# Anexos

**TECHNISCHE UNIVERSITÄT DRESDEN**

**FAKULTÄT ELEKTROTECHNIK UND INFORMATIONSTECHNIK**

**Vodafone Chair Mobile Communications Systems**

# Final Project

Title: Laguerre filtering applied to long impulse responses of ultra wide band channels

Submitted by: Javier Jalle Ibarra

Born on: 28.10.1988 in Zaragoza, Spain

Supervisor: Dr.-Ing. habil. Wolfgang Rave
Responsible Professor: Prof.Dr.-Ing. Gerhard Fettweis
Submitted on: November 16$^{th}$, 2011

## Acknowledgement

I hereby declare that the present work submitted to the examination board from the Fakultät Elektrotechnik und Informationstechnik with the Final Project entitled

**Laguerre filtering applied to long impulse responses of ultra wide band channels**

that has been completely written by me, the undersigner, and that I have not used any other document other than the specified means, materials and references. All references and resources used by me in this work are properly cited within it.

Dresden, November 16$^{\text{th}}$, 2011

Javier Jalle Ibarra

## Abstract

Channel estimation is the approximation of the discrete impulse response of an unknown channel by using the minimum mean square error criterion (MSE). In a classic approach, this approximation can be obtained by adaptively adjusting the weights of a finite impulse response (FIR) transversal filter, also known as 'training'.

This study is a theoretical approach to characterize the viability of the Laguerre filters in a UWB communication system.

Channel estimation in ultra wide band (UWB) channels has to deal with long impulse responses due to the high sampling rate and usually the energy is concentrated in a small fraction of time intervals. With a classical approach with a FIR transversal filter, a large order filter is needed for correctly approximating the long impulse response and most of the filter coefficients gather little or no energy because of the concentration of the energy in small fractions of time intervals.

To understand the source of this problem it has to be explained that an approximation is made by representing a function as a weighted sum of a complete orthonormal set (basis functions), and then truncating these sum to a fixed number of terms (filter order). In the classical FIR approach, these complete orthonormal set is the Kronecker delta ($\delta(k-j)$), also known as canonical basis. The problems with UWB channels arise because of the extremely short time spread of the FIR basis functions.

In the present study, the approximation by means of another complete orthonormal set, the Laguerre sequences, which are a compromise between FIR and infinite impulse response (IIR) systems and can be seen as a generalization of the FIR filters. These sequences are used to estimate the channel response of some UWB-test channels and realizations of the UWB stochastic model given by the IEEE 802.15.4a standard, showing that they offer a better performance, in terms of MSE, as a generalization of FIR filters, however this improvement is not very significant when the channel presents abrupt changes, due to the low-pass behavior of Laguerre filters, as occurs in UWB channels.

Adaptive methods, such as the recursive least squares (RLS), are also covered in the study.

## Resumen

Por estimación de canales entendemos la aproximación de un canal desconocido usando el criterio de mínimo error cuadrático medio (MSE). En un enfoque clásico, esta aproximación es obtenida ajustando de manera adaptativa los coeficientes de un filtro FIR transversal, también conocido como 'entrenamiento'.

La estimación de canales en canales de banda ultra ancha (UWB) tiene que hacer frente a respuestas impulsionales largas debido a la alta frecuencia de muestreo y usualmente la energía está concentrada en una pequeña fracción de los intervalos de tiempo. En un enfoque clásico, usando un filtro transversal FIR, se requiere un orden elevado para poder estimar de manera correcta la larga respuesta impulsional y la mayoría de los coeficientes del filtro reúnen poca o ninguna energía debido a la concentración de la energía en pequeñas fracciones de intervalos de tiempo.

La fuente de este problema viene del hecho de que una aproximación consiste en representar una función como una suma ponderada de una base ortonormal completa (también llamadas funciones base), y luego truncar esta suma a un número fijo de términos (equivalente al orden del filtro). En el enfoque clásico del filtro FIR transversal, las funciones base se corresponden con la delta de Kronecker ($\delta(k-j)$), también conocida como la base canónica. El problema con los canales UWB surgen debido a extremadamente corta extensión temporal de las funciones base de los filtros FIR.

En este estudio, la aproximación es llevada a cabo por medio de otra base ortonormal, las secuencias de Laguerre, las cuales forman un compromiso entre los sistemas FIR y los IIR; y además pueden ser consideradas como una generalización de los filtros FIR. Estas secuencias son utilizadas para la estimación del canal para varios canales de prueba en el ámbito UWB y para realizaciones del modelo estocástico de canales UWB proporcionado por el estándar IEEE 802.15.4a, mostrando que ofrecen mejor rendimiento, en términos del error cuadrático medoi (MSE), como generalización de los filtros FIR, sin embargo esta mejora no es muy significante cuando el canal presenta cambios abruptos, debido al comportamiento paso bajo de los filtros de Laguerre, como ocurre en el caso de canales UWB.

Métodos adaptativos, como el algoritmo RLS; y el problema relativo a la ecualización de canales también están cubiertos en el estudio.

Este estudio es una aproximación teórica para caracterizar la viabilidad de los filtros de Laguerre en un sistema de comunicación UWB.

# Contents

# List of Tables

# List of Figures

# Symbols

| | | |
|---|---|---|
| $\mathbb{N}_0$ | ..... | set of natural numbers and zero |
| $\ell^2(\mathbb{N}_0)$ | ..... | Hilbert space defined on the set $\mathbb{N}_0$ |
| $.^T$ | ..... | matrix transposion |
| $.^*$ | ..... | complex conjugation for scalars |
| $.^H$ | ..... | Hermitian transposition |
| j | ..... | imaginary unit defined as $\sqrt{-1}$ |
| $\mathbf{x}$ | ..... | boldface letter denotes a vector variable |
| $\mathbf{X}$ | ..... | boldface capital letter denotes a matrix |
| $x$ | ..... | letter in normal font denotes a scalar in Euclidean space |
| $\langle \mathbf{x}, \mathbf{y} \rangle$ | ..... | scalar product or dot product between $\mathbf{x}$ and $\mathbf{y}$ |
| $\mathsf{E}\mathbf{x}$ | ..... | expected value of the random variable $\mathbf{x}$ |
| $\mathbf{x} \perp \mathbf{y}$ | ..... | orthogonal variables $\mathbf{x}$ and $\mathbf{y}$ (i.e., $\langle \mathbf{x}, \mathbf{y} \rangle = 0$) |
| $a \stackrel{\triangle}{=} b$ | ..... | quantity $a$ is defined as $b$ |
| $\Re(x)$ | ..... | real part of $x$ |
| $\Im(x)$ | ..... | imaginary part of $x$ |
| $0$ | ..... | zero scalar, vector, or matrix |
| $z^{-1}$ | ..... | unit-time delay operator |
| $t$ | ..... | continuous time variable |
| $k$ | ..... | discrete time variable |
| $X(z)$ | ..... | bilateral z-transform of a scalar sequence $\{x(k)\}$ |
| $X(e^{\mathrm{j}\omega})$ | ..... | discrete-time Fourier transform of $\{x(k)\}$ |
| $x(k)$ | ..... | channel input signal |
| $h(k)$ | ..... | channel impulse response |
| $H(z)$ | ..... | channel transfer function |
| $y(k)$ | ..... | channel output signal |
| $M$ | ..... | order of the filter/approximation |
| $\hat{y}_M(k)$ | ..... | approximation/estimation of signal $y(k)$ of order $M$ |
| $\hat{h}_M(k)$ | ..... | approximation/estimation of system $h$ of order $M$ |
| $\mathbf{x}_M(k)$ | ..... | input data vector or regression vector |
| $w_{M,i}$ | ..... | $i$-th filter weight of a filter of order $M$ |
| $\mathbf{w}_n$ | ..... | filter weights vector of a filter of order $M$ |
| $\sigma_y^2$ | ..... | variance of signal $y$ |

# List of Abbreviations

| | | |
|---|---|---|
| FIR | ..... | Finite Impulse Response |
| IIR | ..... | Infinite Impulse Response |
| CIR | ..... | Channel Impulse Response |
| UWB | ..... | Ultra Wide Band |
| m.m.s.e. | ..... | Minimum Mean Square Error criterion |
| MSE | ..... | Mean Square Error |
| NMSE | ..... | Normalized Mean Square Error |
| LMS | ..... | Least Mean Squares |
| NLMS | ..... | Normalized Least Mean Squares |
| RLS | ..... | Recursive Least Squares |
| LOS | ..... | Line Of Sight |
| NLOS | ..... | Non Line Of Sight |
| Fig. | ..... | Figure |
| NaN | ..... | Not-a-Number |
| ISI | ..... | InterSimbol Interference |
| SNR | ..... | Signal-to-Noise Ratio |
| AWGN | ..... | Additive White Gaussian Noise |
| DTFT | ..... | Discrete Time Fourier Transform |
| ROC | ..... | Region Of Convergence |
| TU Dresden | ..... | Technische Universität Dresden |

# 1. Introduction

## 1.1. Motivation

### 1.1.1. Presentation of the idea

The radio channels in wireless communications suffer from multipath propagation and fading. Multipath is caused by the propagation phenomenon that makes the radio signals to reach the receiving device by two or more paths, causing a constructive or destructive interference which causes a fading effect and intersymbol interference (ISI). Fading can be caused by multipath or shadowing from obstacles. To remove ISI from the received signal, many kinds of equalizers can be used, however most of them require knowledge of the channel impulse response (CIR). To estimate the CIR usually a channel estimator based on a known input sequence is used. A criterion is needed to evaluate the estimation, which in this case it will be the minimum mean square error (m.m.s.e.). In particular, in this study ultra wide band (UWB) channels are estimated, which are characterized by long impulse responses, due to the high sampling rates and usually the energy is concentrated in a small fraction of time intervals or samples. As an example, in Figure 1.1 an UWB channel realization of the IEEE 802.15.4a model typical for Non Line-Of-Sight(NLOS) residential area environments is represented.

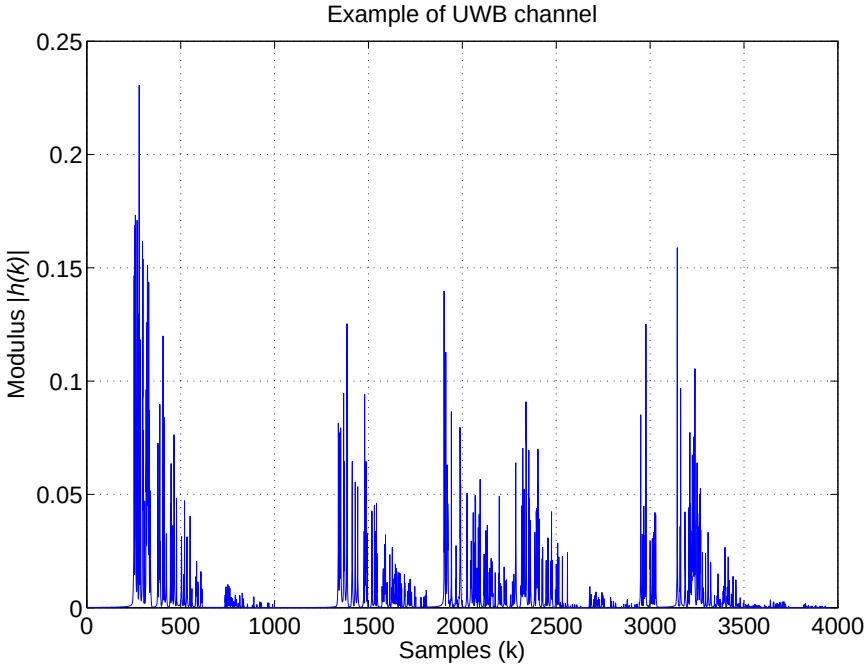



Figure 1.1.: Impulse response of an UWB channel given by the realization of a stochastic model according to [8].

### 1.1.2. Description of the problem

Classically (see [11], [3]) the channel estimation (also known as system identification or channel approximation) is made by adaptively adjusting the weights of a transversal filter. This process is known as 'training' (see []). A finite impulse response (FIR) transversal filter is shown in Figure 1.2. If we consider $x(k)$ as the input of the system, and $w_{M,i}$ as the $i$-th filter weight of a filter of order $M$, the estimated output of the system is given by

$$\hat{y}_M(k) = \sum_{i=0}^{M} w_{M,i} x(k-i) \tag{1.1}$$

and the resulting estimated impulse response (assuming the stationary solution of the filter weights) is the output of the system when fed with an impulse input (Kronecker's delta $\delta(k-i)$), say

$$\hat{h}_M(k) = \sum_{i=0}^{M} w_{M,i} \delta(k-i) \tag{1.2}$$

Taking a look at the resulting estimated impulse response, it can be seen that it corresponds

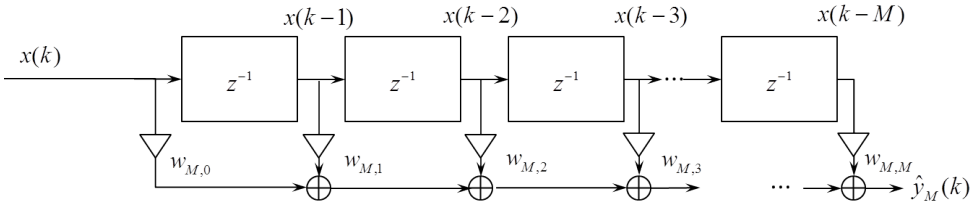



Figure 1.2.: Classic transversal filter

to a finite impulse response (FIR) filter and each filter weight corresponds to a coefficient for one delta function. This causes that when a long impulse response is needed, a large number of filter weights are needed. As an example, let's consider the idealized UWB test channel impulse response of Figure 1.3 which satisfies the two main characteristics of the UWB channels, say, long impulse response and energy concentrated in a small fraction of time intervals. With a classical FIR filter approach as described before, we would need at least some 400 taps for correctly gathering the energy of the CIR. Here come to light two of the main problems of this kind of approach, one is that a large order filter is needed for correctly gathering the energy and the second one is that most of these filter taps contain no energy because the energy is not homogeneously distributed.

The source of this problem is the extremely short time spread of the Kronecker's delta functions, which are the FIR basis functions if the channel impulse response is modeled as an element of the Hilbert space of square integrable functions, $\ell^2(\mathbb{N}_0)$, and the Kronecker's delta function as the basis functions that span the space, i.e., every function of the space can be expressed as a weighted sum of this set, say

$$h(k) = \sum_{i=0}^{+\infty} c_i \delta(k-i). \tag{1.3}$$

An approximation is made by truncating this sum to a fixed number of terms

$$\hat{h}_M(k) = \sum_{i=0}^{M} c_i \delta(k-i)$$

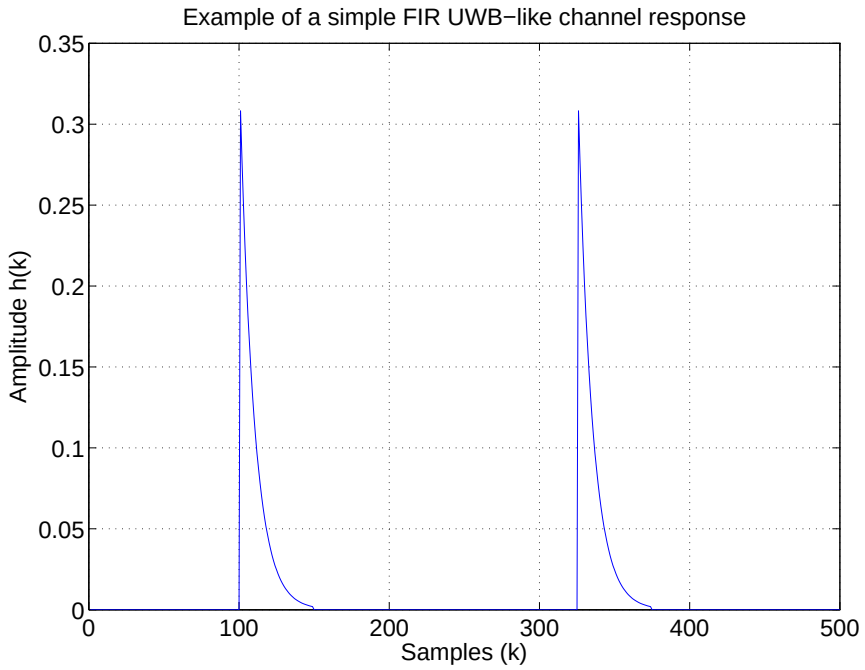



Figure 1.3.: Example of a FIR response that represents in a simple way the featured channel response

which corresponds to the output of an FIR transversal filter of order $M$.

### 1.1.3. Proposed solution

Following with the idea of the basis functions, it has been said that the problem of the estimation of UWB channels with a classic FIR filter is that a large order filter is needed due to the extremely short time spread of the FIR basis functions, the Kronecker's delta. There are many orthonormal sets each of them is appropriate for a specific kind of functions. Examples of commonly used basis functions are

| | |
|---|---|
| $\delta(k-i)$ | Kronecker's delta (Canonical basis) |
| $e^{jwki}$ | Complex exponentials (Fourier series) |

The canonical basis is appropriate for functions with short time spread, functions localized in time, and the Fourier series is appropriate for functions with small frequency spread, localized in frequency. It is clear that these two sets of basis functions represent the two extreme cases, the delta functions have extremely short time spread and infinite frequency spread and the complex exponentials have infinite time spread and extremely short frequency spread. The idea that emerges, is that another set of basis functions in between these two cases can be used with longer time spread in order to reduce the filter order.

The first solution seems to be infinite impulse response (IIR) filters, that can be used to approximate/estimate these impulse responses. However IIR filters have their own problems, the most important are multimodal error surfaces and instability (see [5]). Multimodal error surfaces means that there are local minima which are not global minima, and the problem with instability makes it necessary to develop methods for watching the stability of the system and recovering it when the system is instable.

Laguerre filters are a compromise between FIR and IIR filters. The main reason for that argument is that they have (as it will be seen later on Chapter 2) only one adjustable multiple pole, which allows to keep the stability by keeping this pole inside the unit circle (a generalization of Laguerre filters can be made by adjusting each pole separately instead of one unique pole which leads to a multivariate problem). Another reason is that there are highly efficient RLS Laguerre lattice algorithms available (see [6], [7]). In general, standard RLS algorithms have a complexity $O(M^2)$ per iteration and lattice algorithms only $O(M)$ ($M$ order of the filter), which makes Laguerre filters a feasible option in comparison to standard FIR filters.

## 1.2. Aim of the Work

The aim of this study is to approximate the impulse response of an UWB-like channel, with a Laguerre filter. The quality of the approximation is evaluated using the minimum mean square error (m.m.s.e.) criterion. Ideally we want to achieve an acceptable MSE (mean square error) with a lower order filter as it would be needed with a FIR approach, in order to reduce the computational cost.

## 1.3. Brief Chapters Overview

In Chapter 2, orthonormal sets are introduced, in particular the Laguerre sequences. In chapter 3, the concept of a Laguerre spectrum is presented to analyze how the energy distributes through the Laguerre sequences for a given CIR. Chapter 4 talks about the used filters and how the channel estimation is carried out.
Chapters 5 and 6 present the channel estimation for two idealized channels. Chapter 7 presents the channel estimation with the UWB IEEE 802.15.4a channel model.
Chapter 8 deals with the Laguerre adaptive methods.

# 2. Discrete orthonormal sequences

## 2.1. Introduction

The representation of functions in terms of a sum of orthonormal functions has been shown to be quite useful in many areas. The most widely used example is Fourier analysis, using orthonormalized exponential functions for the terms in the series. The function $x(t)$ (a continuous function of time), is expanded in terms of the complex exponentials in a least-squares sense

$$x(t) = \sum_{i=-\infty}^{+\infty} c_n e^{\mathrm{j}nt}.$$

with

$$c_n = \frac{1}{2\pi} \int_{-\infty}^{+\infty} x(t) e^{-\mathrm{j}nt}\, dt.$$

and

$$\mathrm{j} = \sqrt{-1}$$

One of the advantages of the Fourier analysis is that there are efficient algorithms to implement it, like the FFT (Fast Fourier Transform).

In practice, the infinite sum is truncated to a finite sum of terms by accepting some error in the representation. This error is evaluated in some sense with a criterion, and the resulting $\mathrm{M}^{\mathrm{th}}$-truncated weighted sum is an approximation $x_M(t)$ of the function $x(t)$. The approximation is considered to be representative of $x(t)$ and is used in further analysis. The main advantage of this kind of representation is that (if everything has been done efficiently and the basis functions fit to the function) the number of coefficients is small but still closely representative.

## 2.2. Orthonormal sequences

This section presents common system theory concepts found in [12], [9] or [13]. All signals in real systems are finite energy transients, aperiodic and zero for all values lower than an arbitrary starting time denoted as $t = 0$. With the boom of digital systems, due to their low cost and large scale integration, most of these signals are sampled in order to produce a digital signal. The samples of a signal are defined as those values of the continuous signal attained at the discrete times $t = 0, T, 2T, \cdots$. That is to say $x(t = 0)$, $x(t = T)$, $x(t = 2T)$, $\cdots = x(0)$, $x(1)$, $x(2)$, $\cdots$ . The discrete signal consist of an ordered set of these samples and the set of preceding zeros which can be written as either $\cdots, 0, 0, x(0), x(1), x(2), \cdots$ or more compactly as a set $\{x(k)\} = x(k)$ where $k$ ranges from $-\infty$ to $+\infty$, and is referred as the discrete time variable. This kind of set is called a sequence, which is defined in mathematics as an ordered list of elements.

This sequence can be represented in terms of a set of discrete orthonormal sequences $\{\phi_i(k)\}$ where the index $i = 0, 1, 2, \cdots$ denotes the index of the function set. The expansion is given by

$$x(k) = \sum_{i=0}^{+\infty} c_i \phi_i(k) \tag{2.1}$$

and

$$c_i = \langle \phi_i(k), x(k) \rangle \tag{2.2}$$

where $\langle \cdot, \cdot \rangle$ is defined as the dot product or scalar product of two sequences that gives as result a scalar value, given by

$$\langle f(k), g(k) \rangle = \sum_{k=-\infty}^{+\infty} f(k) g^*(k). \tag{2.3}$$

An approximation is made by truncating this sum to a fixed number of terms

$$\hat{x}_M(k) = \sum_{i=0}^{M} c_i \phi_i(k). \tag{2.4}$$

### 2.2.1. Requirements for orthogonality

A set of sequences $\{\phi_i(k)\},\ i = 0,\ 1,\ 2,\ \cdots$ is orthonormal if the scalar product of any sequence of the set with another one has a value of zero and the scalar product of any sequence of the set with itself has a value of unity. Expressed mathematically we have

$$\langle \phi_i(k), \phi_j(k) \rangle = \sum_{k=-\infty}^{+\infty} \phi_i(k) \phi_j^*(k) = \delta_{ij} \tag{2.5}$$

with $\delta_{ij}$ is called the Kronecker's delta and is defined by

$$\delta_{ij} = \begin{cases} 1 & : i = j \\ 0 & : i \neq j \end{cases}$$

in a geometric interpretation the principle of orthogonality means that every function of the set adds a new orthogonal dimension to the space spanned by the set of functions.

## 2.3. Sequences in z-Domain

The $z$-Domain is a complex frequency-domain representation. The $z$-transform converts a discrete sequence into a complex frequency-domain representation, which is given by

$$X(z) = \mathcal{Z}\{x(k)\} = \sum_{k=-\infty}^{+\infty} x(k) z^{-k} \tag{2.6}$$

where $z$ is, in general, a complex number

$$z = A e^{\mathrm{j}\varphi}$$

where $A$ is the magnitude of $z$, j is the imaginary unit, and $\varphi$ is the phase in radians. Note that when evaluated on the unit circle, i.e., for $z = e^{\mathrm{j}\omega}$ the discrete-time Fourier transform (DTFT) can be found, which makes the $z$-transform a generalization of the DTFT.

### 2.3.1. Zeros and Poles

The representation of a linear system (based on a autoregressive moving average equation) is given by

$$y(k) = \underbrace{\sum_{q=0}^{Q} x(k-q)\beta_q}_{MA(Q)} - \underbrace{\sum_{p=1}^{P} y(k-p)\alpha_p}_{AR(P)}. \tag{2.7}$$

taking the $z$-transform of the above equation yields

$$Y(z)\sum_{p=0}^{P} z^{-p}\alpha_p = X(z)\sum_{q=0}^{Q} z^{-q}\beta_q \tag{2.8}$$

and rearranging the terms, the transfer function of the system is given as

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{q=0}^{Q} z^{-q}\beta_q}{\sum_{p=0}^{P} z^{-p}\alpha_p} = \frac{\beta_0 + z^{-1}\beta_1 + z^{-2}\beta_2 + \cdots + z^{-Q}\beta_q}{1 + z^{-1}\alpha_1 + z^{-2}\alpha_2 + \cdots + z^{-P}\alpha_p} \tag{2.9}$$

The transfer function $H(z)$ can be represented as a function of the numerator and denominator roots. Applying the fundamental theorem of algebra, the numerator has $Q$ roots (called zeros of $H(z)$) and the denominator has $P$ roots (called poles of $H(z)$).

$$H(z) = \frac{(1-q_1z^{-1})(1-q_2z^{-1})\cdots(1-q_Qz^{-1})}{(1-p_1z^{-1})(1-p_2z^{-1})\cdots(1-p_Pz^{-1})} = \frac{\prod_{i=1}^{Q}(1-q_iz^{-1})}{\prod_{i=1}^{P}(1-p_iz^{-1})} \tag{2.10}$$

where $q_i$ is the $i^{th}$ zero and $p_i$ is the $i^{th}$ pole.

### 2.3.2. Stability and causality

To compute the inverse $z$-transform, a region of convergence (ROC) must be given in addition to the $z$-transform itself. Each ROC corresponds to a different signal. The real systems must be causal and stable systems, that makes the ROC to extend outward from the outermost pole and the stability requires that this ROC contains the unit circle. This conditions leads to the condition that all poles of the system must be inside the unit circle, that is to say

$$|p_i| < 1, \forall i \tag{2.11}$$

## 2.4. Laguerre sequences

### 2.4.1. Introduction

The Laguerre sequences are the discrete version of the Laguerre functions, called Laguerre because they are closely related to the Laguerre polynomials as it will be seen later. They consist of a polynomial multiplied by an exponential function, and look like the functions represented in Figure 2.1. They have one only multiple adjustable pole (denoted by '$p$' for the continuous functions and by '$a$' for the discrete case), what makes them to be considered as a compromise between IIR and FIR systems.

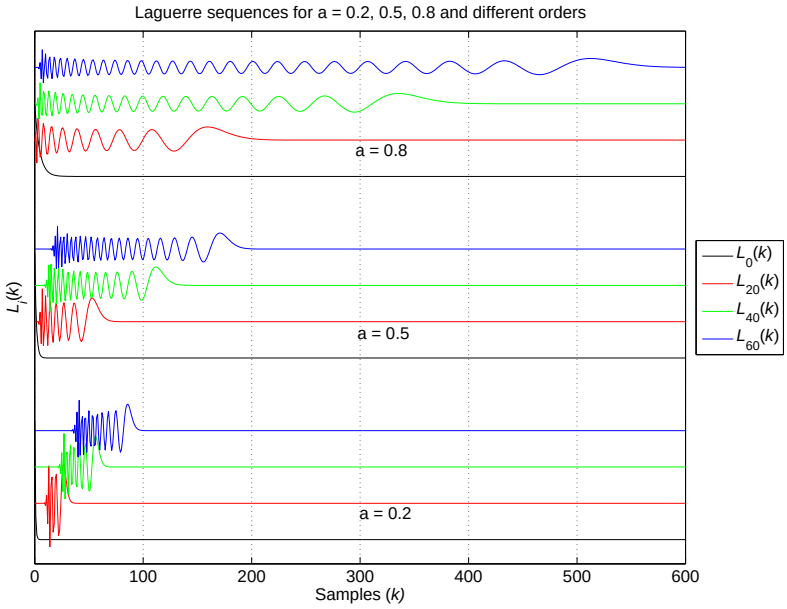



Figure 2.1.: Example of Laguerre sequences, for different orders and pole positions.

### 2.4.2. Laguerre polynomials

The Laguerre polynomials are solutions $L_n(x)$ to the Laguerre differential equation.

$$xy'' + (1-x)y' + \lambda y = 0, \qquad \lambda \in \mathbb{R} \tag{2.12}$$

which are defined by the Rodrigues formula

$$L_n(x) = \frac{e^x}{n!}\frac{d^n}{dx^n}\left(e^{-x}x^n\right) \tag{2.13}$$

The first few Laguerre polynomials are

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= -x + 1 \\ L_2(x) &= \frac{1}{2}\left(x^2 - 4x + 2\right) \\ L_3(x) &= \frac{1}{6}\left(-x^3 + 9x^2 - 18x + 6\right) \\ &\dots \end{aligned}$$

They are illustrated in Figure 2.2 for $x \in [-1, 5]$ and $n = 1, 2, \dots, 5$.

### 2.4.3. Laguerre functions

Laguerre functions are closely related to the Laguerre polynomials. They are obtained by orthonormalizing (as seen in [1]) the function $t^i e^{-pt}$, $i \in \mathbb{N}_0$, $p > 0$, with $p$ denoting the unique

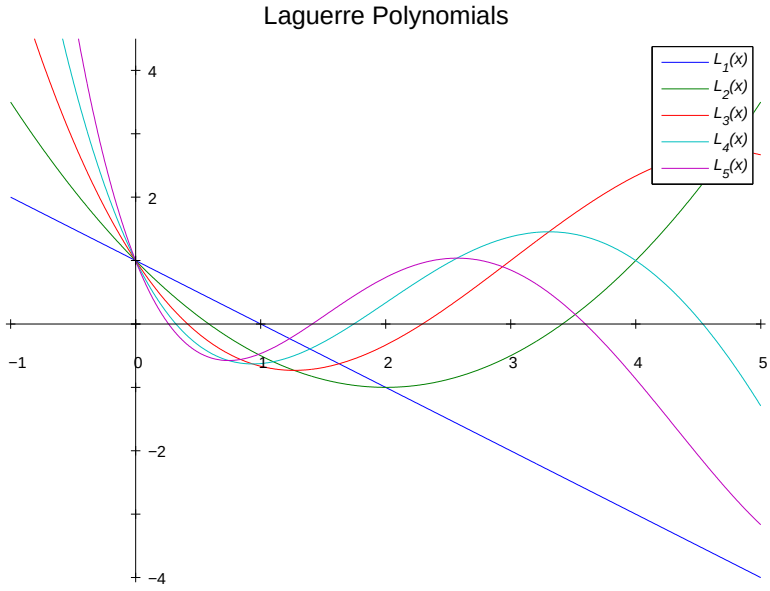



Figure 2.2.: Laguerre Polynomials

multiple pole (according to Equation 2.17). They are given by the formula,

$$l_i(t,p) = \sqrt{2p}e^{-pt}\sum_{j=0}^{i}\binom{i}{j}\frac{(-2pt)^j}{j!}. \tag{2.14}$$

They can be obtained by substituting $x$ by $2pt$ in the Laguerre polynomials and multiplying by $\sqrt{2p}e^{-pt}$.

$$\begin{aligned}
L_0(t,p) &= \sqrt{2p}e^{-pt}\\
L_1(t,p) &= \sqrt{2p}e^{-pt}\left(-2pt+1\right)\\
L_2(t,p) &= \frac{\sqrt{2p}2e^{-pt}}{2}\left(4p^2t^2-8pt+2\right)\\
L_3(t,p) &= \frac{\sqrt{2p}2e^{-pt}}{6}\left(-8p^3t^3+36p^2t^2-36pt+6\right)\\
&\dots
\end{aligned}$$

and form a complete orthonormal set on $\ell^2(\mathbb{R}^+)$, the Hilbert space of square integrable functions. That is to say that every square integrable function, $h(t)$, can be expressed as a weighted sum of these basis functions as

$$h(t) = \sum_{i=0}^{\infty} c_i(p)l_i(t,p) \tag{2.15}$$

with $c_i(p)$ defined as the inner product between the function and the basis functions, say

$$c_i(p) = \langle h(t), l_i(t,p)\rangle = \int_0^{\infty} h(t)l_i(t,p)\,dt \tag{2.16}$$

Despite of the complicated description of Laguerre functions in time domain, they have very simple Laplace transforms given by

$$L_i(s,p) = \sqrt{2p}\frac{(s-p)^i}{(s+p)^{i+1}} \tag{2.17}$$

This simple Laplace transform is one of the attractive characteristics of the Laguerre functions because it can be decomposed into a low-pass filter, followed by a series of all-pass filters by defining $L_0(s,p) = \frac{\sqrt{2p}}{(s+p)}$ and $L_A(s,p) = \frac{s-p}{s+p}$.

$$L_i(s,p) = \underbrace{\left[\frac{\sqrt{2p}}{s+p}\right]}_{low-pass}\underbrace{\left[\frac{s-p}{s+p}\right]^i}_{all-pass} = L_0(s,p)[L_A(s,p)]^i \tag{2.18}$$

This kind of decomposition allows to easily generate Laguerre functions using the structure in Figure 2.3 excited by a unit pulse. The output of the first low-pass section gives the zeroth

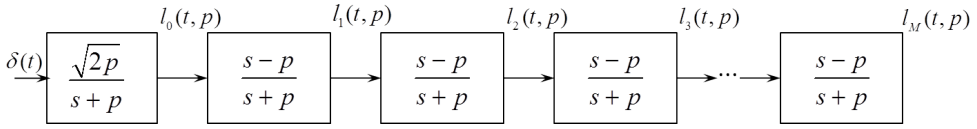
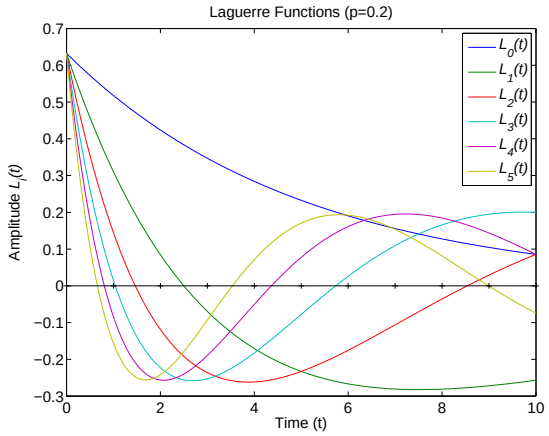


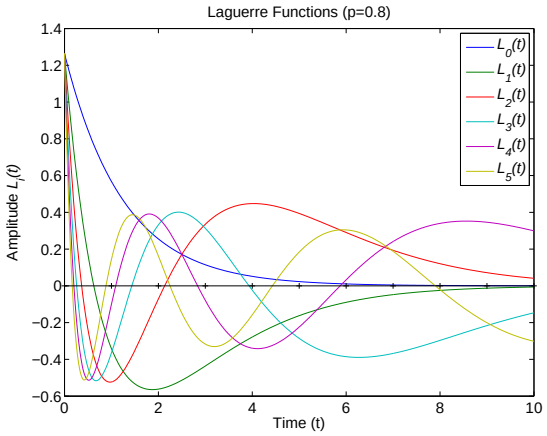

Figure 2.3.: Generation of Laguerre functions

order Laguerre function, then the output of the first all-pass section gives the first order Laguerre function, the second all-pass gives the second order Laguerre functions and so on. If a weighted sum of the different filter outputs is formed, the structure is known in the context of signal processing as Laguerre transversal filter.

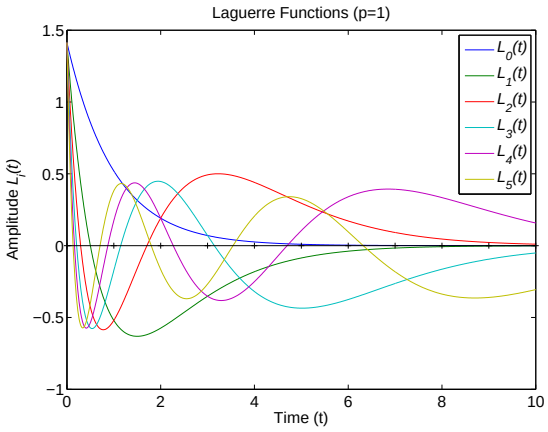
In Figure 2.4, the first six Laguerre functions are represented, for different values of the pole position ($p$), noting that the pole position scales the decay of the Laguerre functions i.e., their time spread.




(a) p=0.2




(b) p=0.8




(c) p=1

Figure 2.4.: Laguerre Functions with different value of the pole position $p$.

### 2.4.4. Laguerre Sequences

Laguerre sequences are the discrete version of the Laguerre functions and given by the following expression (note: now the pole position is denoted as $a$)

$$\begin{aligned} l_i(k,a) &= \sqrt{1-a^2}\sum_{j=0}^{i}(-1)^{i+j}\binom{i}{j}\binom{k+i-j}{i}a^{k+i+2j} \\ &= \sqrt{1-a^2}\sum_{j=0}^{i}(-1)^{i+j}\frac{(k+i-j)!}{(i-j)!j!(k-j)!}a^{k+i+2j} \end{aligned} \tag{2.19}$$

The pole position is varied on the real axis within the unity circle, i.e., from $-1$ to 1 in order to keep the stability as seen in Figure 2.5, emphasizing the range of real valued pole positions that lead to stable behavior. The first Laguerre sequences are derived from this formula

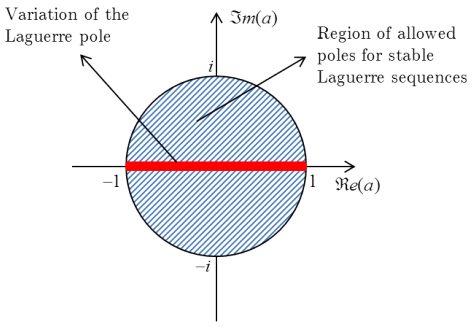



Figure 2.5.: Variation of the pole position

$$
\begin{aligned}
l_0(k,a) &= \sqrt{1-a^2}\frac{k!}{0!k!}a^k = \sqrt{1-a^2}a^k \\
l_1(k,a) &= \sqrt{1-a^2}\left(-\frac{(k+1)!}{0!k!}a^{k+1} + \frac{k!}{1!(k-1)!}a^{k-1}\right) = \sqrt{1-a^2}\left(-(k+1)a^{k+1} + ka^{k-1}\right) \\
l_2(k,a) &= \sqrt{1-a^2}\left(\frac{(k+2)!}{2!k!}a^{k+2} - \frac{(k+1)!}{1!(k-1)!}a^k + \frac{k!}{2!(k-2)!}a^{k-2}\right) \\
&= \sqrt{1-a^2}\left(\frac{k^2+3k+2}{2}a^{k+2} - (k^2+k)a^k + \frac{(k^2-k)}{2}a^{k-2}\right) \\
&\dots
\end{aligned}
$$

From this mathematical representation in discrete time, it can be concluded that $a$ is the decay parameter of the exponential sequence, and the order of the sequence indicates the number of exponential sequences that are superimposed.

Their $z$-transform is

$$
L_i(z,a) = \sqrt{1-a^2}\frac{(z^{-1}-a)^i}{(1-az^{-1})^{i+1}} = \underbrace{\frac{\sqrt{1-a^2}}{1-az^{-1}}}_{low-pass}\underbrace{\left[\frac{z^{-1}-a}{1-az^{-1}}\right]^i}_{all-pass} = L_0(z,a)[L_A(z,a)]^i \tag{2.20}
$$

by defining $L_0(s,p) = \frac{\sqrt{1-a^2}}{1-az^{-1}}$ and $L_A(s,p) = \frac{z^{-1}-a}{1-az^{-1}}$. This simple $z$-transform description, allows to generate Laguerre sequences with the structure shown in Figure 2.6, as the impulse response of a low-pass filter followed by a cascade of all-pass filters.

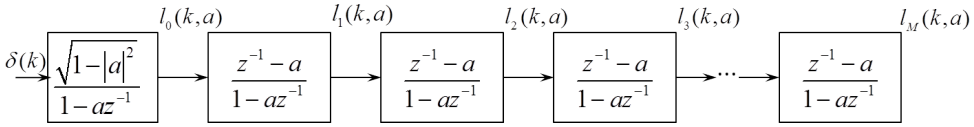
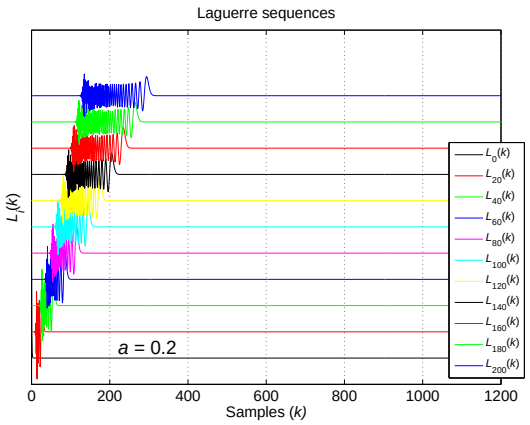


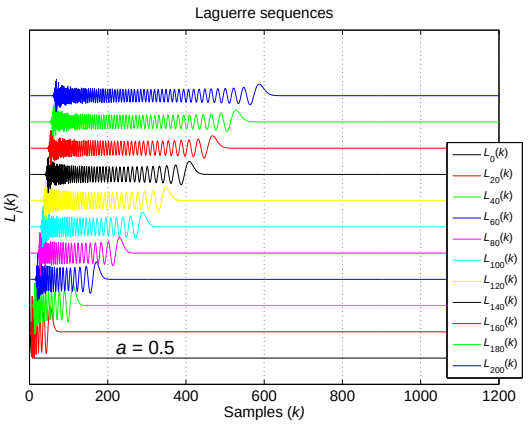

Figure 2.6.: Generation of Laguerre sequences

In Figure 2.7 different Laguerre sequences from order 0 to 200 are represented. Note the dependence of the time spread of each sequence due to the two main parameters, the pole position $a$, which defines the decay of the Laguerre sequence; and the order of the sequence. The closer to the unit circle the pole is ( $|a|$ is close to 1), the slower the sequence converges and the longer the time spread of the sequence. In the opposite case: the closer to zero the pole is the faster converges the sequence and the shorter the time spread of the sequence. It can also

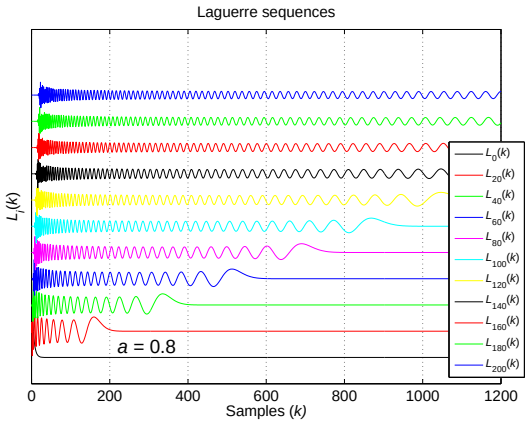
be seen, how the time spread of the sequences increases with the order of the sequence. With a negative pole an oscillatory behavior can be observed due to the $(-1)^k$ factor that a negative pole introduces.



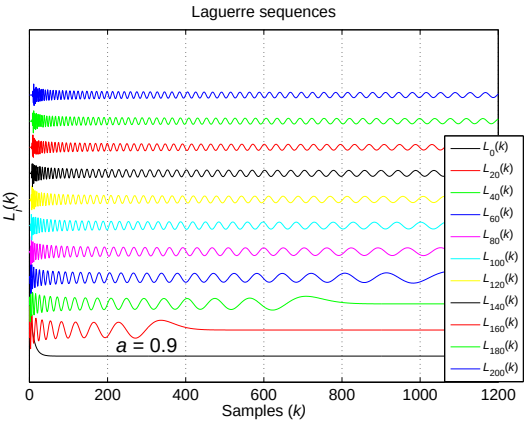

(a) $a = 0.2$



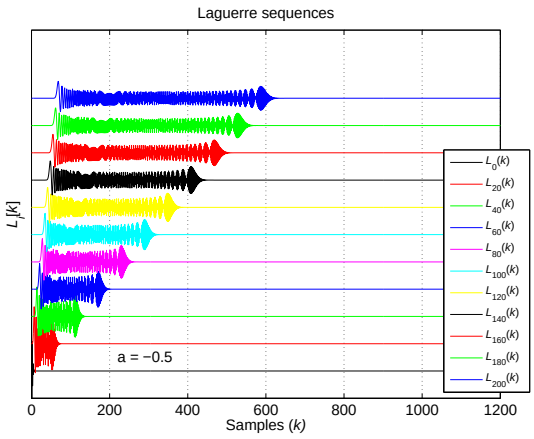

(b) $a = 0.5$



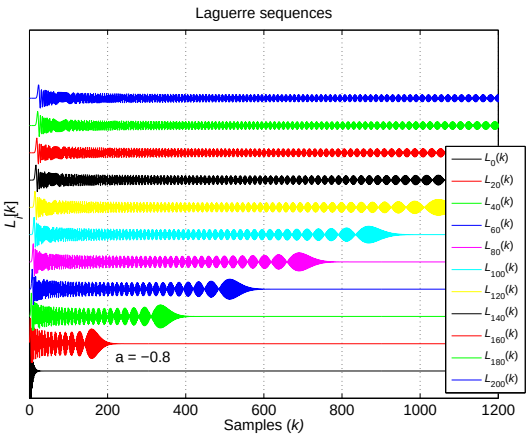

(c) $a = 0.8$




(d) $a = 0.9$




(e) $a = -0.5$




(f) $a = -0.8$

Figure 2.7.: Laguerre sequences with different 'a' value

### 2.4.5. Length covered by the Laguerre sequences

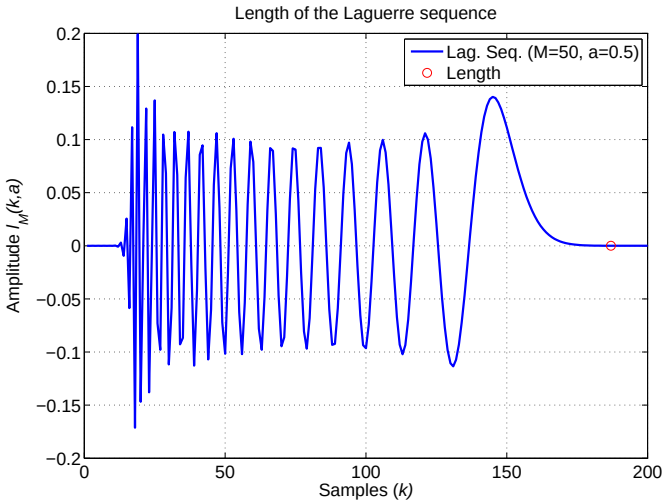



Figure 2.8.: Length of a Laguerre sequence

An interesting aspect of the Laguerre sequences is the length that they cover in samples, because intuitively for estimating a long impulse response with a fixed order, a reasonable choice for the pole position is the one that makes the Laguerre sequences to cover approximately the same length of the impulse response. For that reason if we can model the length covered by the Laguerre sequences, we can make an estimate of a good pole position by using this model with fixed length and filter order.

The length of a Laguerre sequence is defined by locating the sample when the last exponential decay is below a threshold as seen in the example in Figure 2.8. The resulting length as a function of the pole position and the sequence order is shown in Figure 2.9. From a heuristic point of view, it can be expected that the length follows a $\frac{1}{1-a}$ law, and is linear with the filter order $M$.

If we model the true length of a Laguerre sequence as a variable $l(M,a)$, the approximated length is defined by

$$\hat{l}(M,a) = \frac{k_1 M + k_2}{1-|a|} = k_1 \frac{M}{1-|a|} + k_2 \frac{1}{1-|a|} \tag{2.21}$$

we can determine the coefficients $k_1$ and $k_2$ by solving an easy least squares problem. If we calculate the length of the Laguerre sequences $l(M_j,a_j)$ for some pairs of values $\{M_j,a_j\}$ we have the problem

$$\mathbf{l} = \begin{bmatrix} l(M_1,a_1) \\ l(M_2,a_2) \\ \vdots \\ l(M_N,a_N) \end{bmatrix} = \begin{bmatrix} \frac{M_1}{1-|a_1|} & \frac{1}{1-|a_1|} \\ \frac{M_2}{1-|a_2|} & \frac{1}{1-|a_2|} \\ \vdots & \vdots \\ \frac{M_N}{1-|a_N|} & \frac{1}{1-|a_N|} \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \mathbf{H}\mathbf{k} \tag{2.22}$$

then the minimum square error solution for $\mathbf{k}$ is given by (see [11] for more details about the solution of the least squares problem)

$$\mathbf{k} = (\mathbf{H}^H\mathbf{H})^{-1}H^H\mathbf{l} = \begin{bmatrix} k_1 = 1,8402 \\ k_2 = 20,0563 \end{bmatrix} \tag{2.23}$$

which has been calculated using real Laguerre length values from sequences with orders from 1 to 200 and pole positions from 0 to 0.95.

The resulting approximation law for the length of a Laguerre sequence is given by

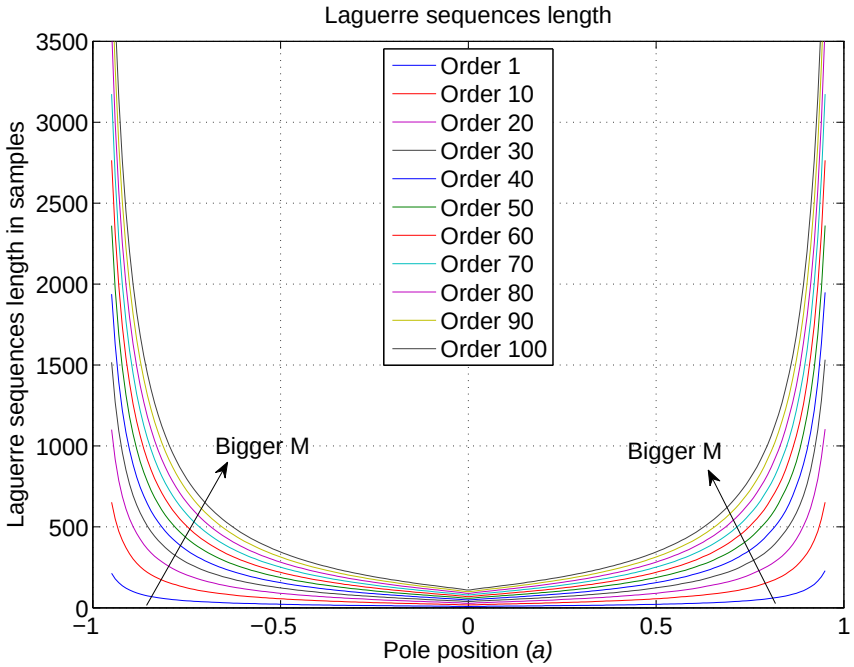
$$\hat{l}(M,a) = \frac{k_1 M + k_2}{1 - |a|}, \text{ with } k_1 = 1,8402 \text{ and } k_2 = 20,0563. \tag{2.24}$$

and a comparison between the true length and its estimate is shown in Figure 2.10.

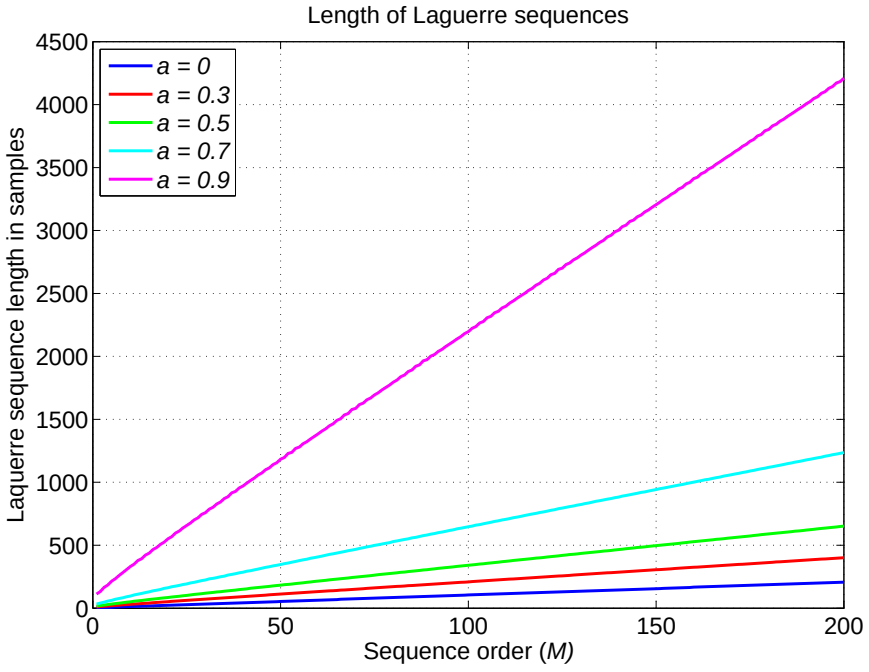
The idea behind this approximation model of the Laguerre length is to estimate an optimal pole position when the typical channel length, $l_h$ is known and the filter order, $M$ is fixed by solving $a$ from Equation 2.24 as

$$|\hat{a}|^{\text{Opt}} = \left| 1 - \frac{k_1 M + k_2}{l_h} \right| \tag{2.25}$$

which is the pole position that makes the Laguerre sequences to cover a length approximately equal to the channel length.

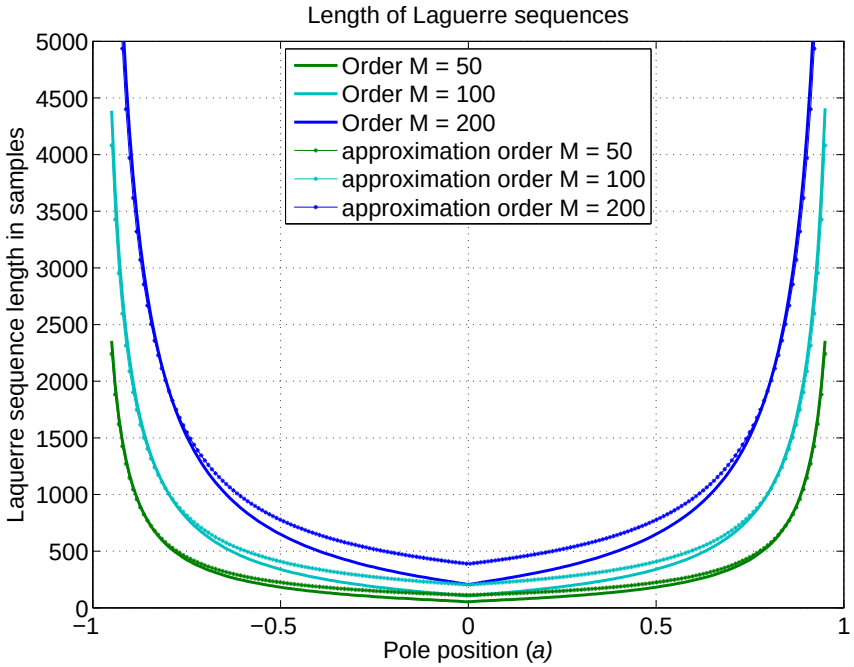



(a) Pole position

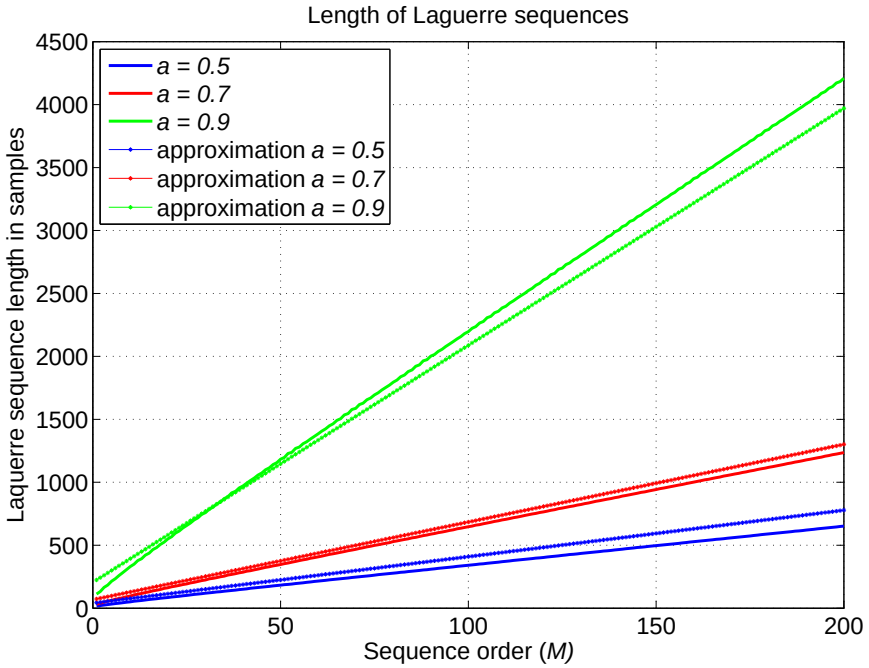



(b) Sequence order

Figure 2.9.: Length of the Laguerre sequences as function of sequence order or pole position.




(a) Function of pole position




(b) Function of sequence order

Figure 2.10.: Length (real and approximated) of the Laguerre sequences.

# 3. Laguerre spectrum

## 3.1. Introduction

To better understand the approximation using Laguerre sequence it is interesting to analyze how energy distributes on the Laguerre coefficients. For this purpose, the concept of Laguerre spectrum is developed. The idea behind it is the same concept as magnitude and power spectrum in Fourier analysis. The power spectrum describes how the energy (or variance) is distributed with frequency. A sequence $h(k)$ can be described as a weighted sum of Laguerre sequences, because they form a complete set, given by the Laguerre expansion

$$h(k) = \sum_{i=0}^{\infty} c_i(a) l_i(k, a), \tag{3.1}$$

and an approximation is made by truncating this sum to a fixed number of terms,

$$\hat{h}_M(k, a) = \sum_{i=0}^{M} c_i(a) l_i(k, a), \tag{3.2}$$

the energy of this approximation is obtained as follows

$$\begin{aligned} P_{\hat{h}_M}(a) &= \left\langle \hat{h}_M(k,a), \hat{h}_M^*(k,a) \right\rangle = \left\langle \sum_{i=0}^{M} c_i(a) l_i(k,a), \left( \sum_{i=0}^{M} c_i(a) l_i(k,a) \right)^* \right\rangle \\ &= \sum_{i=0}^{M} \sum_{j=0}^{M} c_i(a) c_j^*(a) \left\langle l_i(k,a), l_j^*(k,a) \right\rangle = \sum_{i=0}^{M} \sum_{j=0}^{M} c_i(a) c_j^*(a) \delta(i-j) \\ &= \sum_{i=0}^{M} |c_i(a)|^2 \end{aligned} \tag{3.3}$$

It can be observed that the energy that each Laguerre sequence $l_i(k, a)$ gathers corresponds to $|c_i(a)|^2$, because of the fact that Laguerre sequences form an orthonormal set, which makes that each sequence contributes with a new orthogonal dimension and unit energy. The Laguerre spectrum can be defined as

$$\Phi_{\hat{h}_M}(i, a) = |c_i(a)|^2 \tag{3.4}$$

This representation of the energy density or energy spectrum allows to observe how the energy is distributed with the coefficients. If energy is distributed in some structured manner, only some Laguerre sequences (possibly not in sequential order) could be used.

## 3.2. Discussion

To analyze how the energy is distributed along the Laguerre sequences a test case is used, consisting of a delayed exponential power delay profile as the one shown in Figure 3.1. The

expression for this test case is given by

$$h(k) = 0.6^{k-\tau} u[k-\tau],$$

This test case has a delay value ($\tau$) that varies and the Laguerre spectrum is calculated for

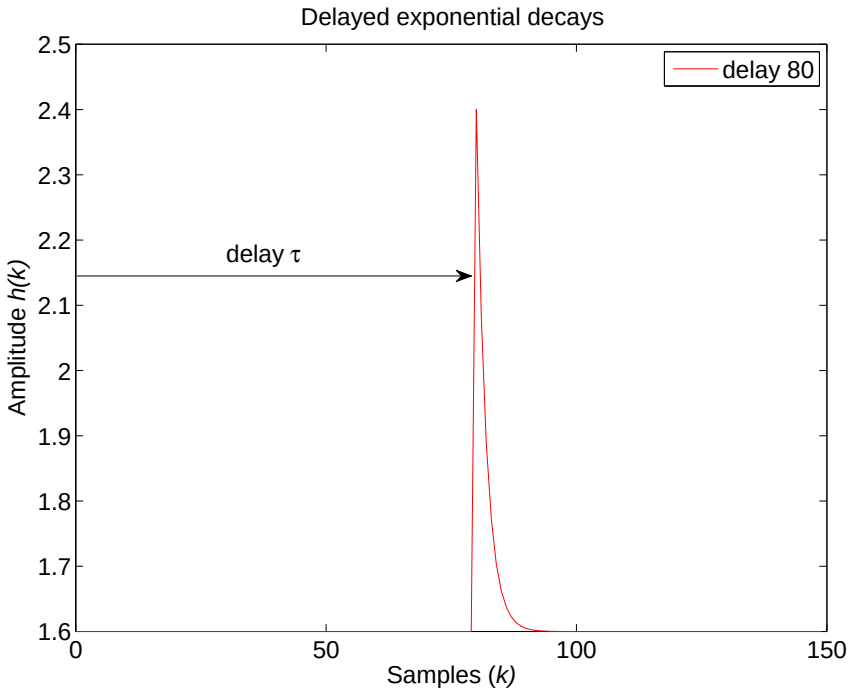
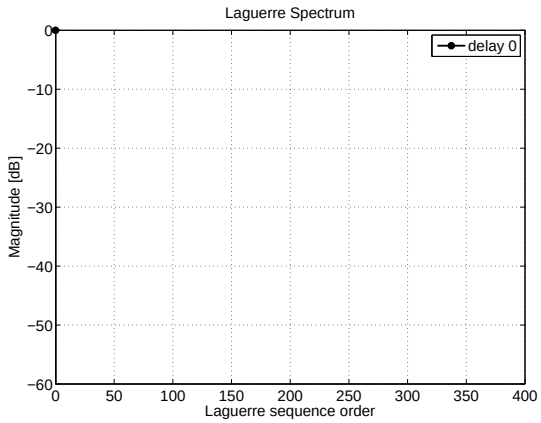



Figure 3.1.: Delayed power delay profile utilized to compute the Laguerre spectrum.

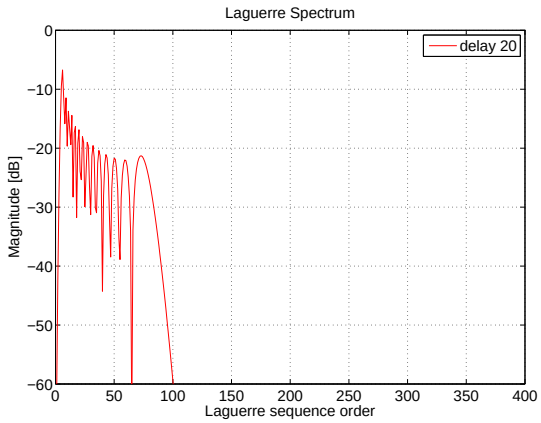
each delay value.

In Figure 3.2, the Laguerre spectrum for various delay values are represented for Laguerre sequences with $a = 0.6$ (the time constant of the system). It can be seen that for the trivial case of $\tau = 0$, which corresponds to an exponential decay, only one Laguerre sequence is needed (remember that Laguerre sequences consist of exponential functions). For the rest of the cases, it can be seen that there is some threshold order in which Laguerre sequences begin to gather energy (in most of the cases this value is small), at this point most of the energy is gathered however the energy is almost homogeneously distributed in the following sequences until it reaches an order value in which all the energy is gathered and the order increase collects no further energy. There are two main conclusions for this test case. One is that there is no exploitable distribution of the energy through Laguerre sequences; the second one is that with Laguerre sequences, energy can be gathered more uniformly than with FIR systems.

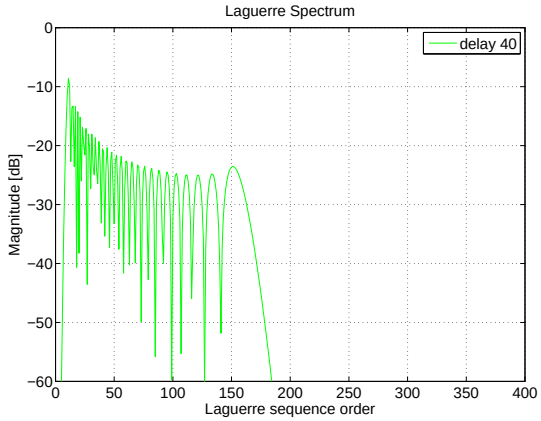
To see these effects, the approximation with Laguerre sequences of a delayed exponential power delay profile with $\tau = 80$ for different orders is shown in Figure 3.3. Here it can be seen that the main problem of the approximation quality lies in the abrupt change of the function (discontinuities).



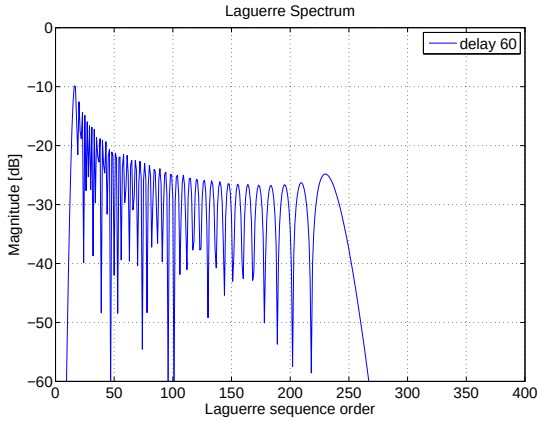

(a) delay $\tau = 0$



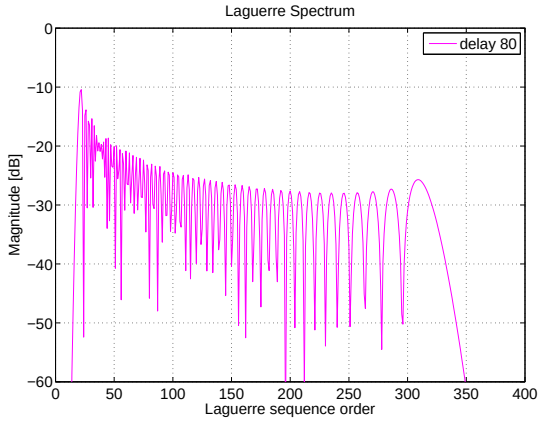

(b) delay $\tau = 20$



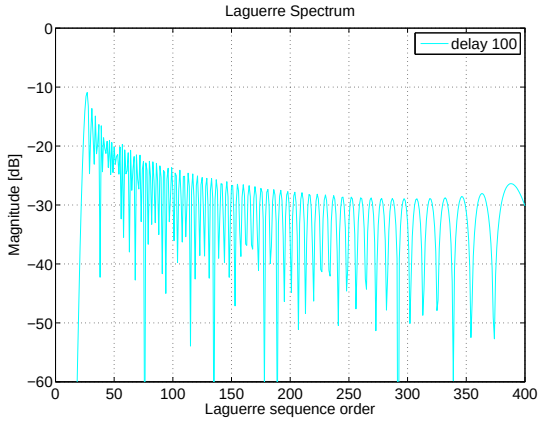

(c) delay $\tau = 40$




(d) delay $\tau = 60$

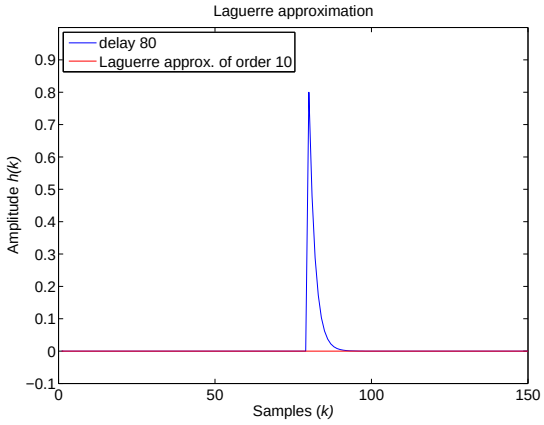



(e) delay $\tau = 80$

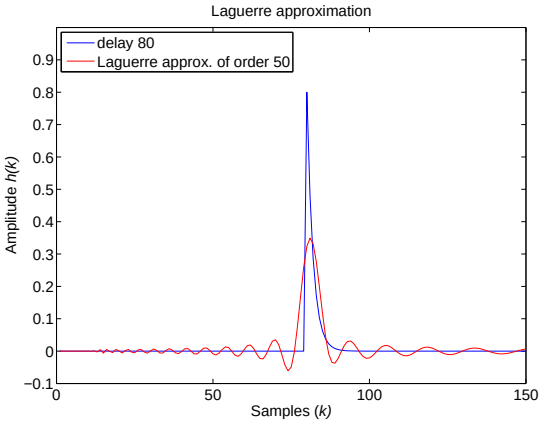



(f) delay $\tau = 100$

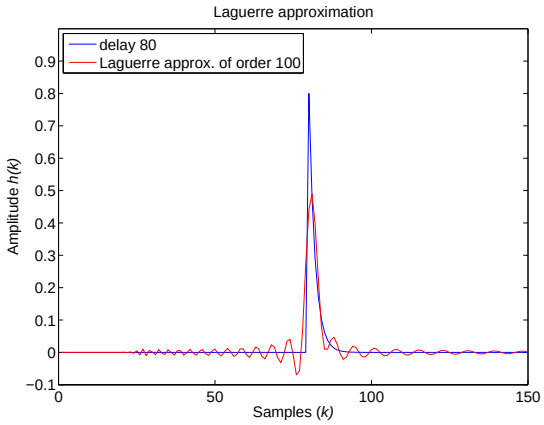
Figure 3.2.: Laguerre spectrum of a delayed exponential power delay profile with pole position equal to the time constant of the system.



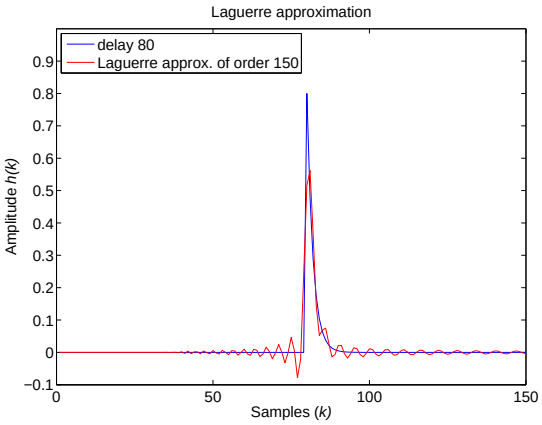

(a) delay $\tau = 80$, Approximation of order 10



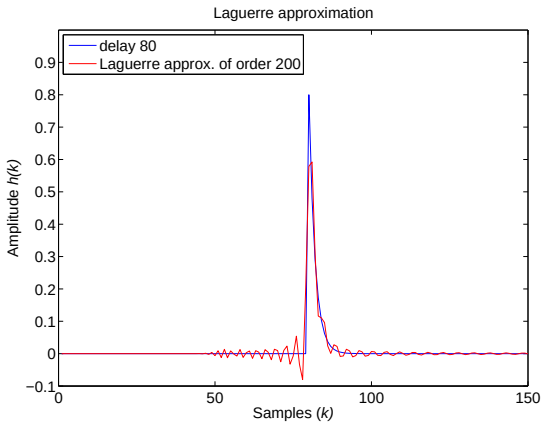

(b) delay $\tau = 80$, Approximation of order 50



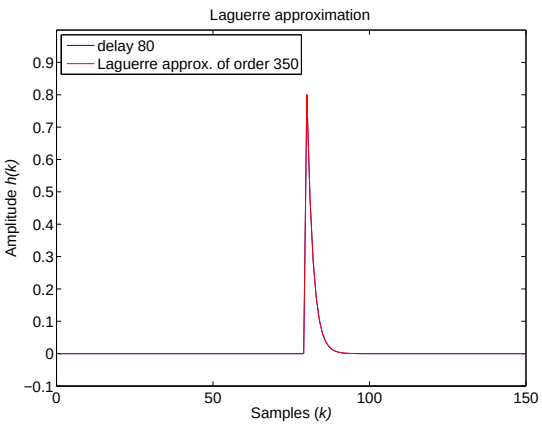

(c) delay $\tau = 80$, Approximation of order 100




(d) delay $\tau = 80$, Approximation of order 150




(e) delay $\tau = 80$, Approximation of order 200




(f) delay $\tau = 80$, Approximation of order 350

Figure 3.3.: Laguerre approximation of a delayed exponential power delay profile with pole position equal to the time constant of the system and different approximation orders.

To see what is the influence of the pole position on the Laguerre spectrum, the same delayed exponential power delay profile with $\tau = 80$ is used to generate the Laguerre spectrum for various pole positions, as shown in Figure 3.5. Here it can be seen that the pole positions plays a role as a compression factor. For values close to zero, the spectrum is more narrow but for values closer to one we have a wider spectrum, however some energy can be gathered with sequences of lower order. If we take a look at the spectrum for $a = 0$ in Figure 3.5, it can be seen that is the most narrower spectrum but until the sequence order reaches 80, no energy can be gathered. However with a pole position of large magnitude the spectrum is wider but some energy can be gathered with lower order sequences.

A compression effect of the Laguerre sequences has proven to be useful in some applications such as spectral estimation (see [5]), which can be derived from the fact that a Laguerre transform can be expressed by means of the Fourier transform doing a variable change, assuming $x(k) = \sum_{i=0}^{\infty} c_i(a) l_i(k, a)$

$$X(e^{\mathrm{j}\omega}) = \frac{\sqrt{1-|a|^2}}{1-ae^{-\mathrm{j}\omega}} \sum_{i=0}^{\infty} c_i(a) \left( \frac{e^{-\mathrm{j}\omega} - a}{1-ae^{-\mathrm{j}\omega}} \right)^i \tag{3.5}$$

with the variable change

$$e^{\mathrm{j}\theta} = \frac{e^{\mathrm{j}\omega} - a}{1-ae^{\mathrm{j}\omega}} \implies e^{\mathrm{j}\omega} = \frac{e^{\mathrm{j}\theta} + a}{1+ae^{\mathrm{j}\theta}} \tag{3.6}$$

then it results

$$\frac{\sqrt{1-|a|^2}}{1+ae^{-\mathrm{j}\theta}} X\left(\frac{e^{\mathrm{j}\theta} + a}{1+ae^{\mathrm{j}\theta}}\right) = \sum_{i=0}^{\infty} c_i(a) e^{-\mathrm{j}\theta i} \tag{3.7}$$

which shows that the Laguerre coefficients $c_i(a)$ are equivalent to the Fourier coefficients of a frequency compressed version of the Fourier transform with a frequency compression that depends on the parameter $a$ as represented in Figure 3.4.

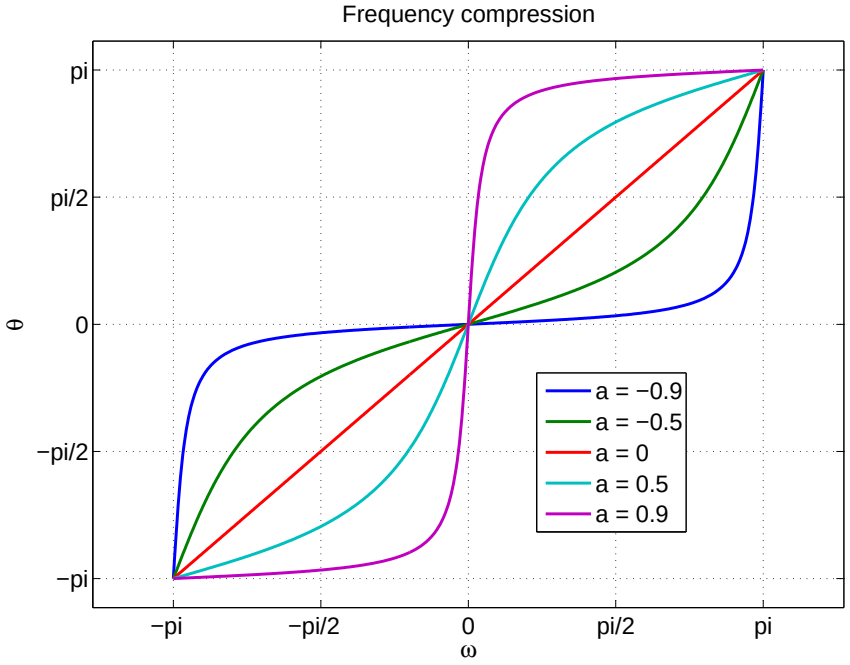
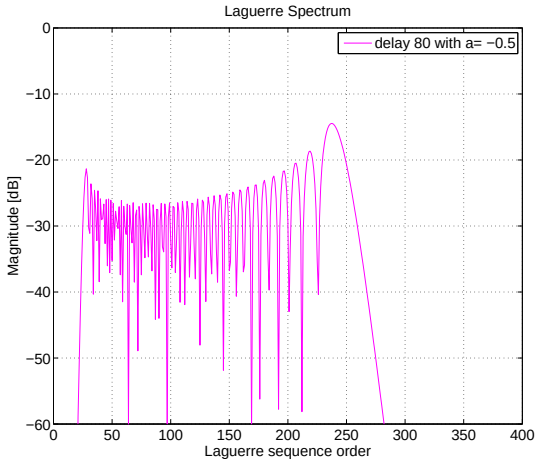
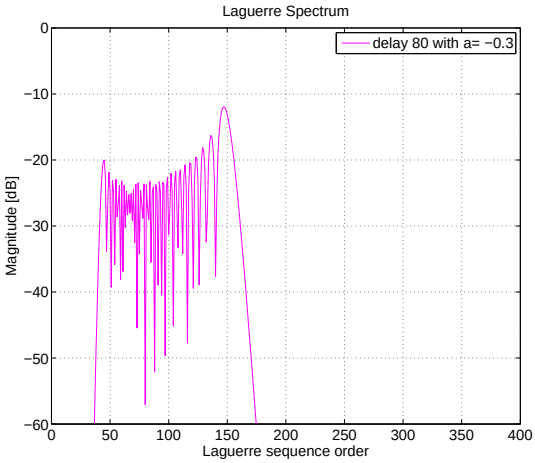


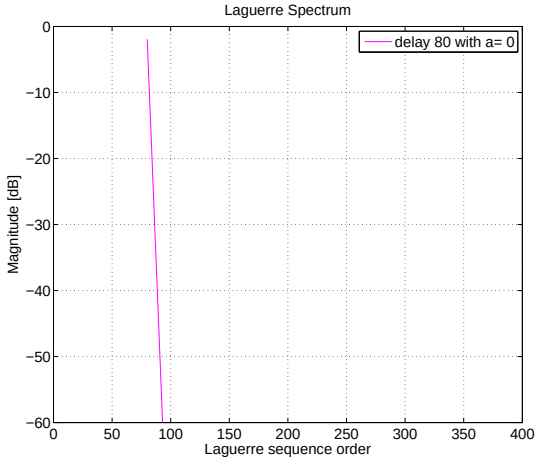

Figure 3.4.: Frequency compression produced by the parameter $a$.



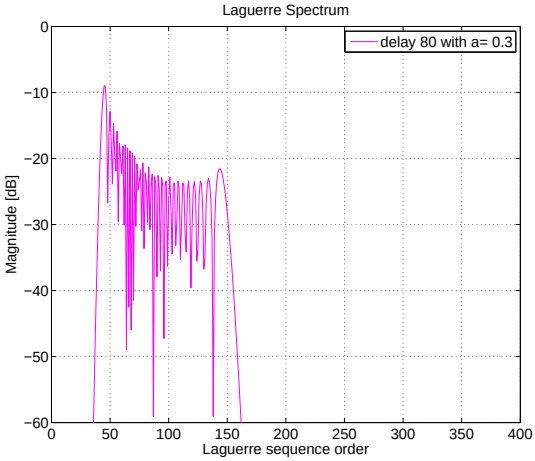

(a) $a = -0.5$



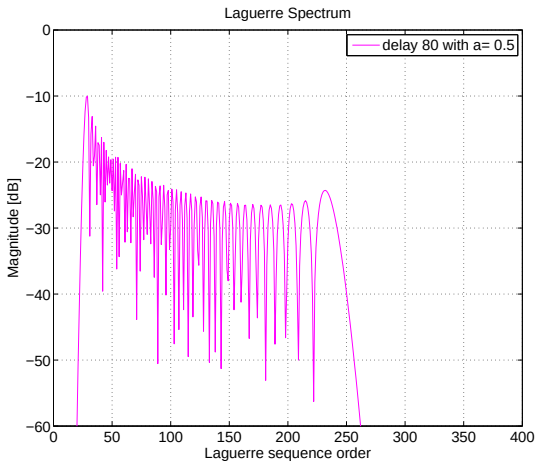

(b) $a = -0.3$



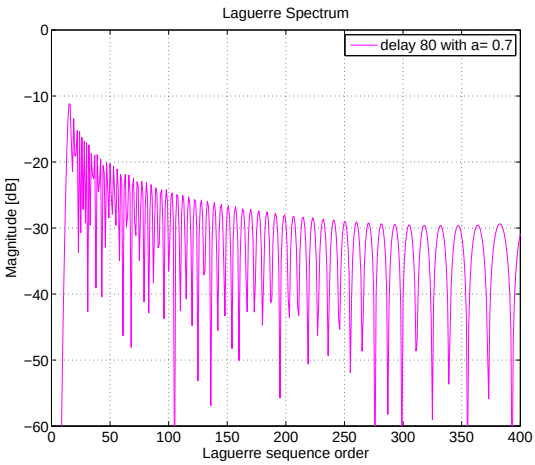

(c) $a = 0$




(d) $a = 0.3$




(e) $a = 0.5$




(f) $a = 0.7$

Figure 3.5.: Laguerre spectrum of a delayed exponential power delay profile ($\tau = 80$) with different pole positions.

# 4. Transversal Filter

## 4.1. Introduction

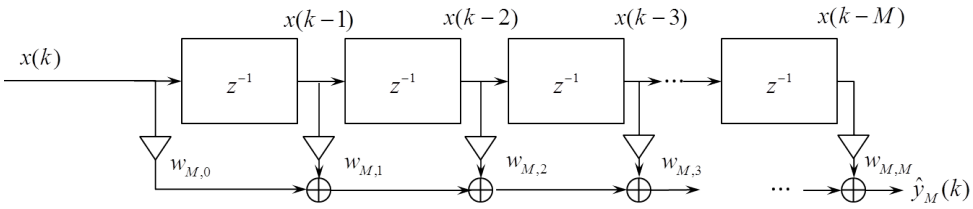



Figure 4.1.: Classic transversal filter

The transversal filter is the most commonly used filter for modeling linear systems, especially if adaptive methods are used. The source of this popularity is their simplicity, stability and the existence of fast and efficient adaptive algorithms. There are also some related structures, such as the lattice filter, which are very popular. Some applications for this kind of filters are found on system identification, linear prediction, channel equalization and echo cancellation. A classic transversal filter is shown in Figure 4.1.

## 4.2. FIR transversal filter

The FIR transversal filter is the well-known transversal filter of Figure 4.1, whose output is given by

$$\hat{y}_M(k) = \sum_{i=0}^{M} w_{M,i} x(k-i) = \mathbf{w}_M \mathbf{x}_M(k) \tag{4.1}$$

with $\mathbf{w}_M$, the weights vector of length $M+1$ and $\mathbf{x}_M(k)$ defined as the input data vector of length $M+1$, which in this case corresponds to the different delayed inputs, given by.

$$\mathbf{w}_M = \begin{bmatrix} w_{M,0} & w_{M,1} & \cdots & w_{M,M} \end{bmatrix}, \tag{4.2}$$

$$\mathbf{x}_M(k) = \begin{bmatrix} x_0(k) \\ x_1(k) \\ \vdots \\ x_M(k) \end{bmatrix} = \begin{bmatrix} x(k) \\ x(k-1) \\ \vdots \\ x(k-M) \end{bmatrix} \tag{4.3}$$

The reason of denoting the output of the system by $\hat{y}_M(k)$ is that the filter is intended to approximate a system response. So the filter output is an approximation $\hat{y}_M(k)$ of order $M$ of the desired output $y(k)$. The approximated impulse response of the system is given by the output of the system to an impulse input, i.e.,

$$\hat{h}_M(k) = \sum_{i=0}^{M} w_{M,i} \delta(k-i) \tag{4.4}$$

As it can be seen, this impulse response is a FIR system, whose representation in time domain leads directly to the implementation of the transversal filter. Treating this impulse response as an approximation in Hilbert space, it is clear that FIR basis functions are the Kronecker delta ($\delta(k-i)$), also known as the canonical basis in $\ell^2$. The finite duration of the impulse response of the transversal filter allows most of the advantages like stability and simplicity but also causes that when a long impulse response is required, the number of delays of the filter required can be quite high.

For this reason, IIR filters can be a possible solution to this problem, however they have their own problems, specially if adaptive methods are used. There are two main possible problems multimodal surfaces and instability (see [5]). The first of them means that the error surface can have local minimum that do not correspond to a global minima. The second one, instability, related to the adaptation of the poles of this kind of filters which makes it necessary to develop methods to keep the stability, i.e., to check if the stability is maintained and recover the stability when the system is instable (poles stay within the unit circle during adaptation).

## 4.3. Laguerre transversal filter

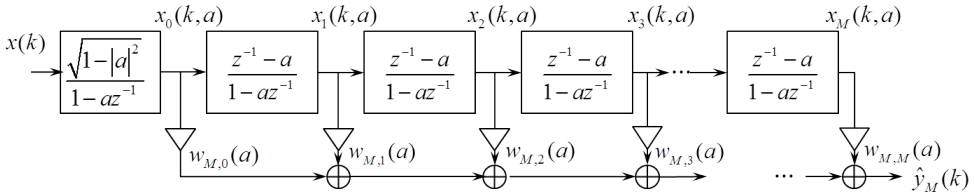



Figure 4.2.: Laguerre transversal filter

As it has been seen, FIR transversal filter requires large order filters when the impulse response has to be long, and IIR filter have their own problems. The idea is to find some filters between these two extreme cases. As seen in Chapter 2, Laguerre sequences have only one multiple adjustable pole, which allows to guarantee the stability of the system. Laguerre filters are considered as a compromise between IIR filters and FIR filter, actually they are a generalization of the FIR filters. Laguerre sequences can be represented in z-domain as

$$L_i(z,a) = \sqrt{1-a^2}\frac{(z^{-1}-a)^i}{(1-az^{-1})^{i+1}} = \underbrace{\frac{\sqrt{1-a^2}}{1-az^{-1}}}_{low-pass}\underbrace{\left[\frac{z^{-1}-a}{1-az^{-1}}\right]^i}_{all-pass} = L_0(z,a)[L_A(z,a)]^i \tag{4.5}$$

with $L_0(z,a)$ a low-pass filter, and $L_A(z,a)$ an all-pass section. This representation leads naturally to the Laguerre transversal filter shown in Figure 4.2 consisting of a low-pass filter followed by a cascade of all-pass sections, with output and impulse response given by

$$\hat{y}_M(k) = \sum_{i=0}^{M} w_{M,i}(a)x_i(k,a) = \mathbf{w}_M(a)\mathbf{x}_M(k,a) \tag{4.6}$$

$$\hat{h}_M(k) = \sum_{i=0}^{M} w_{M,i}(a)l_i(k,a) \tag{4.7}$$

where $\mathbf{x}_M(k,a)$ is the vector with the output signals of Laguerre filters of different orders from 0 to $M$. The pole position, denoted by $a$, is taken as a parameter and not a variable, i.e., the pole

position is chosen a priori. The variation of this pole position is only allowed on the real axis within the unit circle, that is to say, from $-1$ to 1, as seen in Figure 4.3. This allows to keep this pole position inside the unit circle, keeping the stability of the system. This is one of the reasons why Laguerre filters are considered a compromise between IIR and FIR filters, because they have only one multiple adjustable pole. It is also important that when the pole position is equal to zero ($a = 0$), the Laguerre transversal filter degenerates to the FIR transversal filter, i.e., Laguerre filters can be considered as a generalization of the transversal filter. The variation of this pole position allows to control the rate of decay of its impulse response, which is useful to approximate systems with long impulse responses.

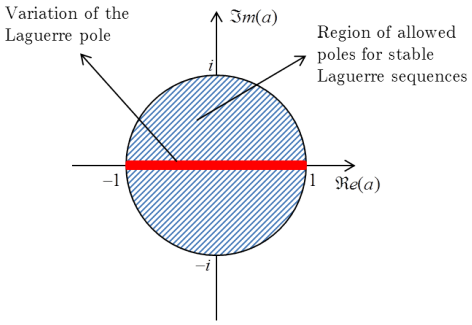



Figure 4.3.: Variation of the pole position

## 4.4. Linear estimation theory

### 4.4.1. Estimation model

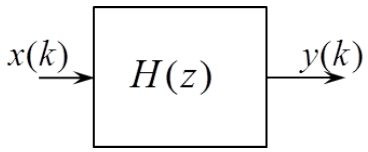



Figure 4.4.: Simple channel model

The channel model for this study is the simple model shown in Figure 4.4, defining $x(k)$ as the input signal, the channel as $H(z)$ with impulse response $h(k)$ and the output as $y(k)$. The reason not to include noise in this model in order to be more realistic is that this is a theoretical study and noise will only add some limitations to the quality of the approximation. In fact, if additive white Gaussian noise (AWGN) is added to the system, it would cause that the MSE will have an asymptotic value limited by the effective SNR of the system.

The aim is to approximate (identify, estimate) the a priori unknown channel impulse response. For this purpose the model seen in Figure 4.5 is proposed. New signals are introduced, the approximated system of order $M$, $\hat{H}_M(z)$, the approximation of order $M$ of the output $y(k)$, i.e. $\hat{y}_M(k)$, and the error signal $e_M(k)$ which is the difference between the desired signal $y(k)$ and its approximation $\hat{y}_M(k)$. The output of the approximated system, $\hat{y}_M(k)$, is supposed to be a linear combination of $M + 1$ observations or data $x_i(k)$. This data is considered to be some linear transformation of the input data $x(k)$, which is a generalized view of all kinds of filters in order to keep the notation compact (in our case only the FIR transversal filter and Laguerre transversal filter will be used, which perfectly match this representation, in the case of the FIR transversal filter the observations are the different delayed samples of the input and

in the Laguerre transversal filters the observations are the outputs of the different transversal filter sections). The output is given by

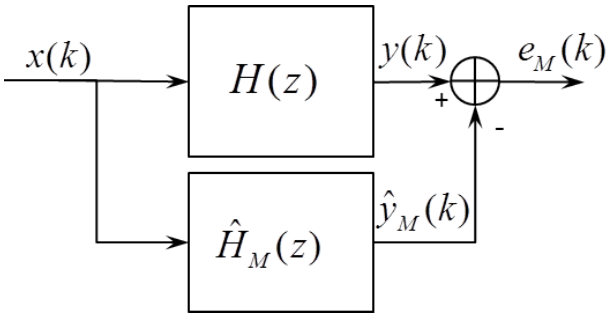



Figure 4.5.: Channel estimation model

$$\hat{y}_M(k) = \sum_{i=0}^{M} w_{M,i} x_i(k) = \mathbf{w}_M \mathbf{x}_M(k) \tag{4.8}$$

and the error signal

$$e_M(k) = y(k) - \hat{y}_M(k) \tag{4.9}$$

It can be noticed that the error signal is defined in terms of $y(k)$ and not $h(k)$, which is the global objective, however when the resulting error is normalized by the energy of $y(k)$, $\sigma_y^2$, the resulting normalized error corresponds to the error of the system approximation.

### 4.4.2. Minimum mean square error criterion

Once the error signal is defined, a criterion is needed in order to compare different approximations and differentiate it. The chosen criterion is the minimum mean square error (m.m.s.e) criterion, which establishes as cost function ($J$) the mean square error (MSE)

$$J(\mathbf{w}_M) = E\left[e_M(k) e_M^*(k)\right] = \underbrace{E\left[|e_M(k)|^2\right]}_{MSE} \tag{4.10}$$

and the m.m.s.e minimizes this cost function by correctly choosing the filter weights $\mathbf{w}_M$

$$\mathbf{w}_M = \underset{\mathbf{w}_M}{\operatorname{argmin}}(J(\mathbf{w}_M))$$

The minimization of the cost function can be achieved through two main approaches, the first is solving the equations with perfect knowledge of the second order statistics of the signals and the second one by means of a recursive solution, i.e., an adaptive method. Later in Chapter 8, adaptive methods will be developed and studied, however now the normal equations will be solved by perfectly knowing the signal statistics.

The minimum of the cost function is obtained by differentiating $J(\mathbf{w}_M)$ with respect to $\mathbf{w}_M$ and setting the derivatives to zero. First of all, the cost function is rewritten as a function of the filter weights $\mathbf{w}_M$

$$\begin{aligned} J(\mathbf{w}_M) &= \mathsf{E}\left[e_M(k) e_M^*(k)\right] = \mathsf{E}\left[(y(k) - \hat{y}_M(k))(y(k) - \hat{y}_M(k))^*\right] \\ &= \mathsf{E}\left[y(k) y^*(k)\right] + \mathsf{E}\left[\hat{y}_M(k) \hat{y}_M^*(k)\right] - \mathsf{E}\left[\hat{y}_M(k) y^*(k)\right] - \mathsf{E}\left[y(k) \hat{y}_M^*(k)\right] \\ &= \mathsf{E}\left[y(k) y^*(k)\right] + \mathbf{w}_M \mathsf{E}\left[\mathbf{x}_M(k) \mathbf{x}_M^H(k)\right] \mathbf{w}_M^H - \mathbf{w}_M \mathsf{E}\left[\mathbf{x}_M(k) y^*(k)\right] - \mathsf{E}\left[y(k) \mathbf{x}_M^H(k)\right] \mathbf{w}_M^H \end{aligned}$$

introducing the following definitions

$$\begin{aligned} \sigma_y^2 &\triangleq \mathsf{E}\left[y(k)y^*(k)\right] \\ \mathbf{R}_x &\triangleq \mathsf{E}\left[\mathbf{x}_M(k)\mathbf{x}_M^H(k)\right] \\ \mathbf{p}_{yx} &\triangleq \mathsf{E}\left[y(k)\mathbf{x}_M^H(k)\right] \end{aligned}$$

the resulting cost function is given by

$$J(\mathbf{w}_M) = \sigma_y^2 + \mathbf{w}_M\mathbf{R}_x\mathbf{w}_M^H - \mathbf{w}_M\mathbf{p}_{yx}^H - \mathbf{p}_{yx}\mathbf{w}_M^H \tag{4.11}$$

now differentiating with respect to $\mathbf{w}_M^H$

$$\nabla_{\mathbf{w}_M^H} J(\mathbf{w}_M) = \mathbf{w}_M\mathbf{R}_x - \mathbf{p}_{yx} \tag{4.12}$$

setting the gradient equal to zero gives the so called normal equations

$$\boxed{\mathbf{w}_M\mathbf{R}_x = \mathbf{p}_{yx}} \tag{4.13}$$

They are called normal equations because the optimal weights $\mathbf{w}_M$ correspond to an approximation $\hat{y}_M$ that fulfills an important orthogonality condition, which is more clearly seen, if the terms are rearranged in the following manner

$$\begin{aligned} \mathbf{w}_M\mathbf{R}_x = \mathbf{p}_{yx} \quad &\Longrightarrow \quad \mathbf{w}_M\mathsf{E}\left[\mathbf{x}_M(k)\mathbf{x}_M^H(k)\right] = \mathsf{E}\left[y(k)\mathbf{x}_M^H(k)\right] \\ &\Longrightarrow \quad \mathsf{E}\left[\left(\mathbf{w}_M\mathbf{x}_M(k) - y(k)\right)\mathbf{x}_M^H(k)\right] = 0 \end{aligned}$$

which gives the equivalent representation of the normal equations

$$\boxed{\mathsf{E}\left[e_M(k)\mathbf{x}_M^H(k)\right] = 0} \tag{4.14}$$

These equations have the important geometric interpretation: It states an orthogonality condition that means that the estimation error, $e_M(k)$ is orthogonal to the data, $\mathbf{x}_M(k)$, in fact, to any linear transformation of the data. This means that no further linear transformation of the data can extract additional information about the desired signal or reduce the estimation error on average.

$$e_M \perp x_M$$

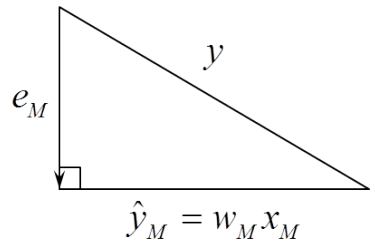



Figure 4.6.: Geometric interpretation of the orthogonality condition

### 4.4.3. Correlation coefficients

In order to calculate the coefficients $w_{M,i}$ of the filter, the normal equations (see Eq. 4.13) must be solved. To solve the problem, two quantities are needed, the autocorrelation matrix of the outputs of the transversal filter sections $\mathbf{R}_x$ and the cross correlation vector between the desired signal and the outputs of the transversal filter sections, $\mathbf{p}_{yx}$ given by

$$\begin{aligned} \mathbf{R}_x &= \mathsf{E}[\mathbf{x}_M(k,a)\mathbf{x}_M^H(k,a)] && (4.15)\\ \mathbf{p}_{yx} &= \mathsf{E}[y(k)\mathbf{x}_M^H(k,a)] && (4.16) \end{aligned}$$

Each correlation coefficients of the Laguerre transversal filter are given by

$$\begin{aligned} r_{ij} &= \mathsf{E}[x_i(k,a)x_j^*(k,a)] && (4.17)\\ p_i &= \mathsf{E}[y(k)x_i^*(k,a)] && (4.18) \end{aligned}$$

These coefficients can be calculated numerically in time domain by taking an arithmetic mean based on a finite number of available data points $N$

$$\begin{aligned} r_{ij} &\simeq \frac{1}{N}\sum_{k=0}^{N-1} x_i(k,a)x_j^*(k,a)\\ p_i &\simeq \frac{1}{N}\sum_{k=0}^{N-1} y(k)x_i^*(k,a) \end{aligned}$$

However, as it has been said in Chapter 2, Laguerre sequences have a very simple frequency domain representation, which intuitively can provide a better way to calculate these coefficients. Remembering Parseval's theorem, which allows to go from a time domain to a frequency domain representation

$$\mathsf{E}[f(k)g^*(k)] = \langle f(k), g(k)\rangle = \frac{1}{2\pi}\int_{-\pi}^{+\pi} F(e^{\mathrm{j}\omega})G^*(e^{\mathrm{j}\omega})\,d\omega \tag{4.19}$$

applying it to the correlation coefficients

$$r_{ij} = \mathsf{E}[x_i(k,a)x_j^*(k,a)] = \frac{1}{2\pi}\int_{-\pi}^{+\pi} X_i(e^{\mathrm{j}\omega},a)X_j^*(e^{\mathrm{j}\omega},a)\,d\omega \tag{4.20}$$

the frequency representation of the i$^{\text{th}}$ output of the Laguerre transversal filter, $x_i(k,a)$ is given by

$$X_i(e^{\mathrm{j}\omega},a) = X(e^{\mathrm{j}\omega})L_i(e^{\mathrm{j}\omega},a) = X(e^{\mathrm{j}\omega})L_0(e^{\mathrm{j}\omega},a)[L_A(e^{\mathrm{j}\omega},a)]^i \tag{4.21}$$

noting that $L_A^*(e^{\mathrm{j}\omega},a) = \frac{1}{L_A(e^{\mathrm{j}\omega},a)}$

$$X_j^*(e^{\mathrm{j}\omega},a) = X^*(e^{\mathrm{j}\omega})L_0^*(e^{\mathrm{j}\omega},a)[L_A^*(e^{\mathrm{j}\omega},a)]^j = X^*(e^{\mathrm{j}\omega})L_0^*(e^{\mathrm{j}\omega},a)[L_A(e^{\mathrm{j}\omega},a)]^{-j} \tag{4.22}$$

now the coefficients can be expressed as

$$\begin{aligned} r_{ij} &= \frac{1}{2\pi}\int_{-\pi}^{+\pi} X_i(e^{\mathrm{j}\omega},a)X_j^*(e^{\mathrm{j}\omega},a)\,d\omega\\ &= \frac{1}{2\pi}\int_{-\pi}^{+\pi} X(e^{\mathrm{j}\omega})X^*(e^{\mathrm{j}\omega})|L_0(e^{\mathrm{j}\omega},a)|^2[L_A(e^{\mathrm{j}\omega},a)]^{i-j}\,d\omega \end{aligned} \tag{4.23}$$

and defining the power spectral density of $x(k)$ as $\phi_{xx}(e^{\mathrm{j}\omega}) = X(e^{\mathrm{j}\omega})X^*(e^{\mathrm{j}\omega})$,

$$r_{ij} = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \phi_{xx}(e^{\mathrm{j}\omega})|L_0(e^{\mathrm{j}\omega},a)|^2[L_A(e^{\mathrm{j}\omega},a)]^{i-j}\,d\omega \tag{4.24}$$

the dependence on $i-j$ only leads to a Toeplitz structure of the autocorrelation matrix $\mathbf{R}_x$ and thereby to the possibility to apply the Levinson-Durbin algorithm to solve the normal equations more efficiently (as seen in next section) by a recursion that order-updates the solution.

Proceeding in the same way for the cross correlation coefficients

$$\begin{aligned} p_i &= \frac{1}{2\pi}\int_{-\pi}^{+\pi} Y(e^{\mathrm{j}\omega})X_i^*(e^{\mathrm{j}\omega},a)\,d\omega = \frac{1}{2\pi}\int_{-\pi}^{+\pi} Y(e^{\mathrm{j}\omega})X^*(e^{\mathrm{j}\omega})L_0^*(e^{\mathrm{j}\omega},a)[L_A(e^{\mathrm{j}\omega},a)]^{-i}\,d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{+\pi} H(e^{\mathrm{j}\omega})X(e^{\mathrm{j}\omega},a)X^*(e^{\mathrm{j}\omega},a)L_0^*(e^{\mathrm{j}\omega},a)[L_A(e^{\mathrm{j}\omega},a)]^{-i}\,d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{+\pi} \phi_{xx}(e^{\mathrm{j}\omega})H(e^{\mathrm{j}\omega})L_0^*(e^{\mathrm{j}\omega},a)[L_A(e^{\mathrm{j}\omega},a)]^{-i}\,d\omega \end{aligned} \tag{4.25}$$

substituting the transfer functions for Laguerre filters $L_0(e^{\mathrm{j}\omega})$ and $L_A(e^{\mathrm{j}\omega})$ by their expression

$$r_{ij} = \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{+\pi} \frac{\Phi_{xx}(e^{\mathrm{j}\omega})(1-a^2)}{|1-ae^{\mathrm{j}\omega}|^2}\left(\frac{e^{\mathrm{j}\omega}-a}{1-ae^{\mathrm{j}\omega}}\right)^{j-i} d\omega \tag{4.26}$$

$$p_i = \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{+\pi} \frac{H(e^{\mathrm{j}\omega})\Phi_{xx}(e^{\mathrm{j}\omega})\sqrt{1-a^2}}{1-ae^{\mathrm{j}\omega}}\left(\frac{e^{\mathrm{j}\omega}-a}{1-ae^{\mathrm{j}\omega}}\right)^{i} d\omega \tag{4.27}$$

### 4.4.4. Solution method for the Normal equations

First of all, the correlation coefficients are calculated for a given pole position $a$ and a filter order $M$. It has been seen that Laguerre filters provide a Toeplitz matrix, which implies that for correctly characterizing the autocorrelation matrix $\mathbf{R}_x$, only a vector $\mathbf{r}$ is needed. They are calculated in frequency domain using the following formulas ([12])

$$r_i = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \frac{\Phi_{xx}(e^{\mathrm{j}\omega})(1-a^2)}{|1-ae^{\mathrm{j}\omega}|^2}\left(\frac{e^{\mathrm{j}\omega}-a}{1-ae^{\mathrm{j}\omega}}\right)^{i} d\omega,\ \ i=0,1,\cdots M \tag{4.28}$$

$$p_i = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \frac{H(e^{\mathrm{j}\omega})\Phi_{xx}(e^{\mathrm{j}\omega})\sqrt{1-a^2}}{1-ae^{\mathrm{j}\omega}}\left(\frac{e^{\mathrm{j}\omega}-a}{1-ae^{\mathrm{j}\omega}}\right)^{i} d\omega,\ \ i=0,1,\cdots M \tag{4.29}$$

Once the two correlation vectors are calculated, $\mathbf{r}$, first row of $\mathbf{R}_x$ and $\mathbf{p}_{yx}$, the normal equations given by

$$\mathbf{w}_M\mathbf{R}_x = \mathbf{p}_{yx} \tag{4.30}$$

are solved. In order to solve the equations, the direct approach is to use matrix inversion. However this approach is not efficient, because the inverse matrix has to be calculated for each order separately with the problems related with inverting a matrix. The method used is by means of a modified version of the Levinson Durbin algorithm, which allows to solve the normal equations in a recursive manner. This means that for the solution of the problem for an order of $M$, the solution of the problem of order $M-1$ is used which increases the efficiency of the method and avoids the matrix inversion. The method is described in the appendix. The Levinson Durbin algorithm is executed until it reaches the desired filter order $M$ and gives as result the coefficients $c_i$ of the orthonormal expansion of $\hat{y}_M(k)$ assuming some orthonormal sequences $\phi_i(k)$ (in fact these orthonormal sequences $\phi_i(k)$ are the backward prediction error

sequences $e_i^b(k)$ as defined and explained in the appendix)

$$\hat{y}_M(k) = \sum_{i=0}^{M} c_i \phi_i(k) \tag{4.31}$$

with

$$c_i = \langle y, \phi_i \rangle$$

note that the coefficients $c_i$ do not depend on the filter order because they are the coefficients of the orthonormal expansion of $y(k)$. The MSE for each order $M$ is calculated as follows

$$MSE = J_{\min} = \mathsf{E}|e_M|^2 = \langle e_M, e_M \rangle = \langle (y - \hat{y}_M), e_M \rangle$$

remembering that the error signal (because of the normal equations) is orthogonal to all linear combinations of the data (i.e., $\hat{y}_M$)

$$MSE = J_{\min} = \langle y, e_M \rangle = \sigma_y^2 - \langle y, \hat{y}_M \rangle = \sigma_y^2 - \sum_{i=0}^{M} c_i \langle y, \phi_i \rangle = \sigma_y^2 - \sum_{i=0}^{M} |c_i|^2$$

This calculated MSE is related to the signal $y(k)$ but we want to calculate the MSE for the system approximation, that is $h(k)$. Normalizing this result with respect to the power of $y(k)$ to obtain the normalized MSE (NMSE) which is the NMSE of the system approximation

$$NMSE = 1 - \sum_{i=0}^{M} \frac{|c_i|^2}{\sigma_y^2}$$

# 5. IIR System

## 5.1. Introduction

Now that the method to calculate the MSE is known, in order to validate the MATLAB code and as an introduction to Laguerre filtering, an IIR system is estimated with a Laguerre transversal filter, in particular the example of the study of Tomas Oliveira e Silva in [12]. It is clear that this scenario only serves to validate our MATLAB code and is favorable for Laguerre filters at the expense of FIR filters because the impulse response is a smooth rational function, so this section is only for illustrating purposes.

## 5.2. System Description

An elliptic third order low-pass IIR filter is used, taken from [12], with the following transfer function,

$$H(z) = \frac{0.01624(1 + z^{-1})(1 - 1.7313z^{-1} + z^{-2})}{(1 - 0.8957z^{-1})(1 - 1.8445z^{-1} + 0.9282z^{-2})}. \tag{5.1}$$

The frequency response is shown in 5.2, allowing to appreciate its low pass behavior. To see the impulse response of this filter, a Kronecker delta is fed into the input of the system, giving the impulse response of the system as output. This impulse response is shown in Figure 5.1.

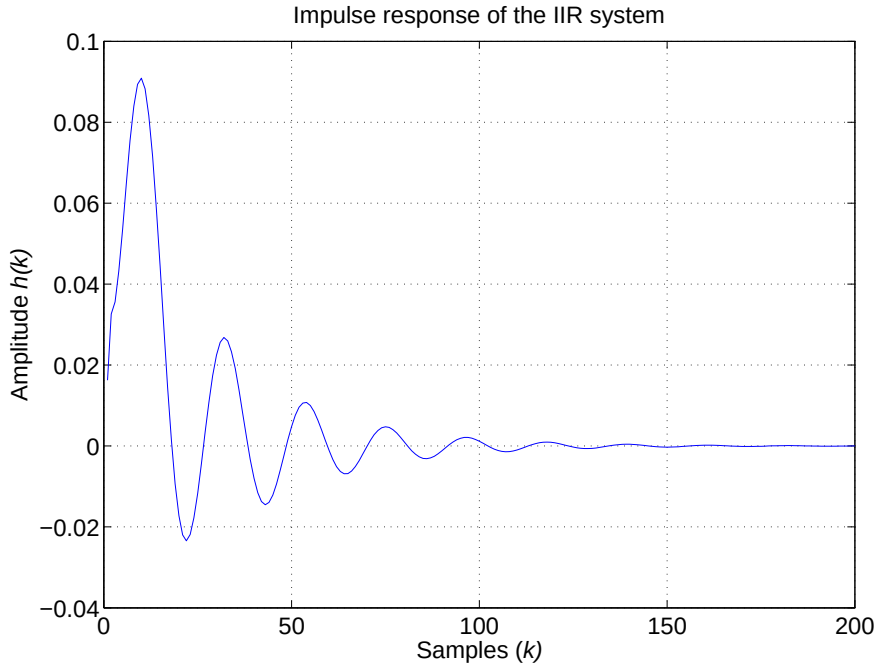



Figure 5.1.: Impulse response of the IIR system.

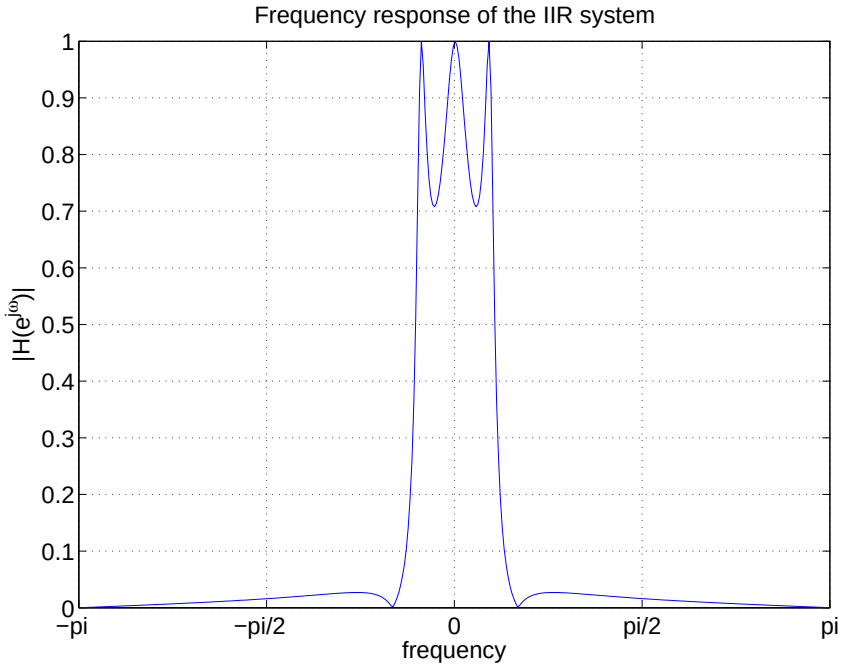



Figure 5.2.: Frequency response of the 3rd order elliptic filter.

## 5.3. Mean Square Error

The interpretation of the MSE is the energy of an unknown system that is not captured by an approximation with a system of order $M$. The process to calculate the Mean Square Error (MSE) is described in Chapter 4. The input training signal of the system is white Gaussian noise of unit variance which is used to produce an approximation with a Laguerre transversal filter of orders 0 to 10. This approximation is made by perfectly knowing the second order statistics of the system, i.e., the autocorrelation matrix $\mathbf{R}_x$ and the cross-correlation vector $\mathbf{p}_{yx}$. The MSE is normalized with respect to the output signal power, to obtain the Normalized MSE (NMSE), which varies from 0 dB to $-\infty$ dB. A 0 dB NMSE means that the normalized error is 1, i.e., there is no approximation at all; a $-\infty$ dB NMSE means that the normalized error is 0, i.e., the approximation is perfect, so the less NMSE the better is the approximation/estimation.

The resulting MSE of the approximation for the IIR channel is shown in Figure 5.3 and is a reproduction of Fig. 7 in [12]. This figure represents the normalized MSE of the Laguerre filter for orders 0 (top) up to 10 (bottom) as a function of the pole position (denoted by $a$). This pole position varies on the real axis within the unit circle, from $-1$ to $+1$ as represented in Figure 4.3. When $a = 0$, as it has been said, the Laguerre filter degenerates into a FIR transversal filter, which allows to see the Laguerre filter as a generalization of the FIR transversal filter. What can be seen in this Figure is that the filter performs much better with a pole position close to 0.7 than a FIR filter, i.e., for the pole position $a = 0$. It also shows that for this kind of impulse response, the pole position is not very critical, it only has to be close to the optimal pole position.

Values of the MSE for the Laguerre filter of order 10 can be seen in table 5.1. As seen in the table, a better behavior of the filter can be obtained with a Laguerre filter by adjusting the pole position close to 0.7.

By reorganizing the information in Figure 5.3, the NMSE can be represented as a function of the filter order for some fixed pole positions, as shown in figure 5.4. Here the performance of the

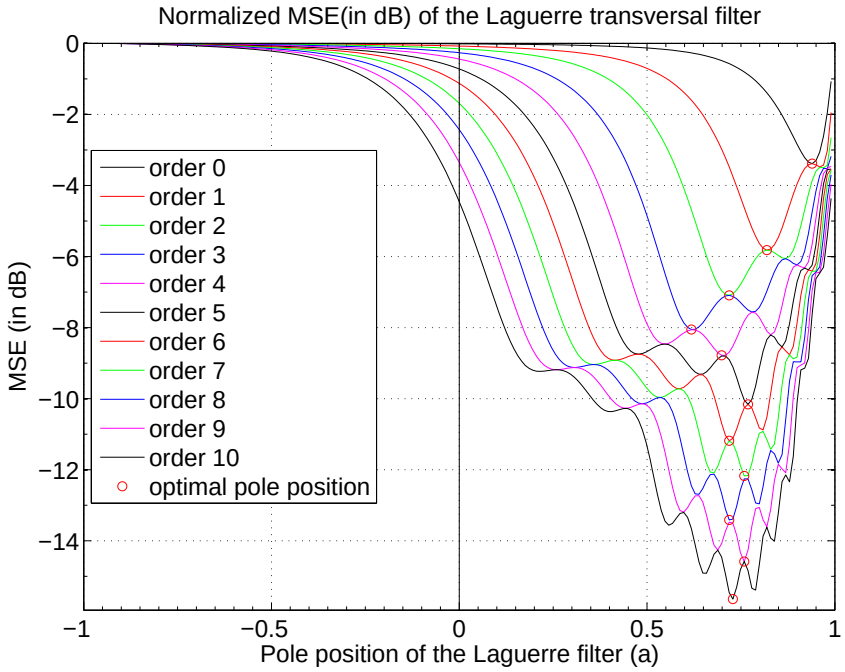



Figure 5.3.: Normalized MSE of IIR channel estimate as a function of the pole position for different filter orders (note: $a = 0 \stackrel{\triangle}{=}$ FIR filter ).

Table 5.1.: NMSE for Laguerre filter of order 10 with different pole positions

| filter | pole position ($a$) | NMSE (in $dB$) |
|---|---|---|
| FIR transversal filter | 0 | $-4.4\,dB$ |
| Laguerre filter | 0.5 | $-11.3\,dB$ |
| Laguerre filter | 0.7 | $-14.5\,dB$ |
| Laguerre filter | 0.73 (optimal pole position) | $-15.6\,dB$ |

Laguerre filter can be seen more accurately. The line corresponding to the optimal pole position, corresponds to a filter with the best (optimal) pole position for each order. This optimal pole position is given in figure 5.5.

To see how Laguerre filters approximate the impulse response, Figure 5.6 represents the result of the approximation for a filter of order 10 for some pole positions. Here it can be clearly seen the problem with FIR filters because only the first samples of the channel response can be approximated which for long impulse responses leads to a bad approximation, while Laguerre sequences due to their longer time spread can approximate a longer section of the impulse response.

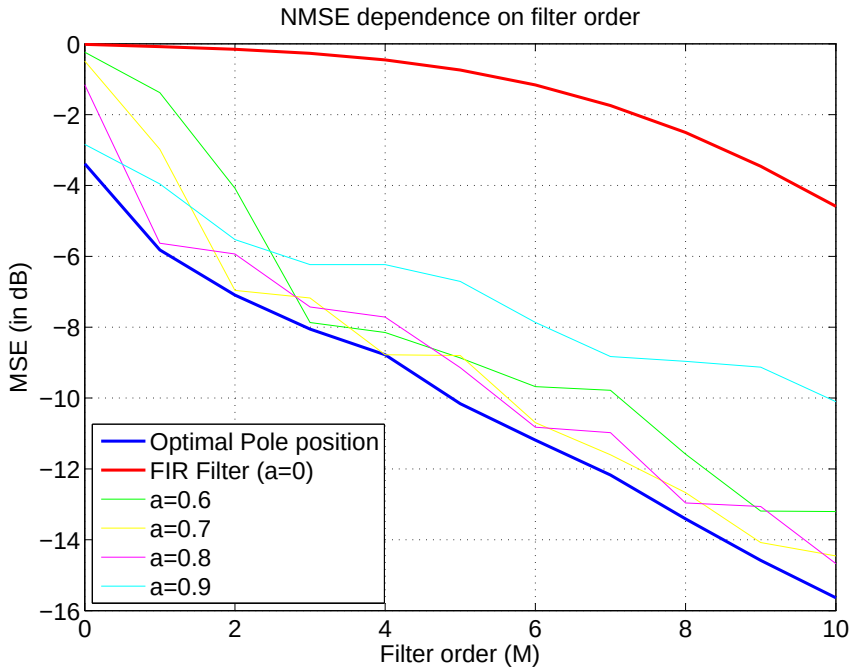



Figure 5.4.: Normalized MSE of IIR channel estimate as a function of filter order for different pole positions (note: $a = 0 \triangleq$ FIR filter ).

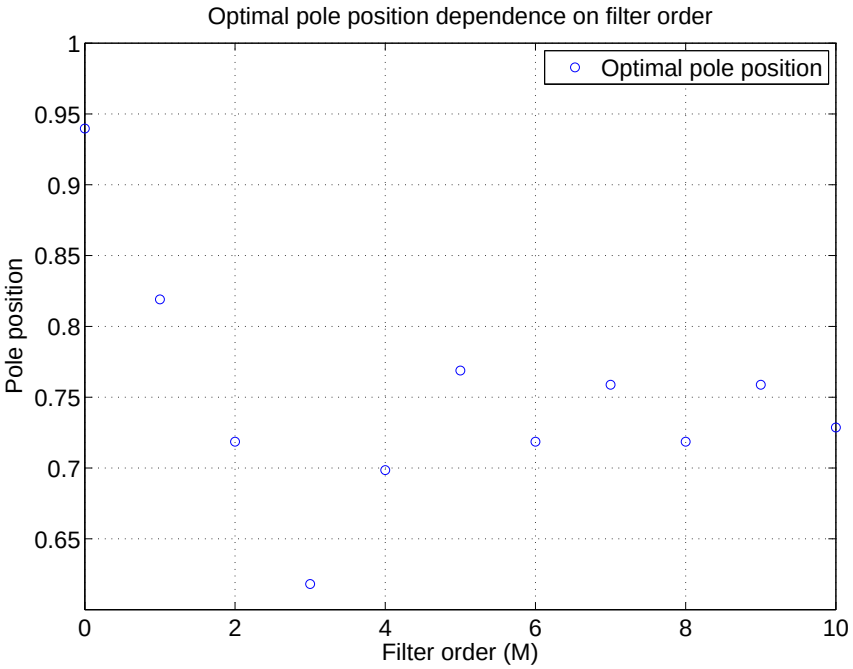



Figure 5.5.: Optimal pole position for the IIR channel as a function of the filter order.

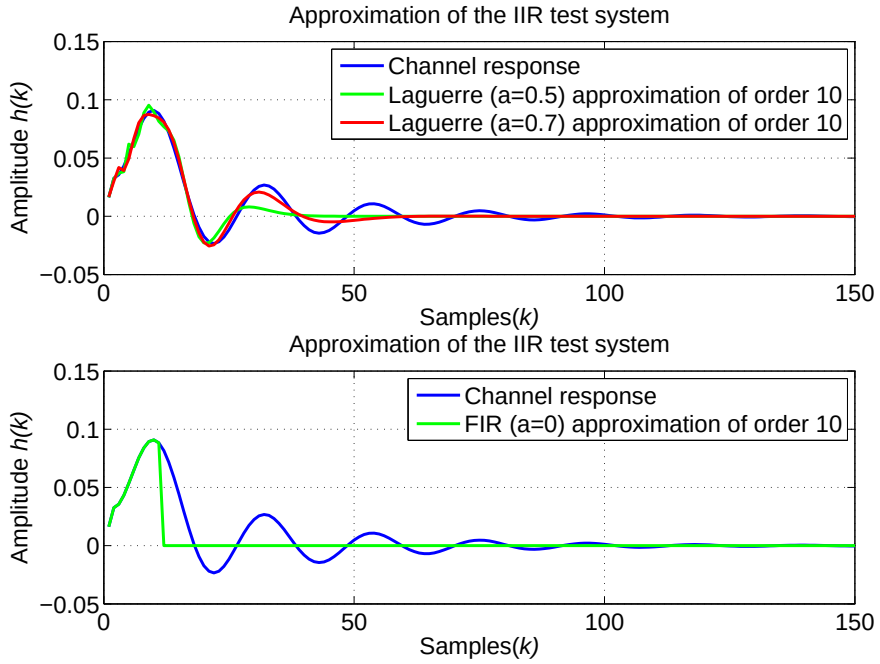



Figure 5.6.: Approximation of the IIR channel impulse response for FIR and Laguerre filters of order 10.

# 6. Approximation of an idealized UWB FIR channel

## 6.1. Introduction

Now we want to study the approximation of impulse responses that are closer to the UWB channels. UWB channels are modeled by the standard IEEE 802.15.4a (see [8]), and this model is based on a clustered behavior of the different echo arrival times and an exponential power delay profile of the clusters. These channels represent a long impulse response and for NLOS environments it can happen that there is no clearly dominant echo. The most simple channel impulse response regarding to this description is a FIR channel consisting of two delayed exponential power delay profiles. The reason for not using the UWB stochastic model yet is that this analysis is intended to provide illustrative results of how good Laguerre filters behave by approximating UWB channels and a stochastic model is not yet appropriate because it makes the interpretation of the results more difficult and will therefore be used only afterwards.

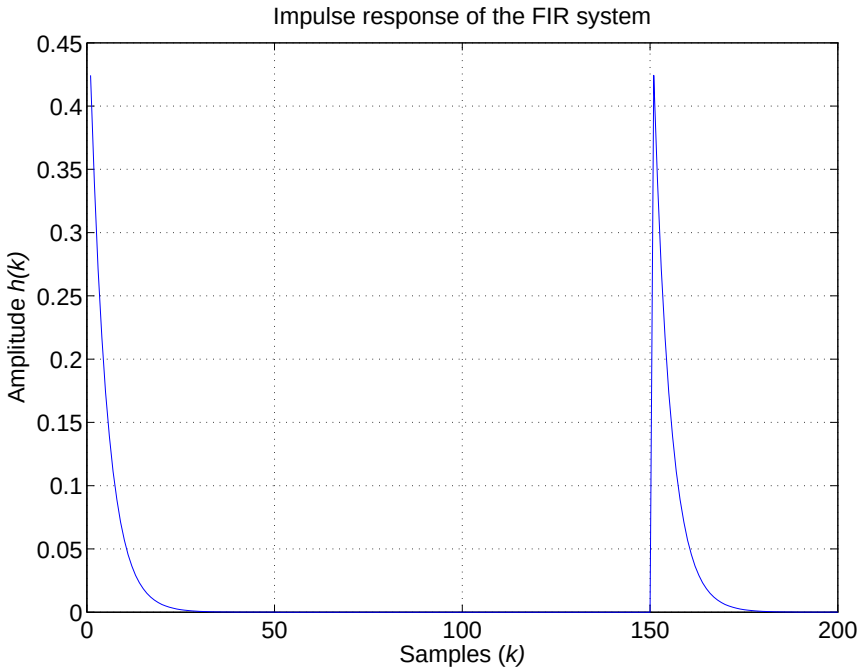



Figure 6.1.: Impulse response of the FIR channel

## 6.2. Channel Description

As it has been said, a FIR channel consisting of two delayed exponential power delay profiles is used as an UWB-like test channel. This impulse response is given by,

$$h(k) = 0.8^k u(k) + 0.8^{k-150} u(k-150),\ k = 0,\ 1,\ \cdots 200 \tag{6.1}$$

$$H(z) = \left[\sum_{k=0}^{+\infty} (0.8)^k z^{-k}\right] \underbrace{\left(1 + z^{-150}\right)}_{\text{sin-function}} \tag{6.2}$$

denoting by $u(k)$ the step function

$$u(k) = \begin{cases} 1 & : k \geq 0 \\ 0 & : k < 0 \end{cases}$$

the impulse response is represented on Figure 6.1 and its frequency response on figure 6.2.

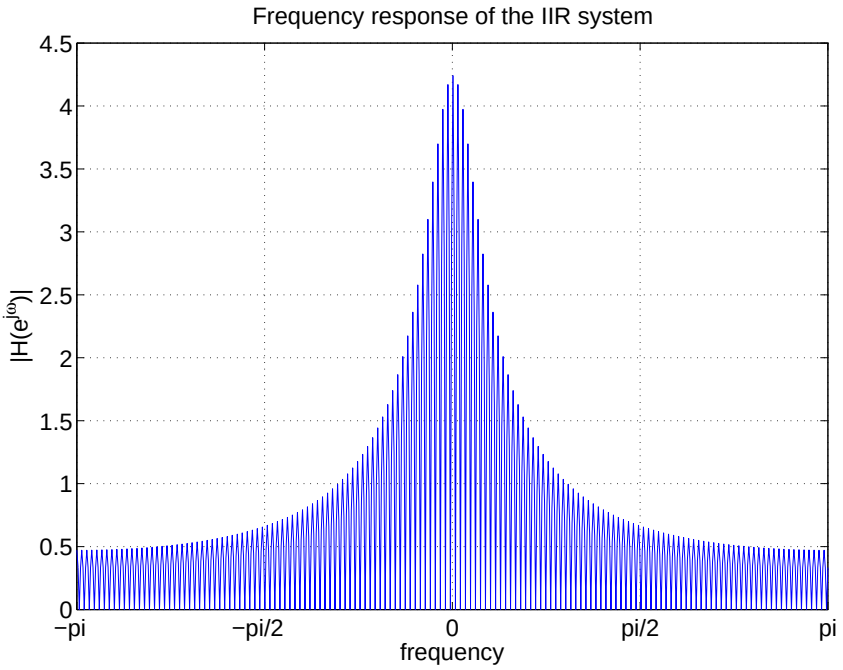



Figure 6.2.: Frequency response of the FIR channel

## 6.3. Mean Square Error

In Figure 6.3 can be seen the normalized MSE for the FIR channel presented before. As seen before, a normalized MSE of 0 dB corresponds to an error of unity, i.e., there is no approximation and a normalized MSE of $-\infty$ dB correspond to an error of zero, i.e., perfect approximation. The pole position varies from -1 to 1, noting that a pole position of 0 corresponds to the FIR filter case.

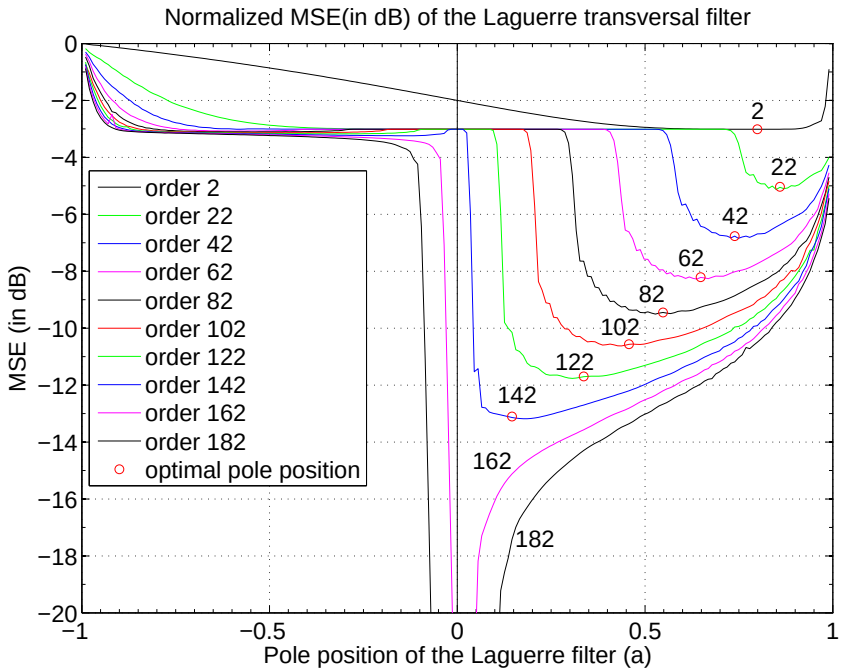



Figure 6.3.: Normalized MSE of FIR channel estimate as a function of the pole position for different filter orders (note: $a = 0 \triangleq$ FIR filter ).

Table 6.1.: NMSE for Laguerre filter of order 60 with different pole positions

| filter | pole position ($a$) | NMSE (in dB) |
|---|---|---|
| FIR transversal filter | 0 | $-3$ dB |
| Laguerre filter | 0.5 | $-7.2$ dB |
| Laguerre filter | 0.7 | $-8.1$ dB |
| Laguerre filter | 0.64 (optimal pole position) | $-8.2$ dB |

There are many interesting things in this Figure. The first of them is that for the FIR case ($a = 0$) the NMSE is stuck in $-3$ dB until the order reaches 150, which is rather easy to understand intuitively because taking a look at Figure 6.1, with a FIR filter of order lower than 150 only the first power delay profile can be gathered, which corresponds to the half of the channel energy, i.e., 3 dB. Another interesting thing is that the optimal pole position begins in the channel time constant, in this case 0.8, (see Equation 6.2) and then moves closer to unity and slowly decreases until it reaches the zero pole position when the order of the filter is comparable to the channel length, which follows from the fact that for small filter orders, longer sequences are needed in order to correctly approximating the long impulse response and therefore large pole positions are best suited and in the opposite case for filter orders close to the channel length, small pole positions are better because they offer a finer temporal resolution which results in better approximation. To see more clearly this results, the information is reorganized to produce Figure 6.4, which instead of a function of pole position the NMSE is represented as a function of filter order for different pole positions. Also the optimal pole position is extracted and represented in Figure 6.5.

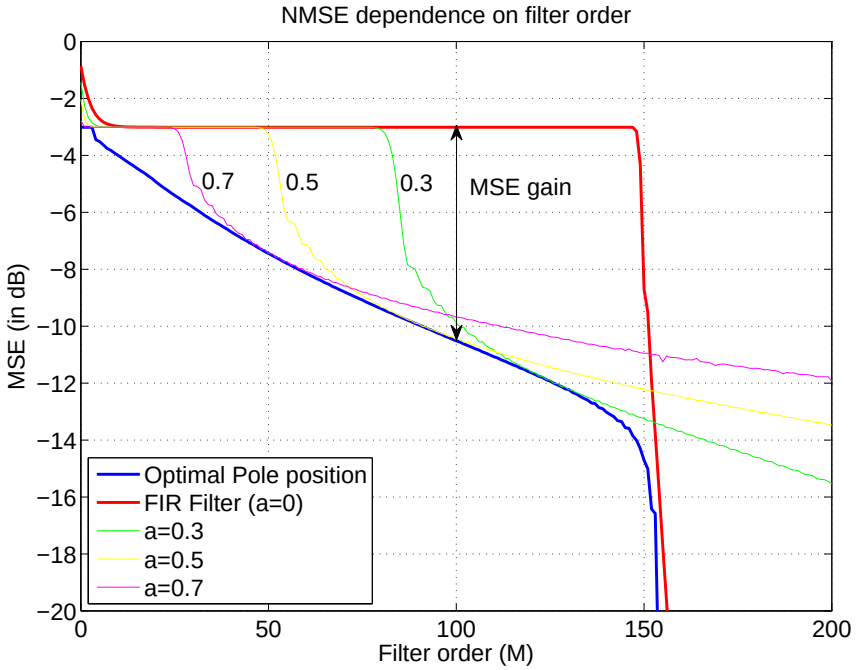



Figure 6.4.: Normalized MSE of FIR channel estimate as a function of the filter order for different pole positions (note: $a = 0 \stackrel{\triangle}{=}$ FIR filter ).

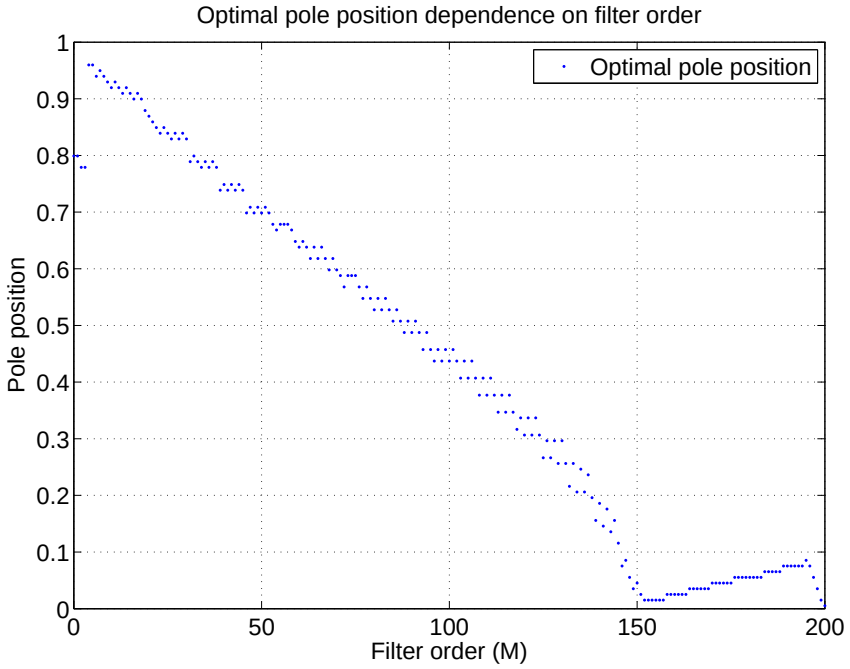



Figure 6.5.: Variation of the optimal pole position on the real axis of the Laguerre filter as a function of filter order.

Regarding to Figure 6.4, it can be seen that there is an interval in which the Laguerre filter performs better than the FIR filter, this interval is delimited by the filter order which is comparable to the channel length ( 150 in this case). Taking a look the optimal pole position in Figure

6.5, can be seen how it varies from a value close to the unity when the filter order is small (with a small order filter, longer sequences are needed to approximate a long channel response which correspond with pole position close to the unity, intuitively speaking) to a zero value when the order is equal or comparable with the impulse response length.

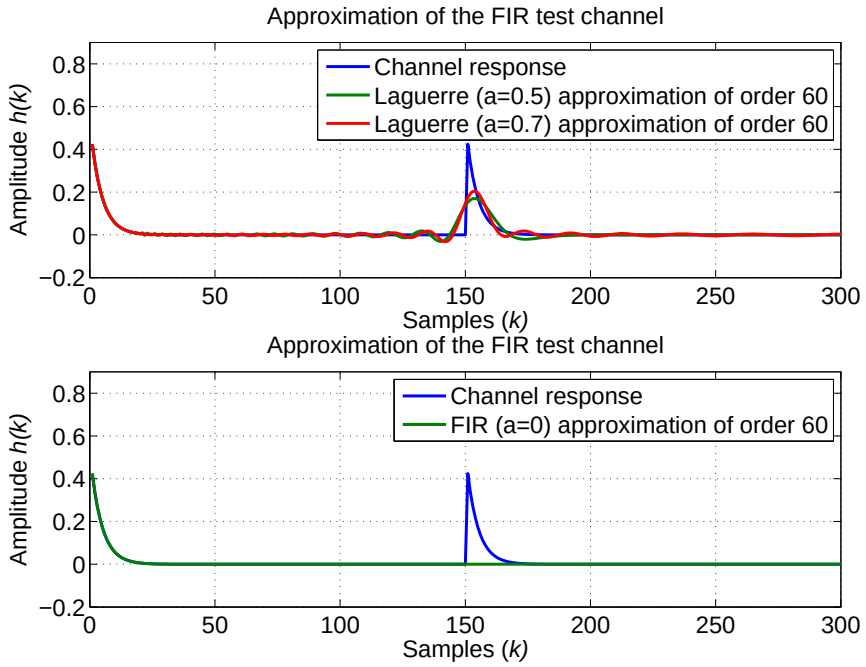



Figure 6.6.: Approximation of the FIR channel impulse response for FIR and Laguerre filter of order 60.

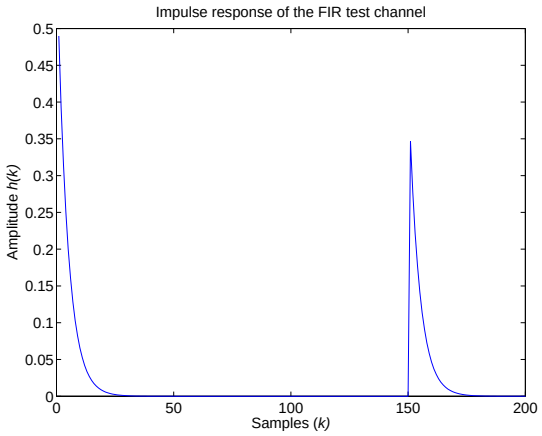
In Figure 6.6, it can be seen how the Laguerre and the FIR filter approximate the channel response with a filter order of 60. The FIR channel can only gather the first power delay profile, but the Laguerre filters can reach the second power delay profile, because they have a longer time spread. It also can be appreciated one problem with Laguerre filter regarding to the abrupt changes (jumps or discontinuities) in the channel response, generating some previous and posterior oscillations that should not be there. The source of this problem is the low-pass behavior of the Laguerre filters (remembering that they are constructed as a low-pass filter, followed by a cascade of all-pass filters), and the fact that an abrupt change contains intrinsically high frequencies. This fact can be seen in the figure because the Laguerre approximation has big oscillations near the abrupt change of the channel deteriorating the resulting MSE. This phenomenon it similar to the Gibbs phenomenon in Fourier analysis.

If we try with the same test channel but with the amplitude ( or energy) of the second cluster reduced by a factor of $\alpha$ (or $\alpha^2$ for the energy), given by

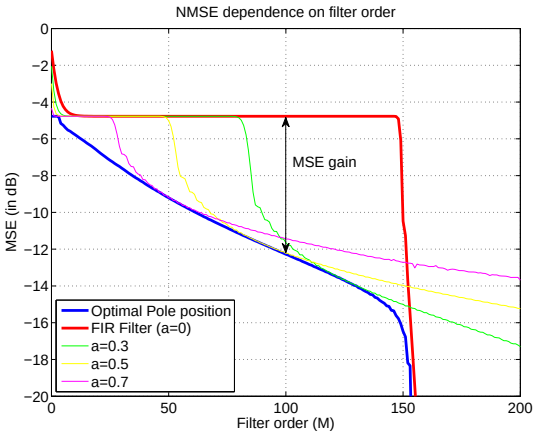
$$h(k) = 0.8^k u(k) + \alpha 0.8^{k-150} u(k-150),\ k = 0,\ 1,\ \cdots 200 \tag{6.3}$$

then the NMSE suffers from a displacement as seen in Figure 6.7.

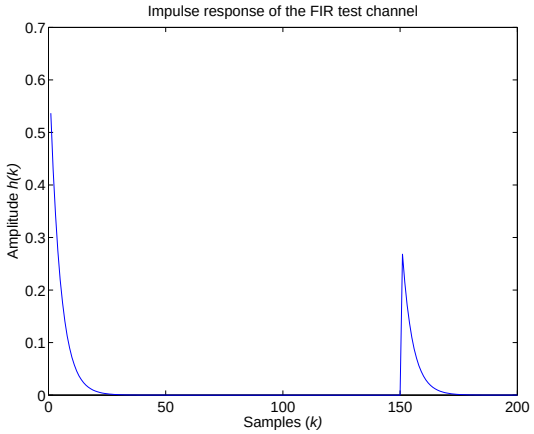
The UWB channels generated using the standard model are complex channels due to a random phase term applied to each multipath component of the channel, as it will be seen in the next Chapter. This makes the real and imaginary part of the channel to have an oscillatory behavior. A test channel closer to the UWB channels is the one shown in Figure 6.8, which consist of the same channel used in this Chapter but with a random sign which gives the channel an oscillatory



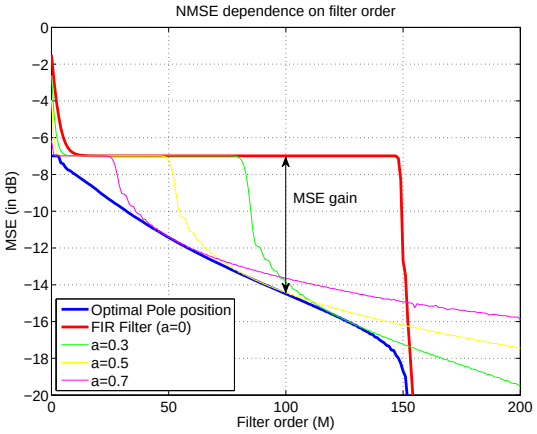

(a) Impulse response for $\alpha = \frac{1}{\sqrt{2}}$




(b) NMSE for $\alpha = \frac{1}{\sqrt{2}}$




(c) Impulse response for $\alpha = 0.5$




(d) NMSE for $\alpha = 0.5$

Figure 6.7.: Normalized MSE of FIR channel estimate as a function of the filter order for different pole positions (note: $a = 0 \triangleq$ FIR filter ).

behavior.

The resulting MSE is shown in Figures 6.9 and 6.10, which shows that the negative poles are best suited for this kind of oscillatory channels.

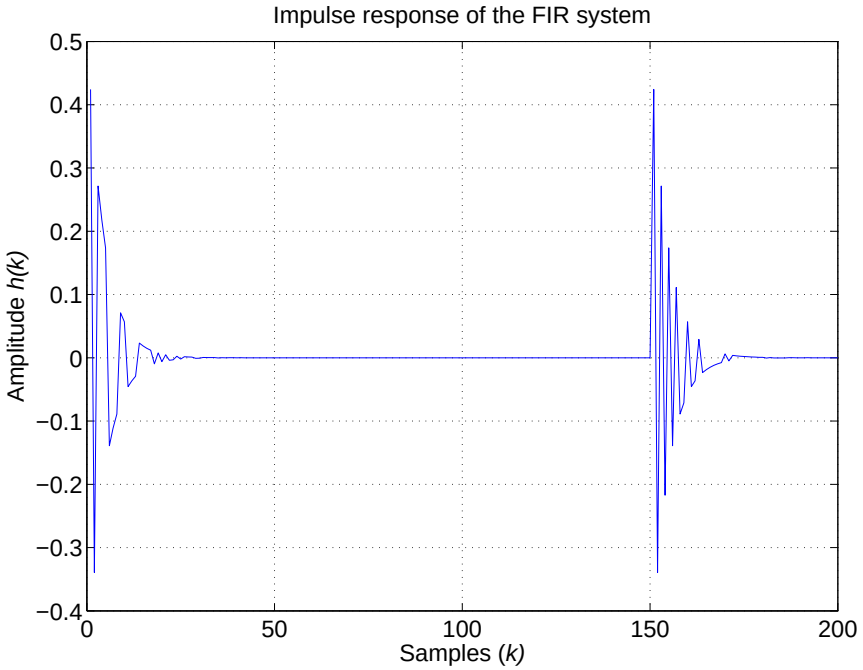



Figure 6.8.: Impulse response of the FIR channel

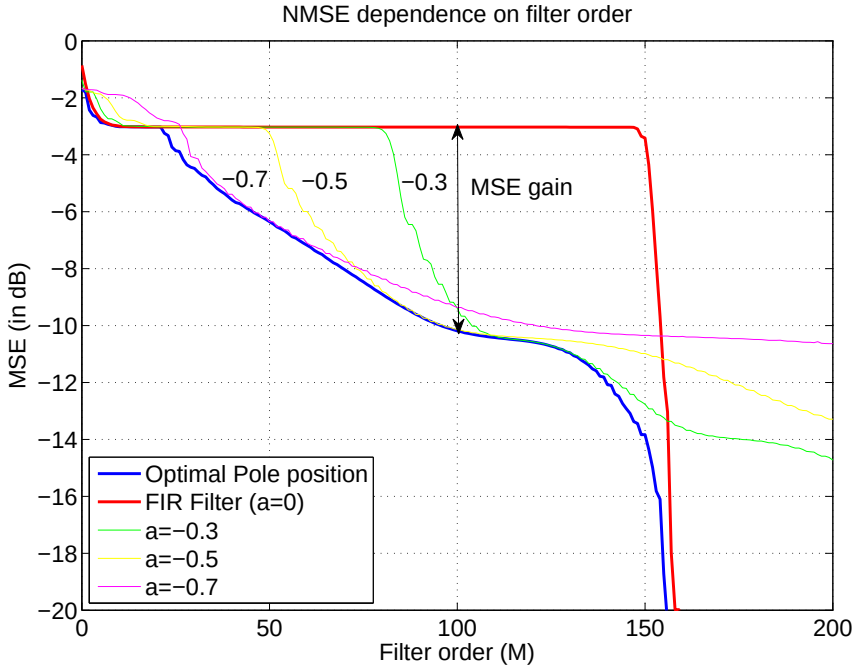



Figure 6.10.: Normalized MSE of FIR channel estimate as a function of the filter order for different pole positions (note: $a = 0 \triangleq$ FIR filter ).

What can be derived from all these results is that there is a filter order interval in which Laguerre filters provide a lower MSE than FIR filters just choosing the correct position of the pole of the system. This interval is delimited by the filter order which is comparable with the

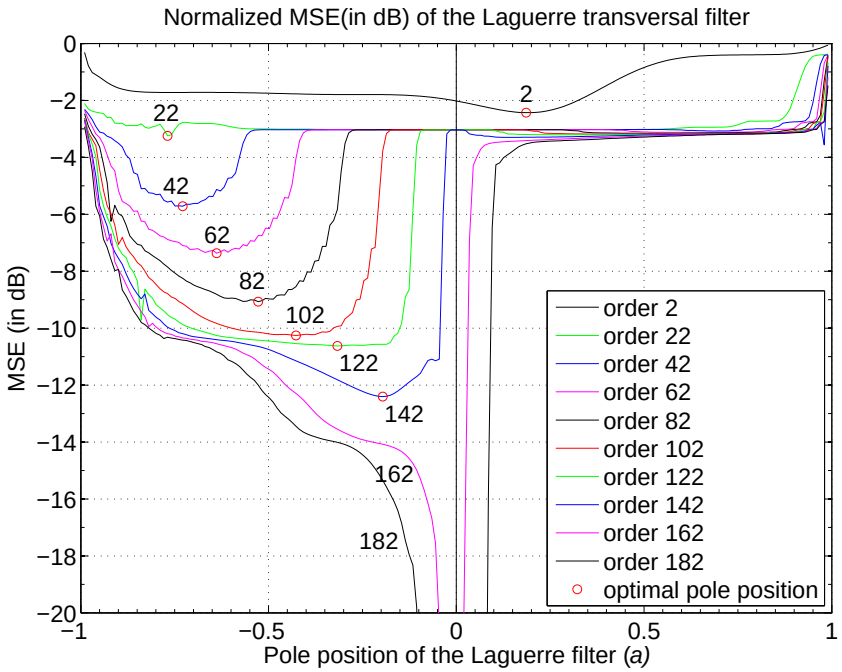



Figure 6.9.: Normalized MSE of FIR channel estimate as a function of the pole position for different filter orders (note: $a = 0 \stackrel{\triangle}{=}$ FIR filter ).

length of the channel. The pole position changes with the filter order, going from a value close to the unity for low orders (in order to have longer sequences allowing to approximate longer impulse responses) to a value close to zero for orders comparable with the channel length (in which the best filter is the FIR filter because it provides the finest time resolution).

# 7. UWB Channels

## 7.1. Introduction

Now that we made a theoretical approach with the analysis of the idealized UWB FIR test channel of Chapter 6, we will present the UWB channels and analyze the behavior of the Laguerre filters. Ultrawideband (UWB) communications use wide transmission bandwidths and very high sampling rates, which results in some desirable capabilities, like accurate position location and ranging, high multiple access capability, covert communications, .... They are in constant development in the recent years for both the business and the research community.

Because of their power restrictions on transmission they are best suited for sensor networks and personal area networks (PANs), that is, short-range communications. On February 21, 2007, the Commission of the European Communities published its "Commission Decision" on the implementation of an UWB radio regulatory framework for the European Union (see [4]) and so did the American Federal Communications Commission (FCC) for the US in February 2001 (see [2]), allowing the transmission of UWB signals if certain transmission power restrictions are fulfilled.

The IEEE established a standardization group, IEEE 802.15.4, which specifies the physical layer and media access control for low-rate wireless personal area networks (LR-WPANs). IEEE 802.15.4a is an amendment to IEEE 802.15.4 which specifies two additional physical layers added to the original. One of these two additional physical layers is the one corresponding to the UWB channel.

There are many channel models available in the literature for wireless communications. These models are based on a signal that has propagated through a wireless channel consisting of multiple echoes (replicas) of the originally transmitted signal, also known as multipath propagation. Each multipath component has its own delay and attenuation value. The main focus of the channel modeling is the correct description and modeling of the parameters that characterize the multipath component. As an example, for narrowband wireless systems the most widely used channel model is the flat Rayleigh-fading channel.

In UWB systems the signals cover a bandwidth of almost 10 GHz, which causes new effects. One can be that only few multipath components overlap because of the better resolution of the system (resolvable run length is 3 cm), what makes that the statistics are no longer Rayleigh, and there can be some sampling intervals that are "empty" (contain no receive energy).

## 7.2. UWB IEEE 802.15.4a Standard model

This channel model (given by the document [8]) provides a model for UWB channels covering the frequency range from 2 to 10 GHz and for indoor residential, indoor office, industrial, outdoor, and open outdoor environments ( with distinction between line of sight (LOS) and non line of sight (NLOS) in most cases). The document also provides a MATLAB® implementation of the model.

There are some primary characteristics of the channel that the model has to match. These characteristics are

- Delay spread

- Power delay profile
- Number of multipath components

The model finally adopted is based on a Saleh-Valenzuela model (see [10]). In particular, the multipath arrivals were modeled in clusters rather than in a continuum, as is common for narrowband channels. This is a result of the short sampling interval. In addition, the multipath arrivals were grouped into two different categories: a cluster arrival and a ray arrival within a cluster. Also the power delay profile is modeled as an exponential decay within the clusters and for the clusters themselves.

## 7.3. Channel estimation

In this study the environments for residential NLOS and office NLOS are used for the channel estimation. These two environments match with the desired channel behavior consisting of clustered echoes and without a dominant peak that motivated the FIR UWB-like test channel of chapter 6. The channel estimation is carried out with the method used in Chapters 5 and 6, described in Chapter 4. One remarkable thing to add is that the model provides complex channels with a random phase, so the represented channels are represented in modulus most of the cases.

### 7.3.1. Residential NLOS

The channel model for the environment of residential NLOS is based on measurements that cover a range from 7-20m, up to 10 GHz. Several realizations of the stochastic model can be viewed in 7.1

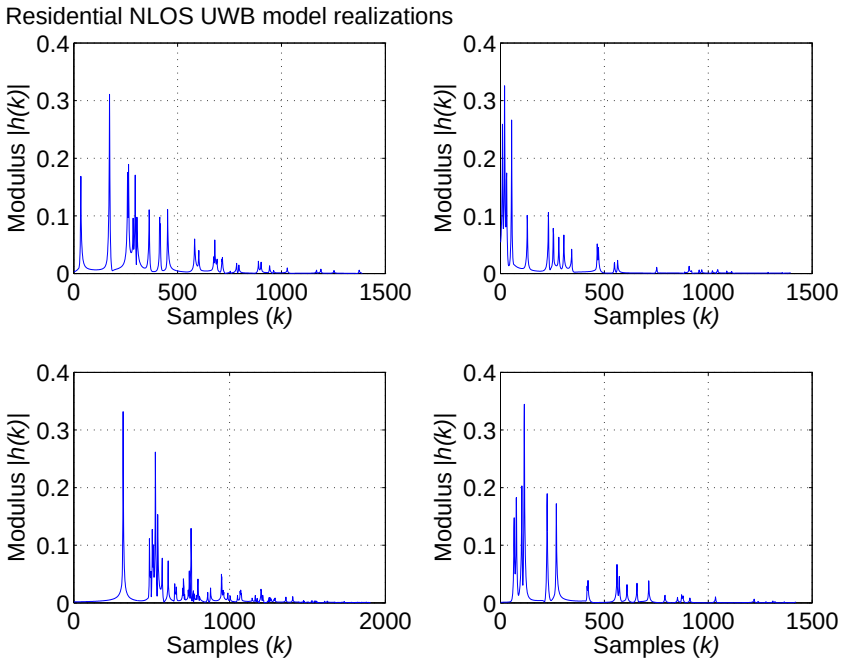



Figure 7.1.: Channel realizations for the Residential NLOS environment of the UWB channel model

We focus first on one realization of the model, like the one shown in Figure 7.2, we can calculate the MSE with Laguerre filters with different orders or pole position. Looking at Figure 7.2, it can be seen that the effective length of the channel impulse response is about 1000 samples (which will allow to estimate a good choice for the pole position $a$, as seen on Chapter 2). A filter of order 200 or 300 should be enough to correctly approximate this impulse response.

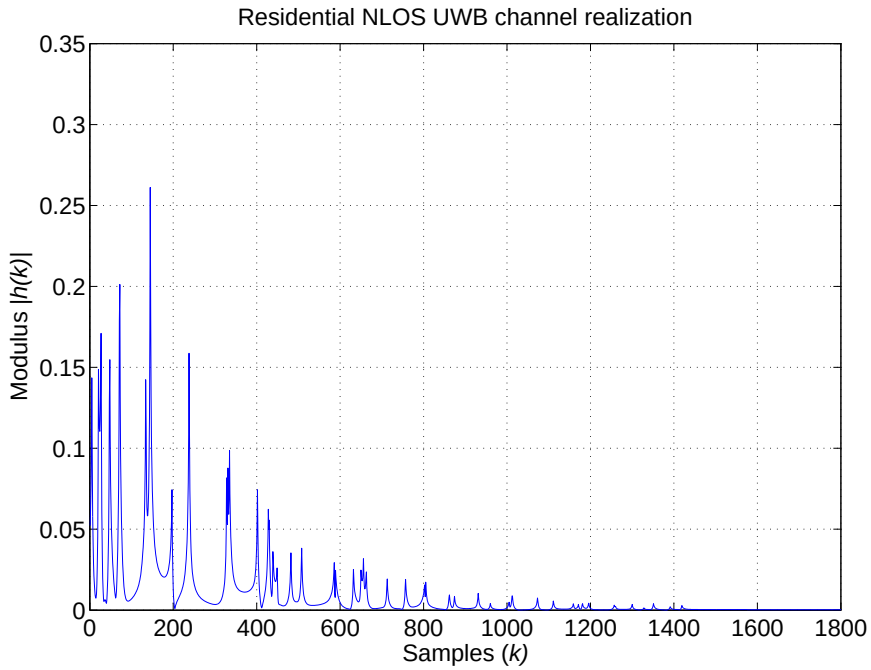



Figure 7.2.: Channel realizations for the Residential NLOS environment of the UWB channel model

In Figure 7.3 the NMSE for this channel model realization for filters with several orders from 50 to 300 are represented as a function of the pole position $a$. It can be seen how Laguerre filters with an appropriate pole position can perform better than the FIR filter for this realization of the channel model. It also can be seen that the optimal pole positions are negative, this occurs due to the fact that the stochastic model gives a complex channel with a random complex phase term for each multipath component which makes the real and imaginary parts of the channel to have an oscillatory behavior and for that reason a negative pole is best suited for this channel because the negative pole introduces a $(-1)^k$ term that generates an oscillatory behavior as seen on Figure 7.4.

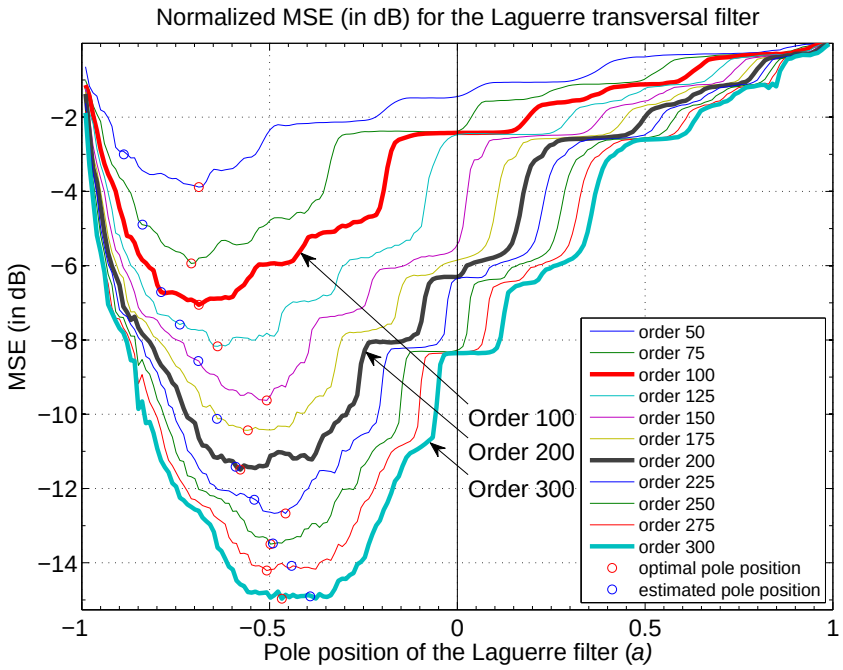
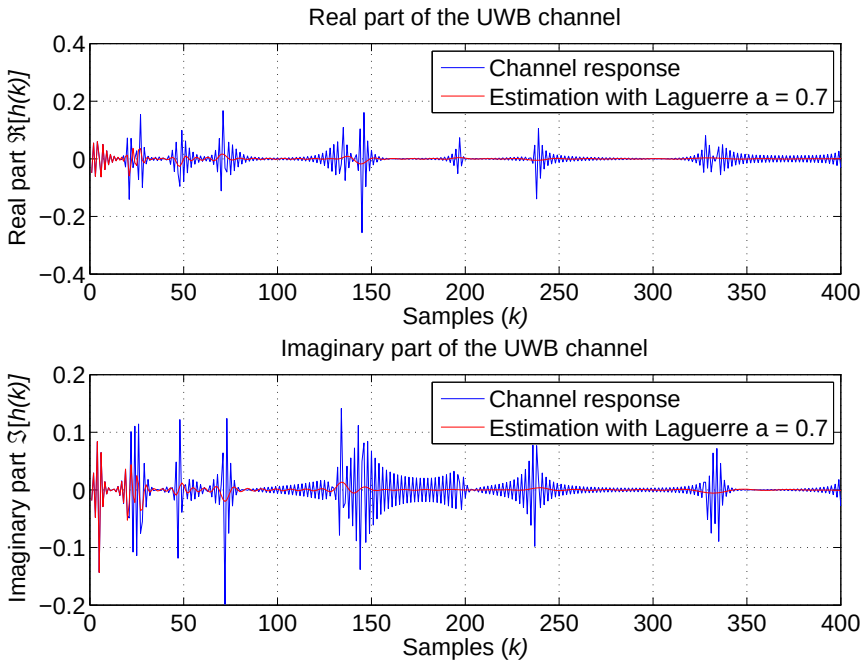
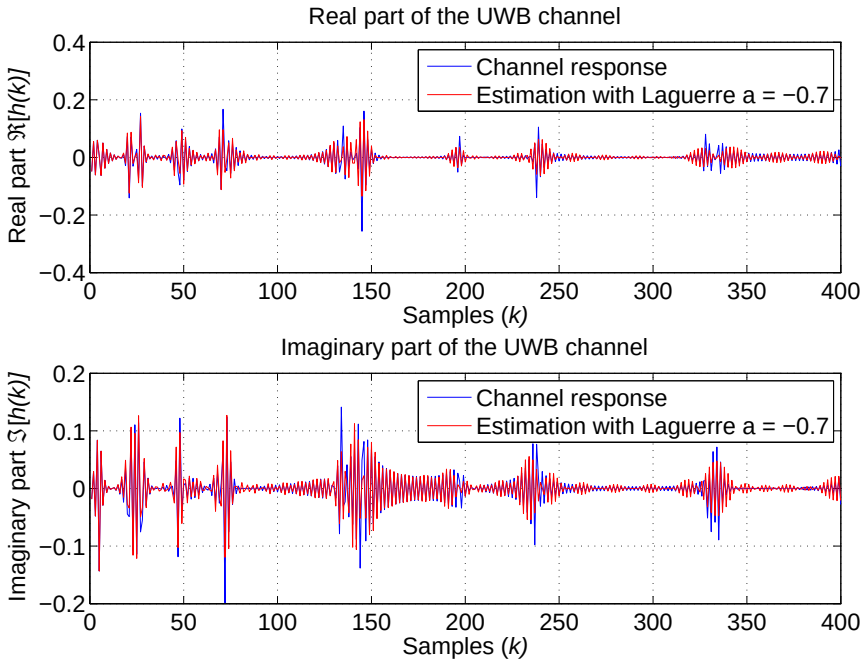



Figure 7.3.: Normalized mean MSE of UWB channels estimate as a function of the pole position for different filter orders (note: $a = 0 \triangleq$ FIR filter ).




(a) Estimation for positive pole $a = 0.7$




(b) Estimation for negative pole $a = -0.7$

Figure 7.4.: Effect of the sign of the pole position for a complex channel estimation with a Laguerre filter of order 100.

Back in Figure 7.3, it also can be seen that the optimal pole positions and the estimated ones (using the criterion described before in Chapter 2, which takes advantage of the previous knowledge of the channel length and makes the Laguerre sequences to cover that length), which reveals that the estimation for the optimal pole is not perfect but gives an adequately close

value which results in a small deterioration of the resulting MSE. An equivalent representation of Figure 7.3 as function of the filter order is shown in Figure 7.5, which allows to see these results more clearly. Here an improvement of almost 5 dB in the MSE can be observed and it also can be appreciated how good is the estimation of the pole position because the MSE which results from the choice of the estimated optimal pole position for each filter order is close to the optimal pole position MSE curve.

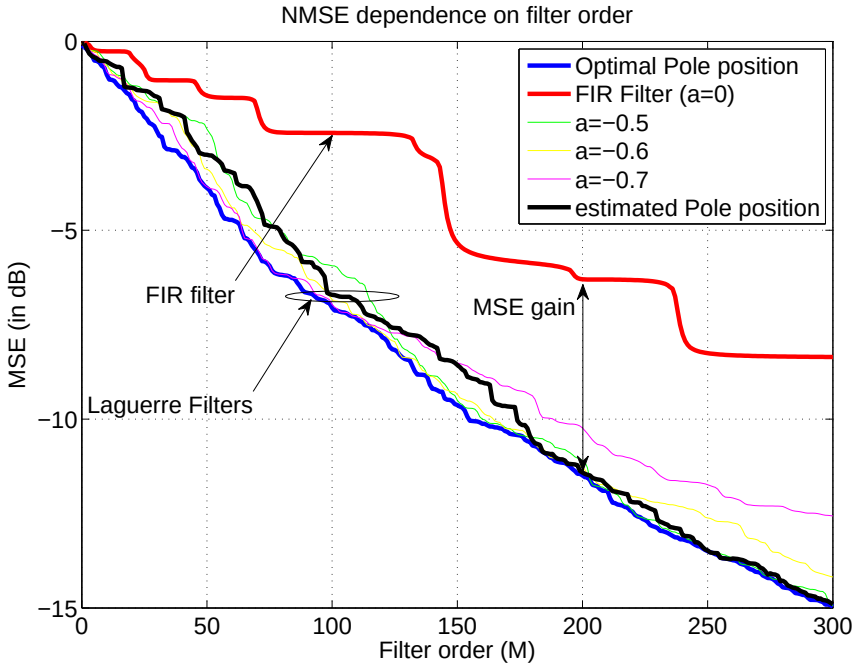



Figure 7.5.: Normalized MSE of UWB realization estimate as a function of the filter order for different pole positions.(note: $a = 0 \triangleq$ FIR filter ).

Some numerical values are shown in Table 7.1, which contains the NMSE values for the filters of orders 100, 200 and 300 with different pole positions (remember that FIR filter is a Laguerre filter with pole position $a = 0$). If we calculate the MSE gain by the different Laguerre filters in comparison with the FIR filter as $\Delta_{\text{MSE}} = \text{MSE}^{\text{FIR}} - \text{MSE}^{Lag}$, the results can be found in Table 7.2, which shows that for the estimated pole position, a MSE gain of about 5 dB can be obtained with Laguerre filters.

Table 7.1.: NMSE achieved by approximating an UWB channel realization with filters of different orders.

| filter | pole position ($a$) | Order 100 | Order 200 | Order 300 |
|---|---|---|---|---|
| FIR transversal filter | 0 | -2.4 dB | -6.3 dB | -8.4 dB |
| Laguerre filter | $-0.5$ | -5.9 dB | -11.1 dB | -14.9 dB |
| Laguerre filter | $-0.7$ | -7.0 dB | -10.2 dB | -12.6 dB |
| Laguerre filter | *Estimated* | -6.7 dB | -11.4 | -14.9 dB |
| Laguerre filter | *Optimal* | -7.1 dB | -11.5 | -15.0 dB |

Table 7.2.: $\Delta_{\text{MSE}}$ or MSE gain achieved by approximating an UWB channel realization with filters of different orders and pole positions.

| pole position ($a$) | Order 100 | Order 200 | Order 300 |
|---|---|---|---|
| $-0.5$ | 3.5 dB | 4.8 dB | 6.5 dB |
| $-0.7$ | 4.6 dB | 3.9 dB | 4.2 dB |
| *Estimated* | 4.3 dB | 5.1 dB | 6.5 dB |
| *Optimal* | 4.7 dB | 5.2 dB | 6.6 dB |

The optimal pole position for each filter order is shown in Figure 7.6 in addition to the estimated pole position.

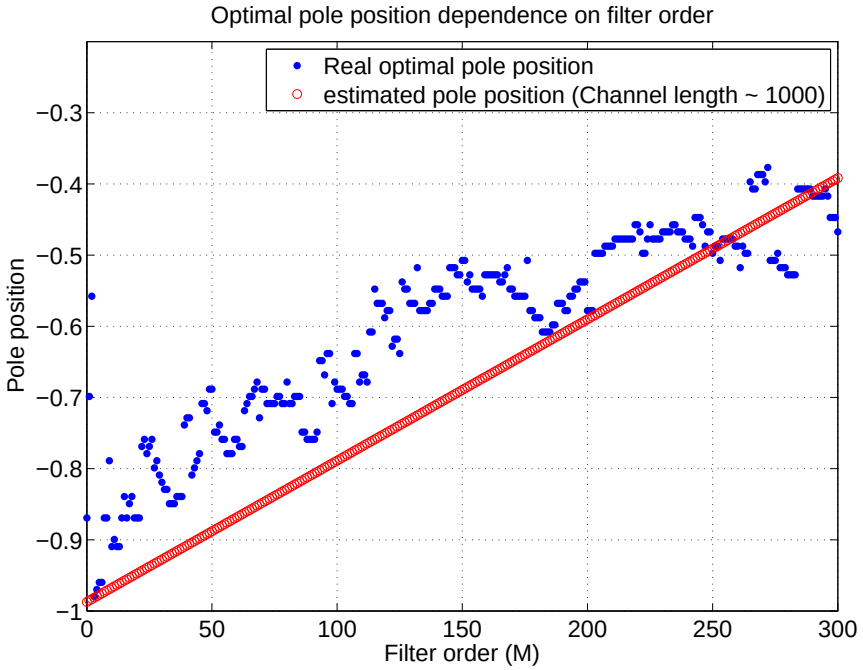



Figure 7.6.: Variation of the optimal (and estimated) pole position for the Laguerre filter as a function of the filter order.

Now that the individual case of a channel model realization is known, a more general case can be obtained by analyzing many realizations of the model as a stochastic experiment. The estimation is accomplished by a Laguerre transversal filter of order up to 300 for 500 realizations of the UWB stochastic channel model.

First of all, the mean NMSE for the 500 realizations of the channel model is represented as a function of the pole position for different orders in Figure 7.7 and as a function of the filter order for different pole positions in Figure 7.8. Here an MSE decrement of about 4-5 dB for the Laguerre filters can be appreciated. Once again the optimal pole positions are negative for the same reasons explained before and the estimated pole position seems to be a good choice because the optimum pole position is close to the estimated one and results in a small deterioration of the MSE. If we take a look at Figure 7.8, the MSE curve for the estimated pole position is close to the MSE curve for the optimal pole position. The optimal pole position and the estimated are shown in Figure 7.9.

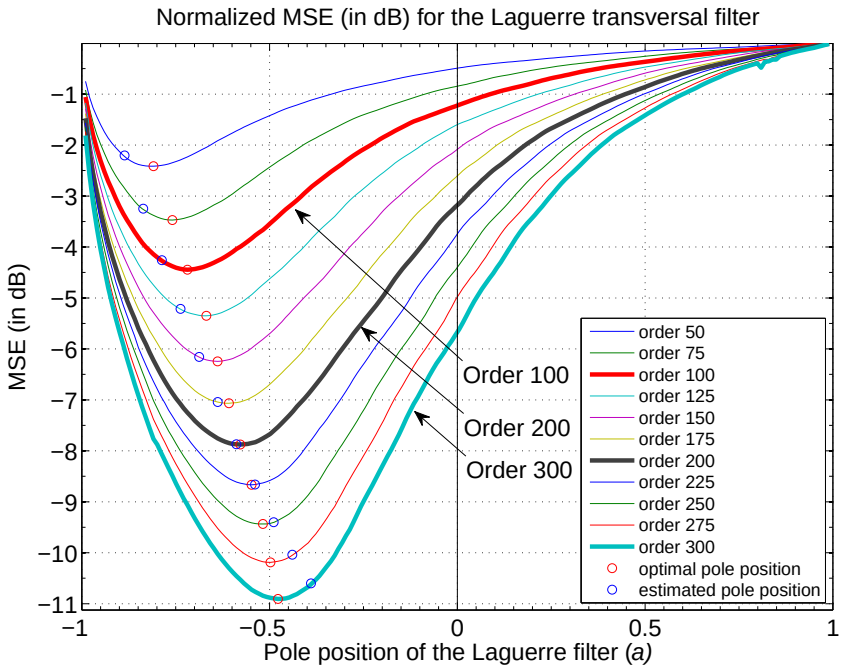



Figure 7.7.: Normalized mean MSE of UWB channels estimate as a function of the pole position for different filter orders (note: $a = 0 \triangleq$ FIR filter ).

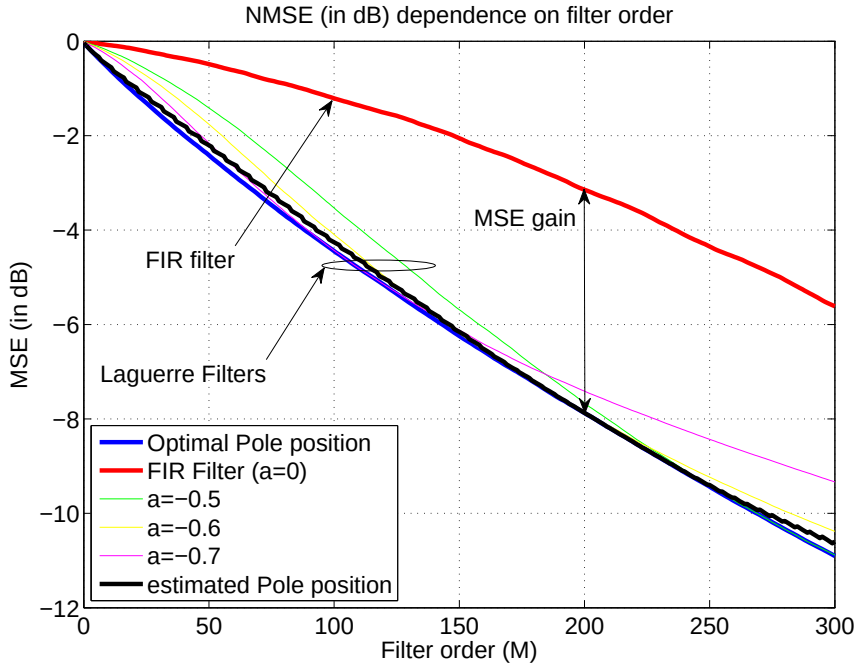



Figure 7.8.: Normalized mean MSE of UWB channels estimate as a function of the filter order for different pole positions.(note: $a = 0 \stackrel{\triangle}{=}$ FIR filter ).

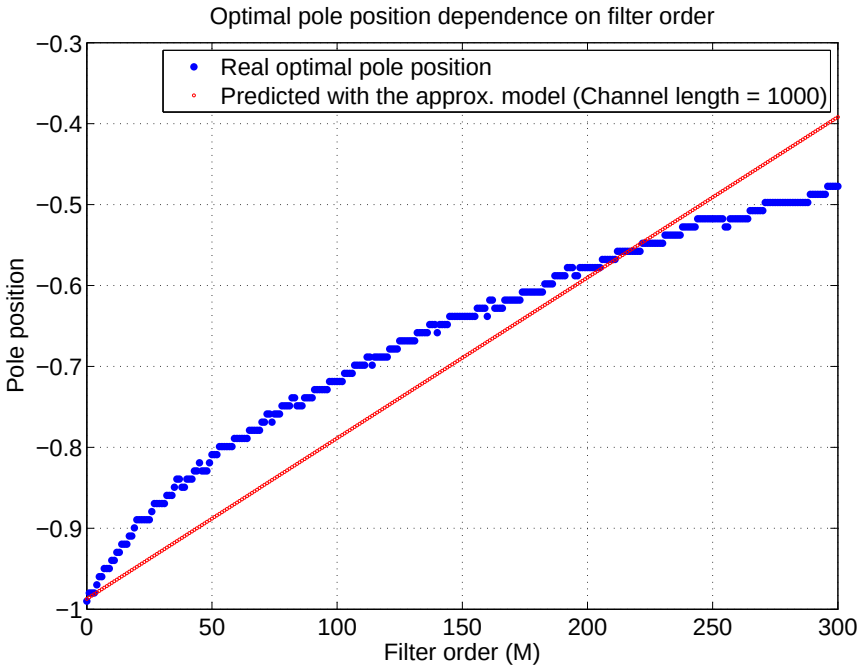



Figure 7.9.: Variation of the optimal (and estimated) pole position for the Laguerre filter as a function of the filter order.

The experimental probability density function (pdf) for the MSE is obtained for filters of orders 100, 200 and 300 and pole positions $a$ =0, -0.5, -0.7 and the estimated and optimal for each order (note that $a = 0$ corresponds to the FIR transversal filter). The reason to choose -0.7 and -0.5 as pole positions is that seeing the mean MSE behavior in Figure 7.7 a pole of -0.7 or -0.5 seems to be a wise election, and another important thing to note is that the filter corresponding to $a = -0.5$ involves an efficient computational implementation because a multiplication/division by a power of 2, is quite simple in digital systems (it's only adding zero bits or dropping bits). This results are summarized in Table 7.3, and represented in Figures 7.10, 7.11 and 7.12 for filter of orders 100, 200 and 300 respectively. The figures show that Laguerre filters in the general case indeed perform better than FIR filter by offering a smaller MSE value for a given filter order, giving a mean improvement of 3.2 dB, 4.7 dB and 5 dB for filters of order 100, 200 and 300 respectively with an optimal pole position as seen in Table 7.4 in comparison with FIR filters.

Table 7.3.: NMSE (in dB) achieved by approximating UWB channels with filters of different orders for several pole positions.

| | | Order 100 | | Order 200 | | Order 300 | |
|---|---|---|---|---|---|---|---|
| filter | pole position ($a$) | Mean | Std | Mean | Std | Mean | Std |
| FIR | 0 | -1.5 dB | 1.7 dB | -3.7 dB | 2.5 dB | -6.6 dB | 3.0 dB |
| Laguerre | -0.5 | -4.0 dB | 2.0 dB | -8.3 dB | 2.4 dB | -11.6 dB | 2.5 dB |
| Laguerre | -0.7 | -4.7 dB | 1.7 dB | -7.7 dB | 1.9 dB | -9.6 dB | 2.0 dB |
| Laguerre | Estimated | -4.5 dB | 1.5 dB | -8.4 | 2.1 dB | -11.5 dB | 2.7 dB |
| Laguerre | Optimal | -4.7 dB | 1.7 dB | -8.4 | 2.2 dB | -11.6 dB | 2.5 dB |

Table 7.4.: Mean $\Delta_{\text{MSE}}$ or MSE gain achieved by approximating UWB channels with filters of different orders and pole positions.

| pole position ($a$) | Order 100 | Order 200 | Order 300 |
|---|---|---|---|
| -0.5 | 2.5 dB | 4.6 dB | 5 dB |
| -0.7 | 3.2 dB | 4.0 dB | 3 dB |
| Estimated | 3.0 dB | 4.7 dB | 4.9 dB |
| Optimal | 3.2 dB | 4.7 dB | 5 dB |

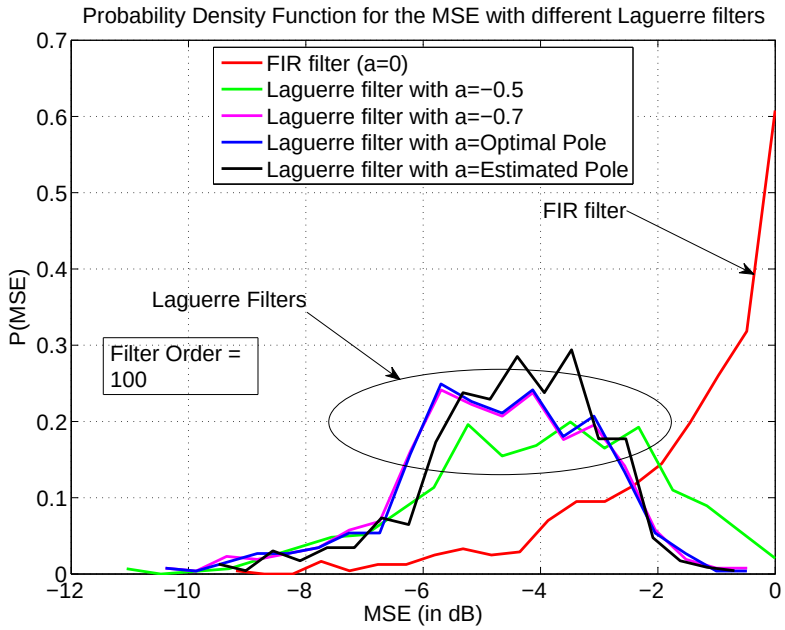



Figure 7.10.: Normalized MSE probability density function for a Laguerre filter of order 100 for different pole positions.

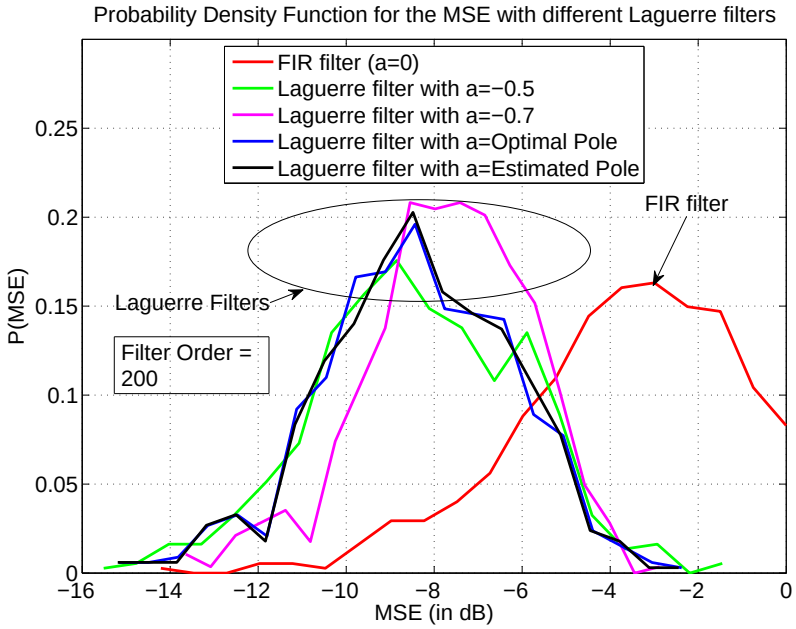



Figure 7.11.: Normalized MSE probability density function for a Laguerre filter of order 200 for different pole positions.

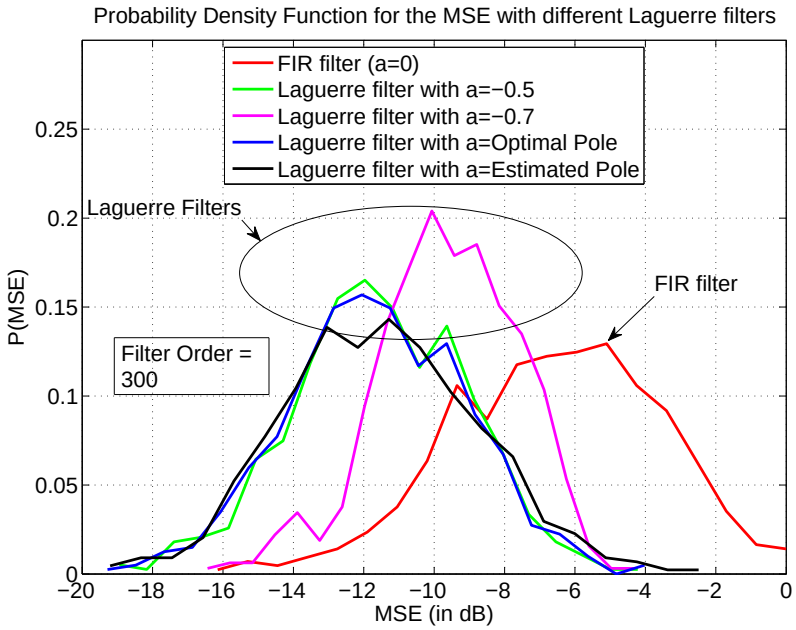
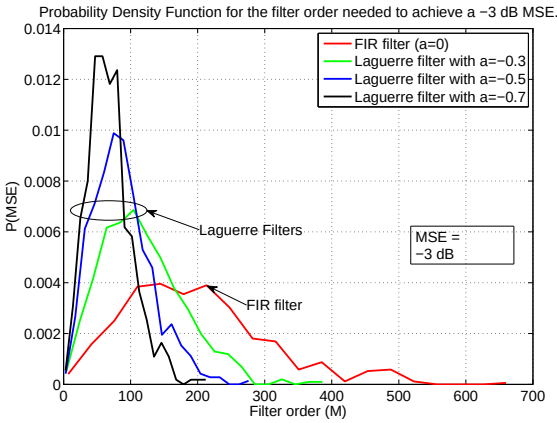



Figure 7.12.: Normalized MSE probability density function for a Laguerre filter of order 300 for different pole positions.

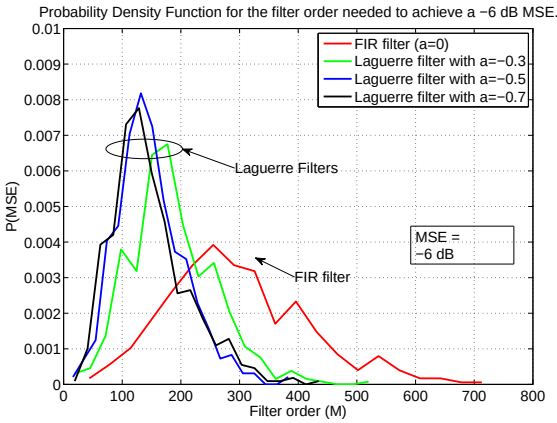
Table 7.5.: Mean filter order needed for Laguerre filters in order to achieve a specific NMSE.

| | | -3dB NMSE | | -6dB NMSE | | -10dB NMSE | |
|---|---|---|---|---|---|---|---|
| filter | pole position ($a$) | Mean | Std | Mean | Std | Mean | Std |
| FIR transversal filter | 0 | 197 | 106 | 296 | 118 | 425 | 124 |
| Laguerre filter | $-0.3$ | 118 | 62 | 187 | 73 | 285 | 83 |
| Laguerre filter | $-0.5$ | 88 | 44 | 150 | 58 | 260 | 89 |
| Laguerre filter | $-0.7$ | 69 | 33 | 148 | 66 | 327 | 116 |

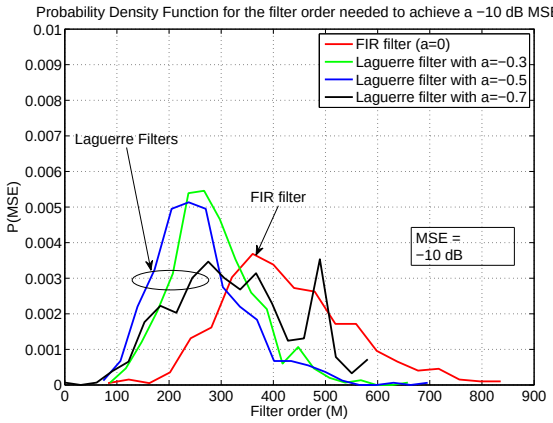
If the order needed to reach some specific MSE is calculated for a given pole position, the following results are obtained (see Figure 7.13 and table 7.5).




(a) NMSE = -3dB




(b) NMSE = -6dB




(c) NMSE = -10dB

Figure 7.13.: Order needed of the different filters for achieving a specific NMSE (-3dB, -6dB and -10 dB).

### 7.3.2. Office NLOS

This channel model is based on measurements that cover a distance from 3-28m, and a frequency range of 2-8 GHz. Several realizations of the stochastic model can be viewed in 7.14. This environment presents a less sparse channel with more chaotic behavior which can produce a worse performance of the Laguerre filter due to its low pass behavior.

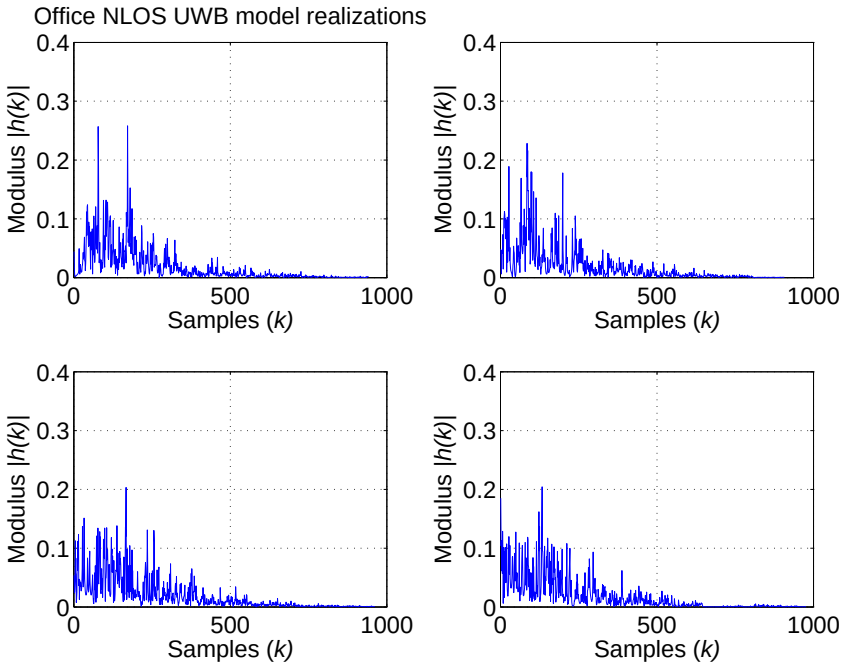



Figure 7.14.: Channel realizations for the Office NLOS environment of the UWB channel model

We focus first on one realization of the model, like the one shown in Figure 7.15, we can calculate the MSE with Laguerre filters with different orders or pole position. Looking at Figure 7.15, it can be seen that the effective length of the channel impulse response is about 500 samples. A filter of order 100 or 200 should be enough to correctly approximate this impulse response.

In Figure 7.16 the NMSE for this channel model realization for filters with several orders from 25 to 200 are represented as a function of the pole position $a$. Once again negative poles are best suited for this channel, for the same reasons explained before. This time, the Laguerre filters perform better than FIR filters, as a generalization of FIR filters, but the MSE improvement is about 2 dB, and it can be seen how the optimal pole position goes quickly to zero, i.e. the FIR filter. This fact is caused by the shorter length of the channel and the chaotic behavior which requires a finer resolution, remembering that the FIR filter provides the finest resolution. The estimated pole position seems to be a good estimate here too. An equivalent representation as a function of the filter order is shown in Figure 7.17.

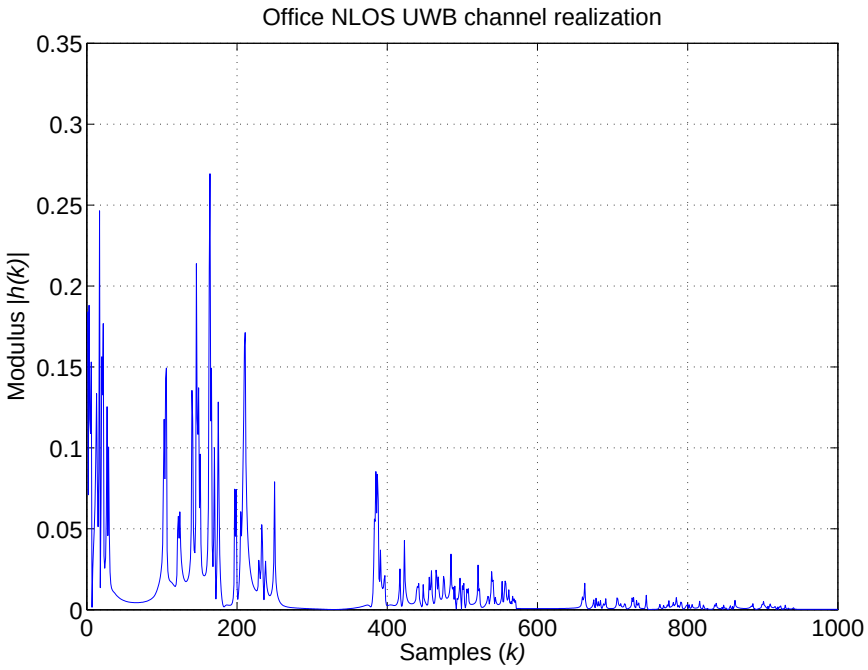



Figure 7.15.: Channel realizations for the Office NLOS environment of the UWB channel model

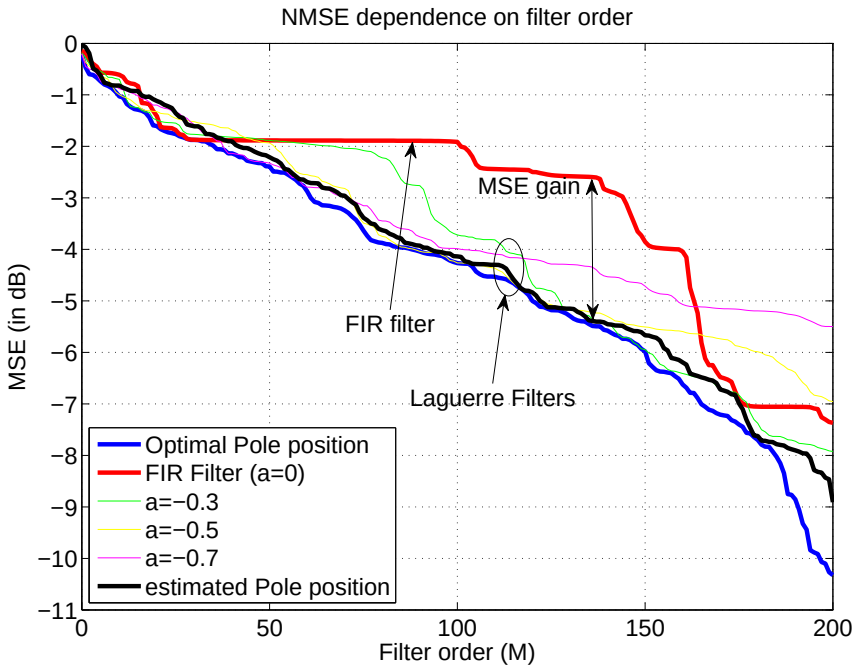



Figure 7.17.: Normalized MSE of UWB realization estimate as a function of the filter order for different pole positions.(note: $a = 0 \triangleq$ FIR filter ).

Some values are shown in Table 7.6, which contains the NMSE values for the filters of orders 50, 100 and 200 with different pole positions (remember that FIR filter is a Laguerre filter with pole position $a = 0$), and the MSE gain is represented in Table 7.7, showing a small decrement of the MSE for the Laguerre filters in comparison to the FIR filters. The optimal pole position

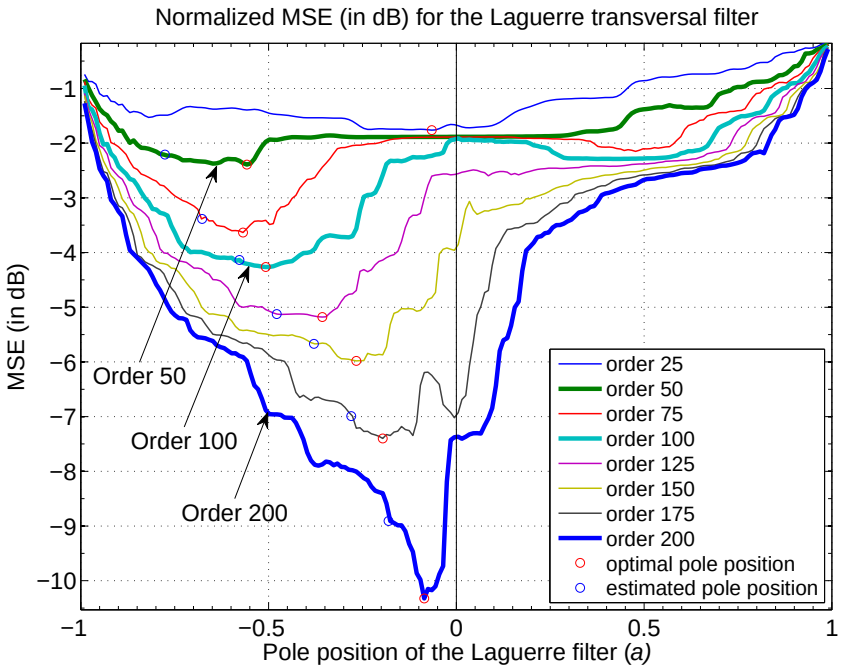



Figure 7.16.: Normalized mean MSE of UWB channels estimate as a function of the pole position for different filter orders (note: $a = 0 \stackrel{\triangle}{=}$ FIR filter ).

and the estimated one for each filter order is shown in Figure 7.18.

Table 7.6.: NMSE achieved by approximating an UWB channel realization with filters of different orders.

| filter | pole position ($a$) | Order 50 | Order 100 | Order 200 |
|---|---|---|---|---|
| FIR transversal filter | 0 | -1.9 dB | -1.9 dB | -7.4 dB |
| Laguerre filter | $-0.3$ | -1.9 dB | -3.7 dB | -7.9 dB |
| Laguerre filter | $-0.5$ | -1.9 dB | -4.3 dB | -7.0 dB |
| Laguerre filter | *Estimated* | -2.2 dB | -4.1 dB | -8.9 dB |
| Laguerre filter | *Optimal* | -2.4 dB | -4.3 dB | -10.3 dB |

Table 7.7.: $\Delta_{\text{MSE}}$ or MSE gain achieved by approximating an UWB channel realization with filters of different orders and pole positions.

| pole position ($a$) | Order 50 | Order 100 | Order 200 |
|---|---|---|---|
| $-0.3$ | 0.0 dB | 1.8 dB | 0.5 dB |
| $-0.5$ | 0.0 dB | 2.4 dB | -0.4 dB |
| *Estimated* | 0.1 dB | 2.2 dB | 1.5 dB |
| *Optimal* | 0.5 dB | 2.4 dB | 2.9 dB |

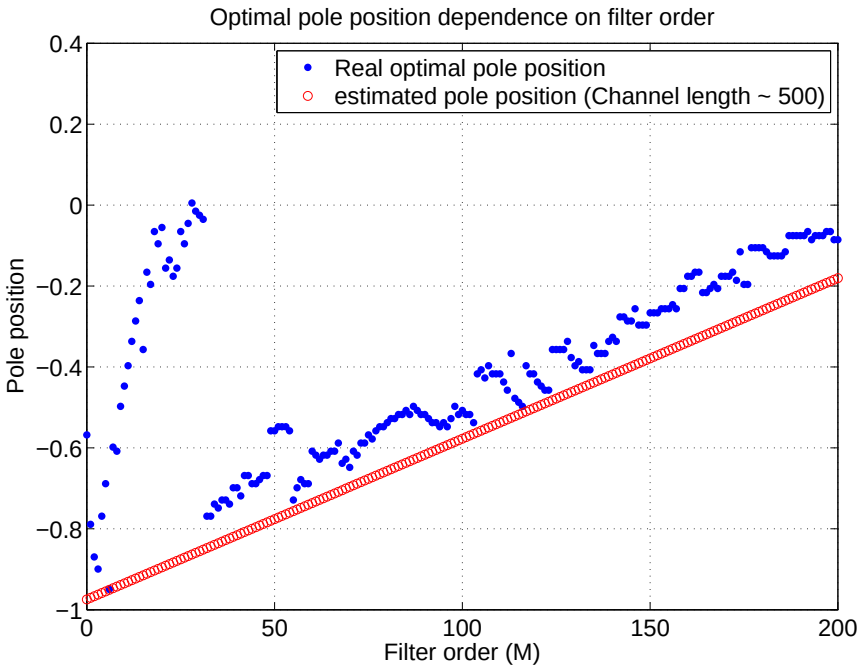



Figure 7.18.: Variation of the optimal (and estimated) pole position for the Laguerre filter as a function of the filter order.

Now that the individual case of a channel realization is known, a more general case can be obtained by analyzing many realizations of the model as a stochastic experiment. The estimation is accomplished by a Laguerre transversal filter of order up to 200 for 500 realizations of the UWB stochastic channel model.

In Figure 7.19 the mean NMSE is represented as a function of the pole position for different filter orders and as a function of the filter order for various pole positions in Figure 7.20. Both figures show that Laguerre filters can achieve a MSE improvement of about 1 dB in the channel estimation and that the estimated pole position provides a pole position close to the optimal one which results in MSE value close to the optimal. Another remarkable feature is the quick decrease of the optimal pole position towards zero due to the shorter effective length of the channels which is about 500 samples. In Figure 7.21, the optimal and estimated pole position

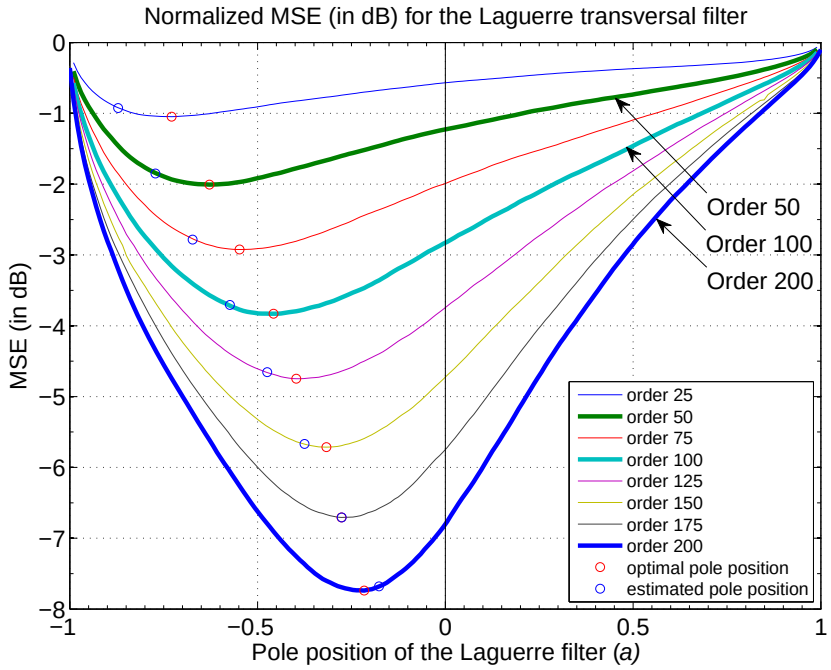



Figure 7.19.: Normalized mean MSE of UWB channels estimate as a function of the pole position for different filter orders (note: $a = 0 \triangleq$ FIR filter ).

are represented. Laguerre filters perform better than FIR filters for these channels but they give a small MSE improvement (about 1 dB) which can be explained due to the poor fit of the Laguerre sequences to this type of channels due to its smaller length and chaotic behavior (irregular and non-sparse CIR).

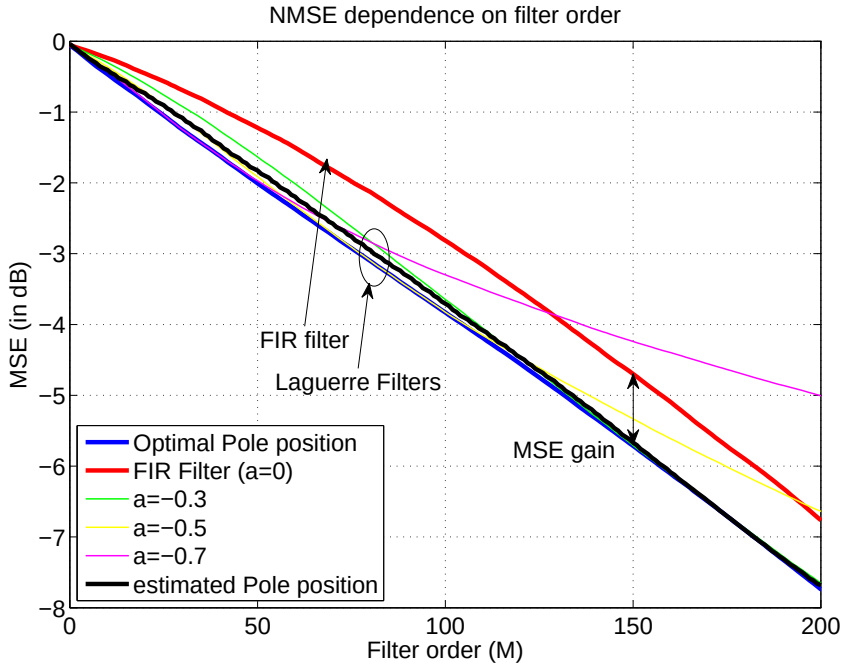



Figure 7.20.: Normalized mean MSE of UWB channels estimate as a function of the filter order for different pole positions.(note: $a = 0 \stackrel{\triangle}{=}$ FIR filter ).

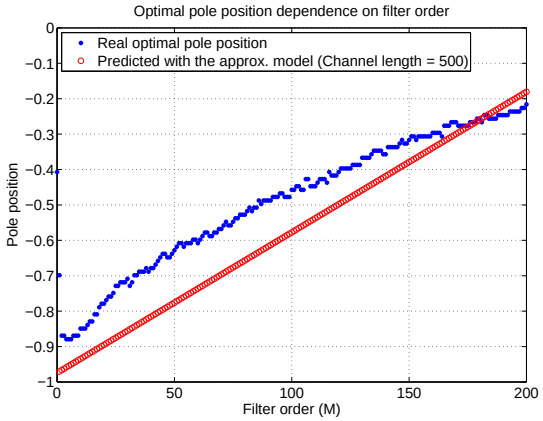



Figure 7.21.: Variation of the optimal (and estimated) pole position for the Laguerre filter as a function of the filter order.

The experimental probability density function (pdf) for the MSE is obtained for filters of orders 50, 100 and 200 and pole positions $a$ =0, -0.3, -0.5 and the optimal and estimated for each order (note that $a = 0$ corresponds to the FIR transversal filter). The reason to choose -0.3 and -0.5 as pole positions is that seeing the mean MSE behavior in Figure 7.19 a pole of -0.3 or -0.5 seems to be a wise election. This results are summarized in Table 7.8, and represented in Figures 7.22, 7.23 and 7.24 for filter of orders 50, 100 and 200 respectively. The mean MSE gain obtained with each filter in comparison to the FIR filter is shown in Table 7.9

Table 7.8.: NMSE (in dB) achieved by approximating UWB channels with filters of different orders for several pole positions.

| | | Order 50 | | Order 100 | | Order 200 | |
|---|---|---|---|---|---|---|---|
| filter | pole position ($a$) | Mean | Std | Mean | Std | Mean | Std |
| FIR | 0 | -1.3 dB | 0.8 dB | -3 dB | 1.4 dB | -7.1 dB | 2.0 dB |
| Laguerre | -0.3 | -1.7 dB | 0.9 dB | -3.8 dB | 1.2 dB | -7.9 dB | 1.5 dB |
| Laguerre | -0.5 | -2.0 dB | 0.8 dB | -3.9 dB | 1.0 dB | -6.7 dB | 1.3 dB |
| Laguerre | Estimated | -1.9 dB | 0.5 dB | -3.8 dB | 0.9 dB | -8.0 dB | 1.8 dB |
| Laguerre | Optimal | -2.1 dB | 0.6 dB | -3.9 dB | 1.0 dB | -8.0 dB | 1.7 dB |

Table 7.9.: Mean $\Delta_{\text{MSE}}$ or MSE gain achieved by approximating UWB channels with filters of different orders and pole positions.

| pole position ($a$) | Order 50 | Order 100 | Order 200 |
|---|---|---|---|
| $-0.3$ | 0.4 dB | 0.8 dB | 0.8 dB |
| $-0.5$ | 0.7 dB | 0.9 dB | -0.4 dB |
| *Estimated* | 0.6 dB | 0.8 dB | 0.9 dB |
| *Optimal* | 0.8 dB | 0.9 dB | 0.9 dB |

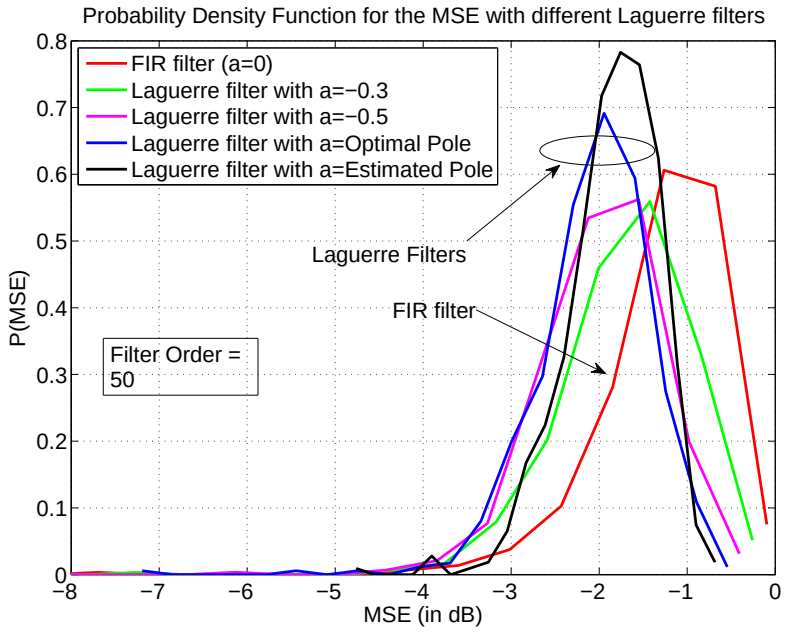



Figure 7.22.: Normalized MSE probability density function for a Laguerre filter of order 50 for different pole positions.

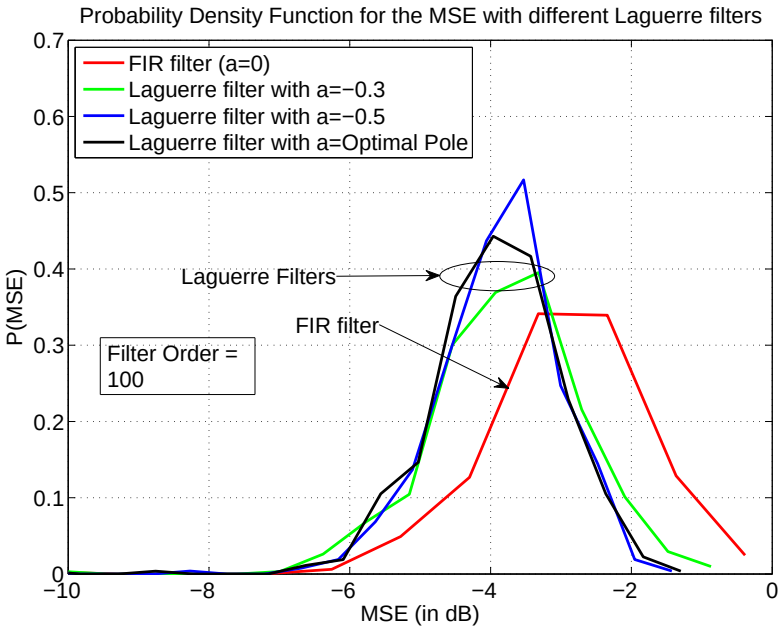



Figure 7.23.: Normalized MSE probability density function for a Laguerre filter of order 100 for different pole positions.

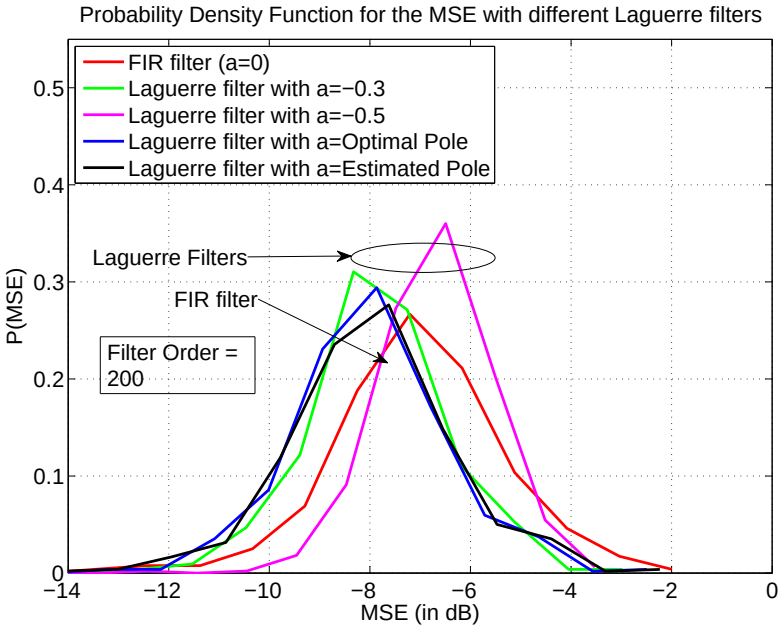
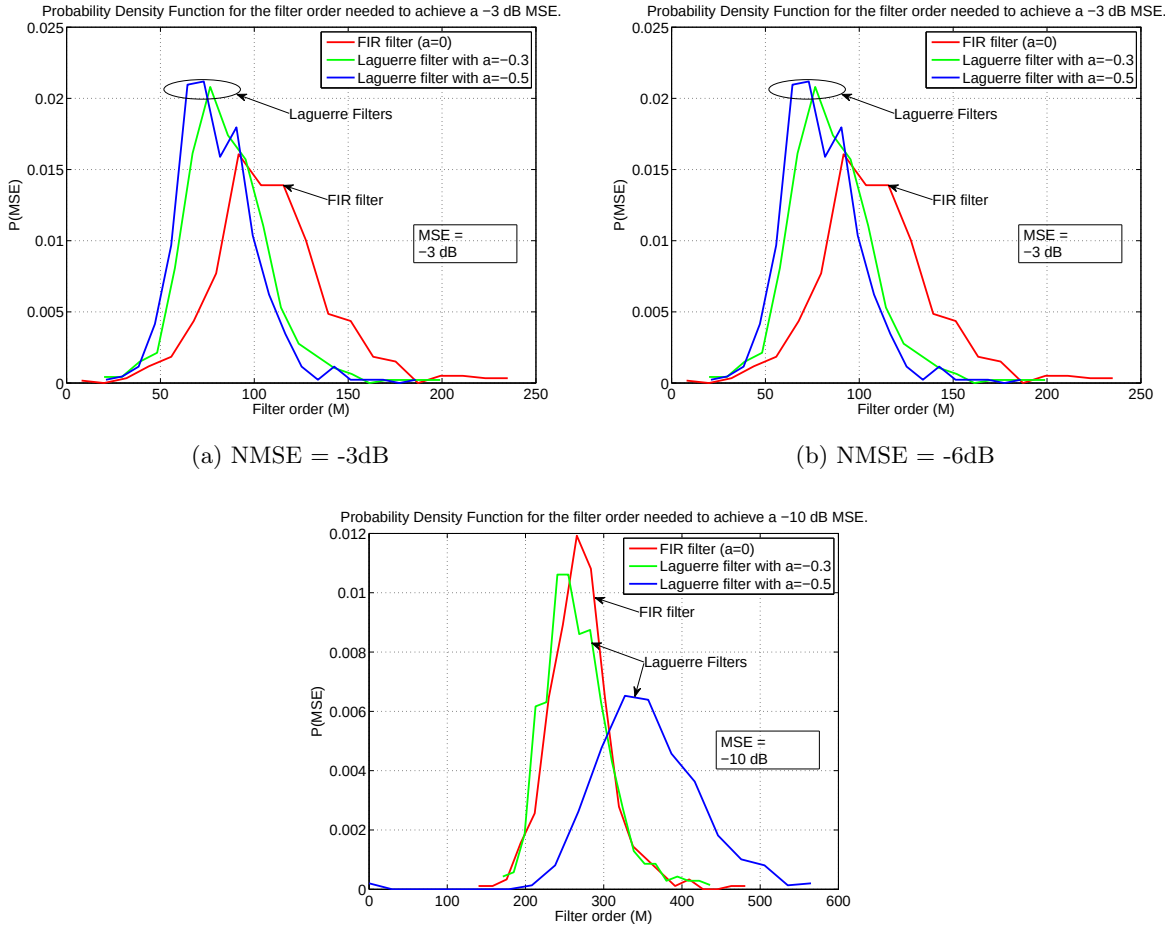



Figure 7.24.: Normalized MSE probability density function for a Laguerre filter of order 200 for different pole positions.

If the order needed to reach some specific MSE is calculated for a given pole position, the following results are obtained (see Figure 7.25 and table 7.10).

Table 7.10.: Mean filter order needed for Laguerre filters in order to achieve a specific NMSE.

| | | -3dB NMSE | | -6dB NMSE | | -10dB NMSE | |
|---|---|---|---|---|---|---|---|
| filter | pole position ($a$) | Mean | Std | Mean | Std | Mean | Std |
| FIR transversal filter | 0 | 109 | 31 | 182 | 39 | 270 | 41 |
| Laguerre filter | $-0.3$ | 86 | 23 | 159 | 31 | 266 | 42 |
| Laguerre filter | $-0.5$ | 81 | 21 | 179 | 40 | 355 | 69 |




(a) NMSE = -3dB

(b) NMSE = -6dB

(c) NMSE = -10dB

Figure 7.25.: Order needed of the different filters for achieving a specific NMSE (-3dB, -6dB and -10 dB).

What can be derived from all these results, is that for this kind of more chaotic channels, Laguerre filers perform better than FIR filters, as a generalization of the FIR filters, but the benefits obtained are small because a fine time resolution is needed to correctly gather the energy of the channel due to the chaotic behavior, this makes that with Laguerre filters a high order is needed to achieve this better time resolution, which derives from the intrinsic low-pass behavior of the Laguerre filters.

# 8. Laguerre adaptive methods

## 8.1. Introduction

In the last chapters, the estimation problem was solved by perfectly knowing the second order statistics of the signals involved in the system, which is of theoretical interest. In practice, the problem of estimating a channel is solved by adaptively adjusting the weights of the filter regarding to some iterative solution of the minimization of the mean square error criterion. The filter that we are using is the one shown in Figure 8.1, which is a generalization of the transversal filter which represents both, the FIR and Laguerre transversal filter (actually, this generalization is not really needed for this study because we have seen before that the Laguerre transversal filter is a generalization of the FIR transversal filter). Note that now the time index $k$ is changed to $i$, because now the discrete time is interpreted as an iteration.

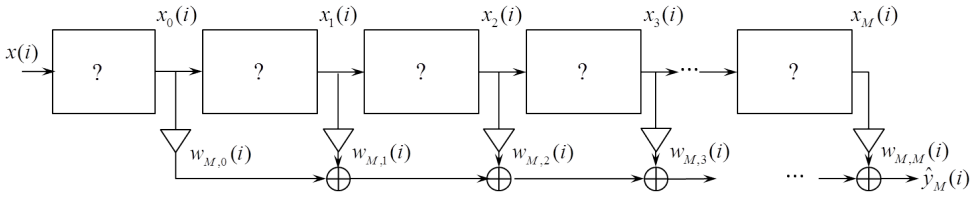



Figure 8.1.: Generalized Transversal Filter

$$\hat{y}_M(i) = \sum_{j=0}^{M} w_{M,j}(i) x_j(i) = \mathbf{w}_M(i)\mathbf{x}_M(i) \tag{8.1}$$

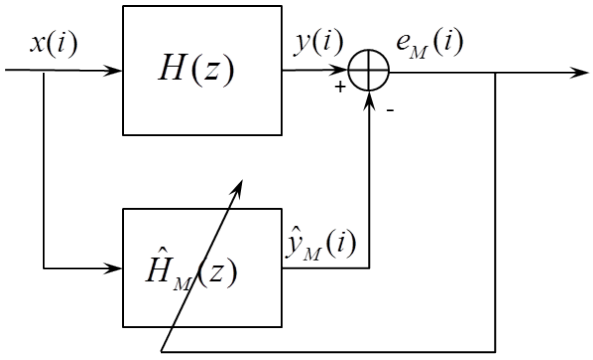



Figure 8.2.: Model used for adaptation

The adaptation model used is shown in figure 8.2. Common concepts in adaptive methods are convergence ( $\mathsf{E}[e_n(i)] \xrightarrow[i\to+\infty]{} J_{\min}$, with $J_{\min}$ the theoretical/asymptotic MSE, which is the MSE calculated in the previous chapters ) and convergence speed ( how quick it reaches this theoretical/asymptotic value)

The method chosen for this section is the Recursive Least Squares (RLS) method, which is one of the most used adaptive algorithms, exhibits fast convergence and there are many efficient

(low computational complexity) algorithms, such as the RLS lattice algorithms (see [6], [7]). The accurate description and derivation of the method can be found in any good adaptive filters book, like [11] and [3].

## 8.2. RLS Method

The RLS algorithm can be viewed as a stochastic-gradient solution for the steepest descent methods that employs a more sophisticated approximation for the autocorrelation matrix $\mathbf{R}_x$, given by a exponentially weighted average

$$\hat{\mathbf{R}}_x = \frac{1}{i+1}\sum_{j=0}^{i}\lambda^{i-j}\mathbf{x}_M(i)\mathbf{x}_M^H(i)$$

instead of an instantaneous approximation like the Least Mean Squares (LMS) algorithm.

However, the RLS algorithm is the exact solution of a well defined Least Squares (LS) problem, given by

$$\min_{\mathbf{w}}\left[\lambda^{i+1}(\mathbf{w}-\bar{\mathbf{w}})^*\mathbf{\Pi}(\mathbf{w}-\bar{\mathbf{w}}) + \sum_{j=0}^{i}\lambda^{i-j}|y(j)-\mathbf{x}_M(j)\mathbf{w}|^2\right] \tag{8.2}$$

which the RLS algorithm solves in a recursively manner. The RLS recursion is given by

$$\begin{aligned}
e_M(i) &= y(i) - \mathbf{x}_M\mathbf{w}_{i-1}, && (8.3)\\
\gamma_i &= \frac{1}{1+\frac{1}{\lambda}\mathbf{x}_M\mathbf{P}_{i-1}\mathbf{x}_M^H}, && (8.4)\\
\mathbf{g}_i &= \frac{1}{\lambda}\mathbf{P}_{i-1}\mathbf{x}_M^H\gamma_i, && (8.5)\\
\mathbf{w}_i &= \mathbf{w}_{i-1} + \mathbf{g}_i e_M(i), && (8.6)\\
\mathbf{P}_i &= \frac{1}{\lambda}\mathbf{P}_{i-1} - \frac{\mathbf{g}_i\mathbf{g}_i^H}{\gamma_i}. && (8.7)
\end{aligned}$$

To evaluate the performance of the RLS algorithm, the ensemble-average learning curve is used. The learning curve for an algorithm is given by the value of the cost function ($J$) as a function of the iteration ($i$)

$$J(i) = |e_M(i)|^2. \tag{8.8}$$

This cost function is calculated for a number of iterations, $0 \leq i \leq N$, which is large enough to observe convergence by computing the error sequence and the corresponding squared-error curve. This cost function as result of the first experiment is denoted by

$$\left\{J^{(1)}(i)\right\} \tag{8.9}$$

with the superscript $^{(1)}$ used to indicate the first experiment. The experiment is repeated several times with the same initial conditions obtaining $L$ cost functions. The ensemble-average learning curve over the interval $0 \leq i \leq N$ is defined as the sample average over the $L$ experiments:

$$\hat{J}(i) \triangleq \frac{1}{L}\sum_{j=1}^{L}J^{(j)}(i),\ 0 \leq i \leq N. \tag{8.10}$$

For the following experiments, $L = 100$.

### 8.2.1. Estimation of UWB channels

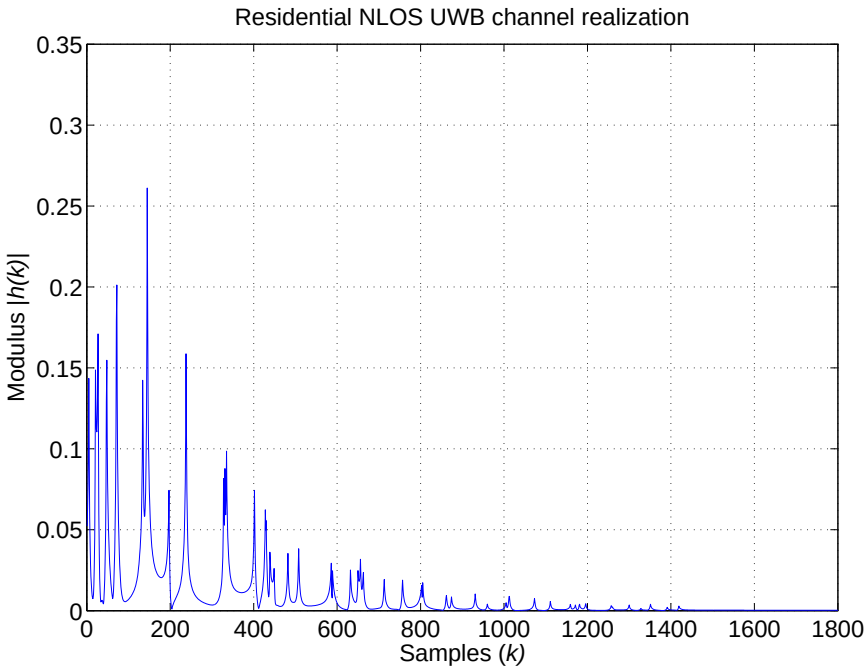
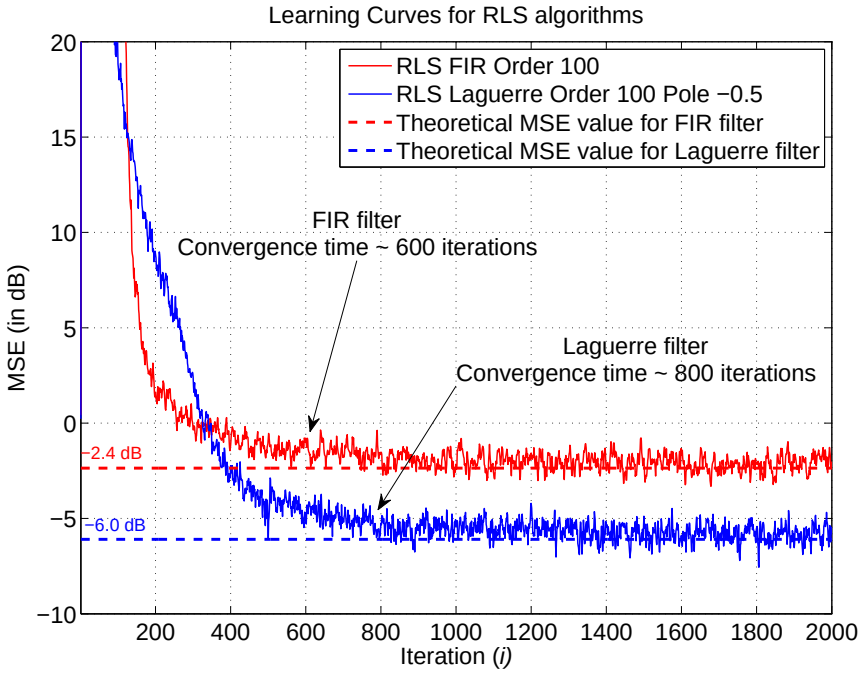



Figure 8.3.: Channel realizations for the Residential NLOS environment of the UWB channel model

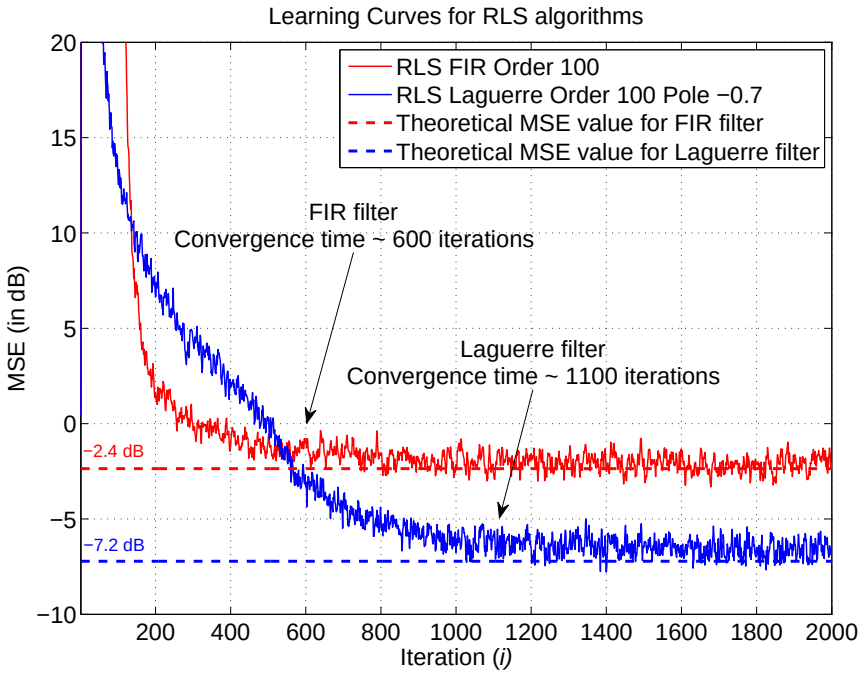
Now that all the theoretical results are known (see Chapter 7), is time to prove the theoretical results by simulating the RLS adaptive problem without previous knowledge of the channel impulse response (or its statistics), and watching other topics not discussed yet like convergence and training sequence length.

The first channel to prove is the realization of the UWB stochastic model for the Residential NLOS environment used before, shown in Figure 8.3.

As seen on Chapter 7 in Figure 7.3, for a filter order of 100, a pole position of -0.7 or -0.5 are a good choice. The learning curves for a=-0.5 and a=-0.7 are shown in Figure 8.4, which reveals a dependence of the convergence time in the pole position. This can be explained taking a look at Figure 8.5, which shows that for a large magnitude pole position a longer part of the impulse response can be gathered by the estimate which makes necessary a longer convergence time.



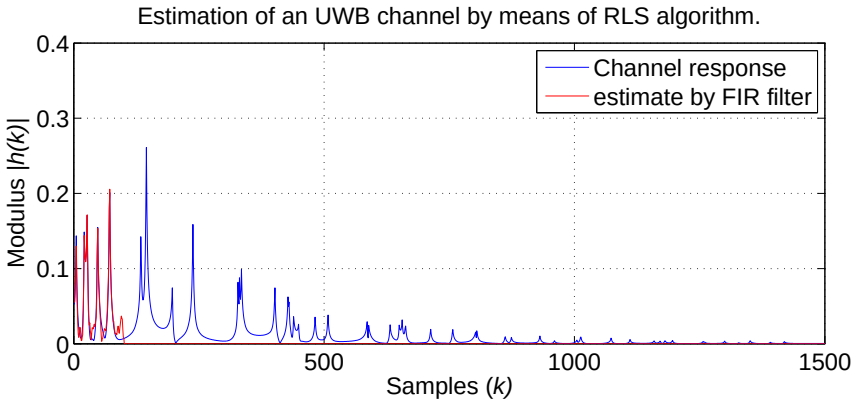

(a) $a = -0.5$



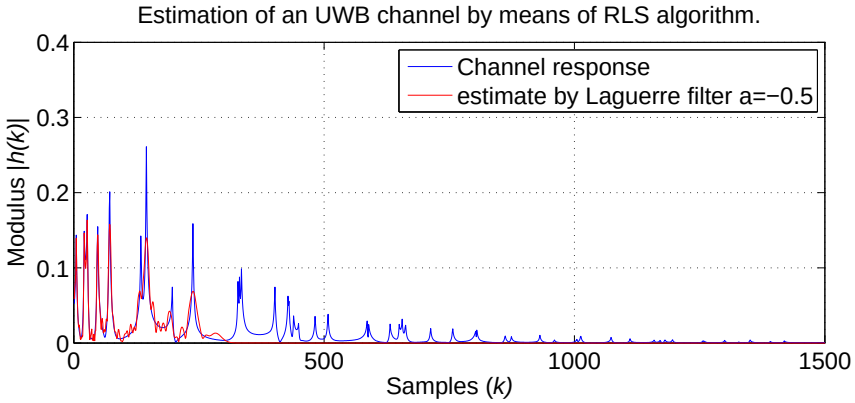

(b) $a = -0.7$

Figure 8.4.: Learning curves of the RLS algorithm for Laguerre and FIR filters.

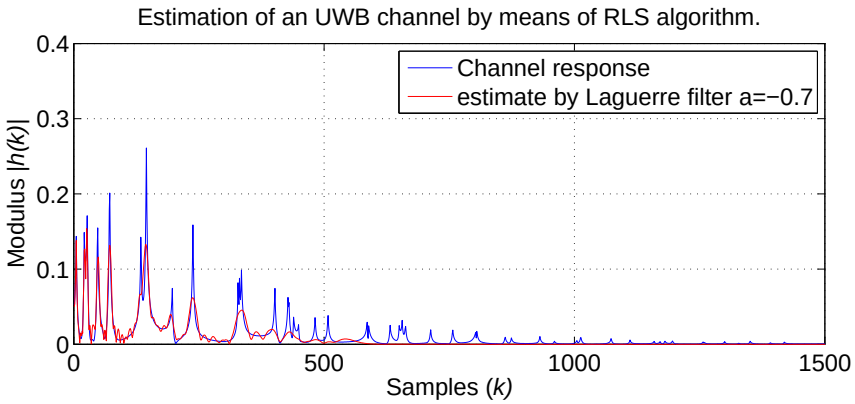
The resulting estimated system response in the steady state of the filters (when the convergence/asymptotic behavior is achieved) is shown in Figure 8.5, which shows the main feature of the channel estimation with Laguerre filters in comparison with FIR filter, by estimating a longer part of the channel impulse response which the FIR filter can not gather.




(a) $a = 0.0$ (FIR filter)




(b) $a = -0.5$




(c) $a = -0.7$

Figure 8.5.: Approximated impulse response obtained by Laguerre ($a = -0.5$ and -0.7) and FIR filter of order 100 with a RLS algorithm.

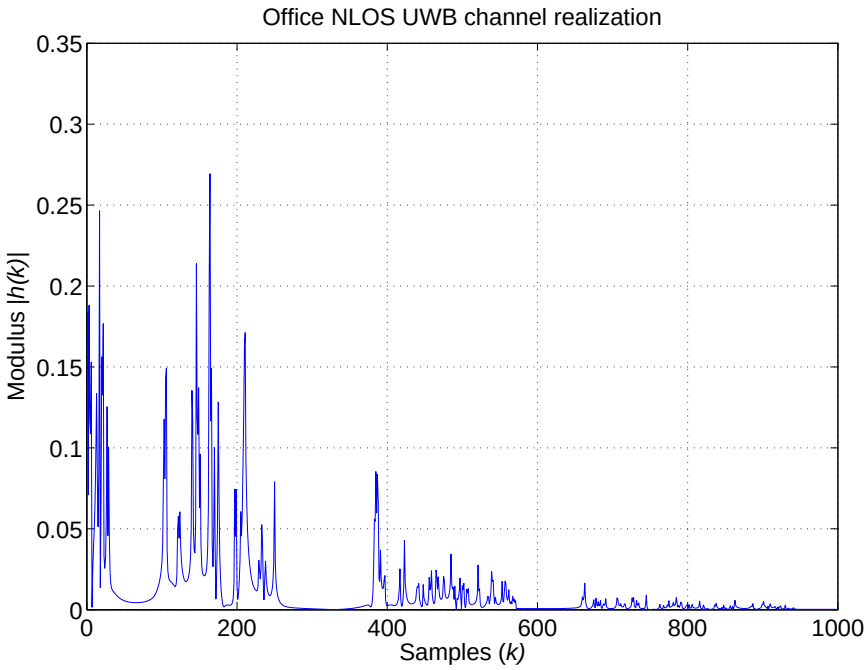
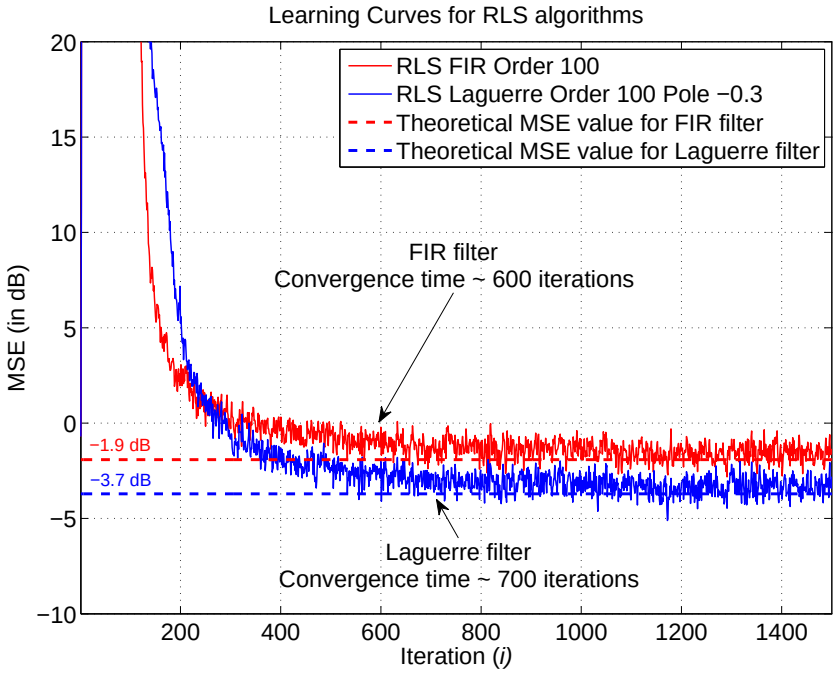


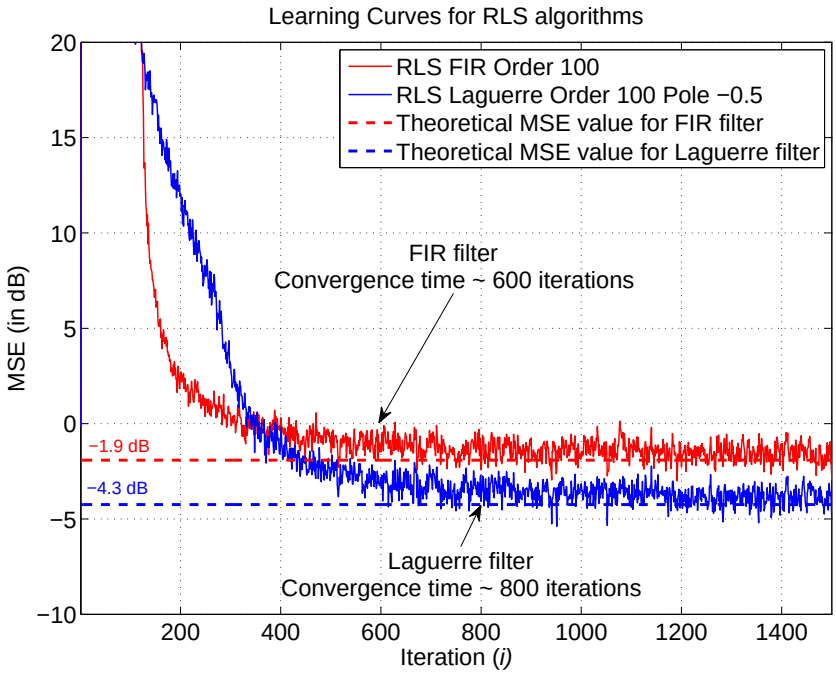

Figure 8.6.: Channel realizations for the Office NLOS environment of the UWB channel model

The second channel is the realization of the UWB stochastic model for the Office NLOS environment used before, shown in Figure 8.6.

Taking a look back to Chapter 7, in Figure 7.16, for a filter order of 100, a pole position of -0.3 or -0.5 are a good choice. The learning curves for a=-0.3 and a=-0.5 are shown in Figure 8.7



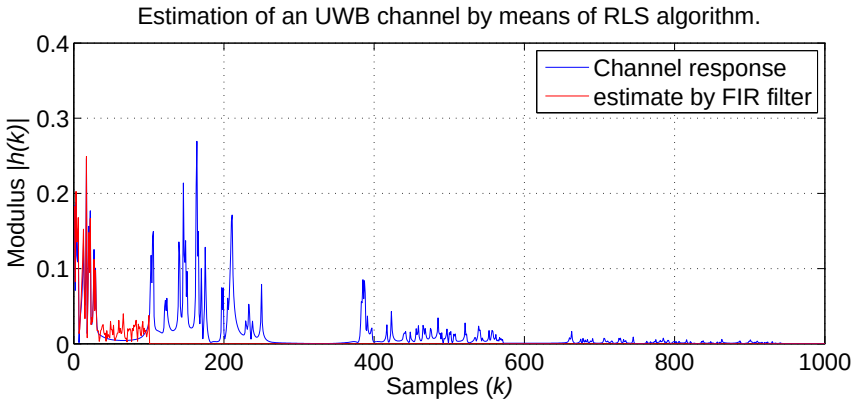

(a) $a = -0.5$



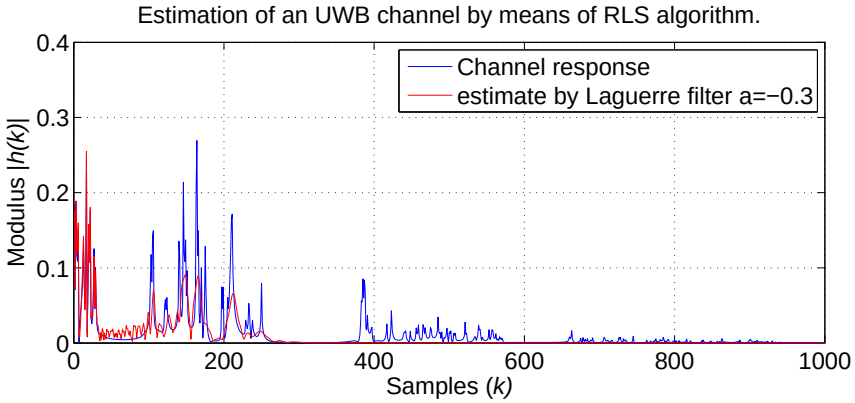

(b) $a = -0.7$

Figure 8.7.: Learning curves of the RLS algorithm for Laguerre and FIR filters.

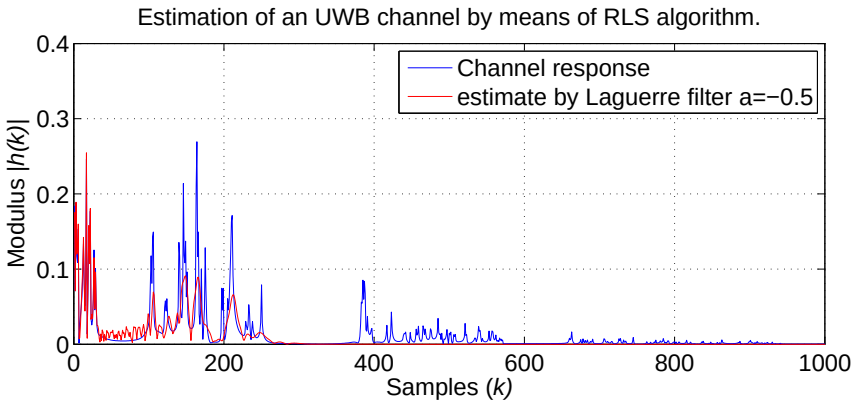
The estimated response of the steady state are shown in Figure 8.8




(a) $a = 0.0$ (FIR filter)




(b) $a = -0.3$




(c) $a = -0.5$

Figure 8.8.: Approximated impulse response obtained by Laguerre ($a = -0.3$ and -0.5) and FIR filter of order 100 with a RLS algorithm.

# 9. Conclusions and future work

Laguerre filters as a generalization of the FIR filters perform better or equal than FIR filters, in particular for UWB channels, we have seen that we can achieve a mean improvement of about 5 dB in the MSE for the residential NLOS environment of the channel model and a moderate 1 dB improvement for the Office NLOS environment. The conclusion is that Laguerre sequences are not specially suited for this kind of channels, specially if they present chaotic or non sparse behavior (like the Office NLOS environment channels generated by the channel model), they fit best smooth channels like channels with a rational transfer function, however they still represent a good choice when the channel length is large and the filter order must be small. This unsuitability for this kind of channels can be explained due to the intrinsic low-pass behavior of the Laguerre filters remembering that they consist of a low-pass section followed by a cascade of all-pass sections.

Another interesting thing is the selection of the pole position of the Laguerre filter. We have seen that the optimal pole position varies for different filter orders, starting from a large magnitude pole position for small orders in order to have longer sequences, and smaller ones for larger filter orders what makes the sequences to have smaller temporal extension but finer temporal resolution. The optimal pole position can be estimated with relative success by applying the easy idea that the Laguerre sequences have to cover a temporal extension approximately equal to the typical channel length. This idea in addition to the linear modeling of the Laguerre sequences length has proved to be a quite good estimate of the optimal pole position for a given typical channel length and a fixed filter order, what makes the selection of the Laguerre pole an easy task for UWB channels.

For future work in this field, there are two generalizations that can be applied to the Laguerre filter. The first one is to have a complex pole position (always inside the unit circle) instead of restricting to a real pole (varying from $-1$ to 1, as in this study) with the structure shown in Figure 9.1. Adding a complex phase to the pole does not affect the temporal extension of the sequences.

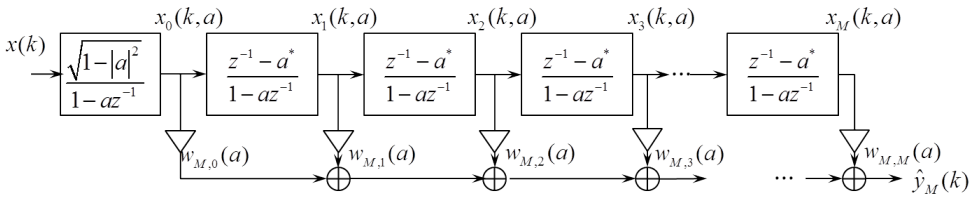



Figure 9.1.: Laguerre transversal filter with complex pole position.

Another generalization is using different poles for the sections of filters instead of only one pole. This new problem deals with the orthonormal structure shown in Figure 9.2, which results in a multivariate optimization problem for choosing the optimal pole positions with its consequent difficulty.

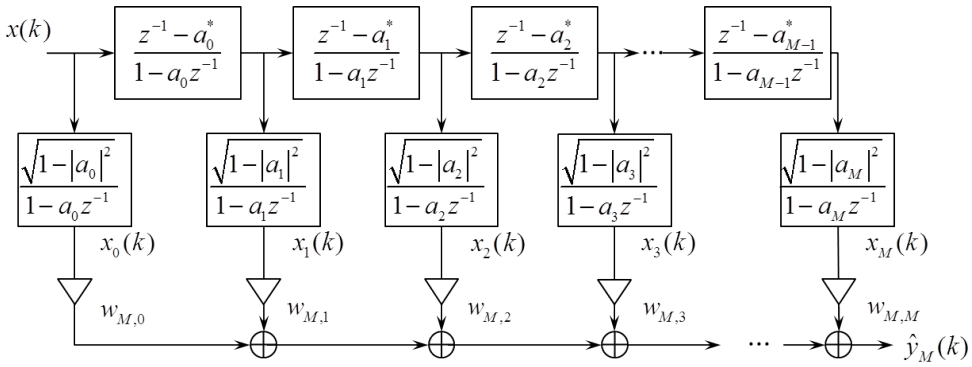



Figure 9.2.: A transversal orthonormal structure with multiple pole positions.

# Bibliography


[1] Back, A.D. ; Tsoi, A.C.: Nonlinear system identification using discrete Laguerre functions. In: *Journal of Systems Engineering* 6 (1996), Nr. 3, S. 194–207

[2] (FCC), Federal Communications C.: Revision of Part 15 of the Commission's rules regarding ultra-wideband transmission systems. In: *ET Docket 98-153* FCC 02-48 (2002), feb.

[3] Haykin, Simon: *Adaptive filter theory (3rd ed.)*. Upper Saddle River, NJ, USA : Prentice-Hall, Inc., 1996. – ISBN 0–13–322760–X

[4] Hirt, W.: The European UWB Radio Regulatory and Standards Framework: Overview and Implications. In: *Ultra-Wideband, 2007. ICUWB 2007. IEEE International Conference on*, 2007, S. 733 –738

[5] Lagunas-Hernandez, M. ; Figueiras-Vidal, A. ; Marino-Acebal, J. ; Vilanova, A.: A linear transform for spectral estimation. In: *Acoustics, Speech and Signal Processing, IEEE Transactions on* 29 (1981), oct, Nr. 5, S. 989 – 994. – ISSN 0096–3518

[6] Merched, R. ; Sayed, A.H.: Order-recursive RLS Laguerre adaptive filtering. In: *Signal Processing, IEEE Transactions on* 48 (2000), nov, Nr. 11, S. 3000 –3010. – ISSN 1053–587X

[7] Merched, R. ; Sayed, A.H.: RLS-Laguerre lattice adaptive filtering: error-feedback, normalized, and array-based algorithms. In: *Signal Processing, IEEE Transactions on* 49 (2001), nov, Nr. 11, S. 2565 –2576. – ISSN 1053–587X

[8] Molisch, Andreas F. ; Balakrishnan, Kannan ; Chong, Chia chin ; Emami, Shahriar ; Fort, Andrew ; Karedal, Johan ; Kunisch, Juergen ; Schantz, Hans ; Schuster, Ulrich ; Siwiak, Kai: IEEE 802.15.4a channel model - final report. In: *Converging: Technology, work and learning. Australian Government Printing Service*, Online]. Available, 2004

[9] Proakis, J.G. ; Manolakis, D.G.: *Digital signal processing.* Pearson Prentice Hall, 2007. – ISBN 9780131873742

[10] Saleh, A ; Valenzuela, R: A Statistical Model for Indoor Multipath Propagation. In: *IEEE Journal on Selected Areas in Communications* 5 (1987), Nr. 2, S. 128–137

[11] Sayed, Ali H.: *Adaptive Filters.* Wiley-IEEE Press, 2008. – ISBN 0470253886

[12] Silva, Tomás Oliveira e: Laguerre Filters — An Introduction. In: *Revista do DETUA* 1 (1995), Januar, Nr. 3, S. 237–248

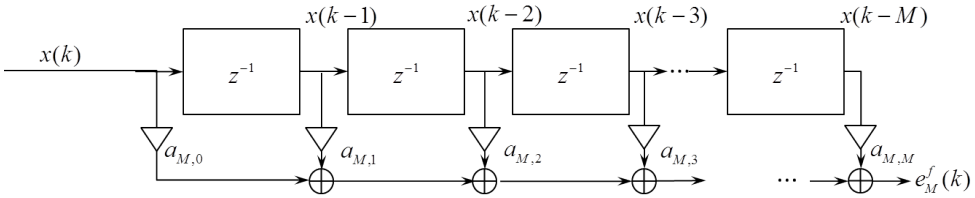
[13] W., Broo P.: Discrete Orthonormal Sequences. In: *J. ACM* 12 (1965), April, S. 151–168. – ISSN 0004–5411

# Appendix

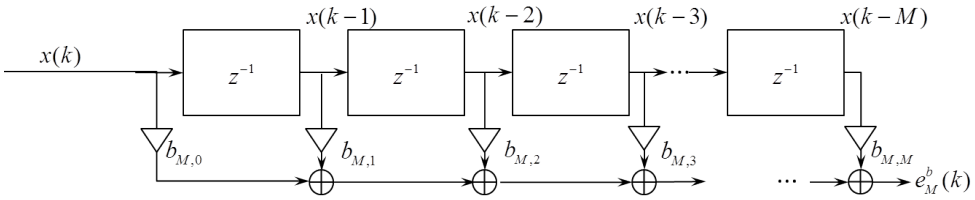
# A. Modified Levinson Durbin algorithm

When the autocorrelation matrix $\mathbf{R}_x$ is an Hermitian Toeplitz matrix ( a Toeplitz matrix or diagonal-constant matrix, named after Otto Toeplitz, is a matrix in which each descending diagonal from left to right is constant), which will be the case of this study as it will be seen on next chapter, the normal equations can be recursively solved in an efficient manner, exploiting the data structure. This efficient way to solve the normal equations is the Levinson-Durbin algorithm (based on the book [3] ).

If we consider the forward and backward error prediction filters, shown in Figure A.1. Here




(a) Forward error prediction filter




(b) Backward error prediction filter

Figure A.1.: Forward and Backward error prediction filters

the forward and backward error predictor weigth vectors are defined as (the pole position $a$ must no be confused with the forward error prediction filter weights $a_{M,i}$)

$$\mathbf{a}_M = \begin{bmatrix} a_{M,0} \\ a_{M,1} \\ \vdots \\ a_{M,M} \end{bmatrix}, \tag{A.1}$$

$$\mathbf{b}_M = \begin{bmatrix} b_{M,0} \\ b_{M,1} \\ \vdots \\ b_{M,M} \end{bmatrix}, \tag{A.2}$$

note that $a_{M,0} = 1.0$ and $b_{M,M} = 1.0$. The forward and backward prediction errors are $e^f_M(k)$

and $e_M^b(k)$ respectively, defined by

$$e_M^f(k) = \sum_{j=0}^{M} a_{M,j} x(k-j), \tag{A.3}$$

$$e_M^b(k) = \sum_{j=0}^{M} b_{M,j} x(k-j). \tag{A.4}$$

Assuming that we have stationary signals, the forward and backward predictions errors will have the same variance, say $\sigma_M^2$. The normal equations for these filters, also known as the augmented Wiener-Hopf equations are

$$\begin{aligned} \mathbf{R}_{M+1}\mathbf{a}_M &= \begin{bmatrix} \sigma_M^2 \\ \mathbf{0}_M \end{bmatrix} \\ \mathbf{R}_{M+1}\mathbf{b}_M &= \begin{bmatrix} \mathbf{0}_M \\ \sigma_M^2 \end{bmatrix} \end{aligned} \tag{A.5}$$

where $\mathbf{0}_M$ is the $M$-by-1 null vector. Only for this section the autocorrelation matrix $\mathbf{R}_x$ will be denoted as $\mathbf{R}_{M+1}$, with $M+1$ the dimension of the matrix in order to notice that it is a $(M+1)$-by-$(M+1)$ matrix. The subscripts refer to their dimensions in $\mathbf{R}_{M+1}$ and $\mathbf{0}_M$, and to the prediction order in $\mathbf{a}_M$ and $\sigma_M^2$,

$$\begin{aligned} \mathbf{R}_{M+1} &= \mathsf{E}\left[\mathbf{x}_M(k)\mathbf{x}_M^H(k)\right] \\ &= \begin{bmatrix} r(0) & r(1) & \cdots & r(M) \\ r^*(1) & r(0) & \cdots & r(M-1) \\ \vdots & \vdots & \ddots & \vdots \\ r^*(M) & r^*(M-1) & \cdots & r(0) \end{bmatrix} \end{aligned} \tag{A.6}$$

Now it can be seen that the corresponding backward prediction-error filter $\mathbf{b}_M$ is obtained by (in view of equations A.5 ) backward rearrangement of the elements of $\mathbf{a}_M$ and their complex conjugation. The combined effect of these two operations is denoted by $\mathbf{a}_m^{B*}$, defining the operation $\cdot^B$ as the backward rearrangement of the elements of the vector, say

$$\mathbf{b}_M = \begin{bmatrix} a_{M,M}^* \\ a_{M,M-1}^* \\ \vdots \\ a_{M,0}^* \end{bmatrix} = \mathbf{a}_M^{B*} \tag{A.7}$$

where the $l$-th tap weight of a forward prediction-error filter of order $M$ is denoted by $a_{M,l}$ , and likewise for $a_{M,l}^*$.

The Levinson Durbin algorithm solves the augmented Wiener Hopf equations, given in A.5 for prediction-error filters of order $M$ using the solution of the same problem for filters of order of $M-1$ (i.e., one order higher, in a recursive manner). The main goal of the Levinson Durbin algorithm is the computational efficiency, in terms of numbers of operations and storage locations in comparison to the standard methods, such as the Gaussian elimination method (matrix inversion).

The main idea behind the Levinson Durbin algorithm is to exploit the Toeplitz structure of $R_{M+1}$ by relating the order-updated solution to the already available one. An intelligent a priori

solution for the forward prediction filter is

$$\mathbf{a}_M = \begin{bmatrix} \mathbf{a}_{M-1} \\ 0 \end{bmatrix} + \kappa_M \begin{bmatrix} 0 \\ \mathbf{b}_{M-1} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_{M-1} \\ 0 \end{bmatrix} + \kappa_M \begin{bmatrix} 0 \\ \mathbf{a}^{B*}_{M-1} \end{bmatrix} \tag{A.8}$$

where $\kappa_M$ is a constant. The scalar version of this order update is

$$a_{M,l} = a_{M-1,l} + \kappa_M a^*_{M-1,l-1}, \qquad l = 0, 1, \ldots, M,$$

Note that $a_{M-1,0} = 1$ and $a_{M-1,M} = 0$.

If we substitute Equation A.8 in A.5, we have

$$\begin{bmatrix} \sigma^2_M \\ \mathbf{0}_M \end{bmatrix} = \mathbf{R}_{M+1} \begin{bmatrix} \mathbf{a}_{M-1} \\ 0 \end{bmatrix} + \kappa_M \mathbf{R}_{M+1} \begin{bmatrix} 0 \\ \mathbf{a}^{B*}_{M-1} \end{bmatrix} \tag{A.9}$$

now we can partition the autocorrelation matrix $\mathbf{R}_{M+1}$, in view of Equation A.6, by defining the vector $\mathbf{r}$ as

$$\mathbf{r}_M = \begin{bmatrix} r(1) \\ r(2) \\ \vdots \\ r(M) \end{bmatrix} = \begin{bmatrix} r^*(-1) \\ r^*(-2) \\ \vdots \\ r^*(-M) \end{bmatrix} \tag{A.10}$$

and the matrx $\mathbf{R}_{M+1}$ can be partitioned as

$$\mathbf{R}_{M+1} = \begin{bmatrix} \mathbf{R}_M & \mathbf{r}^B_M \\ \mathbf{r}^{BH}_M & r(0) \end{bmatrix} = \begin{bmatrix} r(0) & \mathbf{r}^T_M \\ \mathbf{r}_M & \mathbf{R}_M \end{bmatrix}, \tag{A.11}$$

now substituting in A.9, the equation looks like

$$\begin{aligned} \begin{bmatrix} \sigma^2_M \\ \mathbf{0}_M \end{bmatrix} &= \begin{bmatrix} \mathbf{R}_M & \mathbf{r}^B_M \\ \mathbf{r}^{BH}_M & r(0) \end{bmatrix} \begin{bmatrix} \mathbf{a}_{M-1} \\ 0 \end{bmatrix} + \kappa_M \begin{bmatrix} r(0) & \mathbf{r}^T_M \\ \mathbf{r}_M & \mathbf{R}_M \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{a}^{B*}_{M-1} \end{bmatrix} \Longrightarrow \\ \begin{bmatrix} \sigma^2_M \\ \mathbf{0}_M \end{bmatrix} &= \begin{bmatrix} \mathbf{R}_M \mathbf{a}_{M-1} \\ \mathbf{r}^{BH}_M \mathbf{a}_{M-1} \end{bmatrix} + \kappa_M \begin{bmatrix} \mathbf{r}^T_M \mathbf{a}^{B*}_{M-1} \\ \mathbf{R}_M \mathbf{a}^{B*}_{M-1} \end{bmatrix} \end{aligned} \tag{A.12}$$

The set of normal equations for the forward prediction-error filter of order $M-1$, regarding to Eq. A.5, is

$$\mathbf{R}_M \mathbf{a}_{M-1} = \begin{bmatrix} \sigma^2_{M-1} \\ \mathbf{0}_{M-1} \end{bmatrix} \tag{A.13}$$

$$\mathbf{R}_M \mathbf{b}_{M-1} = \mathbf{R}_M \mathbf{a}^{B*}_{M-1} = \begin{bmatrix} \mathbf{0}_{M-1} \\ \sigma^2_{M-1} \end{bmatrix}, \tag{A.14}$$

then the following scalar value is defined

$$\Delta_{M-1} = \mathbf{r}^{BH}_M \mathbf{a}_{M-1} = \sum_{l=0}^{M-1} r^*(M-l) a_{M-1,l} = \sum_{l=0}^{M-1} r(l-M) a_{M-1,l} \tag{A.15}$$

and noting that

$$\Delta_{M-1} = \mathbf{r}^{BH}_M \mathbf{a}_{M-1} = \left(\mathbf{r}^T_M \mathbf{a}^{B*}_{M-1}\right)^*$$

So the equation looks like

$$\begin{bmatrix}\sigma_M^2\\ \mathbf{0}_M\end{bmatrix} = \begin{bmatrix}\sigma_{M-1}^2\\ \mathbf{0}_{M-1}\\ \Delta_{M-1}\end{bmatrix} + \kappa_M \begin{bmatrix}\Delta_{M-1}^*\\ \mathbf{0}_{M-1}\\ \sigma_{M-1}^2\end{bmatrix} \tag{A.16}$$

Here two important relations can be obtained from the first and last entries of Equation A.16

$$\sigma_M^2 = \sigma_{M-1}^2 + \kappa_M \Delta_{M-1}^*, \tag{A.17}$$

and

$$\begin{aligned} 0 = \Delta_{M-1} + \kappa_M \sigma_{M-1}^2 \implies \\ \kappa_M = -\frac{\Delta_{M-1}}{\sigma_{M-1}^2} \end{aligned} \tag{A.18}$$

and eliminating $\Delta_{M-1}$ from equations

$$\sigma_M^2 = \sigma_{M-1}^2(1 - |\kappa_M|^2). \tag{A.19}$$

Sumarizing the results for the Levinson Durbin algorithm

$$\begin{aligned} \Delta_{M-1} &= \mathbf{r}_M^{BT}\mathbf{a}_{M-1} = \sum_{l=0}^{M-1} r(l-M)a_{M-1,l}, \\ \kappa_M &= -\frac{\Delta_{M-1}}{\sigma_{M-1}^2}, \\ \sigma_M^2 &= \sigma_{M-1}^2(1-|\kappa_M|^2), \\ a_{M,l} &= a_{M-1,l} + \kappa_M a_{M-1,l-1}^*, \qquad l = 0, 1, \ldots, M. \end{aligned}$$

Now the problem of forward/backward prediction-error is solved, however, this was not our original goal. Initially we wanted to solve the normal equations, that is

$$\mathbf{w}_M \mathbf{R}_{M+1} = \mathbf{p}_{yx} \tag{A.20}$$

and to obtain the estimate

$$\hat{y}_M = \sum_{i=0}^{M} w_{M,i}\mathbf{x}_i(k). \tag{A.21}$$

However, if one takes into account the generated backward prediction error signals, it can be demonstrated through the normal equations for the backward prediction filter, that these signals are orthogonal to each other, thus enabling to express the estimation as a weighted sum of the backward prediction error signals.

$$\hat{y}_M = \sum_{i=0}^{M} d_i e_i^b(k) = \sum_{i=0}^{M} w_{M,i}x_i(k) \tag{A.22}$$

Note that the set of coefficients $d_i$ does not depend on the filter order ($M$) because the signals $e_i^b(k)$ are orthogonal. They are obtained by taking the scalar product between $\hat{y}_M$ and $e_i^b(k)$,

and applying the orthogonality property, say

$$d_i = \langle y, e_i^b(k) \rangle \tag{A.23}$$

In order to make these signals $e_i^b(k)$ orthonormal, they only have to be normalized through their standard deviation, i.e., a new set of coefficients, $c_i$, is defined for the orthonormal signals $e_i^{bo}(k) = e_i^b(k)/\sigma_i$. The coefficients $c_i$ can be defined as

$$\hat{y}_M = \sum_{i=0}^{M} c_i e_i^{bo}(k) \tag{A.24}$$

$$c_i = \langle y, e_i^{bo}(k) \rangle = \frac{\langle y, e_i^b(k) \rangle}{\sigma_i} \tag{A.25}$$

from the definition of the error prediction filters, it can be written that

$$e_i^b(k) = \sum_{j=0}^{i} b_{i,j} x_j(k) = \sum_{j=0}^{i} a_{i,i-j}^* x_j(k) \tag{A.26}$$

so the resulting coefficients can be written as

$$c_i = \frac{\sum_{j=0}^{i} a_{i,i-j}^* \langle y, x_j(k) \rangle}{\sigma_i} = \frac{\sum_{j=0}^{i} a_{i,i-j}^* p(j)}{\sigma_i} \tag{A.27}$$

with $p(j)$ denoting the $j$-th element of the cross correlation vector $\mathbf{p}_{yx}$.

Now, the coefficients from an orthonormal expansion of $\hat{y}_M(k)$ which is the m.m.s.e. estimate of $y(k)$ are $c_i$, and the resulting MSE is given by

$$MSE = J_{\min} = \mathsf{E}|e_M|^2 = \langle e_M, e_M \rangle = \langle (y - \hat{y}_M), e_M \rangle$$

remembering that the error signal (because of the normal equations) is orthogonal to all linear combinations of the data (i.e., $\hat{y}_M$)

$$MSE = J_{\min} = \langle y, e_M \rangle = \sigma_y^2 - \langle y, \hat{y}_M \rangle = \sigma_y^2 - \sum_{i=0}^{M} c_i \langle y, e_i^{bo} \rangle = \sigma_y^2 - \sum_{i=0}^{M} |c_i|^2$$

Normalizing this result with respect to the power of $y(k)$

$$NMSE = 1 - \sum_{i=0}^{M} \frac{|c_i|^2}{\sigma_y^2}$$