A Novel Iterative OFDMA Channel Estimation Technique for DOCSIS 3.1 Uplink Channels

Tung T. Nguyen, Student Member, IEEE, Brian Berscheid, Member, IEEE, Ha H. Nguyen, Senior Member, IEEE, and J. Eric Salt, Member, IEEE

Abstract—This paper presents an orthogonal frequency division multiple access (OFDMA) channel estimation technique that jointly considers the effects of coarse timing error and multipath propagation. Many conventional approaches only consider an optimistic scenario where timing synchronization is perfect and each of the channel delays is an integer number of system samples. In realistic scenarios timing offsets and echo delays are not integer multiples of the system’s sampling period. This leads to poor estimation and consequently reduces the system’s overall performance. This paper proposes a novel iterative channel estimation technique, which considers the practical scenario of fractional timing error and nonsample-spaced echo delays. The proposed method does not require channel state information (e.g., second-order statistics of the channel impulse responses or the noise power). Moreover, timing error can be conveniently obtained with the proposed technique. Simulation results show that, when comparing OFDMA channel estimation techniques under realistic data over cable service interface specification 3.1 channel conditions, the proposed algorithm significantly outperforms all conventional methods known to the authors.

Index Terms—OFDMA, DOCSIS, channel estimations, inter-symbol interference, timing error estimation.

I. INTRODUCTION

THE DEMAND for data services has steadily increased putting continuous pressure on data service providers to increase the data throughput of their networks. Cable television (CATV) networks are governed by a set of DOCSIS standards that place hard limits on bandwidth and data rates. The latest version of the DOCSIS standard, DOCSIS 3.1 was released in October 2013. DOCSIS 3.1 increases the bandwidth and data throughput available in CATV networks by up to 10 Gbps downstream and 1 Gbps upstream.

Fig. 1 shows a simplified structure of a cable network. There are two main components in the network: a cable modem (CM), which is located at the customer premises, and a cable modem termination system (CMTS), located at the cable company’s headend. The CMTS is responsible for managing a large number of cable modems residing in subscribers’ homes. The CMTS allocates channel resources for each CM and schedules times for sending and receiving packets. The CMTS has both Ethernet and Radio Frequency (RF) interfaces, such that the Internet traffic can be routed from the Ethernet interface, through the CMTS and then onto the RF ports that are connected to coaxial cable networks, eventually reaching the cable modem in the subscriber’s home.

DOCSIS 3.1 is markedly different from prior versions of the standard in that Orthogonal Frequency Division Multiplexing (OFDM) is used in the downstream and OFDMA is used in the upstream direction. By effectively modulating signals on narrow-band carriers, OFDM/OFDMA can mitigate intersymbol interference while simplifying the structure of the channel equalizer.

Cable plants generate a number of channel impairments, some of which differ from typical OFDM/OFDMA systems discussed in the literature such as 3G, LTE, or WiFi. The greatest difference is in the transmission medium, which is coaxial cable for CATV systems. Unlike wireless channels, coaxial cable channels vary slowly, staying essentially constant for long period of time. The signal from a CM reaches a CMTS via a main path and possibly several echo paths, resulting in some frequency selectivity. Aside from using DOCSIS 3.1-compliant modulators/demodulators, cable operators must upgrade their cable infrastructure to satisfy the standard imposed by DOCSIS 3.1. For example, while DOCSIS 3.0 allows the power in a single echo to be 10 dB lower than the main path’s power, DOCSIS 3.1 restricts...
the power in an echo to be at least 16 dB lower than the main path’s power. The latest DOCSIS standard opens new opportunities as well as presenting issues and challenges for researchers to make contributions into the newly specified cable systems.

While OFDMA systems have been studied for many years, the DOCSIS 3.1 upstream standard is unique in that it tries to maximize the overall spectral efficiency of the newly upgraded cable system (i.e., to operate as close to the Shannon limit as conditions allow). In particular, the OFDMA upstream multicarrier system contains up to 3800 data sub-carriers occupying a total bandwidth of 96 MHz. Moreover, the DOCSIS 3.1 upstream specifications introduce very high order modulation schemes, with mandatory support up to 1024-QAM and optional support for constellations up to 4096-QAM, which allows the network operators to maximize their network capacity and take advantage of signal quality improvements.

In order to effectively demodulate a spectrally efficient signal, it is necessary to employ coherent demodulation, which involves estimation and tracking of the multipath channel. With a dense constellation such as 4096-QAM, the signal can be easily distorted by a small impairment, therefore demodulation must be aided with a highly accurate channel estimator. To serve this purpose, DOCSIS 3.1 specifies a pilot-based wideband probing mode, where the sub-carriers of an OFDMA symbol are dedicated to channel estimation. The method of estimating such a channel from the pilot signal is unspecified, as the standard leaves opportunities for DOCSIS 3.1 device manufacturers to come up with their own algorithms. Since approaches that had been used for typical OFDMA systems either have inadequate performance, are unsuitable for the coaxial cable model or are too complex, there is a need for new efficient designs specifically for DOCSIS 3.1 systems.

There are several channel estimation techniques that have been studied for pilot-based estimation. The simplest one, which is Least Square (LS) estimation [1], does not require any channel state information (CSI). LS estimators work with samples in the frequency domain and are relatively low in complexity. However, they suffer from relatively high mean-square error, which is proportional to the power of additive white Gaussian noise (AWGN).

A better technique, which also performs estimation in the frequency domain, is linear minimum mean-square error (LMMSE) estimation [2]. This technique yields much better performance than the LS estimator, especially under low signal-to-noise ratio (SNR) scenarios. The major drawback of the LMMSE estimator is that it requires knowledge of the channel auto-correlation matrix and the noise variance, which are usually unknown at the receiver. The computational complexity of the LMMSE estimator is also very high as it requires a matrix inversion. Many have attempted to reduce the complexity of the LMMSE estimator [3], [4] at the expense of a small sacrifice in estimation accuracy.

Another very good approach uses discrete Fourier transform (DFT) based channel estimation. The DFT-based method firstly employs an LS estimator to obtain the channel’s frequency response (CFR). Then the discrete-time channel impulse response (CIR) is obtained by performing an inverse discrete Fourier transform (IDFT) on the CFR. Since the energy of the CIR is typically concentrated in a few taps having short delays, the algorithm’s performance can be improved if a few taps whose power is significantly higher than noise are preserved while the rest are forced to zero [5]. This operation is commonly referred to as denoising. After denoising, the CIR is transformed back to the frequency domain to obtain the estimated CFR. Consequently, the DFT approach helps to remove the noise power from the LS-estimated CFR. In general, DFT-based methods have moderate complexity thanks to Fast Fourier Transform (FFT) algorithms and perform much better than the LS estimator [6] at low SNRs.

However, with the DFT method, performance degradation can occur due to leakage to neighboring samples in the discrete-time CIR. There are two sources of leakage. The first is multipath components that have non sample-spaced delays. In the case of non sample-spaced delays, the energy from a single multipath component is spread over multiple samples in the discrete time CIR. The spreading is such the noise-only samples can not be eliminated without removing portions of the leakage energy. The second type of leakage emerges if not all sub-carriers are used for channel estimation. In particular, in a typical OFDMA system, the sub-carriers at both ends of the spectrum are left null to form guard bands. Not using the end sub-carriers degrades the performance of DFT-based techniques as this is equivalent to placing a rectangular window in the frequency domain which translates to convolution with a sinc-like function in the time domain. This causes the energy of the CIR to spread out in time. Denoising cuts off the tails of the sinc-like functions causing ripples around the edge sub-carriers when the denoised CIR is converted back to the frequency domain [7]. This phenomenon is often referred to as an “edge effect” or “border effect” and results in estimation errors not being equally distributed over all sub-carriers.

To date studies that effectively address the two leakage issues of the DFT-based techniques have not been found.

The usefulness of the standard channel estimation techniques discussed above is somewhat limited in DOCSIS 3.1 systems, as upstream wideband probing has a subcarrier skipping option. In subcarrier skipping mode, multiple upstream users transmit wideband probing signals on different subcarriers of the same OFDMA symbol. Each user transmits on a different set of subcarriers that are spaced $K$ sub-carriers apart, where $K$ is the number of simultaneous users. The use of subcarrier skipping with $K$ simultaneous users allows a $K$-fold increase in the efficiency of the wideband probing process as compared to a single user probing scheme. However, it places additional computational burden on the receiver, which must generate an estimate of the entire channel for each user despite receiving pilots on only every $K$th subcarrier.

This paper proposes a novel channel estimation technique that can be applied effectively to generic OFDMA systems, but was designed with DOCSIS 3.1 in mind. The proposed technique is capable of accurately interpolating between received pilot symbols, making it well-suited to channel probing schemes involving DOCSIS 3.1-style subcarrier skipping. It is pointed out that preliminary results of this work are published in a conference paper in [8]. Compared to [8],
the current paper not only offer more detailed description and analysis of the main iterative channel estimation technique, but also proposes major improvements, which substantially enhance the channel estimation accuracy, and at the same time, help to accurately detect the timing error.

The rest of this paper is organized as follows. Section II describes the DOCSIS 3.1’s OFDMA channel model. Conventional channel estimation techniques that can be used for OFDMA systems are discussed in Section III. Section IV proposes an iterative channel estimation approach, which is the main contribution of this paper. Section V presents and discusses simulation results. Section VI concludes the paper.

II. OFDMA Channel Model

Consider the complex baseband-equivalent OFDMA system shown in Fig. 2. The channel bandwidth is divided into $N$ sub-carriers, $M$ of which are used for channel estimation and are assigned BPSK preamble symbols $X[m]$, $m = 0, 1, \ldots, M - 1$. According to DOCSIS 3.1 specifications [9], there are guard bands at both ends of the allocated spectrum that cannot be used for data transmission. This restricts the location of the pilot sub-carriers, which are usually centered at the middle of the allocated spectrum. Denote the vector of pilot sub-carrier indexes as $\mathbf{S} = [S(0), S(1), \ldots, S(M - 1)]$. The vector elements are related by $S(m) = S(0) + mK$, $m = 0, 1, \ldots, M - 1$, where $S(0)$ is the “start” sub-carrier and $K$ is sub-carrier “skipping” factor [9].

The OFDMA transmitter employs an IDFT module of size $N$ for modulation. The standard IDFT/DFT is not used here, but rather the transform pair specified in DOCSIS 3.1, where the sub-carrier indexing is shifted by $-N/2$ sub-carriers. Using the DOCSIS 3.1 IDFT, the baseband complex equivalent of the transmitted time-domain samples are written as $x_d[n] = \frac{1}{\sqrt{N}} \sum_{m=0}^{M-1} X[m] \exp\left(\frac{2\pi i (S(m) - N/2)}{N} n\right)$, where $n = 0, 1, \ldots, N - 1$ denotes the sample index. To avoid inter-symbol interference (ISI), a cyclic prefix (CP) consisting of $N_{CP}$ samples is prefixed to the OFDMA symbol. After performing parallel to serial (P/S) conversion, the time-domain samples are serially passed through a DAC clocked at sampling rate $F_s$ and filtered with an image rejection filter to generate the continuous-time signal. Assuming ideal digital to analog (D/A) conversion, the complex baseband continuous-time signal can be expressed as

$$x_a(t) = \frac{1}{\sqrt{N}} \sum_{m=0}^{M-1} X[m] \exp\left(\frac{2\pi i (S(m) - N/2)(t - T_g)}{NT_s}\right),$$

$$0 \leq t \leq NT_s + T_g,$$

where $T_s = 1/F_s$ is the sampling period and $T_g$ is the guard interval in seconds. $T_s$ is the duration of the CP which is $N_{CP} T_s$. It is obvious that after the CP is inserted $x_a(t) = x_a(t + NT_s)$, $\forall \ t \in [0, T_g]$. In general, the validity of (1) depends on how well the up-conversion is performed.

A channel in a coaxial cable distribution network consists of many paths created by impedance mismatches among terminals and ports of devices that make up the network. Each path is characterized by a gain factor $\alpha_i$ and an associated delay $\epsilon_i$ normalized to sampling period $T_s$. Without loss of generality, $\epsilon_0$ is taken to be 0 and $\epsilon_i$ is the delay of path $i^{th}$ relative to...
the delay of path $0^\text{th}$ in unit samples. The impulse response of the baseband-equivalent of the multipath channel is given by

$$h_c(t) = \alpha_0 \delta(t) + \sum_{i=1}^{L-1} \alpha_i \delta(t - \epsilon_i T_s), \quad (2)$$

where $\delta$ is the Dirac delta function and $\alpha_i$ is a complex constant where magnitude represents the path loss and whose angle represents the phase difference in the local oscillators used for up and down conversion. Furthermore, the parameter $L$ is the number of paths in the multipath channel. The channel’s delay spread in seconds is $\epsilon_{max} T_s$, where $\epsilon_{max} = \max_{i=1, \ldots, L-1} \epsilon_i$ which is the delay of the longest multipath component relative to the first.

The continuous time signal received at the receiver is the convolution of the transmitted signal and the impulse response of the multipath channel. That is

$$y_a(t) = \int_{0}^{\infty} h_c(\theta) x_a(t - \tau_0 T_s - \theta) d\theta + w(t) = \alpha_0 x_a(t - \tau_0 T_s) + \sum_{i=1}^{L-1} \alpha_i x_a(t - \tau_0 T_s - \epsilon_i T_s) + w(t), \quad (3)$$

where $w(t)$ is a zero-mean AWGN noise process and $\tau_0$ is the timing offset (normalized to sampling period $T_s$) introduced by error in detecting the start time of the received OFDMA symbol. There are many coarse timing estimation techniques, e.g., [10]-[13], that can detect the start time of the received OFDMA frame. With coarse timing, the detection error is typically on the order of a few samples. Assume the system is well designed so that the length of the CP is greater than the channels delay spread, i.e., $T_C > \epsilon_{max} T_s$, as illustrated in Fig. 3. To ensure the DFT output is ISI-free, the error in the coarse timing detection must not exceed the slack in the CP. This places the following constraint on coarse timing error:

$$0 \leq \tau_0 \leq N_{CP} - \epsilon_{max}, \quad (4)$$

where $\tau_0$ is the error in coarse timing in samples. By defining $\tau_i = \epsilon_i + \tau_0$, $i = 1, 2, \ldots, L-1$, the timing error can be incorporated into the base-band channel to get the more realistic impulse response given by:

$$h_{\tau}(t) = \sum_{i=0}^{L-1} \alpha_i \delta(t - \tau_i T_s). \quad (5)$$

Then (3) simplifies to

$$y_a(t) = \sum_{i=0}^{L-1} \alpha_i x_a(t - \tau_i T_s) + w(t), \quad 0 \leq t \leq NT_s + T_g. \quad (6)$$

The continuous time signal is band-limited and digitally sampled with a complex A/D converter at the receiver at sampling rate $F_s$. After coarse timing detection is performed, the CP is removed. The discrete-time samples after CP removal are given by

$$y_d[n] = y_a(t) \bigg|_{t = n T_s + T_g} = \sum_{i=0}^{L-1} \alpha_i x_a(n T_s + T_g - \tau_i T_s) + w[n], \quad n = 0, 1, \ldots, N-1, \quad (7)$$

where $w[n]$ is $w(t)$ sampled at $t = n T_s + T_g$ after it has been band-limited. $w[n]$ is complex white Gaussian noise with zero mean and variance $\sigma_w^2$. To recover the data, an $N$-point DFT block transforms the time-domain sequence back to the frequency-domain:

$$Y[m] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} y_d[n] \exp\left(-j\frac{2\pi (S(m) - N/2)n}{N}\right)$$

$$= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \sum_{k=0}^{M-1} \sum_{i=0}^{L-1} X[k] \exp\left(\frac{j2\pi (S(k) - N/2)(n T_s - \tau_i T_s)}{NT_s}\right) \times \exp\left(-j2\pi (S(m) - N/2)n\right) + W[m],$$

$$m = 0, 1, \ldots, M-1, \quad (8)$$

where

$$W[m] = \sum_{n=0}^{N-1} w[n] \exp\left(-j\frac{2\pi (S(m) - N/2)n}{N}\right). \quad (9)$$

is complex Gaussian noise with zero mean and variance $\sigma_w^2$. Then

$$Y[m] = \frac{1}{\sqrt{N}} \sum_{k=0}^{M-1} X[k] \sum_{i=0}^{L-1} \sum_{n=0}^{N-1} \sum_{i=0}^{N-1} \exp\left(\frac{j2\pi (S(k) - S(m))n}{N}\right) \times \exp\left(-j2\pi (S(m) - N/2)\tau_i\right) + W[m]$$

$$= X[m] \sum_{i=0}^{L-1} \alpha_i \exp\left(-j\frac{2\pi (S(m) - N/2)\tau_i}{N}\right) + W[m]$$

$$= X[m] H[m] + W[m], \quad (10)$$

where $H[m]$ is the multipath channel’s frequency response at sub-carrier $S(m)$, given as

$$H[m] = \sum_{i=0}^{L-1} \alpha_i \exp\left(-j\frac{2\pi (S(m) - N/2)\tau_i}{N}\right). \quad (11)$$

With the input/output model of (10), the signal-to-noise ratio (SNR) of the received signal is defined as

$$\text{SNR} = \frac{\mathbb{E}\{|X[m]H[m]|^2\}}{\mathbb{E}\{|W[m]|^2\}} = \frac{\sum_{i=0}^{L-1} |\alpha_i|^2}{\sigma_w^2}. \quad (12)$$
The task of channel estimation is to obtain the frequency response of the *entire* channel, which is ideally given as

\[ F[k] = \sum_{i=0}^{L-1} \alpha_i \exp\left( -j2\pi \frac{(k - N/2)\tau_i}{N} \right), \]

\[ k = 0, 1, \ldots, N - 1 \]  

(13)

from known values of \( X[m] \) and observed values of \( Y[m] \). Conventional methods estimate \( M \) values of \( H[m] \) and then interpolate between them to get the entire frequency response, \( F[k] \). Alternatively, since \( F[k] \) is a function of the 2\( L \) unknown parameters \( \{\alpha_i, \tau_i\}_{i=0}^{L-1} \), an estimate of \( F[k] \) may be obtained from estimates of the 2\( L \) unknown parameters. In this paper we present a novel method to obtain 2\( L \) values of \( \{\alpha_i, \tau_i\}_{i=0}^{L-1} \). As long as 2\( L \ll M \), it will be demonstrated in the following section that estimating \( \{\alpha_i, \tau_i\}_{i=0}^{L-1} \) directly provides a better estimate of \( F[k] \).

### III. Conventional Channel Estimation Techniques

This section will discuss several methods that have been used in conventional OFDM/OFDMA systems. The performance measure used for comparison is the mean squared error between the ideal and estimated channel frequency responses.

#### A. Least Square (LS) Estimation

This is the simplest estimation technique, and it can be performed without any knowledge of the channel statistics. In particular, the LS technique estimates the frequency response of the channel at \( M \) frequencies from \( M \) observations. The \( M \) estimates, \( \{\hat{H}_{LS}[m]; \ m = 0, 1, \ldots, M - 1\} \) minimize \( \sum_{m=0}^{M-1} |Y[m] - \hat{H}_{LS}[m]X[m]|^2 \) [2]. The first step is to estimate the frequency response of the channel at the frequencies of the \( M \) pilot tones. The frequency response for pilot tone \( m \), which is sub-carrier \( S(m) \), is calculated by

\[ \hat{H}_{LS}[m] = \frac{Y[m]}{X[m]} = \frac{H[m]X[m] + W[m]}{X[m]} = H[m] + \frac{W[m]}{X[m]}, \]

\[ m = 0, 1, \ldots, M - 1. \]  

(14)

To be consistent with the literature in the field, the index \( m \) in \( \hat{H}_{LS}[m] \) is the pilot tone number, not the more commonly used sub-carrier number. The preamble symbols \( X(m) \) are drawn from a BPSK constellation and are taken to be \( \pm 1 \) without loss of generality. Then \( W[m] = \frac{W[m]}{X[m]} = \pm W[m] \) are Gaussian distributed random variables with zero mean and variance \( \sigma_w^2 \).

The second step is to apply interpolation on \( \hat{H}_{LS}[m] \) to obtain the estimated channel transfer function, \( \hat{F}_{LS}[k] \). There are a number of well-known interpolation techniques, including linear interpolation and Spline Cubic interpolation [14]. A detailed discussion of the merits of various interpolation techniques is outside the scope of this paper. Ideally, with zero interpolation error, the MSE performance of the LS estimator can be approximated in closed form as:

\[ \text{MSE}_{LS} = \mathbb{E}\left[ \left| H[m] - \hat{H}_{LS}[m] \right|^2 \right] = \mathbb{E}\left[ \left| \hat{W}[m] \right|^2 \right] = \sigma_w^2. \]  

(15)

#### B. DFT-Based Estimation

The DFT-based algorithm was designed for channels with an impulse response that has most of its energy concentrated in a small number of taps [6], [15], [16]. The DFT-based method reduces the noise in the estimate of the impulse response by setting the low energy taps to zero. Therefore, the performance of the DFT-based method depends on the energy distribution in the impulse response. If the energy is concentrated in a few taps then it will outperform the LS estimation.

The method involves a series of steps beginning with an \( M \)-point IDFT on \( \hat{H}_{LS}[m] \) to obtain the approximation of the time-domain channel response of the LS estimated channel,

\[ g[n] = \frac{1}{M} \sum_{m=0}^{M-1} Y[m] \frac{\exp\left( j2\pi m n \right)}{X[m]}, \]

\[ = \frac{1}{M} \sum_{m=0}^{M-1} H[m] \frac{\exp\left( j2\pi m n \right)}{M} \]

\[ + \frac{1}{M} \int_{\alpha}^{\beta} \exp\left( j2\pi m n \right) \left| \frac{\exp\left( j2\pi m n \right)}{M} \right| \]

\[ = d[n] + \eta[n], \quad n = 0, 1, \ldots, M - 1, \]  

(16)

where \( \eta[n] \) is AWGN with zero mean and variance \( \sigma_\eta^2 = \frac{\sigma_w^2}{M^2} \). \( d[n] \), which is the signal component of \( g[n] \), can be shown to be

\[ d[n] = \frac{1}{M} \sum_{m=0}^{M-1} \sum_{i=0}^{L-1} \alpha_i \exp\left( -j2\pi (S(m) - N/2)\tau_i \right) \exp\left( j2\pi m n \right) \]

\[ = \frac{1}{M} \sum_{m=0}^{M-1} \sum_{i=0}^{L-1} \alpha_i \exp\left( -j2\pi (S(0) + mK - N/2)(n/K) \right) \]

\[ \times \exp\left( j2\pi (S(0) + mK - N/2)(n/K) \right) \]

\[ \times \exp\left( -j2\pi (S(0) - N/2)(n/K) \right) \]

\[ \times \exp\left( j2\pi (S(0) - N/2)(n/K) \right) \]

\[ = \left( -j2\pi (S(0) - N/2)n \right) \sum_{i=0}^{L-1} \frac{\alpha_i}{M} \]

\[ \times \exp\left( j2\pi (S(0) + mK - N/2)(n/K) \right) \]

\[ \times \exp\left( j2\pi (S(0) - N/2)(n/K) \right) \]

\[ \times \exp\left( j2\pi (S(0) - N/2)(n/K) \right) \]

\[ \times \exp\left( j2\pi (S(0) - N/2)(n/K) \right) \]

\[ \times \exp\left( j2\pi (S(0) - N/2)(n/K) \right) \]

\[ \times \exp\left( j2\pi (S(0) - N/2)(n/K) \right) \]

\[ = \left( -j2\pi (S(0) - N/2)n \right) \sum_{i=0}^{L-1} \frac{\alpha_i}{M} \]

\[ \times \exp\left( j2\pi (S(0) + mK - N/2)(n/K - M\tau_i) \right) \]

\[ \times \exp\left( j2\pi (S(0) - N/2)(n/K - M\tau_i) \right) \]

\[ \times \exp\left( j2\pi (S(0) - N/2)(n/K - M\tau_i) \right) \]

\[ \times \exp\left( j2\pi (S(0) - N/2)(n/K - M\tau_i) \right) \]

\[ \times \exp\left( j2\pi (S(0) - N/2)(n/K - M\tau_i) \right) \]

\[ \times \exp\left( j2\pi (S(0) - N/2)(n/K - M\tau_i) \right) \]

\[ \times \exp\left( j2\pi (S(0) - N/2)(n/K - M\tau_i) \right) \]

\[ \exp\left( j\Delta \left(n - MK\tau_i/K \right) \sin(\pi (n - MK\tau_i/K)) / (\pi (n - MK\tau_i/K)) \right), \]  

(17)

where \( \Delta = 2S(0) - N + (M - 1)K \). It should be noted that if delay \( \tau_i \) is an integer multiple of \( N/\{MK\} \), the \( i \)-th channel path

\[ \text{Only under a set of special conditions } g[n] \text{ becomes the impulse response of the physical channel, i.e., } g[n] = h_c(t)|_{t=nT_s}. \]  

1
energy is concentrated in a single tap located at \( n = MK\tau / N \), otherwise the energy is distributed among all the taps of \( d[n] \).

The second step is to select the significant taps of \( g[n] \). A tap is deemed to be significant if its magnitude is much larger than the standard deviation of the of the noise, i.e., \( |g[n]| \gg \sigma \). Specifically, the sequence \( \hat{g}[n] \) of significant taps is constructed from \( g[n] \) using:

\[
\hat{g}[n] = \begin{cases} 
  g[n], & \text{if } |g[n]| > \Lambda \th \n  0, & \text{otherwise},
\end{cases}
\]

(18)

where the constant \( \Lambda \th \) is a threshold used to select the most significant taps of \( g[n] \).

The third step is to estimate the frequency response at all sub-carriers by taking the \( N \)-point DFT of \( \hat{g}[n] \). This yields:

\[
\hat{F}_\text{DFT}[k] = \sum_{n=0}^{M-1} \hat{g}[n] \exp\left(-j2\pi knN/MK\right),
\]

\[
= \sum_{n=0}^{M-1} g[n] \exp\left(-j2\pi kn/MK\right), \quad k = 0, 1, \ldots, N - 1.
\]

(19)

The MSE of the DFT-based estimator can be conveniently expressed as

\[
\text{MSE}_{\text{DFT}} = E\left\{ \frac{1}{N} \sum_{k=0}^{N-1} |\hat{F}_\text{DFT}[k] - F[k]|^2 \right\}
\]

\[
= E\left\{ \sum_{n=0}^{M-1} |\hat{g}[n] - d[n]|^2 \right\}
\]

\[
= E\left\{ \sum_{n:|g[n]| > \Lambda \th} |g[n] - d[n]|^2 \right\}
+ E\left\{ \sum_{n:|g[n]| \leq \Lambda \th} |d[n]|^2 \right\}
\]

\[
= \frac{\sigma_n^2 L_{\text{sig}}}{M} + P_{\text{prune}}
\]

(20)

where \( L_{\text{sig}} \) is the number of significant taps and \( P_{\text{prune}} \) is the power in the \( N - L_{\text{sig}} \) taps of \( d[n] \) that were pruned in the construction of \( \hat{g}[n] \). Raising the threshold \( \Lambda \th \) reduces \( L_{\text{sig}} \), thus reducing \( \frac{\sigma_n^2 L_{\text{sig}}}{M} \), but increases \( P_{\text{prune}} \). Since both contribute to the MSE, there is an optimum value of \( \Lambda \th \) that minimizes the MSE [6].

To reduce the complexity of the DFT method, system designers should select parameter \( M \) as a power of 2 to make use of the Fast Fourier Transform that can be applied to calculate (16), otherwise (16) will be computationally demanding.

C. LMMSE Estimator

The LMMSE estimator is a linear estimator that employs the second-order statistics of the channel to minimize the MSE [1], [3]. Since it uses information about the channel, the LMMSE estimator performs much better than the LS estimator. This extra information is the auto covariance function of the channel.

The LMMSE estimation of the channel can be done in either the time domain or the frequency domain [3]. In the time-domain, from (16) and (17), one has

\[
g = d + \eta,
\]

(21)

where \( g, d \) and \( \eta \) are defined as the column vectors:

\[
g = [g[0], g[1], \ldots, g[M-1]]^T,
\]

(22)

\[
d = [d[0], d[1], \ldots, d[M-1]]^T,
\]

(23)

\[
\eta = [\eta[0], \eta[1], \ldots, \eta[M-1]]^T.
\]

(24)

The problem becomes estimating \( d \) from \( g \). The best linear estimator in the mean-square error sense is given as [1]–[3]

\[
\hat{g} = \frac{R_{dd}}{R_{dd} + \Gamma M^2} \sigma_n^2 g.
\]

(25)

where \( R_{dd} \) is the auto-covariance matrix of \( d \), which is assumed to be known apriori, and \( I_M \) is the \( M \times M \) identity matrix. From (17), the vector \( d \) can be expressed as

\[
d = \begin{bmatrix}
  D_{0,0} & D_{0,1} & \cdots & D_{0,L-1} \\
  D_{1,0} & D_{1,1} & \cdots & D_{1,L-1} \\
  \vdots & \vdots & \ddots & \vdots \\
  D_{M-1,0} & D_{M-1,1} & \cdots & D_{M-1,L-1}
\end{bmatrix} \begin{bmatrix}
  \alpha_0 \\
  \alpha_1 \\
  \vdots \\
  \alpha_{L-1}
\end{bmatrix} = D \alpha,
\]

(26)

where \( D \) is the \( M \times L \) matrix with elements \( D_{n,l} = 0, 1, \ldots, M-1, \quad l = 0, 1, \ldots, L-1 \), defined as

\[
D_{n,l} = \exp\left(-j\frac{2\pi l (S(0) - N/2)n}{MK}\right) \times \exp\left(j\pi n (n - MK\tau / N) / \sin(\pi (n - MK\tau / N) / M)\right).
\]

(27)

It is pointed out that matrix \( DD^H \), where the superscript \( H \) denotes the conjugate transpose operation, is generally known as the leakage matrix [17], [18], since it describes the correlation among the discrete channel taps. When the echoes in the channel have delays that are integer multiples of \( \text{NT}/(MK) \), \( DD^H \) becomes a size \( M \times M \) diagonal matrix.

The delays in the multipath propagation environment in CATV networks change slowly, or not at all, over time. Assuming the delays and magnitudes of the echoes are known, i.e., \( \tau_i \) and \( |\alpha_i| \) are known, the channel auto-covariance matrix can be written as

\[
R_{dd} = \text{E} \left[ d d^H \right] = \text{E} \left[ D \alpha \alpha^H D^H \right] = D \text{E} \left[ \alpha \alpha^H \right] D^H
\]

\[
\begin{bmatrix}
  |\alpha_0|^2 & 0 & 0 & \cdots \\
  0 & |\alpha_1|^2 & 0 & \cdots \\
  0 & 0 & \ddots & \cdots \\
  0 & 0 & \cdots & |\alpha_{L-1}|^2
\end{bmatrix}
\]

(28)
The mean squared estimation error can be expressed as [19]

$$\text{MSE}_{\text{LMMSE}} = E \left\{ \| \hat{g} - d \|^2 \right\} = \sigma_n^2 \text{Trace} \left( \frac{R_{dd}}{R_{dd} + \frac{1}{M} \sigma_n^2} \right) = \frac{\sigma_n^2}{M} \sum_{i=0}^{M-1} \frac{\lambda_i}{\lambda_i + \sigma_n^2/M} = \frac{\sigma_n^2}{M} \sum_{i=0}^{M-1} \frac{|\alpha_i|^2}{|\alpha_i|^2 + \sigma_n^2/M}$$

(29)

where $\lambda_i, \ i = 0, 1, \ldots, M - 1$ are the eigenvalues of $R_{dd}$.

There are many approaches to finding a good approximation for $R_{dd}$, such as those presented in [15], [16], [18], and [20]. All require additional observations or some apriori knowledge of the channel. Clearly, the accuracy of the LMMSE channel estimator depends on quality of the estimate used for $R_{dd}$. The MSE given by (29) assumes $R_{dd}$ is known exactly. Since $R_{dd}$ is rarely if ever known exactly, (28) serves as an optimistic performance target for channel estimators.

IV. Iterative Channel Estimation Algorithm

In this section, a novel iterative channel estimation algorithm is developed for channels where the only distortion is due to multipath propagation. While the algorithm is novel, the idea of using an iterative approach is not. For example, the SAGE algorithm in [21] uses an iterative algorithm in a very different set of circumstances. This iterative algorithm is developed sequentially in three sub-sections. Section IV-A constructs a basic estimator that illustrates the principle of operation. The algorithm is revised in Section IV-B to significantly enhance the accuracy and then generalized in Section IV-C. Section IV-D then summarizes the algorithm. Finally Section IV-E provides complexity analysis of the final algorithm.

A. Iterative Channel Estimation

The iterative algorithm assumes a multipath channel that has a finite number of paths and is designed to estimate the channel parameters, which are time delays $\tau_i$ and complex amplitudes $\alpha_i, \ i = 0, 1, \ldots, L - 1$, of the paths. The estimated parameters, denoted as $\hat{\tau}_i$ and $\hat{\alpha}_i$ can be used to obtain the frequency response of the entire channel with the following equation:

$$\hat{F}_{\text{ICE}}[k] = \sum_{i=0}^{L-1} \hat{\alpha}_i \exp \left( -\frac{j2\pi (k - N/2) \hat{\tau}_i}{N} \right), \quad (30)$$

where $k = 0, 1, \ldots, N - 1$. The proposed iterative channel estimation technique starts by performing an inverse Fourier transform on $\hat{F}_{\text{ICE}}[m]$. The transform has a length of $NU/K$, where $U$ is a parameter which controls the resolution of the resulting time domain vector $q[u]$, given as

$$q[u] = \frac{1}{M} \sum_{m=0}^{M-1} Y[m] \exp \left( \frac{j2\pi (m + (S(0) - N/2)/K) u}{NU/K} \right) \exp \left( \frac{j\pi(M-1)u}{N} \right).$$

(31)

Note that $U$ is not necessarily an integer, but rather a number that is chosen to make $NU/K$ an integer. One suggestion is to make $NU/K$ a power of two so that the complexity of (31) can be reduced through the use of Fast Fourier Transform algorithm. Equation (31) can be simplified as follows:

$$q[u] = \frac{1}{M} \sum_{m=0}^{M-1} Y[m] \exp \left( \frac{j2\pi (m + (S(0) - N/2)/K) u}{NU/K} \right) \exp \left( \frac{j\pi(M-1)u}{N} \right) \times \exp \left( \frac{j2\pi(M-1)u}{N} \right) = \frac{1}{M} \sum_{m=0}^{M-1} Y[m] \exp \left( \frac{j2\pi (m + (S(0) - N/2)/K) u}{NU/K} \right), \quad (32)$$

where $\rho[u]$ is the AWGN noise component given by

$$\rho[u] = \frac{1}{M} \sum_{m=0}^{M-1} W[m] \exp \left( \frac{j2\pi (m + (S(0) - N/2)/K) u}{NU/K} \right), \quad (33)$$

which has zero mean and variance $\sigma_n^2 = \sigma_w^2$.

Since $M \leq N/K$, the complexity of (31) is equivalent to an $NU/K$-point IDFT. The signal component of (31) can be expressed in a more meaningful form as

$$b[u] = \frac{1}{M} \sum_{m=0}^{M-1} \sum_{i=0}^{L-1} \alpha_i \exp \left( -\frac{j2\pi (S(m) - N/2) \tau_i}{N} \right) \times \exp \left( \frac{j2\pi (m + (S(0) - N/2)/K) u}{NU/K} \right) = \sum_{i=0}^{L-1} \frac{1}{M} \sum_{m=0}^{M-1} \exp \left( -\frac{j2\pi (S(0) + mK - N/2) \tau_i}{N} \right) \times \exp \left( \frac{j2\pi (mK + S(0) - N/2) u}{N} \right) \times \exp \left( \frac{j2\pi (S(0) - N/2) (u - \tau_i)}{N} \right) \times \exp \left( \frac{j2\pi (u - \tau_i)}{N} \right) - 1 \times \exp \left( \frac{j2\pi (u - \tau_i)}{N} \right) - 1 \times \exp \left( \frac{j2\pi (u - \tau_i)}{N} \right) = \sum_{i=0}^{L-1} \sin \left( \pi \frac{(u - \tau_i)MK}{N} \right) \times \exp \left( \frac{j\pi (2S(0) - N + (M - 1)K)(u - \tau_i)}{N} \right) \times \exp \left( \frac{j\pi (u - \tau_i)}{N} \right) \times \exp \left( \frac{j\pi (u - \tau_i)}{N} \right) - 1 \times \exp \left( \frac{j\pi (u - \tau_i)}{N} \right) - 1 \times \exp \left( \frac{j\pi (u - \tau_i)}{N} \right) = \sum_{i=0}^{L-1} \alpha_i \gamma \left( \frac{u}{U} - \tau_i \right), \quad u = 0, 1, \ldots, \langle NU/K \rangle - 1, \quad (34)$$

where $b[u]$ is represented as a summation of several channel path kernel functions, $\gamma(\cdot)$, that are delayed by $\tau_i$ and scaled in amplitude by $\alpha_i$. Each $\gamma(\cdot)$ function represents a path in the channel.
The shape of the $\Upsilon(\cdot)$ function is more clear when it is expressed as

$$\Upsilon(x) = \exp\left(\frac{\pi \Delta x}{N}\right) \text{psinc} \left(\frac{xK}{N}, M\right).$$

(35)

where $\Delta = 2S(0) - N + (M - 1)K$ and \text{psinc}(x, M) is the Dirichlet or periodic sinc function, defined as

$$\text{psinc}(x, M) = \begin{cases} \sin(\pi Mx), & x \notin \mathbb{Z} \\ \frac{\sin(\pi x)}{M \sin(\pi x)}, & x \in \mathbb{Z}. \end{cases}$$

(36)

The $\Upsilon(x)$ function has zero-crossings at integer multiplies of $N/(MK)$, and therefore the width of the main lobe is $2N/(MK)$, which is inversely proportional to $KM$, as illustrated in Fig. 4. This relatively narrow main lobe ensures the magnitude of $\Upsilon(x)$ has a sharp peak at $x = 0$. Furthermore, the peak value of $\Upsilon(x)$ is 1. The iterative channel estimation technique centers on peak detection of $q(u)$. Without loss of generality, the path indices are defined based on path strength such that

$$|\alpha_0| \geq |\alpha_1| \geq \ldots \geq |\alpha_{L-1}|.$$  

(37)

Provided $U$ is chosen large enough for

$$\left|\frac{\tau_i U}{U} - \tau_i\right| < \left|\frac{\tau_j U}{U} - \tau_j\right|, \quad \forall j \neq i, \ i = 0, 1, 2, \ldots, L - 1,$$

(38)

where $\langle \cdot \rangle$ indicates rounding, then from (35), (37) and (38) it follows that

$$|\alpha_0 \Upsilon\left(\frac{\tau_0 U}{U} - \tau_0\right)| > |\alpha_i \Upsilon\left(\frac{\tau_0 U}{U} - \tau_i\right)|,$$

$$\forall i = 1, 2, \ldots, L - 1. \quad (39)$$

This indicates $\alpha_0 \Upsilon(\langle \tau_0 U \rangle / U - \tau_0)$ is the dominant magnitude contributor to $b(\langle \tau_0 U \rangle)$. The estimates of $\alpha_i$ and $\tau_i$ can be found iteratively starting with a rough approximation of the parameters of the first path, $\alpha_0$ and $\tau_0$ as follows:

$$u_0^{[1]} = \arg\max_{u=0,1,\ldots,NU/K-1} |q(u)|,$$

$$\hat{\tau}_0^{[1]} = u_0^{[1]} / U \quad \text{and} \quad \hat{\alpha}_0^{[1]} = q\left[u_0^{[1]}\right],$$

(40)

where the super script $[1]$ indicates that the value was found on the first iteration. Rough estimates of the parameters of the second path can then be generated by subtracting the estimated contribution of the first path, $\hat{\alpha}_0^{[1]} \Upsilon(u/U - \hat{\tau}_0^{[1]})$, from $q(u)$ to yield:

$$u_1^{[1]} = \arg\max_{u=0,1,\ldots,NU/K-1} |q(u) - \hat{\alpha}_0^{[1]} \Upsilon(u/U - \hat{\tau}_0^{[1]})|,$$

$$\hat{\tau}_1^{[1]} = u_1^{[1]} / U \quad \text{and} \quad \hat{\alpha}_0^{[1]} = q\left[u_1^{[1]}\right] - \hat{\alpha}_0^{[1]} \Upsilon\left(u_1^{[1]} / U - \hat{\tau}_0^{[1]}\right).$$

(41)

Similarly, rough estimates for $\alpha_i$, $\tau_i$, $i = 2, 3, \ldots, L - 1$, are found using

$$u_i^{[1]} = \arg\max_{u=0,1,\ldots,NU/K-1} |q(u) - \sum_{k=0}^{i-1} \hat{\alpha}_k^{[1]} \Upsilon(u/U - \hat{\tau}_k^{[1]})|,$$

$$\hat{\tau}_i^{[1]} = u_i^{[1]} / U \quad \text{and} \quad \hat{\alpha}_i^{[1]} = q\left[u_i^{[1]}\right] - \sum_{k=0}^{i-1} \hat{\alpha}_k^{[1]} \Upsilon\left(u_i^{[1]} / U - \hat{\tau}_k^{[1]}\right).$$

(42)

After a set of rough estimates are obtained, the estimates for $\alpha_0$ and $\tau_0$ can be improved by removing the estimated contributions from paths 1 to $L - 1$ from $q(u)$. The improved estimates of $\alpha_0$ and $\tau_0$ in this 2nd iteration are given by

$$u_0^{[2]} = \arg\max_{u=0,1,\ldots,NU/K-1} |q(u) - \sum_{k=1}^{L-1} \hat{\alpha}_k^{[1]} \Upsilon(u/U - \hat{\tau}_k^{[1]})|,$$

$$\hat{\tau}_0^{[2]} = u_0^{[2]} / U \quad \text{and} \quad \hat{\alpha}_0^{[2]} = q\left[u_0^{[2]}\right] - \sum_{k=1}^{L-1} \hat{\alpha}_k^{[1]} \Upsilon\left(u_0^{[2]} / U - \hat{\tau}_k^{[1]}\right).$$

(43)

In a similar manner, the better estimates of $\alpha_0$ and $\tau_0$ can be used to produce better estimates of $\alpha_1$ and $\tau_1$. The estimates can be continually improved in this iterative fashion. Specifically, the impulse response of path $i$ on iteration $v$ is approximated by

$$q_i^{[v]}(u) \overset{\text{def}}{=} q(u) - \sum_{l=0}^{i-1} \hat{\alpha}_l^{[v-1]} \Upsilon\left(u/u - \hat{\tau}_l^{[v-1]}\right) - \sum_{k=1}^{L-1} \hat{\alpha}_k^{[v-1]} \Upsilon\left(u/u - \hat{\tau}_k^{[v-1]}\right) \approx \alpha_i \Upsilon(u/U - \tau_i),$$

(44)

and the improved estimates of $\alpha_i$ and $\tau_i$ are given by

$$u_i^{[v]} = \arg\max_{u=0,1,\ldots,NU/K-1} |q_i^{[v]}(u)|,$$

$$\hat{\tau}_i^{[v]} = u_i^{[v]} / U \quad \text{and} \quad \hat{\alpha}_i^{[v]} = q_i^{[v]}\left[u_i^{[v]}\right].$$

(45)

After several iterations the estimates reach, or at least nearly reach, steady state values, which are denoted as $\hat{\alpha}_i$ and $\hat{\tau}_i$. These steady state values are used in (30) to obtain the estimated frequency response of the channel.

The accuracy of the aforementioned approximation is significantly affected by the error in $\hat{\tau}_i$. The proposed approach ideally estimates $\hat{\tau}_i = \langle \tau_i U \rangle / U$. Thus the estimation error
in the worst case could be 0.5/U, i.e., |τ̂i − τi| ≤ 0.5/U. Obviously, larger U reduces the error at the cost of increasing the estimator’s complexity. Furthermore, the error in τ̂i has the corresponding effect of diminishing the magnitude of ̂ai by up to |Y(0.5/U)|. For the case of M = N, the reduction can be as much as -4 dB, -0.9 dB and -0.2 dB for U = 1, U = 2 and U = 4, respectively.

B. Further Enhancing Estimation Accuracy

The estimate for τ used in Section IV-A is rather crude. The accuracy of this estimator, which simply rounds the argument of (45) to the sample nearest to the peak, can be improved by interpolating between a few samples nearest the peak, thus eliminating the rounding error, as shown in Fig. 5. There are different ways to interpolate between samples. A common approach is to approximate the vicinity of a peak using a parabolic function [22]. Such an approach would be very accurate provided the three digital samples used for approximation are in close proximity to the peak, which implies a high resolution factor U. This subsection applies the log-domain interpolation technique [1] that uses only two digital samples in the vicinity of the peak, while providing better interpolation accuracy, especially in the case of low resolution factor U.

The method used here is to find the parameters ̂τi and ̂ai that force ̂aiY(̂ui/U − ̂τi) and ̂aiY((̂ui + 1)/U − ̂τi) to equal the samples on each side of the peak of the argument, where the pair of indices u = ̂ui and u = ̂ui + 1 can be found, or at least estimated as

\[ \hat{u}_{i,\text{max}} = \arg\max_{u=0,1,...,NU/K-1} \left| q_i^{[u]}[u] \right|, \]

\[ \hat{a}_i = \begin{cases} \hat{a}_{i,\text{max}}, & \text{if } \left| q_i^{[u]}[\hat{a}_{i,\text{max}} + 1] \right| \geq \left| q_i^{[u]}[\hat{a}_{i,\text{max}}] \right| \end{cases}, \]

\[ \hat{u}_i = \hat{u}_{i,\text{max}} - 1, \quad \text{otherwise}. \] (46)

As shown in Fig. 4, the Y(·) function has multiple extrema. For the interpolation to yield a unique and accurate solution, the two samples used must be on opposite sides of the main lobe peak, i.e., ̂ui/U ≤ ̂τi ≤ (̂ui + 1)/U. Under this constraint, the solutions for ̂ai and ̂τi can be found by simultaneously solving

\[ \begin{align*} \hat{a}_i Y(\hat{u}_i/U - \hat{\tau}_i) &= q_i^{[\hat{u}_i]}, \\
\hat{a}_i Y((\hat{u}_i + 1)/U - \hat{\tau}_i) &= q_i^{[\hat{u}_i + 1]}, \end{align*} \] (47)

The solution for ̂τi from (47) is given by ̂τi = (̂ui + ̂θi)/U, where ̂θi is the root of the following equation:

\[ \frac{q_i^{[\hat{u}_i]}}{q_i^{[\hat{u}_i + 1]}} = \frac{\gamma(-\theta_i/U)}{\gamma((1 - \theta_i)/U)} = \exp\left(\frac{-j\pi \Delta}{NU} \sin\left(\frac{\pi M \theta_i}{NU} \right) \sin\left(\frac{\pi (1 - \theta_i)}{NU} \right) \right), \] (48)

where \( \theta_i \in [0, 1] \) and \( \Delta \) is defined in (35). Since \( \sin(x) > 0 \) \( \forall x \in (0, \pi) \), equation (48) can be simplified to

\[ \begin{align*} \left| q_i^{[\hat{u}_i]}/q_i^{[\hat{u}_i + 1]} \right| &= \sin\left(\frac{\pi M \theta_i}{NU} \right) \sin\left(\frac{\pi (1 - \theta_i)}{NU} \right) \end{align*} \] (49)

Denoting \( \kappa = \frac{|q_i^{[\hat{u}_i]}|}{|q_i^{[\hat{u}_i + 1]}|} \), \( \theta_i \) can be found as

\[ \theta_i = \Gamma^{-1}(\kappa), \quad 0 \leq \theta_i \leq 1. \] (50)

Solving (50) in real-time is possible, but very costly due to the complexity of \( \Gamma^{-1}(\cdot) \). Moreover, the precision of the computation must be very high when \( \theta_i \) is close to 0 or 1, i.e., when the denominator of (49) approaches zero.

Fortunately, (50) can be modified to yield a hardware friendly form. Since \( \kappa = e^{\ln(\kappa)}, \theta_i \) can be expressed as the function of \( \ln(\kappa) \) defined as \( \Gamma^{-1}(\ln(\kappa)) = \Gamma^{-1}(e^{\ln(\kappa)}) = \Gamma^{-1}(\kappa) \). The simplicity of the logarithmic form is illustrated in Fig. 6, where \( \theta_i \) is plotted as a function of \( \ln(\kappa) \) for 3 values of \( U \). It is apparent from Fig. 6 that for \( U \geq 2, \Gamma^{-1}(\cdot) \) is nearly linear and can be approximated by \( \Gamma^{-1}(\ln(\kappa)) \approx \beta \ln(\kappa) + 0.5 \),
where $\beta$ is a coefficient that can be pre-computed from system parameters $K, M, N$ and $U$ as
\[
\beta = \frac{-0.5}{\ln(\sin(\frac{\pi M K}{N_T})/M) - \ln(\sin(\frac{\pi K}{N_T})).}
\] (51)

$\theta_i$ can then be approximated by
\[
\theta_i \approx \beta \ln(\kappa) + 0.5
\]
\[
= \beta \left( \ln \left( \left| q_i^{[v]}[\tilde{a}_i] \right| \right) - \ln \left( \left| q_i^{[v]}[\tilde{a}_i + 1] \right| \right) \right) + 0.5,
\] (52)

for $U \geq 2$. Note that (52) does not require a division operation, and is therefore significantly more hardware friendly than (50).

There are two ways to find the echo strength $\hat{\alpha}_i$. The straightforward approach is to substitute $\theta_i$ into (47), yielding the following expression:
\[
\hat{\alpha}_i = \frac{q_i^{[v]}[\tilde{a}_i]}{\Upsilon(-\theta_i/U)}.
\] (53)

Although (53) is computationally simple, it amplifies the noise and ISI present in $q_i^{[v]}[\tilde{a}_i]$. In the worst-case scenario, this results in the noise and ISI being increased by a factor of $|\Upsilon(1/U)\Upsilon^{-1}|$. This factor decreases as the resolution factor $U$ increases. For the system parameters shown in Fig. 6, the worst case amplification values are 11.12dB, 1.66dB and 0.39dB for $U = 1, 2$ and 4, respectively. Alternatively, $\hat{\alpha}_i$ can be found as:
\[
\hat{\alpha}_i = \frac{1}{M} \sum_{m=0}^{M-1} \hat{H}_L[m] \exp \left( \frac{j2\pi (S(m) - N/2)\hat{\tau}_i}{N} \right)
\]
\[
- \sum_{k=0, k \neq i}^{L-1} \hat{\alpha}_k \Upsilon(\hat{\tau}_i - \hat{\tau}_k),
\] (54)

which is computationally more expensive than (53), but it prevents the noise and interference amplification effect. Consequently, for $U \leq 2$, it is advisable to use (54) in order to avoid severe performance degradation due to noise amplification. For $U > 2$, the amplification effect is minimal, so the more computationally efficient (53) is preferred.

C. Estimating the Number of Channel Paths

The ICE technique does not require any channel information, except for an initial estimation of the number of channel paths, denoted as $\hat{L}$, which must be determined before performing the iterative channel estimation. In reality, the parameter $L$ in (34) should be replaced by $\hat{L}$ so that the ICE algorithm will estimate $2\hat{L}$ channel parameters, $\{\alpha_i, \tau_i\}_{i=0}^{L-1}$, instead of $2L$. Therefore it is reasonable to expect the best performance achieved when $\hat{L} = L$.

In some cases, such as CATV networks, the plant is maintained to limit the number of dominant echo paths. Historically, cable network have been designed to have no more than 3 echoes ($L \leq 4$). To facilitate higher data rates, DOCSIS 3.1 networks are expected to eventually be upgraded such that $L \leq 2$. Therefore it may be reasonable to fix parameter $\hat{L} = 2$ in equipment used in DOCSIS 3.1 upstream transmission.

Although the ICE technique was initially designed for DOCSIS 3.1 systems, it applies to general OFDMA systems, where the parameter $L$ is not so constrained and the initial guesstimations of $\hat{L}$ affects the channel estimation performance. In particular, with the proposed ICE technique, if the number of paths in the channel is under-detected, i.e., $\hat{L} < L$, there will be performance degradation as the model is unable to compensate for the least significant channel paths. If the number of path is over-detected, i.e., $\hat{L} > L$, the ICE technique would interpret noise samples as channel paths. Since the power of noise is much less than the power of an echo, e.g., $|\alpha_{\tau_{j-1}}|^2 > \sigma_n^2$, performance degradation due to over-detection is generally less than the degradation caused by under-detection. Therefore, it is better to err on the side of over-detection.

Moreover, the performance degradation caused by over-detection can be mitigated as the significant power difference between echoes and noise can be exploited to suppress the over-detected paths. In particular, a threshold can be employed to differentiate the channel paths from the noise. The thresholding process replaces (53) and (54) with:
\[
\hat{\alpha}_i = \begin{cases} 0, & \text{if } |q_i^{[v]}[\tilde{a}_i]| \leq \lambda_T \\ \frac{1}{M} \sum_{m=0}^{M-1} \hat{H}_L[m] \exp \left( \frac{j2\pi (S(m) - N/2)\hat{\tau}_i}{N} \right) \\ - \sum_{k=0, k \neq i}^{L-1} \hat{\alpha}_k \Upsilon(\hat{\tau}_i - \hat{\tau}_k), & \text{if } |q_i^{[v]}[\tilde{a}_i]| > \lambda_T \& U \leq 2 \end{cases}
\] (55)

where $\lambda_T$ is the threshold level. With any threshold level decision, there is always some probability of a false alarm where a noise sample is declared as an echo. The threshold can be set to obtain a false alarm probability of $P_e$, using
\[
\lambda_T = \sqrt{-\frac{\sigma_n^2}{\ln(P_e)}} = \sqrt{-\frac{\sigma_n^2}{M \ln(P_e)}}.
\] (56)

Simulation results show that the estimation performance is not particularly sensitive to threshold level $\lambda_T$. A reasonable threshold is obtained by setting the false alarm probability to $P_e = 10^{-3}$.

D. Summary

The iterative channel detection procedure is summarized below:

1) Perform an $\frac{N_L}{K}$-point IDFT on $\hat{H}_L[m]$ to obtain a transformation of the channel response, $q[u]$, as given in (31).

2) Conservatively guesstimating a value for $\hat{L}$ based on the assumed characteristics of the channel, making sure $\hat{L} \geq L$.

3) Initialize the iteration number and channel path parameters, i.e., $\nu = 1$ and $\hat{\alpha}_i^{[0]} = \hat{\tau}_i^{[0]} = 0$; $i = 0, 1, \ldots, \hat{L} - 1$.

4) Subtract the effect of the estimated channel paths from $q[u]$ using (44) with $\hat{\tau}_i$ and $\hat{\alpha}_i$ in place of $\hat{\tau}_i$ and $\hat{\alpha}_i$. 

5) Estimate $\hat{g}_i^{[v]}$ as in (46) then use (50) to calculate $\tilde{z}_i^{[v]}$ if $U = 1$. If $U > 1$, use (52) instead to simplify the computation.
6) Use (55) to obtain $\tilde{g}_i^{[v]}$.
7) Repeat steps 4) to 6) $\tilde{L}$ times, starting with $i = 0$ and ending with $i = \tilde{L} - 1$ to complete one iteration.
8) Increase $v$ by 1 and repeat from step 4). Stop the process after a preset number of iterations.
9) Finally, the transfer function of the multipath channel can be computed similarly to (30) with channel parameters $\tilde{\alpha}_i$ and $\tilde{\epsilon}_i$ obtained from the last iteration.

E. Complexity Analysis

The complexity of the proposed algorithm will be described in terms of complex multiplies required per input sample. The number of multiplies to calculate a length-$NU$ IDFT is $\frac{NU\log_2 NU}{K}$. At each iteration, the subtraction in (44) requires $L(L-1)NU/K$ multiplies, the path delay approximation in (52) and the path gain approximation in (55) require $L$ and $L(N/K+L-1)$ multiplies, respectively. Finally, calculating the transfer function of the multipath channel, i.e., (30), requires $NL$ multiplies. Therefore, the total number of multiplies required to produce an estimate of the channel is $\frac{NU\log_2 NU}{K} + N\bar{L}((L-1)NU/K + N + L/K) + NL$, where $N\bar{L}$ is the number of iterations. Since the algorithm is performed block-wise, which take into account $N$ samples of an OFDMA symbol to process, thus its complexity in terms of complex multiplies required per input sample is

$$U \log_2 \frac{NU}{K} + \frac{N\bar{L}((L-1)\frac{NU}{K} + N + L/K)}{N} + L. \quad (57)$$

V. Simulation Results

This section investigates the performance of the proposed channel estimation algorithm. At first, the single echo channel model in [9] is considered, which fixes $\bar{L} = L = 2$. Coarse timing error $t_0$ is modeled as a random variable that is uniformly distributed between 0 and 10, i.e., $t_0 \sim U(0, 10)$. The echo delay in seconds, i.e., $\epsilon_1 Ts$, is uniformly distributed between 0 and 0.5$\mu$s. The power of the micro-reflection is $-16$dBc relative to the main path, which is the worst case specified in DOCSIS 3.1. The sampling rate of the system is $F_s = 102.4$MHz. The signal is generated using an $N = 2048$ point IFFT, and has $M = 1900$ pilots indexed by $S(m) = m + 74$, $m = 0, 1, \ldots, 1899$ (no sub-carrier skipping), which leaves 74 unused carriers as guard bands at both ends of the spectrum. Fig. 7 presents the performance of the proposed iterative channel estimator (ICE) in a DOCSIS 3.1 channel and compares it with the performance of the conventional estimators discussed previously. It can be seen that performance of the simple LS approach (no priori channel information) is the worst among all, followed by the DFT method. Despite having superior performance over the LS method in the low SNR region, the performance of the DFT method is worse than the LS method for SNR greater than 33 dB. It is apparent that 33 dB marks the location where the leakage power in the estimation is greater than noise power and becomes the main source of performance degradation for the DFT method (see Eq. (20)).

Performance of the ICE technique is illustrated for three resolution factors: $U = 1$, $U = 2$ and $U = 4$. The ICE method clearly outperforms both the LS and DFT techniques, even with $U = 1$. However, with $U = 1$, the ICE method hits an error floor of $4 \cdot 10^{-5}$ for SNRs above 20 dB that can not be reduced by increasing the number of iterations.

When increasing the resolution factor to $U = 2$, it can be seen that the error floor depends on the number of iterations. As can be seen in Fig. 7, there is considerable improvement between 2 and 20 iterations. In particular, the MSE level at 20 iterations is 10 times lower than that achieved with 2 iterations. However, only marginal performance gains are obtained by exceeding 20 iterations. Interestingly, a further increase of the resolution factor to 4 has little effect on the algorithm performance. Specifically, the MSE level with $U = 4$ and 20 iterations is almost identical to that with $U = 2$ and 20 iterations. The simulation results provide strong evidence that a resolution factor of $U = 2$ and 20 iterations should be taken as an upper limit.

It can be seen that the proposed algorithm outperforms the conventional methods, especially in the low-SNR region, where the ICE estimator is $30$dB better than the LS. In addition, the performance of the ICE asymptotically approaches that of the LMMSE, but does not need apriori knowledge of the auto-covariance of the channel. Furthermore, the ICE method requires only a single OFDMA symbol to achieve this level of performance.

Fig. 8 illustrates the MSE of the ICE estimator for different echo delays where $U = 2$ and the SNR is 20 and 40 dB. Simulations were run $10^5$ times with different echo delays that are uniformly distributed between 0 and 10 samples. The more realistic fractional delay channel model shows that the leakage becomes significant when $\epsilon_1 \leq 1$. As can be seen
from the figure, the estimation error reaches its peak when two paths are separated by about half the sample period. When only one iteration is performed, the estimation error gradually declines as the echo delay increases, as the influence of the main path on the echo and vice versa is lessened when the delay between them increases. Interestingly, when the estimator employs more than 5 iterations, the MSE quickly drops to error floors of $10^{-5}$ and $10^{-7}$ for SNR = 20 and 40 dB, respectively, which indicates that the residual leakage is successfully suppressed.

It is notable that when the echo delay is less than a sample period, the estimation error is very high as compared to the error caused by larger echo delay. Specifically, when $\epsilon_1 < 1$, the performance of the ICE technique is limited to around $10^{-4}$ regardless of SNR level. That observation indicates that the error floor of the ICE method shown in Fig. 7 is likely caused by scenarios in which the echo delays are less than 1 sample duration. However, even in such a challenging scenario, the estimation performance is still very reasonable (MSE $\leq 10^{-4}$ after 20 iterations, which is 20dB better than the LS). Note that the algorithm requires a larger number of iterations to accurately detect echoes with delay less than one sample. Therefore, the estimation process can be further simplified by adaptively performing more iterations for echoes with short delays and fewer iterations for echoes with longer delays. Simulations indicate that for any echo delay greater than 6 samples, only 2 iterations are needed to accurately detect the multipath channel’s coefficients. Since $t_0$ is the delay of the main path, it is also the timing error. Therefore the proposed technique not only estimates the channel’s frequency response, but also detects the timing error. Fig. 9 illustrates the variance of timing estimation error normalized to a sample period, i.e., $\mathbb{E}(|\hat{t}_0 - t_0|^2)$, for various SNR values. The simulation parameters are the same as in Fig. 7, where the resolution factor $U$ is fixed at 2 and a maximum of 20 iterations were executed for each detection. To avoid the most difficult detection scenario discussed previously, $\epsilon_1$ is constrained to be greater or equal to 1. As can be seen in Fig. 9, at 20 iterations, the timing MSE decays exponentially with SNR, indicated by the straight line in the figure.

Conventional OFDMA timing detection techniques, such as [10]–[13], are all timing-metric based estimation techniques, which limit detection resolution to a sample period. Therefore, timing offset variance of these conventional techniques is inherently greater than $1/12$, which is easily outperformed by the proposed algorithm. The initial ICE algorithm of Section IV-A has a timing offset that is uniformly distributed between $-0.5/U$ and $0.5/U$. Therefore, the timing variance of the initial ICE algorithm would asymptotically approach $U^{-2}/12$, which are the four horizontal lines in Fig. 9 tagged with $U = 1$, $U = 4$, $U = 32$ and $U = 2048$. As such, the interpolation method introduced in Section IV-B is very important as it provides excellent timing estimates for a small computational cost. Fig. 10 illustrates the performance for a 6 echo scenario ($L=7$), where the 6 echoes have strengths and
NGUYEN et al.: NOVEL ITERATIVE OFDMA CHANNEL ESTIMATION TECHNIQUE FOR DOCSIS 3.1 UPLINK CHANNELS 373

Fig. 10. Performance of the proposed technique for the case of 6 echoes.

TABLE I
MICRO-REFLECTION CHARACTERISTICS FOR MULTI-ECHO SCENARIO

<table>
<thead>
<tr>
<th>Echo #</th>
<th>Power</th>
<th>Delay in seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>-16 dBc</td>
<td>≤ 0.5 μs (≈ 51 T_s)</td>
</tr>
<tr>
<td>2nd</td>
<td>-22 dBc</td>
<td>≤ 1.0 μs (≈ 102 T_s)</td>
</tr>
<tr>
<td>3rd</td>
<td>-29 dBc</td>
<td>≤ 1.5 μs (≈ 154 T_s)</td>
</tr>
<tr>
<td>4th</td>
<td>-35 dBc</td>
<td>≤ 2.0 μs (≈ 205 T_s)</td>
</tr>
<tr>
<td>5th</td>
<td>-42 dBc</td>
<td>≤ 3.0 μs (≈ 307 T_s)</td>
</tr>
<tr>
<td>6th</td>
<td>-51 dBc</td>
<td>≤ 4.5 μs (≈ 461 T_s)</td>
</tr>
</tbody>
</table>

delays as specified in Table I. ϵ1 is constrained to be greater or equal to 1 and \( \hat{L} = L = 7 \). The simulation was run for two different skipping factors: \( K = 1 \), i.e., no sub-carrier skipping, and \( K = 4 \) which limits the number of pilot sub-carriers to \( M = 1900/4 = 475 \). The simulation used \( U = 2 \) and 20 iterations for each detection. As predicted, when \( \epsilon_1 > 1 \), performance of the ICE algorithm shows no apparent error floor. Furthermore, significant improvements are seen after the first few iterations, which is consistent with observations made from Fig. 7 and Fig. 8.

As shown Fig. 10, the performance for \( K = 4 \) is about 6 dB worse than for \( K = 1 \). This makes sense since the noise power \( \sigma^2_\rho = \sigma^2_w \) is inversely proportional to \( M \), so it increases by 6 dB when the number of pilot sub-carriers is reduced from 1900 to 475. Therefore, experimental evidence suggests that the performance of the ICE technique scales well with the number of pilot sub-carriers. Fig. 11 shows the MSE performance when the estimated number of channel paths, \( \hat{L} \), differs from the true number of paths, \( L \). The simulation parameters are the same as in Fig. 10 for the 7 paths channel described in Table I. The skipping factor is \( K = 1 \) and the SNR is varied from 10 to 40 dB. When SNR = 10 dB, the best performance is achieved when \( \hat{L} = 4 \). This is because the power of the 4th, 5th and 6th echoes are close to or less than the noise power, \( \sigma^2_\rho \), and thus can not be detected properly. A similar trend is observed when the SNR increases to 25 dB, in which case \( \hat{L} = 6 \) gives the best performance since only the last channel tap can not be detected due to noise.

The detection performance behaves differently in the high SNR region, i.e., when SNR ≥ 30 dB. In particular, the best performance is observed when \( \hat{L} \) is given the exact number of channel paths, e.g., when \( \hat{L} = 7 \). As expected, when \( \hat{L} > 7 \), under-detection significantly reduces channel estimation performance. However, only negligible performance degradation is observed with over-detection, e.g., when \( \hat{L} > 7 \), due to the threshold that mitigates the possibility of misidentifying noise samples as channel paths.

Finally, the raw bit error rate performance (i.e., without channel coding) of the proposed estimator under the 6-echo channel model is shown in Fig. 12. The parameters for the
4 designs considered are shown in the figure. The other simulation parameters are \( N = 4096, M = 3800 \) and \( U = 2 \). The modulation orders for data sub-carriers are 1024-QAM and 4096-QAM, which gives the highest throughput that DOCSIS 3.1 can achieve. Design 1 has the worst performance since such a large skipping factor leaves a large gap between pilot sub-carriers (i.e., 400kHz). Consequently it has problem with aliasing when the channel has a delay spread of 4.5\( \mu \)s. Performance of designs 2, 3 and 4 are quite similar to each other, and about 1 dB worse than the perfect channel estimator. Design 3 marginally outperforms designs 2 and 4, but may not be preferred in practice, as it requires both a large number of iterations and a large number of pilot subcarriers.

VI. CONCLUSION

The iterative channel estimation (ICE) algorithm proposed in this paper estimates the frequency response of a multipath channel by analyzing a single OFDMA symbol. Its performance in terms of variance of estimation error is markedly better than the conventional LS and DFT-based estimators, especially when the transport delays along the paths are fractional, i.e., the delays are not integer multiplies of the system sampling period. The performance of the ICE approaches that of the LMMSE, without needing apriori CSI.

In addition to estimating the frequency response of the channel, the ICE algorithm provides, at no cost, estimates of parameters that would otherwise be computed elsewhere in the demodulator. It provides a high precision estimate of the time of arrival of the main path, which is needed to trim the CP. It also provides estimates of echo strengths that can potentially be analyzed to identify the location of degraded or faulty components in the cable network.

REFERENCES


Ha H. Nguyen (M’01–SM’05) received the B.Eng. degree from the Hanoi University of Technology, Hanoi, Vietnam, in 1995, the M.Eng. degree from the Asian Institute of Technology, Bangkok, Thailand, in 1997, and the Ph.D. degree from the University of Manitoba, Winnipeg, MB, Canada, in 2001, all in electrical engineering. He joined the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK, Canada, in 2001, where he became a Full Professor in 2007. His research interests include diversity techniques in wireless communications, wireless relay networks, and error-control coding. He was an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2007 to 2011. He currently serves as an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and the IEEE WIRELESS COMMUNICATIONS LETTERS. He was the Technical Program Co-Chair for the Wireless Track, IEEE Vehicular Technology Conference, Vancouver, BC, Canada, in Fall 2014, the Multiple Antenna Systems and Space-Time Processing Track, IEEE Vehicular Technology Conferences, Ottawa, ON, Canada, and Quebec, QC, Canada, in Fall 2010 and Fall 2012, respectively, and the 14th Canadian Workshop on Information Theory, St. John’s, Canada, in 2015. He has co-authored with E. Shwedyk, a book A First Course in Digital Communications (Cambridge University Press). He is a Registered Member of the Association of Professional Engineers and Geoscientists of Saskatchewan.

J. Eric Salt (M’88) received the B.Sc degree in electrical engineering from the University of Saskatchewan, Saskatoon, SK, Canada, in 1974, the M.Eng. degree in electrical engineering from Carleton University, Ottawa, ON, Canada, in 1978, and the Ph.D. degree in electrical engineering from the University of Saskatchewan, in 1987. He joined the Department of Electrical Engineering and Computer Engineering, University of Saskatchewan, in 1985, where he currently holds the rank of Professor. His current research interests are in DSP algorithms and architectures for implementing CATV headend.