

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

Yimin Jiang

Feng-Wen Sun

John S. Baras

Abstract—This paper addresses the channel estimation problem for slow frequency-selective fading channels using training sequences and the maximum likelihood (ML) approach. In the literature people usually assume a symbol period spaced delay-tapped-line model and additive white Gaussian noise (AWGN). Due to the 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., the timing information is unavailable), the AWGN model is not valid anymore. A more general ML channel estimation method using the discrete Fourier transform (DFT) is derived for colored Gaussian noise and over-sampling. A similar idea can be adopted to derive the ML joint carrier phase and timing offsets estimation algorithm.

I. INTRODUCTION

For burst-transmission digital communication systems, channel estimation is required for ML sequence estimation receivers [1]. 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 problems are widely addressed in the literature [2]–[4]. Estimation can be implemented using a Wiener filter or the DFT. For example, [2]–[4] consider channel estimation given a known TS. The authors of [2] addressed the problem of selecting the optimum TS for channel estimation by processing in the frequency domain. Optimum unbiased channel estimation given white noise is considered in [3] using a ML approach. Following the least-squares (LS) philosophy, [4] presents algorithms for optimal unbiased channel estimation with aperiodic spread spectrum signals for white or nonwhite noise.

Previous works [2]–[4] assumed a symbol period delay-tapped-line model or AWGN noise [2]. Due to the pre-

filtering in the receiver front end, this model is not accurate enough and will cause aliasing or leakage in the frequency domain. Since typical pulse shaping rolloff factors in wireless communications 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 receivers [1]. When the sampling rate is higher than one sample per symbol or timing information is unknown, the 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 [4] to deal with the colored noise, which actually follows a more general idea in Van Trees' book [5]. In this paper, we shall show that a direct optimum estimator can be derived without preliminary processing [5, p.289].

In what follows we take a ML approach and derive an optimal channel estimation algorithm in the frequency domain. The key issue is how to tackle the colored Gaussian noise. It is well known that the autocovariance matrix of a colored Gaussian noise process is a Toeplitz matrix which was thoroughly studied in [6]–[8]. We show that the inverse of Toeplitz matrices can be substituted by a circular matrix asymptotically under certain condition in [7], [8]. This leads to the frequency domain approach because of the fact that the eigendecomposition of a circular matrix is equivalent to the DFT.

The rest of the paper is organized as follows. Section II models the slow frequency-selective fading channel and formulates the ML estimation problem mathematically. Section III first revisits some properties of Toeplitz matrices and then derives the channel estimator. As a special example case, the ML joint carrier phase and timing offsets estimator is addressed in Section IV that also shows some simulation results. Section V concludes the paper.

II. PROBLEM FORMULATION

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

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

Yimin Jiang and Feng-Wen Sun are with Hughes Network Systems Inc., 11717 Exploration Lane, Germantown, MD 20876, USA. E-mails: {yjiang, fsun}@hns.com.

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

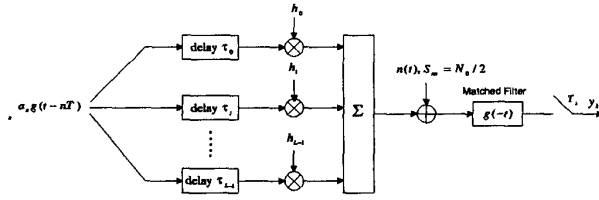


Fig. 1. Modeling of the Slow Frequency-Selective Fading Channel and the Matched Filter

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

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

where $\{a_n\}$ is the TS, $g(t) = g_T(t) \otimes f(t)$, $g_T(t)$ is the transmitter shaping function, $f(t)$ is the prefilter in the receiver, $n(t)$ 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 $y(t)$ is given by follows

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

where $r(t) = g(t) \otimes g(-t)$, $N(t) = n(t) \otimes g(-t)$. The noise process $N(t)$ is colored due the prefilter and matched filter. Let $\mathbf{y}_k = \mathbf{y}(kT_s)$ and $N_k = N(kT_s)$. The likelihood function of $\{h_l, \tau_l\}$ is the pdf of a Gaussian r.v. The mean of \mathbf{y}_k given $\{h_l, \tau_l\}$ is $\mathbf{m}_y(k)$ that is equal to

$$\mathbf{m}_y(k) = \sqrt{E_s} \sum_{l=0}^{L-1} \sum_{n=-N/2}^{N/2-1} h_l a_n r(kT_s - nT - \tau_l T), \quad (4)$$

where $k \in [-K/2, K/2 - 1]$ with $K = M(N + R)$ that is the total number of digital samples. Variable R models the ideal case in which a shaping pulse modulated only by the TS $\{a_n\}$ is transmitted and used to estimate the channel response, and the receiver observes $\mathbf{y}(t)$ beyond the TS portion in order to collect the statistics of the noise process $N(t)$. The autocovariance matrix of the observation vector \mathbf{y} ($K \times 1$) is

$$\text{cov}[\mathbf{y}|\mathbf{h}, \boldsymbol{\tau}] = \frac{N_0}{2} \Lambda \quad (5)$$

where Λ is a $K \times K$ Hermitian Toeplitz matrix with jk th element equal to

$$[\Lambda]_{j,k} = r((j - k)T_s). \quad (6)$$

The log likelihood function of $(\mathbf{h}, \boldsymbol{\tau})$ is given by

$$\begin{aligned} l(\mathbf{y}|\mathbf{h}, \boldsymbol{\tau}) &= -\frac{1}{N_o} [-\mathbf{y}^H Q \mathbf{m}_y - \mathbf{m}_y^H Q \mathbf{y} + \mathbf{m}_y^H Q \mathbf{m}_y] \\ &\quad - \left(\frac{1}{N_o} \mathbf{y}^H Q \mathbf{y} + \log \left[(2\pi)^{K/2} \left| \frac{N_o}{2} \Lambda \right|^{1/2} \right] \right) \end{aligned} \quad (7)$$

where Q is the inverse of Λ , \mathbf{m}_y is the mean vector of \mathbf{y} , \mathbf{h} and $\boldsymbol{\tau}$ are the channel response vectors. The ML estimate of channel response $(\mathbf{h}, \boldsymbol{\tau})$ is

$$(\mathbf{h}, \boldsymbol{\tau}) = \arg \max_{\mathbf{h}, \boldsymbol{\tau}} l(\mathbf{y}|\mathbf{h}, \boldsymbol{\tau}). \quad (8)$$

The computation of (7) and (8) involves the analysis of the inverse of a Toeplitz matrix.

III. ML CHANNEL ESTIMATOR IN THE FREQUENCY DOMAIN

One of the difficulties in deriving the ML channel estimator arises from the fact that inverse of a Toeplitz matrix is no longer Toeplitz. In this section, we first present some results on Toeplitz matrices and then apply these results to tackle this problem.

A. On the Inverse of Toeplitz Matrices

Intrigued by the observation that the information through wireless channels is conveyed by uniformly spaced pulses that are some kind of *distorted convolution* of data symbols and a shaping pulse, we try to design the estimation algorithm in the frequency domain. Fortunately the special characteristic of Toeplitz matrices provides a solution. The general idea is as follows. It is shown in [6] that a family of Toeplitz matrices converges to a circular matrix in the *weak sense*. However in practice, stronger convergence (e.g. quadratic forms that appear in the likelihood function (7)) is desired. Furthermore the inverse of Toeplitz matrices is not Toeplitz in general. In [7] and [8], we observe that under certain condition the central part of the inverse matrix converges to a circular matrix in the strong sense, which leads to the definition of *finite term strong sense convergence*. The eigendecomposition of a circular matrix is equivalent to the DFT, which diagonalizes the inverse matrix and provides the frequency domain approach. The above conclusion is formulated mathematically in the sequel.

A family of Toeplitz matrices $\{T_n\}$ (where n is the dimension of the matrix) is defined by a sequence of complex numbers

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

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

$$[T_n]_{i,j} = t_{i-j}, \quad (9)$$

where $t_{-i} = t_i^*$, i.e., we restrict our discussion to the Hermitian case. Its corresponding circular matrix is defined by the discrete time Fourier transform (DTFT) of $\{t_n\}$. Let $\mathcal{F}(\lambda)$ denote the DTFT of t_n , i.e.,

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

Let U_n denote the DFT matrix defined as

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

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

$$D_n = \text{diag}[\mu_{1,n}, \mu_{2,n}, \dots, \mu_{n,n}]. \quad (11)$$

Then the circular matrix $U_n^H D_n U_n$ is defined as C_n . We can introduce the finite term strong sense convergence as follows.

Definition 1 For two families of Hermitian matrices A_n and B_n , consider the quadratic form

$$\max \left| \frac{x^H (A_n - B_n) x}{x^H x} \right|, \quad (12)$$

where the maximum is over all the n -dimensional vector of the form

$$x = (0, \dots, 0, x_{-\mathcal{L}}, \dots, x_0, \dots, x_{\mathcal{L}}, 0, \dots, 0).$$

If (12) converges to zero for any given \mathcal{L} , we shall call A_n converges to B_n in the finite term strong sense.

If x corresponds to an observation within the window $[-\mathcal{L}, \mathcal{L}]$ (where \mathcal{L} does not increase with n) 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. We have the following theorem:

Theorem 1 Let $\{T_n\}$ be a family of Hermitian Toeplitz matrices associated with the sequence $\{t_n\}$, and $F(z)$ be the z -transform of $\{t_n\}$. If $|F(z)|$ is continuous and does not have any zero on the unit circle, and $\sum_{k=-\infty}^{\infty} |kt_k| \leq \infty$, T_n^{-1} converges to C_n^{-1} in the strong sense for finite term quadratic forms and

$$\|T_n^{-1} - C_n^{-1}\| \leq O(1/\sqrt{n}). \quad (13)$$

B. ML Channel Estimator

In our problem the inverse matrix is Q . Because Λ is the autocorrelation matrix of the noise process $N(t)$, it is non-negative definite. Furthermore if we assume Q exists for all K , i.e., Λ is positive definite for all K , it is easy to verify that the z transform of $\{r(kT_s)\}$ is positive on the unit circle. In the over-sampling case, the technique in [5, p.289] can be applied to guarantee our operation will be meaningful. In typical communication receivers, the shaping pulse $\{r(kT_s)\}$ usually degrades faster than $O(1/|k|^2)$, e.g., the magnitude of the raised-cosine shaping pulse is less than $O(1/|k|^3)$. In engineering practice, it is under the system designer's control to make the training portion \underline{m}_y satisfy the finite term condition through packing zeros along with the training sequence $\{a_n\}$. Following the finite term strong sense convergence theorem, we can replace the inverse matrix Q by a circular matrix $C^{-1} \triangleq U_K^H D^{-1} U_K$, where U_K is defined in (10). The k th ($k = 0, \dots, K-1$) diagonal element of D is equal to

$$\mathcal{F}_r[k] = \frac{1}{T_s} \sum_{l=-\infty}^{\infty} \mathcal{R} \left(\frac{2\pi k}{KT_s} - \frac{2\pi l}{T_s} \right) \quad (14)$$

where $\mathcal{F}_r[k]$ is the DTFT of $\{r(kT_s)\}$, $\mathcal{R}(\omega)$ is the Fourier transform of $r(t)$.

From (8) we have

$$\begin{aligned} & (\underline{h}, \underline{\tau}) \\ & = \arg \max_{\underline{h}, \underline{\tau}} \left\{ -\frac{1}{N_0} [-\underline{y}^H Q \underline{m}_y - \underline{m}_y^H Q \underline{y} + \underline{m}_y^H Q \underline{m}_y] \right\}. \end{aligned} \quad (15)$$

When K is large enough, i.e., R is large enough, and the sampling rate M satisfies the Nyquist sampling theorem, we obtain

$$\begin{aligned} \underline{y}^H Q \underline{m}_y &= \frac{\sqrt{E_s}}{K} \sum_{m=-K/2}^{K/2-1} \frac{\mathcal{F}_y[m]^* \mathcal{F}_r[m] \mathcal{A}[m] H[m]}{\mathcal{F}_r[m]} \\ &= \frac{\sqrt{E_s}}{K} \sum_{m=-K/2}^{K/2-1} \mathcal{F}_y[m]^* \mathcal{A}[m] H[m], \end{aligned} \quad (16)$$

where $\mathcal{A}[m] = \sum_{n=-N/2}^{N/2-1} a_n e^{-j2\pi mn/(N+R)}$ is the $N+R$ point DFT of TS $\{a_n\}$, $H[m] = \sum_{l=0}^{M-1} h_l e^{-j2\pi m\tau_l/(N+R)}$

that is the $N + R$ point DFT of the channel response, $\mathcal{F}_y[m]$ is the K point DFT of \underline{y} , i.e., $\mathcal{F}_y[m] = \sum_{k=-K/2}^{K/2-1} y(kT_s)e^{-j2\pi mk/K}$. Similarly we have

$$\underline{m}_y^H Q \underline{m}_y = \frac{E_s}{K} \sum_{m=-K/2}^{K/2-1} \mathcal{F}_r[m] |\mathcal{A}[m]|^2 |H[m]|^2. \quad (17)$$

After some arithmetic, it can be shown that the ML estimate of channel response is given by

$$\begin{aligned} & (\underline{h}, \underline{\tau}) \\ & = \arg \min_{\underline{h}, \underline{\tau}} \\ & \sum_{m=-K/2}^{K/2-1} \left| E_s^{1/4} \sqrt{\mathcal{F}_r[m]} \mathcal{A}[m] H[m] - \frac{\mathcal{F}_y[m]}{E_s^{1/4} \sqrt{\mathcal{F}_r[m]}} \right|^2 \\ & = \arg \min_{\underline{h}, \underline{\tau}} \sum_{m=-K/2}^{K/2-1} \left| \frac{\sqrt{E_s} \mathcal{F}_r[m] \mathcal{A}[m] H[m] - \mathcal{F}_y[m]}{E_s^{1/4} \sqrt{\mathcal{F}_r[m]}} \right|^2. \end{aligned} \quad (18)$$

In summary, the ML estimate of $(\underline{h}, \underline{\tau})$ has the following DFT

$$H[m] = \begin{cases} \mathcal{F}_y[m] / (\sqrt{E_s} \mathcal{F}_r[m] \mathcal{A}[m]) & \text{if } \mathcal{F}_r[m] \neq 0, \\ 0 & \text{if } \mathcal{F}_r[m] = 0. \end{cases} \quad (19)$$

We have the following observations:

- When K is large enough, we can apply the time-shift property of Fourier transform to separate the shaping pulse, TS and channel response in the frequency domain.
- People are only interested in the channel response within the passband of the shaping function, i.e., when $\mathcal{F}_r[m] \neq 0$, (19) 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 (19). Therefore the channel estimator (19) is unbiased.
- Because $H[m]$ is just the DTFT of \underline{h} and $\underline{\tau}$, there are many possible \underline{h} and $\underline{\tau}$ that have the same $H[m]$. If the time domain response is desired, $H[m]$ can be treated as an intermediate result.
- In practice, large K can be handled in the following manner: if TS length N is large enough, R can be dropped with negligible performance loss; if N is a small number compared with the shaping pulse length, proper number (R) of zeros can be packed along with the TS to provide the channel estimator enough statistical information of the noise process $N(t)$.

IV. AN EXAMPLE CASE AND SIMULATION RESULTS

The frequency domain approach introduced by the circular matrix approximation simplifies the estimator design, a similar idea can be applied to design joint carrier phase and timing offsets estimator.

A. The Data-Aided ML Joint Carrier Phase and Timing Offset Estimator

Consider the following special case of the channel model (1): there is only one path, i.e. $L = 1$ and $h_0 = e^{j\theta}$. The frequency-selective fading channel estimation problem becomes the joint carrier phase and timing offsets estimation problem. Variables ϕ and τ are used to model the carrier phase and timing offsets between the transmitter and receiver respectively. It is easy to verify that the ML joint carrier phase and timing offsets estimator is as follows

$$\begin{aligned} (\tau, \phi) & = \arg \max_{\tau, \phi} \left\{ \frac{1}{N_0} \Re(\underline{y}^H Q \underline{m}_y) \right\} \\ & = \arg \max_{\tau, \phi} \left\{ \frac{\sqrt{E_s}}{N_0 K} \right. \\ & \quad \left. \Re \left(\sum_{k=-K/2}^{K/2-1} \mathcal{F}_y[k] \mathcal{A}[k]^* e^{j(2\pi k\tau / (N+R) - \phi)} \right) \right\}. \end{aligned} \quad (20)$$

Define $\mu(\tau)$

$$\mu(\tau) = \frac{1}{K} \sum_{k=-K/2}^{K/2-1} \mathcal{F}_y[k] \mathcal{A}[k]^* e^{j2\pi k\tau / (N+R)}, \quad (21)$$

where $\mu(\tau)$ is the cross-correlation between the time-shifted \underline{y} and TS \underline{a} in the frequency domain. The ML timing offset estimate is given by

$$\hat{\tau} = \arg \max_{\tau} |\mu(\tau)|, \quad (22)$$

and the ML phase offset estimate is given by

$$\hat{\phi} = \arg(\mu(\hat{\tau})). \quad (23)$$

Parseval relation serves a bridge to connect the time domain processing with the frequency domain. Because K -point DFT is an orthonormal transform, we obtain

$$\mu(\tau) = \sum_{n=-N/2}^{N/2-1} y(nT + \tau T) a_n^*. \quad (24)$$

Therefore the ML estimate of τ is the argument that maximizes the magnitude of the cross-correlation between the received samples and the TS either in the frequency domain or in the time domain. In fact the ML estimator (24) was proposed in [1], which was derived based on other techniques.

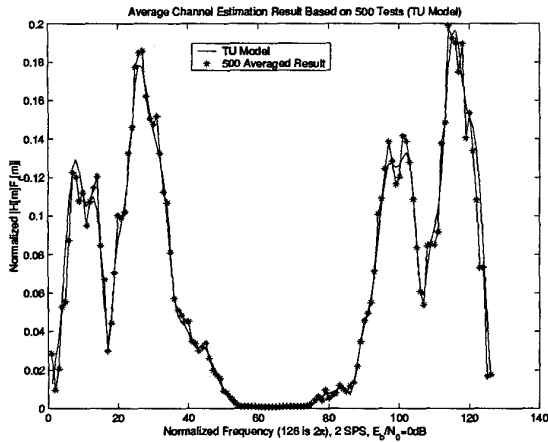


Fig. 2. Channel Estimation Result Averaged over 500 Tests, $M = 2$

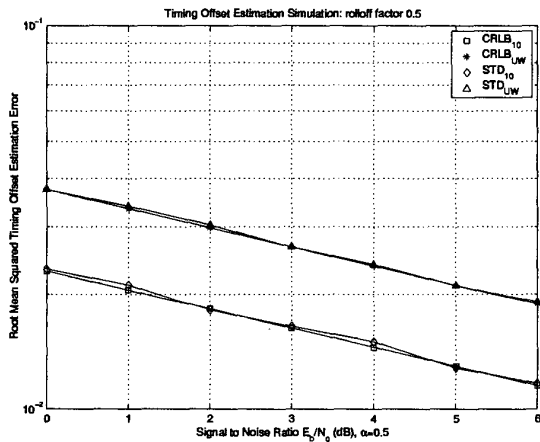


Fig. 3. The RMS Timing Offset Estimation Error with $\alpha = 0.50$

B. Simulation Results

Computer simulations were conducted to test the channel estimator and the joint carrier phase and timing offsets estimator. An M -sequence with length 63 (as the TS), a sampling rate $M = 2$ with $K = 63M$, a carrier at 800 MHz, and a 6-ray typical urban (TU) channel model were applied as in the channel estimation simulation. Figure 2 shows the averaged estimation result (over 500 tests) at 0dB, where the x -axis is the normalized frequency, and the y -axis is the normalized magnitude response $|H[m]F_r[m]|$. We can see that the estimator is unbiased.

A simplified algorithm (based on (22) that uses *curve-fitting* technique [9] was applied in our joint timing and carrier phase estimation simulation. The following conditions were assumed: $N = 48$, $M = 4$ and the rolloff factor $\alpha = 0.5$. Figure 3 shows the root mean squared (RMS) timing estimation error versus the Cramer-Rao lower bound (CRB) that was derived in [8]. Two TS

patterns were tested: the *dotting* sequence (one-zero sequence) and a pseudo-random sequence. Simulation shows that the estimation performance approaches the CRB.

V. CONCLUSIONS

In this paper, a ML channel estimator for slow frequency-selective fading channels was derived in the frequency domain given general colored Gaussian noise and over-sampling conditions. The key issue is to resolve the inverse of Toeplitz matrices that are introduced by the prefilter and matched filter in the receiver front end. A similar idea can be applied to derive the ML joint carrier phase and timing offsets estimator.

REFERENCES

- [1] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers, Synchronization, Channel Estimation, and Signal Processing*, Wiley, New York, 1998.
- [2] C. Tellambura, M. G. Parker, Y. J. Guo, S. J. Shepherd, and S. K. Barton, "Optimal sequences for channel estimation using discrete Fourier transform techniques," *IEEE Transaction on Communications*, vol. 47, no. 2, pp. 230–238, February 1999.
- [3] A. P. Clark, Z. C. Zhu, and J. K. Joshi, "Fast start-up channel estimation in the presence of noise," *IEE Proceedings*, vol. 131, pp. 375–381, July 1984.
- [4] T. Felhauer, "Digital signal processing for optimum wideband channel estimation," *IEE Proceedings*, vol. 140, no. 3, pp. 179–186, June 1993.
- [5] H. L. Van Trees, *Detection, Estimation and Modulation Theory, Part I*, Wiley, New York, 1968.
- [6] U. Grenander and G. Szego, *Toeplitz Forms and Their Applications*, University of California Press, Berkeley California, 1958.
- [7] F. W. Sun, Y. Jiang, and J. S. Baras, "On the convergence of the inverse of Toeplitz matrices and its applications," *submitted to IEEE Transaction on Information Theory*, June 2000.
- [8] Yimin Jiang, *Synchronization and Channel Parameter Estimation in Wireless Communications*, Ph.D. thesis, University of Maryland at College Park, July 2000.
- [9] Y. Jiang, F. B. Verahrami, W. Ting, R. L. Richmond, and J. S. Baras, "VLSI implemented ML joint carrier phase and timing offsets estimator for QPSK/OQPSK burst modems," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC'99)*, New Orleans, September 1999.