Published in IET Communications
Received on 6th December 2011
Revised on 10th August 2012
Accepted on 26th October 2012
doi: 10.1049/iet-com.2011.0910

ISSN 1751-8628

Biased estimators with adaptive shrinkage targets for orthogonal frequency division multiple access channel estimation

Sheetal Kalyani[1], *Raghavendran Lakshminarayanan*[2], *Krishnamurthy Giridhar*[1]

[1]*Department of Electrical Engineering, Indian Institute of Technology Madras, Adyar, Chennai 600036, India*
[2]*Department of Avionics, IIST, Trivandrum, Thiruvananthapuram 695547, India*
E-mail: skalyani@ee.iitm.ac.in

Abstract: In orthogonal frequency division multiple access-based systems where channel frequency response (CFR) estimation has to be carried out using only the user-specific (localised) pilots within a small time frequency block, the accuracy of the estimates suffer because of the limited number of pilots and imperfect knowledge of the channel statistics. A biased estimator is proposed for the estimation of CFR over the time frequency block. Hypothesis tests are designed to ascertain the time and frequency selectivity of the CFR within the region of interest, and the outcome of these tests are used to determine a vector shrinkage target for the biased estimator. Simulation results indicate that the performance of the proposed estimator is comparable to that of the optimal minimum mean square error estimator, even though it does not have any knowledge of the channel statistics.

1 Introduction

The focus of this paper is on obtaining accurate channel estimates in practical orthogonal frequency division multiple access (OFDMA)-based systems based on standards such as IEEE 802.16m and 3GPP LTE-A when only user-specific (localised) pilots are used. Although there is a vast amount of literature on channel estimation and/or channel frequency response (CFR) estimation in orthogonal frequency division multiple (OFDM) systems, most of the work has focused on systems which have finite but large number of pilots [1, 2] and/or wideband pilots [3]. In both IEEE 802.16m and 3GPP LTE-A standards, data transmission is in terms of small units called resource blocks (RBs) where a RB consists of Q subcarriers per OFDM symbol and R OFDM symbols. In the user-specific pilot mode, CFR estimation has to be necessarily carried out using only the user-specific pilots in the allocated RB. When such localised pilots are used, one cannot apply estimation methods such as the modified least squares in [3], since it requires wideband pilots. The two-dimensional (2D)-minimum mean square error (2D-MMSE) methods [1, 2] can be applied using the pilots in the RB. However, optimal MMSE estimator requires the knowledge of the channel statistics which are seldom known accurately at the receiver.

One approach to not having any knowledge of the channel statistics is to use the robust 2D-MMSE filter, which is designed assuming an ideally band-limited and time-limited uniform scattering function [1, 2]. For the proposed estimator and the robust 2D-MMSE estimator considered here, it is assumed that channel statistics such as Doppler and the power delay profile (PDP) are unknown. However, the noise variance and hence the maximum likelihood (ML) covariance matrix is assumed to be known for both the proposed approach and the robust 2D-MMSE approach. However, the performance of the robust 2D-MMSE estimator degrades when compared with the optimal MMSE estimator if the robust filter has finite number of taps [2] (this becomes especially pronounced when the number of taps of the robust filter is very small as in the case of CFR estimation using only pilots within a RB). This points to the possibility of developing new estimators that can bridge the gap between the robust 2D-MMSE and the optimal MMSE approach.

In this work, a CFR estimation method is proposed which applies biased estimation to obtain improved CFR estimates on the pilots, and then interpolates them using the robust 2D-MMSE filter to obtain the CFR estimates within the RB. Prior work on biased estimation, specific to signal processing problems has been reported in [4–6]. To the best of our knowledge, biased estimator have not been used in the context of localised channel estimation for OFDMA systems. Our main contributions are (i) applying biased estimation techniques for localised CFR estimation, (ii) adaptively choosing a vector shrinkage target for the biased estimator that best reflects the the time–frequency selectivity of the CFR using the aid of multiple hypothesis tests and (iii) choosing the thresholds for the hypothesis tests such that the probability of incorrect choice of shrinkage targets is bounded. Although the optimal MMSE estimator requires complete knowledge of the channel statistics, our approach requires only the knowledge of the noise variance and yet

provides a comparable performance. Although both the proposed method and the robust 2D-MMSE method do not assume any knowledge of channel statistics, the performance of the robust 2D-MMSE approach is very poor compared with the proposed approach when the actual channel is only moderately frequency selective. Simulations results in Section 4 show that the proposed estimator significantly outperforms the robust 2D-MMSE estimator and approaches the performance of the optimal MMSE filter even though it does not have any knowledge of the channel statistics. An analysis of the computational complexity of the proposed estimator has also been provided and it indicates that the complexity of the proposed method is comparable to that of the robust 2D-MMSE method.

2 OFDM system description

In the vector notation, the OFDMA system representation on the pilots within an RB is given by

$$\boldsymbol{Y}_{\mathrm{p}} = \boldsymbol{X}_{\mathrm{p}}\boldsymbol{H}_{\mathrm{p}} + \boldsymbol{V}_{\mathrm{p}}, \quad \boldsymbol{V}_{\mathrm{p}} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{C}_{\mathrm{v}}) \tag{1}$$

where $\{\boldsymbol{Y}_{\mathrm{p}}, \boldsymbol{H}_{\mathrm{p}}, \boldsymbol{V}_{\mathrm{p}} \in C^{P\times 1}\}$, $\boldsymbol{X}_{\mathrm{p}} \in C^{P\times P}$ is diagonal matrix containing the pilot symbols, and $\boldsymbol{H}_{\mathrm{p}}$ is the vector of the CFR seen at these locations.

2.1 Channel/CFR estimation problem

The ML estimate of the CFR at pilot locations (when Gaussian noise $\boldsymbol{V}_{\mathrm{p}}$ has covariance $\boldsymbol{C}_{\mathrm{v}}$) is

$$\widehat{\boldsymbol{H}}_{\mathrm{ML,p}} = (\boldsymbol{X}_{\mathrm{p}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{v}}^{-1}X_{\mathrm{p}})^{-1}\boldsymbol{X}_{\mathrm{p}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{v}}^{-1}\boldsymbol{Y}_{\mathrm{p}} \tag{2}$$

whose covariance matrix is given by $\boldsymbol{C} = (\boldsymbol{X}_{\mathrm{p}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{v}}^{-1}\boldsymbol{X}_{\mathrm{p}})^{-1}$. Here the covariance matrix of the ML estimate $\boldsymbol{C} = \sigma^2\boldsymbol{I}_P$ when $\boldsymbol{C}_{\mathrm{v}} = \sigma^2\boldsymbol{I}_P$ and $\boldsymbol{X}_{\mathrm{p}}^{\mathrm{H}}\boldsymbol{X}_{\mathrm{p}} = \boldsymbol{I}_P$, where $\boldsymbol{I}_n$ is a $n \times n$ identity matrix. The optimal MMSE estimator of the CFR over the entire RB is the 2D-Wiener smoother that utilises the correlations along time and frequency [2]. Vectorised optimal MMSE estimate of the CFR matrix over the RB, namely, $\widehat{\boldsymbol{H}}_{\mathrm{MMSE}} \in \mathbb{C}^{QR\times 1}$ can be obtained by interpolating the ML estimates on the pilots using the Wiener smoother $\boldsymbol{W}_{\mathrm{opt}} \in \mathbb{C}^{QR\times P}$ thus

$$\widehat{\boldsymbol{H}}_{\mathrm{MMSE}} = \boldsymbol{W}_{\mathrm{opt}}\widehat{\boldsymbol{H}}_{\mathrm{ML,p}} \tag{3}$$

where $\boldsymbol{W}_{\mathrm{opt}} = \boldsymbol{R}_{\boldsymbol{H},\widehat{\boldsymbol{H}}_{\mathrm{ML,p}}}\boldsymbol{R}^{-1}_{\widehat{\boldsymbol{H}}_{\mathrm{ML,p}},\widehat{\boldsymbol{H}}_{\mathrm{ML,p}}}$ and $\boldsymbol{R}_{\boldsymbol{H},\widehat{\boldsymbol{H}}^{\mathrm{H}}_{\mathrm{ML,p}}} \in \mathbb{C}^{QR\times P}$ and $\boldsymbol{R}_{\widehat{\boldsymbol{H}}_{\mathrm{ML,p}},\widehat{\boldsymbol{H}}_{\mathrm{ML,p}}} = E[\widehat{\boldsymbol{H}}_{\mathrm{ML,p}}\widehat{\boldsymbol{H}}^{\mathrm{H}}_{\mathrm{ML,p}}] = \boldsymbol{R}_{\boldsymbol{H}_{\mathrm{p}},\boldsymbol{H}_{\mathrm{p}}} + \boldsymbol{C} \in \mathbb{C}^{P\times P}$ are the ML channel estimate cross-correlation (between actual CFR over the RB and ML CFR estimate at pilots) and auto-correlation matrices, respectively, and $E[.]$ is the expectation operator.

The correlation functions used in (3) require the knowledge of the channel PDP and fade characteristics, which are not often known as the wireless system, is put to use under different operating conditions. One option is the robust 2D-MMSE approach which assumes a uniform scattering function for designing the Wiener smoother, which is denoted by $\boldsymbol{W}_{\mathrm{rob}}$. Further, for reducing complexity the $QR \times P$ matrix $\boldsymbol{W}_{\mathrm{rob}}$ is designed with real coefficients [2]. For computing $\boldsymbol{W}_{\mathrm{rob}}$, the correlation function between kth subcarrier in nth symbol and lth subcarrier in mth symbol, is given by [2]

$$r_{\mathrm{rob}}(k-l, n-m) = \mathrm{sinc}\left(\frac{2\pi L(k-l)}{N}\right)\mathrm{sinc}(2\pi f'_{\mathrm{D}}(n-m)) \tag{4}$$

Since the multipath delay spread length L is not known, it is set equal to the cyclic prefix length L_{CP}. $f'_{\mathrm{D}} = f_{\mathrm{D}}T_{\mathrm{s}}$ is the normalised Doppler frequency and $\mathrm{sinc}(x) = \sin(x)/x$. Here, f_{D} is assumed maximum Doppler frequency and T_{s} is the OFDM symbol duration. Vectorised robust 2D-MMSE estimate of the CFR matrix, namely, $\widehat{\boldsymbol{H}}_{\mathrm{rob}} \in \mathbb{C}^{QR\times 1}$ is given by $\widehat{\boldsymbol{H}}_{\mathrm{rob}} = \boldsymbol{W}_{\mathrm{rob}}\widehat{\boldsymbol{H}}_{\mathrm{ML,p}}$

It has been pointed out in [7] that the task of obtaining the CFR estimates over the entire RB can be split into two simpler problems of (i) designing a CFR estimator on the pilots and (ii) designing an interpolation filter to interpolate the CFR estimates obtained from the pilots. Further, Auer and Karipidis [7] has shown that if the CFR estimator on the pilots has a low mean square error (MSE), one can use the robust 2D-MMSE interpolator on these pilot estimates to obtain a performance close to the optimal MMSE estimator. Hence, the focus here is on obtaining improved CFR estimates on the pilots by using the proposed method and the interpolating these improved CFR estimates using the robust 2D-MMSE filter.

3 Proposed estimator

Minimum variance unbiased estimators (MVUE) minimise the MSE with zero bias as the constraint and achieve the Cramer–Rao bound, however they do not achieve the smallest possible MSE. Optimal MMSE estimator achieve the smallest MSE possible, however they require the knowledge of the channel statistics. An approach, which falls in between the ML and MMSE approaches, is the biased estimation approach wherein the ML estimate is biased in order to reduce its variance at the cost of a small increase in bias. This reduces the estimation MSE when compared with the ML estimator and at the same time does not require any apriori knowledge about the channel statistics. It can be shown that the optimal MMSE filter (when applied for CFR estimation only on pilots) becomes the MVUE when $\boldsymbol{R}^{-1}_{\boldsymbol{H}_{\mathrm{p}},\boldsymbol{H}_{\mathrm{p}}} \to \boldsymbol{0}$. For any $\boldsymbol{R}^{-1}_{\boldsymbol{H}_{\mathrm{p}},\boldsymbol{H}_{\mathrm{p}}} \geq \boldsymbol{0}$ the optimal MMSE filter can be thought of as 'down-weighting' the observation $\boldsymbol{Y}_{\mathrm{p}}$ compared with the MVUE. Equivalently, down-weighting the MVUE estimate suitably would result in a lower MSE. Since the MVUE is an unbiased estimator, the effect of down-weighting is looked upon as 'biasing' the unbiased estimate for obtaining better MSE. The subject of biased estimation deals with designing estimators that compute an appropriate down-weighting factor for obtaining a better MSE when $\boldsymbol{R}_{\boldsymbol{H}_{\mathrm{p}},\boldsymbol{H}_{\mathrm{p}}}$ (channel statistics) is not known.

Stein's work on estimating the mean of a multivariate Gaussian distribution [8] resulted in the development of a family of biased non-linear estimators, which dominate the unbiased estimators uniformly for all the values of a general parameter vector provided that the dimension of the parameter vector being estimated is strictly greater than two. Applying the James–Stein (JS) estimation method [8, 9] to CFR estimation gives us the estimates on the pilot locations as

$$\widehat{\boldsymbol{H}}_{\mathrm{JS,p}} = \left(1 - \frac{(\tilde{P}-2)}{\widehat{\boldsymbol{H}}^{\mathrm{H}}_{\mathrm{ML,p}}\boldsymbol{C}^{-1}\widehat{\boldsymbol{H}}_{\mathrm{ML,p}}}\right)^{+}\widehat{\boldsymbol{H}}_{\mathrm{ML,p}} \tag{5}$$

doi: 10.1049/iet-com.2011.0910

where $(x)^+ = x$ if $x > 0$; else it is equal to zero. Here $\tilde{P} = ((\mathrm{Tr}(\boldsymbol{C}))/(\lambda_{\max}(\boldsymbol{C})))$, with $\mathrm{Tr}(\boldsymbol{C})$ and $\lambda_{\max}(\boldsymbol{C})$ denoting the trace and the largest eigenvalue of $\boldsymbol{C}$, respectively, where $\boldsymbol{C}$ is the covariance matrix of $\widehat{\boldsymbol{H}}_{\mathrm{ML,p}}$.

When the parameter vector being estimated lies close to the origin, the JS estimator exhibits a markedly improved performance when compared with the ML estimator since it shrinks the ML estimate towards the origin. Hence, to further improve the performance of the JS estimator one should shrink the ML estimate towards a vector value which is close to the actual value of the CFR at the pilots. This modified JS estimate of the CFR is given by

$$\widehat{\boldsymbol{H}}_{\mathrm{JS,p}} = \tilde{\boldsymbol{H}} + \left(1 - \frac{(\tilde{P}-2)}{(\widehat{\boldsymbol{H}}_{\mathrm{ML,p}} - \tilde{\boldsymbol{H}})^{\mathrm{H}} \boldsymbol{C}^{-1} (\widehat{\boldsymbol{H}}_{\mathrm{ML},p} - \tilde{\boldsymbol{H}})}\right)^+ (\widehat{\boldsymbol{H}}_{\mathrm{ML,p}} - \tilde{\boldsymbol{H}}) \quad (6)$$

where $\tilde{\boldsymbol{H}}$ is the $P \times 1$ vector shrinkage target vector towards which one wants to shrink the ML estimate. The closer the shrinkage target is to $\boldsymbol{H}_{\mathrm{p}}$, the better is the estimate $\widehat{\boldsymbol{H}}_{\mathrm{JS,p}}$. Our main contribution in this work is in determining an appropriate vector shrinkage target $\tilde{\boldsymbol{H}}$, which varies adaptively with CFR variation. It is proposed to formulate multiple hypothesis tests to decide the extent of frequency selectivity and time selectivity in a RB, and then use the outcome of these hypothesis tests to decide between different shrinkage targets. The test for frequency selectivity and time selectivity can be independent. Although this approach can exploit both the time and frequency selectivity of the CFR for determining the shrinkage target, in the rest of the paper the focus is on a case where the time selectivity is negligible.

Based on the pilot locations and the dimensions of the RB, the pilots can be grouped into r disjoint clusters $\{Z_i\}$ each consisting of p_i pilots for $1 \le i \le r$. The grouping of pilots into clusters depends on whether one wants to focus on time and/or frequency selectivity of CFR and to what extent one wants to quantify the selectivity of the CFR (for example, a cluster can include pilot subcarriers adjacent in frequency or if one wants to differentiate between selectivity seen in adjacent subcarriers, a cluster can comprise of just one subcarrier). From (2), it follows that the sample mean m_i of the ML CFR estimates in the ith cluster is normally distributed with $E[m_i] = (1/p_i)\sum_{k,n \in Z_i} H_{k,n}$. When m_i, $\forall i$ are statistically similar, it implies negligible frequency selectivity in the RB. When all the cluster means differ 'significantly' from each other, it implies significant frequency selectivity in that RB. If only a subset of the means are statistically equal, then it implies that the CFR is moderately selective within the RB. To mathematically differentiate between flat, highly selective and moderately selective CFRs, we propose to use hypothesis tests for testing equality of the sample means of the different clusters.

3.1 Hypothesis test to determine frequency/time selectivity over the RB

Testing for frequency selectivity within a RB constitutes a multiple test of hypothesis checking for pairwise equality of the mean value of the ML CFR estimate of each of the r clusters with every other cluster [Here each cluster has (P/r)] samples. Define $\boldsymbol{m}_i = [\mathrm{real}(m_i)\ \mathrm{imag}(m_i)]^{\mathrm{T}}$ where real (.) denotes the real part and imag(.) denotes the imaginary part. The following null hypothesis $H_{0,\,(i,\,j)}$ for the equality of means is tested

$$H_{0,(i,j)} : \boldsymbol{m}_i = \boldsymbol{m}_j \text{ against } H_{1,(i,j)} : \boldsymbol{m}_i \neq \boldsymbol{m}_j \ \forall \quad i \neq j \quad (7)$$

The statistical U_{ij} for testing equality of the mean of the ith cluster with respect to the mean of the jth cluster is given by

$$U_{ij} = (\boldsymbol{m}_i - \boldsymbol{m}_j)^{\mathrm{T}} \left(\frac{\boldsymbol{R}_i + \boldsymbol{R}_j}{(P/r)}\right)^{-1} (\boldsymbol{m}_i - \boldsymbol{m}_j) = \frac{P}{\sigma^2 r} (\boldsymbol{m}_i - \boldsymbol{m}_j)^{\mathrm{T}} (\boldsymbol{m}_i - \boldsymbol{m}_j) \quad (8)$$

where $\boldsymbol{R}_i = (\sigma^2/2)\boldsymbol{I}_2\ \forall i$ is the covariance matrix of any sample in the ith cluster. A sample in any of the clusters is nothing but a real vectorised representation of one of the ML estimates. Since each ML estimate has the variance σ^2 (Since $\boldsymbol{C} = \sigma^2 \boldsymbol{I}$), therefore the real vectorised version (the 2×1 vector) of the ML estimate has the covariance matrix $(\sigma^2/2)\boldsymbol{I}_2$.

U_{ij} follows the $\chi^2(2)$ distribution under the null hypothesis $H_{0,(i,\ j)}$ and the non-central χ^2 distribution under the alternative hypothesis $H_{1,(i,\ j)}$ with non-centrality parameter λ_{ij}. Pairwise testing for the equality of cluster means results in $\binom{r}{2}$ number of tests, whose outcome is used to decide the shrinkage targets. Depending on the deployed OFDM system (for example LTE-A or 802.16m), the number of clusters, their size and location will vary. As the number of clusters increases, the number of shrinkage targets will increase, and the strategy for shrinkage target selection will become more complex. The decision strategy to choose the shrinkage targets in the case of $r = 3$ is given below. (see (9)) Here, $\tilde{m}_i$ is the shrinkage target for the ith cluster. Note that the decision strategy in (9) covers all possible outcomes of the three hypothesis tests. If all the tests, or at least two out of the three tests satisfy the null hypothesis, then all the m_i's are statistically similar implying that either the CFR is flat, or the noise is high enough to mask its frequency selectivity in the RB. Consequently, averaging the CFR estimates over all the pilots results in an improved noise averaged CFR estimate and therefore the ML CFR estimates should be shrunk towards the mean value of all the clusters. Although, if the tests do not satisfy the null hypothesis for even one pair of means, then there is significant frequency selectivity within the RB and hence the shrinkage target for each cluster should be only its own cluster mean. If the null hypothesis is satisfied for only one pair of means, then the

$$\tilde{m}_i = \begin{cases} \dfrac{1}{r}\sum_{i=1}^{r} m_i & H_{0,(i,j)} : \boldsymbol{m}_i = \boldsymbol{m}_j, \text{ for at least } r-1 \text{ combination of the cluster means} \\ m_i & H_{0,(i,j)} : \boldsymbol{m}_i = \boldsymbol{m}_j \text{ is not satisfied form}_i \text{ for any } j \neq i \\ \dfrac{m_i + m_j}{2} & H_{0,(i,j)} : \boldsymbol{m}_i = \boldsymbol{m}_j \text{ is satisfied for only one pair } m_i, m_j \end{cases} \quad (9)$$

shrinkage target for the corresponding pair of clusters should be the mean of CFR values of both those clusters, whereas the shrinkage target for the third cluster should be its own mean CFR value.

Although using multiple hypothesis tests to infer about the frequency and time selectivity of the CFR in the RB, the error in the actual decision comprises of the combined error in the multiple tests. In the section that follows the Types I and II error probability for a single test is computed and then the probability of making an error while choosing a specific shrinkage target is derived.

3.2 Trade-off between Types I and II error and its impact

Having chosen a threshold γ_{ij} for $H_{0,(i,\ j)}$, the probability of erroneously deciding against the equality of cluster means, (Type I error) is $P(U_{ij} > \gamma_{ij}/H_{0,(i,\ j)})$. Also, the probability of incorrectly accepting the null hypothesis (Type II error), whereas cluster means are unequal is $P(U_{ij} < \gamma_{ij}/H_{1,(i,\ j)})$. Specific to our application, a Type I error indicates that the test statistical U_{ij} implies that the CFR over a region of the RB is selective even though it is actually flat. On the other hand in case of Type II error, U_{ij} indicates that the CFR is flat over a region when it actually is selective. At high signal-to-noise ratios (SNRs), Type II error will dominate the performance since when noise is low exploiting the structure of the CFR is important, whereas at low SNRs Type I error will be more prominent.

To compute the probabilities of Types I and II errors the probability density function (pdf) of U_{ij} is required. Since m_i is 2×1 vector, the pdf of U_{ij} is given by

$$f_{U_{ij}}(u) = \begin{cases} \frac{1}{2}\mathrm{e}^{-(u/2)}, & \text{under } H_{0,(i,j)} \\ \frac{1}{2}\mathrm{e}^{-((\lambda_{ij}+u)/2)} I_0\left(\sqrt{\lambda_{ij} u}\right), & \text{under } H_{1,(i,j)} \end{cases} \tag{10}$$

Here, $I_0(.)$ is the zeroth order modified Bessel function and λ_{ij} is the non-centrality parameter which depends on the pdf of $E[m_i - m_j] = (1/p_j)\sum_{k,n\in Z_i} H_{k,n} - (1/p_j)\sum_{k,n\in Z_j} H_{k,n}$. To simplify the analysis, it is assumed that CFR is nearly constant in a cluster, $(1/p_i)\sum_{k,n\in Z_i} H_{k,n} \simeq H_{p,n}$, where $H_{p,\ n}$ is a CFR value such that $\{p,\ n\} \in Z_i$ and similarly $(1/p_j)\sum_{k,n\in Z_j} H_{k,n} \simeq H_{q,m}$, where $H_{q,\ m}$ is a CFR value such that $\{q,\ m\} \in Z_j$. Assuming negligible time selectivity within the RB, the time indices m and n are dropped. Both H_p and H_q are identically distributed but correlated Gaussian random variables with mean zero and variance $2\sigma^2_{\mathrm{PDP}} = \sum_{i=1}^{L} \sigma_i^2$, where $\sigma_i^2 = E[h_i h_i^*]$ and L is the number of channel taps. Since the channel is assumed to unit energy $2\sigma^2_{\mathrm{PDP}} = 1$. Here $m_{ij} = E[m_i - m_j]$ is given by, (see (11))

$$m_{ij} = \begin{bmatrix} m_{ij}(1) \\ m_{ij}(2) \end{bmatrix} = \begin{bmatrix} \sum_{k=0}^{L-1}\sum_{i=0}^{L-1} h_{i\mathrm{R}} \cos\left(\frac{2\pi ip}{N}\right) + h_{i\mathrm{I}} \sin\left(\frac{2\pi ip}{N}\right) - h_{k\mathrm{R}} \cos\left(\frac{2\pi kq}{N}\right) - h_{k\mathrm{I}} \sin\left(\frac{2\pi kq}{N}\right) \\ \sum_{k=0}^{L-1}\sum_{i=0}^{L-1} -h_{i\mathrm{R}} \sin\left(\frac{2\pi ip}{N}\right) + h_{i\mathrm{I}} \cos\left(\frac{2\pi ip}{N}\right) + h_{k\mathrm{R}} \sin\left(\frac{2\pi kq}{N}\right) - h_{k\mathrm{I}} \cos\left(\frac{2\pi kq}{N}\right) \end{bmatrix} \tag{11}$$

where N denotes the fast Fourier transform (FFT) size, $h_{i\mathrm{R}}$ is the real part of the ith channel tap, $h_{i\mathrm{I}}$ is the imaginary part and $0.5\sigma_i^2 = E[h_{i\mathrm{R}} h_{i\mathrm{R}}] = E[h_{i\mathrm{I}} h_{i\mathrm{I}}]$. Here, $\sum_{i=0}^{L-1} h_{i\mathrm{R}} \cos(2\pi ip/N) + h_{i\mathrm{I}} \sin(2\pi ip/N)$ and $\sum_{k=0}^{L-1} h_{k\mathrm{R}} \cos(2\pi kq/N) + h_{k\mathrm{I}}$ $\mathcal{N}(0, \sigma^2_{\mathrm{PDP}})$ are Gaussian random variables with distribution $\mathcal{N}(0, \sigma^2_{\mathrm{PDP}})$.

The difference between two correlated Gaussian variables is a Gaussian random variable with mean given by the difference between the two means and variance given by the sum of individual variances minus twice the covariance between the two correlated variables [10, p. 158]. Therefore elementary calculations for the difference of the cluster means gives $m_{ij}(1) \sim \mathcal{N}(0, 2\sigma^2_{\mathrm{PDP}} - 2\sigma^2_{ij})$ and $m_{ij}(2) \sim \mathcal{N}(0, 2\sigma^2_{\mathrm{PDP}} - 2\sigma^2_{ij})$, where σ^2_{ij} (the covariance between the two correlated Gaussian random variables) is defined as

$$\sigma^2_{ij} = \sum_{k=0}^{L-1}\sum_{l=0}^{L-1}\left(h_{l\mathrm{R}} \cos\left(\frac{2\pi lp}{N}\right) + h_{l\mathrm{I}} \sin\left(\frac{2\pi lp}{N}\right)\right) \times \left(h_{k\mathrm{R}} \cos\left(\frac{2\pi kq}{N}\right) + h_{k\mathrm{I}} \sin\left(\frac{2\pi kq}{N}\right)\right) \tag{12}$$

Using the fact that $E[h_{i\mathrm{R}} h_{i\mathrm{I}}] = 0\ \forall i$, $E[h_{i\mathrm{R}} h_{j\mathrm{R}}] = 0\ \forall i \neq j$ and $E[h_{i\mathrm{I}} h_{j\mathrm{I}}] = 0\ \forall i \neq j$, we obtain

$$\sigma^2_{ij} = \sum_{k=0}^{L-1} 0.5\sigma_k^2 \cos\left(\frac{2\pi k(p-q)}{N}\right) \tag{13}$$

Furthermore, using the fact that the channel taps are uncorrelated one can show that $E[m_{ij}(1) m_{ij}(2)] = 0$, and hence $m_{ij}(1)$ and $m_{ij}(2)$ are independent and identically distributed Gaussian random variables. Therefore the non-centrality parameter $\lambda_{ij} = ((m_{ij}(1)^2 + m_{ij}(2)^2)/(0.5\sigma^2))$ is an exponentially distributed random variable with pdf given by

$$f_{\lambda_{ij}}(x) = \frac{0.5\sigma^2}{(\sigma^2_{\mathrm{PDP}} - \sigma^2_{ij})} \mathrm{e}^{-((0.5\sigma^2 x)/(\sigma^2_{\mathrm{PDP}} - \sigma^2_{ij}))} \tag{14}$$

Note that λ_{ij} is a function of the distance (in terms of subcarriers) between the two clusters whose means are being compared. Having obtained the parameters of the pdf $f_{U_{ij}}(u)$ in (10), we now proceed to compute the probability of Types I and II errors. Since the pdf of U_{ij} is given by (10), the corresponding probability of Types I and II errors is given by

$$P_{1,ij} = P(U_{ij} > \gamma_{ij}/H_{0,(i,j)}) = \Gamma\left(1, \frac{\gamma_{ij}}{2}\right) = \mathrm{e}^{-(\gamma_{ij}/2)} \tag{15}$$

$$P_{2,ij} = P(U_{ij} < \gamma_{ij}/H_{1,(i,j)}) = 1 - Q_1\left(\sqrt{\lambda_{ij}}, \sqrt{\gamma_{ij}}\right) \tag{16}$$

where $Q_u(x, y)$ is the generalised Marcum Q-function. Since λ_{ij} is a random variable, we are interested in the average value of the Type II error, which is given by

$$E[P_{2,ij}] = E[P(U_{ij} < \gamma_{ij}/H_{1,(i,j)})] = \int_0^\infty \left(1 - Q_1\left(\sqrt{\lambda_{ij}}, \sqrt{\gamma_{ij}}\right)\right) f_{\lambda_{ij}}(\lambda)\mathrm{d}\lambda \tag{17}$$

doi: 10.1049/iet-com.2011.0910

which evaluates to

$$E[P_{2,ij}] = 1 - \mathrm{e}^{-((\sigma^2\gamma_{ij})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{ij}))} \tag{18}$$

It is apparent from (18) that to compute the probability of Type II error one requires knowledge of the channel PDP. However, since one does not know the channel PDP, $E[P_{2,ij}]$ is computed assuming that the alternate hypothesis $H_{1,(i,\,j)}$ corresponds to the channel having an uniform PDP, that is

$$\sigma^2_{ij} = \frac{1}{2L}\sum_{k=0}^{L-1}\cos\left(\frac{2\pi k(p-q)}{N}\right)$$

3.3 Impact of multiple hypothesis tests

Since we use multiple hypothesis tests to determine a shrinkage target for the JS estimator, the thresholds for these tests, γ_{ij}, should be chosen such that the probability of incorrectly choosing any of the shrinkage targets is bounded by a small value. The decision strategy in (9) leads to five shrinkage targets, and we will study the probability of each shrinkage target being incorrectly chosen because of Type I and/or Type II error in each of the three hypothesis tests. Here $(1/r)\sum_{i=1}^{r} m_i$ is chosen as the shrinkage target when the outcomes of the hypothesis tests indicate that at least two pair of cluster means are equal. Hence, the probability of incorrectly choosing $(1/r)\sum_{i=1}^{r} m_i$ as a shrinkage target $\forall i$ is given by the probability of at least two Type II errors in the three hypothesis tests. This error probability can be computed as (see (19))

It is assumed that in the absence of 'prior' information, both the null and alternate hypothesis are equi-probable (In a Bayesian framework, in the absence of prior knowledge both outcomes are assumed to be equi-probable. However, depending on factors such as inter-cluster distance, information about the terrain, any information about the channel statistics, one can choose different values for $P(H_{0,(i,j)})$ and $P(H_{1,(i,j)})$.) and hence $P(H_{0,(i,j)}) = P(H_{1,(i,j)}) = (1/2)$. Using (15) and (18), P_1 can be rewritten as (see (20)) where

$\mathcal{S}_2 = \{\{i,j\},\{j,k\},\{k,i\} : ij \neq jk \neq ki, \{i,j\} \in \{\{1,2\},\{2,3\},\{3,1\}\}, \{j,k\} \in \{\{1,2\},\{2,3\},\{3,1\},\{k,i\}\} \in \{1,2\},\{2,3\},\{\{3,1\}\}\}$.

Similarly, the probability of incorrectly choosing the shrinkage target as $\tilde{m}_i = m_i\ \forall i$ is given by the probability of making one or more Type I errors in the the three hypothesis tests which is given by (see (21))

In the case of the other three shrinkage targets given by the decision strategy outlined in (9), the probability of error in choosing the shrinkage target is the probability of making one or more Type I and/or Type II errors. We have given below the probability of erroneously choosing the shrinkage target $\tilde{m}_1 = \tilde{m}_2 = ((m_1 + m_2)/2)$ and $\tilde{m}_3 = m_3$ (see (22)).

The probability of making an error in choosing the other two shrinkage targets, that is, P_4 and P_5 can be computed in a similar fashion. Using these expressions for error probabilities one can find the values of $\gamma_{1,\,2}$, $\gamma_{2,\,3}$ and $\gamma_{3,\,1}$, which lead to an acceptable (small) value for error probability for each of the shrinkage targets. Note that if one wants to decrease the value of P_1 then it leads to an increase in P_2 and there exists a trade-off between the error probabilities for the five shrinkage targets. The accuracy with which $\tilde{\boldsymbol{H}}$ reflects the actual CFR at the pilot locations is a function of P_i, $i = 1, \ldots, 5$. Smaller the P_i, $i = 1, \ldots, 5$ values closer is $\tilde{\boldsymbol{H}}$ to the actual CFR. Since the value of P_i,

$$\begin{aligned}
P_1 =& E[P(U_{12} < \gamma_{12}|H_{1,(1,2)})]E[P(U_{23} < \gamma_{23}|H_{1,(2,3)})]E[P(U_{31} < \gamma_{31}|H_{1,(3,1)})]P(H_{1,(1,2)})P(H_{1,(2,3)})P(H_{1,(3,1)}) \\
&+ E[P(U_{12} < \gamma_{12}|H_{1,(1,2)})]E[P(U_{23} < \gamma_{23}|H_{1,(2,3)})]P(U_{31} < \gamma_{31}|H_{0,(3,1)})P(H_{1,(1,2)})P(H_{1,(2,3)})P(H_{0,(3,1)}) \\
&+ E[P(U_{12} < \gamma_{12}|H_{1,(1,2)})]P(U_{23} < \gamma_{23}|H_{0,(2,3)})E[P(U_{31} < \gamma_{31}|H_{1,(3,1)})]P(H_{1,(1,2)})P(H_{0,(2,3)})P(H_{1,(3,1)}) \\
&+ P(U_{12} < \gamma_{12}|H_{0,(1,2)}E[P(U_{23} < \gamma_{23}|H_{1,(2,3)})]E[P(U_{31} < \gamma_{31}|H_{1,(3,1)})]P(H_{0,(1,2)})P(H_{1,(2,3)})P(H_{1,(3,1)})
\end{aligned} \tag{19}$$

$$P_1 = \frac{1}{8}\left(\sum_{\mathcal{S}_2}(1-\mathrm{e}^{-((\sigma^2\gamma_{ij})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{ij}))})(1-\mathrm{e}^{-((\sigma^2\gamma_{jk})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{jk}))})(1-\mathrm{e}^{-((\gamma_{ki})/2)}) + \prod_{\mathcal{S}_1}(1-\mathrm{e}^{-((\sigma^2\gamma_{ij})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{ij}))})\right) \tag{20}$$

$$\begin{aligned}
P_2 = \frac{1}{8}\Bigg(&\prod_{\mathcal{S}_1}\mathrm{e}^{-((\gamma_{ij})/2)} + \sum_{\mathcal{S}_2}\mathrm{e}^{-((\gamma_{ij})/2)}\mathrm{e}^{-((\gamma_{jk})/2)}\mathrm{e}^{-((\sigma^2\gamma_{ki})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{ki}))} \\
&+ \sum_{\mathcal{S}_2}\mathrm{e}^{-((\gamma_{ij})/2)}\mathrm{e}^{-((\sigma^2\gamma_{jk})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{jk}))}\mathrm{e}^{-((\sigma^2\gamma_{ki})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{ki}))}\Bigg)
\end{aligned} \tag{21}$$

$$P_3 = \frac{1}{8}\left(\begin{array}{l}
(1-\mathrm{e}^{-((\gamma_{12})/2)})(\mathrm{e}^{-((\gamma_{23})/2)}(\mathrm{e}^{-((\gamma_{31})/2)} + \mathrm{e}^{-((\sigma^2\gamma_{31})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{31}))}) + \mathrm{e}^{-((\gamma_{31})/2)}\mathrm{e}^{-((\sigma^2\gamma_{23})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{23}))}) \\
+(1-\mathrm{e}^{-((\sigma^2\gamma_{12})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{12}))}) \\
\times\left(\begin{array}{l}
\mathrm{e}^{-((\sigma^2\gamma_{23})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{23}))}\mathrm{e}^{-((\sigma^2\gamma_{31})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{31}))} + \mathrm{e}^{-((\sigma^2\gamma_{23})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{23}))}\mathrm{e}^{-((\gamma_{31})/2)} \\
+\mathrm{e}^{-((\sigma^2\gamma_{31})/(0.5\sigma^2+4\sigma^2_{\mathrm{PDP}}-4\sigma^2_{31}))}\mathrm{e}^{-((\gamma_{23})/2)} + \mathrm{e}^{-((\gamma_{23})/2)}\mathrm{e}^{-((\gamma_{31})/2)}
\end{array}\right)
\end{array}\right) \tag{22}$$

$i = 1, \ldots, 5$ depend on the choice of γ_{ij}, the accuracy of the shrinkage target is a function of the γ_{ij} values.

Once an appropriate shrinkage target $\tilde{\boldsymbol{H}}$ is chosen based on the outcome of the hypothesis tests and the JS estimates $\widehat{\boldsymbol{H}}_{\mathrm{JS,p}}$ are computed on the pilot location using the selected $\tilde{\boldsymbol{H}}$ in order to compute (6), then these estimates are interpolated onto the rest of the time–frequency grid in the RB using the robust 2D-MMSE filter. The final CFR estimate $\widehat{\boldsymbol{H}}_{\mathrm{JS-2D}} \in \mathbb{C}^{QR\times 1}$ over the entire RB is given by

$$\widehat{\boldsymbol{H}}_{\mathrm{JS-2D}} = \boldsymbol{W}_{\mathrm{rob}}\widehat{\boldsymbol{H}}_{\mathrm{JS,p}} \tag{23}$$

The efficacy of the hypothesis test with the appropriate γ_{ij} values in detecting the frequency selectivity of the CFR within the RB is evident from the simulation results presented in Section 4. These hypothesis tests help to choose the shrinkage target $\tilde{\boldsymbol{H}}$ which best represents the structure of the CFR within the RB for the operating SNR. Even for the same channel model, different RBs may see selectivity to a varying extents and performing hypothesis tests for each RB will lead to maximum exploitation of the time and frequency varying nature of the CFR.

3.4 Computational complexity

The computational complexity of the proposed estimator is compared with the robust 2D-MMSE estimator. Assuming that the $\boldsymbol{W}_{\mathrm{rob}}$ matrix can be precomputed and stored for a range of SNRs, the computation complexity of the robust 2D-MMSE estimator is due to the multiplications and additions in $\boldsymbol{W}_{\mathrm{rob}}\widehat{\boldsymbol{H}}_{\mathrm{ML,p}}$ which is equivalent to 2QRP multiplications and 2QRP additions since $\boldsymbol{W}_{\mathrm{rob}}$ is a real matrix. By multiplications, additions and divisions we mean real multiplications, real additions and real divisions, respectively. All complex multiplications and complex addition have been converted to their equivalent real multiplications and real additions. The computational cost for the proposed estimator consists of the complexity of selecting the shrinkage target $\tilde{\boldsymbol{H}}$, the complexity of the computation in (6) and the complexity of (23). To select the appropriate $\tilde{\boldsymbol{H}}$, one needs to compute U_{ij} [i.e. (8)] for all the cluster pairs, compare the computed U_{ij} values with the corresponding thresholds γ_{ij}, and then finally compute $\tilde{m}_i$ using (9). Note that similar to the robust 2D-MMSE estimator the thresholds γ_{ij} can be precomputed and stored for a range of SNRs.

Computing U_{ij} for a pair of clusters involves ten additions, eight multiplications and one division when $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$ are not diagonal matrices and additionally $2P$ additions which is common across clusters. In our case where $\boldsymbol{R}_i = \boldsymbol{R}_j = (\sigma^2/2)\boldsymbol{I}_2$, the U_{ij} computation for a pair of clusters involves three additions and two multiplications and one division and additionally $2P$ additions which is common across all the clusters. Comparing the computed U_{ij} values with the corresponding thresholds γ_{ij} requires one comparison per pair of clusters. The number of operations for computing $\tilde{m}_i$ depends on the decision strategy. For the decision strategy given by (9) for $r = 3$, on an average (i.e. under the assumption that all the five shrinkage are equi-probable when one averages over all channel conditions the computations involved for each shrinkage target is divided by five) (8/5) additions, (6/5) multiplications and (2/5) divisions are required. Hence, for $r = 3$ clusters the total computational complexity for calculating $\tilde{\boldsymbol{H}}$ when $\boldsymbol{R}_i = \boldsymbol{R}_j = (\sigma^2/2)\boldsymbol{I}_2$ is $2P + 3r + (8/5)$ additions, $2r + (6/5)$ multiplications, $r + (2/5)$ divisions and r comparisons.

The computation in (6) for the case when $\boldsymbol{C} = \sigma^2\boldsymbol{I}$ requires $5P + 2$ additions, $4P + 1$ multiplications, 1 division and 1 comparison. Finally, (23) requires $2QRP$ multiplications and $2QRP$ additions. Hence, the total number of operations required for the proposed estimator when $r = 3$ is $11P + 26 + (16/5) + 4QRP$ operations. An operation is any one of the following: addition, subtraction, multiplication, finite-precision division, and comparison of two numbers. For example, 50 additions and 100 multiplications are equivalent to 150 operations. For the RB structure assumed in the simulation, $P = 12$, $Q = 18$ and $R = 6$ and in this case the computational complexity of the proposed estimator is 5345.2 operations, whereas the complexity of the robust 2D-MMSE is $4\mathrm{QRP} = 5184$ operations. Hence, the proposed estimator is only 3% more complex than the robust 2D-MMSE estimator even though its performance is significantly better.

4 Simulation results

We have considered a system with two transmit and two receive antennas with one of the transmit antennas being virtual. The RB structure shown in Fig. 1 is based on IEEE 802.16m standards and it is of size 18 subcarriers-by-6 OFDM symbols with 12 pilots per RB (in the case of the virtual antenna mode). The cyclic prefix length $L_{\mathrm{CP}} = 32$. As Fig. 1 indicates, the pilots have been clustered into three disjoint groups and hence the number of hypothesis tests are three. The MSE and mean absolute phase error (PE) values are computed by averaging 10 000 random channel realisations for every SNR value for every channel model. For the case of the clusters in Fig. 1, the decision strategy given in (9) results in five possible shrinkage targets, which are given in Table 1.

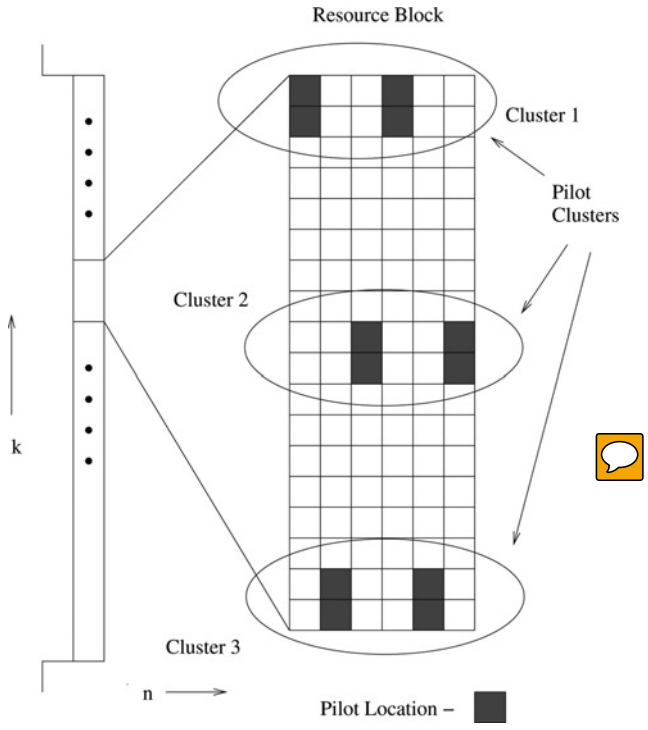

Fig. 1 *OFDM resource block structure*

doi: 10.1049/iet-com.2011.0910

Table 1 Various shrinkage targets

Sl. No	Shrinkage target $\tilde{\boldsymbol{H}}$	Corresponding outcome of hypothesis tests
(i)	$\mathbf{1}_P(1/P)\sum_{i=1}^{P}\widehat{\boldsymbol{H}}_{\mathrm{ML,p},i}$	H_0:$\boldsymbol{m}_i=\boldsymbol{m}_j$, satisfied for at least two pairs of cluster means
(ii)	$[m_1\mathbf{1}_2^{\mathrm{T}}\ \ m_3\mathbf{1}_2^{\mathrm{T}}\ \ m_2\mathbf{1}_2^{\mathrm{T}}\ \ m_1\mathbf{1}_2^{\mathrm{T}}\ \ m_3\mathbf{1}_2^{\mathrm{T}}\ \ m_2\mathbf{1}_2^{\mathrm{T}}]^{\mathrm{T}}$	H_0 not satisfied for even one pair of cluster means
(iii)	$[((m_1+m_2)/2)\mathbf{1}_2^{\mathrm{T}}\ \ m_3\mathbf{1}_2^{\mathrm{T}}\ \ ((m_1+m_2)/2)\mathbf{1}_4^{\mathrm{T}}\ \ m_3\mathbf{1}_2^{\mathrm{T}}\ ((m_1+m_2)/2)\mathbf{1}_2^{\mathrm{T}}]^{\mathrm{T}}$	H_0:$\boldsymbol{m}_i=\boldsymbol{m}_j$ is satisfied for only the pair m_1, m_2.
(iv)	$[((m_1+m_3)/2)\mathbf{1}_4^{\mathrm{T}}\ \ m_2\mathbf{1}_2^{\mathrm{T}}\ \ ((m_1+m_3)/2)\mathbf{1}_4^{\mathrm{T}}\ \ m_2\mathbf{1}_2^{\mathrm{T}}]^{\mathrm{T}}$	H_0:$\boldsymbol{m}_i=\boldsymbol{m}_j$ is satisfied for only the pair m_1, m_3.
(v)	$[m_1\mathbf{1}_2^{\mathrm{T}}\ \ ((m_2+m_3)/2)\mathbf{1}_4^{\mathrm{T}}\ \ m_1\mathbf{1}_2^{\mathrm{T}}\ \ ((m_2+m_3)/2)\mathbf{1}_4^{\mathrm{T}}]^{\mathrm{T}}$	H_0:$\boldsymbol{m}_i=\boldsymbol{m}_j$ is satisfied for only the pair m_2, m_3.

In Table 1, (i) is suited for the case of negligible statistical variation of the CFR, (ii) better suited for significant variation of the CFR and (iii), (iv) and (v) are suitable for the case when the CFR variation is neither highly significant nor completely flat over the RB. Proposed-JS denotes the $\widehat{\boldsymbol{H}}_{\mathrm{JS-2D}}$ estimate which uses the hypothesis tests (7) to decide between the shrinkage targets given in Table 1 for each RB in every trial. In all the figures, JS-2DMMSE denotes the $\widehat{\boldsymbol{H}}_{\mathrm{JS-2D}}$ estimate for which $\tilde{\boldsymbol{H}}$ is always given by (i) in Table 1 and JS1-2DMMSE denotes the $\widehat{\boldsymbol{H}}_{\mathrm{JS-2D}}$ estimate for which $\tilde{\boldsymbol{H}}$ is always given by (ii) in Table 1. In other words, whereas for proposed-JS, $\tilde{\boldsymbol{H}}$ can vary between any of the five vectors given in Table 1 depending on the hypothesis tests, for both JS-2DMMSE and JS1-2DMMSE the shrinkage target $\tilde{\boldsymbol{H}}$ is fixed. Although optimal MMSE refers to the estimator that uses the true PDP and Doppler profile, the plot denoted by robust 2D-MMSE utilises the uniform PDP with L_{CP} number of taps and uniform Doppler profile.

The proposed methods have been tested on a variety of channel models such as: (i) a model with exponential PDP given by $\mathrm{e}^{-[0:L_{\mathrm{CP}}]/2}$ with L_{CP} number of taps, (ii) the Pedestrian A channel model and (iii) the Vehicular A channel model. These are commonly used channel models in the evaluation methodology for LTE 802.16m and WiMaX systems. The MSE is calculated over the entire RB and then averaged over the number of subcarriers in the RB. The fade rate is 5 Hz and both optimal MMSE and robust 2D-MMSE know the fade rate.

Figs. 2–4 study the impact of varying γ_{ij} on the MSE performance of proposed-JS for different SNR values in the case of Pedestrian A, Pedestrian B and Vehicular A channel models. For simplicity we have taken $\gamma_{12}=\gamma_{23}=\gamma_{31}$. It is apparent from these figures that with varying γ_{ij} the MSE

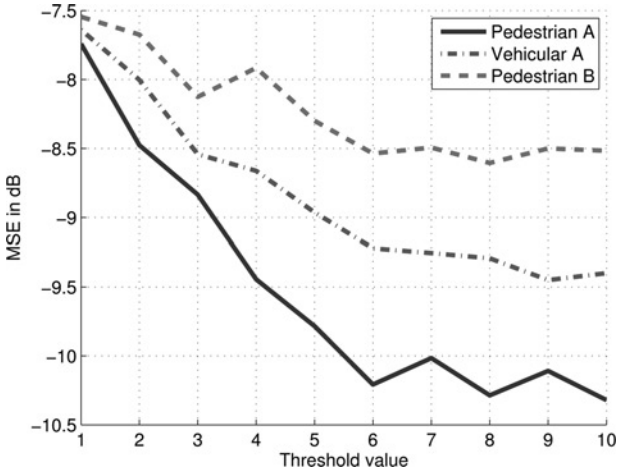

Fig. 2 *MSE performance of proposed estimator as a function of γ_{ij} when SNR = 0 dB*

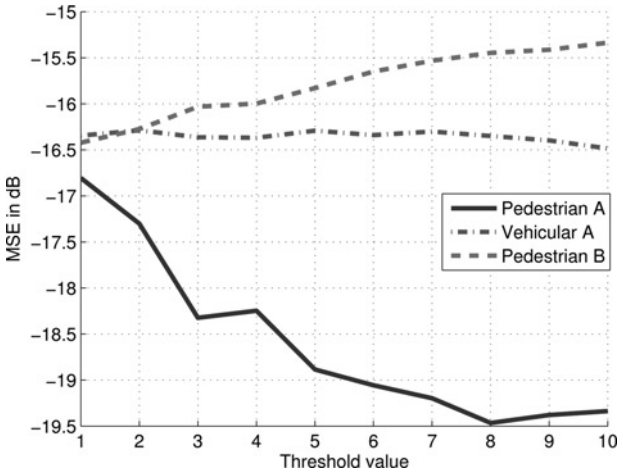

Fig. 3 *MSE performance of proposed estimator as a function of γ_{ij} when SNR = 10 dB*

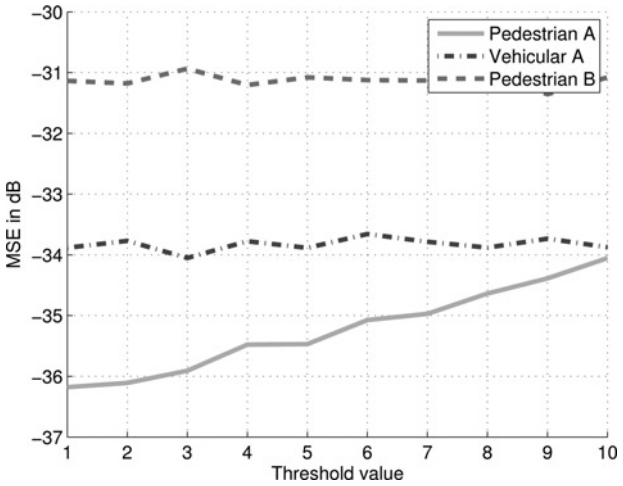

Fig. 4 *MSE performance of proposed estimator as a function of γ_{ij} when SNR = 30 dB*

can either increase or decrease depending on the channel model. For example since Pedestrian A has only two dominant taps, it is relatively flat and in such a case Type I error dominates over Type II error and hence the higher the γ_{ij} values the smaller the MSE over a wide range of SNR. On the other hand, Pedestrian B is much more selective and hence the impact of Type II error is higher and hence a smaller γ_{ij} is preferred. Furthermore with increasing SNR, Type II error dominates the performance and lower γ_{ij} values are preferred, that is, even for Pedestrian A at an SNR of 30 dB, Type II error becomes dominant and the

MSE increases with increasing γ_{ij}. Since we do not know the channel statistics the γ_{ij}s are computed using the expressions for P_1, P_2, P_3, P_4 and P_5 given in Section 3.3 such that the probability of incorrectly choosing any of the shrinkage targets is low. For example, at a SNR of 10 dB the values $\gamma_{12} = \gamma_{23} = \gamma_{31} = 6$ leads to $P_1 = 0.026$, $P_2 = 0.011$ and $P_3 = P_4 = P_5 = 0.03$ which are very small values. At a specific SNR, further increasing γ_{ij} value leads to a higher value of P_1 and a lower value for P_2.

Fig. 5 shows the MSE performance of the various estimators for the case of Pedestrian A channel. The MSE gap between the optimal MMSE and robust 2D-MMSE is between 5 and 10 dB. The JS-2DMMSE, and proposed-JS methods are very close to the optimal MMSE method. The Pedestrian A channel is nearly flat over an RB and hence the JS-2DMMSE method is nearly as good as the proposed-JS method, whereas the JS1-2DMMSE method has poor performance.

Fig. 6 shows the PE performance for the Pedestrian A channel. Here, the absolute PE at a subcarrier is defined as the absolute value of the phase angle between the actual CFR and the estimated CFR at that subcarrier. This value is averaged over the RB to obtain the PE (in radians). The PE is an important statistic since the bit-error rate (BER) is dominated by the PE (and not the MSE) of the CFR in the case of M-PSK modulation. It is apparent from the Fig. 6 that the PE of both the robust 2D-MMSE and the JS1-2DMMSE is very high at low SNRs and this will lead to high BER for M-PSK systems. The PE of the JS-2DMMSE and proposed-JS method are very close to the PE of the optimal MMSE method.

Fig. 7 shows the MSE performance for the case of Vehicular A channel. The proposed-JS and the optimal MMSE methods have similar performance. Since Vehicular A channel is more selective than the Pedestrian A channel the proposed-JS is also significantly better than the JS-2DMMSE especially at SNRs greater than 10 dB wherein the structure of the CFR dominates over the noise. The MSE performance for the case of exponential channel is similar to the MSE performance in Vehicular A channel (see Fig. 8). We have similar results for other channel models such as Pedestrian B channel. For all the

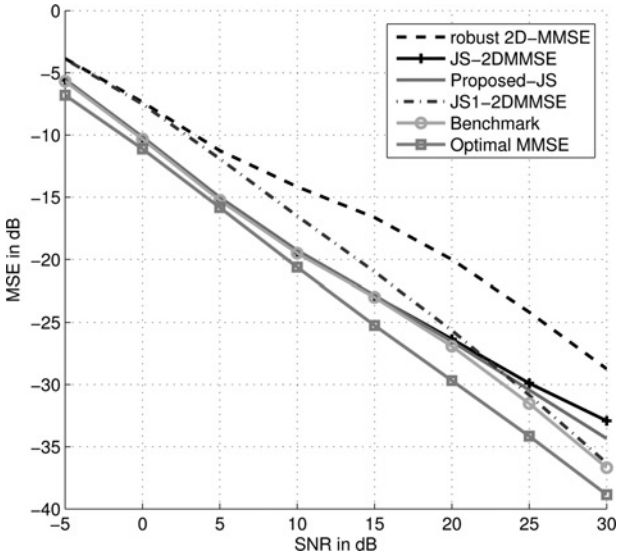

Fig. 5 *MSE performance as a function of SNR for the Pedestrian A channel model*

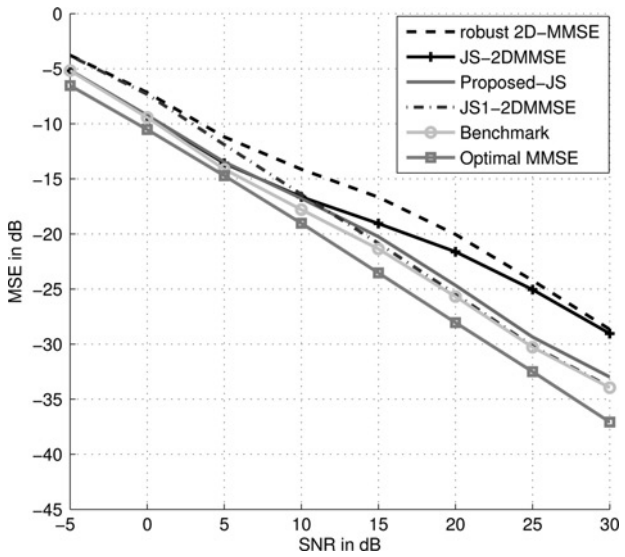

Fig. 7 *MSE performance as a function of SNR for the Vehicular A channel model*

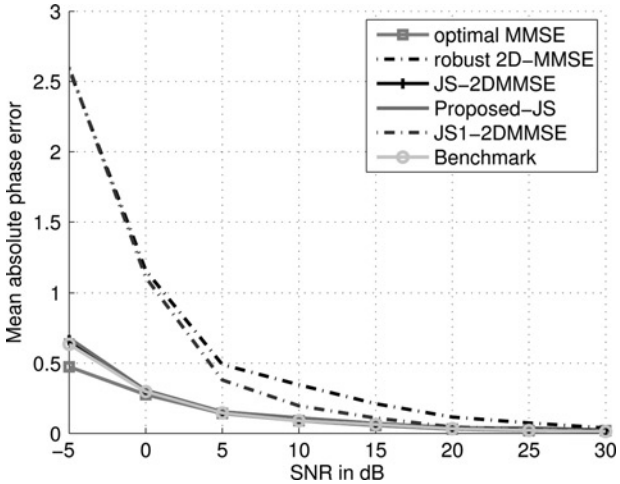

Fig. 6 *PE performance as a function of SNR for the Pedestrian A channel model*

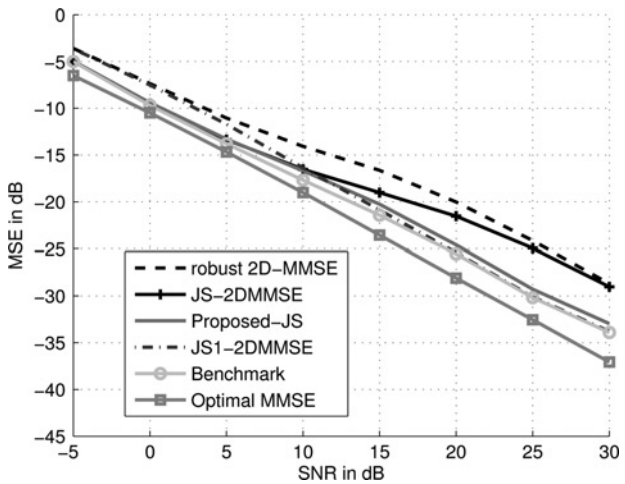

Fig. 8 *MSE performance as a function of SNR for the exponential channel model*

doi: 10.1049/iet-com.2011.0910

channel models considered here, the PE performance of the proposed-JS method is very close to the optimal MMSE method.

The curve marked Benchmark in all the plots refer to the performance obtained if one could choose in each trial from the three methods JS-2DMMSE, JS1-2DMMSE and proposed-JS the method which has the minimum MSE in that particular trial. This curve just serves as a benchmark, since it is not realisable in practice. However, from the figures it is encouraging to see that the proposed-JS method is very close to the benchmark curve, which points to the fact that the hypothesis tests almost always pick the best shrinkage target in every trial.

Fig. 9 shows block error rate (BLER) performance for Pedestrian A channel in the presence of Gaussian noise. A block which comprises of four RBs is considered to be in error if even one data symbol in the block is in error. A BLER of $>10^{-1}$ is mandated by the IEEE 802.16m standard. It is apparent from the figure that the performance of the proposed method is comparable to the optimal MMSE method and 1 dB better than the robust MMSE method. For reference, the BLER performance of the Genie receiver has been included which has perfect knowledge of the channel.

Fig. 10 show the BLER performance in Pedestrian A channel when the pilots see a single interferer with 0 dB relative power, whereas the data symbols see negligible interference because of interference cancellation/mitigation techniques on data. QPSK modulation is used and the BLER in this plot reflects the impact of channel estimation. The proposed method is more than 3 dB better than the robust 2D-MMSE channel estimator at a BLER of 10^{-1} and its performance is comparable with the optimal MMSE performance. Such a gain in BLER performance can be expected since the BLER for QPSK modulation is dominated by the PE and the PE of the robust 2D-MMSE estimator is significantly higher than that of the proposed method at low SNR or low signal-to-interference ratio (SIR) (see Fig. 6 for PE performance).

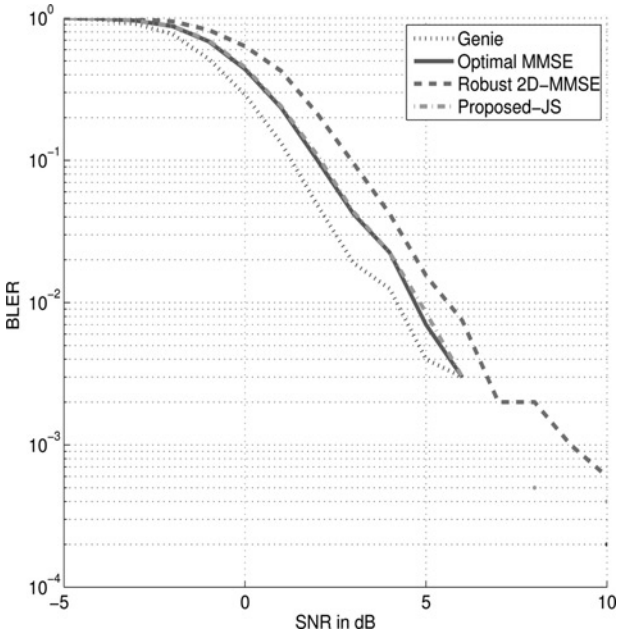

Fig. 9 *BLER performance in Pedestrian A channel for QPSK modulation*

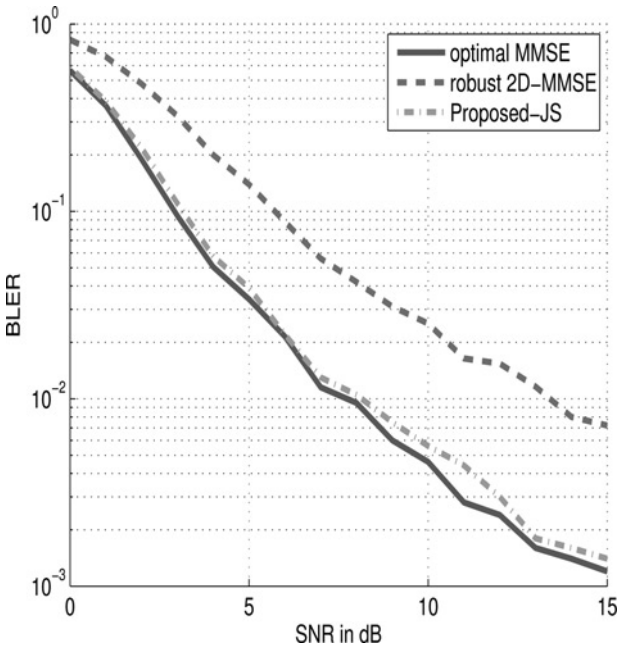

Fig. 10 *BLER performance in Pedestrian A channel for QPSK modulation in the presence of a dominant interferer with SIR 0 dB*

It is apparent from the plots that for different channel models and SNRs, different shrinkage targets are suitable and hence using one single/fixed shrinkage target (for example JS1-2DMMSE and JS-2DMMSE) is not a good idea. Therefore the proposed-JS which chooses between different shrinkage targets adaptively based on both the SNR and the time and frequency selectivity within the RB shows very good performance for a wide range of operating conditions. Furthermore, it is interesting to observe that although the RB considered here spans only 18 subcarriers, it can still have non-negligible selectivity and exploiting this is essential to obtain a performance comparable to that of the optimal MMSE method. Note that even in the high SNR region, the MSE performance of the optimal MMSE estimator is much better than the MSE performance of the robust 2D-MMSE estimator. However from the viewpoint of BLER performance, once the MSE is below a certain threshold both the estimators have a similar performance for the same modulation scheme. If on the other hand adaptive modulation and coding is considered then this MSE gap can be exploited even at high SNRs.

5 Conclusions

A novel method for CFR estimation in OFDMA systems has been proposed for the case when one is restricted to using only the pilots available with an RB. The proposed method uses hypothesis tests on ML estimates at the pilot locations to determine the selectivity of the channel within the RB, and uses the outcome of the tests to design a biased estimator whose shrinkage target varies adaptively with variation in the CFR. Then these biased estimates are interpolated using the simple robust 2D-MMSE filter to give the CFR estimates over the RB. The proposed method shows significant improvement over the robust 2D-MMSE method and is very close to the optimal MMSE method even though it does not require any knowledge of channel statistics. Furthermore, the method has negligible increase in computational complexity when compared with the

robust 2D-MMSE method even though it provides as much as 5–10 dB gain over the robust 2D-MMSE method.

6 Acknowledgments

The financial support of the Department of Science and Technology (DST) India, under the auspices of the Indo-UK Advanced Technology Centre (IU-ATC) in Wireless Communications is gratefully acknowledged. The authors would also like to thank the anonymous reviewers and the Editor for their valuable comments.

7 References

1 Li, Y., Cimini, L., Sollenberger, N.: 'Robust channel estimation for OFDM systems with rapid dispersive fading channels', *IEEE Trans. Commun.*, 1998, **46**, pp. 902–915
2 Hanzo, L., Munster, M., Choi, B.J., Keller, T.: '*OFDM and MC-CDMA for broadband multi-user communications, WLANs and broadcasting*' (IEEE Press, 2002)
3 van de Beek, J.-J., Edfors, O., Sandell, M., Wilson, S., Börjesson, P.: 'On channel estimation in OFDM systems'. Proc. IEEE Vehicular Technology Conf. (VTC'95), July 1995, vol. 2, pp. 815–819
4 Ben-Haim, Z., Eldar, Y.C.: 'Blind minimax estimation', *IEEE Trans. Inf. Theory*, 2007, **53**, pp. 3145–3157
5 Alkhaldi, W., Iskander, D.R., Zoubir, A.M.: 'The applicability of biased estimation in model and model order selection'. Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, ICASSP 2009, April 2009, pp. 3461–3464
6 Lazaro-Gredilla, M., Azpicueta-Ruiz, L.A., Figueiras-Vidal, A.R., Arenas-Garcia, J.: 'Adaptively biasing the weights of adaptive filters', *IEEE Trans. Signal Process.*, 2010, **58**, pp. 3890–3895
7 Auer, G., Karipidis, E.: 'Pilot aided channel estimation for OFDM: a separated approach for smoothing and interpolation'. Proc. of IEEE Int. Conf. on Communications, ICC 2005, May 2005, pp. 2173–2178
8 James, W., Stein, C.: 'Estimation with quadratic loss'. Proc. Fourth Berkeley Symp. on Mathematical Statistics, Berkeley, 1961, vol. 1, pp. 361–379
9 Judge, G.G., Bock, M.E.: 'The statistical implications of pre-test and Stein-rule estimators in econometrics' (North-Holland, Amsterdam, 1978)
10 Papoulis, A.: 'Probability, random variables, and stochastic processes' (McGraw-Hill, New York, 1991, 3rd edn.)

doi: 10.1049/iet-com.2011.0910