ON DESIGNING TIME-MULTIPLEXED PILOTS FOR DOUBLY-SELECTIVE CHANNEL ESTIMATION USING DISCRETE PROLATE SPHEROIDAL BASIS EXPANSION MODELS

Liying Song and *Jitendra K. Tugnait*

Department of Electrical & Computer Engineering, Auburn University, Auburn, AL 36849, USA

Email: songliy@auburn.edu, tugnajk@eng.auburn.edu

ABSTRACT

Channel estimation for single-input single-output frequency- and time-selective channels is considered using time- multiplexed training. The time-varying channel is assumed to be well-described by a basis expansion model using discrete prolate spheroidal sequences as the bases (DPS-BEM). First, the popular linear least-squares approach is exploited to estimate the basis expansion coefficients. Then the issue of training power allocation is addressed. Finally, computer simulation examples are presented where the channel is generated via Jakes' model.

***Index Terms*—** Doubly-selective channels, channel estimation, basis expansion models, training design

1. INTRODUCTION

Consider a doubly-selective single-input single-output (SISO) finite impulse response (FIR) linear channel with symbol-rate impulse response $h(n;l)$ (channel response at time n to a unit input at time $n-l$). In a basis expansion representation over a time-block $n = \{0, 1, \cdots, N-1\}$, it is assumed that [2]

$$h(n;l) = \sum_{q=0}^{Q} w_q(l)\psi_q(n) \tag{1}$$

where $\psi_q(n)$ is the q-th basis function and the basis expansion parameter $w_q(l)$ is fixed over the data block. With $s(n)$ denoting the transmitted symbol sequence, the received sequence in the presence of additive noise $\eta(n)$ is given by

$$x(n) = \sum_{q=0}^{Q} \psi_q(n) \left[\sum_{l=0}^{L} w_q(l)s(n-l)\right] + \eta(n). \tag{2}$$

Let T_s denote the symbol interval. For a channel with a multipath delay spread of τ_d sec and a Doppler spread of f_d Hz, in the complex exponential basis expansion model (CE-BEM) [7, 3] one takes $\psi_q(n) := e^{j\omega_q n}$, $\omega_q := 2\pi[q - Q/2]/N$, $L := \lfloor \tau_d/T_s \rfloor$ and $Q := 2\lceil \nu_{\mathrm{Dmax}} N\rceil$ where $\nu_{\mathrm{Dmax}} := f_d T_s$ is the maximum normalized Doppler bandwidth. In the discrete prolate spheroidal sequence-based BEM (DPS-BEM), the DPS vectors $\boldsymbol{\psi}_q \in \mathbb{R}^N$ (called Slepian sequences in [10], which are time-windowed DPS sequences) with elements $\psi_q(n)$ for $n \in \{0, ..., N-1\}$, are eigenvectors of the matrix $\boldsymbol{C} \in \mathbb{R}^{N\times N}$, fulfilling $\boldsymbol{C}\boldsymbol{\psi}_q = \lambda_q \boldsymbol{\psi}_q$ where λ_q are eigenvalues of matrix $\boldsymbol{C}$ with its (y, z)th entry given by $\frac{\sin[2\pi(y-z)\nu_{\mathrm{Dmax}}]}{\pi(y-z)}$ [8].

The Fourier basis expansion (i.e. CE-BEM) has the major drawback that the rectangular window associated with the truncated discrete Fourier transform (DFT) introduces spectral leakage [6], which results in a floor in the bit error rate (BER) in CE-BEM-based approaches [10]. DPS/Slepian sequences are a good alternative as a basis set to approximate bandlimited channels alleviating the spectral leakage of CE-BEM [10]. In this case one takes $Q = \lceil 2\nu_{\mathrm{Dmax}} N\rceil$ [10].

To acquire the channel state information (CSI) at the receiver, training symbols are usually periodically inserted during transmission, which is known as pilot symbol aided modulation (PSAM) [1]. Optimization of PSAM for CE-BEM based doubly-selective channel models has been considered in [4] where training sequence is designed to minimize the channel estimation MSE and furthermore, an estimated channel-based average capacity lower bound is maximized to select certain training parameters such as training power allocation. No such considerations are to be found in [10] where it is shown that DPS-BEM-based approaches significantly outperform CE-BEM-based approaches for doubly-selective channel estimation and data detection. The main objective of this paper is to consider certain aspects of PSAM parameter design for DPS-BEM, following the CE-BEM results of [4].

In this paper, we consider channel estimation for doubly-selective SISO channel described by DPS-BEM. A linear least-squares (LS) estimator is presented. Using the developed channel estimation variance expression, we cast the power allocation problem as one of optimizing a signal-to-noise ratio. Relationship to existing results is discussed in some detail in Sec. 5.

Notation: Superscripts H and T denote the complex conjugate transpose and transpose operations, respectively. $\delta(\cdot)$ is the Kronecker delta function and $\boldsymbol{I}_N$ is the $N \times N$ identity matrix. The symbol $\otimes$ denotes the Kronecker product.

2. LEAST-SQUARES CHANNEL ESTIMATION

2.1. Model Development

We consider block transmission as in [4], where transmitted symbols are collected into $N\times 1$ blocks with $\boldsymbol{s} = [s(0), s(1), \cdots, s(N-1)]^T$ as the 0th block and received $x(n)$'s are also collected into blocks with $\boldsymbol{x} = [x(0), x(1), \cdots, x(N-1)]^T$ as 0th block. To avoid inter-block interference (IBI), as in [4], L guard zeros are inserted in $\boldsymbol{s}$ (per block) at the transmitter. Then (2) is rewritten as

$$\boldsymbol{x} = \sum_{q=0}^{Q} \boldsymbol{D}_{\psi_q} \boldsymbol{W}_q \boldsymbol{s} + \boldsymbol{\eta}, \tag{3}$$

where $\boldsymbol{\eta}$ is defined similarly to $\boldsymbol{x}$, $\boldsymbol{D}_{\psi_q} = \mathrm{diag}[\boldsymbol{\psi}_q]$ with $\boldsymbol{\psi}_q := [\psi_q(0), \psi_q(1), ..., \psi_q(N-1)]$, and $\boldsymbol{W}_q$s are $N \times N$ lower triangular Toeplitz matrices with 1st column $[w_q(0), w_q(1), \cdots, w_q(L), 0, \cdots, 0]^T$. Since $\psi_q(n)$ are known at receiver, the objective of channel estimator is to find basis expansion parameters in Eq. (1) from

This work was supported by NSF under Grant ECS-0424145.

the received samples corresponding to the training symbols. The proposed channel estimation relies on time multiplexing training symbols at known positions.

As in [4], each transmitted block $\boldsymbol{s}$ consists of J segments (subblocks) of training and information symbols $b(n)$ and $c(n)$, respectively, and each segment has the same length. Then the general structure of $\boldsymbol{s}$ is

$$\boldsymbol{s} = [\boldsymbol{b}_1^T, \boldsymbol{c}_1^T, ..., \boldsymbol{b}_J^T, \boldsymbol{c}_J^T]^T, \tag{4}$$

where $\boldsymbol{b}_j$ with length N_b and $\boldsymbol{c}_j$ with length $N_c, \forall j \in [1, J]$, denote training and information symbol subblocks, respectively. Therefore, $N = J(N_b + N_c)$ with $N_b > L$. Let $M = N_b + N_c$ denote the subblock size. Obviously, the first L symbols in the "training part" of the j-th subblock of the received signal are contaminated by information symbols in the preceding $(j-1)$th subblock. In the similar way, the first L symbols in the "information part" of the j-th subblock of the received signal are also contaminated by the last L training symbols in the current j-th subblock. In order to avoid the inter-subblock interference (ISBI) so that channel estimation is decoupled from data detection, we will choose the first and the last L symbols in each training subblock to be zeros, as in [4].

Define $\boldsymbol{b}_j := [b((j-1)M), b((j-1)M+1), ..., b((j-1)M + N_b - 1)]^T$. Further define $\bar{\boldsymbol{D}}_{\psi_{q,j}} = \text{diag}\{\bar{\boldsymbol{\psi}}_{q,j}\}$ where $\bar{\boldsymbol{\psi}}_{q,j} = [\psi_q((j-1)M + L), \psi_q((j-1)M + L + 1), ..., \psi_q((j-1)M + N_b - 1)]^T$. Then the ISBI free received subblock can be written as

$$\bar{\boldsymbol{x}}_{b,j} = \sum_{q=0}^{Q} \bar{\boldsymbol{D}}_{\psi_{q,j}} \bar{\boldsymbol{W}}_q \boldsymbol{b}_j + \bar{\boldsymbol{\eta}}_{b,j}, \tag{5}$$

where $\bar{\boldsymbol{x}}_{b,j} := [x_b((j-1)M+L), x_b((j-1)M+L+1), ..., x_b((j-1)M + N_b - 1)]^T$, $\bar{\boldsymbol{\eta}}_{b,j}$ is defined similarly, and $(N_b - L) \times N_b$ matrix $\bar{\boldsymbol{W}}_q$ is

$$\bar{\boldsymbol{W}}_q = \begin{bmatrix} w_q(L) & \dots & w_q(0) & & \\ & \ddots & \vdots & \ddots & \\ & & w_q(L) & \dots & w_q(0) \end{bmatrix}.$$

Gathering training symbols per block, we obtain

$$\bar{\boldsymbol{x}}_b = \sum_{q=0}^{Q} \begin{bmatrix} \bar{\boldsymbol{D}}_{\psi_{q,1}} \bar{\boldsymbol{W}}_q \boldsymbol{b}_1 \\ \vdots \\ \bar{\boldsymbol{D}}_{\psi_{q,J}} \bar{\boldsymbol{W}}_q \boldsymbol{b}_J \end{bmatrix} + \bar{\boldsymbol{\eta}}_b. \tag{6}$$

According to the commutativity property of convolution, we have $\bar{\boldsymbol{W}}_q \boldsymbol{b}_j = \boldsymbol{B}_j \boldsymbol{w}_q$ with $\boldsymbol{w}_q := [w_q(0), ..., w_q(L)]^T$ and $\boldsymbol{B}_j$ a $(N_b - L) \times (L+1)$ Toeplitz matrix given by

$$\boldsymbol{B}_j := \begin{bmatrix} b_j(L) & \dots & b_j(0) \\ \vdots & \ddots & \vdots \\ b_j(N_b - 1) & \dots & b_j(N_b - L - 1) \end{bmatrix}, \tag{7}$$

where $b_j(l) := b((j-1)M + l)$. Therefore, (6) can be rewritten as

$$\bar{\boldsymbol{x}}_b = \boldsymbol{\Phi} \boldsymbol{w} + \bar{\boldsymbol{\eta}}_b, \tag{8}$$

with simple substitutions, where the $[J(N_b - L)] \times [(Q+1)(L+1)]$ matrix

$$\boldsymbol{\Phi} := \begin{bmatrix} \bar{\boldsymbol{D}}_{\psi_{0,1}} \boldsymbol{B}_1 & \dots & \bar{\boldsymbol{D}}_{\psi_{Q,1}} \boldsymbol{B}_1 \\ \vdots & \ddots & \vdots \\ \bar{\boldsymbol{D}}_{\psi_{0,J}} \boldsymbol{B}_J & \dots & \bar{\boldsymbol{D}}_{\psi_{Q,J}} \boldsymbol{B}_J \end{bmatrix}, \tag{9}$$

$$\boldsymbol{w} := [\boldsymbol{w}_0^T, \boldsymbol{w}_1^T, \dots, \boldsymbol{w}_Q^T]^T. \tag{10}$$

2.2. Least-Squares Channel Estimation

The linear least-squares (LS) channel estimator based on (8) is

$$\hat{\boldsymbol{w}} = \boldsymbol{\Lambda} \bar{\boldsymbol{x}}_b, \tag{11}$$

where $\boldsymbol{\Lambda} = (\boldsymbol{\Phi}^H \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^H$. Define the channel estimation error as $\tilde{\boldsymbol{w}} := \boldsymbol{w} - \hat{\boldsymbol{w}}$. Then the covariance matrix of $\tilde{\boldsymbol{w}}$ is ($\sigma_\eta^2 = E\{|\eta(n)|^2\}$) $\boldsymbol{R}_{\tilde{w}} := E[\tilde{\boldsymbol{w}} \tilde{\boldsymbol{w}}^H] = \sigma_\eta^2 \left(\boldsymbol{\Phi}^H \boldsymbol{\Phi}\right)^{-1}$. Using [5, Lemma 1], the MSE of $\tilde{\boldsymbol{w}}$ is lower-bounded as ($S := (Q+1)(L+1)$)

$$\sigma_{\tilde{w}}^2 := \text{tr}(\boldsymbol{R}_{\tilde{w}}) \geq \sigma_\eta^2 \sum_{i=1}^{S} \frac{1}{[\boldsymbol{\Phi}^H \boldsymbol{\Phi}]_{i,i}}, \tag{12}$$

with equality iff $\boldsymbol{\Phi}^H \boldsymbol{\Phi}$ is a diagonal matrix. By the arithmetic-geometric mean inequality,

$$\sum_{i=1}^{S} \frac{1}{[\boldsymbol{\Phi}^H \boldsymbol{\Phi}]_{i,i}} \geq S \left(\prod_{i=1}^{S} \frac{1}{[\boldsymbol{\Phi}^H \boldsymbol{\Phi}]_{i,i}} \right)^{1/S} \tag{13}$$

where the equality holds iff $\left[\boldsymbol{\Phi}^H \boldsymbol{\Phi}\right]_{i,i}$ are all equal. Equivalently, we need $\boldsymbol{\Phi}^H \boldsymbol{\Phi}$ to be a diagonal matrix with all its diagonal components equal. Suppose that we pick $\boldsymbol{\Phi}^H \boldsymbol{\Phi} = \alpha \boldsymbol{I}_{(Q+1)(L+1)}$ for some $\alpha > 0$. Then by (9) we must have $\sum_{j=1}^{J} \boldsymbol{B}_j^H \bar{\boldsymbol{D}}_{\psi_{q_1,j}}^H \bar{\boldsymbol{D}}_{\psi_{q_2,j}} \boldsymbol{B}_j = \alpha \boldsymbol{I} \delta(q_1 - q_2)$. It turns out that

$$\sum_{j=1}^{J} \bar{\boldsymbol{D}}_{\psi_{q_1,j}}^H \bar{\boldsymbol{D}}_{\psi_{q_2,j}} \approx M^{-1} \boldsymbol{I}_{N_b - L} \delta(q_1 - q_2) \tag{14}$$

whereas in case of CE-BEM the corresponding result is exactly true. To justify (14), consider [9, Sec. 5.3.1] where it is "shown" (heuristically) that for large N, the solutions $\boldsymbol{\psi}_q$ to $\boldsymbol{C}\boldsymbol{\psi}_q = \lambda_q \boldsymbol{\psi}_q$ (see Sec. 1 for DPS vectors $\boldsymbol{\psi}_q$) can be approximated by $\tilde{\boldsymbol{\psi}}_q = [1, e^{-2\pi(q-Q/2)/N}, e^{-4\pi(q-Q/2)/N}, \cdots, e^{-2\pi(q-Q/2)(N-1)/N}]^T$ with $\lambda_q = 1$ for $0 \leq q \leq Q = \lceil 2\nu_{\text{Dmax}} N \rceil$ and $\lambda_q = 0$ for $q \geq Q+1$. That is, we have CE-BEM for which (14) holds exactly. [To obtain real-valued DPS vectors one can create sine and cosine functions from conjugate pairs of $\tilde{\boldsymbol{\psi}}_q$.] Alternatively, for "smaller" record lengths, one can numerically calculate (14). We have done so and found that for the parameters used for simulations in this paper, the (normalized average) error norm in (14) is less than 1% for $q_1 = q_2$ and less than 10% for $q_1 \neq q_2$ as ν_{Dmax} varies from 0.001 to 0.01, where we average over all admissible values of q, $0 \leq q \leq Q$, and normalize with the norm of $M^{-1}\boldsymbol{I}_{N_b - L}$.

Under (14), following [4], we pick $N_b = 2L+1$ with $\boldsymbol{b}_j^T = [\boldsymbol{0}_L^T, b, \boldsymbol{0}_L^T]^T$ where $\boldsymbol{0}_L$ is a size L null column, in which case $\boldsymbol{\Phi}^H \boldsymbol{\Phi} = \frac{b^2}{M} \boldsymbol{I}_{(Q+1)(L+1)}$. Note that $b^2/M = Jb^2/N$. Under this optimal choice, the lower bound in (12) is achieved, yielding

$$\sigma_{\tilde{w}}^2 = \frac{N\sigma_\eta^2}{\mathcal{P}_b}(L+1)(Q+1), \quad \mathcal{P}_b := Jb^2, \tag{15}$$

where $\mathcal{P}_b$ denotes total training power in the given data block.

3. TRAINING POWER ALLOCATION

We assume that the time-varying channel $h(n;l)$ is zero-mean, WSS complex Gaussian with the same variance σ_h^2 for each tap. We also assume that the channel taps are mutually independent, i.e. $h(n;l)$ is

WSSUS. The received information symbols at the received antenna can be expressed as

$$x_c(n) = \underbrace{\sum_{l=0}^{L} \hat{h}(n;l)c(n-l)}_{:=x_s(n)}$$

$$+ \underbrace{\sum_{l=0}^{L}[h(n;l) - \hat{h}(n;l)]c(n-l) + \eta(n)}_{:=x_\eta(n)} \tag{16}$$

where $\hat{h}(n;l) = \sum_{q=0}^{Q} \hat{w}_q(l)\psi_q(n)$ is used for data detection. Therefore, the signal power is given by

$$\sigma_{xs}^2(n) := E\left\{|x_s(n)|^2\right\} =$$

$$\bar{\mathcal{P}}_c \sum_{l=0}^{L} \left[E_{\boldsymbol{w}} \left\{ E\left\{ \left| h(n;l) - \hat{h}(n;l) \right|^2 |\boldsymbol{w} \right\} \right\} + E\left\{|h(n;l)|^2\right\} \right]$$

$$= \bar{\mathcal{P}}_c \left[\sigma_{\tilde{h}}^2(n) + (L+1)\sigma_h^2\right], \ \ \bar{\mathcal{P}}_c := E\{|c(n)|^2\}, \tag{17}$$

and the effective noise power is

$$\sigma_{x\eta}^2(n) = E\left\{ \left| \sum_{l=0}^{L}[h(n;l) - \hat{h}(n;l)]c(n-l) \right|^2 \right\}$$

$$+E\left\{|\eta(n)|^2\right\} = \bar{\mathcal{P}}_c \sigma_{\tilde{h}}^2(n) + \sigma_\eta^2 \tag{18}$$

where $\sigma_{\tilde{h}}^2(n) := \sum_{l=0}^{L} E_{\boldsymbol{w}} \left\{ E\{|h(n;l) - \hat{h}(n;l)|^2 |\boldsymbol{w}\} \right\}$ and $\bar{\mathcal{P}}_c$ is the average power of information symbols. Define

$$\mathcal{W}_l := [w_0(l), w_1(l), ..., w_Q(l)]^T,$$

$$\mathcal{W} := [\mathcal{W}_0, \mathcal{W}_1, ..., \mathcal{W}_L]^T. \tag{19}$$

By (10) and (19), we have

$$\mathcal{W} = \Omega \boldsymbol{w}, \tag{20}$$

for some permutation matrix Ω (satisfying $\Omega^H\Omega = \mathbf{I}$). We may rewrite $\sigma_{\tilde{h}}^2(n)$ as

$$\sigma_{\tilde{h}}^2(n) = \sum_{l=0}^{L} E_{\mathcal{W}} \left\{ E\left\{ \left| h(n;l) - \hat{h}(n;l) \right|^2 |\mathcal{W} \right\} \right\}$$

$$= \operatorname{tr}\left\{ \Psi(n) E_{\mathcal{W}} \left\{ \operatorname{cov}\{\hat{\mathcal{W}}, \hat{\mathcal{W}}|\mathcal{W}\} \right\} \Psi(n)^H \right\}, \tag{21}$$

where $\Xi(n) := [\psi_0(n), \psi_1(n), ..., \psi_Q(n)]$ and $\Psi(n) := \boldsymbol{I}_{(L+1)} \otimes \Xi(n)$. Based on (19) and (20), we have

$$E_{\mathcal{W}} \left\{ \operatorname{cov}\{\hat{\mathcal{W}}, \hat{\mathcal{W}}|\mathcal{W}\} \right\} = \Omega \underbrace{E_{\mathcal{W}} \left\{ \operatorname{cov}\{\hat{\boldsymbol{w}}, \hat{\boldsymbol{w}}|\mathcal{W}\} \right\}}_{=:\boldsymbol{R}_{\tilde{w}}} \Omega^H. \tag{22}$$

Therefore, the time-average of $\bar{\sigma}_{\tilde{h}}^2$ over information subblocks in the current block is

$$\bar{\sigma}_{\tilde{h}}^2 := (N - JN_b)^{-1} \sum_{n} \sigma_{\tilde{h}}^2(n) \approx N^{-1} \sum_{n=0}^{N-1} \sigma_{\tilde{h}}^2(n)$$

$$= \frac{1}{N} \operatorname{tr}\left\{ \Omega \boldsymbol{R}_{\tilde{w}} \Omega^H \right\} = \frac{1}{N} \sigma_{\tilde{w}}^2 \tag{23}$$

where we have used $\sum_{n=0}^{N-1} \Psi^H(n)\Psi(n) = \boldsymbol{I}_{(L+1)(Q+1)}$. Similarly, the time averaged signal and noise powers turn out to be $\bar{\sigma}_{xs}^2 = \bar{\mathcal{P}}_c[\bar{\sigma}_{\tilde{h}}^2 + (L+1)\sigma_h^2]$ and $\bar{\sigma}_{x\eta}^2 = \bar{\mathcal{P}}_c \bar{\sigma}_{\tilde{h}}^2 + \sigma_\eta^2$. Therefore, we obtain an effective average SNR for (16) as

$$\mathrm{SNR}_d = \bar{\sigma}_{xs}^2 / \bar{\sigma}_{x\eta}^2. \tag{24}$$

Define the total information power and received signal power $\mathcal{P}_c = JN_c\bar{\mathcal{P}}_c$ and $\mathcal{P} := \mathcal{P}_b + \mathcal{P}_c$, respectively. Define the training power overhead

$$\beta := \mathcal{P}_b / \left[\mathcal{P}_c + \mathcal{P}_b\right]. \tag{25}$$

Our objective is to maximize SNR with respect to β under the constraint of a fixed $\mathcal{P}$. Thus, incorporating those constraint-carrying variables into (24) and using the developed expression for average signal and noise powers in (24), we obtain the unconstrained cost

$$\mathrm{SNR}_d(\beta) = \frac{\frac{(1-\beta)\mathcal{P}}{JN_c}[\bar{\sigma}_{\tilde{h}}^2 + (L+1)\sigma_h^2]}{\frac{(1-\beta)\mathcal{P}}{JN_c}\bar{\sigma}_{\tilde{h}}^2 + \sigma_\eta^2}. \tag{26}$$

Using the lower bound of LS estimator in (15), we can explicitly write (24) as

$$\mathrm{SNR}_d(\beta) = \left[f_1\beta^2 + f_2\beta + f_3\right]\left[g_1\beta + g_2\right]^{-1}, \tag{27}$$

where $f_1 = -[\mathcal{P}(L+1)\sigma_h^2]/[N_cJ]$, $f_2 = [\mathcal{P}(L+1)\sigma_h^2 - (L+1)(Q+1)\sigma_\eta^2]/[N_cJ]$, $f_3 = [(L+1)(Q+1)\sigma_\eta^2]/[N_cJ]$, $g_1 = \sigma_\eta^2 - [(L+1)(Q+1)\sigma_\eta^2]/[N_cJ]$ and $g_2 = [(L+1)(Q+1)\sigma_\eta^2]/[N_cJ]$. Setting the first derivative of $\mathrm{SNR}_d(\beta)$ with respect to β to zero, we obtain a quadratic equation in β with two roots, one of which is negative ($\beta < 0$), and hence is excluded, and the other root is given by

$$\beta_{\mathrm{opt}} = \frac{g_2}{g_1}\left[-1 + \sqrt{1 + \frac{g_1(f_3g_1 - f_2g_2)}{g_2^2 f_1}}\right]. \tag{28}$$

4. SIMULATION EXAMPLES

In the following examples we consider a doubly-selective channel with $L = 2$ in (2). We use binary phase shift keying (BPSK) modulation. Each transmitted block has $J = 10$ subblocks, and each subblock has $N_c = 30$ information symbols and $N_b = 2L+1 = 5$ training symbols with optimal structure $[0,0,b,0,0]^T$, ($b > 0$). A doubly-selective Rayleigh fading channel $h(n;l)$ is simulated according to [10, 11] with channel order $L = 2$, carrier frequency of 2 GHz, data rate of 40 kbps, and thus, symbol duration $T_s = 25\mu$s. Therefore, each tap of the generated time-variant channel has a Jakes' spectrum; it is not generated using the assumed BEM modeling. The 3 taps of the channel are mutually independent and the channel power is also normalized to unity. The SNR refers to the average signal power per block divided by the average noise power per block, where the signal power includes both training and information symbols.

Example 1: LS Channel Estimation. In this case we pick $b = 1$ leading to $\mathcal{P}_b = 10$ and $\mathcal{P}_c = 300$. We picked f_d =40 Hz leading to $\nu_{\mathrm{Dmax}} = f_dT_s$= 0.001 . The LS estimator (11) is used to estimate $\boldsymbol{w}$, and then the channel is estimated as $\hat{h}(n;l) = \sum_{q=0}^{Q} \hat{w}_q(l)\psi_q(n)$. Based on M_r Monte Carlo runs, the channel estimation MSE is calculated as

$$\mathrm{MSE} = (M_rN)^{-1} \sum_{i=1}^{M_r} \sum_{n=0}^{N-1} \sum_{l=0}^{L} |\hat{h}^{(i)}(n;l) - h^{(i)}(n;l)|^2 \tag{29}$$

where superscript (i) denotes the results of the i-th run. In Fig. 1, the lower bound in (15) is compared with the simulation results (averaged over 200 Monte Carlo runs); also shown are $\pm\sigma$ bounds on the simulation averages. It can be seen that the theoretical results are consistent with the simulation results indicating that the optimal pilot design does minimize channel MSE when using DPS-BEM.

Example 2: Training Power Allocation Here we vary training power (by varying b) with fixed total transmitted power. The BER versus β (see (25)) based on simulation results (averaged over 1000 Monte Carlo runs) is shown in Fig. 2 for SNR of 15 dB where we used a Viterbi detector based on the estimated channel for data detection. We also varied ν_{Dmax}. In Fig. 3, we plot the optimum theoretical values of β (derived in (28)) versus the received signal SNR. Comparing Figs. 2 and 3, we see that the two show mutually consistent results supporting our theoretical results: the optimal (simulations based) β inferred from Fig. 2 is in good agreement with the theoretical β_{opt} of Fig. 3.

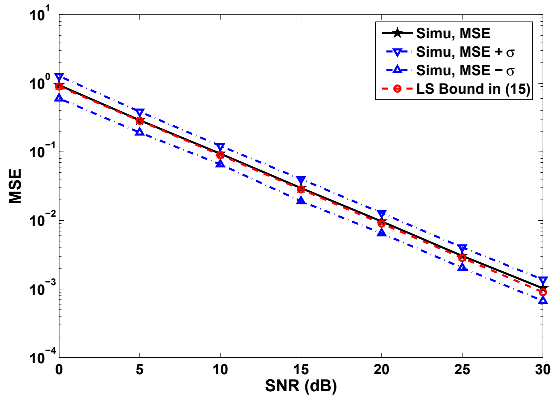

Fig. 1. Comparison between channel estimation MSE lower bound in (15) and simulation-based results. ν_{Dmax}=0.001 i.e. f_d= 40 Hz

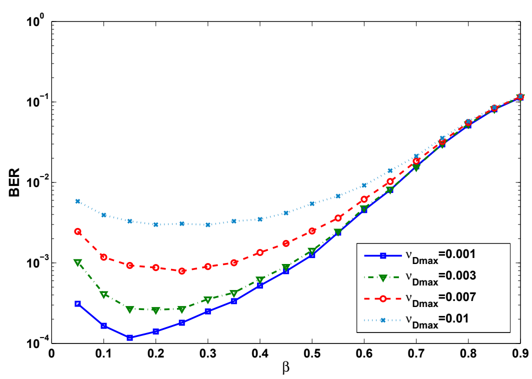

Fig. 2. Simulations-based BER versus β (25) for SNR=15 dB.

5. COMPARISONS WITH EXISTING RESULTS

Our model development in Sec. 2.1 is exactly as in [4] except that instead of CE-BEM as in [4] we use DPS-BEM. In [4] linear MMSE channel estimator is used which requires knowledge of the noise variance and of the covariance matrix $E\{\boldsymbol{w}\boldsymbol{w}^H\}$ of the channel BEM coefficient; while the former may be known at the receiver, the latter is seldom known. In [4] $E\{\boldsymbol{w}\boldsymbol{w}^H\}$ is assumed to be known and diagonal. For Jakes' model, the BEM coefficients for a given tap are not mutually uncorrelated, hence the diagonal assumption does not always hold true. In this paper we do not need to know $E\{\boldsymbol{w}\boldsymbol{w}^H\}$ or make any assumption regarding its nature. [On the other hand, performance of linear MMSE channel estimator is better than that of the LS channel estimator; however, the difference is negligible at SNRs $\geq$10 dB for the model considered in this paper.] In [4] training power allocation is carried out by maximizing a lower bound on average capacity whereas we do so by optimizing an effective SNR "for equalization." In [4] channel MMSE is linked to a channel capacity lower bound by using the fact that linear MMSE estimator is also MMSE estimator for jointly Gaussian processes; we do not (yet?) have any such result. Ref. [10] is the first to apply DPS-BEM for doubly-selective channel modeling and estimation. However, [10] does not consider pilot design.

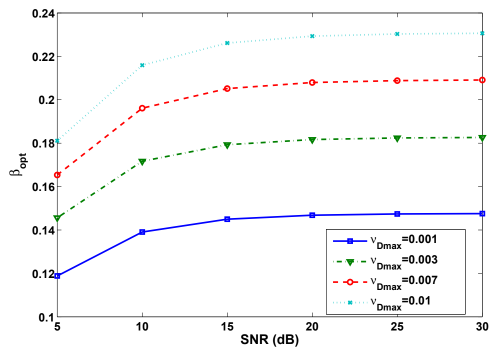

Fig. 3. Theoretical β_{opt} (28) versus received signal SNR

6. REFERENCES

[1] J.K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686-693, Nov. 1991.

[2] G.B. Giannakis and C. Tepedelenlioğlu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE*, vol. 86, pp. 1969-1986, Nov. 1998.

[3] G. Leus, S. Zhou and G.B. Giannakis, "Orthogonal multiple access over time- and frequency- selective channels," *IEEE Trans. Info. Theory*, vol. 49, pp. 1942-1950, Aug. 2003.

[4] X. Ma, G.B. Giannakis and S. Ohno, "Optimal training for block transmissions over doubly-selective wireless fading channels," in *IEEE Trans. Signal Proc.*, vol. 51, pp. 1351-1366, May 2003.

[5] S. Ohno and G.B. Giannakis, "Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels," *IEEE Trans. Inform. Theory.*, vol. 50, pp. 2138-2145, Sept. 2004.

[6] J.G. Proakis and D.G. Manolaks, *Digital Signal Processing*, Third Ed., Englewood Cliffs, NJ: Prentice-Hall, 1996.

[7] A.M. Sayeed, A. Sendonaris and B. Aazhang, "Multiuser detection in fast-fading multipath environment," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1691-1701, Dec. 1998.

[8] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty - V: The discrete case," *Bell Syst. Tech. J.*, vol. 57, pp. 1371-1430, May-June 1978.

[9] P. Stoica and R. Moses, *Introduction to Spectral Analysis*, Upper Saddle River, NJ: Prentice-Hall, 1997.

[10] T. Zemen and C.F. Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Trans. Signal Proc.*, vol. 53, pp. 3597-3607, Sept. 2005.

[11] Y.R. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 51, pp. 920-928, June 2003.