ICI Mitigation for Pilot-Aided OFDM Mobile Systems

Yasamin Mostofi, *Member, IEEE* and Donald C. Cox, *Fellow, IEEE*

***Abstract*—Orthogonal frequency-division multiplexing (OFDM) is robust against frequency selective fading due to the increase of the symbol duration. However, for mobile applications channel time-variations in one OFDM symbol introduce intercarrier-interference (ICI) which degrades the performance. This becomes more severe as mobile speed, carrier frequency or OFDM symbol duration increases. As delay spread increases, symbol duration should also increase in order to maintain a near-constant channel in every frequency subband. Also, due to the high demand for bandwidth, there is a trend toward higher carrier frequencies. Therefore, to have an acceptable reception quality for the applications that experience high delay and Doppler spread, there is a need for ICI mitigation within one OFDM symbol. We introduce two new methods to mitigate ICI in an OFDM system with coherent channel estimation. Both methods use a piece-wise linear model to approximate channel time-variations. The first method extracts channel time-variations information from the cyclic prefix. The second method estimates these variations using the next symbol. We find a closed-form expression for the improvement in average signal-to-interference ratio (SIR) when our mitigation methods are applied for a narrowband time-variant channel. Finally, our simulation results show how these methods would improve the performance in a highly time-variant environment with high delay spread.**

***Index Terms*—Channel estimation, intercarrier-interference (ICI) mitigation, mobility, orthogonal frequency-division multiplexing (OFDM).**

I. Introduction

ORTHOGONAL frequency-division multiplexing (OFDM) handles frequency selective fading resulting from delay spread by expanding the symbol duration [1]–[4]. By adding a guard interval to the beginning of each OFDM symbol, the effect of delay spread (provided that there is perfect synchronization) would appear as a multiplication in the frequency domain for a time-invariant channel.[1] This feature allows for higher data rates and has resulted in the selection of OFDM as a standard for digital audio broadcasting (DAB [5]), digital video broadcasting (DVB [6]), and wireless local area networks (802.11a).

Manuscript received November 7, 2002; revised July 7, 2003; accepted November 7, 2003. The editor coordinating the review of this paper and approving it for publication is A. Scaglione. Part of this work was presented in at the International Communications Conference 2003.

Y. Mostofi is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: yasi@cds.caltech.edu).

D. C. Cox is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: dcox@spark.stanford.edu).

Digital Object Identifier 10.1109/TWC.2004.840235

[1]Adding the guard interval will also prevent inter-OFDM symbol-interference.

Transmission in a mobile communication environment is impaired by both delay and Doppler spread. As delay spread increases, symbol duration should also increase for two reasons. First, most receivers require a near-constant channel in each frequency subband. As delay spread increases, this can be achieved by an increase of the symbol length. Second, to prevent inter-OFDM symbol-interference, the length of the guard interval should increase as well. Therefore, to reduce redundancy, the symbol length should increase [7]. OFDM systems become more susceptible to time-variations as symbol length increases. Time-variations introduce ICI, which must be mitigated to improve the performance in high delay and Doppler spread environments.

In [8] and [9], authors analyzed the effect of ICI by modeling it as Gaussian noise. A simplified bound on ICI power has also been derived [10]. To mitigate the introduced ICI, techniques using receiver antenna diversity have been proposed [8], [11]. However, sensitivity analysis has shown that as normalized Doppler spread (defined as the maximum Doppler spread divided by the sub-carrier spacing) increases, antenna diversity becomes less effective in mitigating ICI in OFDM mobile systems [12].

Jeon and Chang have proposed another method for ICI mitigation which assumes a linear model for channel variations [13]. However, they assumed that some of the coefficients of the channel matrix are negligible, which is only the case under low Doppler and delay spread conditions. For instance, their results showed performance improvement under normalized Doppler of up to 2.72% and delay spread of 2 μs for a two-tap channel. In high-mobility applications that require ICI mitigation, however, delay spread can be much longer. For instance, the delay spread can be as high as 40 μs for single frequency network (SFN) channels[2] and 20 μs for cellular applications. Furthermore, normalized Doppler can get as high as 10% depending on the carrier frequency. Their method also relies on the information of adjacent OFDM symbols for channel estimation, which increases processing delay.

To improve the performance in high delay and Doppler spread environments, we present two new ICI mitigation methods in this paper. Unlike the method of Jeon *et al.*, our methods can mitigate ICI in considerably high delay and Doppler spread applications such as SFN and cellular networks. Furthermore, in Method I, we mitigate ICI without relying on the adjacent symbols. Both of our methods are based on a piece-wise linear approximation for channel time-variations.

[2]SFN refers to DAB and DVB type environments in which adjacent base stations transmit in the same frequencies to save the bandwidth.

1536-1276/$20.00 © 2005 IEEE

Authorized licensed use limited to: Perkins Coie LLP. Downloaded on August 12,2020 at 15:25:59 UTC from IEEE Xplore. Restrictions apply.

ZyXEL Communications Corporation Ex. 1027
Page 1 of 10

In the absence of time-variations, frequency domain pilot tones or differential modulation should be used to remove the effect of channel frequency-variations. As the delay spread increases, differential modulation across adjacent subbands degrades the performance. As mobility and/or the length of the OFDM symbol increase, differential modulation across adjacent symbols leads to performance loss as well. Therefore, we use frequency domain pilot tones in this paper since we are dealing with high delay and Doppler spread environments. The minimum number of pilot tones required in each symbol exceeds normalized channel delay spread[3] by one [14]. These pilot tones should be equally spaced in the frequency domain to minimize noise enhancement [14].

In the presence of Doppler spread, however, these pilot tones can not estimate channel time-variations. In this work, we show how to estimate these variations utilizing either the cyclic prefix or the next symbol. Finally, our analysis and simulation results show performance improvement in high delay and Doppler spread environments.

II. System Model

Fig. 1 shows the discrete baseband equivalent system model. We assume perfect timing synchronization in this paper. More information on timing synchronization for a pilot-aided OFDM system can be found in [15]. The available bandwidth is divided into N subchannels and the guard interval spans G sampling periods. We assume that the normalized length of the channel is always less than or equal to G in this paper. X_i represents the transmitted data point in the ith frequency subband and is related to the time domain sequence, x, as follows:

$$X_i = \sum_{k=0}^{N-1} x_k e^{-\frac{j2\pi ki}{N}} \quad 0 \le i \le N-1. \tag{1}$$

$\vec{x}_p$ is the cyclic prefix vector with length G and is related to x as follows:

$$\vec{x}_p(i) = x_{N-G+i} \quad 0 \le i \le G-1. \tag{2}$$

Let T be the time duration of one OFDM symbol after adding the guard interval. Then, $h_k^{(i)}$ represents the kth channel tap at time instant $t = i \times T_s$ where $T_s = T/(N+G)$ is the sampling period. A constant channel is assumed over the time interval $i \times T_s \le t < (i+1) \times T_s$ with $t = 0$ indicating the start of the data part of the symbol. $h_k^{(i)}$ for $-G \le i \le -1$ and $0 \le i \le N-1$ represents the kth channel tap in the guard and data interval respectively.

The channel output y can then be expressed as follows:

$$y_i = \sum_{k=0}^{G} h_k^{(i)} x_{((i-k))_N} + w_i \quad 0 \le i \le N-1. \tag{3}$$

In (3), $((\quad))_N$ represents a cyclic shift in the base of N and w_i represents a sample of additive white Gaussian noise. Then,

[3]Normalized channel delay spread refers to the channel delay spread divided by the sampling period.

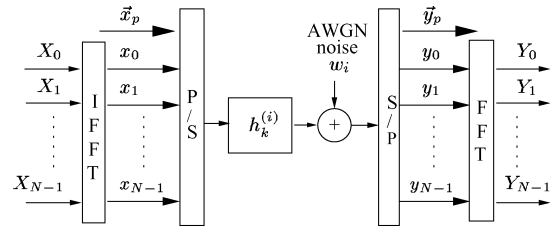

Fig. 1. Discrete baseband equivalent system model.

Y, the fast Fourier transform (FFT) of sequence y, will be as follows:

$$Y_i = H_{i,0}X_i + \underbrace{\sum_{d=1}^{N-1} H_{i,d} X_{((i-d))_N}}_{ICI} + W_i \quad 0 \le i \le N-1 \tag{4}$$

where W denotes the FFT of w and the second term on the right hand side of (4) represents ICI. Define F_m as the FFT of the mth channel tap with respect to time-variations:

$$F_m(k) = \sum_{u=0}^{N-1} h_m^{(u)} e^{-\frac{j2\pi uk}{N}} \quad 0 \le m \le G \quad \& \quad 0 \le k \le N-1. \tag{5}$$

Then $H_{i,d}$ can be defined as

$$H_{i,d} = \frac{1}{N} \sum_{m=0}^{G} F_m(d) e^{-\frac{j2\pi m(i-d)}{N}} \quad 0 \le i, d \le N-1. \tag{6}$$

Furthermore, $H_{i,0} = \sum_{m=0}^{G} h_m^{\text{ave}} e^{\frac{-j2\pi mi}{N}}$ where $h_m^{\text{ave}} = (1/N)\sum_{u=0}^{N-1} h_m^{(u)}$ is the average of the mth channel tap over the time duration of $0 \le t < N \times T_s$. Therefore, $H_{i,0}$ represents the FFT of this average (note that h_m^{ave} solely refers to a time averaging over symbol data part and is different from channel ensemble average).

As was noted by previous work, the ICI term on the right-hand side of (4) can not be neglected as the maximum Doppler shift, f_d, increases (e.g., [8]).

III. Pilot Extraction

Let $\nu \le G$ be the maximum predicted normalized length of the channel. In this paper, we assume that the normalized length of the channel is always smaller than ν. We insert $L \ge \nu + 1$ equally spaced pilots, P_{l_i}, at subchannels $l_i = (i \times N)/L$ for $0 \le i \le L-1$. An estimate of $H_{i,0}$ can then be acquired at pilot tones as follows:

$$\hat{H}_{l_i,0} = \frac{Y_{l_i}}{P_{l_i}} = H_{l_i,0} + \frac{(I_{l_i} + W_{l_i})}{P_{l_i}} \quad 0 \le i \le L-1. \tag{7}$$

In (7), I_{l_i} denotes ICI [marked in (4)] at l_ith subcarrier. Through an IFFT of length L, the estimate of h_k^{ave} would be

$$\hat{h}_k^{\text{ave}} = \frac{1}{L} \sum_{i=0}^{L-1} \hat{H}_{l_i,0} e^{\frac{j2\pi ik}{L}} \quad 0 \le k \le L-1. \tag{8}$$

Authorized licensed use limited to: Perkins Coie LLP. Downloaded on August 12,2020 at 15:25:59 UTC from IEEE Xplore. Restrictions apply.

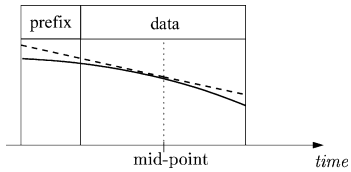

Fig. 2. Piece-wise linear model in one received OFDM symbol. Solid curve: real or imaginary part of a channel path. Dashed line: piece-wise linear model.

In the absence of mobility, L pilots would have been enough to estimate the channel. However in the presence of Doppler, due to the ICI term of (4), using the estimate of $\hat{h}_k^{\text{ave}}$ for data detection results in poor performance. This motivates the need to mitigate the resultant ICI.

IV. Piece-Wise Linear Approximation

In this paper, we approximate channel time-variations with a piece-wise linear model with a constant slope over the time duration T (Fig. 2). For normalized Doppler of up to 20%, linear approximation is a good estimate of channel time-variations and the effect on correlation characteristics is negligible. To see this, Appendix A shows how this approximation affects the correlation function as normalized Doppler increases.

In this section, we will derive the frequency domain relationship, similar to (4), when the linear approximation is applied. Let α_k denote the slope of the kth channel tap in the current OFDM symbol. To perform the linearization, knowledge of the channel at one time instant in the symbol is necessary. Let $E(z)$ represent the average of z. Then for the kth channel tap, $E(|h_k^{\text{ave}} - h_k^{(s)}|^2)$ is minimized for $s = \left(\frac{N}{2} - 1\right)$ as is shown in Appendix B. Therefore, we approximate $h_k^{\left(\frac{N}{2}-1\right)}$ with the estimate of h_k^{ave}. We will have

$$\hat{h}_k^{\left(\frac{N}{2}-1\right)} = \hat{h}_k^{\text{ave}}. \tag{9}$$

Consider linearization around $h_k^{\left(\frac{N}{2}-1\right)}$. Then, $h_k^{(i)}$ can be approximated as follows:

$$h_k^{(i)} \approx h_k^{\left(\frac{N}{2}-1\right)} + \left(i + 1 - \frac{N}{2}\right) \times \alpha_k \times T_s \quad 0 \le i \le N-1. \tag{10}$$

Inserting (10) into (3), we will have

$$\vec{y} \approx (\mathbf{h}_{\text{mid}} + \mathbf{M} \times \mathbf{A}) \times \vec{x} + \vec{w} \tag{11}$$

where $\vec{y}$, $\vec{x}$, and $\vec{w}$ are $N \times 1$ vectors containing samples of y_i, x_i, and w_i for $0 \le i \le N-1$. Furthermore, for $1 \le k, m \le N$, we will have

$$\mathbf{h}_{\text{mid}}(k,m) = \begin{cases} h_{((k-m))_N}^{\left(\frac{N}{2}-1\right)} & 0 \le k-m \le G \quad \& \\ & -(N-1) \le k-m \le -(N-G) \\ 0 & else \end{cases} \tag{12}$$

$$\mathbf{A}(k,m) = \begin{cases} \alpha_{((k-m))_N} & 0 \le k-m \le G \quad \& \\ & -(N-1) \le k-m \le -(N-G) \\ 0 & else. \end{cases} \tag{13}$$

$\mathbf{M}$ is a diagonal matrix with diagonal elements of $\mathbf{M}(k,k) = \beta_{k-1} = T_s \times (k - \frac{N}{2})$ for $1 \le k \le N$. It is shown in Appendix C that taking an FFT of $\vec{y}$ will result in the following frequency domain relationship:

$$\vec{Y} \approx \mathbf{H}_{\text{mid}}\vec{X} + \mathbf{C} \times \mathbf{H}_{\text{slope}}\vec{X} + \vec{W} \tag{14}$$

where

$$\mathbf{H}_{\text{mid}} = \text{diag}\left\{FFT\left(\left[h_0^{\left(\frac{N}{2}-1\right)} \; h_1^{\left(\frac{N}{2}-1\right)} \; \cdots \; h_G^{\left(\frac{N}{2}-1\right)} \; 0 \; \cdots \; 0\right]\right)\right\} \tag{15}$$

$$\mathbf{H}_{\text{slope}} = \text{diag}\{FFT([\alpha_0 \; \alpha_1 \; \cdots \; \alpha_G \; 0 \; \cdots \; 0])\}. \tag{16}$$

Here, $FFT(\vec{z})$ represents the FFT of the vector $\vec{z}$. $\vec{W}$ is a vector containing the FFT of noise samples and $\mathbf{C}_{n,m} = \frac{B_{n-m}}{N}$ where B is the FFT of β and is defined as follows:

$$B_k = T_s \times N \times \begin{cases} -\frac{1}{1-e^{-\frac{j2\pi k}{N}}} & k \ne 0 \\ 0.5 & k = 0 \end{cases} \tag{17}$$

To solve (14) for $\vec{X}$, both $\mathbf{H}_{\text{mid}}$ and $\mathbf{H}_{\text{slope}}$ should be estimated. Matrix $\mathbf{C}$ is a fixed matrix that is precalculated and stored in the receiver. An estimate of $\mathbf{H}_{\text{mid}}$ is readily available from (7)–(9) and (15). In the following subsections, we show how to estimate $\mathbf{H}_{\text{slope}}$ with our two methods. In Method I, this is done by utilizing the redundancy of the cyclic prefix while in Method II the information of the next symbol is used.

A. Method I: ICI Mitigation Using Cyclic Prefix

The output prefix vector, $\vec{y}_p$ of Fig. 1, can be written as follows:

$$\vec{y}_p = \mathbf{Q} \times \vec{p} + \vec{w}_p. \tag{18}$$

In (18), $\vec{p} = \begin{bmatrix} \vec{x}_p^{pre} \\ \vec{x}_p \end{bmatrix}$, $\vec{w}_p$ contains AWGN samples and

$$\mathbf{Q}(i,j) = h_{((i-j+G))_{2\times G}}^{(-G+i-1)} \quad 1 \le i \le G \quad 1 \le j \le 2\times G. \tag{19}$$

Since $\nu \le G$, $h_k^{(i)} = 0$ for $k > G$ in matrix $\mathbf{Q}$. $\vec{x}_p$ is a $G \times 1$ vector defined in (2). $\vec{x}_p^{pre}$ is similarly defined for the transmitted cyclic prefix of the previous OFDM symbol and is already known to the receiver. Define $\vec{\zeta}$ as a vector containing slopes of all the taps

$$\vec{\zeta} = [\alpha_0 \quad \alpha_1 \quad \cdots \quad \alpha_G]^t. \tag{20}$$

Inserting $h_k^{(i)}$ from (10) in $\mathbf{Q}$, it can be easily shown that (18) can be written as follows:

$$\vec{y}_p - \mathbf{R} \times \vec{p} \approx \mathbf{D} \times \mathbf{X}_{p_{\text{mat}}} \times \vec{\zeta} + \vec{w}_p \tag{21}$$

Here, $\mathbf{R}(i,j) = h_{((i-j+G))_{2\times G}}^{\left(\frac{N}{2}-1\right)}$ for $1 \le i \le G$ and $1 \le j \le 2 \times G$ and can be estimated from (7)–(9). $\mathbf{D}$ is a predetermined

Authorized licensed use limited to: Perkins Coie LLP. Downloaded on August 12,2020 at 15:25:59 UTC from IEEE Xplore. Restrictions apply.

TABLE I
PROCEDURE FOR METHOD I

Step	*Command*
1	Set the initial estimate of $\mathbf{H}_{slope}$ to zero
2	Use Eq. 7-9 & 15 to estimate $\mathbf{H}_{mid}$ from pilots
3	Solve Eq. 14 for $\vec{X}$
4	Solve Eq. 21 for $\vec{\zeta}$
5	Use Eq. 16 to estimate $\mathbf{H}_{slope}$
6	If not converged or timed out, go to step 3

diagonal matrix with $\mathbf{D}(i,i) = T_s \times (-\frac{N}{2} + i - G)$ for $1 \le i \le G$ and is stored in the receiver. $\mathbf{X}_{p_{\text{mat}}}$ is defined as follows:

$$\mathbf{X}_{p_{\text{mat}}} = \begin{bmatrix} Rev\,(\vec{p}^t[1:G+1]) \\ Rev\,(\vec{p}^t[2:G+2]) \\ \vdots \\ Rev\,(\vec{p}^t[G:2\times G] \end{bmatrix} \tag{22}$$

where $Rev(\vec{J}[i:j])$ represents the vector produced by reversing the order of elements i through j of vector $\vec{J}$ and $\vec{J}^t$ denotes transpose of $\vec{J}$. Equations (14) and (21) provide enough information to solve for $\vec{X}$. There are two sets of unknowns, $\vec{X}$ and $\vec{\zeta}$ ($\mathbf{H}_{\text{slope}}$ is formed from FFT of $\vec{\zeta}$). It is possible to combine both equations to form a new one that is only a function of $\vec{X}$. However, the complexity of solving such an equation would be high. Therefore, we use a simpler iterative approach to solve for $\vec{X}$. We start with an initial estimate for $\vec{X}$ and $\vec{\zeta}$. In each iteration, we improve the estimate of $\vec{X}$ using (14) and then improve the estimate of $\vec{\zeta}$ using (21). This procedure is summarized in Table I.

B. Method II: ICI Mitigation Utilizing Adjacent Symbols

It is possible to acquire channel slopes without using the redundancy of the cyclic prefix. This can be done by utilizing either the previous symbol or both adjacent symbols. A constant slope is assumed over the time duration of $T + (N/2) \times T_s$ for the former and T for the latter. Therefore, the former can handle lower Doppler values while adding no processing delay. On the other hand, the latter would have a better performance at the price of delay of reception of the next symbol. Since we are interested in ICI mitigation in high mobility environments, we utilize both adjacent symbols to acquire channel slopes. This is shown in Fig. 3. Pilots of the current symbol provide an estimate of the channel at the mid-point of the current symbol, $\hat{h}_k^{(\frac{N}{2}-1)}$. This estimate is stored in the system. Upon processing of the next symbol, an estimate of the channel at midpoint of the next symbol, $\hat{h}_k^{(\frac{N}{2}-1),\text{next}}$, becomes available. Estimate of the slopes in region 2 (see Fig. 3) can then be obtained as follows:

$$\hat{\alpha}_k^{r_2} = \frac{\hat{h}_k^{(\frac{N}{2}-1),next} - \hat{h}_k^{(\frac{N}{2}-1)}}{T} \quad 0 \le k \le G \tag{23}$$

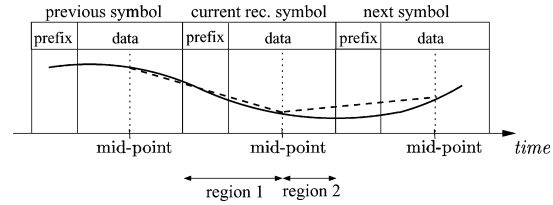

Fig. 3. Piece-wise linear model for method II. Solid curve: real or imag. part of a channel path. Dashed line: piece-wise linear model.

where $\alpha_k^{r_2}$ represents the slope of the kth channel tap in region 2. Similarly, $\alpha_k^{r_1}$, the slope in region 1, is estimated while processing the previous OFDM symbol and is stored in the receiver. Utilizing two slopes introduces a minor change in (11). It can be shown that in this case, we will have

$$\vec{y}_{\text{method}_{II}} \approx (\mathbf{h}_{\text{mid}} + \mathbf{M}^{r_1} \times \mathbf{A}^{r_1} + \mathbf{M}^{r_2} \times \mathbf{A}^{r_2}) \times \vec{x} + \vec{w}. \tag{24}$$

In (24), $\mathbf{A}^{r_m}$ represents channel slope matrix of (13) in the mth region ($1 \le m \le 2$) with $\mathbf{M}^{r_1}$ and $\mathbf{M}^{r_2}$ defined as follows:

$$\mathbf{M}^{r_1}(i,j) = \begin{cases} \mathbf{M}(i,j) & i=j \ \& \ 0 \le i \le \frac{N}{2}-1 \\ 0 & else \end{cases}$$
$$\mathbf{M}^{r_2}(i,j) = \begin{cases} \mathbf{M}(i,j) & i=j \ \& \ \frac{N}{2} \le i \le N-1. \\ 0 & else \end{cases} \tag{25}$$

Following the same procedure of Appendix C, it can be easily shown that the frequency domain relationship will be

$$\vec{Y}_{\text{method}_{II}} \approx \mathbf{H}_{\text{mid}}\vec{X} + \left(\mathbf{C}^{r_1} \times \mathbf{H}_{\text{slope}}^{r_1} + \mathbf{C}^{r_2} \times \mathbf{H}_{\text{slope}}^{r_2}\right) \times \vec{X} + \vec{W} \tag{26}$$

In (26), $\mathbf{H}_{\text{slope}}^{r_m}$ is the diagonal matrix defined in (16) for the slopes of the mth region and can be formed from $\hat{\alpha}_k^{r_m}$. $\mathbf{C}^{r_1}$ and $\mathbf{C}^{r_2}$ are fixed matrices. It can be easily shown that

$$\mathbf{C}^{r_1}(n,m) = T_s \times \begin{cases} \frac{-.5}{1-e^{\frac{-j2\pi(n-m)}{N}}} + \frac{1-(-1)^{n-m}}{N \times \left(1-e^{\frac{-j2\pi(n-m)}{N}}\right)^2} & n \ne m \\ \frac{1}{4} - \frac{N}{8} & n = m \end{cases}$$
$$\mathbf{C}^{r_2}(n,m) = T_s \times \begin{cases} \frac{-.5}{1-e^{\frac{-j2\pi(n-m)}{N}}} - \frac{1-(-1)^{n-m}}{N \times \left(1-e^{\frac{-j2\pi(n-m)}{N}}\right)^2} & n \ne m \\ \frac{1}{4} + \frac{N}{8} & n = m \end{cases}. \tag{27}$$

An estimate of X can then be obtained from (26).

C. Complexity Analysis

In general, solving (14) and (21) in case of Method I and (26) or (24) in case of Method II requires matrix inversion which could increase receiver complexity. For Method I, since the size of (21) is smaller, the main complexity is in solving (14). This requires an $N \times N$ matrix inversion. In general, any matrix inversion algorithm can be used. Also, (14) and (26) show a special structure. For instance, in (14) we need to invert a sum of $diagonal + toeplitz \times diagonal$. The special structure can be used to reduce the complexity in iterative methods. Comparing

Authorized licensed use limited to: Perkins Coie LLP. Downloaded on August 12,2020 at 15:25:59 UTC from IEEE Xplore. Restrictions apply.

with the method proposed in [13], Methods I and II can handle considerably higher delay and Doppler spread (see Section I) at the price of higher computational complexity (by neglecting some of the channel coefficients, the complexity of the method proposed in [13] is reduced to $N-q$ inversions of a matrix of size $q \times q$, where q is smaller than N). However, depending on the computational power of the receiver, other less complex methods like conjugate gradient can be used for matrix inversion. A good survey of such methods and their complexity analysis can be found in [16], [17]. Furthermore, Section VI shows how adding a noise/interference reduction mechanism can further reduce the complexity.

Another important issue is the convergence property of Method I. In general, for the range of Doppler values that the piece-wise linear approximation can be applied, channel slopes are small enough that the initial estimate of zero in the first step of Method I results in the convergence of the method after a few iterations. A more detailed analysis of convergence properties of such iterative methods is beyond the scope of this paper but can be found in [17], [18].

V. Mathematical Analysis of the Effect of Linearization

In this section, we provide a mathematical analysis of the effect of piece-wise linear approximation in mitigating ICI. We assume a narrowband time-variant channel to make the analysis tractable and leave the case of wideband channels to our simulations in Section VII. We define SIR_{ave} as the ratio of average signal power to the average interference power. Our goal is to calculate SIR_{ave} when ICI is mitigated and compare it to that of the "no mitigation" case. Consider a narrowband time-variant channel, $h^{(i)}$. Note that we drop the index k of $h_k^{(i)}$ in this section under narrowband channel assumption. Then, in the absence of noise, (3) can be simplified as follows:

$$y_i = h^{(i)} \times x_i \tag{28}$$

The estimate of x_i will be

$$\hat{x}_i = \frac{y_i}{\hat{h}^{(i)}} = x_i + \frac{e_i}{\hat{h}^{(i)}} \times x_i \tag{29}$$

In (29), $e_i = h^{(i)} - \hat{h}^{(i)}$. Since $\hat{h}^{(i)}$ is the sum of a considerable number of uncorrelated random variables as estimated from pilot tones, we approximate its distribution with a complex Gaussian. In practice if $\hat{h}^{(i)}$ is having near to zero values, the received signal will not be divided by it. From theoretical standpoint, if the cases of near to zero $\hat{h}^{(i)}$ are not excluded, the variance of $\hat{x}_i$, the estimate of x_i, will be infinite (this can be seen from the results of this section). Therefore, we need to exclude the probability of a near to zero $\hat{h}^{(i)}$ to make the analysis meaningful. This can be done by introducing a slight modification in the pdf of $|\hat{h}^{(i)}|^2$. Let $prob(|\hat{h}^{(i)}|^2 \leq \mu) = \epsilon$. Then, for an ϵ near zero, we take the pdf of $|\hat{h}^{(i)}|^2$ to be zero for $|\hat{h}^{(i)}|^2 \leq \mu$.

Taking an FFT of (29), we will have

$$\hat{X}_i = X_i + A_i \tag{30}$$

In (30), $a_i = \frac{e_i}{\hat{h}^{(i)}} \times x_i$ and A_i is the FFT of it. A_i is not purely interference and contains a term that depends on X_i as well. However, it can be shown that the power of that term is considerably small. Therefore, to reduce the complexity of the analysis we take A_i as the interference term which makes the analysis a tight approximation. SIR_{ave} can then be defined as follows:

$$SIR_{\text{ave}} = \frac{\sigma_X^2}{\sigma_A^2} \tag{31}$$

where σ_X^2 is the average power of X and σ_A^2 can be calculated as follows:

$$\sigma_A^2 = E\left(A_i A_i^*\right) = \sum_{n=0}^{N-1}\sum_{m=0}^{N-1} E\left(a_n a_m^*\right) e^{-\frac{j2\pi i(n-m)}{N}}. \tag{32}$$

Since $E(x_n x_m^*) = (\sigma_X^2/N)\delta_{n,m}$, then

$$\sigma_A^2 = \frac{\sigma_X^2}{N} \times \sum_{m=0}^{N-1} E\left(\left|\frac{e_m}{\hat{h}^{(m)}}\right|^2\right) \tag{33}$$

and we will have

$$SIR_{\text{ave}} = \frac{N}{\sum_{m=0}^{N-1} E\left(\frac{|e_m|^2}{|\hat{h}^{(m)}|^2}\right)}. \tag{34}$$

Both e_m and $\hat{h}^{(m)}$ have complex Gaussian distributions. Furthermore, they are jointly Gaussian (since a linear combination of them is the sum of a considerable number of uncorrelated random variables) and correlated. Then, $E\left(\left|\frac{e_m}{\hat{h}^{(m)}}\right|^2\right)$ will be as follows (derived in Appendix D):

$$E\left(\left|\frac{e_m}{\hat{h}^{(m)}}\right|^2\right) = \frac{\rho_m^2 \sigma_{e_m}^2}{\sigma_{\hat{h}^{(m)}}^2} + \sigma_{e_m^2}\left(1-\rho_m^2\right) E\left(\frac{1}{\left|\hat{h}^{(m)}\right|^2}\right). \tag{35}$$

In (35), $\sigma_{\hat{h}^{(m)}}^2 = E(|\hat{h}^{(m)}|^2)$, $\sigma_{e_m}^2 = E(|e_m|^2)$ and $\rho_m = E(e_m \hat{h}^{(m)*})/\sigma_{e_m}\sigma_{\hat{h}^{(m)}}$. These parameters are functions of channel correlation characteristics (or Doppler spectrum) and are derived in Appendix D. Also, it can be easily shown that $E(1/|\hat{h}^{(m)}|^2) = Ei(-ln(1-\epsilon))/\sigma_{\hat{h}^{(m)}}^2$ where Ei and ln stand for the exponential integral and logarithm in the base of e respectively. Inserting (35) in (33) will result in the following SIR_{ave}:

$$SIR_{\text{ave}} = \frac{N}{\sum_{m=0}^{N-1} \frac{\sigma_{e_m}^2 \times \rho_m^2 + Ei(-ln(1-\epsilon)) \times (1-\rho_m^2) \times \sigma_{e_m}^2}{\sigma_{\hat{h}^{(m)}}^2}}. \tag{36}$$

Fig. 4 shows SIR_{ave} of (36) as a function of $f_{d,\text{norm}}$ at $\epsilon = 10^{-6}$. $f_{d,\text{norm}}$ is defined as f_d (maximum Doppler) divided by the sub-carrier spacing. Channel power spectrum is Jakes spectrum [19] for this result. This means that function R of Appendix D is $R(t) = J_0(2\pi f_d t)$ with J_0 representing

Authorized licensed use limited to: Perkins Coie LLP. Downloaded on August 12,2020 at 15:25:59 UTC from IEEE Xplore. Restrictions apply.

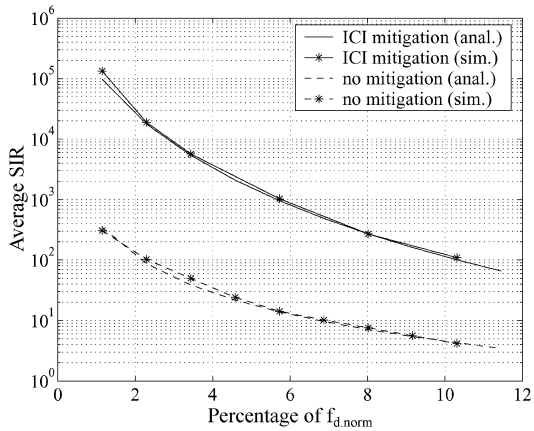

Fig. 4. Average SIR versus % of $f_{d,\mathrm{norm}}$ for a narrowband channel.

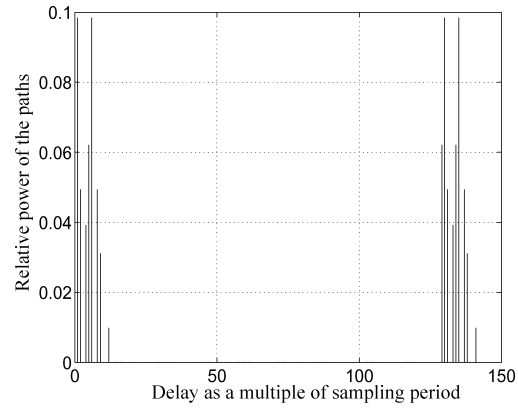

Fig. 5. Power-delay profile of channel#2.

zero-order Bessel function. This analytic result matches the corresponding simulation as can be seen from the graph. For comparison, SIR_{ave} for the case of "no mitigation" is plotted as well. Appendix D shows how $\sigma^2_{\hat{h}(m)}$, $\sigma^2_{e_m}$ and ρ_m can be found for the case of "no mitigation". The graph indicates how ICI mitigation through linearization improves SIR_{ave}.

VI. Noise/Interference Reduction

In Section III, an estimate of h_k^{ave} was acquired using equally-spaced pilots. In most cases, the number of active channel taps will be less than $\nu + 1$, especially in an SFN environment (see Fig. 5 as an example). In such cases, some of the estimated channel taps would be noise/interference samples after the IFFT in (8). Therefore, if these taps can be removed, the effect of noise/interference will be reduced. To do so, estimated channel taps are compared with a $Threshold$. If the value of a tap is below the $Threshold$, it will be zeroed:

$$if \quad \left|\hat{h}_k^{\mathrm{ave}}\right| \leq Threshold \quad => \quad \hat{h}_k^{\mathrm{ave}} = 0, \;\; 0 \leq k \leq L-1 \tag{37}$$

The optimum way to define the $Threshold$ is to relate it to the received signal-to-noise plus interference ratio (*SNIR*) such that the taps comparable to or below noise/interference level are zeroed. However, this requires estimation of the power of noise/interference which may not be feasible in a high mobility environment. Instead, we define a simple but effective $Threshold$. Upon estimation of $\hat{h}_k^{\mathrm{ave}}$ from (8), the tap with maximum absolute value is detected. Let MAV represent this maximum. Then, all the estimated taps with absolute values smaller than $\frac{MAV}{\gamma}$ for some $\gamma \geq 1$ will be zeroed. Choosing small γ increases the chance of losing channel taps with significant values and only improves the performance if the noise/interference level is high. On the other hand, choosing a high γ will reduce the risk of losing taps with considerable values at the price of less efficiency in high noise/interference cases. In general, in the absence of knowledge of *SNIR*, it is better to choose γ such that losing the taps below the $Threshold$ does not introduce considerable performance loss. Following the same criteria, we choose $\gamma = 10$ in our simulations in the next section. In the case of losing a tap, the power of such a tap is less than 1% of the power of the strongest channel tap. This will lead to a slight performance loss at very high *SNIR* which should not be a problem since these cases already have a very low error rate.

TABLE II
Parameters of the Simulated System

Parameter	*Value*
Input Modulation	8PSK
bit rate (excluding the redundancy)	7.3Mbps
# of sub-carriers (N)	892
# of pilots (L)	223
length of the guard interval (T_g)	$44.4\mu s$
length of one OFDM symbol (T)	$273.5\mu s$
$T_s = \frac{T}{N+L}$	$.26\mu s$
$G = \frac{T_g}{T_s}$	173

Furthermore, the number of nonzero channel taps can be estimated from $\hat{h}_k^{\mathrm{ave}}$ after (37) is applied. Let $N_p \leq L$ represent this estimate in the current OFDM symbol. Therefore, we only need to estimate N_p slopes. This will reduce the complexity of both algorithms. For instance it will reduce the number of unknowns from G to N_p in step 4 of Method I. This reduction can be considerable for SFN channels.

VII. Simulation Results

We simulate an OFDM system in a time-variant environment with high delay spread. System parameters[4] are summarized in Table II. We simulate two power-delay profiles. The power-delay profile of channel#1 has two main taps that are separated by 20 μs. Power-delay profile of channel#2 is shown in Fig. 5

[4]Parameters are based on Sirius Radio second-generation system specification proposal for an SFN environment.

Authorized licensed use limited to: Perkins Coie LLP. Downloaded on August 12,2020 at 15:25:59 UTC from IEEE Xplore. Restrictions apply.

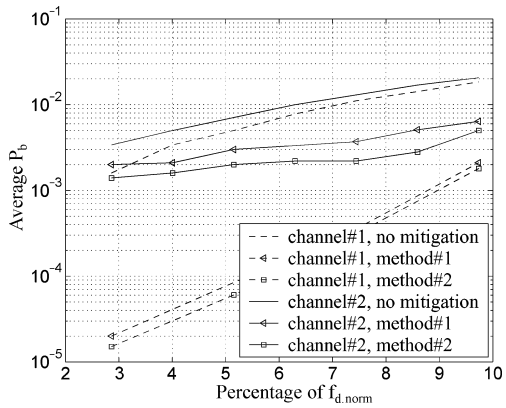

Fig. 6. Error floor versus maximum normalized Doppler.

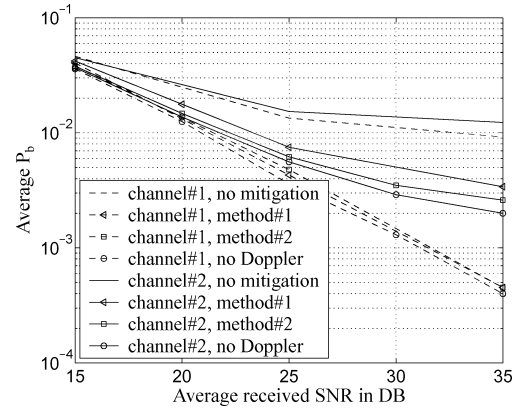

Fig. 7. Average bit-error rate versus average received SNR for $f_{d,\mathrm{norm}} = 6.5\%$.

and has two main clusters with the total delay of 36.5 μs to represent a case of reception from two adjacent base stations in an SFN environment. Each channel tap is generated as a random process with Rayleigh distributed amplitude and uniformly distributed phase using Jakes model [19]. Therefore, the auto-correlation of each tap is a zero-order Bessel function. For both channels, the power of channel taps is normalized to result in a total power of one. To see how ICI mitigation methods reduce the error floor, Fig. 6 shows the average bit-error rate, P_b (before decoding), in the absence of noise for both channels. In the "no mitigation" case, pilots are used to estimate h_k^{ave} which is then used to detect transmitted data without any estimation of time-variations. As can be seen from Fig. 6, average P_b increases considerably for the "no mitigation" case. Both of the proposed methods reduce the error floor considerably. Method II shows a slightly better performance than Method I. This is due to the iterative way of solving for unknowns in Method I. Also, channel#1 results in a lower error floor due to its shorter delay and smaller number of taps, as expected. To see the effect of noise, Fig. 7 shows average P_b (before decoding) as a function of average received SNR for $f_{d,\mathrm{norm}} = 6.5\%$. Average received SNR is defined as the ratio of the average total signal power received through all the channel paths to the average received noise power. Average error rate for the ideal case of no Doppler is also plotted for comparison. It can be seen from Fig. 7 that ICI mitigation reduces the error rate considerably for both channels. In particular for channel#1, the error rate is almost reduced to that of the case with no Doppler.

To see how ICI mitigation methods reduce the required received SNR for achieving a specific pre-decoding bit error rate, Fig. 8 shows the required received SNR for reaching an average $P_b = 0.02$ before decoding. The graph shows how ICI mitigation saves power. For comparison, the required received SNR for the case of no Doppler is 17.6 dB for both channels. It can be seen that both methods reduce the amount of required power to a level close to that of the no Doppler case. For instance, at $f_{d,\mathrm{norm}} = 8\%$, the amount of power saving is around 4 dB. Compared to the "no mitigation" case, the amount of power saving increases considerably as $f_{d,\mathrm{norm}}$ increases.

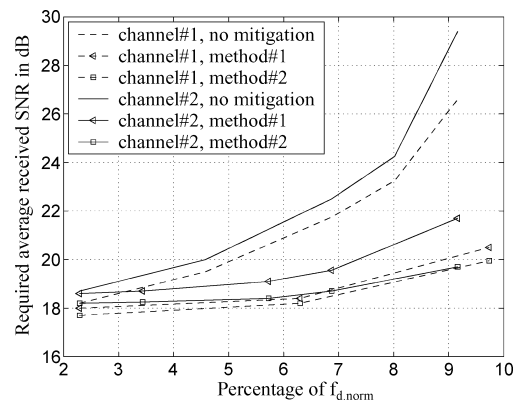

Fig. 8. Required average received SNR to achieve average $P_b = .02$.

VIII. Conclusion

In this paper, we proposed two new methods for ICI mitigation in pilot-aided OFDM mobile systems. Both methods used a piece-wise linear approximation to estimate channel time-variations in each OFDM symbol. Performance improvement was shown analytically by deriving SIR_{ave} formulas in a narrowband mobile environment. In high delay and Doppler spread environments, our simulation results showed considerable performance improvement. They illustrated that applying these methods would reduce average P_b or the required received SNR to a value close to that of the case with no Doppler. The power savings become considerable as $f_{d,\mathrm{norm}}$ increases.

Appendix A

Consider a wide-sense stationary process h and its piece-wise linear approximation h_{lin} shown in Fig. 9. h can represent any of the channel taps. $h_{\mathrm{lin}}^{(m)}$ represents h_{lin} at time instant $m \times T_s$. It is of interest to compare the auto-correlation function of h_{lin},

Authorized licensed use limited to: Perkins Coie LLP. Downloaded on August 12,2020 at 15:25:59 UTC from IEEE Xplore. Restrictions apply.

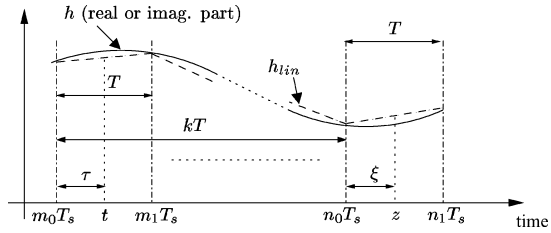

Fig. 9. Piece-wise linear approximation of a random process h.

$R_{\text{lin}}(t = mT_s, z = nT_s) = E(h_{\text{lin}}^{(m)} h_{\text{lin}}^{(n)*})$, to that of h. From Fig. 9, $h_{\text{lin}}^{(m)}$ and $h_{\text{lin}}^{(n)}$ can be expressed as follows:

$$h_{\text{lin}}^{(m)} = h_{\text{lin}}^{(m_0)} + \tau \times \frac{h_{\text{lin}}^{(m_1)} - h_{\text{lin}}^{(m_0)}}{T} \quad m_0 \le m \le m_1$$
$$h_{\text{lin}}^{(n)} = h_{\text{lin}}^{(n_0)} + \xi \times \frac{h_{\text{lin}}^{(n_1)} - h_{\text{lin}}^{(n_0)}}{T} \quad n_0 \le n \le n_1. \quad (38)$$

For simplicity, we assume that $h_{\text{lin}}^{(x)} = h^{(x)}$ for[5] $x = m_0, m_1, n_0$ and n_1. Then, $R_{\text{lin}}(t, z)$ will be

$$\begin{aligned} R_{\text{lin}}(t = mT_s, z = nT_s) &= \left(1 - \frac{\tau}{T}\right)\left(1 - \frac{\xi}{T}\right) \times E\left(h^{(m_0)} h^{(n_0)*}\right) \\ &+ \left(\frac{\xi}{T} - \frac{\tau\xi}{T^2}\right) \times E\left(h^{(m_0)} h^{(n_1)*}\right) \\ &+ \left(\frac{\tau}{T} - \frac{\tau\xi}{T^2}\right) \times E\left(h^{(m_1)} h^{(n_0)*}\right) \\ &+ \frac{\tau\xi}{T^2} \times E\left(h^{(m_1)} h^{(n_1)*}\right) \end{aligned} \quad (39)$$

Let $R(t = mT_s, z = nT_s) = R((m-n) \times T_s) = E(h^{(m)} h^{(n)*})$ represent the auto-correlation function of process h. Then, we will have

$$\begin{aligned} R_{\text{lin}}(t, z) &= R_{\text{lin}}(\tau, \xi, k) \\ &= \left(\frac{\tau}{T} - \frac{\tau\xi}{T^2}\right) \times R((k-1)T) \\ &+ \left(\frac{\xi}{T} - \frac{\tau\xi}{T^2}\right) \times R((k+1)T) \\ &+ \left(1 - \frac{\tau}{T} - \frac{\xi}{T} + 2\frac{\tau\xi}{T^2}\right) \times R(kT). \end{aligned} \quad (40)$$

Since h is a wide-sense stationary process, we will have $R(t, z) = R(t - z) = R(\tau - \xi - kT)$. Define P as the average power of the difference of $R(t, z)$ and $R_{\text{lin}}(t, z)$ over D OFDM symbols

$$P = \sum_{k=0}^{k=D-1} \int_0^T \int_0^T \left(R_{\text{lin}}(\tau, \xi, k) - R(\tau - \xi - kT)\right)^2 d\tau d\xi \quad (41)$$

[5]Due to the presence of noise/interference, $h_{\text{lin}}^{(x)}$ may differ from $h^{(x)}$ for $x = m_0, m_1, n_0$ and n_1. Extending the analysis to include this difference should be a straightforward extension of the work in this appendix.

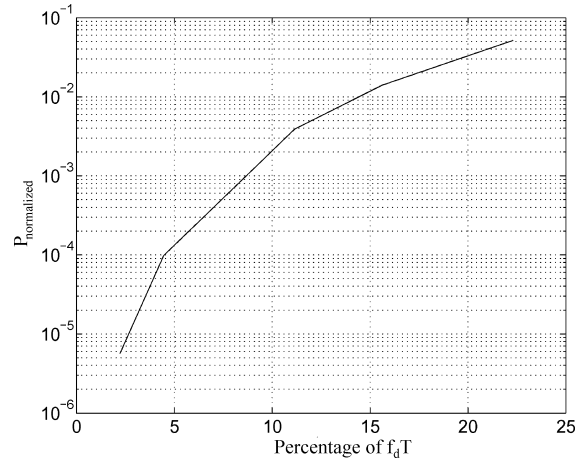

Fig. 10. $P_{\text{normalized}}$ versus % of $f_d \times T$.

We are interested in $P_{\text{normalized}}$ which can be defined as follows:

$$P_{\text{normalized}} = \frac{P}{\sum_{k=0}^{k=D-1} \int_0^T \int_0^T R^2(\tau - \xi - kT) d\tau d\xi} \quad (42)$$

For $R(t) = J_0(2\pi f_d t)$, we characterize $P_{\text{normalized}}$ of (42). Fig. 10 shows $P_{\text{normalized}}$ as a function of $f_d \times T$. We pick D large enough, i.e. $D \gg 1/(f_d \times T)$, so that $R(DT)$ becomes negligible. The graph suggests that for $f_d \times T$ of up to 20%, $P_{\text{normalized}}$ is negligible. For instance, for $f_d \times T = 15\%$, $P_{\text{normalized}}$ is 1%.

APPENDIX B

Call $\text{F(s)} = E(|h_k^{\text{ave}} - h_k^{(s)}|^2)$. The goal is to minimize F over $0 \le s \le N-1$. F can be written as follows:

$$\begin{aligned} F(s) &= \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{u'=0}^{N-1} E\left(h_k^{(u)} h_k^{*(u')}\right) + E\left(\left|h_k^{(s)}\right|^2\right) \\ &- \frac{1}{N} \sum_{u=0}^{N-1} E\left(h_k^{(u)} h_k^{*(s)}\right) \\ &- \frac{1}{N} \sum_{u=0}^{N-1} E\left(h_k^{*(u)} h_k^{(s)}\right) \end{aligned} \quad (43)$$

$$\begin{aligned} F(s) &= \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{u'=0}^{N-1} R_k(u - u') + R_k(0) \\ &- \frac{2}{N} \sum_{u=0}^{N-1} Re\left\{R_k(u - s)\right\}. \end{aligned} \quad (44)$$

Here, $Re\{\}$ stands for the real part of argument and $R_k(z) = E(h_k^{(u)} h_k^{*(u-z)})$. From (44), minimization of F is equivalent to maximization of $F' = \sum_{u=0}^{N-1} Re\{R_k(u - s)\}$ over s. Without loss of generality, assume $R_k(z) = \sigma_k^2 J_0(2\pi f_d z T_s)$ where J_0 represents zero-order Bessel function. Since $NT_s \le (1/f_d)$ (which means that the length of the symbol data part is less

Authorized licensed use limited to: Perkins Coie LLP. Downloaded on August 12,2020 at 15:25:59 UTC from IEEE Xplore. Restrictions apply.

than channel coherence time), it can be easily seen that F' is maximized for $s = (N/2) - 1$ and $s = N/2$ (N is assumed even). Therefore, we approximate $h_k^{(\frac{N}{2}-1)}$ with the estimate of h_k^{ave}.

APPENDIX C

Let $\vec{y}_1 = \mathbf{A}\vec{x}$. Since $\mathbf{A}$ is a circular toeplitz matrix, taking an FFT of $\vec{y}_1$ will result in a multiplication by a diagonal matrix in the frequency domain. Therefore, we will have $\vec{Y}_1 = \mathbf{H}_{\text{slope}}\vec{X}$ with Y_1 representing FFT of y_1 and $\mathbf{H}_{\text{slope}}$ as defined in (16). Let $\vec{y}_2 = \mathbf{M}\vec{y}_1$. Taking an FFT of it, we will have $Y_2(k) = \frac{B_k \otimes_N Y_1(k)}{N}$ where B is the FFT of β and $\otimes_N$ represents circular convolution in the base of N. Taking an FFT of β, it can be easily calculated that B is as defined in (17) [20]. Therefore, we will have

$$\vec{Y} \approx \mathbf{H}_{\text{mid}}\vec{X} + \vec{Y}_2 = \mathbf{H}_{\text{mid}}\vec{X} + \mathbf{C} \times \mathbf{H}_{\text{slope}}\vec{X} + \vec{W} \quad (45)$$

APPENDIX D

A. Proof of (35)

Let $e_m = a_m + jb_m$ and $\hat{h}^{(m)} = c_m + jd_m$ where a_m, b_m, c_m and d_m are zero mean Gaussian variables with independent inphase and quadrature parts. We will have

$$\begin{aligned} E\left(\left|\frac{e_m}{\hat{h}^{(m)}}\right|^2\right) &= E\left(\frac{a_m^2 + b_m^2}{c_m^2 + d_m^2}\right) \\ &= E\left(\frac{E\left(a_m^2 + b_m^2\right)|_{c_m,d_m}}{c_m^2 + d_m^2}\right) \\ &= E\left(\frac{E\left(a_m^2\right)|_{c_m} + E\left(b_m^2\right)|_{d_m}}{c_m^2 + d_m^2}\right). \quad (46) \end{aligned}$$

It can be easily shown [21] that $E(a_m^2)|_{c_m} = \sigma_{a_m}^2(1 - r_m^2) + (r_m^2\sigma_{a_m}^2/\sigma_{c_m}^2)c_m^2$ with $\sigma_{a_m}^2 = E(a_m^2) = .5\sigma_{e_m}^2$, $\sigma_{c_m}^2 = E(c_m^2) = .5\sigma_{\hat{h}^{(m)}}^2$ and r_m, the correlation coefficient of a_m and c_m, being equal to ρ_m (defined in Section V). $E(b_m^2)|_{d_m}$ can be similarly calculated. Inserting $E(a_m^2)|_{c_m}$ and $E(b_m^2)|_{d_m}$ in (46) will result in (35).

B. Finding $\sigma_{e_m}^2$, $\sigma_{\hat{h}^{(m)}}^2$ and ρ_m

1) The Case of ICI Mitigation: From (7), (8), in the absence of noise and for a narrowband channel, we will have $\hat{h}^{\text{ave}} = h^{\text{ave}} + \Delta_{h^{\text{ave}}}$, where $\Delta_{h^{\text{ave}}}$ is the estimation noise with variance of $\sigma_{\Delta_{h^{\text{ave}}}}^2$. It can be easily shown, using (7) and (8), that $\sigma_{\Delta_{h^{\text{ave}}}}^2 = \frac{N^2R(0) - NR(0) - 2\sum_{i=1}^{N-1}(N-i)R(iT_s)}{N^2L}$. R is the auto-correlation function of the narrowband channel as defined in Appendix A. Furthermore, it can be shown that $E(\Delta_{h^{\text{ave}}} \times h^{(m)*}) = 0$ due to the independency of the transmitted data points and channel. Similarly, in the next OFDM symbol, $\hat{h}^{\text{ave,next}} = h^{\text{ave,next}} + \Delta_{h^{\text{ave,next}}}$, where $\Delta_{h^{\text{ave,next}}}$ is the estimation noise with $\sigma_{\Delta_{h^{\text{ave,next}}}}^2 = \sigma_{\Delta_{h^{\text{ave}}}}^2$. Define $\lambda_m = (m - .5N + 1)(T_s/T)$ and $\eta_m = 1 - \lambda_m$. Using Method II, we will have

$$\hat{h}^{(m)} = \eta_m\hat{h}^{\text{ave}} + \lambda_m\hat{h}^{\text{ave,next}} \quad m \in Region2. \quad (47)$$

A similar formula can be written for estimation in Region 1. Using these formulas, after some lengthy but straightforward computations, the following formulas can be derived: (see (48) located at bottom of page).

2) The Case of No Mitigation: In this case, we will have $\hat{h}^{(m)} = \hat{h}^{\text{ave}}$ for $0 \le m \le N-1$. Then, the parameters can be easily derived as follows:

$$\begin{aligned} \sigma_{\hat{h}^{(m)}}^2 &= \frac{NR(0) + 2\sum_{i=1}^{N-1} R(iT_s)}{N^2} + \sigma_{\Delta_{h^{\text{ave}}}}^2 \\ \sigma_{e_m}^2 &= R(0) + \sigma_{\hat{h}^{(m)}}^2 - \frac{2}{N}\sum_{k=0}^{N-1} R\left((m-k)T_s\right) \\ \rho_m &= \frac{\frac{1}{N}\sum_{k=0}^{N-1} R\left((m-k)T_s\right) - \sigma_{\hat{h}^{(m)}}^2}{\sigma_{\hat{h}^{(m)}}\sigma_{e_m}}. \end{aligned} \quad (49)$$

$$\begin{aligned} \sigma_{\hat{h}^{(m)}}^2 &= \left(\eta_m^2 + \lambda_m^2\right) \times \left(\frac{NR(0) + 2\sum_{i=1}^{N-1} R(iT_s)}{N^2} + \sigma_{\Delta_{h^{\text{ave}}}}^2\right) \\ &\quad + 2\frac{\eta_m\lambda_m}{N^2}\sum_{i=\frac{T}{T_s}-N+1}^{\frac{T}{T_s}+N-1}\left(N - \left|i - \frac{T}{T_s}\right|\right) \times R(iT_s) \\ \sigma_{e_m}^2 &= R(0) + \sigma_{\hat{h}^{(m)}}^2 - \frac{2}{N} \times \left(\eta_m\sum_{k=0}^{N-1} R\left((m-k)T_s\right) + \lambda_m\sum_{k=\frac{T}{T_s}}^{\frac{T}{T_s}+N-1} R\left((m-k)T_s\right)\right) \\ \rho_m &= \frac{\frac{1}{N} \times \left(\eta_m\sum_{k=0}^{N-1} R\left((m-k)T_s\right) + \lambda_m\sum_{k=\frac{T}{T_s}}^{\frac{T}{T_s}+N-1} R\left((m-k)T_s\right)\right) - \sigma_{\hat{h}^{(m)}}^2}{\sigma_{\hat{h}^{(m)}}\sigma_{e_m}} \end{aligned} \quad (48)$$

Authorized licensed use limited to: Perkins Coie LLP. Downloaded on August 12,2020 at 15:25:59 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. COMM-33, pp. 665–675, July 1985.

[2] Weinstein and Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun.*, vol. COMM-19, pp. 628–634, Oct. 1971.

[3] A. R. S. Bahai and B. R. Saltzberg, *Multi-Carrier Digital Communications—Theory and Applications of OFDM*. Norwell, MA: Kluwer, 1999.

[4] R. D. J. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Boston, MA: Artech House, 1998.

[5] B. Le Floch, R. Hallbert-Lasalle, and D. Castellain, "Digital audio broadcasting to mobile receivers," *IEEE Trans. Consum. Electron.*, vol. 35, pp. 493–503, Aug. 1989.

[6] H. Sari, G. Karam, and J. Janclaude, "Transmission techniques for digital TV broadcasting," *IEEE Commun. Mag.*, vol. 36, pp. 100–109, Feb. 1995.

[7] Q. Sun and D. C. Cox, "Fundamental limits on symbol rate in frequency selective continuous fading channels," in *Proc. IEEE Vehicular Technology Conf.*, vol. 2, Fall 2001, pp. 1205–1209.

[8] M. Russell and G. Stuber, "Interchannel interference analysis of OFDM in a mobile environment," in *Proc. IEEE Vehicular Technology Conf.*, 1995, pp. 820–824.

[9] P. Robertson and S. Kaiser, "The effects of Doppler spreads in OFDM(A) mobile radio systems," in *Proc. IEEE Vehicular Technology Conf.*, Fall 1999, pp. 329–333.

[10] Y. Li and L. Cimini, "Bounds on the inter-channel interference of OFDM in time-varying impairments," *IEEE Trans. Commun.*, vol. 49, pp. 401–404, Mar. 2001.

[11] S. Sung and D. Brady, "Spectral spatial equalization for OFDM in time-varying frequency-selective multipath channels," in *Proc. IEEE Workshop Sensor Array Multichannel Signal Processing*, 2000, pp. 434–438.

[12] A. Hutter, E. Carvalho, and J. Cioffi, "On the impact of channel estimation for multiple antenna diversity reception in mobile OFDM systems," in *Proc. 34th Asilomar*, vol. 2, 2000, pp. 1820–1824.

[13] W. G. Jeon and K. H. Chang, "An equalization technique for orthogonal frequency-division multiplexing systems in time-variant multipath channels," *IEEE Trans. Commun.*, vol. 47, pp. 27–32, Jan. 1999.

[14] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consumer Electron.*, vol. 44, pp. 1122–1128, Aug. 1998.

[15] Y. Mostofi, D. Cox, and A. Bahai, "Effect of frame synchronization errors on pilot-aided channel estimation in OFDM: analysis and solution," in *Proc. 5th Int. Symp. Wireless Personal Multimedia Communications*, Oct. 2002, pp. 1309–1313.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[17] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*. Philadelphia, PA: Frontiers in Applied Mathematics, 1995.

[18] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*. San Diego, CA: Academic, 1970.

[19] W. Jakes, *Microwave Mobile Communications*. Piscataway, NJ: IEEE Press, 1993.

[20] A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[21] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd ed. New York: McGraw-Hill International, 1991.

Yasamin Mostofi (S'98–M'04) received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 1997, and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 1999 and 2003, respectively.

She has worked on several aspects of wireless communications for Lucent Technologies–Bell Labs, Murray Hills, NJ, in 1999 and National Semiconductor, Fremont, CA, in 2002. Since January 2004, she has been a Postdoctoral Scholar with the California Institute of Technology, Pasadena. Her research interests include mobile communications, sensor networks, cross-layer designs, and advanced signal processing.

Donald C. Cox (M'61–SM'72–F'79) received the B.S. and M.S. degrees in electrical engineering from the University of Nebraska, Lincoln, in 1959 and 1960, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1968. He received an Honorary Dr.Sci. degree from the University of Nebraska in 1983.

From 1960 to 1963, he was with Wright-Patterson AFB, OH, where he reasearched microwave communications system design. From 1963 to 1968, he was with Stanford University researching tunnel diode amplifier design and research on microwave propagation in the troposphere. From 1968 to 1973, his research at Bell Laboratories, Holmdel, NJ, in mobile radio propagation and on high-capacity mobile radio systems provided important input to early cellular mobile radio system development, and is continuing to contribute to the evolution of digital cellular radio, wireless personal communications systems, and cordless telephones. From 1973 to 1983, he was Supervisor of a group at Bell Laboratories that did innovative propagation and system research for millimeter-wave satellite communications. In 1978, he pioneered radio system and propagation research for low-power wireless personal communications systems. At Bell Laboratories in 1983, he organized and became Head of the Radio and Satellite Systems Research Department that became a Division in Bell Communications Research (Bellcore) with the breakup of the Bell System on January 1, 1984. He was Division Manager of the Radio Research Division until it again became a department in 1991. He continued as Executive Director of the Radio Research Department where he championed, led, and contributed to research on all aspects of low-power wireless personal communications entitled Universal Digital Portable Communications (UDPC). He was instrumental in evolving the extensive research results into specifications that became the U.S. Standard for the Wireless or Personal Access Communications System (WACS or PACS). In September 1993, he became a Professor of Electrical Engineering and Director of the Center for Telecommunications, Stanford University, where he continues to pursue research and teaching of wireless mobile and personal communications. He was appointed Harald Trap Friis Professor of Engineering at Stanford in 1994. He is author or coauthor of many papers and conference presentations, including many invited and several keynote addresses, and books. He has been granted 15 patents.

Dr. Cox was a member of the Administrative Committee of the IEEE Antennas and Propagation Society (1986–1988), and was an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION (1983–1986). He received the 1983 IEEE Vehicular Technology Society Paper of the Year Award, the IEEE 1985 Morris E. Leeds Award, the 1990 Communications Magazine Prize Paper Award, the IEEE Communications Society 1992 L. G. Abraham Prize Paper Award, the IEEE 1993 Alexander Graham Bell Medal "for pioneering and leadership in personal portable communications," and the IEEE Third Millennium Medal in 2000. He was also a corecipient of the 1983 International Marconi Prize in Electromagnetic Wave Propagation (Italy), and received the Bellcore Fellow Award in 1991. He is a Fellow of AAAS and the Radio Club of America and a member of the National Academy of Engineering, Commissions B, C and F of USNC/URSI, Sigma Xi, Sigma Tau, Eta Kappa Nu, and Phi Mu Epsilon. He was a member of the URSI Intercommission Group on Time Domain Waveform Measurements from 1982 to 1984. He is a Registered Professional Engineer in Ohio and Nebraska.

Authorized licensed use limited to: Perkins Coie LLP. Downloaded on August 12,2020 at 15:25:59 UTC from IEEE Xplore. Restrictions apply.