Spatio-Temporal Channel Identification and Equalization in the Presence of Strong Co-Channel Interference

Chong-Meng Samson See, Colin F.N. Cowan and Arye Nehorai

C.M.S. See is with DSO NATIONAL LABORATORIES, 20 Science Park Drive, SINGAPORE 118230. Tel: 065-3732115. Fax: 065-2728418. Email: schongmc@dsio.org.sg. A. Nehorai is with the University of Illinois, Chicago, USA. Tel: (312)996-2778. Fax:(312)413-0024. Email: nehorai@eecs.uic.edu. The work of A. Nehorai was supported by the Air Force of Scientific Research under Grant F49620-97-1-0481, the National Science Foundation under Grant MIP-9615590 and the office of Naval Research under Grant N00014-96-1-1078. C.F.N. Cowan is with Electrical and Electronics Engnr Dept., Queen’s University of Belfast, BT9 5AH. Email: c.f.n.cowan@ee.qub.ac.uk. C.F.N. Cowan holds a personal research chair supported by the Royal Academy of Engineering and Nortel. A shorter version of this paper will be presented as “Spatio-Temporal Channel Identification and Equalization in the Presence of Strong Co-Channel Interference” (Invited Paper) at 32nd Asilomar Conference on Signals, Systems and Computers, Nov. 1998 by C.M.S. See, A. Nehorai, C.F.N. Cowan.

December 16, 1998

DRAFT
Abstract

We address the problem of identifying and equalizing communication channels in the presence of strong co-channel interference (CCI). In this paper, we consider the interference and noise as colored noise of unknown covariance. In addition to the finite alphabet property of the information sequence and the inherent algebraic structure of the data model, we show empirically that potential improvement in sequence detection and channel estimation accuracy can be achieved from temporal diversity where the channel outputs are observed from a tapped delay line. To exploit these underlying structures and constraints under a common framework, the proposed algorithm optimizes a weighted least square cost function using an iterative reweighting alternating minimization procedure. Numerical examples are presented and showed that the proposed algorithm is capable of achieving reliable channel identification and equalization in the presence of strong CCI at moderate SNR.

I. INTRODUCTION

The radio propagation channels in wireless mobile communications are extremely harsh. Under the constraints of limited signal bandwidth and high signaling rate, signals propagating through the multipath radio channels will suffer impairments due to intersymbol interference (ISI). In TDMA cellular systems like GSM, the multipath induced ISI needs to be mitigated to achieve reliable sequence detection.

The notion of frequency reuse plays a key role in the improvement of the spectral efficiency in many cellular systems. The extent of the spectral efficiency improvement is however limited by the frequency reuse distance which determined by the tolerable radio frequency interference such as co-channel interference (CCI). In addition, the choice of frequencies in the adjacent cells are also affected by the level of tolerable adjacent channel interferences (ACI) due to the limitation of the receiver’s front end filtering. Unlike thermal noise, the effects of CCI as well as ACI cannot be removed by increasing the signal power. Increasing the signal transmission power in one cell will result in the correspondingly increase of the CCI/ACI level in the neighbouring cells. Current cellular systems mitigate and control these effects through careful frequency planning and allocation.

The spectral efficiencies of current cellular systems need to be improved in order to
meet future demands driven by the increasing number of subscribers and the continual need to enhance the wireless services. But with limited bandwidth resource, increasing the signalling rate will worsen the channel impairments due to multipath induced ISI and the reduction of frequency reuse distance will increase the level of CCI/ACI. In order to achieve the desired spectral efficiency, it is important to develop effective algorithms to curtail the signal impairments due to ISI and CCI.

This paper addresses the problem of channel identification and equalization of TDMA based digital cellular mobile communications in the presence of signal impairments due to ISI and CCI/ACI. In the past few years, a number of algorithms have been suggested. For example in [7] [8], the problem is approached from a classification perspective by exploiting the finite alphabet property of the information sequence. While the assumptions made in their formulations are very general, these algorithms typically require long training sequence. This may render them unsuitable for mobile communications applications where short burst formats are used. Optimum diversity combining algorithms that mitigate the effects of ISI and CCI using minimum mean square criterion were suggested in [15] [16] [21].

With the assumption of prior knowledge of CCI parameters, Wales uses a superstate trellis approach to jointly estimate the data sequences of the signal of interest (SOI) and co-channel interference [12]. In [13] [14], space-time algorithms are proposed for the case of unknown channel. Therein, the SOI and CCI channels and their associated data sequences jointly are identified and detected. Although these algorithms provides optimal solution, they require the SOI and CCI to be bit-synchronous. Also, their receiver structures can become prohibitively complex when the channels are highly dispersive or when the number of dominant CCI is large.

In [18], Bottomley and Jamal proposed a space-only maximum likelihood sequence detection algorithm that uses the second order statistics of the radio frequency interference and noise. The significance of this approach is that the CCI and SOI need not be synchronous nor does it requires prior knowledge of the number of CCI and their modulation waveforms. However it requires the base-stations of the SOI and CCI to be time-slot synchronized. As pointed out in [19], this assumption can easily be met in pico- and micro
cellular applications. Under the ideal condition of Gaussian interference, this approach achieves optimal sequence detection [20].

In most applications of [18], the channel matrix and the interference+noise covariance matrix need to be estimated. In [19], the channel matrix is least square estimated from the training data without assuming the presence of CCI. This is followed by the estimation of interference+noise covariance matrix based on the least squares estimated channel parameters. Under mild conditions, these estimates are asymptotically unbiased. Such asymptotical results are not very useful practically as most TDMA cellular systems use short data formats. When short training sequences are available in order to conserve bandwidth, the estimated channel parameters and interference+noise covariance matrix (as in [19]) become highly biased in the presence of strong CCI. This can degrade the sequence detection performance.

In this paper, we propose a spatial-temporal algorithm in the spirit of [18] [19] [20] whereby the radio frequency interferences are considered as stochastic processes. Central to our approach is the recognition of the following basic ideas.

- Temporal diversity can be gained by observing the channel outputs from a tapped delay line. Geometrically, the temporal diversity can be interpreted as the widening of the noiseless channel outputs.
- The FA properties or discreteness of the information sequence and the inherent algebraic (Toeplitz) structures of the channel output vectors and data sequence can offer additional information to the estimation of the channel parameters and the noise+covariance matrix.

The paper is organized as follows. We first present the data model, assumptions and the problem statement. In the next section, we describe the proposed algorithm based on the ideas described earlier. Simulation results are presented in section III. In section IV, we analyze the complexity of the proposed algorithm. Finally, section V summarizes the paper. A shorter version of this paper will be presented in [34].
II. DATA MODEL AND PROBLEM FORMULATION

The signal received at the $i^{th}$ antenna of a $N$-element antenna array is given by

$$x_i(t) = \sum_k h_i(t - kT_b) s_k + n_i(t).$$

(1)

where $h_i(t)$ is the combined impulse response (having finite support of length $LT_b$) which includes the transmit pulse shaping, multipath radio channel and receiver filter. $s_k$ is the information symbol belonging to a finite alphabet set $\Omega$ and $T_b$ is the symbol duration. Assuming the received signal is oversampled $M$ times of the baud rate and collecting the received samples from $N$-antenna, the channel output vector in presence of co-channel interference can be written as

$$\hat{x}(kT_b) = \Xi \hat{s}_k + \hat{i}(kT_b) + \hat{n}(kT_b)$$

(2)

where

$$\hat{x}(kT_b) = \begin{bmatrix} x_1(kT_b), \ldots, x_1((k + \frac{M-1}{M})T_b), \\
\ldots, x_N(kT_b), \ldots, x_N((k + \frac{M-1}{M})T_b) \end{bmatrix}^T$$

(3)

$$\hat{s}_k = [s_k, \ldots, s_{k-L+1}]^T$$

(4)

$$\hat{n}(kT_b) = \begin{bmatrix} n_1(kT_b), \ldots, n_1((k + \frac{M-1}{M})T_b), \\
\ldots, n_N(kT_b), \ldots, n_N((k + \frac{M-1}{M})T_b) \end{bmatrix}^T.$$  

(5)

$\Xi$ is the channel matrix given by

$$\Xi = \begin{bmatrix} h_1(0) & \cdots & h_1(-(L-1)T_b) \\
h_1(\frac{1}{M}T_b) & \cdots & h_1(-(L-1) + \frac{1}{M})T_b) \\
\vdots & \vdots & \vdots \\
h_N(\frac{M-1}{M}T_b) & \cdots & h_M(-(L-1) + \frac{M-1}{M})T_b) \end{bmatrix}.$$

(6)

The term $\hat{i}(t)$ is due to the radio frequency interference term. In cellular communications where CC1 is the dominating RFI, we can write

$$\hat{i}(t) = \sum_{i=1}^{N_{CC1}} G_i b_i(t)$$

(7)
where $G_i$ and $b_i(t)$ are the channel matrix and symbol vector associated with the $i^{th}$ CCI. Compact expressions relating the channel parameters $\Xi$ to the multipath strength, directions and times of arrival were derived in [9][11] and listed in Appendix I for completeness.

In this paper, we shall assume the following:

- **A1**: The desired and interference channel response remains stationary over the time-slot. This assumption is valid in most GSM radio channels and indoor applications. The extension to time-varying channels is beyond the scope of this paper. However, the approach will be briefly described at the end of this paper.
- **A2**: The desired and interfering signal are time-slot synchronized. This is a reasonable assumption, particularly in micro and picocells wireless applications where time-slot (not symbol) synchronized base-stations can be easily maintained.
- **A3**: The signal of interest and the CCI are uncorrelated.
- **A4**: The interference is a zero mean process.

With the model of (2) in hand, the problem considered in this paper can be succinctly stated as follows:

"Given the sampled channel output vectors collected over the time-slot, determine the channel parameters and the information sequence."

III. PROPOSED CHANNEL IDENTIFICATION AND EQUALIZATION METHOD

The spatial-temporal measurements can be written as

\[ x(kT_b) = C(\Xi)s_k + I(kT_b) + n(kT_b) \]

\[ = C(\Xi)s_k + w(kT_b) \]

(8)

where

\[ x(kT_b) = [\tilde{x}^T(T) \cdots \tilde{x}^T((k-m+1)T_b)]^T \]

(9)

\[ s_k = [s(t) \cdots s(t-L-m+2)]^T \]

(10)

\[ I(kT_b) = [\tilde{i}^T(T) \cdots \tilde{i}^T((k-m+1)T_b)]^T \]

(11)

\[ n(kT_b) = [\tilde{n}^T(kT_b) \cdots \tilde{n}^T((k-m+1)T_b)]^T \]

(12)

\[ w(kT_b) = I(kT_b) + n(kT_b). \]

(13)
Note the \( \mathbf{C}(\Phi) \) is a matrix operator that generates a block Toeplitz matrix from \( \Phi \). The advantage of utilizing spatial temporal measurements will be explained later in this paper. By \( \mathbf{A1} \), the \( N \)-snapshots collected over the time-slot can be written as

\[
\mathbf{X} = \mathbf{C}(\mathbf{\Xi})\mathbf{S} + \mathbf{W}
\]  

(14)

where

\[
\mathbf{X} = [\mathbf{x}(kT_b) \mathbf{x}((k+1)T_b) \cdots \mathbf{x}((k+N-1)T_b)]
\]  

(15)

\[
\mathbf{S} = [\mathbf{s}_k \mathbf{s}_{k+1} \cdots \mathbf{s}_{k+N-1}]
\]  

(16)

\[
\mathbf{W} = [\mathbf{w}(kT_b) \mathbf{w}((k+1)T_b) \cdots \mathbf{w}((k+N-1)T_b)].
\]  

(17)

We begin by assuming the interference+noise to be zero mean Gaussian process of unknown covariance \( \mathbf{R}_{ww} \). The maximum likelihood (ML) estimation of the channel parameter matrix \( \mathbf{\Xi} \) and the symbol matrix \( \mathbf{S} \) can be achieved by optimizing the following likelihood functional:

\[
\{\mathbf{\Xi}, \hat{\mathbf{S}}, \hat{\mathbf{R}}_{ww}\} = \arg \min_{\mathbf{\Xi}, \mathbf{S}, \mathbf{R}_{ww}} \mathbf{V}(\mathbf{\Xi}, \mathbf{S}, \mathbf{R}_{ww})
\]  

(18)

where

\[
\mathbf{V}(\mathbf{\Xi}, \mathbf{S}, \mathbf{R}_{ww}) = \text{Tr}( (\mathbf{C}(\mathbf{\Xi})\mathbf{S} - \mathbf{X})^H \mathbf{R}_{ww}^{-1} (\mathbf{C}(\mathbf{\Xi})\mathbf{S} - \mathbf{X}) ) + N \log |\mathbf{R}_{ww}|.
\]  

(19)

As one may have noticed, the matrix structure and FA properties of the unknown symbol matrix offer useful structural constraint to the problem addressed here. The FA property of the transmitted data limits the noiseless channel outputs to a discrete set and related to the symbol matrix through the channel parameter matrix. The Toeplitz FA symbol matrix limits the transition of the symbol vectors from one time instant to another, henceforth, the transition to another noiseless channel output vector will also be limited to an even smaller set. The interferences may not follow the Gaussian statistics and the \textit{true} likelihood function is not likely to be available in practice. Hence, the likelihood function derived in (18) can only be perceived as an approximation to the \textit{true} likelihood function. Geometrically, the interference and noise can be approximated to be ellipsoidally distributed around the noiseless channel outputs. This approximation has demonstrated to
be useful in similar applications such as [17] [7] [8] [29]. In this case, the joint identification of the channel parameters and the detection of the data sequence can be perceived geometrically as finding “valid” sequence of noiseless channel outputs generated from the hypothesized channel parameter and symbol matrices, and interference+noise covariance that minimize the Mahalanbosis distance from the sequence of observed channel outputs vectors. Mathematically speaking, we have

$$\left\{ \Xi, \hat{\Theta}, \hat{Q} \right\} = \arg\min \text{Tr} \left( (X - C(\Xi)S)^H \hat{Q} (X - C(\Xi)S) \right) + N \log |\hat{Q}|$$

(20)

where \( Q \) is an appropriate weighting matrix describing the ellipsoidal distribution of the interference and noise. The second term in (20) can be perceived as the “regularizing” term since choosing \( Q = 0 \) results in a trivial solution. Naturally, the optimal choice of \( Q \) for Gaussian interference is \( Q = R_{ww}^{-1} \).

The motivation of utilizing the spatial temporal measurements is to exploit the temporal diversity gain from observing the channel outputs through a tapped delay line. We shall illustrate this by the following example. The channel impulse responses of the signal of interest, \( \Xi_{SOI} \), and the co-channel interference, \( \Xi_{CCI} \), are given by

$$\Xi_{SOI}(z) = [0.3 \ 0.8 \ 0.3]$$

(21)

$$\Xi_{CCI}(z) = [0.2 \ 0.9 \ 0.5]$$

(22)

respectively. We assume the transmitted symbols are limited to the FA set \( \Omega \in \{-1, 1\} \).

The Mahalanbosis distance between two channel outputs is defined as

$$d_{i,j} = (c_i - c_j)^H R_{ww}^{-1} (c_i - c_j)$$

(23)

$$c_{(i)} = C(\Xi_{CCI}) s^{(i)}.$$  

(24)

$$R_{ww} = \alpha_{CCI} C(\Xi_{CCI}) C\Xi_{CCI} \sigma_s^2 + \sigma_n^2 I.$$  

(25)

\( s^{(i)} \in \Omega^{m+L-1} \cdot \alpha_{CCI} \) is the level of co-channel interference. \( \sigma_s^2 \) and \( \sigma_n^2 \) are the signal and noise power, respectively. Due to the Toeplitz structure of the symbol matrix, the channel outputs at the next time instant, say from \( x(kT_b) \) to \( x((k + 1)T_b) \), will be limited by the valid transitions of the symbol vectors, \( s_k \) to \( s_{k+1} \). For example, the transition of \( s_k = [1 \ 1 \ 1 \ 1 \ 1]^T \) to \( s_{k+1} \) will be limited to \( [1 \ 1 \ 1 \ 1 \ 1]^T \) and \( [-1 \ 1 \ 1 \ 1 \ 1]^T \). Figure 1
Fig. 1. Temporal Diversity Gain versus Temporal Window Length

illustrates the temporal diversity gain as a function of $m$ with $\alpha_{CCI} = 1$. The normalized temporal diversity gain is defined as the ratio of distance of temporal window $m$ with the distance with $m = 1$. We note from Figure 1 that with increasing $m$, the distance between the noiseless channel outputs becomes larger. However this is a diminishing gain. In this example, the temporal diversity gain diminishes after $m \geq 4$. Figure 2 plots the temporal diversity gain as a function of the $\alpha_{CCI}$ with temporal window fixed at $m = 3$. It shows the advantage of temporal diversity in the presence of strong CCI. Note that the gain diminishes as $\alpha_{CCI} \rightarrow 0$. This indicates that in the absence of CCI, observing the channel output from a tapped delay line will not offer any advantage but further complicate the computational structure of the receiver.

In summary, we approach the problem of joint channel identification and sequence detection by optimizing the likelihood function in (20) with the exploitation of temporal diversity, the inherent matrix structure and the FA property of the unknown symbol matrix. Central to our approach is espoused by the recognition that

- **R1**: Temporal diversity can be gained from observing the channel outputs vectorally from a tapped delay line.
- **R2**: The FA property of the unknown data sequence and its inherent algebraic structure
and relationships with the noiseless channel outputs provide a powerful constraint.

It suffices to note that many TDMA based wireless cellular communications systems utilize short data format. For example, GSM and IS-54 transmits with time-slots approximately equivalent to 150 symbols. With short training sequence length, the employment of asymptotics may no longer be useful. The approach proposed herein can alleviate this problem by utilizing R1 and exploiting R2 to jointly identify the unknown channel parameter, interference+noise covariance and symbol matrix. In particular, the channel parameter and interference+noise covariance matrix are estimated with data not limited to the training sequence but from the intrinsic information embedded entire data packet.

A. An Estimation Algorithm

The likelihood function (18) is a highly nonlinear function of continuous and FA constrained parameters. Such optimization problem can be solved by applying the alternating minimization (AM) procedure. Basically, AM minimizes the cost function by optimizing it with respect to (w.r.t) to a subset of the parameters while the remaining parameters, previously estimated, are held fixed. By cyclicly minimizing w.r.t each parameter subsets, the cost function will monotonically decrease to a local minima. In general, global con-
vergence will depend on the choice of initial estimates. This is a widely used approach for optimizing highly nonlinear multi-value cost function in many applications such as [25][26][27][28].

In this paper, we adapt the AM concept and derive an iterative reweighting alternating minimization (IRAM) algorithm to solve the problem represented in (18). IRAM comprises of the following basic steps to estimate $\mathbf{Z}$, $\mathbf{S}$ and $\mathbf{R}_{ww}$. Beginning with initial estimates of $\mathbf{Z}$ and $\mathbf{R}_{ww}$ (for example, computed from the training sequence), the symbol matrix $\mathbf{S}$ is computed by optimizing

**Step 1**

$$\hat{\mathbf{S}} = \arg \min_{\mathbf{S} \in \mathbb{C}^{(m+L-1) \times N}} \text{Tr} \left( (\mathbf{C}(\hat{\mathbf{Z}})\mathbf{S} - \mathbf{X})^H \mathbf{R}_{ww}^{-1} (\mathbf{C}(\hat{\mathbf{Z}})\mathbf{S} - \mathbf{X}) \right)$$

(26)

subject to $\mathbf{C1}$ and $\mathbf{C2}$.

**Step 2** Next (18) is updated with the estimated symbol matrix and by minimizing $\mathbf{Z}$ w.r.t (18), we have

$$\text{vec} (\hat{\mathbf{Z}}) = \mathbf{P} \hat{\mathbf{S}} \text{vec}(\mathbf{R}_{ww}^{-1} \mathbf{X})$$

(27)

where

$$\mathbf{P} \hat{\mathbf{S}} = \left( \hat{\mathbf{S}}^T \otimes \mathbf{R}_{ww}^{-1} \right) \mathbf{P}$$

(28)

($\cdot)^\dagger$ and $\otimes$ denote pseudo-inverse and kronecker product, respectively. $\text{vec}(\mathbf{A})$ vectorizes $\mathbf{A}$ by stacking its columns into a vector and $\mathbf{P}$ is a full rank selection matrix such that $\text{vec}(\mathbf{C}(\hat{\mathbf{Z}})) = \mathbf{P} \text{vec}(\mathbf{Z})$.

**Step 3** The interference+noise covariance matrix can be updated from the error residual based on the latest estimates of $\mathbf{Z}$ and $\mathbf{S}$ by

$$\hat{\mathbf{W}} = \mathbf{X} - \mathbf{C}(\hat{\mathbf{Z}}) \hat{\mathbf{S}}$$

$$\hat{\mathbf{R}}_{ww} = \frac{1}{N} \hat{\mathbf{W}} \hat{\mathbf{W}}^H.$$  

(29)

After **Step 1** and **2**, $\hat{\mathbf{Z}}$ and $\hat{\mathbf{S}}$ will converge towards, but not necessary to the final converged values. In **Step 3**, $\hat{\mathbf{R}}_{ww}$ is refined thereby conditioning the cost functional towards better estimation of $\mathbf{Z}$ and $\mathbf{S}$. Thus, by alternatingly minimizing (18) with **Step 1** and **2**, and refining the cost function by reweighting with the interference+noise covariance matrix computed in **Step 3**, $\hat{\mathbf{Z}}$ and $\hat{\mathbf{S}}$ will monotonically converge to a local minima solution. A detailed discussion on the convergence is deferred to Appendix II.
**Sequence Detection**

Symbol detection by exploiting the Toeplitz structure of $\mathbf{S}$ can be achieved with the application of the Viterbi algorithm. This can be computationally expensive to implement, in particular with the spatial-temporal channel matrix. For example with $m = 3$, $L = 3$ and $K = 2$ level signalling, the number of states in the trellis will amount to $K^{m+L-1} = 32$. In this paper, the Toeplitz structure of the symbol matrix is slightly relaxed. We use a symbol by symbol detection approach that constrains $\mathbf{S}$ to be partially Toeplitz\cite{5,6} in the following manner:

Given the tentative detection of the $k^{th}$ symbol vector

$$\hat{s}_k = [\hat{s}_k \cdots \hat{s}_{k-L-m+2}]^T$$

and estimated channel matrix $\hat{\mathbf{E}}$, the tentative detection of the $(k+1)^{th}$ symbol vector is given by

$$\hat{s}_{k+1} = \begin{bmatrix} \hat{\phi}_f \\ \hat{\phi}_{fb} \end{bmatrix}$$

where

$$\hat{\phi}_f = \min_{\phi_f \in \mathbf{O}^{mL}} \text{Tr} \left( \mathbf{I}^H \hat{\mathbf{R}}_{\mathbf{wW}}^{-1} \mathbf{I} \right)$$

$$\mathbf{I} = \mathbf{x}((k+1)T_b) - \hat{\mathbf{E}} \left[ \begin{array}{c} \hat{\phi}_f \\ \hat{\phi}_{fb} \end{array} \right].$$

$$\hat{\phi}_{fb} = [\hat{s}_k]_{k-n_{fb}+1 \cdots k-L-m+3,1}$$

with $[\mathbf{A}]_{i \cdots k,l}$ denoting a vector whose elements are extracted from the $l^{th}$ column and $i^{th}$ to $k^{th}$ row of the matrix $\mathbf{A}$.

**Proposed Algorithm**

The proposed method is summarized as follows:

1. The received channel output matrix and the corresponding symbol matrix consist of the information (of length $P$) and training sequence (of length $N^{Tr}$), and are given by $\mathbf{X} = [\mathbf{X}_{Tr} \ \mathbf{X}_{Data}]$ and $\mathbf{S} = [\mathbf{S}_{Tr} \ \mathbf{S}_{Data}]$. 

DRAFT

December 16, 1998
2. Based on previously estimated $\mathbf{R}_{ww}$ and $\mathbf{\Xi}_{WLS}$, estimate $\hat{\mathbf{S}}_{Data}$ with constraints C2 and C3 from

$$
\hat{\mathbf{S}}_{Data} = \arg \min_{\mathbf{S}_{Data}} \text{Tr} \left( \Delta^H \mathbf{R}_{ww}^{-1} \Delta \right)
$$

$$
\Delta = C(\mathbf{\Xi}_{WLS}) \mathbf{S}_{Data} - \mathbf{X}_{Data}
$$

by

- For $k = 0 \ldots P - 1$
- Estimate $\hat{s}_{k+1} = [\phi_j^T \phi_{f_0}^T]^T$ from (33) with $\mathbf{\Xi} = \mathbf{\hat{\Xi}}_{WLS}$ and $\mathbf{R}_{ww} = \mathbf{\hat{R}}_{ww}$.
- End

3. Estimate $\mathbf{\Xi}_{WLS}$ from $\mathbf{X}$ with constraint C1 by

$$
\text{vec}(\mathbf{\hat{\Xi}}_{WLS}) = \Phi^\dagger(\hat{\mathbf{S}}) \text{vec}(\mathbf{\hat{R}}_{ww}^{-1} \mathbf{X})
$$

4. Compute

$$
\mathbf{\hat{W}} = \mathbf{X} - C(\mathbf{\hat{\Xi}}_{WLS}) \hat{\mathbf{S}}
$$

and update

$$
\mathbf{\hat{R}}_{ww} = \frac{1}{N} \mathbf{\hat{W}} \mathbf{\hat{W}}^H.
$$

5. Repeat step 3-5 until convergence.

The channel parameters $\mathbf{\hat{\Xi}}_{WLS}$ and the interference+noise covariance can be initialized with its least square estimates from the training data that assumes the absence of CCI:

$$
\text{vec}(\mathbf{\hat{\Xi}}_{WLS}) = (\mathbf{S}_{Tr} \otimes \mathbf{I}) \text{vec}(\mathbf{X}_{Tr})
$$

where $\mathbf{X}_{Tr}$ are the channel output vectors due to the training preamble $\mathbf{S}_{Tr}$. The interference+noise covariance, $\mathbf{R}_{ww}$, is then computed as in (36) and (37).

A similar procedure is used in [19] to estimate the channel parameters and the interference+noise covariance from the training data\(^1\). The estimates of $\mathbf{\Xi}$ and $\mathbf{R}_{ww}$ are asymptotically unbiased. When $N^{Tr} \to \infty$, we have

$$
E \left( \mathbf{\hat{\Xi}}_{LS} \right) = E \left( \mathbf{X S}^\dagger \right)
$$

$$
= E \left( (\mathbf{\Xi S} + \mathbf{W}) \mathbf{S}^\dagger \right)
$$

$$
= \mathbf{\Xi} + E \left( \mathbf{W S}^\dagger \right) = \mathbf{\Xi}.
$$

\(^1\)The approach used in [19] did not exploit temporal diversity
The equality results from A3 and 4 where

$$E\left(WS^\dagger\right) = E(W) E\left(S^\dagger\right) = 0. \quad (40)$$

Hence the estimated interference+noise covariance, $\hat{R}_{uw}$, is also asymptotically unbiased. As remarked earlier, such asymptotical results are not useful in many mobile applications as the time-slot of the data packets are very short. Hence, incorporating long training sequence will severely limit the system’s spectral efficiency. On the other hand, if $\Xi$ and $R_{uu}$ are not determined accurately, the estimation errors can result in poor sequence detection.

IV. SIMULATION EXAMPLES

We present the results of simulations in this section. We use the TU channel profile described in [22] and assume a two element antenna array. To illustrate the core ideas, we shall assume the data to be BPSK, $\Omega \in \{-1, 1\}$, transmitting at bit rate of 277kbs. The fading is assumed to be independent between paths and held constant over each burst. However, they are varied independently from burst to burst. The following parameters are fixed unless otherwise stated. We set training and data sequence length to be $N^{Tr} = 15$ and $P = 150$, respectively. We assume the SOI and a CCI impinge the antenna array from mean direction of arrival of 0 and 60 degrees. The multipath angles are randomly generated. We also assume that each signal path (angle) is angularly spread with distribution of $N(0, 30)$ due to local scatterers. The channel parameter matrix is generated based on the structured channel model described in Appendix I. The signal to noise ratio is 10dB.

In Figure 3, we show the BER achieved by IRAM with CIR ranging from $-15dB$ to 7dB. Algorithms based on MMSE criterion[30], interference ratio combining (IRC)[18] and hybrid 2-Stage approach[31] are used to benchmark the proposed algorithm. The results are averaged directly over 5000 independent trials. The performance gains of the proposed algorithm are encouraging. For example, at 2% raw BER, IRAM achieved 5dB gain over IRC. Figure 3 also displays the BER performance of the conventional Viterbi based sequence detector. It clearly demonstrates the significant performance loss due to the presence of CCI. Note that the MMSE linear equalizer outperforms the conventional Viterbi sequence detector over the range of CIR. This is due to the dominance of the CCI
Fig. 3. BER versus SNR. ×: IRAM, □: Hybrid 2-Stage Approach, …: IRC, *: MLSD and +: MMSE

over ISI.

The convergence properties of the proposed algorithm is examined next. In Figure 4, we plot a typical example of the cost function, channel estimation error and the number of erroneous detection as a function of iterations. Notice the incremental improvement of the channel estimates and sequence detection with the number of iterations.

In Figure 5, we examine the BER performance of IRAM and IRC with exact and estimated initialization over a range of CIR. It is interesting to note that IRAM with estimated channel parameter matrix and interference+noise covariance has similar performance compared to IRC with exact initialization. It also suffices to note that significant performance can be gained (≈ 5dB) with better estimation of the channel parameters and the interference+noise covariance matrix.

Figure 6 displays the BER performance of the IRAM and IRC as a function of training length overhead, $N_T^r$. This graph demonstrates the importance of good initial estimates of the channel parameter matrix and interference+noise covariance. We observe an order of magnitude of gain is achieved in IRAM when the training length is 50 as compared to 15. This result shows that better initial estimates lead to better BER performance suggesting the need for more data efficient approaches of estimating these parameters. One approach
to achieve this is to exploit possible priors like the modulation waveforms first proposed in [32] and later in [33]. Another approach is to apply the parametric channel estimator suggested in [9][10][3]. If permitted at system level, the training waveforms of SOI and CCI may be made as orthogonal as possible to allow for better estimation of the channel parameter matrix and the interference+noise covariance matrix.

Figure 7 shows the BER performance of the IRAM improves with the length of the information sequence, \( P \). The result indicates that little improvement in performance can be achieved, perhaps only when the information sequence is extremely long. In Figure 8, we investigate the effects of direction of arrival of the CCI. The results show that the performance of the IRAM and IRC remained somewhat constant over principal quadrant. This may not be the case for DF-beamformer approaches.

V. CONCLUDING REMARKS

In this paper, we consider the problem of channel identification and equalization in the presence of strong CCI. We propose an algorithm that identifies and equalizes ISI channels in the presence of strong multipath based on a least squares framework. The proposed algorithm does not require knowledge of the number of interferers nor their structural information but requires time-slot synchronization. This requirement however can be easily achieved in cellular systems using small cells such as micro- and pico-cells. The potential performance gains suggested by the simulation results have been encouraging.

APPENDIX

I. STRUCTURED CHANNEL MODEL

The multipath channel can be approximated and modelled parametrically by a finite number of signal paths described by path gain, direction and time of arrival. The noiseless channel outputs sampled \( M \) times of the symbol rate is given by

\[
x(kT_b) = \sum_{i=0}^{m} \beta_i(kT_b) \left( a(\theta_i) \otimes G^T(\tau_i) \right) s_k
\]

where

\[
x(kT_b) = [x_1(kT_b), \ldots, x_1((k + \frac{M-1}{M})T_b), \ldots, x_n((k + \frac{1}{M})T_b) \cdots x_n((k + \frac{M-1}{M})T_b)]^T
\]
\[ G(\tau) = \left[ g(0, \tau), \ldots, g\left(\frac{M-1}{T_b}, \tau\right) \right] \]
\[ g(t, \tau) = [g(t - \tau), g(t - T_b - \tau), \ldots, g(t - (L-1)T_b - \tau)]^T \]
\[ s_k = [s_k, \ldots, s_{k-1} \cdots s_k-L+1]^T \]

\( x_i(kT_b) \) is the sampled output of the \( i^{th} \) array channel output. \( \mathbf{a}(\theta_i) \) is the antenna array response to a signal from direction \( \theta_i \). \( \beta_i(t) \) and \( \tau_i \) are the complex gain and delay of the \( i^{th} \) signal path. Expressed compactly in matrix form

\[ \mathbf{x}(kT_b) = \mathbf{U}(\eta) (\mathbf{\beta}(kT_b) \otimes \mathbf{I}_L) \mathbf{s}_k \]

where

\[ \mathbf{U}(\eta) = [\mathbf{u}(\theta_1, \tau_1), \ldots, \mathbf{u}(\theta_m, \tau_m)] \]
\[ \mathbf{u}(\theta_i, \tau_i) = \mathbf{a}(\theta_i) \otimes \mathbf{G}^T(\tau_i) \]
\[ \mathbf{\beta}(t) = [\beta_1(t) \cdots \beta_m(t)]^T \]
\[ \eta = [\theta_1, \ldots, \theta_m, \tau_1, \ldots, \tau_m]^T. \]

\( \mathbf{I}_L \) denotes the identity matrix of dimension \( L \).

II. Convergence Analysis

Let \( \mathbf{S}_c \) and \( \mathbf{\Xi}_c \) be the solution of a local minima. Also, let \( \mathbf{S}^{(i)}, \mathbf{\Xi}^{(i)} \) and \( \mathbf{Q}^{(i)} \) denote, respectively, the estimated symbol matrix, channel matrix and interference+noise covariance after the \( i^{th} \) iteration. In the next iteration, we have

\[ \mathbf{S}^{(i+1)} = \min \ \text{Tr} \left( (\mathbf{C}(\mathbf{\Xi}^{(i)}) \mathbf{S} - \mathbf{X})^H \mathbf{\tilde{Q}}^{(i)} (\mathbf{C}(\mathbf{\Xi}^{(i)}) \mathbf{S} - \mathbf{X}) \right) \]

and followed by

\[ \mathbf{\Xi}^{(i+1)} = \min \ \text{Tr} \left( (\mathbf{C}(\mathbf{\Xi}) \mathbf{S}^{(i+1)} - \mathbf{X})^H \mathbf{\tilde{Q}}^{(i)} (\mathbf{C}(\mathbf{\Xi}) \mathbf{S}^{(i+1)} - \mathbf{X}) \right). \]

We have

\[ |\mathbf{S}_c - \mathbf{S}^{(i+1)}| \leq |\mathbf{S}_c - \mathbf{S}^{(i)}| \] (51)
\[ |\mathbf{\Xi}_c - \mathbf{\Xi}^{(i+1)}| \leq |\mathbf{\Xi}_c - \mathbf{\Xi}^{(i)}|. \] (52)
The matrix $Q^{(i+1)}$ is updated by

$$Q^{(i+1)} = \left( \frac{1}{N} (C(\hat{S}^{(i+1)})S^{(i+1)} - X)(C(\hat{S}^{(i+1)})S^{(i+1)} - X)^H \right)^{-1}. \quad (53)$$

It follows that

$$|Q_e - Q^{(i+1)}| \leq |Q_e - Q^{(i)}|. \quad (54)$$

From (52), (52) and (54), the estimated $\hat{S}$, $\hat{E}$ and $\hat{Q}$ by iterative reweighted alternating minimization will approach their corresponding converged values.

**References**


Fig. 4. (a) Detection Error versus Iterations, (b) Cost Function versus Iterations
Fig. 5. BER versus SNR. ×: IRAW with exact initialization, o: IRAW with estimated initialization, +: IRC with exact initialization, *: IRC with estimated initialization.

Fig. 6. BER versus SNR. ×: IRAW and +: IRC
**Fig. 7.** BER versus Information Sequence Length. ×: IRAM and +: IRC

**Fig. 8.** BER versus Offset Angle. ×: IRAM