TECHNICAL RESEARCH REPORT

Maximum Likelihood Slow Frequency-Selective Fading Channel Estimation Using Frequency Domain Approach

by Y. Jiang and John S. Baras

CSHCN T.R. 2000-13
(ISR T.R. 2000-35)

Maximum Likelihood Slow Frequency-Selective Fading Channel Estimation Using Frequency Domain Approach

Yimin Jiang   Feng-Wen Sun *   John S. Baras†

Abstract

This paper addresses the channel estimation problem for slow frequency-selective fading channel using training sequence and maximum likelihood (ML) approach. Traditional works assumed symbol period spaced delay-tapped line model and additive white Gaussian noise (AWGN). Because of pre-filtering in the receiver front end, if the sampling rate is larger than one sample per symbol or sampling epoch is unknown (i.e., timing information is not available), AWGN model is not valid anymore. A more general ML channel estimation method using discrete Fourier transform (DFT) is derived with the assumption of colored Gaussian noise and over sampling. Similar idea can be adopted to derive the ML joint timing and phase estimation algorithm.

Subject Area: Communication Theory, Contact Author: Yimin Jiang

*Yimin Jiang and Feng-Wen Sun are with Hughes Network Systems Inc. 11717 Exploration Lane, Germantown, MD 20874 USA, E-mails: {yjiang, fsun }@hns.com, Tel: 301-601-6494, Fax: 301-428-7177

†John S. Baras is with the Institute for Systems Research, University of Maryland, College Park, MD 20742, USA, E-mail: baras@isr.umd.edu

1 Introduction

For burst-transmission digital communication systems, channel estimation is required for maximum-likelihood sequence estimation receivers [1] [2]. A typical data burst consists of several blocks of user data and a predetermined training sequence (TS) which is used to estimate the channel impulse response. Channel estimation can be done using a Wiener filter or the discrete Fourier transform (DFT). For example, [3] - [5] consider channel estimation given a known training sequence. The authors of [3] addressed the problem of selecting the optimum training sequence for channel estimation by processing in the frequency domain. Optimum unbiased channel estimation given white noise is considered in [4] following a maximum-likelihood approach. Following the least-squares (LS) philosophy, [5] presents algorithms for optimal unbiased channel estimation with aperiodic spread spectrum signals for white or nonwhite noise.

Previous works [3] - [5] assumed symbol period delay-tapped line model or AWGN noise [3]. Because of pre-filtering in the receiver front end, this model is not accurate enough and will cause aliasing or leakage on the spectrum. Since typical pulse shaping rolloff factors in wireless communication range between 0.2 and 0.7, a sampling frequency larger than one symbol rate is required to prevent aliasing. Typically a nominal sampling rate of two samples per symbol period is used in wireless receiver [1]. When sampling rate is higher than one sample per symbol or timing information is unknown, AWGN model is not valid. Therefore a more general model is desired to accommodate colored Gaussian noise and a higher sampling rate. Felhauer proposed a whitening matched filter approach in [5] to deal with colored noise, which actually follows the general idea in Van Trees' classical work in [6]. In this paper, we will show that a direct optimum estimator can be derived without preliminary processing [6] (p.289).

This paper takes a ML approach and derives an optimal channel estimation algorithm in the frequency domain. The auto-covariance matrix of a colored Gaussian noise is a Toeplitz matrix. Toeplitz matrix was thoroughly studied in [7]. In their book, Grenander and Szego showed that a Toeplitz matrix converged to a circular matrix in the "weak sense". More engineering descriptions can be found in [8] [9]. It is well known that the inverse of a Toeplitz matrix is not Toeplitz generally. Kobayashi showed that the inverse of a Toeplitz matrix was asymptotically Toeplitz [11], similar methodology was adopted in Meyr's book [1]. However we found that it was not generally true and there was some condition to apply this idea in our work [12]-[13]. The condition is that there should be no zeros of the Z transform of the function that defines a Toeplitz matrix on unit circle. In another word, the discrete time Fourier transform (DTFT) of this function has no zeros within the frequency band that is smaller than the sampling rate. If the above condition satisfies, the inverse of a Toeplitz matrix converges to a circular matrix in the "finite boundary strong sense" (refer to [12]).

It is well understood that the unitary matrix used in the eigendecomposition of a circular matrix is a DFT matrix which will be defined later. We adopt the eigendecomposition of the inverse of auto-covariance matrix, the DFT approximation based our research result, and the time shift property of DFT here. This paper is organized as follows: Section 2 describes our channel model and derives the likelihood function. The ML channel estimator is presented in Section 3. Section 4 addresses one special case of this idea, the ML joint carrier phase and timing offset estimator.

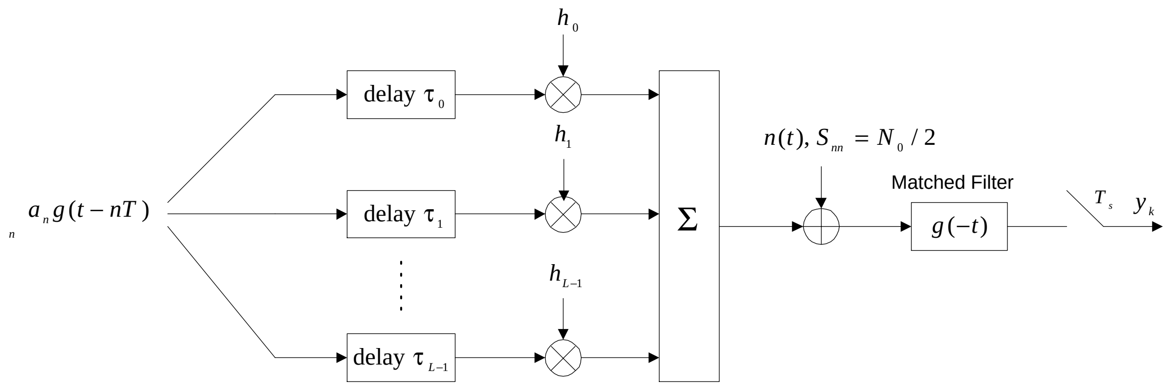

Figure 1: Modeling of Slow Frequency-Selective Fading Channel and Matched Filter

Some computer simulation results are shown in Section 4 too.

2 Problem Formulation

Without limitations on the number of paths and delay of each path in our problem, the following channel model is assumed:

$$h(t) = \sum_{l=0}^{L-1} h_l \delta(t - \tau_l T) \tag{1}$$

where L is the total number of paths, which is unknown, h_l and τ_l are the attenuation and delay factor of path l respectively, T is the symbol period. In our model, we assume slow and frequency selective fading channel, i.e., h_l and τ_l remain constant within the observation window, τ_l is comparable with symbol period. The baseband received signal is modeled as the following:

$$x(t) = \sqrt{E_s} \sum_{l-0}^{L-1} \sum_{n=-N/2}^{N/2-1} h_l a_n g(t - nT - \tau_l T) + n(t) \tag{2}$$

where $\{a_n\}$ is the training sequence, $n(t)$ is the AWGN noise with two-sided power spectrum density (PSD) $N_0/2$. The received signal $x(t)$ is passed through a matched filter with response $g(-t)$, then sampled at the rate $1/T_s$ with $T_s = T/M$ (M is the sampling rate in samples per symbol). The output of the matched filter is defined as $y(t)$ that is given by:

$$y(t) = \sqrt{E_s} \sum_{l=0}^{L-1} \sum_{n=-N/2}^{N/2-1} h_l a_n r(g - nT - \tau_l T) + N(t) \tag{3}$$

where $r(t) = g(t) \otimes g(-t)$, $N(t) = n(t) \otimes g(-t)$.

The likelihood function of $\{h_l, \tau_l\}$ is the pdf of a Gaussian r.v. The mean of y_k given $\{h_l, \tau_l\}$ is

$$m_y(k) = \sqrt{E_s} \sum_{l=0}^{L-1} \sum_{n=-N/2}^{N/2-1} h_l a_n r(kT_s - nT - \tau_l T) + N(kT_s), \quad k \in [-K/2, K/2-1] \tag{4}$$

with $K = M(N + R)$ is the total number of digital samples, N is used to model the central portion of the TS, R is used to model the remaining observation including the shaping pulse tail. It is reasonable to assume that $m_y(k) = 0$ when $|k| > \mathcal{K}$ in engineering problems. The physical explanation is that all the information of $\underline{m}_y$ is included. The auto-covariance matrix of vector $\underline{y}$ is

$$\mathrm{cov}[\underline{y}|\underline{h}, \underline{\tau}] = \frac{N_0}{2}\Lambda \tag{5}$$

where Λ is a K by K Hermitian and Toeplitz matrix defined as

$$\Lambda = \begin{bmatrix} r(0T_s) & r(-T_s) & \cdots & r(-(K-1)T_s) \\ r(T_s) & r(0T_s) & \cdots & r(-(K-2)T_s) \\ \vdots & \vdots & \ddots & \vdots \\ r((K-1)T_s) & r((K-2)T_s) & \cdots & r(0T_s) \end{bmatrix} \tag{6}$$

The likelihood function of $\underline{y}$ given $\underline{a}, \underline{h}, \underline{\tau}$ is

$$f(\underline{y}|\underline{h}, \underline{\tau}) = \frac{\exp\left\{-\frac{1}{2}(\underline{y} - \underline{m}_y)^H \left(\frac{N_o}{2}\Lambda\right)^{-1} (\underline{y} - \underline{m}_y)\right\}}{(2\pi)^{K/2}|\frac{N_o}{2}\Lambda|^{1/2}} \tag{7}$$

The log likelihood function is given by

$$\begin{aligned} l(\underline{y}|\underline{h}, \underline{\tau}) &= \log(f(\underline{y}|\underline{h}, \underline{\tau})) \\ &= -\frac{1}{N_o}\left[-\underline{y}^H Q \underline{m}_y - \underline{m}_y^H Q \underline{y} + \underline{m}_y^H Q \underline{m}_y\right] \\ &\quad - \left(\frac{1}{N_o}\underline{y}^H Q \underline{y} + \log\left[(2\pi)^{K/2} \left|\frac{N_o}{2}\Lambda\right|^{1/2}\right]\right) \end{aligned} \tag{8}$$

where Q is the inverse of Λ.

3 ML Channel Estimator in Frequency Domain

Before going through the derivation of the channel estimator, we review some mathematical results based on our research.

3.1 On the Toeplitz Matrices

A family of Toeplitz matrices T_n (n is the dimension of the matrix) are defined by a sequence of complex numbers

$$t_i; \{i = \cdots, -1, 0, 1, \cdots\}$$

such that the elements of T_n at the ith row and jth column is equal to t_{i-j}, i.e.,

$$T_n = \{t_{i-j}\} \tag{9}$$

Furthermore, we restrict our discussion to the case that $t_{-i} = t_i^*$, the conjugate of t_i. With this restriction, T_n becomes Hermitian. Toeplitz matrix in this form plays pivotal role in many signal processing issues. Actually, often what is more relevant is the inverse of such matrix rather than the matrix itself. For instance, if t_i represents the correlation of a stationary random process (in our case Λ). The inverse of T_n (in our case Q) is associated with the joint probability function of n consecutive samples of the random process.

One of the difficulties in analyzing the inverse matrix arises from the fact that the inverse of a Toeplitz is no longer Toeplitz, though it was shown in [10] that such inverse can be decomposed into multiplication and summation of Toeplitz matrices. One widely used technique to tackle the problem is to substitute the Toeplitz matrix T_n with a circular matrix. Such circular matrix can be defined by the Fourier Transform of sequence t_n. Let $\mathcal{F}(\lambda)$ denote the Fourier Transform of t_n, i.e.,

$$\mathcal{F}(\lambda) = \sum_{k=-\infty}^{\infty} t_k e^{-j\lambda k}.$$

Let U_n denote the unitary matrix defined as

$$U_n = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{-j(2\pi/n)} & \cdots & e^{-j(2\pi(n-1)/n)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j(2\pi(n-1)/n)} & \cdots & e^{-j(2\pi(n-1)(n-1)/n)} \end{bmatrix} \tag{10}$$

and D_n denote the diagonal matrix with the ith diagonal element $\mu_{i,n} = \mathcal{F}(2\pi(i-1)/n)$, i.e.,

$$D_n = \begin{bmatrix} \mu_{1,n} & 0 & \cdots & 0 \\ 0 & \mu_{2,n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_{n,n} \end{bmatrix}. \tag{11}$$

Then the circular matrix $U_n^H D_n U_n$ is defined as C_n. Substituting T_n with C_n is based on the well-known fact that T_n converges to C_n in the weak sense as long as $|\mathcal{F}(\lambda)|$ is bounded*.

In many applications, the quadratic form $x^H T_n^{-1} x$ can be limited to the case that x has only finite nonzero terms in the middle of the vector, i.e.

$$x = (0, \cdots, 0, x_{-k}, x_{-k+1}, \cdots, x_0, \cdots, x_k, 0, \cdots, 0) \tag{12}$$

and k does not increase with n. We shall call this finite boundary quadratic form.

Definition: For two families of Hermitian matrices A_n, B_n, consider the quadratic form

$$\max_x \left| \frac{x^H (A_n - B_n) x}{x^H x} \right|, \tag{13}$$

where the maximum is over all the n-dimensional vector of the form (12). If (13) converges to zero for any given k, we shall call A_n converges to B_n in the finite boundary strong sense convergence.

If x corresponds to an observation within the window $[-k, k]$ and with negligible leakage outside the observation window, we are able to replace A_n with B_n asymptotically in evaluating the quadratic forms. Many practical applications fall into this category.

*In [7], C_n is defined through the inverse Fourier Transform, i.e., the ith diagonal element of D_n is equal to $\mathcal{F}(-2\pi i/n)$ and U_n is replaced by U_n^T. The current notation is more consistent with engineering conventions.

Theorem: Let T_n be a family of Hermitian Toeplitz matrix associated with the sequence $\{t_n\}$ of finite order, i.e., $t_s = 0$ for $|s| > W$ ([8], p.23), and $\mathcal{F}(z)$ be the z-transform of $\{t_n\}$, that is

$$\mathcal{F}(z) = \sum_{k=-\infty}^{\infty} t_k z^{-k}.$$

If $|\mathcal{F}(z)|$ does not have any zero on the unit circle, T_n^{-1} converges to C_n^{-1} in the strong sense for finite boundary quadratic form.

Refer to [12] for proof details.

3.2 The ML Channel Estimator

Following the finite boundary strong sense convergence theorem, we exam if our problem fits the condition to apply this theorem. First, if $r(t)$ is the raised cosine shaping pulse, the DFTF of $r(kT_s)$ has no cross-zero points in its passband. One question comes out naturally, if $r(t)$ is over-sampled, $F_r(z)$ will have zeros on the unit circle, does this theorem still apply? The answer is positive. Actually it is a classical question related to building the likelihood function when over-sampling is applied. If $r(t)$ is over-sampled, as n large, T_n is singular, which means some rows are the linear combination of other rows (through interpolation) and T_n^{-1} does not exist at all. To prevent this situation from happening, the technique mentioned in [6] (p.289) can be applied to get around this problem and guarantee that our operation will be meaningful. Second, $\underline{m}_y$ that is the mean of $\underline{y}$ satisfies the finite boundary condition in our model. The physical explanation of large K is that in ideal case the observation of $\underline{y}$ includes both training portion and the tail of all the response.

We can use C^{-1} to approximate Q after carefully checking the convergence condition, then we get,

$$Q \sim C^{-1} = U_K^H D^{-1} U_K \tag{14}$$

as K large enough, which means $N+R$ large enough

$$\mu_{n,K} \rightarrow \mathcal{F}_r[k] \tag{15}$$

where $\mathcal{F}_r$ is the power spectral density (PSD) of the colored noise process $\{N_k\}$, which is expressed

$$\mathcal{F}_r[k] = \frac{1}{T_s} \sum_{l=-\infty}^{\infty} \mathcal{R}\left(\frac{k}{KT_s} - \frac{l}{T_s}\right) \tag{16}$$

The ML estimate of channel response $\{\underline{h}, \underline{\tau}\}$ is

$$(\underline{h}, \underline{\tau}) = \arg\max_{\underline{h},\underline{\tau}} l(\underline{y}|\underline{h}, \underline{\tau}) \tag{17}$$

Because all the information related to channel response is in $\underline{m}_y$, therefore

$$(\underline{h}, \underline{\tau}) = \arg\max_{\underline{h},\underline{\tau}} \left\{ -\frac{1}{N_o} \left[-\underline{y}^H Q \underline{m}_y - \underline{m}_y^H Q \underline{y} + \underline{m}_y^H Q \underline{m}_y \right] \right\} \tag{18}$$

After some arithmetic,

$$U_K^H \underline{m}_y = \sqrt{\frac{E_s}{K}} \Pi \tag{19}$$

where Π is a K by 1 vector with mth element equal to (as K large enough, and sampling rate satisfies Nyquist sampling theroem)

$$
\begin{aligned}
\pi[m] &= \sum_{k=-K/2}^{K/2-1} \sum_{l=0}^{L-1} \sum_{n=-N/2}^{N/2-1} h_l a_n r(kT_s - nT - \tau_l T) e^{-j(2\pi mk/K)} \\
&\approx \sum_{l=0}^{L-1} \sum_{n=-N/2}^{N/2-1} h_l a_n \mathcal{F}_r[m] e^{-j(2\pi m(n+\tau_l)/(N+R))} \\
&= \sum_{l=0}^{L-1} \mathcal{F}_r[m] \mathcal{A}[m] h_l e^{-j(2\pi m\tau_l/(N+R))} \\
&= \mathcal{F}_r[m] \mathcal{A}[m] H[m]
\end{aligned} \tag{20}
$$

where $\mathcal{A}[m]$ ($\mathcal{A}[m] = \sum_{n=-N/2}^{N/2-1} a_n e^{-j(2\pi mn/(N+R))}$) is $N+R$ point DFT of training sequence $\{a_n\}$, $H[m] = \sum_{l=0}^{L-1} h_l e^{-j(2\pi m\tau_l/(N+R))}$ that is the DFT of sampled channel response.

Similarly, we can get

$$
\begin{aligned}
\underline{y}^H Q \underline{m}_y &= \frac{\sqrt{E_s}}{K} \sum_{m=-K/2}^{K/2-1} \frac{\mathcal{F}_y[m]^* \mathcal{F}_r[m] \mathcal{A}[m] H[m]}{\mathcal{F}_r[m]} \\
&= \frac{\sqrt{E_s}}{K} \sum_{m=-K/2}^{K/2-1} \mathcal{F}_y[m]^* \mathcal{A}[m] H[m]
\end{aligned} \tag{21}
$$

where $\mathcal{F}_y[m]$ is the K-point DFT of $\underline{y}$, i.e.,

$$
\mathcal{F}_y[m] = \sum_{k=-K/2}^{K/2-1} y(kT_s) e^{-j(2\pi mk/K)} \tag{22}
$$

Also we can get

$$
\underline{m}_y^H Q \underline{m}_y = \frac{E_s}{K} \sum_{m=-K/2}^{K/2-1} \mathcal{F}_r[m] |\mathcal{A}[m]|^2 |H[m]|^2 \tag{23}
$$

Therefore, the ML estimate of channel response is given by

$$(\underline{h}, \underline{\tau}) = \arg \max_{\underline{h},\underline{\tau}} \{ \frac{\sqrt{E_s}}{N_0 K} \sum_{m=-K/2}^{K/2-1} [\mathcal{F}_y[m]\mathcal{A}[m]^* H[m]^* + \mathcal{F}_y[m]^* \mathcal{A}[m] H[m] \\ - \sqrt{E_s}\mathcal{F}_r[m]|\mathcal{A}[m]|^2 |H[m]|^2]\} \tag{24}$$

The content in the parenthesis on the RHS of eq. (24) can be reorganized as

$$\begin{aligned} RHS &= \sum_{m=-K/2}^{K/2-1} [\mathcal{F}_y[m]\mathcal{A}[m]^* H[m]^* + \mathcal{F}_y[m]^* \mathcal{A}[m] H[m] \\ &\quad - \sqrt{E_s}\mathcal{F}_r[m]|\mathcal{A}[m]|^2 |H[m]|^2 - \frac{|\mathcal{F}_y[m]|^2}{\sqrt{E_s}\mathcal{F}_r[m]} + \frac{|\mathcal{F}_y[m]|^2}{\sqrt{E_s}\mathcal{F}_r[m]}] \\ &= \sum_{m=-K/2}^{K/2-1} \left[\frac{|\mathcal{F}_y[m]|^2}{\sqrt{E_s}\mathcal{F}_r[m]} - \left(E_s^{1/4}\sqrt{\mathcal{F}_r[m]}\mathcal{A}[m]H[m] - \frac{\mathcal{F}_y[m]}{E_s^{1/4}\sqrt{\mathcal{F}_r[m]}} \right) \cdot \right. \\ &\quad \left. \left(E_s^{1/4}\sqrt{\mathcal{F}_r[m]}\mathcal{A}[m]^* H[m]^* - \frac{\mathcal{F}_y[m]^*}{E_s^{1/4}\sqrt{\mathcal{F}_r[m]}} \right) \right] \\ &= \sum_{m=-K/2}^{K/2-1} \left[\frac{|\mathcal{F}_y[m]|^2}{\sqrt{E_s}\mathcal{F}_r[m]} - \left| E_s^{1/4}\sqrt{\mathcal{F}_r[m]}\mathcal{A}[m]H[m] - \frac{\mathcal{F}_y[m]}{E_s^{1/4}\sqrt{\mathcal{F}_r[m]}} \right|^2 \right] \end{aligned} \tag{25}$$

Because $\mathcal{F}_y[m]$ and $\mathcal{F}_r[m]$ do not contain $\{\underline{h}, \underline{\tau}\}$, therefore, the ML estimate of channel response becomes

$$\begin{aligned} (\underline{h}, \underline{\tau}) &= \arg \min_{\underline{h},\underline{\tau}} \sum_{m=-K/2}^{K/2-1} \left| E_s^{1/4}\sqrt{\mathcal{F}_r[m]}\mathcal{A}[m]H[m] - \frac{\mathcal{F}_y[m]}{E_s^{1/4}\sqrt{\mathcal{F}_r[m]}} \right|^2 \\ &= \arg \min_{\underline{h},\underline{\tau}} \sum_{m=-K/2}^{K/2-1} \left| \frac{\sqrt{E_s}\mathcal{F}_r[m]\mathcal{A}[m]H[m] - \mathcal{F}_y[m]}{E_s^{1/4}\sqrt{\mathcal{F}_r[m]}} \right|^2 \end{aligned} \tag{26}$$

Thing becomes clear now, the ML estimate of $\{\underline{h}, \underline{\tau}\}$ is the $\{\underline{h}, \underline{\tau}\}$ that has the following DFT

$$H[m] = \begin{cases} \mathcal{F}_y[m]/(\sqrt{E_s}\mathcal{F}_r[m]\mathcal{A}[m]) & \text{if } \mathcal{F}_r[m] \neq 0, \\ 0 & \text{if } \mathcal{F}_r[m] = 0. \end{cases} \tag{27}$$

Remarks:

- Clearly, we are only interested in the channel response within the pass-band of shaping function, we just need to calculate those when $\mathcal{F}_r[m] \neq 0$, (27) follows.

- If there is no noise, i.e., $N_0 = 0$, it is straightforward to verify that the real channel response $H[m]$ is exactly equal to (27). The channel estimator (27) is unbiased.

- Because $H[m]$ is just the DTFT of $\underline{h}$ and $\underline{\tau}$, there are a lot of possible $\underline{h}$ and $\underline{\tau}$ that has the same $H[m]$. If the time domain response is of more interest, $H[m]$ can be treated as an intermediate result. With the help of some physical modeling on $\underline{h}$ and $\underline{\tau}$, we can get the time domain response from $H[m]$.

- According to the finite boundary strong sense convergence theorem, when $K \to \infty$, the inverse matrix Q approaches to a circular matrix. How to handle it in reality? We have two observations based on (27): first, if the sampling rate M is larger than Nyquist sampling rate, all the information about the channel response is preserved. Second, the variable $\mathcal{F}_y[m]$ (22) is the DFT of $\underline{y}$, y_k can be treated as zero with negligible leakage as $|k|$ large enough when calculating (22), and $\mathcal{F}_y[m]$ obtained this way contains most of useful information. In real engineering implementation, when N is not large enough to make $\underline{y}$ contain most information about $\underline{h}$, $\underline{\tau}$ noise $\underline{N}$, R can be chosen to cover the tail of shaping pulse; when N is sufficiently large (according to specific application), then R can be dropped. From simulation we observe that the performance loss due to this simplification is negligible.

- If the PSD of $\underline{N}$ has no cross zero points within the frequency band of interest, the ML channel estimator is given by (27). Compared with other frequency domain estimation methods mentioned in [3], higher sampling rate is adopted in our algorithm, and the color noise is compensated in the denominator in (27). Also we observe that the training sequence length N has to be large enough to apply the old estimation methods.

4 An Example and Simulation Results

The substitution of Q by C_K^{-1} simplifies the channel estimation problems a lot, and gives us all the advantages of frequency domain approaches.

4.1 The Data-Aided ML Joint Phase and Timing Offset Estimator

As a special case, let us consider the following problem: $L = 1$ which means there is only one path, and $h = e^{j\theta}$, which means that the frequency-selective fading channel estimation problem is simplified to joint carrier phase and timing offsets estimation problem. Timing and phase recovery is a very important synchronization function in coherent demodulation. The variables ϕ and τ are used to model the carrier phase and timing offsets between the transmitted and received signals respectively.

The channel response $H[m]$ is equal to $e^{j\theta}e^{-j(2\pi\tau m/(N+R))}$, the following holds

$$e^{-j(2\pi\tau m/(N+R)+\theta)} = \frac{\mathcal{F}_y[m]}{\sqrt{E_s}\mathcal{F}_r[m]\mathcal{A}[m]} \tag{28}$$

The RHS of (28) is equal to an exponential wave, our objective is to estimate the frequency and phase of this exponential wave. Therefore the timing estimation problem becomes a frequency estimation problem, which has been studied for many years. For example, based on the RHS of (28), the linear regression on the phase of the RHS is proportional to τ [9].

Another estimator can be derived from (18). After some arithmetic, it is easy to verify that the ML estimator is equivalent to the following

$$\begin{aligned}(\tau,\phi) &= \arg\max_{\tau,\phi}\left\{\frac{1}{N_0}\mathrm{Re}(\underline{y}^H Q\underline{m}_y)\right\} \\ &= \arg\max_{\tau,\phi}\left\{\frac{\sqrt{E_s}}{N_0 K}\mathrm{Re}\left(e^{j\phi}\sum_{k=-K/2}^{K/2-1}\mathcal{F}_y[k]^*\mathcal{A}[k]e^{-j(2\pi k\tau/(N+R))}\right)\right\}\end{aligned} \tag{29}$$

As we mentioned before, as the training sequence length N large enough $N+R\approx N$. Define $\mu(\tau)$

$$\mu(\tau) = \sum_{k=-K/2}^{K/2-1}\mathcal{F}_y[k]^*\mathcal{A}[k]e^{-j(2\pi k\tau/N)} \tag{30}$$

two-dimension maximization can be downsized to one-dimension search

$$(\tau,\phi) = \arg\max_{\tau,\phi}\left\{|\mu(\tau)|\mathrm{Re}\left(e^{j(\phi+\arg(\mu(\tau)))}\right)\right\} \tag{31}$$

From (30), we can see that $\mu(\tau)$ is the cross-correlation between $\underline{y}$ and the training sequence $\underline{a}$ in the frequency domain. The Parsevel relation serves a bridge to connect the time domain and the frequency domain. As a result of dropping R, $K\approx MN$ for large N, (30) is equal to the following equation:

$$\mu(\tau) = \sum_{l=-N/2}^{N/2-1} y(lT_s-\tau T)^* a_l \tag{32}$$

Therefore the ML estimate of τ is the argument that maximizes the magnitude of the cross-correlation between the received samples and training sequence in either frequency domain or time domain. Actually the ML estimator with the definition of $\mu(\tau)$ (32) is the same as that in [1], which is derived based on other techniques.

From this example, we can see that frequency domain approaches give people a lot of advantages to analyze problems in both time and frequency domains, especially in colored Gaussian noise environment.

4.2 Simulation Results

Computer simulations were conducted to test the ML channel estimator and the ML data-aided joint timing and phase estimator.

An M-sequence with length 63 was used in our slow frequency-selective fading channel estimator simulation. M sequence is good for channel magnitude response estimation, because its PSD (i.e. $|\mathcal{A}[m]|$) is a constant except for DC component. For more information on training sequence design, refer to [3]. Square root raised-cosine shaping pulse with rolloff factor 0.75 was adopted in both transmitter and receiver. In our case, $K = 63M$, from simulation we can see that this K is reasonably large to apply our theorem. A 800 MHz carrier was assumed, and we used a 6-ray typical urban (TU) channel model. Different sampling rates (M) were tested.

Computer simulation results are shown in Figure 2, 3 and 4, where the X-the is the frequency with 63M equal to 2π, the Y-axis is the joint normalized magnitude response of $H[m]\mathcal{F}_r[m]$. Figure 2 shows the averaged estimation result (over 500 tests) at 0dB and $M = 4$, it is clearly that the estimator is unbiased. Figure 3 shows similar result with lower sampling rate ($M = 2$). We can see that higher sampling rate does not provide extra information when it is larger than the Nyquist sampling rate. Figure 4 shows one test result at 0dB.

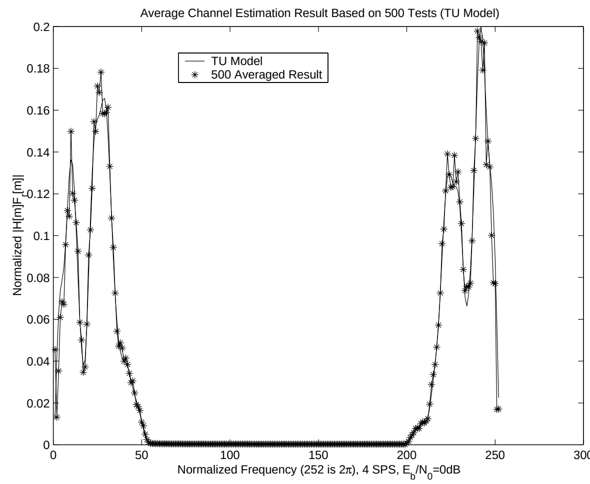

Figure 2: Channel Estimation Result Averaged over 500 Tests, 4 Samples Per Symbol

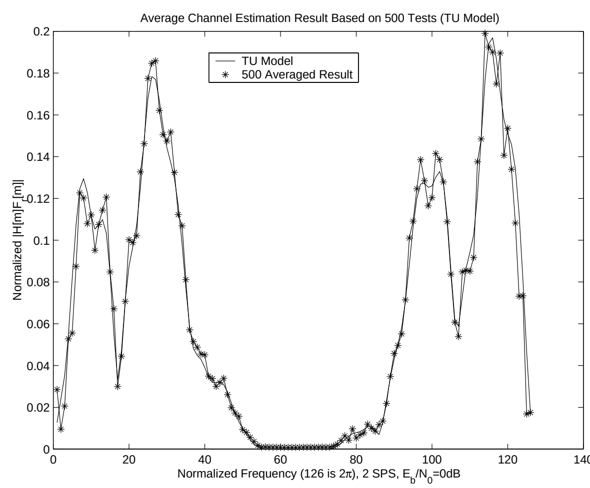

Figure 3: Channel Estimation Result Averaged over 500 Tests, 2 Samples Per Symbol

We also ran computer simulations for the ML joint timing and phase estimator derived here. Actually a simplified algorithm that used *curve-fitting* technique based on (32) was used in our simulation [14]. The following conditions were applied: 4 samples per symbol, $N = 48$, rolloff-factor α was equal to 0.5. Two kind of training sequences were tested in our timing recover test

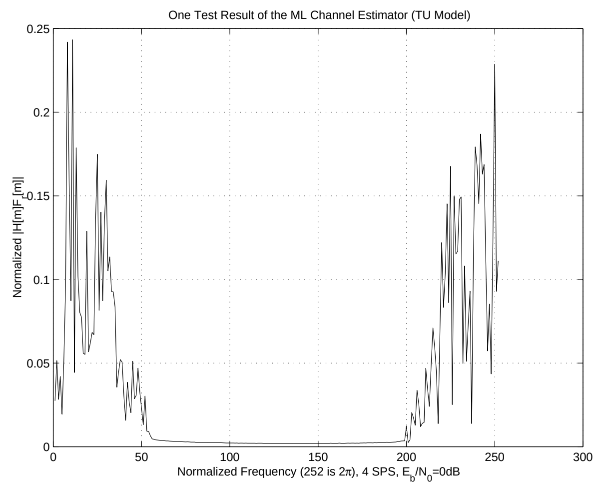

Figure 4: One Channel Estimation Result, 4 Samples Per Symbol

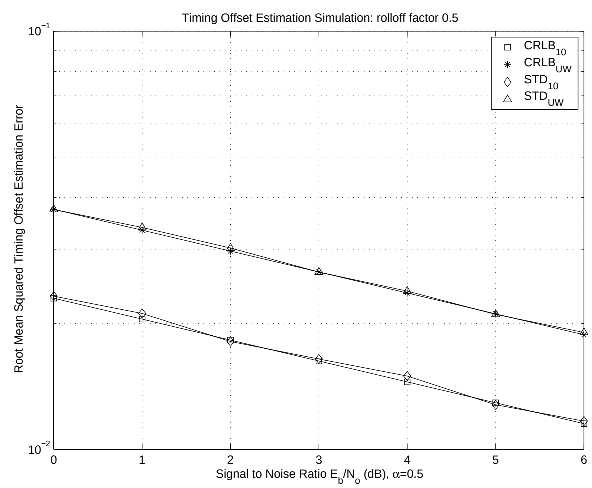

Figure 5: The RMS Timing Offset Estimation Error with $\alpha = 0.50$

because the performance of timing offset estimation is closely related to the data pattern [13]. One data pattern is alternating one-zero pattern; the other is pseudo-random data patter. Figure 5 shows the root mean square (RMS) timing offset estimation error versus the Cramer-Rao lower bound (CRB) derived by us in [13]. Figure 6 shows that of the phase estimation. From these two

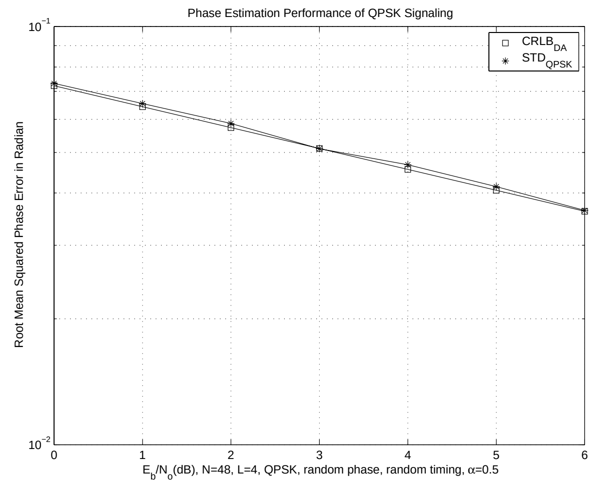

Figure 6: The RMS Phase Offset Estimation Error with $\alpha = 0.5$

figures, we can see that the performance of our estimation algorithms meets the CRB.

5 Conclusions

In this paper, we derived the frequency domain ML channel estimator with the general Gaussian noise and over-sampling assumption. The derivation is based on *the finite boundary strong sense convergence theorem for the inverse of Toeplitz matrices*. With the help of this theorem, many good algorithms that take advantages of transform domain can be derived. As a special case, the ML joint carrier phase and timing offsets estimator is presented. Simulations show that the performance of our algorithm meet the CRB at low signal to noise ratio (SNR).

References

[1] H. Meyr, M. Moeneclaey, S. Fechtel, *Digital Communication Receivers, Synchronization, Channel Estimation, and Signal Processing*, New York: Wiley, 1998.

[2] J. G. Proakis, *Digital Communications*, 3rd ed., New York: McGraw Hill, 1995.

[3] C. Tellambura, M. G. Parker, Y. J. Guo, S. J. Shepherd, S. K. Barton, *"Optimal sequences for channel estimation using discrete fourier transform techniques"*, *IEEE Trans. Comm.*, vol. 47, No. 2, pp.230-238, Feb. 1999.

[4] A. P. Clark, Z. C. Zhu, and J. K. Joshi, *"Fast start-up channel estimation in the presence of noise"*, *Proc. Inst. Elect. Eng.*, vol. 131, pp. 375-381, pt. F, Jul. 1984.

[5] T. Felhauer, *"Digital signal processing for optimum wideband channel estimation"*, *Proc. Inst. Elect. Eng.*, vol. 140, No. 3, pp. 179-186, pt. F, Jun. 1993.

[6] H. L. Van Trees, *Detection, Estimation and Modulation Theory, Part I*, New York: Wiley, 1968.

[7] U. Grenander, G. Szgeo, *Toeplitz Forms and Their Applications*, Berkeley California: University of California Press, 1958.

[8] R. M. Gray, *"Toeplitz and circulant matrices"*, Tech. Rep. 6502-1, Information Systems Laboratory, Center for Systems Research, Stanford University, Jun. 1971.

[9] S. Kay, *Modern Spectral Estimation Theory and Application*, New Jersey: Prentice Hall, 1988.

[10] T. Kailath, A. Vieira and M. Morf, *Inverses of Toeplitz Operators, Innovations, and Orthogonal Polynominals*, SIAM Reviw, Vol.20, No.1 January 1978, pp106-119.

[11] K. Kobayashi, *"Simultaneous adaptive estimation and decision algorithms for carrier modulated data transmission systems"*, *IEEE Trans. Comm.*, vol. 37, pp.268-280, Jun. 1971.

[12] F. W. Sun, Y. Jiang, J. S. Baras, *"On the convergence of Toeplitz matrices and its application"*, in preparation, Jan. 2000.

[13] Y. Jiang, F. W. Sun, J. S. Baras, *"On the true Cramer-Rao lower bound for DA joint estimation of carrier phase and timing offsets"*, *Accepted by IEEE ICC'2000*, Sept. 1999.

[14] Y. Jiang, F. B. Verahrami, W. Ting, R. L. Richmond, J. S. Baras, "VLSI implemented ML joint carrier phase and timing offsets estimator for QPSK/OQPSK burst modems", *Proc. IEEE WCNC'99*, Sept. 1999.