OFDM CIR ESTIMATION WITH UNKNOWN LENGTH VIA BAYESIAN MODEL SELECTION AND AVERAGING

Ido Nevat†, Gareth W. Peters§ and Jinhong Yuan†

†School of Electrical Engineering and Telecommunications
§CSIRO Mathematical and Information Sciences, University of NSW, Sydney, Australia
§School of Mathematics and Statistics, University of NSW, Sydney, Australia
email: ido@student.unsw.edu.au, peterga@maths.unsw.edu.au, J.Yuan@unsw.edu.au

ABSTRACT
This work presents new CIR estimators for OFDM systems over frequency selective channels, where the length of the CIR is unknown a-priori. We derive the MMSE estimator for this model using two different criteria, namely, Bayesian model averaging and Bayesian model order selection. Estimation of the CIR length enables adapting the CP length to the changing propagation environment, resulting in increased throughput due to shorter CP. Simulation results under different channel conditions demonstrate the robustness of the estimators.

I. INTRODUCTION
Orthogonal frequency division multiplexing (OFDM) [1] has received a considerable amount of attention over the last few years due to its ability to transform a frequency selective channel into a group of low-rate parallel flat channels, thereby increasing the symbol duration. The concatenation of a cyclic prefix (CP) at the beginning of each OFDM symbol cancels the inter block interference (IBI). The cyclically extended guard interval also converts linear convolution of the signal and the channel into circular convolution, leading to a simple one-tap equalization. One drawback of OFDM systems is the loss of spectral efficiency due to CP insertion. The length of the CP needs to be longer than the length of the channel impulse response (CIR) and does not convey any useful information [1]. For example, in IEEE 802.11a standard [2], a fixed proportion of 20% of the energy and time is spent on CP. This leads to reduced spectral efficiency. The CIR length varies according to the environment in which the system operates. Since the acceptable channel estimation error is a predefined system requirement, knowing the CIR length possesses another advantage: The number of pilot symbols can also be adapted based on the knowledge of the CIR length. When the CIR length is relatively long, more pilot symbols are required to achieve the required estimation error. When the CIR length is small, reduced number of pilot symbols can be used, leading to increased throughput. While overestimating the CIR length may decrease bandwidth efficiency, underestimation of the CIR length will cause inter symbol interference and inter carrier interference due to loss in orthogonality, this is a more serious situation which needs to be avoided. It is therefore desirable to design an OFDM transceiver for which the length of the CIR is continuously estimated and adapts the length of the CP. In IEEE 802.16a [3], an attempt to have an adaptive form of the CP has been suggested. The Base Station (BS) can choose the optimal CP. However, once the BS has decided on the optimal length, all users within the specific cell will use the same CP length. Existing approaches include most significant taps (MST) idea developed in [4] to estimate the channel order. In [5] an auxiliary function is used to distinguish between real taps and the noise contribution. We present two new CIR estimators with unknown length based on the Bayesian paradigm. First, a Bayesian model averaging (BMA) based CIR estimator is considered. This approach considers the entire ensemble of statistical models, and provides a coherent mechanism for model uncertainty. Alternatively, a Bayesian model order selection (BMOS) CIR estimator is developed. In contrast to the BMA approach, this method makes use of a point estimate of the model length, which is optimal under the hit-miss criterion (Maximum a
Posteriori (MAP). In presenting a Bayesian approach we are free to introduce priors for length which penalize underestimation of CIR length in order to avoid ISI and ICI problems.

The following notation is used throughout: boldface upper case letters denote frequency domain column vectors, boldface lower case letters denote time domain column vectors, and standard lower case letters denote scalars. The superscripts (.)\(^T\) and (.)\(^H\) denote Hermitian and Transpose, respectively. By \(\mathbf{I}\) we denote the identity matrix. The functions \(p(x)\), \(p(x|y)\) and \(E\{\cdot\}\) denote the probability distribution function (PDF) of \(x\), the PDF of \(x\) given \(y\), and the expectation, respectively. The covariance matrix of \(a\) is defined as \(\Sigma_a = E\{(a - E(a))(a - E(a))^H\}\) and the cross covariance matrix of \(a\) and \(b\) is defined as \(\Sigma_{ab} = E\{(a - E(a))(b - E(b))^H\}\)

**II. SYSTEM MODEL**

A sequence of information bits is mapped to complex-valued symbols of an M-ary modulation alphabet set \(A = \{a_1, ..., a_{|A|}\}\). The data symbols are multiplexed to OFDM subcarriers. After modulation of the OFDM symbols via an \(N\) point IDFT (Inverse Discrete Fourier Transform), the signal is transmitted over a frequency selective, time varying channel. The assumptions made about the channel are that it is statistically independent taps, and covariance matrix \(\Sigma_h\). \(W_L\) is an \(N \times L\) partial DFT matrix, defined as

\[
W_L = \left\{ e^{-j2\pi nl/N} \right\}_{n=0,...,N-1;l=0,...,L-1}. \tag{2}
\]

The \(N \times 1\) channel frequency response (CFR), \(H_n\), with elements \(H_n^k\), \(1 \leq k \leq N\), can be written as

\[
H_n^k = \sum_{l=0}^{L-1} h_n[l] e^{-j2\pi kl/N}. \tag{3}
\]

In this setup, a one-tap simple equalizer can be applied for each subcarrier in order to perform detection. In future we omit time dependency \(n\) when it is not needed. As mentioned before, the CP does not convey any useful information and only used to prevent ISI. We can define the SNR loss due to CP as

\[
SNR_{loss} = -10 \log_{10} \left( 1 - \frac{T_{cp}}{T} \right), \tag{4}
\]

where \(T_{cp}\) is the length of the CP and \(T = T_{cp} + T_s\) is the length of the transmitted symbol. It is therefore desirable to minimize the CP length while making sure it’s longer than the channel length in order to avoid ISI.

**III. CIR ESTIMATION WITH KNOWN CHANNEL LENGTH**

If the CIR length \(L\) is known a priori, the minimum mean squared error (MMSE) estimate of \(h\), derived in [6], is given by

\[
\hat{h} = E\{h|Y, L\} = E\{hY^H\} E^{-1}\{YY^H\} Y = \Sigma_h (DW_L)^H (DW_L \Sigma_h (DW_L)^H + \sigma_w^2 I)^{-1} Y, \tag{5}
\]

and the corresponding estimation covariance error matrix is

\[
E\left\{ (h - \hat{h}) (h - \hat{h})^H \right\} = \Sigma_h - \Sigma_h Y \Sigma_{YY}^{-1} \Sigma_y Y^H. \tag{6}
\]

The overall estimation error is

\[
Error = Trace \left\{ E\left\{ (h - \hat{h}) (h - \hat{h})^H \right\} \right\}. \tag{7}
\]

It is clear that the smaller \(L\) is, the smaller the overall error. In the rest of the paper we will derive the MMSE CIR estimation without prior knowledge of \(L\).
IV. CIR ESTIMATION WITH UNKNOWN CHANNEL LENGTH USING BAYESIAN MODEL AVERAGING

In this section we develop CIR estimation without a-priori knowledge of CIR length, \( L \), using Bayesian model averaging (BMA) [7]. BMA accounts for model uncertainty by averaging over possible or plausible models, where each model’s weight is given by its posterior model probability. Using BMA, one calculates the quantity of interest under each model and then averages according to how likely each model is.

The quantity of interest is the Bayesian MMSE channel estimation, that can be written as [6]

\[
\hat{h} = E \{ h | Y \}. \tag{8}
\]

First, the posterior distribution of \( h \) given the data \( Y \) is

\[
p(h | Y) = \sum_{l=1}^{L_{\text{max}}} p(h | Y, l) p(l | Y). \tag{9}
\]

The posterior mean of \( h \) can be written as

\[
\hat{h} = E \{ h | Y \} = \int h p(h | Y) \, dh = \int \left[ \sum_{l=1}^{L_{\text{max}}} \{ h p(h | Y, l) p(l | Y) \} \right] \, dh = \sum_{l=1}^{L_{\text{max}}} p(l | Y) E \{ h | Y, l \},
\]

where \( \hat{h} \) is a weighted average of all possible posterior means under each of the models considered \( E \{ h | Y, l \} \), weighted by their posterior model probability \( p(l | Y) \). Given the prior for \( l, p(l) \), the posterior model probability for \( l \) can be expressed as

\[
p(l | Y) = \frac{p(Y | l) p(l)}{\sum_{l=1}^{L_{\text{max}}} p(Y | l) p(l)}, \tag{11}
\]

where the marginal likelihood \( p(Y | l) \) measures how well the model predicts the data and it can be written as

\[
p(Y | l) = \int p(Y | h, l) p(h | l) \, dh,
\]

where \( N(m, \Sigma) \) is the Normal distribution with mean equal to \( m \) and covariance matrix \( \Sigma \).

We now evaluate \( E \{ h | Y, l \} \). Conditional on \( l, Y \) and \( h \) are zero-mean jointly Gaussian, and the MMSE estimator for \( h \) given in [6] is

\[
E \{ h | Y, l \} = E \{ h Y^H \} E^{-1} \{ Y Y^H \} Y
\]

\[
= \sum_{h} (D W_l)^H \left( \frac{1}{\Sigma_{Y | l}} \right)^{-1} \left( \frac{(D W_l)^H + \sigma^2 I}{\Sigma_{Y | l}} \right)^{-1} Y.
\]

Substituting (11), (12) and (13) into (10), the BMA estimator of \( \hat{h} \) can be expressed as

\[
\hat{h} = \sum_{l=1}^{L_{\text{max}}} \frac{1}{|\Sigma_{Y | l}|^{1/2}} \exp \left\{ -\frac{1}{2} Y^H \Sigma_{Y | l}^{-1} Y \right\} E \{ h | Y, l \}
\]

\[
= \sum_{l=1}^{L_{\text{max}}} \frac{1}{|\Sigma_{Y | l}|^{1/2}} \exp \left\{ -\frac{1}{2} Y^H \Sigma_{Y | l}^{-1} Y \right\} - E \left\{ \log \left( \sum_{l=1}^{L_{\text{max}}} P(h | Y, l) P(l | Y) \right) \right\} \tag{14}
\]

Averaging over all the possible models \( l = 1, \ldots, L_{\text{max}} \) provides a better estimate than any single model under the logarithmic scoring criterion [8]

\[
- E \left\{ \log \left( \sum_{l=1}^{L_{\text{max}}} P(h | Y, l) P(l | Y) \right) \right\} \leq - E \{ \log \{ P(h | Y, j) \} \}, j = 1, \ldots, L_{\text{max}}.
\]

V. CIR ESTIMATION WITH UNKNOWN CHANNEL LENGTH USING MODEL ORDER SELECTION

In this section we develop the CIR estimation without a-priori knowledge of the CIR length \( L \) using a model order selection approach. Using this approach we first find the most probable model using the MAP estimate of the model order, \( \hat{l}_{\text{MAP}} \), and then condition on this estimate to obtain the MMSE estimate of \( h \).

This procedure is composed of the following 2 steps:

1) MAP estimate of the CIR length

\[
\hat{l}_{\text{MAP}} = \arg \max_l p(l | Y) = \arg \max_l p(Y | l) p(l).
\]

2) MMSE estimate of the CIR

\[
\hat{h} = E \{ h_{\hat{l}_{\text{MAP}}} Y^H \} E^{-1} \{ Y Y^H \} Y
\]

\[
= \sum_{h_{\hat{l}_{\text{MAP}}}} (D W_{\hat{l}_{\text{MAP}}})^H \left( \frac{1}{\Sigma_{Y | \hat{l}_{\text{MAP}}}} \right)^{-1} Y.
\]

\[
\tag{17}
\]
VI. COMPLEXITY ISSUES

The complexity of the BMA MMSE in (14) and of the BMOS MMSE in (17) is actually quite low, as most of the terms can be precalculated. For example, when calculating (14) a significant amount of precalculation can be performed. Excluding the outermost right matrix multiplication (Y), for each model I, all the remaining multiplications can be performed in advance and stored in the system’s memory.

VII. SIMULATION RESULTS

The OFDM system setup is \( N = 64 \) subcarriers employing QPSK symbols, \( L_{\text{max}} = N/4 \) (typical CP length value in OFDM systems). The channel is modeled as block Rayleigh fading and the channel length \( L \) follows a truncated Poisson distribution

\[
P(l) = \frac{\lambda^l \exp(-\lambda)}{C l!},
\]

where \( \lambda = 8 \), and \( C = \sum_{l=1}^{L_{\text{max}}} \frac{\lambda^l \exp(-\lambda)}{l!} \).

The realization of the channel length \( L \) is 8 and \( \sigma^2_h = \frac{1}{L} \). One OFDM pilot symbol was used for every frame of length 128, and 5000 frames were transmitted for each SNR. The channel estimation mean squared error (MSE) and bit error rate (BER) results for different estimators are depicted in Figs. 1 and 2, respectively. The following estimators have been compared: An estimator which has the knowledge of \( l \) (labeled as Known L), and serves as the system’s lower bound. An estimator based on BMA and is given by eq. (14) and an estimator based on BMOS and is given by eq. (17). BMA and BMOS show comparable results (labeled as Est L). The last estimator assumes \( l = L_{\text{max}} \) (labeled as \( L_{\text{max}} \)) and does not employ CIR length estimation. As a lower bound for the BER, a detector with known CIR (labeled as CSI) is provided. The results show that the proposed algorithms operate very close to the detector with known CIR length. In Fig. 3, the histogram of the CIR length estimate (eq. (16)) for SNR=10 dB is depicted. This result shows that over 80% of the times the MAP estimation is correct and demonstrates the robustness of the MAP estimate in (16). We also studied the performance of this estimator for different number of subcarriers \( (N = 32, 64, 128) \). The average channel length estimate for different SNR values is depicted in Fig. 4. These results are evident of the robustness of the MAP estimator, even in low SNR values. It is also evident that the larger \( N \) is the better the MAP estimate. In high SNR values, regardless of the value of \( N \), the MAP estimator converges to the actual channel order.

We also studied two scenarios of model misspecification, in which the prior distribution for model order was a truncated Poisson. In scenario I we specified the true model order significantly less than the prior mean, and vise versa in scenario II. The outcome was for small \( N (N < 16) \), small difference in performance was observed between BMOS and BMA at low SNR \( (SNR < 10 \text{ dB}) \). For large \( N \) no difference was evident for any level of SNR. These two effects can be explained that in high SNR there is very weak prior influence. In low SNR with small \( N \) there is sufficient prior influence to induce a performance gain of BMA over BMOS of around \( \frac{1}{2} \text{ dB} \).

VIII. CONCLUSIONS

In this work we developed two time-domain channel estimators for OFDM systems with no knowledge of the CIR length. Using Bayesian model averaging and model selection strategies two CIR estimators have been derived. The simulation results demonstrate the effectiveness of these methods. We show that the estimators not only perform better than a conventional estimator, they enable the use of adaptive CP for OFDM systems, providing a bandwidth efficient system.
**IX. REFERENCES**


