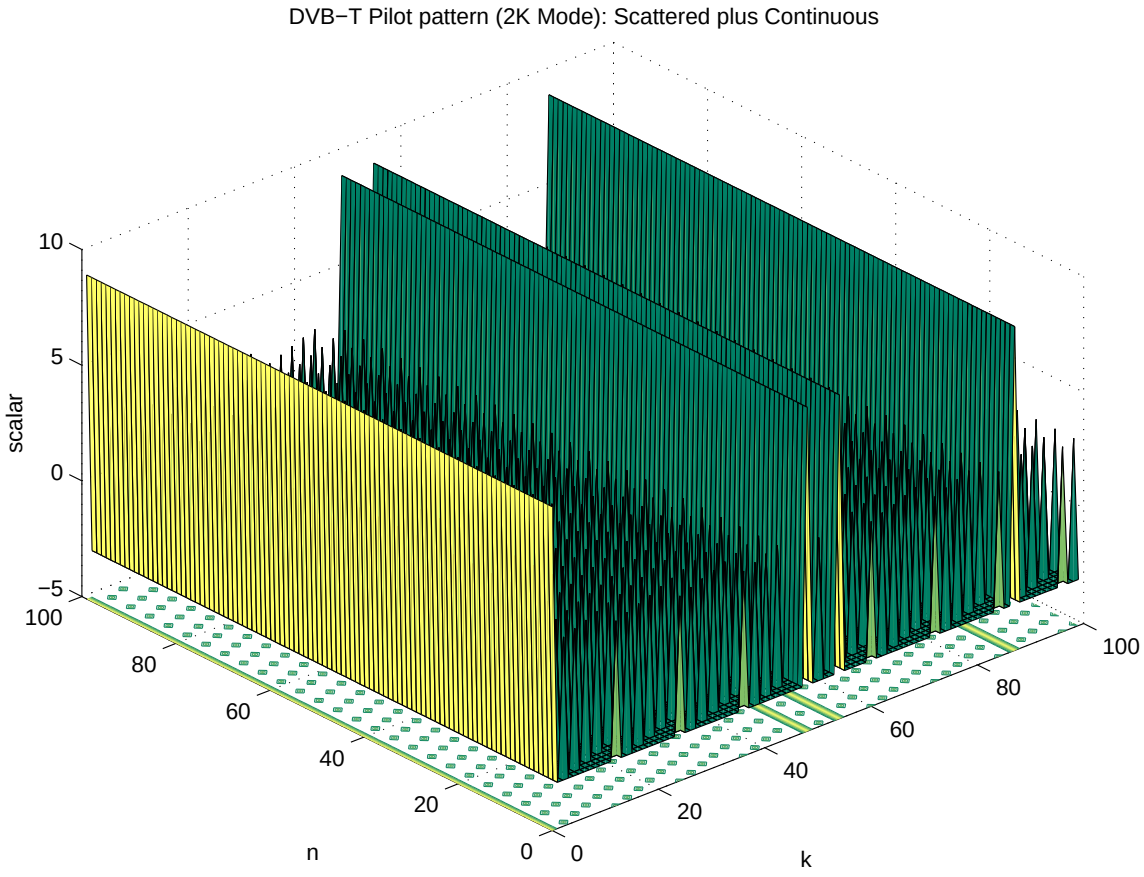

Figure 17. DVB-T pilots: Scattered pilots shown in magenta and data carriers in green. Continuous pilots shown in blue markers.It may be observed that continuous pilots may coincided with scattered pilots at the continuous pilot tones. Only a segment of the time frequency grid shown in figure.

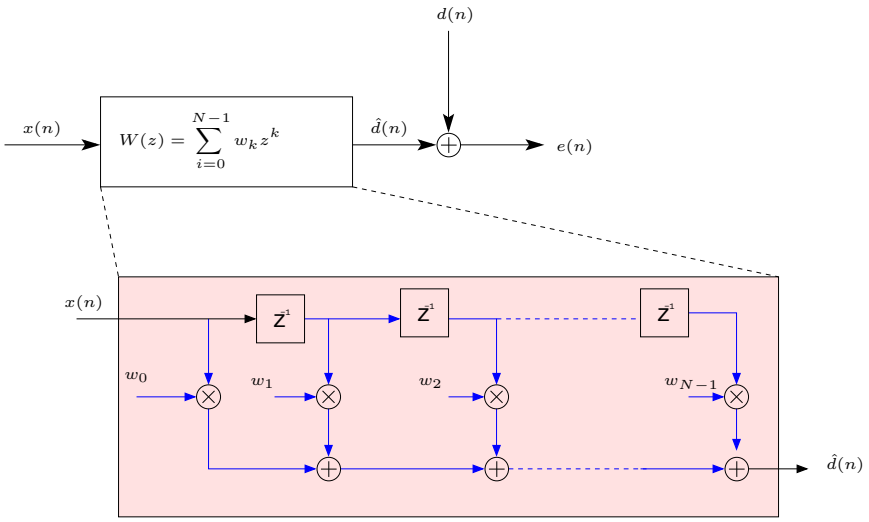

Figure 18. Wiener filter setup: Prediction

$$\underbrace{\begin{bmatrix} r_x(0) & r_x^*(1) & r_x^*(2) & \dots & r_x^*(N-1) \\ r_x(1) & r_x(0) & r_x^*1 & \dots & r_x^*(N-2) \\ r_x(2) & r_x(1) & r_x(0) & \dots & r_x^*(N-3) \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ r_x(N-1) & r_x(N-2) & r_x(N-3) & \dots & r_x(0) \end{bmatrix}}_{\mathbb{R}_{xx}} \underbrace{\begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_{N-1} \end{bmatrix}}_{w} = \underbrace{\begin{bmatrix} r_x(1) \\ r_x(2) \\ r_x(3) \\ \vdots \\ r_x(N) \end{bmatrix}}_{r_{dx}}$$

Optimum (Wiener/LMMSE) solution is,

$$\hat{w} = \mathbb{R}_{xx}^{-1} r_{dx} \tag{32}$$

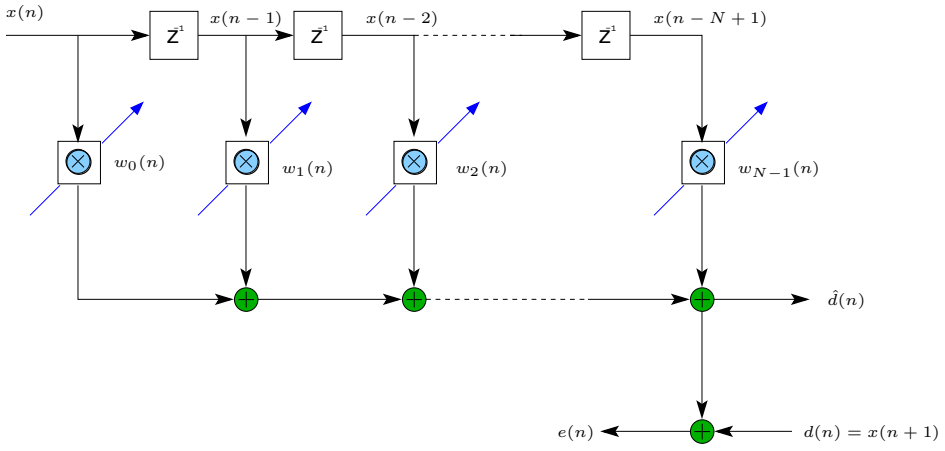

Figure 19. LMS for adaptive prediction

The LMS weight vector adaptation equation is given by,

$$w(n+1) = w(n) + \mu E\left[e(n)x^*(n)\right] \tag{33}$$

Individual tap update (say tap weight w_i) can be written similarly,

$$w_i(n+1) = w_i(n) + \mu E\left[e(n)x^*(n-i)\right] \tag{34}$$

NLMS (Normalised LMS) algorithm has a normalization[6] factor in the adaptation. NLMS update equation can thus be written as follows.

$$w(n+1) = w(n) + \mu \frac{x^*(n)}{\|x(n)\|^2} e(n) \tag{35}$$

$$w_i(n+1) = w_i(n) + \mu \frac{x^*(n-i)}{\|x(n-i)\|^2} e(n) \tag{36}$$

To aoid the gradient noise amplification (when $x(n)$ is high), a slight modified version is used.

$$w(n+1) = w(n) + \mu \frac{x^*(n)}{\epsilon + \|x(n)\|^2} e(n) \tag{37}$$

$$w_i(n+1) = w_i(n) + \mu \frac{x^*(n-i)}{\epsilon + \|x(n-i)\|^2} e(n) \tag{38}$$

8. SIMULINK IMPLEMENTATION

TBD later

[6]Here the adaptation coefficient μ is not the same as that used in LMS. Same notation is used to limit the usage of too many symbols

REFERENCES

[1] ETSI Standard: EN 300 744 V1.5.1 (2004-11), Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television, download from ETSI.
[2] Goldsmith, Andreas ., *Wireless Communications*, (Cambridge University Press, 2005).
[3] Tse, David and Viswanath, Pramod *Fundamentals of Wireless Communication*, (Cambridge University Press, 2005).
[4] Henrik Schulze and Christian Luders, Theory and applications of OFDM and CDMA : wideband wireless communications, Chichester, West Sussex, England Hoboken, NJ : John Wiley, 2005.
[5] Khaled Fazel, Stefan Kaiser, Multi-carrier and Spread Spectrum Systems, John Wiley and Sons, 2003, ISBN 0470848995, 9780470848999

DEMOD IP TECHNOLOGY GROUP
E-mail address: rethna.p@gnss.com
E-mail address: ratnuu@gmail.com