

Maximum Likelihood Blind Channel Estimation for Space-Time Coding Systems

Hakan A. Çirpan

*Department of Electrical Engineering, Istanbul University, Avcilar, 34850 Istanbul, Turkey
Email: hcirpan@istanbul.edu.tr*

Erdal Panayırçı

*Department of Electronic Engineering, IŞIK University, Maslak, 80670 Istanbul, Turkey
Email: eepanay@isikun.edu.tr*

Erdinc Çekli

*Department of Electrical Engineering, Istanbul University, Avcilar, 34850 Istanbul, Turkey
Email: erdinc@istanbul.edu.tr*

Received 30 May 2001 and in revised form 7 March 2002

Sophisticated signal processing techniques have to be developed for capacity enhancement of future wireless communication systems. In recent years, space-time coding is proposed to provide significant capacity gains over the traditional communication systems in fading wireless channels. Space-time codes are obtained by combining channel coding, modulation, transmit diversity, and optional receive diversity in order to provide diversity at the receiver and coding gain without sacrificing the bandwidth. In this paper, we consider the problem of blind estimation of space-time coded signals along with the channel parameters. Both conditional and unconditional maximum likelihood approaches are developed and iterative solutions are proposed. The conditional maximum likelihood algorithm is based on iterative least squares with projection whereas the unconditional maximum likelihood approach is developed by means of finite state Markov process modelling. The performance analysis issues of the proposed methods are studied. Finally, some simulation results are presented.

Keywords and phrases: blind channel estimation, conditional and unconditional maximum likelihood.

1. INTRODUCTION

The rapid growth in demand for a wide range of wireless services is a major driving force to provide high-data rate and high quality wireless access over fading channels [1]. However, wireless transmission is limited by available radio spectrum and impaired by path loss, interference from other users and fading caused by destructive addition of multipath. Therefore, several physical layer related techniques have to be developed for future wireless systems to use the frequency resources as efficiently as possible. One approach that shows real promise for substantial capacity enhancement is the use of diversity techniques [2]. Diversity techniques basically reduce the impact of fading due to multipath transmission and improve interference tolerance which in turn can be traded for increase capacity of the system. In recent years, the use of antenna array at the base station for transmit diversity has become increasingly popular, since it is difficult to deploy more than one or two antennas at the portable unit. Transmit diversity techniques make several replicas of the signal

available to the receiver with the hope that at least some of them are not severely attenuated. Moreover, the methods of transmitter diversity combined with channel coding have been employed at the transmitter, which is referred to as space-time coding, to introduce temporal and spatial correlation into signals transmitted from different antennas [2, 3]. The basic idea is to reuse the same frequency band simultaneously for parallel transmission channels to increase channel capacity [2, 3].

Unfortunately, employing antenna diversity at the transmitter is particularly challenging, since the signals are combined in space prior to reception. Moreover, estimation of fading channels in space-time systems is further complicated, since the receiver estimates the path gain from each transmit antenna to each receive antenna. It is also important to note that space-time decoding requires multi-channel state information. Thus the achievable diversity gain comes at the price of proportional increase in the amount of training which results in efficiency loss, especially in a rapidly varying environment. Clearly, the practical advantages of eliminating

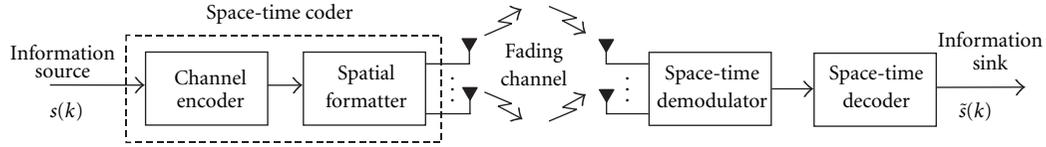


FIGURE 1: Space-time coding and decoding system.

the need for a training sequence numerous. This motivates the development of receiver structures with blind channel estimation capabilities. There has been considerable work reported in the literature on the estimation of channel information to improve performance of space-time coded systems operating on fading channels [4, 5, 6, 7]. In this paper, we consider the problem of blind estimation of space-time coded signals along with the matrix of path gains. We propose two different approaches based on the assumptions on the input sequences. Our proposed approaches also exploit the finite alphabet property of the space-time coded signals. We treat both conditional and unconditional maximum likelihood (ML) approaches. The first approach (conditional ML) results in joint estimation of the channel matrix and the input sequences, and is based on the iterative least squares and projection [8]. The second approach, which is known as unconditional ML, treats the input sequence as stochastic independent identically distributed (i.i.d.) sequences. In contrast, the unconditional ML approach formulates the blind estimation problem in discrete-time finite state Markov process framework [9, 10, 11]. Since the proposed algorithms obtain ML estimates of channel matrix and the space-time coded signals, they enjoy many attractive properties of the ML estimator including consistency and asymptotic normality. Moreover, it is asymptotically unbiased and its error covariance approaches Cramér-Rao lower bound (CRB).

The performance of the proposed ML approaches are explored based on the evaluation of CRB. The CRB is a well-known statistical tool that provides benchmarks for evaluating the performance of actual estimators. For the conditional estimator, the CRB derived in [12], is adapted to the present scenario. In unconditional case, since, the computation of the exact CRB is analytically intractable, some alternative methods must therefore be considered for simplifying CRB calculation [13]. The derivation technique used for unconditional ML have the advantage of eliminating the need to evaluate computationally intractable averaging over all possible input sequences. However, it provides a looser bound which is not as tight as the exact CRB, but it is computationally easier to evaluate.

The outline of the paper is as follows. In Section 2, we describe a basic model for a communication system that employs space-time coding with n transmit and m receive antennas. In Section 3, we derive both conditional and unconditional ML estimators for the blind estimation of space-time coded signals along with the channel matrix. In Section 4, we develop CRB for the covariance of the estimation errors for the achievable variance of any unbiased estimator

for these parameter set. Finally, we present some numerical examples that illustrate the performance of the ML estimators in Section 5.

Notations used in this paper are standard. Symbols for matrices (in capital letter) and vector (lower case) are in boldface. $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^*$, and \otimes denote transpose, Hermitian, conjugate, and Kronecker product, respectively. The symbol \mathbf{I} stands for identity matrix with proper dimension; $\hat{\boldsymbol{\theta}}$ denotes the estimate of parameter vector $\boldsymbol{\theta}$; and $\|\cdot\|$ denotes the 2-norm.

2. SYSTEM MODEL

In the sequel, we consider a mobile communication system equipped with n transmit antennas and optional m receive antennas. A general block diagram for the systems of interest is depicted in Figure 1. In this system, the source generates bit sequence $s(k)$, which are encoded by an error control code to produce codewords. The encoded data are parsed among n transmit antennas and then mapped by the modulator into discrete complex-valued constellation points for transmission across channel. The modulated streams for all antennas are transmitted simultaneously. At the receiver, there are m receive antennas to collect the transmissions. Spatial channel link between each transmit and receive antenna is assumed to experience statistically independent fading.

The signals at each receive antenna is a noisy superposition of the faded versions of the n transmitted signals. The constellation points are scaled by a factor of E_s , so that the average energy of transmitted symbols is 1. Then we have the following complex base-band equivalent received signal at receive antenna j :

$$r_j(k) = \sum_{i=1}^n \alpha_{i,j}(k) c_i(k) + n_j(k), \quad (1)$$

where $\alpha_{i,j}(k)$ is the complex path gain from transmit antenna i to receive antenna j , $c_i(k)$ is the coded symbol transmitted from antenna i at time k , $n_j(k)$ is the additive white Gaussian noise sample for receive antenna j at time k .

Equation (1) can be written in a matrix form as

$$\mathbf{r}(k) = \boldsymbol{\Omega}(k)\mathbf{c}(k) + \mathbf{n}(k), \quad (2)$$

where $\mathbf{r}(k) = [r_1(k), \dots, r_m(k)]^T \in \mathbb{C}^{m \times 1}$ is the received signal vector, $\mathbf{c}(k) = [c_1(k), \dots, c_n(k)]^T \in \mathbb{C}^{n \times 1}$ is the code vector transmitted from the n transmit antennas at time k , $\mathbf{n}(k) = [n_1(k), \dots, n_m(k)]^T \in \mathbb{C}^{m \times 1}$ is the noise vector at the receive antennas, and $\boldsymbol{\Omega}(k) \in \mathbb{C}^{m \times n}$ is the fading channel gain matrix

given as

$$\mathbf{\Omega}(k) = \begin{bmatrix} \alpha_{1,1}(k) & \cdots & \alpha_{1,n}(k) \\ \vdots & \dots & \vdots \\ \alpha_{m,1}(k) & \cdots & \alpha_{m,n}(k) \end{bmatrix}. \quad (3)$$

We impose the following assumptions on model (2) for the rest of the paper:

- (AS1) the coded symbol $c_i(k)$ is adopting finite complex values;
 (AS2) the noise vector $\mathbf{n}(k) = [n_1(k), \dots, n_m(k)]^T$ is Gaussian distributed with zero-mean and

$$\mathbb{E}[\mathbf{n}(k)\mathbf{n}^H(l)] = \sigma^2 \mathbf{I} \delta_{k,l}, \quad \mathbb{E}[\mathbf{n}(k)\mathbf{n}^T(l)] = \mathbf{0}, \quad (4)$$

where \mathbb{E} denotes expectation operator and $\delta_{k,l}$ is the Kronecker delta ($\delta_{k,l} = 1$ if $k = l$ and 0 otherwise). Thus $\mathbf{n}(k)$ is assumed to be uncorrelated both temporally and spatially;

- (AS3) the fading channel is assumed to be quasi-static flat fading, so that during the transmission of L codeword symbols across anyone of the links, the complex path gains do not change with time k , but are independent from one codeword transmission to the next, that is,

$$\alpha_{i,j}(k) = \alpha_{i,j}, \quad k = 1, 2, \dots, L. \quad (5)$$

The problem of estimating matrix of path gains along with the space-time coded signals from noisy observations $\mathbf{r}(L) = [\mathbf{r}^T(1), \dots, \mathbf{r}^T(L)]^T$ is the main concern of the paper. The traditional solution to this problem is to first estimate $\boldsymbol{\theta} = [\boldsymbol{\Omega}, \sigma^2]$ from training sequence embedded in the input signal, and then use these estimates as if they were the true parameters to obtain estimates of input sequence. As an alternative, we propose ML blind approaches based on finite alphabet property of the space-time coded signals. Then we derive ML cost functions for our proposed approaches in the next section.

3. ML ESTIMATION

Regarding the input sequence, two different assumptions can be considered: (i) conditional model which assumes the input sequences to be deterministic unknown parameters and (ii) unconditional model which assumes the input sequences to be stochastic processes. These two signal models lead to corresponding ML solutions. In the first approach, the input sequences are treated as unknown but deterministic quantities, therefore they are part of the set of unknown parameters. The number of unknown parameters in deterministic case grows with the increase in the number of observations which usually results in inconsistent estimates. In contrast, under the unconditional signal model, the input sequences are treated as random quantities, and are not included in the parameter set. As a result, the number of unknown parameters is fixed and it is therefore possible to obtain consistent estimates. Now we develop corresponding ML estimation algorithms.

3.1. Conditional ML approach

In this section, an ML approach is developed under (AS1), (AS2), (AS3), and the conditional signal model assumption. The log-likelihood function is then given by

$$\mathcal{L} = -\text{const} - mL \log \sigma^2 - \frac{1}{\sigma^2} \sum_{k=1}^L \|\mathbf{r}(k) - \mathbf{\Omega} \mathbf{c}(k)\|^2. \quad (6)$$

The conditional ML estimation can be obtained by jointly maximizing \mathcal{L} over the unknown parameters $\mathbf{\Omega}$ and $\mathbf{c}(L) = [\mathbf{c}^T(1), \dots, \mathbf{c}^T(L)]^T$. After neglecting unnecessary terms, conditional ML yields the following minimization problem:

$$\min_{\mathbf{\Omega}, \mathbf{c}(L)} \|\mathbf{r}(L) - \mathbf{\Omega} \mathbf{c}(L)\|^2. \quad (7)$$

Since the elements of $\mathbf{c}(L)$ are restricted to be finite alphabet, (7) results in a nonlinear separable optimization problem with mixed integer and continuous variables. Typically, the minimization problem in (7) is solved in two steps by alternatively minimizing with respect to $\mathbf{\Omega}$ and $\mathbf{c}(L)$ while keeping other parameters fixed. First, we minimize (7) with respect to $\mathbf{\Omega}$ by the least squares solution. Then substitute $\hat{\mathbf{\Omega}}$ back into (7) and solve it for $\mathbf{c}(L)$. The ML estimate of $\mathbf{c}(L)$ in the second step can be obtained by enumeration. However, this search is computationally very demanding since the number of possible $\mathbf{c}(L)$ matrices that need to be checked grows exponentially both with L and n . Therefore, the iterative approaches attempt to solve this problem with lower computational complexity.

We now adopt a block conditional ML algorithm that has a lower computational complexity [8]. The proposed algorithm is based on iterative least squares and projection (ILSP). It takes advantage of the ML estimator being separable in its continuous and integer variables. Note that the dimension of the channel gain matrix $\mathbf{\Omega}$ is chosen to satisfy $n \leq m$ for this particular approach.

Given an initial estimate $\hat{\mathbf{\Omega}}$ of $\mathbf{\Omega}$, the minimization of (7) with respect to $\mathbf{c}(L)$ is a least squares problem that can be solved in closed form. Each element of the solution is rounded-off to its closest discrete values (coded MPSK signals). Then a better estimate of $\mathbf{\Omega}$ is obtained by minimizing (7) with respect to $\mathbf{\Omega}$, keeping $\hat{\mathbf{c}}(L)$ fixed. This minimization also results in least squares. This process continues until $\mathbf{\Omega}$ converges. In practice, we can stop when the difference $\|\mathbf{\Omega}_i - \mathbf{\Omega}_{i-1}\|$ is within a threshold ϵ .

The following steps summarize the conditional ML algorithm:

- Start with initial estimate $\mathbf{\Omega}_{(0)}$, $i = 0$
- (1) $i = i + 1$
 - $\mathbf{c}_i(L) = (\mathbf{\Omega}_{i-1}^* \mathbf{\Omega}_{i-1})^{-1} \mathbf{\Omega}_{i-1}^* \mathbf{r}(L)$.
 - Project each element of $\mathbf{c}_i(L)$ to closest discrete values.
 - $\mathbf{\Omega}_i = \mathbf{r} \mathbf{c}_i^*(L) (\mathbf{c}_i(L) \mathbf{c}_i^*(L))^{-1}$.
 - (2) Continue until $\|\mathbf{\Omega}_i - \mathbf{\Omega}_{i-1}\| \leq \epsilon$.

Clearly, due to nonlinear operation in projecting $\mathbf{c}_i(L)$ to its closest discrete values, the convergence is not guaranteed.

However, sufficiently good initialization provided from sub-optimal techniques improve the possibility of global convergence and also reduce the number of iterations required.

3.2. Unconditional ML approach

Under (AS2), (AS3), and the signal model (2), we can formulate the probability density function of the received vector \mathbf{r} (given \mathbf{u}) as

$$f_{\theta}(\mathbf{r} | \mathbf{u}) = \frac{1}{(\pi\sigma^2)^{mL}} \prod_{k=1}^L \exp \left\{ -\frac{\|\mathbf{r}(k) - \mathbf{\Omega}\mathbf{g}(\mathbf{u}(k))\|^2}{\sigma^2} \right\}, \quad (8)$$

where $\mathbf{g}(\cdot)$ is the same nonlinear mapping that describes channel coder, spatial formatter, and modulator, $\mathbf{u}(k)$ is the input sequence influencing the space-time coded symbols.

In general, trying to estimate θ and \mathbf{u} jointly from (8) is computationally demanding except for small data alphabet size and small data record. Therefore, the goal is to obtain a cost function that is dependent only on θ , in this way it is possible to avoid least squares based on two step procedures for blind ML estimation. To this end, we therefore consider an unconditional signal model and compute the corresponding ML cost function via the expectation of the conditional ML function with respect to the statistics of the input sequences

$$f_{\theta}(\mathbf{r}) = \mathbb{E}_{\mathbf{u}} [f_{\theta}(\mathbf{r} | \mathbf{u})]. \quad (9)$$

However, the expectation $\mathbb{E}_{\mathbf{u}}$ in (9) leads to complicated cost function. The maximization of this cost function is therefore computationally demanding. At this point, we modified (AS1) for the unconditional case in the following form:

(AS1_u) information sequence $s(k)$ is an i.i.d. sequence adopting equiprobable finite values.

If we exploit the assumption (AS1_u) on the input sequence and use the conditional ML function (8), we can obtain the unconditional ML function specifically for the problem at hand as

$$f_{\theta}(\mathbf{r}) = \frac{1}{2^{(l+t-1)}(\pi\sigma^2)^{mL}} \prod_{k=1}^L \sum_{p=1}^{2^{(l+t-1)}} \exp \left\{ -\frac{\|\mathbf{r}(k) - \mathbf{\Omega}\mathbf{g}(\zeta_p)\|^2}{\sigma^2} \right\}, \quad (10)$$

where $\zeta_p = [s(lk+l-1), \dots, s(lk-t)]^T$ is the input vector influencing the coded symbols at time k , t is the number of memory elements in the encoder, $l = \log_2 M$ is the block length of information bits that are transmitted (if we restrict ourselves to MPSK). Since each element of the ζ_p takes on 2 possible values, $2^{(l+t-1)}$ is the set of all possible $(l+t-1)$ vectors of 2.

The log-likelihood function for the unconditional signal model is then given by

$$\mathcal{L}(\theta) = \sum_{k=1}^L \log \left(\sum_{p=1}^{2^{(l+t-1)}} \exp \left\{ -\frac{\|\mathbf{r}(k) - \mathbf{\Omega}\mathbf{g}(\zeta_p)\|^2}{\sigma^2} \right\} \right) + \text{constant}, \quad (11)$$

and the unconditional ML estimation of θ is the global max-

imizer of $\mathcal{L}(\theta)$. Unfortunately, existence of the globally convergent algorithm for this nonlinear cost function is unlikely. Moreover, the direct maximization of (11) still results in computationally demanding nonlinear optimization problem. In finding the ML estimator, it is quite common to resort numerical techniques of maximization such as the Newton-Raphson and scoring methods. However, the Newton-Raphson and scoring methods may suffer from convergence problems. As an alternative, the problem can be cast in a finite-state Markov chain framework by employing the Baum-Welch algorithm which reduces computational burden significantly. The Baum-Welch algorithm although iterative in nature, is guaranteed under certain mild conditions to converge and at convergence to produce a local maximum.

In the sequel, we exploit finite-state Markov process modelling property of the space-time coded signals and employed associated estimation algorithm to provide computationally efficient solution to resulting optimization problem. Let us then introduce unconditional ML framework based on finite-state Markov process modelling first.

3.2.1 Function of a Markov chain

Many important problems in digital communications such as inter-symbol interference, partial response signalling can be modelled by means of finite-state Markov process with unknown parameters observed in independent noise [10, 11]. Based on (AS1_u), codeword produced by the channel encoder in space-time coder can be characterized as a finite-state Markov process. Moreover, the received signal vector at an antenna array in the presence of spatial formatting, fading channel and noise can also be viewed as a stochastic process (function of Markov chain) that has an underlying Markovian finite-state structure.

The space-time coder is characterized by a memory of length t and $2^{(l+t-1)}$ state trellis, where the state $\zeta(k)$ at time k labels the coder memory $(s(lk+l-2), \dots, s(lk-t))$,

$$\zeta(k) \in \Pi = \{\tau_p, p = 1, \dots, 2^{(l+t-2)}\}. \quad (12)$$

The transition from state $\zeta(k)$ to $\zeta(k+1)$ is represented on the trellis by a branch denoted by the vector

$$\phi(k) = [s(lk+l-1), \dots, s(lk-t)]^T \quad (13)$$

and $\phi(k) \in \Phi = \{\xi_n, n = 1, \dots, 2^{(l+t-1)}\}$. Then both the $\{\zeta(k)\}$ sequence and the $\{\phi(k)\}$ sequence form a first-order finite Markov chains, that is,

$$\Pr [\phi(k) = \xi_n] = \Pr [\zeta(k) = \tau_q, \zeta(k-1) = \tau_s] \quad (14)$$

for some q, s depending on k .

The observation vector $\mathbf{r}(k)$ can therefore be modelled as a probabilistic function of the Markov chain. In the received signal model, the unknown channel matrix $\mathbf{\Omega}$ enter in a linear way, while the nonlinear part of the function $\mathbf{g}(\cdot)$ is due to the space-time coder and is known. Let $\mathbf{g}(\xi_n)$ de-

note the space-time coder output corresponding to the event $\phi(k) = \xi_n$. The sample $\phi(k) = \xi_n$ is a realization of the complex random sample $\mathbf{g}(\phi(k))$ which takes 2^{l+t-1} possible values depending on the $\phi(k) = \xi_n$. Moreover, every realization of a sequence of symbols corresponds to the sequence of branches $\{\mathbf{x}_k\}$ of length L , given as

$$\mathcal{X} = (\mathbf{x}_1, \dots, \mathbf{x}_L), \quad \mathcal{X} \in \Xi \mid \Xi \in 2^{L(l+t-1)}. \quad (15)$$

The underlying Markovian structure of our signal model can then be characterized by the following model parameters:

- (i) $\Pr[\zeta(k) = \tau_q \mid \zeta(k-1) = \tau_s]$ is a predetermined transition probability. If no information about the transmitted sequence is available, all permissible state transitions have the same probability, that is, $\Pr[\zeta(k) = \tau_q \mid \zeta(k-1) = \tau_s] = 1/2^{l+t-1}$, if state τ_s leads to state τ_q ;
- (ii) $\hat{\boldsymbol{\pi}}(0) = [\hat{\pi}_1(0), \dots, \hat{\pi}_{2^{l+t-1}}(0)]$ initial state probability vector. If no assumption on the starting bits is made, the initial probability is same for all states;
- (iii) the conditional density $f(\mathbf{r}(k) \mid \zeta(k) = \tau_q, \zeta(k-1) = \tau_s) = f(\mathbf{r}(k) \mid \phi(k) = \xi_n)$ is that of a Gaussian complex random vector with mean $\boldsymbol{\Omega} \mathbf{g}(\xi_n)$ and variance σ^2 .

Since the state transition probability and the initial state probability vector are predetermined, the only model parameter of the Markov chain left to be estimated is $f(\mathbf{r}(k) \mid \phi(k) = \xi_n)$ for the current model. We therefore devise the Baum-Welch algorithm to estimate the Markov chain model parameter (iii) or equivalently to estimate $\boldsymbol{\theta}$.

3.2.2 Baum-Welch algorithm

The Baum-Welch algorithm is a commonly used iterative technique for estimating the parameters of a probabilistic functions of a Markov chain. It maximizes an auxiliary function related to the Kullback-Leibler information measure instead of the likelihood function [9]. The auxiliary function is defined as a function of two sets of parameters $\boldsymbol{\theta}, \boldsymbol{\theta}'$

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}') = \sum_{\mathcal{X} \in \Xi} f_{\boldsymbol{\theta}}(\mathbf{r}, \mathcal{X}) \log(f_{\boldsymbol{\theta}'}(\mathbf{r}, \mathcal{X})), \quad (16)$$

where $f_{\boldsymbol{\theta}}(\mathbf{r}, \mathcal{X})$ represents the conditional likelihood, given a particular branch sequences \mathcal{X} , weighted by $\Pr[\mathcal{X}]$, the a priori probability of \mathcal{X} (e.g., [10]).

The theorem that forms the basis for the Baum-Welch algorithm explains the reason why Kullback-Leibler information measure can be used instead of the average likelihood.

Theorem 1. The maximization of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}')$ leads to increased likelihood, that is, $Q(\boldsymbol{\theta}, \boldsymbol{\theta}') \geq Q(\boldsymbol{\theta}, \boldsymbol{\theta}) \Rightarrow f_{\boldsymbol{\theta}'}(\mathbf{r}) \geq f_{\boldsymbol{\theta}}(\mathbf{r})$.

For the proof of the theorem, see [9].

To obtain the explicit form of the auxiliary function for the current problem, we start with

$$\log f_{\boldsymbol{\theta}'}(\mathbf{r}, \mathcal{X}) = \log \Pr[\mathcal{X}] + \log f_{\boldsymbol{\theta}'}(\mathbf{r} \mid \mathcal{X}). \quad (17)$$

Since sequences \mathcal{X} have equal probability, the first term $\log \Pr[\mathcal{X}]$ is constant. For the second term, we use the fact that the noise samples are independent and obtain

$$\begin{aligned} & \sum_{k=1}^L \log f_{\boldsymbol{\theta}'}(\mathbf{r}(k), \mathbf{x}_k) \\ &= \sum_{k=1}^L \sum_{p=1}^{2^{l+t-1}} \log f_{\boldsymbol{\theta}'}(\mathbf{r}(k), \mathbf{x}_k = \xi_p) \delta(\mathbf{x}_k, \xi_p), \end{aligned} \quad (18)$$

where $\delta(\mathbf{x}_k, \xi_p) = 1$ when $\mathbf{x}_k = \xi_p$ and 0 otherwise, and

$$\begin{aligned} & \log f_{\boldsymbol{\theta}'}(\mathbf{r}(k), \mathbf{x}_k = \xi_p) \\ &= -\frac{1}{\sigma^2} \|\mathbf{r}(k) - \boldsymbol{\Omega}' \mathbf{g}(\xi_p)\|^2 - \log(\sigma^2). \end{aligned} \quad (19)$$

Substitution of (18) in (16) yields

$$\begin{aligned} & Q(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}') \\ &= C + \sum_{k=1}^L \sum_{p=1}^{2^{l+t-1}} \left\{ \left[-\frac{1}{\sigma'^2} \|\mathbf{r}(k) - \boldsymbol{\Omega}' \mathbf{g}(\xi_p)\|^2 - \log(\sigma'^2) \right] \right. \\ & \quad \left. \times \sum_{\mathcal{X} \in \Xi} f_{\boldsymbol{\theta}^{(i)}}(\mathbf{r}, \mathcal{X}) \delta(\mathbf{x}_k, \xi_p) \right\}. \end{aligned} \quad (20)$$

It was shown in [10], that the sum over Ξ is equal to $f_{\boldsymbol{\theta}^{(i)}}(\mathbf{r}, \phi(k) = \xi_p)$. We thus have

$$\begin{aligned} & Q(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}') = C + \sum_{k=1}^L \sum_{p=1}^{2^{l+t-1}} f_{\boldsymbol{\theta}^{(i)}}(\mathbf{r}, \phi(k) = \xi_p) \\ & \quad \times \left\{ -\frac{1}{\sigma'^2} \|\mathbf{r}(k) - \boldsymbol{\Omega}' \mathbf{g}(\xi_p)\|^2 - \log(\sigma'^2) \right\}, \end{aligned} \quad (21)$$

where $\boldsymbol{\theta}^{(i)}$ is the old parameter estimates obtained at the i th iteration while $\boldsymbol{\theta}' = [\boldsymbol{\Omega}', \sigma'^2]$ is the new parameter set to be estimated at the $(i+1)$ th iteration and $f_{\boldsymbol{\theta}^{(i)}}(\mathbf{r}, \phi(k) = \xi_p)$ is the weighted conditional likelihood. The direct computation of weighted conditional likelihood is computationally intensive. Fortunately, there exist recursive procedures (called forward and backward procedures), for computing $f_{\boldsymbol{\theta}^{(i)}}(\mathbf{r}, \phi(k) = \xi_p)$ whose complexity increases only linearly with data length L [9].

The following explicit expression for the array response matrix is obtained from $\partial Q / \partial \boldsymbol{\Omega}' = 0$:

$$\begin{aligned} & \boldsymbol{\Omega}^{(i+1)} = \left(\sum_{k=1}^L \sum_{p=1}^{2^{l+t-1}} f_{\boldsymbol{\theta}^{(i)}}(\mathbf{r}, \phi(k) = \xi_p) \mathbf{r}(k) \mathbf{g}(\xi_p)^H \right) \\ & \quad \times \left(\sum_{k=1}^L \sum_{p=1}^{2^{l+t-1}} f_{\boldsymbol{\theta}^{(i)}}(\mathbf{r}, \phi(k) = \xi_p) \mathbf{g}(\xi_p) \mathbf{g}(\xi_p)^H \right)^{-1}. \end{aligned} \quad (22)$$

The last equality follows from the definition of the partial derivative with respect to a complex quantity (see, e.g., [14])

$$\frac{\partial Q}{\partial \Omega'_{ij}} = \frac{1}{2} \left[\frac{\partial Q}{\partial \text{Re}\{\Omega'_{ij}\}} + j \frac{\partial Q}{\partial \text{Im}\{\Omega'_{ij}\}} \right], \quad (23)$$

where Ω_{ij} is the ij th element of Ω .

From $\partial Q / \partial \sigma'^2 = 0$, the iterative estimation formula can also be derived for the noise variance

$$\sigma'^2 = \frac{\sum_{k=1}^L \sum_{p=1}^{2^{(l+i-1)}} f_{\theta^{(i)}}(\mathbf{r}, \phi(k) = \xi_p) \|\mathbf{r}(k) - \Omega' \mathbf{g}(\xi_p)\|^2}{\sum_{k=1}^L \sum_{p=1}^{2^{(l+i-1)}} f_{\theta^{(i)}}(\mathbf{r}, \phi(k) = \xi_p)}. \quad (24)$$

Based on this results, the steps of the proposed unconditional ML algorithm are summarized as follows:

Set the parameters to some initial value $\theta^{(0)} = (\Omega^{(0)}, \sigma'^2)^{(0)}$.

- (1) Compute the forward and backward variables to obtain $f_{\theta^{(i)}}(\mathbf{r}, \zeta(k) = \zeta_p)$.
- (2) Compute $\Omega^{(i+1)}$ from (22).
- (3) Compute $\sigma'^2^{(i+1)}$ from (24).
- (4) Repeat steps (1)–(3) until $\|\theta^{(i+1)} - \theta^{(i)}\| < \epsilon$, where ϵ is a predefined tolerance parameter.
- (5) Use $f_{\theta^{(i)}}(\mathbf{r}, \phi(k) = \xi_p)$'s to recover the transmitted symbols.

Since the proposed method exploits the finite alphabet structure of the space-time coded signals and implements a stochastic ML solution, it is expected to exhibit better performance than suboptimal estimation techniques, especially when short data records are available. For a sufficiently good initialization, the proposed algorithm converges rapidly to the ML estimate of $\hat{\theta}$. In practice, however, we did not observe convergence problem when we initialized parameters according to suggestions of [11] (while initial guess on σ'^2 is large enough to avoid overflow, Ω is initialized arbitrarily (e.g., $\Omega^{(0)} \approx \mathbf{0}$)).

4. PERFORMANCE ANALYSIS

The performance of the conditional and unconditional ML methods are assessed here by deriving their CRBs for the unbiased estimates of the nonrandom parameters. The CRB depends on the information on vector parameter θ quantified by the Fisher information matrix (FIM) and provides a lower bound on the variance of the unbiased estimate (i.e., $\mathbb{E}\{\hat{\theta}\} = \theta$). Then the CRB for an unbiased estimator $\hat{\theta}$ is bounded by the inverse of the FIM $\mathbf{J}(\theta)$:

$$\mathbb{E}\left\{(\theta - \hat{\theta})(\theta - \hat{\theta})^T\right\} \geq \mathbf{J}^{-1}(\theta). \quad (25)$$

4.1. Conditional CRB

The derivation of $\mathbf{J}(\theta)$ in (25) follows along the lines of [12]. We start constructing FIM by calculating the derivative of (6) with respect to

$$\boldsymbol{\tau} = [\mathbf{c}_r^T(1) \ \mathbf{c}_c^T(1) \ \cdots \ \mathbf{c}_r^T(L) \ \mathbf{c}_c^T(L) \ \boldsymbol{\alpha}_r^T \ \boldsymbol{\alpha}_c^T]^T, \quad (26)$$

where

$$\begin{aligned} \mathbf{c}_r(k) &= \text{Re} \left\{ [c_1(k), \dots, c_n(k)]^T \right\}, \\ \mathbf{c}_c(k) &= \text{Im} \left\{ [c_1(k), \dots, c_n(k)]^T \right\}, \\ \boldsymbol{\alpha}_r^i &= \text{Re} \left\{ [\alpha_{1,i}, \dots, \alpha_{n,i}]^T \right\}, \\ \boldsymbol{\alpha}_c &= \text{Re} \left\{ [\alpha_1^T, \dots, \alpha_n^T]^T \right\}, \\ \boldsymbol{\alpha}_c^i &= \text{Im} \left\{ [\alpha_{1,i}, \dots, \alpha_{n,i}]^T \right\}, \\ \boldsymbol{\alpha}_c &= \text{Im} \left\{ [\alpha_1^T, \dots, \alpha_n^T]^T \right\}. \end{aligned} \quad (27)$$

Taking the partial derivatives of (6), we then have

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \mathbf{c}_r(k)} &= \frac{\partial}{\partial \mathbf{c}_r(k)} \left(\text{const.} - \frac{1}{\sigma'^2} \sum_{k=1}^L \mathbf{n}^H(k) \mathbf{n}(k) \right) \quad k = 1, \dots, L \\ &= \frac{1}{\sigma'^2} (\boldsymbol{\Omega}^H \mathbf{n}(k) + \boldsymbol{\Omega}^T \mathbf{n}^*(k)) \\ &= \frac{2}{\sigma'^2} \text{Re} \left\{ \boldsymbol{\Omega}^H \mathbf{n}(k) \right\}, \\ \frac{\partial \mathcal{L}}{\partial \mathbf{c}_c(k)} &= \frac{\partial}{\partial \mathbf{c}_c(k)} \left(\text{const.} - \frac{1}{\sigma'^2} \sum_{k=1}^L \mathbf{n}^H(k) \mathbf{n}(k) \right) \quad k = 1, \dots, L \\ &= \frac{1}{\sigma'^2} (-j \boldsymbol{\Omega}^H \mathbf{n}(k) + j \boldsymbol{\Omega}^T \mathbf{n}^*(k)) \\ &= \frac{2}{\sigma'^2} \text{Im} \left\{ \boldsymbol{\Omega}^H \mathbf{n}(k) \right\}, \\ \frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_r^i} &= \frac{2}{\sigma'^2} \sum_{k=1}^L \text{Re} \left\{ c_i^*(k) \mathbf{n}(k) \right\} \quad i = 1, \dots, n, \\ \frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_r} &= \frac{2}{\sigma'^2} \sum_{k=1}^L \text{Re} \left\{ \mathbf{c}^*(k) \otimes \mathbf{n}(k) \right\}, \\ \frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_c^i} &= \frac{1}{\sigma'^2} \left\{ -j c_i^*(k) \mathbf{n}(k) + j c_i(k) \mathbf{n}^*(k) \right\} \quad i = 1, \dots, n \\ &= \frac{2}{\sigma'^2} \sum_{k=1}^L \text{Im} \left\{ c_i(k) \mathbf{n}(k) \right\}, \\ \frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_c} &= \frac{2}{\sigma'^2} \sum_{k=1}^L \text{Im} \left\{ \mathbf{c}^*(k) \otimes \mathbf{n}(k) \right\}. \end{aligned} \quad (28)$$

We need the following assumption and results to obtain FIM, (see [12]):

$$\begin{aligned} E[\mathbf{n}(n) \mathbf{n}^H(m)] &= \sigma'^2 \mathbf{I}, \\ E[\mathbf{n}(n) \mathbf{n}^T(m)] &= 0, \\ E[\mathbf{n}^H(n) \mathbf{n}(n) \mathbf{n}^T(m)] &= 0. \end{aligned} \quad (29)$$

Using (28), (29), and taking expectations, we then obtain the entries of the FIM for the conditional case, which are given by

$$\begin{aligned}
E\left\{\left(\frac{\partial \mathcal{L}}{\partial \mathbf{c}_r(n)}\right)\left(\frac{\partial \mathcal{L}}{\partial \mathbf{c}_r(m)}\right)^T\right\} &= \frac{2}{\sigma^2} \operatorname{Re}\{\boldsymbol{\Omega}^H \boldsymbol{\Omega}\} \delta_{n,m} = \mathbf{A}, \\
E\left\{\left(\frac{\partial \mathcal{L}}{\partial \mathbf{c}_r(n)}\right)\left(\frac{\partial \mathcal{L}}{\partial \mathbf{c}_c(m)}\right)^T\right\} &= -\frac{2}{\sigma^2} \operatorname{Im}\{\boldsymbol{\Omega}^H \boldsymbol{\Omega}\} \delta_{n,m} = \mathbf{B}, \\
E\left\{\left(\frac{\partial \mathcal{L}}{\partial \mathbf{c}_c(n)}\right)\left(\frac{\partial \mathcal{L}}{\partial \mathbf{c}_c(m)}\right)^T\right\} &= \frac{2}{\sigma^2} \operatorname{Re}\{\boldsymbol{\Omega}^H \boldsymbol{\Omega}\} \delta_{n,m}, \\
E\left\{\left(\frac{\partial \mathcal{L}}{\partial \mathbf{c}_r(k)}\right)\left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_r}\right)^T\right\} &= \frac{2}{\sigma^2} \operatorname{Re}\{\boldsymbol{\Omega}^H \otimes \mathbf{c}^H(k)\} = \mathbf{C}_k, \\
E\left\{\left(\frac{\partial \mathcal{L}}{\partial \mathbf{c}_c(k)}\right)\left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_r}\right)^T\right\} &= \frac{2}{\sigma^2} \operatorname{Im}\{\boldsymbol{\Omega}^H \otimes \mathbf{c}^H(k)\} = \mathbf{D}_k, \\
E\left\{\left(\frac{\partial \mathcal{L}}{\partial \mathbf{c}_r(k)}\right)\left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_c}\right)^T\right\} &= -\frac{2}{\sigma^2} \operatorname{Im}\{\boldsymbol{\Omega}^H \otimes \mathbf{c}^H(k)\}, \\
E\left\{\left(\frac{\partial \mathcal{L}}{\partial \mathbf{c}_c(k)}\right)\left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_c}\right)^T\right\} &= \frac{2}{\sigma^2} \operatorname{Re}\{\boldsymbol{\Omega}^H \otimes \mathbf{c}^H(k)\}, \\
E\left\{\left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_r}\right)\left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_r}\right)^T\right\} &= \frac{2}{\sigma^2} \sum_{n=1}^L \sum_{m=1}^L \operatorname{Re}[\mathbf{c}^*(k) \otimes \mathbf{n}(k) \\
&\quad \times \mathbf{n}^H(m) \otimes \mathbf{c}^H(m)], \\
&= \frac{2}{\sigma^2} \sum_{k=1}^L \operatorname{Re}[\mathbf{c}^*(k) \otimes \mathbf{I}_m \\
&\quad \otimes \mathbf{c}^H(k)] = \mathbf{E}, \\
E\left\{\left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_c}\right)\left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_c}\right)^T\right\} &= \frac{2}{\sigma^2} \sum_{n=1}^L \sum_{m=1}^L \operatorname{Re}[\mathbf{c}^*(k) \otimes \mathbf{n}(k) \\
&\quad \times \mathbf{n}^H(m) \otimes \mathbf{c}^H(m)], \\
&= \frac{2}{\sigma^2} \sum_{k=1}^L \operatorname{Re}[\mathbf{c}^*(k) \otimes \mathbf{I}_m \otimes \mathbf{c}^H(k)], \\
E\left\{\left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_r}\right)\left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_c}\right)^T\right\} &= -\frac{2}{\sigma^2} \sum_{n=1}^L \sum_{m=1}^L \operatorname{Im}[\mathbf{c}^*(k) \otimes \mathbf{n}(k) \\
&\quad \times \mathbf{n}^H(m) \otimes \mathbf{c}^H(m)], \\
&= -\frac{2}{\sigma^2} \sum_{k=1}^L \operatorname{Im}[\mathbf{c}^*(k) \otimes \mathbf{I}_m \\
&\quad \otimes \mathbf{c}^H(k)] = -\mathbf{F}. \tag{30}
\end{aligned}$$

Then the FIM can be written in partitioned form as

$$\mathbf{J} = \left[\begin{array}{ccc|c} \mathcal{H} & \mathbf{0} & & \mathcal{C}_1 \\ & \ddots & & \vdots \\ \mathbf{0} & & \mathcal{H} & \mathcal{C}_L \\ \hline \mathcal{C}_1^T & \dots & \mathcal{C}_L^T & \mathcal{E} \end{array} \right], \tag{31}$$

where

$$\mathcal{H} = \begin{bmatrix} \mathbf{A} & -\mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{bmatrix}, \quad \mathcal{C}_k = \begin{bmatrix} \mathbf{C}_k & -\mathbf{D}_k \\ \mathbf{D}_k & \mathbf{C}_k \end{bmatrix}, \quad \mathcal{E} = \begin{bmatrix} \mathbf{E} & -\mathbf{F} \\ \mathbf{F} & \mathbf{E} \end{bmatrix}. \tag{32}$$

The FIM can now be directly constructed. We can numerically compute the variance of individual parameter estimate by inverting the FIM $\operatorname{CRB}(\boldsymbol{\tau}) = \operatorname{diag}\{\mathbf{J}^{-1}(\boldsymbol{\tau})\}$.

4.2. Unconditional CRB

We now turn to the evaluation of the unconditional CRB. Under (AS1_u), the computation of the exact CRB is analytically intractable, we therefore consider an alternative approach for simplifying CRB calculation [13].

The evaluation of the exact form of the unconditional CRB requires the Hessian matrix for the unconditional log-likelihood function. The corresponding log-likelihood function explicitly for the current problem is given by

$$\begin{aligned}
\log [f_{\boldsymbol{\theta}}(\mathbf{r})] &= -nL \log(2) - mL \log(\pi \sigma^2) \\
&\quad + \sum_{k=1}^L \log \left(\sum_{p=1}^{2^{(l+t-1)}} \exp \left\{ -\frac{\|\mathbf{r}(k) - \boldsymbol{\Omega} \mathbf{g}(\zeta_p)\|^2}{\sigma^2} \right\} \right). \tag{33}
\end{aligned}$$

Unfortunately, due to the nature of (33) the evaluation of the Hessian matrix is analytically intractable. However, it is common to adopt (see, e.g., [13]) an approximate log-likelihood function to obtain valid CRB. Due to concavity of the log-likelihood function and Jensen's inequality, we obtain from (33) the following approximate log-likelihood function:

$$\log [f_{\boldsymbol{\theta}}(\mathbf{r})] \leq \sum_{k=1}^L \sum_{p=1}^{2^{(l+t-1)}} \log \left[\exp \left\{ -\frac{\|\mathbf{r}(k) - \boldsymbol{\Omega} \mathbf{g}(\zeta_p)\|^2}{\sigma^2} \right\} \right]. \tag{34}$$

If we further simplify (35), we obtain

$$\log [f_{\boldsymbol{\theta}}(\mathbf{r})] \leq -\frac{1}{\sigma^2} \sum_{k=1}^L \sum_{p=1}^{2^{(l+t-1)}} \|\mathbf{r}(k) - \boldsymbol{\Omega} \mathbf{g}(\zeta_p)\|^2. \tag{35}$$

At this point, we should point out that the Hessian matrix from the approximate log-likelihood function can be easily obtained. However, (35) leads to a CRB called modified CRB(MCRB) which is not as tight as exact CRB, but it is computationally easier to evaluate.

It turns out from the approximate log-likelihood function of (34) that the entries of the FIM are as

$$\mathbf{J}_{\sigma^2, \sigma^2} = \frac{nL}{\sigma^4}, \quad \mathbf{J}_{\sigma^2, \boldsymbol{\Omega}} = \mathbf{0}, \quad \mathbf{J}_{\boldsymbol{\Omega}, \sigma^2} = \mathbf{0}. \tag{36}$$

Moreover, the submatrix $\mathbf{J}_{\boldsymbol{\Omega}, \boldsymbol{\Omega}}$ can also be obtained as

$$\mathbf{J}_{\boldsymbol{\Omega}, \boldsymbol{\Omega}} = \frac{2}{\sigma^2} \sum_{p=1}^{2^{(l+t-1)}} \mathbf{g}(\zeta_p) \mathbf{g}^H(\zeta_p). \tag{37}$$

The i.i.d. input sequence coded with orthogonal space-time codes results in uncorrelated coded sequence. It is therefore possible to further simplify the valid MCRB's. In this case, the valid MCRB can be easily obtained as follows:

$$\mathbf{J}^{-1} = \sigma^2 \begin{bmatrix} \sigma^2 & & \\ & \mathbf{0} & \\ & & 2 \mathbf{I} \\ \mathbf{0} & & & \frac{1}{2^{2(l+t-1)}} \end{bmatrix}. \tag{38}$$

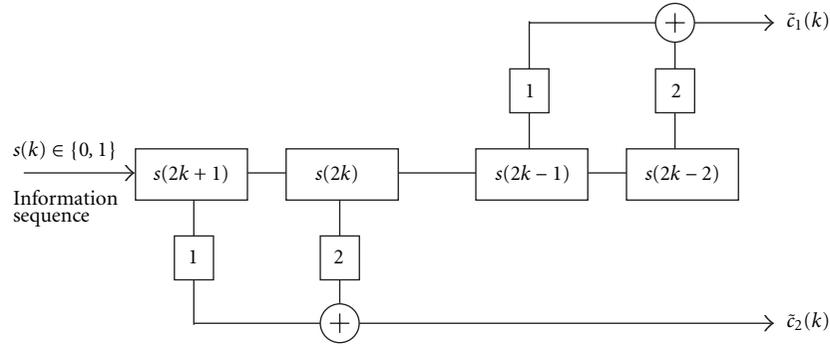


FIGURE 2: 4-state space-time coding system model.

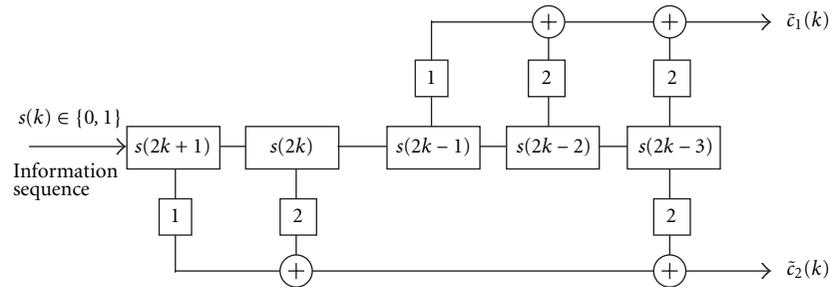


FIGURE 3: 8-state space-time coding system model.

5. SIMULATIONS

In this section, we illustrate some simulation results to evaluate the effectiveness and applicability of the proposed ML approaches. We consider the generator matrix form representation of the space-time coding system [15]. In this representation the stream of coded complex *M*PSK symbols are obtained by applying the mapping function \mathcal{M} to the following matrix multiplication:

$$\mathbf{c}(k) = \mathcal{M}(\mathbf{u}(k) \cdot \mathbf{G}(\text{mod}M)), \quad (39)$$

where $\mathbf{u}(k) = [s(lk+t-1), \dots, s(lk-t)]^T$ and \mathbf{G} is the generator matrix with n columns and $l+s$ rows and \mathcal{M} is a mapping function that maps integer values \tilde{c}_i to the coded *M*PSK symbols, $\mathcal{M}(\tilde{c}_i) = \exp(2\pi j\tilde{c}_i/M)$.

The performance of the proposed methods was evaluated as a function of SNR (signal-to-noise ratio) based on the Monte Carlo simulations. Both conditional and unconditional ML methods were tested for 200 Monte Carlo trials per SNR point across range of SNRs. In each trial, the estimation error of each parameter estimate from conditional and unconditional ML for the channel parameters were recorded. We consider the following two different cases.

Case 1. 4PSK space-time code example shown in Figure 2 is considered with $n = 2$, $t = 2$ and the generator matrix

$$\mathbf{G} = \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 0 & 2 \\ 0 & 1 \end{bmatrix}. \quad (40)$$

In this case, the coded 4PSK symbols obtained from two current information bits are transmitted over the first antenna, whereas the coded 4PSK symbols obtained from two preceding bits are transmitted over the second antenna simultaneously. The coded symbols are then transmitted through quasi-static fading channel matrix.

In Figure 4, we have plotted the estimation error obtained from conditional and unconditional ML for the channel parameters as well as the corresponding CRBs. The estimation error experienced by the proposed estimation procedures at each iteration (SNR = 10 dB) is shown in Figure 6.

Case 2. A slightly more complicated space-time encoder with $n = 2$, $t = 3$ and the generator matrix

$$\mathbf{G} = \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 0 & 2 \\ 0 & 1 \\ 2 & 2 \end{bmatrix} \quad (41)$$

is considered in this case. This example would be an 8-state code as shown in Figure 3.

In Case 2, the coded 4PSK symbols generated from $[s(2k+1), s(2k), s(2k-3)]$ are transmitted over the first antenna, whereas the coded 4PSK symbols obtained from $[s(2k-1), s(2k-2), s(2k-3)]$ are transmitted over the second antenna simultaneously. The coded symbols are then transmitted through the quasi-static fading channel matrix.

Figure 5 shows the experimental estimation error for

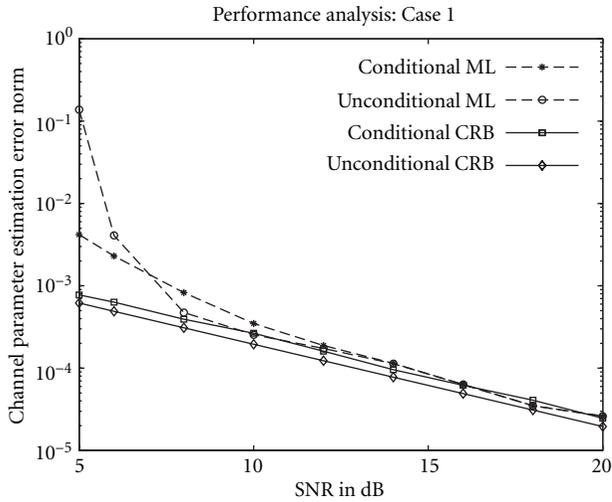


FIGURE 4: Case 1: Channel matrix estimation error norm.

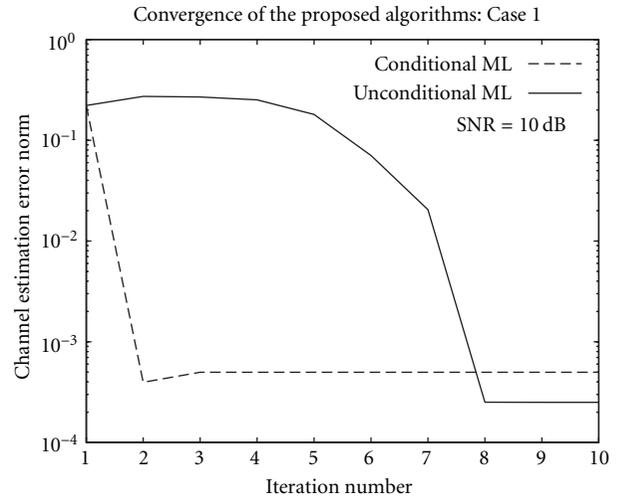


FIGURE 6: Case 1: Convergence of the channel matrix.

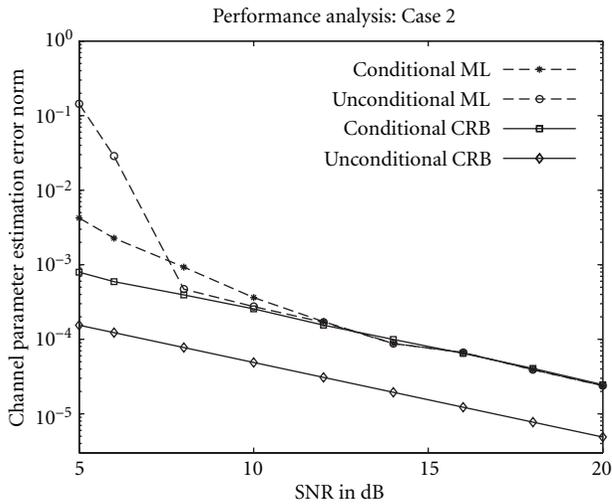


FIGURE 5: Case 2: Channel matrix estimation error norm.

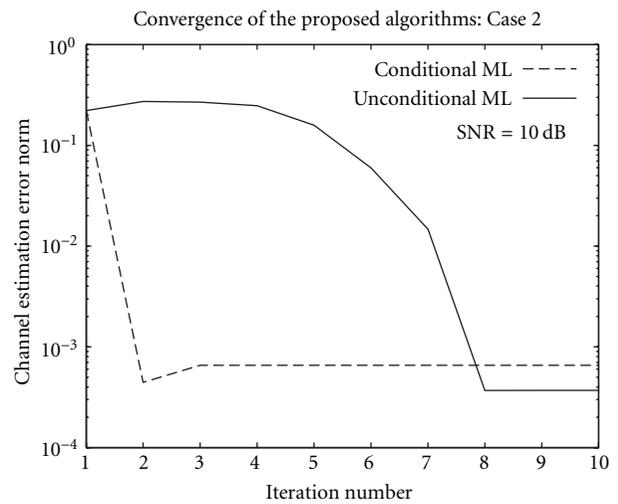


FIGURE 7: Case 2: Convergence of the channel matrix.

both the conditional and unconditional ML together with their corresponding CRB's for a range of SNR's. Figure 7 shows the estimation error experienced by the proposed estimation procedures at each iteration (SNR = 10 dB).

Based on the simulations we made the following observations:

- (i) the proposed conditional and unconditional ML approaches perform almost identically for high SNR values. Moreover, conditional ML achieve conditional CRB for high SNRs;
- (ii) since the unconditional cost function is dominated by only one term for high SNR, it results in exactly the same cost function as one would obtain for conditional ML estimation of θ . It is therefore expected that both conditional and unconditional cost functions yield similar estimates of θ at high SNR. Thus the unconditional ML approach also achieves conditional CRB for high SNR;

- (iii) the unconditional approach requires more iterations than the conditional approach to converge, however, unconditional approach is more successful in reducing channel estimation error norm at convergence for moderate SNR values.

6. CONCLUSIONS

In this paper, we presented the conditional and unconditional approaches to the problem of blind estimation of the channel parameters along with the space-time coded sequence. We derived iterative ML algorithms based on the conditional and unconditional signal models. Furthermore, the performance of the proposed algorithms are explored based on the derivation of their associated CRBs. We also presented Monte Carlo simulations to verify the theoretically predicted estimator's performance. The examples demonstrated that proposed ML approaches achieve the conditional

CRB for high SNR values. Since the unconditional CRB provides a looser bound, it is not as tight as exact CRB.

ACKNOWLEDGMENTS

This work was supported in part by the Research Fund of The University of Istanbul, Project numbers: B-924/12042001, Ö-1032/07062001, 1072/031297 and The Scientific and Technical Council of Turkey (TUBITAK) Project number 100EE006.

REFERENCES

- [1] T. S. Rappaport, *Wireless Communications Principles and Practice*, Prentice Hall, Upper Saddle River, NJ, USA, 1996.
- [2] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.
- [3] A. F. Naguib, V. Tarokh, N. Seshadri, and A. R. Calderbank, "A space-time coding model for high data rate wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1459–1478, 1998.
- [4] Y. Li, G. N. Georghiades, and G. Huang, "EM-based sequence estimation for space-time codes systems," in *ISIT '2000*, p. 315, Sorrento, Italy, June 2000.
- [5] A. F. Naguib and N. Seshadri, "MLSE and equalization of space-time coded signals," in *VTC2000*, pp. 1688–1693, Tokyo, Japan, Spring 2000.
- [6] Z. Liu, X. Ma, and G. B. Giannakis, "Space-time coding and Kalman filtering for diversity transmissions through time-selective fading channels," in *Proc. MILCOM Conf.*, vol. 1, pp. 382–386, Los Angeles, Calif, USA, October 2000.
- [7] C. Cozzo and B. L. Hughes, "Joint channel estimation and data symbol detection in space-time communications," in *IEEE International Conference on Communications*, vol. 1, pp. 287–291, 2000.
- [8] S. Talwar, M. Viberg, and A. Paulraj, "Blind estimation of multiple co-channel digital signals using an antenna array," *IEEE Signal Processing Letters*, vol. 1, no. 2, pp. 29–31, 1994.
- [9] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," *The Annals of Mathematical Statistics*, vol. 41, no. 1, pp. 164–171, 1970.
- [10] G. K. Kaleh and R. Valet, "Joint parameter estimation and symbol detection for linear and nonlinear unknown channels," *IEEE Trans. Communications*, vol. 42, no. 7, pp. 2406–2413, 1994.
- [11] M. Erkuurt and J. G. Proakis, "Joint data detection and channel estimation for rapidly fading channels," in *IEEE Globecom '1992*, pp. 910–914, Orlando, Fla, USA, December 1992.
- [12] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, no. 5, pp. 720–741, 1989.
- [13] A. N. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer-Rao bound and its application to synchronization problems," *IEEE Trans. Communications*, vol. 42, no. 2/3/4, pp. 1391–1399, 1994.
- [14] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1996.
- [15] S. Bärö, G. Bauch, and A. Hansmann, "Improved codes for space-time trellis coded modulation," *IEEE Communications Letters*, vol. 4, no. 1, pp. 20–22, 2000.

Hakan A. Çırpan received his B.S. degree in 1989 from Uludag University, Bursa, Turkey, the M.S. degree in 1992 from the University of Istanbul, Istanbul, Turkey, and the Ph.D. degree in 1997 from the Stevens Institute of Technology, Hoboken, NJ, USA, all in electrical engineering. From 1995–1997, he was a Research Assistant with the Stevens Institute of Technology, working on signal processing for wireless communications. In 1997, he joined the faculty of the Department of Electrical-Electronics Engineering at The University of Istanbul. His current research activities are focused on signal processing and communication concepts with specific attention to channel estimation and equalization algorithms for space-time coding and multicarrier (OFDM) systems. Dr. Çırpan received the Peskin Award from Stevens Institute of Technology as well as Prof. Nazim Terzioğlu award from the Research fund of The University of Istanbul. He is a Member of IEEE and Member of Sigma Xi.



Erdal Panayırıcı received the Diploma Engineering degree in electrical engineering from Istanbul Technical University, Istanbul, Turkey in 1964 and the Ph.D. degree in electrical engineering and system science from Michigan State University, East Lansing, Michigan, USA, in 1970. Between 1970–2000 he has been with the Faculty of Electrical and Electronics Engineering at the Istanbul Technical University, where he was a Professor and Head of the Telecommunications Chair. Currently, he is a Professor and Head of the Electronics Engineering Department at IŞIK University, Istanbul, Turkey. He is engaged in research and teaching in digital communications and wireless systems, equalization and channel estimation in multicarrier (OFDM) communication systems, and efficient modulation and coding techniques (TCM and turbo coding). He spent two years (1979–1981) with the Department of Computer Science, Michigan State University, as a Fulbright-Hays Fellow and a NATO Senior Scientist. From August 1990 to December 1991 he was with the Center for Communications and Signal Processing, New Jersey Institute of Technology, as a Visiting Professor, and took part in the research project on Interference Cancellation by Array Processing. Between 1998–2000, he was Visiting Professor at the Department of Electrical Engineering, Texas A&M University and took part in research on developing efficient synchronization algorithms for OFDM systems. Between 1995–1999, Prof. Panayırıcı was an Editor for IEEE Transactions on Communications in the fields of Synchronizations and Equalizations.



Erdinc Çekli was born in Istanbul, Turkey, on April 19, 1969. He received his B.S., M.S., and Ph.D. degrees in electrical engineering from Istanbul University, Istanbul Turkey, in 1993, 1996, and 2001, respectively. From 1994–2001, he was a Research Assistant at the University of Istanbul. He was a Visiting Researcher at the Technical University of Munich in 1999. Currently, he works as a Research Associate in the Scientific and Technical Council of Turkey (TUBITAK).

