Channel Estimation in a DMT Based Power-Line Communication System Using Sparse Bayesian Regression

Ashraf A. Tahat [*], and Nikolaos. P. Galatsanos [†]

[*] School of Electrical Engineering, Princess Sumaya University for Technology, Amman, Jordan.
[†] Electrical and Computer Engineering Department, University of Patras, Rio, Greece.
tahat@psut.edu.jo, and ngalatsanos@upatras.gr

Abstract: - An enhanced power-line communications channel estimation method in discrete multi-tone (DMT) communication system based on sparse Bayesian regression is presented. By exploiting a probabilistic Bayesian learning framework, the sparse model used provides an accurate model for channel estimation in presence of noise and consequently equalization. We consider frequency domain equalization (FEQ) using the improved channel estimate at both the transmitter and receiver for a power-line system and compare the resulting bit error rate (BER) performance curves for both approaches and various channel estimation techniques. Simulation results show that the performance of the proposed method is superior to previous least squares based techniques.

Key-Words: - SBL, DMT, RVM, Channel estimation, Regression.

1. Introduction

Power-line communications (PLC) have received a lot of attention recently, because they represent efficient and economic solutions for both access and local area networks. PLC have been rapidly developing in the past few years as one of the most promising technologies to provide competitive techniques for numerous in-home communication applications that offer end-users a broad spectrum of services such as fast Internet access, voice-over-IP, multimedia applications, home automation and energy management. The key advantage of powerlines is that they provide a “pre-installed” infrastructure of wires and wall outlets that are easy to access throughout a building. However, a PLC channel is characterized with strong frequency selective and different types of noise [1], [2].

Multi-carrier (MC) based systems are popular for wireline and wireless communications due to their many advantages. At the present time, the MC transmission format as presented in [3] is more commonly known as Discrete Multi-tone (DMT) modulation or Orthogonal Frequency Division Multiplexing (OFDM). Contrary to analog MC systems, an FFT-based DMT implementation permits considerable overlap between subchannels, and results in high bandwidth efficiency. The use of DMT is well suited for PLC channels [4]. One of the advantages of DMT is its flexibility in loading bits onto the different tones taking into account the estimated signal-to-noise ratio (SNR). The measurement of SNR is usually performed during the initialization phase of the modem during which the frequency response of the channel is also estimated. To deliver high bit rates towards the user, DMT modems depend on advanced digital signal processing to mitigate several loop (channel) impairments such as: time dispersion, noise, echo and radio frequency interference (RFI). In this paper, we use Bayesian regression with a sparse model to improve channel estimation when the SNR is low. A model with similar philosophy has been used in machine learning applications also and is known as the relevance vector machine (RVM) [5]. Sparse Bayesian models when used for regression avoid fitting noise in the presented signal by *automatically setting to zero* the appropriate regression coefficients [5], [6]. This is important for channel equalization problems in the moderate-low SNR case [7]. In such cases the available signal contains the desired channel response embedded in noise. We demonstrate with numerical examples that approach improves the probability of error performance as compared with the traditional Least Squares (LS) approach.

2. System Model

Consider the system block diagram shown in *Figure 1*, which represents a Discrete Multi-Tone (DMT) link that includes: a DMT modulator, the channel model, DMT demodulator, and frequency domain equalizer.

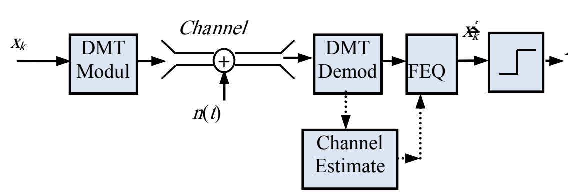

Fig. 1-a: DMT system block diagram with FEQ at receiver

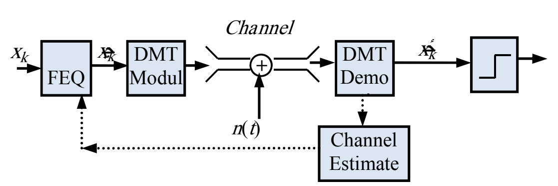

Fig. 1-b: DMT system block diagram with FEQ at transmitter

2.1 The DMT Modulator

DMT transmission [3] splits a high-rate data stream, *R*, into *N* lower rate streams transmitted simultaneously over subcarriers. The symbol rate for any of the *N* data streams is *R/N* symbols/second, which are sent to an 2*N*-point inverse fast Fourier transform (IFFT) block. After zero padding complex conjugate symmetry is created around the center of the IFFT. This converts the *N* frequency-domain complex in general data symbols into 2*N* time-domain real-valued samples as shown in *Figure 2*. A cyclic prefix (CP) of length *v* is pre-pended to the 2*N* point time domain samples to form the cyclically extended DMT symbol.

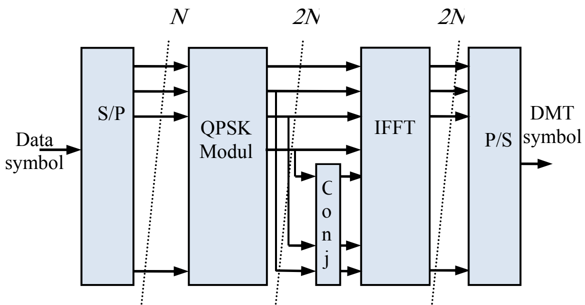

Fig. 2. Block diagram of the DMT modulator

This is similar to the configuration used for DMT-based Asymmetric Digital Subscriber Line (ADSL), but with fixed modulation on each subchannel (i.e., QPSK). The cyclic prefix length, *v*, is chosen to encompass the maximum *delay spread* of the channel to prevent intersymbol interference (ISI) and makes the DMT symbol appear periodic over the time span of interest.

2.2 The PLC Channel Model

The power-line network differs considerably in topology structure, and physical properties from conventional media such as twisted pair, coaxial, or fiber-optic cables. Therefore PLC channels have complex characteristics when high frequency signals are transmitted [2]. They have two undesired properties: The first is low pass behavior caused by attenuation that increases with line length and frequency of transmitted signal. The second is multipath propagation caused by impedance mismatches at the branches and unmatched line ends. In addition, their characteristics depend on variety of other factors such as network structure, type of the lines used, and the locations of the transmitter and receiver. Due to their complexity, a widely accepted model has not been known for PLC channel description. Only recently, models based on measurements were developed for signal transmission in the access domain [2], [8]. In this paper we will use the simplified PLC channel model described in [2] as:

$$H(f) = \sum_{i=1}^{M} g_i \, e^{-(a_0 + a_1 f^k) d_i} \, e^{-j 2\pi f \frac{d_i}{v_p}} \quad (1)$$

where M is the total number of paths in the model, with corresponding path weighting factors g_i, path lengths d_i and phase velocity v_p. Parameters a_0, a_1 and k describe the low pass characteristics of the channel. Channel parameters are obtained through measurements and are defined in [2], for the set of reference channels extracted on the basis of statistically representative scenarios. In the access domain, channels are defined for lengths of: 150m, 250m, and 350m, and quality classes labeled as “good”, “medium” and “bad”, depending on the attenuation and the depth of notches in the frequency band 25MHz. PLC channel noise can be modeled using colored Gaussian noise with higher power values at lower frequencies. The noise power spectral density is given by $N(f) = N_0 + N_1 \cdot e^{-f/f_0}$ [8], where some typical values for the parameters N_0, N_1, and f_0 are equal to -140dB/Hz, -38dB/Hz, and 0.7MHz, respectively. This PLC channel model is simple and suitable for simulation and analysis of PLC system performance.

2.3 Linear System Representation of Signal Transmission

When the channel impulse response (CIR) is smaller than or equal in length to the CP length, *v*, the 2*N* DMT sampled-time points (Baud) and the channel are circularly convolved. This enables straight forward frequency domain channel equalization at the receiver. However, when the order of the CIR is larger than the cyclic prefix, more complex equalization techniques are required [9]. In this paper, the PLC channel that we will consider has a CIR that does not exceed the CP. Hence, insertion of the CP at the transmitter and then discarding the first *v* samples at the receiver gives the received real-valued sampled-time domain samples over one DMT data block (baud) as:

$$\mathbf{r} = \mathbf{C}\mathbf{x} + \mathbf{n}\,, \quad (2)$$

where the vector $\mathbf{r} = [\, r_0, r_1, r_2, \ldots, r_{2N-1}]^T$ contains the received time domain samples, the vector $\mathbf{x} = [\, x_0, x_1, x_2, \ldots, x_{2N-1}]^T$ contains the transmitted time domain samples, and the vector $\mathbf{n} = [\, n_0, n_1, n_2, \ldots, n_{2N-1}]^T$ contains the samples of the band limited additive noise with average power (variance) σ_n^2. The matrix multiplication operator in (2) represents circular convolution without inter symbol interference (ISI). The circulant matrix $\mathbf{C} \in \Re^{2N\times 2N}$

$$\mathbf{C} = \begin{bmatrix} h_0 & 0 & \ldots & 0 & h_L & \cdots & h_1 \\ h_1 & h_0 & & & & & \vdots \\ \vdots & & \ddots & & & & h_L \\ h_L & & & & & & 0 \\ 0 & h_L & & & & & \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_L & \cdots & h_1 & h_0 \end{bmatrix}, \quad (3)$$

can be diagonalized by pre- and post-multiplication with the 2*N*-point discrete Fourier transform (DFT) matrix, $\mathbf{W}_{2N}$, and inverse DFT (IDFT) matrix $(\mathbf{W}_{2N})^{-1}$ [9] i.e.,

$$\mathbf{W}_{2N}\mathbf{C}\,\mathbf{W}_{2N}^{-1} = diag\{[H_0 \dots H_{2N-1}]\}\ , \qquad (4)$$

where [$H_0, H_1, H_2, \ldots, H_{2N-1}]^T$ is the frequency response of the channel.

When the transmitted symbols are generated in the frequency domain (i.e., $\mathbf{x} = \mathbf{W}_{2N}^{-1}$ [$X_0, X_1, X_2, \ldots, X_{2N-1}]^T$), and the received samples are converted to the frequency domain (i.e. , [$R_0, R_1, R_2, \ldots, R_{2N-1}]^T = \mathbf{W}_{2N}\,\mathbf{r}$), we obtain the very simple input-output relationship:

$$\begin{bmatrix} R_0 \\ \vdots \\ R_{2N-1} \end{bmatrix} = \begin{bmatrix} H_0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & H_{2N-1} \end{bmatrix} \begin{bmatrix} X_0 \\ \vdots \\ X_{2N-1} \end{bmatrix} + \begin{bmatrix} N_0 \\ \vdots \\ N_{2N-1} \end{bmatrix}, \qquad (5)$$

or, alternatively as:

$$\begin{bmatrix} R_0 \\ \vdots \\ R_{2N-1} \end{bmatrix} = \begin{bmatrix} X_0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_{2N-1} \end{bmatrix} \begin{bmatrix} H_0 \\ \vdots \\ H_{2N-1} \end{bmatrix} + \begin{bmatrix} N_0 \\ \vdots \\ N_{2N-1} \end{bmatrix}, \qquad (6)$$

where [$N_0, N_1, N_2, \ldots, N_{2N-1}]^T = \mathbf{W}_{2N}\ \mathbf{n}$. From (5) and (6), we can see that each received frequency domain symbol on each of the sub-channels is simply a scaled version of the transmitted frequency domain symbol plus colored Gaussian noise. Moreover, every sub-channel can be processed independently of the other sub-channels. In other words, block transmission with CP has converted a time dispersive channel into $2N$ parallel, narrowband flat sub-channels, or tones each having a channel gain H_k and additive Gaussian noise, hence the term discrete multi-tone transmission (DMT).

3. Channel Estimation

During startup of the DMT modem, the receiver measures the quality of the signals received (SNR) on each tone and reports this information to the transmitter [1], this is repeated periodically for dynamic channels. To estimate the channel during initialization, a pseudo-random frequency-domain training (pilot) sequence, $X_b(k)$, is employed. Where $\mathbf{X_b}$ is some constant modulus frequency-domain training sequence (i.e., $|X_b(k)|$ = constant for k = 0, 1, …, $2N$-1)). Here, without loss of generality, we normalize $\mathbf{X_b}$ so that $|X_b(k)| = 1$. Now, the received pilot symbols can be expressed in vector form as:

$$\mathbf{R}_b = \mathbf{X}_b\mathbf{H} + \mathbf{N}. \qquad (7)$$

In (7), $\mathbf{R}_b$ is the $2N \times 1$ received signal vector, the diagonal matrix $\mathbf{X}_b$ is the transmitted (pilot) signal, $\mathbf{N}$ is a vector containing the complex noise of the $2N$ subcarriers, and $\mathbf{H}$ is a vector that contains the overall complex channel gains between transmitter and receiver.

The problem at hand is to estimate the frequency-domain channel vector, H_k, (k = 0, 1, 2, 3, ···, N-1) as the FFT of L unknown sample-spaced time-domain tap gains, where L is chosen to encompass the maximum expected delay spread, and does not exceed the CP length, v. The frequency-domain model of the channel for each of the $2N$ tones is given by:

$$H_k = \sum_{n=0}^{L-1} h_n\ e^{-j2\pi kn/2N}\ , \qquad (8)$$

where h_n is channel tap at discrete time n.

To perform channel estimation from (7), we start by multiplying each of the frequency-domain symbols by the conjugate of the training symbols to produce the vector $\mathbf{T_b}$ as [7]:

$$\mathbf{T_b} = \mathbf{X_b^H}\mathbf{R}_b = \mathbf{X_b^H}\mathbf{X}_b\mathbf{H}\ + \mathbf{X_b^H}\mathbf{N}\quad , \quad (9)$$

where the superscript H, denotes conjugate-transpose. Because of the property of unitary magnitude of the training symbols, (9) can be expressed as:

$$\mathbf{T_b} = \mathbf{X_b^H}\mathbf{R_b} = \mathbf{H} + \mathbf{N}'. \qquad (10)$$

In the noise-free case, $\mathbf{T_b}$ in (10) will contain the perfect frequency-domain channel gains. Due to channel noise, $\mathbf{T_b}$ contains the channel estimates that are corrupted by the additive noise, $\mathbf{N}'$. This is equivalent to least squares (LS) estimation [11] in which no assumption about the channel impulse response length has been incorporated (i.e., h_n can have values larger than zero for n =0, 1, …, $2N$-1).

The LS estimator for the impulse response, $\mathbf{H}$, minimizes the criterion $(\mathbf{R_b}\text{-}\mathbf{X_b}\,\mathbf{H})^H\,(\mathbf{R_b}\text{-}\mathbf{X_b}\,\mathbf{H})$ and is given by:

$$\mathbf{H}_{LS} = \left(\mathbf{X}_b^H\mathbf{X}_b\right)^{-1}\mathbf{X}_b^H\mathbf{R}_b\ \ , \qquad (11)$$

where due to the unitary magnitude of the training symbols we have

$$\mathbf{H}_{LS} = \mathbf{X}_b^H\mathbf{R}_b = \mathbf{T}_b. \qquad (12)$$

When the values of the channel impulse response are forced to be zero for sample numbers larger than the cyclic prefix duration, v , then this will be equivalent to “modified LS” estimation [11].

Our objective is to improve the performance of the channel estimation at low to moderate SNR after truncation of the channel impulse response to v. Specifically, we will use a model similar to the one used for the relevance vector machine method (RVM), [5] and [6]. We apply this model to DMT channel estimation by first taking the inverse Fourier transform of (10) to fit a regression model as follows:

$$\mathbf{t_b} = \mathbf{h} + \mathbf{n}'\ \ , \qquad (13)$$

where $\mathbf{t_b}$ = [$t_0, t_1, t_2, \ldots, t_{2N-1}]^T$, is the target vector time domain samples, $\mathbf{h}$ = [$h_0, h_1, h_2, \ldots, h_v$, 0, …, 0 $]^T$ is the $2N \times 1$ channel vector, and the vector $\mathbf{n}'$ = [$n_0, n_1, n_2, \ldots, n_{2N-1}]^T$ contains samples of the band-limited additive Gaussian noise with variance σ'^2 .

In order to model $\mathbf{h}$ from the observations $\mathbf{t_b}$ in (13) and avoid the effects of the noise, we will use sparse Bayesian regression. This approach has the property of setting the appropriate regression weights to zero automatically, avoiding the fitting of noise in the signal $\mathbf{t_b}$.

Assume that the channel can be approximated using the function $y(n)$ which is the linear combination of kernel functions as:

$$y(n) = \sum_i w_i \varphi(n-i) \quad . \tag{14}$$

This convolution that can be written in matrix vector as:

$$\mathbf{y} = \mathbf{\Phi w} \quad , \tag{15}$$

where $\mathbf{w} = [\ w_0, w_1, w_2, \ldots, \ w_i\ , \ \ldots, w_v]^{\mathrm{T}}$ are the model weights, and $\mathbf{\Phi}$ is the $v \times (v+1)$ convolution (design) matrix that is created from the kernel we select. One commonly used kernel is the Gaussian shaped given by:

$$\phi(n-i) = \exp\left(\frac{-\|n-i\|^2}{2\sigma^2}\right) . \tag{16}$$

However, there are many other choices [6]. Now, our model becomes:

$$\mathbf{t_b} = \mathbf{y} + \mathbf{e} = \mathbf{\Phi w} + \mathbf{e} . \tag{17}$$

Where $\mathbf{y} = [\ y_0, \ y_1, \ y_2, \ \ldots, \ y_v]^{\mathrm{T}}$ is an approximation function and $\mathbf{e} = [\ e_0, e_1, e_2, \ldots, e_v]^{\mathrm{T}}$ is the vector with the regression model error.

The used Bayesian model assumes that the errors are modeled as independent identically distributed zero-mean Gaussian random variables, with variance σ^2 so:

$$p(\mathbf{e}) = \prod_{i=1}^{v} \mathcal{N}(e_i \mid 0, \sigma^2) \cdot \tag{18}$$

The parameter σ^2 can be set in advance if known, but can be also estimated from the data. This error model implies a multivariate Gaussian likelihood for the target vector $\mathbf{t_b}$:

$$p(\mathbf{t}_b \mid \mathbf{w}, \sigma^2) = (2\pi\sigma^2)^{-2N/2} \exp\left(\frac{-\|\mathbf{t}_b - \mathbf{\Phi w}\|^2}{2\sigma^2}\right) . \tag{19}$$

Maximum-likelihood estimation of $\mathbf{w}$ from (17) is equivalent to the Least Squares solution given by

$$\mathbf{w}_{LS} = \left(\mathbf{\Phi}^T \mathbf{\Phi}\right)^{-1} \mathbf{\Phi}^T \mathbf{t}_b \quad , \tag{20}$$

and leads to severe over-fitting of the data [6].

In order to overcome this problem a flexible Gaussian prior over the weights $\mathbf{w}$ and Bayesian inference will be used [6]. More specifically,

$$p(\mathbf{w}, \boldsymbol{\alpha}) = \prod_{i=1}^{v} \mathcal{N}(w_i \mid 0, \alpha_i^{-1}) \tag{21}$$

with $\boldsymbol{\alpha}$ a vector of $v+1$ hyperparameters [5], [6]. The flexibility of this prior is based on the fact that it uses a *separate hyperparameter* for every weight.

The weights $\mathbf{w}$ are marginalized in according to

$$p(\mathbf{t}_b; \mathbf{a}, \sigma^2) = \int p(\mathbf{t}_b / \mathbf{w}; \sigma^2) p(\mathbf{w}; \mathbf{a}) d\mathbf{w} . \tag{22}$$

This integral can be found in closed form and is given by:

$$p(\mathbf{t}_b \mid \boldsymbol{\alpha}, \sigma^2) = N(\mathbf{t}_b \mid \mathbf{0}, \mathbf{B}^{-1} + \mathbf{\Phi A}^{-1}\mathbf{\Phi}^{\mathrm{T}}), \tag{23}$$

where we have defined $\mathbf{A} = \mathrm{diag}(\ \alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_v\ )$, $\mathbf{B} = \sigma^2 \mathbf{I_v}$, and $\mathbf{I_v}$ is the $v \times v$ identity matrix. The pdf $p(\mathbf{t}_b \mid \boldsymbol{\alpha}, \sigma^2)$ is called the *marginal likelihood* (or *evidence*) for the hyperparameters $\boldsymbol{\alpha}$ and σ^2 [11]. Because the involved pdf's are Gaussian calculation of the *posterior* in closed form is also possible and is given by:

$$p(\mathbf{w} \mid \mathbf{t}_b, \boldsymbol{\alpha}, \sigma^2) = N(\mathbf{w} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}), \tag{24}$$

with

$$\boldsymbol{\Sigma} = (\mathbf{\Phi}^T \mathbf{B \Phi} + \mathbf{A})^{-1} \text{ and } \boldsymbol{\mu} = \boldsymbol{\Sigma} \mathbf{\Phi}^T \mathbf{B}\, \mathbf{t}_b . \tag{25}$$

The regression estimates obtained by Bayesian Inference are given by $\mathbf{\Phi}\boldsymbol{\mu}$ where the hyperparameters $\boldsymbol{\alpha}$ and σ^2 used in $\boldsymbol{\mu}$ are the *Maximum Likelihood* (ML) estimates given the observations $\mathbf{t}_b$ and are found by maximizing $p(\boldsymbol{\alpha}, \sigma^2 \mid \mathbf{t}_b)$ [5]. However, maximization of (23) with respect to $\boldsymbol{\alpha}$ and σ^2, which is termed as the *Type II ML* [6], [11], is equivalent to finding the maximum of $p(\boldsymbol{\alpha}, \sigma^2 \mid \mathbf{t}_b)$ assuming a uniform hyperprior [5], [6].

The values of $\boldsymbol{\alpha}$ and σ^2 which maximize (23) cannot be obtained in closed form. However, they can be found by stationary point analysis of (23) and yield the following iterative update equations [12]:

$$\alpha_i^{new} = \frac{\gamma_i}{\mu_i^2} \text{ and } (\sigma^2)^{new} = \frac{\|\mathbf{t} - \mathbf{\Phi}\boldsymbol{\mu}\|^2}{v - \sum_i \gamma_i} \tag{26}$$

where μ_i is the i^{th} posterior mean weight from (25) and the quantities γ_i are defined by: $\gamma_i = 1 - \alpha_i \Sigma_{ii}$ with Σ_{ii} the i^{th} diagonal element of the posterior weight covariance from (25) computed with the current $\boldsymbol{\alpha}$ and σ^2 values.

The learning algorithm thus proceeds by repeated application of (25) and (26) until some convergence criteria have been satisfied [5], [13]. It can be shown that this iteration is equivalent to an expectation-maximization (EM) algorithm [6]. The E-step is (25) and the M-step is (26). This relation guarantees convergence of the proposed algorithm [6].

An important property of this model is that typically the optimal value of most hyperparameters α_i is *infinite* [5], [13]. From (25) this implies that the corresponding μ_i is zero. Thus, the regressor $y(n) = \sum_{i=1}^{v} \mu_i \phi(n-i)$ is *sparse* since many its weights μ_i are zero. In matrix from, the RVM channel estimate is expressed as:

$$h_{RVM} = \mathbf{\Phi\mu} \tag{27}$$

4. Frequency-Domain Channel Equalization

If the channel is known for the noise free case, we can perform the equalization by inverting (2) assuming **C** is non-singular. The eigenvalues of the $2N \times 2N$ circulant matrix are equal to the DFT coefficients of the first column [14]. Since the first column has the channel coefficients in ascending order, from (7) these eigenvalues are:

$$H_k = \sum_{n=0}^{L-1} h_n \, e^{-j2\pi kn/2N} = \sum_{n=0}^{L-1} h_n \, W^{-kn} \,, \tag{28}$$

where $W = e^{\,j2\pi/2N}$, and H_k represents the channel frequency response. Note that $h_n = 0$ for $L < n < 2N$.
The circulant matrix **C** can be diagonalized with the DFT matrix [13] as was mentioned in (4) as :

$$\mathbf{W}_{2N} C \, \mathbf{W}_{2N}^{-1} = diag\{[H_0 \ldots H_{2N-1}]\} = \Lambda_C \,. \tag{29}$$

The diagonal elements of $[\Lambda_C]^{-1}$ are $1/\ H_k$, can be regarded as a set of DFT-domain equalizers. Of course, in practice channel noise can be amplified by $[\Lambda_C]^{-1}$ if the coefficients, $1/\ H_k$, are large and frequency equalization is carried out at the receiver as in *Figure 1*-a, and in this version, the receiver has all the complexity. The symbols are equalized in the frequency domain before detection as:

$$\hat{\mathbf{X}} = \mathbf{\Lambda}_C^{-1} \, \mathbf{R} \quad , \tag{30}$$

where $\hat{\mathbf{X}}$ is a $2N \times 1$ vector, and **R** is a $2N \times 1$ received signal vector .
If the channel is known, we can move $[\Lambda_C]^{-1}$ to the transmitter [15], yielding a useful configuration for the cases where the receiver is simpler. In addition, channel noise will not be amplified for this second version. This also results in superior performance as will be demonstrated in the simulation section. The symbols are equalized in the frequency domain before modulation and transmission as:

$$\breve{\mathbf{X}} = \mathbf{\Lambda}_C^{-1} \, \mathbf{X} \tag{31}$$

where $\breve{\mathbf{X}}$ is a $2N \times 1$ symbols vector.

5. Simulation Results

MATLAB® simulations are used to evaluate the performance of the DMT/RVM system. Following are the simulation parameters: $2N = 512$ the DMT FFT size, $v = 64$ is the Cyclic Prefix length. The used sub channels consisted of tones 6 to 255, where each of the subcarriers uses a QPSK modulation scheme. The additive channel noise is band limited colored Gaussian with power σ_n^2. The actual power-line channel length is $L = 43$. It was generated based on the model of equation (1) using parameters of a test network with physical length of 212 meters and four path models from [2] as follows: k=1, a_0=0 , a_1=7.8 × 10^{-10} s/m, and path parameters g_i= {0.64, 0.38, -0.15, 0.05} and corresponding d_i/m={200, 222.4, 244.8, 267.5}.

In *Figure 3*, we present an example target vector time domain samples of equation (13), which in the noise-free case should be the actual impulse response of the channel and is superimposed on the same plot

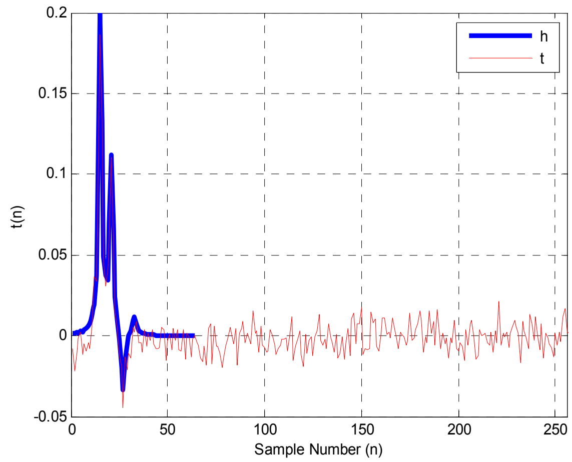

Fig. 3. An example target vector

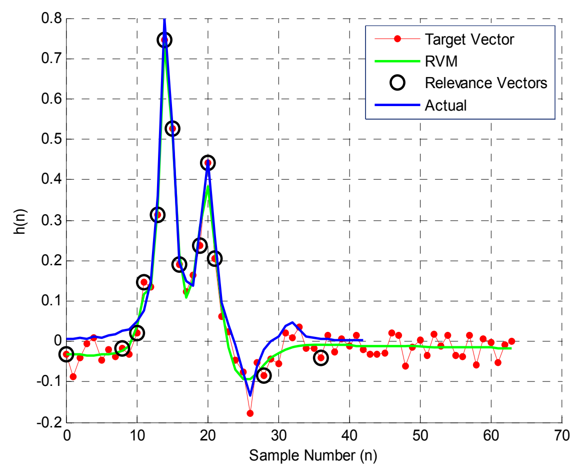

Fig. 4. Channel impulse response and its estimate obtained using the RVM method

The channels were assumed to remain constant after each sounding using the pilot sequence. *Figure 4* depicts the actual sampled-time domain channel in the cyclic prefix range where it is equal to zero for values of time larger than the cyclic prefix. The target vector is also plotted in the same time range which is equivalent to the Least Squares (LS) estimate of the channel. The proposed Bayesian sparse regression algorithm is used to process the target vector samples to filter out the noise at an SNR of 5 dB. The locations of the non-zero regression weights are identified on the plot with circles. Notice that we use *only 13 non-zero weights* from a total of 65. The remaining 52 are set to zero by the sparse model.
To demonstrate the effects of estimating the channel impulse response on the frequency attenuation of each subcarrier (and hence equalization), in *Figure 5* we plot

the frequency response of the estimated channel, the actual channel, and the original noise corrupted channel. From *Figure 5* we observe that in the frequency domain the estimated channel by the sparse model has smoother variation between adjacent tones and is closer to the desired frequency response.

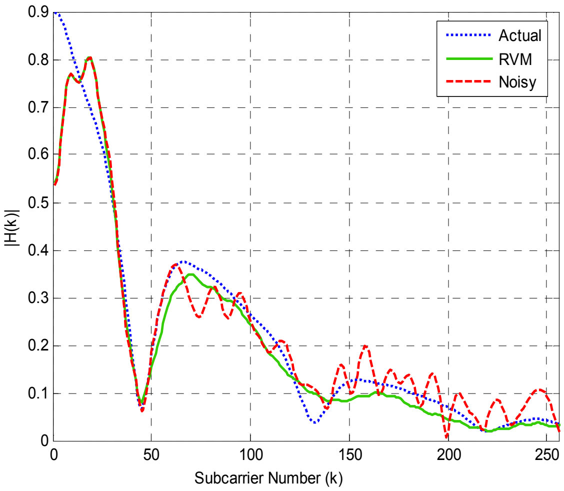

Fig. 5. Frequency response of actual and estimated channel.

When FEQ is implemented at the transmitter using the estimated channel at the receiver, BER performance curves (labeled TX-) demonstrate in *Figure 6* the superior performance attained as a result of the more accurate sparse regression channel (RVM) estimate over the LS one.

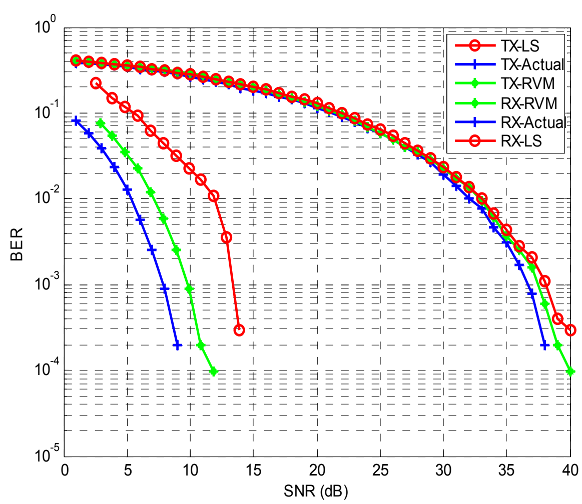

Fig. 6. BER curves with equalization at transmitter and receiver.

Also, the BER performance curve based on the actual channel (perfect estimate) is superimposed on the same plot of *Figure 6* as a performance bench mark. We also compare in *Figure 6* the accuracy in channel estimation when FEQ is implemented at the receiver as compared to that at the transmitter. We can make two observations: First, it is evident that FEQ at the transmitter is far superior to that at the receiver for the same channel estimate. Second, the improved accuracy in channel estimation has limited effect on BER performance when FEQ is implemented at the receiver (labeled RX-), and BER curves are virtually identical for the actual channel, sparse regression channel estimate, and the LS channel estimate. Low SNR of the attenuated subcarriers during the equalization process renders the improved channel accuracy not important in this case.

6. Conclusion

In order to attain full use of available bandwidth in a power-line DMT data transmission system, channel equalization is a prerequisite. In this paper, an enhanced channel estimation method in DMT communication systems based on sparse Bayesian regression via the RVM method was presented. This method provided improved accuracy in channel estimation for frequency equalization. Frequency domain equalization at the transmitter is more sensitive to the accuracy of channel estimation than at the receiver side and yields better BER performance curves. The performance of the proposed sparse regression method is superior to the traditional least squares technique.

7. References

[1] H. Hrasinca, A. Haidine, and R. Lehnert, *Broadband Powerline Communications Networks*. West Sussex, England: Wiley & Sons, 2004.

[2] M. Zimmermann, K. Dostert, "A multipath model for the powerline channel," *IEEE Transactions on Communications*, vol. 50, pp. 553-559, April 2002.

[3] S. Weinstein, and P. Ebert, " Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Transactions on Communications Technology*, vol. 19, pp. 628–634, October 1971.

[4] HomePlug Powerline Alliance, "HomePlug AV White Paper, " 2005

[5] M. Tipping, "Sparse Bayesian and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211-244, September 2001.

[6] C. Bishop, *Pattern Recognition and Machine Learning*. Secaucus, NJ: Springer-Verlag, 2006.

[7] A. Tahat, and N. Galatsanos, "Relevenace Vector Machines for Enhanced BER in DMT Based Systems," *Journal of Electrical and Computer Engineering*, vol. 2010, pp. 1-8, July 2010.

[8] Theoretical postulations of PLC channel model, D4 OPERA project delivery, http://www.ist-opera.org/opera1, April 2005.

[9] T. Pollet, M. Peeters, M. Moonen, and L. Vandendorpe, "Equalization for DMT-based broadband modems," *IEEE Communications Mag*., vol. 38, pp. 106–113, May 2000.

[10] G. Golub and C. Van Loan, *Matrix Computations* (3rd edition). Baltimore, MD: The John Hopkins University Press, 1996.

[11] J. de Beek, et. al., "On channel estimation in OFDM systems," in *Proc. IEEE Vehicular Technology Conference*, Chicago, USA, Sept. 1995. pp. 815 - 819.

[12] J. Berger, *Statistical Decision Theory and Bayesian Analysis*. New York, NY: Springer, 1985.

[13] A. Faul, M. Tipping, "Analysis of sparse Bayesian learning," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2001, pp. 383-389.

[14] A. Papoulis, *Signal Analysis*. New York, NY: McGraw Hill, 1977.

[15] D. Love *et. al.*, "An overview of limited feedback in wireless communication systems," *IEEE J.Sselected areas in Communications*, vol. 26 , pp. 1341–1366, October 2008.