

**MASTER COPY:** PLEASE KEEP THIS "MEMORANDUM OF TRANSMITTAL" BLANK FOR REPRODUCTION PURPOSES. WHEN REPORTS ARE GENERATED UNDER THE ARO SPONSORSHIP, FORWARD A COMPLETED COPY OF THIS FORM WITH EACH REPORT SHIPMENT TO THE ARO. THIS WILL ASSURE PROPER IDENTIFICATION. NOT TO BE USED FOR INTERIM PROGRESS REPORTS; SEE PAGE 2 FOR INTERIM PROGRESS REPORT INSTRUCTIONS.

**MEMORANDUM OF TRANSMITTAL**

U.S. Army Research Office  
ATTN: AMSRL-RO-BI (TR)  
P.O. Box 12211  
Research Triangle Park, NC 27709-2211

Reprint (Orig + 2 copies)

Technical Report (Orig + 2 copies)

Manuscript (1 copy)

Final Progress Report (Orig + 2 copies)

Related Materials, Abstracts, Theses (1 copy)

CONTRACT/GRANT NUMBER:

REPORT TITLE:

is forwarded for your information.

SUBMITTED FOR PUBLICATION TO (applicable only if report is manuscript):

Sincerely,

# REPORT DOCUMENTATION PAGE

Form Approved  
OMB NO. 0704-0188

Public Reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)		2. REPORT DATE <b>May 2006</b>	3. REPORT TYPE AND DATES COVERED <b>Manuscripts 01 June 2005 – 31 July 2006</b>	
4. TITLE AND SUBTITLE <b>Multiple-Pass Decision-Directed Channel Estimation for Highly Mobile MIMO Communications</b>			5. FUNDING NUMBERS <b>W911NF0410224</b>	
6. AUTHOR(S) <b>Michael Larsen and A. Lee Swindlehurst</b>				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Brigham Young University Provo, UT 84604</b>			8. PERFORMING ORGANIZATION REPORT NUMBER <b>N/A</b>	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) <b>U. S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211</b>			10. SPONSORING / MONITORING AGENCY REPORT NUMBER <b>N/A</b>	
11. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.				
12 a. DISTRIBUTION / AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited.</b>			12 b. DISTRIBUTION CODE <b>N/A</b>	
13. ABSTRACT (Maximum 200 words) <p>In multiple-input multiple-output (MIMO) wireless communication systems employing coherent detection, good performance depends on the availability of accurate channel estimates. This is a significant challenge in mobile MIMO systems where the channel coefficients change too rapidly for the effective use of traditional training methods. In this paper, we present a multiple-pass decision-directed scheme for channel estimation and symbol decoding in scenarios for which the MIMO channel varies rapidly. In this scheme, pilot symbols are embedded over a block of data symbols at the transmitter, and the receiver uses this pilot information to iteratively optimize the channel and symbol estimates over the entire symbol block. The optimization is carried out over a posterior distribution which incorporates information about the channel model and noise properties. The conjugate gradient algorithm is used to jointly optimize the channel estimates over the block. Simulations show the benefits of using embedded training and demonstrate the utility of this multiple-pass method.</p>				
14. SUBJECT TERMS <b>N/A</b>			15. NUMBER OF PAGES <b>5</b>	
			16. PRICE CODE <b>N/A</b>	
17. SECURITY CLASSIFICATION OR REPORT <b>UNCLASSIFIED</b>	18. SECURITY CLASSIFICATION ON THIS PAGE <b>UNCLASSIFIED</b>	19. SECURITY CLASSIFICATION OF ABSTRACT <b>UNCLASSIFIED</b>	20. LIMITATION OF ABSTRACT <b>UL</b>	

# Multiple-Pass Decision-Directed Channel Estimation for Highly Mobile MIMO Communications

Michael Larsen and A. Lee Swindlehurst  
Department of Electrical and Computer Engineering  
Brigham Young University, 459 CB, Provo, Utah 84602  
Telephone: (801) 422-5349  
Fax: (801) 422-0201  
Email: mikedl@byu.edu

**Abstract**—In multiple-input multiple-output (MIMO) wireless communication systems employing coherent detection, good performance depends on the availability of accurate channel estimates. This is a significant challenge in mobile MIMO systems where the channel coefficients change too rapidly for the effective use of traditional training methods. In this paper, we present a multiple-pass decision-directed scheme for channel estimation and symbol decoding in scenarios for which the MIMO channel varies rapidly. In this scheme, pilot symbols are embedded over a block of data symbols at the transmitter, and the receiver uses this pilot information to iteratively optimize the channel and symbol estimates over the entire symbol block. The optimization is carried out over a posterior distribution which incorporates information about the channel model and noise properties. The conjugate gradient algorithm is used to jointly optimize the channel estimates over the block. Simulations show the benefits of using embedded training and demonstrate the utility of this multiple-pass method.

## I. INTRODUCTION

In this paper, we consider the problem of channel estimation and symbol decoding in mobile multiple-input multiple-output (MIMO) wireless communication systems. In order for the receiver in a MIMO system to correctly determine what symbols were sent, the system requires an estimate of the channel, either explicitly or implicitly. Most techniques used for channel estimation in MIMO systems rely on the assumption that the channel varies slowly enough so that it may be considered constant over a block of symbols. If this assumption holds, then known pilot symbols sent from the transmitter to the receiver may be used to estimate the channel for the block. Once the channel is known, the remaining data symbols in the block may be decoded. In pilot symbol assisted modulation (PSAM), the most common form of training, the pilot symbols are typically sent at the beginning of each block or at regular intervals over the block. Recently, it has been shown that improved performance can be obtained by techniques that superimpose training data throughout the data block [1]–[3]. A generalized approach to this problem that includes PSAM as a special case is the pilot-embedded data-bearing (PEDB)

This work was supported by the U.S. Army Research Office under the Multi-University Research Initiative (MURI) grant W911NF-04-1-0224, by the National Science Foundation under grant CCF-0313056, and by a National Science Foundation Graduate Research Fellowship.

algorithm [4], in which orthogonal transformations are used to separately encode the data and pilot symbols.

In mobile environments, the channel may vary too rapidly for a block-constant assumption to apply. When symbol decoding techniques are used that rely on block-type training with such channels, poor performance results. Such situations require methods that estimate the time-varying channel more precisely. In [5], the MIMO channel is decomposed into SISO subchannels whose variation are estimated using SISO techniques. However, the method only works on extremely slow fading rates. Other methods, such as [6]–[8], initially estimate the channel using training and then seek to track the channel using Kalman or recursive least-squares (RLS) filters. In [9], a multiple-pass technique is developed for MIMO orthogonal frequency division multiplexing (OFDM) in which two sets of Kalman filters are used to iteratively estimate the channel using pilot information. In this work, we iteratively estimate the transmitted symbols and time-varying MIMO channel using multiple passes of a decision-directed method over the data block. While the use of decision direction in this context is not new (*e.g.*, see [9]–[12]), the novelty here is that we use it in conjunction with superimposed training to estimate the channel at each symbol time. In addition, since direct estimation of the channel samples (*e.g.*, via least-squares) is prohibitively expensive, we employ an approximation to the posterior density and the conjugate gradient algorithm to find a tractable solution.

The remainder of this paper is organized as follows. In Section II, we give the system model and review the PEDB training approach. In Section III, we explicitly formulate the optimization problem, and in Section IV we present a multiple-pass decision-directed solution. Simulation results and concluding remarks are presented in Sections V and VI.

## II. BACKGROUND

### A. System Model

Consider a point-to-point narrow-band MIMO system with  $n_T$  transmit antennas and  $n_R$  receive antennas, whose output is mathematically represented as

$$\mathbf{y}_{k,n} = \mathbf{H}_{k,n} \mathbf{u}_{k,n} + \mathbf{n}_{k,n} \quad (1)$$

where  $k$  is the block index and  $n$  is the symbol time index within the block. If  $M$  is the number of symbol times per block, then the overall discrete sample time is  $t = Mk + n$ . The vector  $\mathbf{u}_{k,n}$  is the  $n_T \times 1$  transmit symbol vector,  $\mathbf{y}_{k,n}$  is the  $n_R \times 1$  received data vector,  $\mathbf{H}_{k,n}$  is the  $n_R \times n_T$  channel matrix, and  $\mathbf{n}_{k,n}$  is an  $n_R \times 1$  additive noise vector. The noise is assumed to be zero-mean and independent of the signal.

### B. Pilot-Embedding Structure

In this work, we will attempt to estimate the channel at all  $M$  symbols times within each symbol block, rather than using a single block-based estimate. In particular, we will exploit the presence of superimposed training data and demonstrate its benefits in obtaining symbol-by-symbol channel estimates. We will couch our development in terms of the general PEDB framework of [4], which we briefly introduce below.

If the channel  $\mathbf{H}_{k,n}$  is initially considered to be quasi-static, i.e., the channel is approximately constant over a block of  $M$  symbol times, then the following model results:

$$\mathbf{Y}_k = [ \mathbf{y}_{k,0} \quad \mathbf{y}_{k,1} \quad \cdots \quad \mathbf{y}_{k,M-1} ] = \mathbf{H}_k \mathbf{U}_k + \mathbf{N}_k \quad (2)$$

where  $\mathbf{Y}_k$  and  $\mathbf{N}_k$  are  $n_R \times M$  matrices, and  $\mathbf{U}_k$  is an  $n_T \times M$  matrix. Using the general PEDB training method of [4], the transmitted symbol block  $\mathbf{U}_k$  is a combination of pilot and data symbols given by

$$\mathbf{U}_k = \mathbf{S}_k \mathbf{A} + \mathbf{P}, \quad (3)$$

where  $\mathbf{S}_k$  is an  $n_T \times N_D$  matrix of data symbols,  $\mathbf{A}$  is the  $N_D \times M$  data spreading matrix, and  $\mathbf{P}$  is an  $n_T \times M$  pilot spreading matrix. Note that  $N_P = M - N_D$  is the effective number of training symbol vectors. The matrices  $\mathbf{A}$  and  $\mathbf{P}$  are constrained to have the following properties:

$$\mathbf{A}\mathbf{P}^T = \mathbf{0}, \quad \mathbf{A}\mathbf{A}^T = \beta\mathbf{I}, \quad \mathbf{P}\mathbf{P}^T = \alpha\mathbf{I}, \quad (4)$$

where  $\beta$  and  $\alpha$  are real-valued nonnegative variables used to control the amount of power allocated to the symbol and pilot portions of  $\mathbf{U}_k$ . Using (3), (2) may be written as

$$\mathbf{Y}_k = \mathbf{H}_k(\mathbf{S}_k \mathbf{A} + \mathbf{P}) + \mathbf{N}_k. \quad (5)$$

Using the properties of (4), the channel may be estimated by post-multiplying  $\mathbf{Y}_k$  by  $\mathbf{P}^T$  and dividing by  $\alpha$  so that

$$\frac{\mathbf{Y}_k \mathbf{P}^T}{\alpha} = \mathbf{H}_k + \frac{\mathbf{N}_k \mathbf{P}^T}{\alpha}. \quad (6)$$

If we post-multiply  $\mathbf{Y}_k$  by  $\mathbf{A}^T$  and divide by  $\beta$ , then

$$\frac{\mathbf{Y}_k \mathbf{A}^T}{\beta} = \mathbf{H}_k \mathbf{S}_k + \frac{\mathbf{N}_k \mathbf{A}^T}{\beta}, \quad (7)$$

and standard symbol detection via maximum likelihood decoding or VBLAST can be used to solve for the data. PSAM can be seen as a special case of this pilot-embedding technique when  $\mathbf{A}$  and  $\mathbf{P}$  are chosen so that  $\mathbf{U}_k = \mathbf{S}_k \mathbf{A} + \mathbf{P} = [ \mathbf{P}' \quad \mathbf{S}_k ]$ , where  $\mathbf{P}'$  is an  $n_T \times N_P$  matrix of pilot symbols.

In situations where the channel is rapidly varying, the quasi-static assumption is no longer applicable. Instead, the received symbols of (1) are given by

$$\mathbf{y}_{k,n} = \mathbf{H}_{k,n}(\mathbf{S}_k \mathbf{a}_n + \mathbf{p}_n) + \mathbf{n}_{k,n} \quad (8)$$

where  $\mathbf{p}_n$  and  $\mathbf{a}_n$  are the  $n$ -th columns of the  $\mathbf{P}$  and  $\mathbf{A}$  matrices, respectively. Equations (6) and (7) may still be used to estimate the channel and data, but the results are not necessarily reliable since the block constant assumption has been violated so that the data and pilot symbols can no longer be completely separated. Note that even though the channel estimate given by (6) is only a single channel estimate for the entire block, an appropriate choice of  $\mathbf{A}$  and  $\mathbf{P}$  often results in this channel estimate being more representative of the entire channel than a PSAM estimate. We will use the PEDB estimates as a starting point for the multiple-pass approach we present later.

For notational simplicity, we will omit the block subscript  $k$  from our equations for the remainder of this paper. All operations will continue to be over the  $k$ -th block of transmitted symbols and/or received data.

## III. THE OPTIMIZATION PROBLEM

Let  $\mathbf{h} = [ \mathbf{h}_1^T \quad \cdots \quad \mathbf{h}_M^T ]^T$ ,  $\mathbf{y} = [ \mathbf{y}_1^T \quad \cdots \quad \mathbf{y}_M^T ]^T$ , and  $\mathbf{s} = \text{vec}(\mathbf{S})$ . Our goal is to jointly estimate  $\mathbf{h}$  and decode  $\mathbf{s}$  using the receiver measurements  $\mathbf{y}$ . To do this, we will construct a posterior probability distribution in  $\mathbf{s}$  and  $\mathbf{h}$  and then find the  $\mathbf{s}$  and  $\mathbf{h}$  that maximize this pdf. The posterior pdf will be based on the following assumptions about the channel and symbols.

### A. Channel Model

It is assumed that the channel evolves continuously in time. This time evolution is driven by movement of the transmit or receive antenna arrays and by changes in the environment over which the transmitted signal travels. We represent this time evolution of the channel in our discrete-time space with the first-order autoregressive (AR) model

$$\mathbf{h}_n = \text{vec}(\mathbf{H}_n) = \mathbf{F}\mathbf{h}_{n-1} + \mathbf{q}_n \quad (9)$$

where  $\mathbf{q}_n$  represents the deviations of the channel from the predicted value. For simplicity,  $\mathbf{q}_n$  is assumed to be a zero-mean proper complex Gaussian random vector whose components are independent from each other and in time. Based on (9),  $\mathbf{h}$  is a zero-mean Gaussian vector with covariance matrix

$$\mathbf{C}_H = E[\mathbf{h}\mathbf{h}^H] = \begin{bmatrix} E[\mathbf{h}_1\mathbf{h}_1^H] & \cdots & E[\mathbf{h}_1\mathbf{h}_M^H] \\ \vdots & \ddots & \vdots \\ E[\mathbf{h}_M\mathbf{h}_1^H] & \cdots & E[\mathbf{h}_M\mathbf{h}_M^H] \end{bmatrix} \quad (10)$$

whose individual expectations are given by

$$E[\mathbf{h}_m\mathbf{h}_n^H] = \sum_{i=0}^{\min(m,n)} \mathbf{F}^{m-i} \mathbf{Q}_i (\mathbf{F}^H)^{n-i}, \quad (11)$$

where  $\mathbf{Q}_i = E[\mathbf{q}_i\mathbf{q}_i^H]$ .

## B. Data Model

It is assumed that the transmitted symbols are drawn uniformly and independently from a finite symbol alphabet of size  $N_c$ . The symbol pdf is, therefore, given as

$$\rho(\mathbf{s}) = \prod_{i=1}^{N_D n_T} \left( \frac{1}{N_c} \right) \sum_{j=1}^{N_c} \delta(s^i - s_c^j) \quad (12)$$

where  $s^i$  is the  $i$ -th symbol of the symbol vector  $\mathbf{s}$  and  $s_c^j$  is the  $j$ -th symbol constellation point.

From (8) we know that  $\mathbf{y} = \mathbf{g}(\mathbf{s}, \mathbf{h}) + \mathbf{n}$  where

$$\mathbf{g}(\mathbf{s}, \mathbf{h}) = \begin{bmatrix} \mathbf{H}_1(\mathbf{S}\mathbf{a}_1 + \mathbf{p}_1) \\ \vdots \\ \mathbf{H}_M(\mathbf{S}\mathbf{a}_M + \mathbf{p}_M) \end{bmatrix}. \quad (13)$$

Using the vectorization identity

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B}), \quad (14)$$

where  $\otimes$  denotes the Kronecker product, we can write  $\mathbf{g}(\mathbf{s}, \mathbf{h})$  linearly in terms of  $\mathbf{h}$  as

$$\mathbf{g}(\mathbf{s}, \mathbf{h}) = \begin{bmatrix} (\mathbf{S}\mathbf{a}_1 + \mathbf{p}_1)^T \otimes \mathbf{I}_{n_R} & & \\ & \ddots & \\ & & (\mathbf{S}\mathbf{a}_M + \mathbf{p}_M)^T \otimes \mathbf{I}_{n_R} \end{bmatrix} \mathbf{h} \equiv \mathbf{B}(\mathbf{s})\mathbf{h}. \quad (15)$$

Also, we can write  $\mathbf{g}(\mathbf{s}, \mathbf{h})$  linearly in terms of  $\mathbf{s}$  as

$$\mathbf{g}(\mathbf{s}, \mathbf{h}) = \begin{bmatrix} \mathbf{a}_1^T \otimes \mathbf{H}_1 \\ \vdots \\ \mathbf{a}_M^T \otimes \mathbf{H}_M \end{bmatrix} \mathbf{s} + \begin{bmatrix} \mathbf{H}_1 \mathbf{p}_1 \\ \vdots \\ \mathbf{H}_M \mathbf{p}_M \end{bmatrix} \equiv \mathbf{P}(\mathbf{h})\mathbf{s} + \mathbf{r}(\mathbf{h}). \quad (16)$$

Thus,  $\mathbf{g}(\mathbf{s}, \mathbf{h})$  is bilinear in  $\mathbf{s}$  and  $\mathbf{h}$ .

## C. The Posterior Distribution

Using the channel and symbol assumptions given above, we can use the theory of [13] to write the posterior distribution as

$$\rho(\mathbf{s}, \mathbf{h} | \mathbf{y}) = K e^{-\frac{1}{2}S(\mathbf{s}, \mathbf{h})} \prod_{i=1}^{N_D n_T} \left( \frac{1}{N_c} \right) \sum_{j=1}^{N_c} \delta(s^i - s_c^j) \quad (17)$$

where the misfit function  $S(\mathbf{s}, \mathbf{h})$  is given by

$$S(\mathbf{s}, \mathbf{h}) = (\mathbf{y} - \mathbf{g}(\mathbf{s}, \mathbf{h}))^H \mathbf{C}_D^{-1} (\mathbf{y} - \mathbf{g}(\mathbf{s}, \mathbf{h})) + \mathbf{h}^H \mathbf{C}_H^{-1} \mathbf{h}, \quad (18)$$

$K$  is a normalizing constant, and  $\mathbf{C}_D$  is the covariance matrix for the additive noise of (8). Direct maximization of (17) is prohibitively expensive even for a moderate number of symbols per block. The pdf is not unimodal, so gradient-based optimization techniques cannot guarantee correct results. However, with a ‘‘good enough’’ starting point, the unimodal assumption may still be reasonable for such algorithms. To further simplify the problem, we will also relax the finite alphabet constraint on the symbols. Instead of maximizing (17), we seek to maximize

$$\rho(\mathbf{s}, \mathbf{h} | \mathbf{y}) = K e^{-\frac{1}{2}S(\mathbf{s}, \mathbf{h})}. \quad (19)$$

## IV. MULTIPLE-PASS ALGORITHM

To find the maximum of (19), we will use the gradient to find the minimum of the misfit function. The posterior pdf has an exponential form, so minimizing the misfit function is equivalent to maximizing the posterior. Because the misfit function of our modified posterior may be written linearly in terms of  $\mathbf{h}$  or  $\mathbf{s}$  using (15) and (16), respectively, we can find the gradient of  $S(\mathbf{s}, \mathbf{h})$  with respect to  $[\mathbf{h}^H \ \mathbf{s}^H]^T$  as

$$\nabla S = \begin{bmatrix} \mathbf{C}_H & \mathbf{0} \\ \mathbf{0} & E_s \mathbf{I} \end{bmatrix} \begin{bmatrix} \frac{\partial S}{\partial \mathbf{h}^*} \\ \frac{\partial S}{\partial \mathbf{s}^*} \end{bmatrix} \quad (20)$$

where  $E_s$  is the symbol energy (assuming all symbols have equal energy),

$$\frac{\partial S}{\partial \mathbf{h}^*} = -\mathbf{B}^H \mathbf{C}_D^{-1} (\mathbf{y} - \mathbf{B}\mathbf{h}) + \mathbf{C}_H^{-1} \mathbf{h}, \quad (21)$$

and

$$\frac{\partial S}{\partial \mathbf{s}^*} = -\mathbf{P}^H \mathbf{C}_D^{-1} (\mathbf{y} - \mathbf{r}) + \mathbf{P}^H \mathbf{C}_D^{-1} \mathbf{P}\mathbf{s}. \quad (22)$$

Setting the gradient equal to zero, we must find the simultaneous solution of the following two equations:

$$(\mathbf{B}^H \mathbf{C}_D^{-1} \mathbf{B} + \mathbf{C}_H^{-1})\mathbf{h} = \mathbf{B}^H \mathbf{C}_D^{-1} \mathbf{y} \quad (23)$$

$$(\mathbf{P}^H \mathbf{C}_D^{-1} \mathbf{P})\mathbf{s} = \mathbf{P}^H \mathbf{C}_D^{-1} (\mathbf{y} - \mathbf{r}). \quad (24)$$

Instead of solving (23) and (24) jointly, our approach is to alternate between solving (23) for  $\mathbf{h}$  and (24) for  $\mathbf{s}$  until convergence. Such solutions are common in the literature, but here we determine different channel estimates at each symbol time using the temporal correlation properties of the channel from (9) and (10). Equation (24) may be solved directly in an efficient manner using the equation

$$\mathbf{y} = \begin{bmatrix} \mathbf{H}_1 & & \\ & \ddots & \\ & & \mathbf{H}_M \end{bmatrix} \text{vec}(\mathbf{S}\mathbf{A} + \mathbf{P}). \quad (25)$$

This equation is solved by removing the effects of each  $\mathbf{H}_i$  individually, unvectorizing the overall result, and applying  $\mathbf{A}^T$  to the right.

Equation (23) is more difficult to solve directly since the matrix  $(\mathbf{B}^H \mathbf{C}_D^{-1} \mathbf{B} + \mathbf{C}_H^{-1})$  may be quite large. Instead, we will use the linear conjugate gradient (CG) algorithm to iteratively approximate the solution to this equation [13]. The CG algorithm requires the gradient of the misfit function and a step size parameter  $\alpha$  for each iteration. We already have the gradient information we need in (20). The step size parameter  $\alpha$  is found by solving  $\alpha = \arg \min_{\alpha > 0} S(\mathbf{s}, \mathbf{h} + \alpha \mathbf{p})$ , where  $\mathbf{p}$  is the negative conjugate gradient vector. Using the Kuhn-Tucker conditions, we find

$$\alpha = \max \left\{ 0, \frac{\text{Re}\{(\mathbf{y} - \mathbf{B}\mathbf{h})^H \mathbf{C}_D^{-1} \mathbf{B}\mathbf{p} - \mathbf{h}^H \mathbf{C}_H^{-1} \mathbf{p}\}}{\mathbf{p}^H (\mathbf{B}^H \mathbf{C}_D^{-1} \mathbf{B} + \mathbf{C}_H^{-1}) \mathbf{p}} \right\}. \quad (26)$$

At first glance, the number of computations required to solve (21) and (26) may look daunting. However, with uncorrelated noise,  $\mathbf{C}_D$  is diagonal, and  $\mathbf{C}_H$  and  $\mathbf{B}$  often have structure that can be taken advantage of to reduce the amount of computation required.

## V. SIMULATION RESULTS

In this section, we present simulation results for the multiple-pass decision-directed algorithm presented above. In all of the simulations, the transmitted symbols are unit-energy quadrature phase-shift keying (QPSK) symbols. The channels are generated using a ray-based single-bounce physical model, although the algorithm is implemented without any specific information about the physical channel by setting  $\mathbf{Q}_i = \sigma_q^2 \mathbf{I}$  and  $\mathbf{F} = \mathbf{I}$ . Improved performance can be obtained by estimating these quantities along with the channel and data symbols. The receiver antenna array moves at 2.64 wavelengths per second with 1000 samples per wavelength to create the channel time variation. For the antenna arrays,  $n_T = 2$  and  $n_R = 3$ . We choose a block length of  $M = 128$  with  $N_P = 2$  training symbol vectors per block. The additive noise is independent identically-distributed Gaussian noise with power level chosen so that the resulting signal-to-noise ratio (SNR) with respect to the symbol power is 30dB. Since we are using a high SNR, almost all of the symbols errors in the simulations are due to poor channel estimation. The PEDB method is used to obtain initial channel and symbol estimates for the multiple-pass algorithm. The rows of the  $\mathbf{A}$  and  $\mathbf{P}$  PEDB matrices are selected from the rows of an  $M \times M$  normalized Walsh-Hadamard matrix.

Figure 1 is a plot of the magnitude and phase of a single element of the channel matrix over a one-wavelength interval. Included in the plot are the actual channel values, the initial PEDB channel estimate, and the final channel estimate using the multiple-pass algorithm. It can be seen that the initial PEDB block estimates can differ significantly from the true channel values. These differences are significant enough to result in several symbol errors per block. The channel estimates produced using the decision-directed algorithm are much better approximations of the true channel values than the PEDB estimates. As a result, the symbol error performance will also be improved.

Previously, we stated that we selected the PEDB training method to allow the pilot symbols to provide our algorithm with information about the channel over an entire block of symbols. Figures 2 and 3 are a comparison of the error performance when using PSAM training versus the PEDB training. Figure 2 is a plot of the symbol errors per block resulting from using the PSAM and PEDB block estimates to decode the symbols over 78 symbol blocks. The two methods have a similar performance at this point, with the PEDB training method resulting in slightly fewer initial symbol errors than the PSAM method. Figure 3 shows the number of symbol errors per block following the use of the multiple-pass algorithm. These results indicate that the PEDB training structure provides a significant performance advantage over PSAM training. When the PSAM method was used, the channel was estimated well early in the block near the pilot symbols, but the estimation quality decreased farther into each block. Since the PEDB method spreads the pilot symbols over the entire block, this difficulty is not encountered.

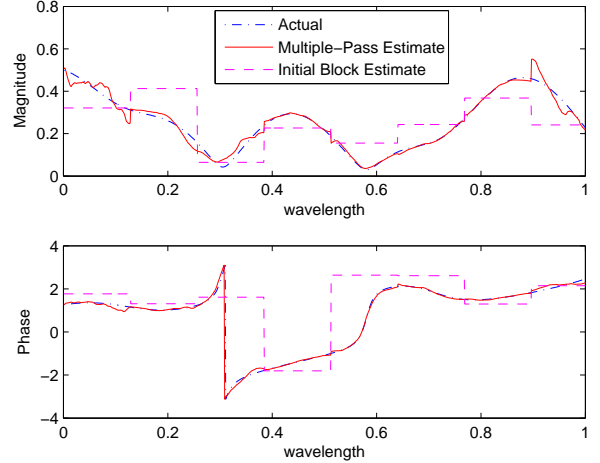


Fig. 1. Magnitude and phase of a single channel coefficient for the actual and estimated channel.

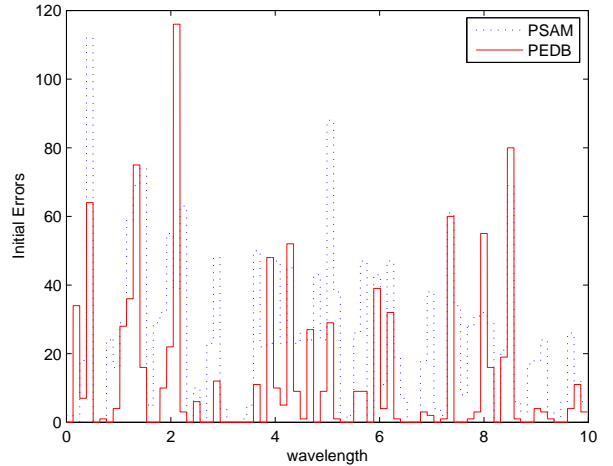


Fig. 2. Initial symbol error counts per block.

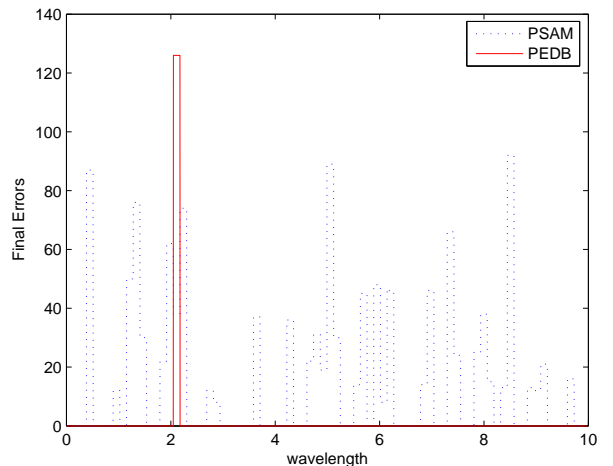


Fig. 3. Symbol error counts per block after the multiple-pass algorithm.

The results presented in Fig. 3 suggest that our multiple-pass algorithm eliminates all of the symbol errors within each symbol block most of the time. However, there are some blocks in which the algorithm is not able to correct the symbol errors. Recall that the original pdf of (17) was not unimodal. These blocks with uncorrected errors correspond to cases in which the algorithm converged to a local instead of global maximum. This generally occurs when there are a large number of initial symbol errors so that the algorithm starting point is far from the optimal point.

To avoid convergence to a local maximum, we can restrict the use of our algorithm to situations where the number of initial symbol errors are below a certain threshold. This is a reasonable assumption in cases where an error detection code can be used to assess the quality of the decoded symbols. For blocks with poor initial decoding performance, retransmission could be used to obtain the correct data. For example, out of 3900 symbol blocks in our simulation, 2680 blocks initially contain errors. For a threshold of 32 initial symbol errors (approximately 1 of 8 symbols in error), 3011 blocks are below the threshold, 1791 of which have errors. The errors in all but one of these blocks are completely eliminated using our algorithm. For a threshold of 50 (approximately 1 of 5 symbols in error), 3348 blocks are below the threshold, 2128 containing errors, of which 2119 can be made error free.

In the above example, a large percentage of blocks were above the thresholds and would require retransmission. Note, however, that the errors in many of these blocks may still be completely eliminated by our algorithm. For example, of the 889 blocks above the 32-error threshold, 657 are made error-free by the algorithm. Of the 552 blocks above the 50-error threshold, 328 are made error-free. An alternative approach would be to process all of the blocks, detect resulting errors, and only request block retransmission for those blocks that still contain errors.

Figure 4 shows the symbol-error-rate performance versus the number of algorithm iterations required for blocks with no final symbol errors, those with a threshold of 32 initial errors, and those with a threshold of 50. For the case where the threshold is 32, only 10 algorithm iterations are required to achieve almost all of the benefits of the algorithm. Our simulations indicate an average of 3, a median of 1, and a maximum of 30 iterations were required to remove the symbol errors from the correctable blocks.

## VI. CONCLUSION

In this paper, we considered the problem of jointly estimating the channel coefficients and data symbols in a mobile MIMO communication system. To do this, we first formulated a posterior distribution to describe the estimation problem. We then proposed a multiple-pass decision-directed algorithm to iteratively estimate the channel and decode the data symbols, using the CG approach to simplify the symbol-by-symbol channel estimation. Simulations indicate that the algorithm may converge to local optima when poor initial estimates are used. However, if the use of the algorithm is restricted to

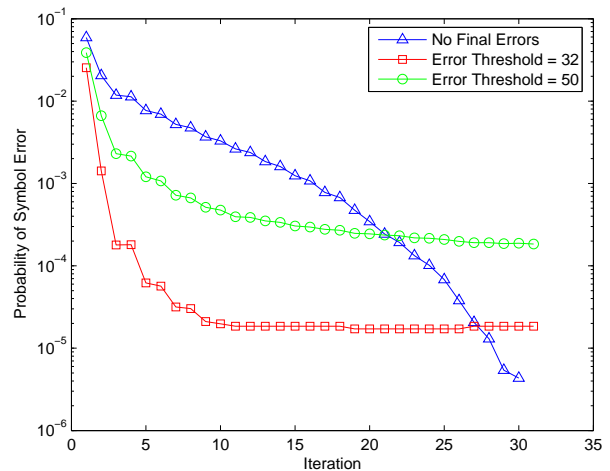


Fig. 4. Probability of symbol error versus the number of algorithm iterations.

symbol blocks where the number of initial errors are below some threshold, the algorithm is able to correct nearly all initial symbol decoding errors.

## REFERENCES

- [1] P. Hoeher and F. Tufvesson, "Channel estimation with superimposed pilot sequence," in *Proc. GLOBECOM*, vol. 4, Rio de Janeiro, Brazil, Dec. 1999, pp. 2162–2166.
- [2] H. Zhu, B. Farhang-Boroujeny, and C. Schlegel, "Pilot embedding for joint channel estimation and data detection in MIMO communication systems," *IEEE Commun. Lett.*, vol. 7, pp. 30–32, Jan. 2003.
- [3] M. Ghogho and A. Swami, "Improved channel estimation using superimposed training," in *Proc. IEEE 5th Workshop on Signal Processing Advances in Wireless Communications (SPAWC'04)*, July 2004, pp. 110–114.
- [4] C. Pirak, Z. J. Wang, K. J. R. Liu, and S. Jitapunkul, "Performance analysis for pilot-embedded data-bearing approach in space-time coded MIMO systems," in *Proc. ICASSP*, Mar. 2005, pp. 593–596.
- [5] M. Huiheng and A. Burr, "Iterative channel estimation for turbo equalization over MIMO double selective fading channels," in *Proc. IEEE Vehicular Technology Conference (VTC'04)*, vol. 2, pp. 1478–1482.
- [6] Y. Sun, M. Yee, and M. Sandell, "Iterative channel estimation with MIMO MMSE-turbo equalization," in *Proc. IEEE Vehicular Technology Conference (VTC'03)*, Oct. 2003, pp. 1278–1282.
- [7] C. Kominakis, C. Fragouli, A. Sayed, and R. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Processing*, vol. 50, pp. 1065–1076, May 2002.
- [8] G. Yanfei and H. Zishu, "MIMO channel tracking based on Kalman filter and MMSE-DFE," in *Proc. International Conference on Communications, Circuits, and Systems*, May 2005, pp. 223–226.
- [9] K. J. Kim, T. Reid, and R. A. Iltis, "Data detection and soft-Kalman filter based semi-blind channel estimation algorithms for MIMO-OFDM systems," in *Proc. IEEE International Conference on Communications (ICC'05)*, vol. 4, May 2005, pp. 2488–2492.
- [10] A. L. Swindlehurst, S. Daas, and J. Yang, "Analysis of a decision directed beamformer," *IEEE Trans. Signal Processing*, vol. 43, pp. 2920–2927, Dec. 1995.
- [11] S. Talwar, M. Viberg, and A. Paulraj, "Blind estimation of multiple co-channel digital signals using an antenna array," *IEEE Signal Processing Lett.*, vol. 1, pp. 29–31, Feb. 1994.
- [12] A. Ranheim, "A decoupled approach to adaptive signal separation using an antenna array," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 676–682, May 1999.
- [13] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2005.