A Joint Channel and Channel Length Estimation Algorithm for MIMO-OFDM

J. Siew, J. Coon, A. Nix, M. Beach, S. Armour, and J. McGeehan
Centre for Communications Research, University of Bristol
Merchant Venturers Building, Woodland Road, Bristol BS8 1UB, UK
Email: {Jiun.Siew, Justin.Coon}@bristol.ac.uk

Abstract—The recent interest in coherently detected MIMO-OFDM systems has fueled research efforts on various channel estimation methods. Few, however, have addressed the practical problem of estimating the instantaneous channel length. This necessary but often overlooked component of channel estimation is crucial to accurate channel estimates in many OFDM channel estimation techniques. In this paper we propose a method for jointly estimating the channel and instantaneous channel length. We describe the algorithm in detail, provide a mathematical framework, and show simulated results, both in terms of the estimation accuracy and the impact it has on a system’s packet error rate (PER) performance.

I. INTRODUCTION

Common methods of estimating the channel for multiple input-multiple output (MIMO) orthogonal frequency division multiplexing (OFDM) systems such as those published in [1, 2], rely on knowing the instantaneous channel length. Most of the published work on channel estimation however, assumes that this information is known at the receiver. In practice, this assumption generally does not hold and some method of obtaining the channel length must be used. Alternatively, the length of the channel can be simply assumed to be the length of the cyclic prefix (CP). Making this assumption however, potentially increases the complexity of the channel estimation method unnecessarily and results in overall performance degradation.

Surprisingly, most of the work in this area has concentrated on estimating the delay spread of the channel, leaving estimation methods of the instantaneous channel length relatively neglected. However, in [3], the authors propose a probabilistic method of estimating the channel length, while in [4] a recursive algorithm that jointly synchronises the system, estimates the channel length, and channel is proposed. This algorithm is based on the least squares error metric. In this paper we address this very topic using a variant of the channel estimation method proposed in [5] to estimate the channel length. More specifically, we iteratively estimate the channel using varying channel lengths until a reasonable channel estimate is found. The novelty in our proposed method lies in how a reasonable estimate is determined.

This paper is organised as follows: Section II details the system model used. The channel estimation algorithm is then described in Section III followed by the channel length estimation algorithm in Section IV. We then investigate the simulated performance of these algorithms in a MIMO-OFDM system with 2 different channels models in Section V before concluding in Section VI.

Notation: In the following sections, upper and lower case bold letters are used to denote matrices and vectors respectively. \( A_{p,q} \) denotes the element of matrix \( A \) in row \( p \), column \( q \). \( I_{N \times N} \) and \( 0_{N \times N} \) denote the \( N \times N \) identity and zero matrices respectively, while \( A^T \) and \( A^H \) denotes the transpose and transpose conjugate. \( E \{ \cdot \} \) denotes the expectation, \( 'tr' \) is the trace of a matrix, \( \text{diag} \{ A \} \) denotes a vector consisting of the diagonal elements of matrix \( A \), while \( \text{diag} \{ a_1, a_2, \ldots \} \) denote a matrix with the vector \( a \) or the values \( (a_1, a_2, \ldots) \) on its main diagonal. \( [A, B] \), and \( [A; B] \) denote the concatenation of \( A \) and \( B \) along columns and rows respectively; the resultant matrix is denoted by an underline.

II. SYSTEM DESCRIPTION

Consider a single antenna OFDM system with \( N \) subcarriers. The transmitted \( N \times 1 \) data vector \( x(t) \), at time \( t \), is multiplied by an \( N \times N \) inverse Fourier matrix, \( F^{-1} \), where \( F^{-1} = (1/\sqrt{N}) e^{j(2\pi ij)/N}, i, j \in \{0, \ldots, N - 1\} \). A CP is inserted at the transmitter by a \( P \times N \) matrix \( T_{CP} \) where \( T_{CP} = [0_{c \times (N-C)}; I_{C \times C}; I_{N \times N}] \). \( C \) is the length of the CP and \( P = C + N \). Thus, for a channel with \( L \) taps, the received OFDM symbol at time \( t \) is given by the equation:

\[
y(t) = H(0) T_{CP} F^{-1} x(t) + H(1) T_{CP} F^{-1} x(t-1) + \eta(t)
\]

where \( H(0) \) is a \( P \times P \) Toeplitz matrix with first column \( [h(0), \ldots, h(L-1), 0, \ldots, 0]^T \) and first row \( [h(0), 0, \ldots, 0]^T \) while \( H(1) \) is a \( P \times P \) Toeplitz matrix with first column \( [0, \ldots, 0]^T \) and first row \( [0, \ldots, h(L-1), h(1)]^T \) and \( \eta(t) \) represents additive white Gaussian noise with zero mean and variance \( \sigma_n^2/2 \) per dimension [6]. The assumption that \( C \geq L - 1 \) is made; implying that interblock interference (IBI) in \( y(t) \) is absent.

At the receiver, the CP is removed by the \( N \times P \) matrix \( R_{CP} = [0_{N \times C}, I_{N \times N}] \), and the resulting sequence is multiplied by the Fourier matrix \( F \). The matrix \( R_{CP} H(0) T_{CP} \), is circulant and thus FFT processing ensures that the resultant channel matrix, \( \hat{H} = FR_{CP} H(0) T_{CP} F^{-1} \) is diagonal. Note that \( R_{CP} H(1) T_{CP} = 0_{N \times N} \), thus maintaining subcarrier orthogonality. The processed received vector can be modeled as:

\[
y(t) = \tilde{H} x(t) + \eta(t)
\]
For a MIMO-OFDM system with \( n_T \) transmit and \( n_R \) receive antennas, the received signal of the \( p \)-th antenna on the \( n \)-th tone can be mathematically described by

\[
y_{p;n} = \sum_{q=1}^{n_T} h_{p,q;n} x_{q;n} + \eta_{p;n} \tag{3}
\]

where \( h_{p,q;n} \) is the channel coefficient between antennas \( p \) and \( q \) on subcarrier \( n \), \( x_{q;n} \) is the transmitted symbol from antenna \( q \) on subcarrier \( n \), and \( \eta_{p;n} \) is the noise term at the receiver on subcarrier \( n \). Obviously, the key difference between the MIMO and single antenna case is the superposition of transmitted symbols at the receiver.

### III. Estimating The Channel

To estimate the channel, a training sequence is designed with pilot symbols spaced every \( n_T \) subcarriers apart. The remaining subcarriers are nulled. Training sequences transmitted from each antenna are made orthogonal to each other by placing the pilot symbols on different subcarriers. So for example, for a 2 transmit-2 receive (2x2) antenna system the training sequence may look something like:

\[
T_1 = [1, 0, -1, 0, 1, 0, \ldots] ;
\]

\[
T_2 = [0, 1, 0, -1, 0, 1, \ldots] ;
\tag{4}
\]

where \( T_q \) is the transmitted training sequence from transmit antenna \( q \), and the elements of \( T_q \) are the transmitted symbols on each subcarrier. More generally, let \( \Omega_q \) be the set of indexes where pilot symbols are placed for the \( q \)-th transmitted training sequence, and let \( \Omega_1 \not\subseteq \Omega_2 \not\subseteq \ldots \not\subseteq \Omega_{n_T} \). Furthermore, we assume that \( n_T \leq N/L \) \cite{7, 8}. Using this training sequence structure reduces the received sequence in (3) to

\[
y_{p;n} = h_{p,q;n} x_{q;n} + \eta_{p;n}, \quad n \in \Omega_q \tag{5}
\]

At the receiver, the channel estimation (CE) algorithm performs the following steps:

1) Obtain the initial channel estimate
2) Multiply the initial channel estimate by the number of transmit antennas, \( n_T \)
3) Convert the channel estimate into the delay domain
4) Window significant taps
5) Convert this delay domain signal back into the frequency domain.

The initial channel estimate \( \hat{h}_{p,q}^{(0)} \) is a \( N \times 1 \) column vector whose elements are given by the equation:

\[
\hat{h}_{p,q}^{(0)} = \left\{ \begin{array}{ll}
y_{p;n} & n \in \Omega_q \\
x_{q;n} & n \not\in \Omega_q, \\
0 & \text{otherwise}
\end{array} \right. \quad n \in \Omega_q \tag{6}
\]

The final channel estimate for all channel paths in vector form, \( \hat{h} = [\hat{h}_{1,1}^T, \hat{h}_{1,2}^T, \ldots, \hat{h}_{n_R,n_T}^T]^T \) is given by

\[
\hat{h} = \mathbf{K} \hat{h}^{(0)} \tag{7}
\]

where \( \mathbf{K} = \text{diag}\{ \mathbf{K}_1, \ldots, \mathbf{K}_{n_R,n_T} \} \) and \( \hat{h}^{(0)} \) is the initial channel estimate for all channel paths. The \( N \times N \) matrix \( \mathbf{K}_{p,q} \) is defined by

\[
\mathbf{K}_{p,q} = n_T \mathbf{W} \mathbf{W}_T \mathbf{T}_{p,q} \mathbf{F}^{-1} \tag{8}
\]

where \( \mathbf{W}_{T,p,q} = \text{diag}\{ 1, 1, \ldots, 0, 0, \ldots \} \) is an \( N \times N \) delay domain windowing matrix with 1’s on the main diagonal where delay taps exist and zeros elsewhere for the channel path between transmit antenna \( q \) and receive antenna \( p \).

Note that scaling the channel estimate by \( n_T \) in step (2) can be physically interpreted as a method of recovering the complete channel energy if pilot symbols were transmitted on all subcarriers.

### IV. Estimating The Channel Length

The instantaneous channel length can be estimated by repeatedly performing the CE algorithm using various windowing sizes and determining the mean square error (MSE) between the current and previous channel estimates. The channel length estimation (CLE) algorithm is described as follows:

1) Initialisation
i. Calculate threshold, \( MSE_q \), based on received signal’s SNR
ii. Set initial window size equal to CP length, \( WNSZ = CP \)

2) Main algorithm for iteration \( i \)
   i. Calculate channel estimate \( \hat{h}^{(i)} \) using \( WNSZ^{(i)} \)
   ii. Decrement \( WNSZ^{(i)} \);
      \( WNSZ^{(i+1)} = WNSZ^{(i)} - 1 \)
   iii. Estimate channel \( \hat{h}^{(i+1)} \) with \( WNSZ^{(i+1)} \)
   iv. Calculate \( MSE_{\hat{h}^{(i)}, \hat{h}^{(i+1)}} \); the MSE between \( \hat{h}^{(i)} \) and \( \hat{h}^{(i+1)} \),

3) Make decision
i. Stop if \( MSE_{\hat{h}^{(i)}, \hat{h}^{(i+1)}} > MSE_q \). The estimated channel length, \( \hat{L} \), is the previous iteration’s window size \( WNSZ^{(i)} \), and the channel estimate is \( \hat{h}^{(i)} \)
ii. Otherwise, repeat main algorithm again.

Mathematically, \( \hat{h}^{(i)}_{p,q} \) can be written as

\[
\hat{h}^{(i)}_{p,q} = n_T \mathbf{W} \mathbf{W}_T^{(i)} \mathbf{T}_{p,q} \mathbf{F}^{-1} \mathbf{W}_q \left( h_{p,q} + \eta \right) = \mathbf{K}^{(i)} \mathbf{W}_q \left( h_{p,q} + \eta \right) \tag{9}
\]

where \( \mathbf{W}_q \left( h_{j,j} \right) = 1, \forall j \in \Omega_q \) and 0 otherwise, while \( h_{p,q} \) is a \( N \times 1 \) column vector containing the channel coefficients between transmit antenna \( q \) and receive antenna \( p \).

Using (9), the threshold value, \( MSE_q \), can be derived directly from the error covariance matrix between \( \hat{h}^{(i)} \) and \( \hat{h}^{(i+1)} \) assuming the channel and noise are uncorrelated:

\[
\begin{align*}
\mathbf{C} &= \mathbf{E} \left\{ \left( \hat{h}^{(i)} - \hat{h}^{(i+1)} \right) \left( \hat{h}^{(i)} - \hat{h}^{(i+1)} \right)^H \right\} \\
&= \mathbf{K}^{(i)} \mathbf{W} \mathbf{W}_T \mathbf{W}_q \mathbf{K}^{(i)} - \mathbf{K}^{(i)} \mathbf{W} \mathbf{W}_T \mathbf{W}_q \mathbf{K}^{(i+1)} \\
&+ \mathbf{K}^{(i+1)} \mathbf{W} \mathbf{W}_T \mathbf{W}_q \mathbf{K}^{(i)} + \mathbf{K}^{(i+1)} \mathbf{W} \mathbf{W}_T \mathbf{W}_q \mathbf{K}^{(i+1)} + \sigma_n^2 \left( \mathbf{K}^{(i)} \mathbf{W} \mathbf{W}_T \mathbf{W}_q \mathbf{K}^{(i)} - \mathbf{K}^{(i)} \mathbf{W} \mathbf{W}_T \mathbf{W}_q \mathbf{K}^{(i+1)} \right) \\
&- \mathbf{K}^{(i+1)} \mathbf{W} \mathbf{W}_T \mathbf{W}_q \mathbf{K}^{(i)} + \mathbf{K}^{(i+1)} \mathbf{W} \mathbf{W}_T \mathbf{W}_q \mathbf{K}^{(i+1)} \tag{10}
\end{align*}
\]

where \( \mathbf{R}_H = \mathbf{E} \left\{ \mathbf{h} \mathbf{h}^H \right\} \) and \( W^{(i)}_{T,j,j} = 1 \forall j \in [1, WNSZ^{(i)}] \) and 0 otherwise. Equation (10) can be viewed as two separate
components: the correlation component and the noise component. The correlation components are the terms containing \( \mathbf{R}_H \) while the noise component are the terms containing \( \sigma^2_\eta \).

Let

\[
\mathbf{C}_R = \mathbf{K}^{(i)} \mathbf{W} \mathbf{R}_H \mathbf{W}^H \mathbf{K}^{(i)H} - \mathbf{K}^{(i)} \mathbf{W} \mathbf{R}_H \mathbf{W}^H \mathbf{K}^{(i+1)H} - \\
\mathbf{K}^{(i+1)} \mathbf{W} \mathbf{R}_H \mathbf{W}^H \mathbf{K}^{(i)H} + \mathbf{K}^{(i+1)} \mathbf{W} \mathbf{R}_H \mathbf{W}^H \mathbf{K}^{(i+1)H}
\]

The total theoretical MSE is determined by

\[
\text{MSE}_{\text{total}} = \frac{\text{tr} \{ \mathbf{C} \}}{N \text{snr}_{\mathcal{R}}},
\]

while the MSE of the correlation component is

\[
\text{MSE}_R = \frac{\text{tr} \{ \mathbf{C}_R \}}{N \text{snr}_{\mathcal{R}}},
\]

and the MSE of the noise components is

\[
\text{MSE}_\eta = \frac{\sigma^2_\eta \text{tr} \{ \mathbf{C}_\eta \}}{N \text{snr}_{\mathcal{R}}^2}.
\]

Plots of \( \text{MSE}_{\text{total}} \), \( \text{MSE}_R \) and \( \text{MSE}_\eta \) for an 8-tap Rayleigh fading channel are shown in Figures 1 to 3 for varying SNR points. From Figure 3, it is clear that for a certain SNR, \( \text{MSE}_\eta \) is constant regardless of \( \text{WNSZ} \) while \( \text{MSE}_R \) decreases significantly once the \( \text{WNSZ} \geq L \) as shown in Figure 2. Hence, the lower bound of MSE between channel estimates is determined by the noise component, \( \text{MSE}_\eta \). Furthermore, since the only varying parameter of \( \text{MSE}_\eta \) is the noise variance, it is possible to precalculate an unscaled value of \( \text{MSE}_\eta \), in order to save computational overhead. The actual determination of \( \text{MSE}_\eta \) at the initialisation stage then becomes a simple scalar multiplication. Obviously, this assumes that the noise variance at the receiver is known.

### A. Algorithm Complexity

The complexity of the algorithm is obviously dependent on the length of the channel or the number of iterations needed to obtain the actual channel length, \( \Gamma \), and obtaining the threshold values, \( \Gamma_{\text{threshold}} \). The approximate number of multiplications for each are given by:

\[
\Gamma(i) \approx 2Nn_Tn_R (\log_2 N) (i) \tag{13}
\]

\[
\Gamma_{\text{threshold}} \approx 28 (Nn_T)^3 + 4 (Nn_T)^2 \tag{14}
\]

The dependency of complexity on \( N \) and \( i \) is shown in Figure 4.
transmitter, a packet of 1024 bits were coded using the 1/2 rate convolutional code specified in [9]. These bits were then randomly interleaved before being mapped to QPSK symbols. The symbols were then buffered into blocks of 64 before being passed to the IFFT. The appropriate CP was then added before being transmitted. At the beginning of every packet, the training sequence for each antenna was transmitted prior to any data bearing symbols. The transmitted symbols were then convolved with an 8-tap i.i.d. Rayleigh fading channel and an 8-tap exponentially decaying channel. Channel power was scaled such that the average energy was 1 and was assumed static over the length of the packet. At the receiver, the received symbols were demodulated by an FFT before the CP was removed. The proposed algorithm was performed on each training sequence and the resulting channel estimates were used to equalise the received symbols with a minimum mean square error (MMSE) equaliser. The symbols were then detected with soft-decisions before being decoded by a Viterbi decoder.

In Figures 7 to 10 the packet error rate (PER) performance of the MIMO system is shown and compared to 3 other cases: perfect channel knowledge (IdealL), perfect channel length knowledge with an estimated channel (IdealL1), and no channel length knowledge (FixedL). The curve labelled AdptvL denotes the PER performance using the CLE algorithm. In the FixedL case, the window size was set to the length of the CP. It is evident from Figures 7 and 8 that there is about a 1dB performance gain in the IdealL and AdptvL case relative to the FixedL case for 2x2 systems, which approaches 2dB for the 4x4 systems shown in Figures 9 and 10. This gain highlights the importance of having accurate channel length knowledge at the receiver. The capability of this algorithm is proven since there is only a slight difference between the performance using the CLE algorithm and the IdealL case.

Note also that the inaccuracies of estimating L as pointed out in Figure 6 results in only minor PER performance loss as compared to the IdealL case for the 2x2 system in Figure 8 and no performance loss whatsoever for the 4x4 system as shown in Figure 10. Furthermore, in Figure 10 the AdptvL performs slightly better than the IdealL case for lower SNR values. This is because although the algorithm is actually estimating the values of L incorrectly as in Figure 6, the last few channel taps that the CLE algorithm might remove in the exponential channel are small and thus the channel estimate is still good enough to adequately detect the symbols. If the CLE algorithm removes more channel taps than is necessary it also inadvertently removes noise as well, thus aiding the accuracy of the channel estimate for low SNR.

### V. Results and Discussion

To test the performance of the CLE algorithm, two different simulations were performed. The first simulation tested how well the CLE algorithm actually estimated the channel length. For this simulation, 10,000 channel realisations were independently generated using two power delay profiles (PDP): an 8-tap i.i.d. Rayleigh fading channel, and an 8-tap exponentially decaying channel. An orthogonal training sequence for a 2x2 system was used as described in (4) and convolved with each channel realisation. The algorithm was then performed on the received training sequence and the accuracy of the algorithm was recorded. From the data, a PDF was constructed and shown in Figures 5 and 6.

From Figure 5, it is evident that the algorithm performs very well, as most of the time it accurately estimates that there are 8 taps in the channel. Performance is somewhat degraded however, in the exponentially decaying channel as shown in Figure 6. This is because in the exponential channel, some of the last taps are very small in amplitude and are mistaken for noise. Hence, for low SNR points, the algorithm does not estimate the channel as accurately.

The second simulation investigated the impact of the CLE algorithm on the overall system performance. For this simulation, coded 2x2 and 4x4 spatial multiplexing systems with 64 subcarriers and a CP length of 16 was used. At the
to the length of the CP. Ongoing work will investigate the issue of obtaining the noise variance in order to estimate the channel length.

ACKNOWLEDGEMENT

The authors would like to thank Toshiba TRL Bristol for financially supporting this work. We are grateful for the technical support provided by Dr. M. Sandell, Dr. M. Yee, Dr. S. Parker, and Dr. Y. Sun.

REFERENCES