SUMMARY To describe joint time- and frequency-selective (doubly-selective) channels in mobile broadband wireless communications, we propose to use the finite parameter model based on the same Bessel functions for each tap (Bessel model). An expression of channel estimation mean squared error (MSE) based on the finite parameter models in Orthogonal Frequency Division Multiplexing (OFDM) systems is derived. Then, our Bessel model is compared with commonly used finite parameter models in terms of the channel estimation MSE. Even if the channel taps have different channel correlations and some of the taps do not coincide with the Bessel function, the channel estimation MSE of the Bessel model is shown to be comparable or outperform existing models as validated by Monte-Carlo simulations over an ensemble of channels in typical urban and suburban environments.

key words: OFDM, channel estimation, parameter model

1. Introduction

The demand for high data rate transmissions calls for broadband communication systems. In multipath propagation, the delay spread of the wireless channels exceeds the sampling period resulting in frequency-selectivity in the channel. Carrier frequency-offset due to the oscillators’ mismatch, together with Doppler shifts attributed to relative motion between the transmitter and the receiver, induces temporal variations in the channel, which is commonly referred to as time-selectivity. This joint time- and frequency-selectivity (doubly-selectivity) in mobile broadband wireless communications introduces intersymbol interference (ISI) and intercarrier interference (ICI) in block transmissions including Orthogonal Frequency Division Multiplexing (OFDM), which impact on the communication performance adversely.

Equalizers can mitigate the doubly-selective channels. But, to design equalizers, reliable and highly accurate channel estimates are required. To facilitate channel estimation, several finite parameter models for doubly-selective channels have been proposed. In [4], the exponential basis expansion model (EBEM) describes the intrablock channel time variations by a superposition of time-invariant coefficients modulated by complex Fourier exponentials, while in [2], the polynomial time-varying model does the same using a small set of polynomial functions based on Taylor series expansion. However, the EBEM fails to track the doubly-selective channel at the edges of the block resulting in significant modeling error [1], [6]. In [1], this problem is partially reduced by employing the oversampled basis expansion model (OBEM) which is an extension of the EBEM by using a period with multiple block lengths. On the other hand, the polynomial model (PM) has a simple deterministic kernel matrix structure that can be generated a priori. In [7], [11], the Slepian model is applied to represent the bandlimited sequences with a small number of basis functions, known as discrete prolate spheroidal (DPS) sequences, that exhibit double orthogonality property over an infinite and finite intervals.

It is well-known that stationary process can be approximated by its Karhunen-Loève (KL) expansion. If the channel is single path time-selective, then the channel can be parameterized by KL expansion of its exact channel correlation as proposed in [9]. In this paper, we propose to use Bessel functions to generate the FPM in order to describe the doubly-selective channels. The Bessel function is a more reasonable candidate than the function used in Slepian model as it assumes a U-shaped spectrum. When all the taps are independent and have the same correlation function accurately described by the Bessel function, our model is reduced to the model based on KL expansion. Following that, based on a general expression for all FPMs, we evaluate the channel estimation mean squared error (MSE) in OFDM systems, from which we show that increasing the model order of the FPM does not necessarily improve the channel estimation MSE, i.e., there exists an optimal model order for each FPM. Excessive numerical results over an ensemble of channels typical of urban and suburban environments as listed in [12] are provided to show that our Bessel model performs well in both channel estimation MSE and BER performance, even if the channel correlation of some of the taps do not coincide with the Bessel function. As a side objective, if the Doppler frequency is low, we also show that the classical block-constant fading model suffices to model the doubly-selective channel.

Notation: Upper (lower) bold face letters denote matrices (vectors). The operator ($\cdot)^H$ represents the complex conjugate transpose and ($\cdot)^T$ the transpose. $E\{\cdot\}$ stands for the statistical expectation, $tr(\cdot)$ the trace, $\| \cdot \|$ the Euclidean
norm and \( \delta(t) \) the Kronecker delta. \( I_K \) and \( 0 \) are respectively the identity matrix of size \( K \times K \) and the all-zero matrix; the subscripts are omitted when the dimension of the matrices is clear from the context. \( \text{diag}(\mathbf{x}) \) denotes a diagonal matrix with \( \mathbf{x} \) on its main diagonal and \( z = (y \mod 4) \) yields the smallest integer \( z \geq 0 \) so that \( y = nz + z \) for an integer \( n \geq 0 \).

2. Overview of the Finite Parameter Models

Assume sampling of the receiver at the information symbol transmission rate, given by \( R_x \) and take \( L \) as the maximum order of the discrete-time baseband equivalent channel. Without loss of generality, the discrete-time baseband equivalent description of the \( i \)th received sample is expressed as [4]

\[
y(i) = \sum_{l=0}^{L} h(i; l)u(i - l) + w(i), \quad l \in [0, L],
\]

where \( h(i; l) \) denotes the \( l \)th channel tap at time \( i \), \( u(i) \) the transmitted sample at time \( i \) and \( w(i) \) the additive white Gaussian noise (AWGN) with zero mean and variance \( \sigma_w^2 \).

In block-by-block processing, data symbols are transmitted in blocks. Block transmission confers the advantage that the interblock ISI can be easily eliminated by either appending cyclic prefix (CP) to the top of the block, or zero-padding to the end of the block. In Orthogonal Frequency Division Multiplexing (OFDM), CP is used. Since OFDM is the more popular mode of multicarrier block transmission, we consider block transmission with CP of length \( G \geq L \).

Denote \( \bar{N} = N + G \), where \( N \) is the block size. For clarity of exposition, we use two arguments \( n \) and \( m \) to describe the serial index \( i = m\bar{N} + n \) for \( n \in [-G, N - 1] \), where \( n \) is the index inside a block and \( m \) is the block index (cf. Fig. 1). We collect the samples \( y(i) \) into \( N \times 1 \) vectors:

\[
y^{(m)} = [y^{(m)}(0), \cdots, y^{(m)}(N - 1)]^T,
\]

where the \((n+1)\)st entry of the \( m \)th block of the received signal in the time domain is denoted as \( y^{(m)}(n) = y(m\bar{N} + n) \) and its corresponding channel coefficient \( h^{(m)}(n; l) \equiv h(m\bar{N} + n; l) \). Similarly, the transmitted signal \( u(i) \) and the noise \( w(i) \) are expressed respectively as

\[
u^{(m)} = [u^{(m)}(0), \cdots, u^{(m)}(N - 1)]^T,
\]

and \( w^{(m)} = [w^{(m)}(0), \cdots, w^{(m)}(N - 1)]^T \).

The matrix-vector counterpart of (1) after the removal of the received portion corresponding to CP at the receiver can be rewritten as

\[
y^{(m)} = H^{(m)}u^{(m)} + w^{(m)},
\]

where \( H^{(m)} \) is an \( N \times N \) “pseudo-circulant” matrix composed of the channel coefficients as

\[
\begin{bmatrix}
h^{(m)}(0; 0) & 0 & h^{(m)}(0; L) & \cdots & h^{(m)}(0; L - 1) \\
0 & h^{(m)}(L; 1) & \cdots & \cdots & \cdots \\
\vdots & \vdots & \ddots & \cdots & \vdots \\
0 & h^{(m)}(L - 1; L - 1) & \cdots & \cdots & 0 \\
0 & h^{(m)}(N - 1; L) & \cdots & \cdots & 0
\end{bmatrix}
\]

In point-to-point OFDM block transmission, prior to CP insertion, inverse fast Fourier transform (FFT) is performed on the serial transmitted data which is converted to slow parallel data. If the channel coefficients in (2) is time-invariant within a block, i.e., block constant fading, \( H^{(m)} \) becomes circulant, which enables the computationally efficient one tap frequency domain equalization to be performed when channel state information (CSI) is available. However, channels are in general time-varying due to relative motion between the transmitter and the receiver. We need to estimate the channel matrix to design equalizers.

In (2), for every \( N + G \) received samples, we need \( N(L + 1) \) channel coefficients to form the channel matrix \( H^{(m)} \). Even if \( u^{(m)} \) is a training vector known a priori at the receiver, we cannot solve for \( N(L + 1) \) channel coefficients as the number of unknown parameters exceeds the known data parameters \( N \). In our case, we approximate the doubly-selective channel with a finite parameter model (FPM) of model order \( K \) which satisfies the condition \((K + 1)(L + 1) \leq N \). Provided that the modeling is accurate, such finitely parameterized expansions render time-varying channel estimation tractable and allows efficient low-complexity channel equalization to be realized.

A general expression of all FPMs in describing the doubly-selective channel for the \( i \)th tap is given by

\[
h^{(m)}(n; l) = \sum_{k=0}^{K} f_k(n)g^{(m)}(k; l), n \in [0, N - 1],
\]

where \( g^{(m)}(k; l) \) represents the \( k \)th finite parameter model coefficient of the \( l \)th tap at block index \( m \), and \( f_k(n) \) the corresponding kernel functions that are common to all blocks and all taps at time \( n \). We can express (3) in matrix-vector form as

\[
H^{(m)} = Qf_{\bar{m}}^{(m)},
\]

where \( H^{(m)} = [h^{(m)}(0; l) \cdots h^{(m)}(N - 1; l)]^T \), \( f_{\bar{m}}^{(m)} = [g^{(m)}(0; l) \cdots g^{(m)}(K; l)]^T \) and \( Q \) denotes the \( N \times (K + 1) \) matrix common to all channel taps.

Each FPM is characterized by its distinct kernel matrix \( Q \) or \( \{f_k(n)\}_{k=0}^{K} \). Specific choices for \( \{f_k(n)\}_{k=0}^{K} \) include Fourier exponential bases (with frequencies on an FFT grid) as in the exponential basis exponential model (EBEM) where \( f_k(n) = e^{j2\pi k - K/2\pi n/N} \) for \( K, N \) even. The demerit of EBEM is that it cannot track the doubly-selective channel at the edges of the observation block even if the period of the
exponential is the same or larger than the block size $[1],[6]$. This means that for a fix $K+1$ bases, even though the EBM is critically sampled to cover the Doppler spread, Gibbs effect and spectral leakage occur which cause large modeling error at both ends of the block. To mitigate this problem, the oversampled basis expansion model (OBEM) or generalized complex exponential model (GCE-BEM) in [1] is developed by employing the EBM with a period equal to a multiple of the block length, i.e., taking $f_k(n) = e^{2\pi(n-K/2)n/(PN)}$ for $P$ a positive integer. We note that EBEM is subsumed in OBEM by putting $P = 1$. For $P = 1$, the kernel matrix for OBEM is orthogonal. On the other hand, for $P > 1$, the kernel matrix is not orthogonal for block size $N$ and to reduce computational complexity due to the non-orthogonality, QR decomposition may be performed. From numerical simulations in [6], the length $K+1$ or number of bases per $PN$ block is decided by $2[Pf_DN] + 1$, where $f_D$ denotes the maximum normalized Doppler frequency.

On the other hand, the polynomial model (PM) in [2] is designed as follows: From Taylor series expansion, an $N \times (K+1)$ Vandermonde kernel matrix $\hat{Q}$ is formed as $\hat{Q} = [\hat{q}^{(0)}, \hat{q}^{(1)}, \ldots, \hat{q}^{(K)}]$, where for $k \in [0,K]$, 

$$
\hat{q}^{(k)} = \left((-N/2)^k, (-N/2 + 1)^k, \ldots, (-1)^k, \delta(k), 1, \ldots, (N/2 - 2)^k, (N/2 - 1)^k\right).
$$

Then, $\hat{Q}$ undergoes QR decomposition to form an $N \times (K+1)$ orthogonal matrix $Q$ which is the PM kernel matrix.

In Slepian model [7],[11], the discrete prolate spheroidal (DPS) sequences are most concentrated within a discrete time interval of block length $N$ and most band-limited to bandwidth $W$. Each column of $Q$ is generated by the $K+1$ eigenvectors corresponding to the $K+1$ largest eigenvalues of an $N \times N$ matrix $R$, where

$$
[R]_{i,j} = \frac{\sin 2\pi W(i-j)}{\pi(i-j)}, \quad i,j \in [1,N],
$$

in which $[R]_{i,j}$ is the $(i,j)$th entry of $R$ and the bandwidth $W$ is equivalent to the maximum normalized Doppler frequency $f_D$. Figure 2 depicts kernel functions of Slepian model for $W = f_D = 0.0183$, $K = 2$, and $N = 64$.

3. Proposed Finite Parameter Model

In this section, we first evaluate the channel modeling mean squared error (MSE) for FPMs. For notational simplicity, since only one block is considered, we ignore the index $m$ in the sequel. For a given orthogonal kernel matrix $Q$ (which is taken to be the same for all taps), let the modeled channel $\tilde{h}_l$ be $\tilde{h}_l = Q\tilde{g}_l$. The channel modeling MSE expression is defined as

$$
J := E \left\{ \sum_{l=0}^{L} \| h_l - \tilde{h}_l \|^2 \right\}.
$$

We assume that $[h_l]_{l=0}^{L}$ are uncorrelated. Let us denote the channel correlation matrix for each tap as $R_l = E(h_l h_l^H)$ for $l \in [0,L]$. Given $Q$, the channel modeling MSE can then be expressed as

$$
J = tr \left( \sum_{l=0}^{L} R_l - Q \tilde{g}_l Q^H - Q \tilde{g}_l Q^H + \tilde{g}_l \tilde{g}_l^H \right).
$$

Clearly, the minimum channel modeling MSE is given by putting $\tilde{g}_l = Q^H h_l$. Consequently, the modeled channel is simply expressed as

$$
h_l = Q Q^H h_l,
$$

while the channel modeling MSE is written in the form

$$
J = tr \left( (I - Q Q^H) \sum_{l=0}^{L} R_l \right).
$$

If there is only one tap, i.e., single path channel, in (8), then the optimal kernel minimizing the channel modeling MSE is obtained by Karhunen-Lo`eve (KL) expansion as reported in [5],[9]. However, we have multiple taps that may have different second order statistics. If we know all $R_l$s, we can construct the optimal kernel matrix that minimizes the channel modeling MSE. Or if we have complete knowledge of the channel output statistics, we can develop a channel model based on KL expansion as in [10]. However, it is difficult to obtain the channel statistics of each tap or channel output statistics in practice. To overcome this problem, we revisit the fundamentals of time-varying wireless channels [3].

The channel impulse response or channel coefficient of each tap $l \in [0,L]$ is often (especially in an urban scenario) approximated by the sum of exponentials as $h(n;l) = \sum_{n=1}^{N_l} a_{l,n} e^{2\pi f_{ps}(\cos \theta_n)n}$, where $N_l$ denotes the total number of scatterers, $a_{l,n}$ the amplitude of each path from a scatterer and $\theta_n$ the angle of arrival of each path. Since $\theta_n$ for each path is impossible to determine, it is reasonable to assume that $\theta_n$ is uniformly distributed between $[-\pi, \pi]$. In this...
case, regardless of the amplitude of each scatterer, the correlation function of each tap is simply a multiple of the same function expressed as

\[ R(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j2\pi f_0 (\cos \theta) n} d\theta = J_0(2\pi f_D n). \tag{9} \]

where \( J_0 \) denotes the zeroth-order Bessel function of the first kind. This corresponds to the correlation function of the Jakes’ model approximated by the zeroth-order Bessel function of the first kind, which follows a U-shaped spectrum.

Back to the Slepian model, the correlation function used is not (9) but (5). Intuitively, instead of using (5) in the Slepian model that assumes a rectangular spectrum, it makes more sense to use the Bessel function to approximate the channel correlation function. Although practical Doppler shapes vary according to the different terrains, since channel tap correlation functions are often well approximated by Bessel functions, it is realistic to use Bessel functions. Thus, we propose to construct our FPM, based on Bessel functions.

Our model is developed as follows: We assume that the velocity of the moving terminal is available so that the maximum normalized Doppler frequency \( f_D \) can be utilized (at least estimated). From the maximum normalized Doppler frequency, we compute an \( N \times N \) matrix \( R \) with entries

\[ [R]_{i,j} = J_0(2\pi f_D |i - j|), \quad i, j \in [1, N]. \tag{10} \]

Subsequently, eigenvalue decomposition is performed on \( R \) to yield \( R = V \Lambda V^H \), where \( \Lambda = \text{diag}([\lambda_n]_{n=1}^{N}) \) denotes the diagonal matrix composed of eigenvalues \( [\lambda_n]_{n=1}^{N} \) of \( R \) arranged in order of decreasing magnitude and \( V \) the unitary matrix formed by the linearly independent column vectors corresponding to the eigenvalues. Each column of the kernel matrix \( Q \) of our Bessel model is then generated by the \( K + 1 \) columns of \( V \) corresponding to the \( K + 1 \) significant eigenvalues of \( R \). We note that the kernel matrix of our model is common to all the taps just like the other FPMs. Figure 2 compares kernel functions of Slepian model and Bessel model for \( f_D = 0.0183 \), \( K = 2 \), and \( N = 64 \), where clear differences between them can be seen.

In the generation of the kernel matrices, the requirement for QR decomposition for OBEM and PM and eigenvalue decomposition for both Slepian model and Bessel model contribute to the additional computation burden compared to EBM. Once the orthogonal kernel matrix is computed, the complexities for channel estimation are all the same except for EBM where computationally efficient FFT can be exploited.

4. Channel Estimation MSE in OFDM

We evaluate the channel estimation MSE in OFDM when Least Squares (LS) estimation is performed on (2) with \( u \) known to the receiver, where \( u \) is a training vector or a detected symbol vector. Our analysis holds true if \( w^{(m)} \) in (2) is white. Since co-channel interferences from adjacent cells can be well approximated by white Gaussian noises, the following MSE expression holds approximately even when co-channel interferences exist. For simplicity, we only consider PSK signaling.

The received signal in (2) can be reorganized as

\[ y = \sum_{l=0}^{L} U_l h_l + w, \tag{11} \]

where the symbol matrix corresponding to the \( l \)th tap is denoted by \( U_l = \text{diag}([u((N - l + n) \mod N)]_{n=0}^{N-1}) \). Substituting (4) into (11) gives

\[ y = \sum_{l=0}^{L} U_l Q g_l + w. \tag{12} \]

Then, the LS estimates of \( g_l \) is simply expressed as

\[ \hat{g}_l = [U_l Q]^H (U_l Q)^{-1} U_l Q h_l \tag{13} \]

Since we draw \( \{u(n)\}_{n=0}^{N-1} \) from a PSK constellation, we have \( U_l^H U_l = \sigma_u^2 I_N \), where \( \sigma_u^2 \) is the variance of the symbols. This means that (13) becomes \( \hat{g}_l = \frac{1}{\sigma_u^2} (U_l Q)^H y \), and the LS channel estimate \( \hat{h}_l \) is expressed according to

\[ \hat{h}_l = Q \hat{g}_l = \frac{1}{\sigma_u^2} Q (U_l Q)^H y. \tag{14} \]

From (11) and (14), the error \( h_l - \hat{h}_l \) is calculated as

\[ h_l - \hat{h}_l = (I - QQ^H) h_l - \frac{1}{\sigma_u^2} Q (U_l Q)^H w. \tag{15} \]

If \( \{u(n)\}_{n=0}^{N-1} \) are assumed independent and identically distributed (i.i.d.), then from the Central Limit Theorem, the following approximation holds true for large \( N \):

\[ Q^H U_l^H U_l Q = \sigma_u^2 I_N + \frac{1}{\sigma_u^2} \delta(i - j). \]

Thus, we obtain \( h_l - \hat{h}_l \approx (I - QQ^H) h_l - \frac{1}{\sigma_u^2} Q (U_l Q)^H w. \) Since \( |h_l|_{l=0}^{L} \) and \( w \) are uncorrelated, from \( E[ww^H] = \sigma_w^2 I_N \), the channel estimation MSE is factored into two components: the channel modeling MSE \( \beta = \text{tr}((I - QQ^H) \sum_{l=0}^{L} R_l) \) given by (8) and the channel identification MSE \( \gamma \) attributable to the AWGN noise as

\[ E \left\{ \sum_{l=0}^{L} ||h_l - \hat{h}_l||^2 \right\} = \beta + \gamma, \tag{16} \]

where

\[ \gamma = E \left\{ \sum_{l=0}^{L} \frac{1}{\sigma_u^2} Q (U_l Q)^H w^2 \right\} = (L + 1)(K + 1) \frac{\sigma_w^2}{\sigma_u^2}. \tag{17} \]

Obviously, to minimize the channel estimation MSE
in (16), we need to minimize both $\beta$ and $\gamma$ in (16). For a fixed $K, L$, as well as SNR, i.e., $\gamma$, we just need to minimize $\beta$ which corresponds to the channel modeling MSE. Hence, the minimization of the channel modeling MSE is equivalent to minimizing the channel estimation MSE. For a channel having correlation function of all its taps expressed exactly by the Bessel function, the Bessel model will offer the minimum mean squared channel estimation error.

Equation (16) is a universal expression of channel estimation MSE for all FPMs with orthogonal kernels, and it serves as our main comparison criterion between the models. For PM, it is shown by numerical example in [2] that the channel estimation MSE performance does not necessarily improve with an increase in the order $K$ of the finite parameter model, which may contradict our intuition. The answer to this question is provided by our theoretical evaluation of the channel estimation MSE. In (16), we show that the channel estimation MSE is governed by the channel modeling MSE $\beta$ and the channel identification MSE $\gamma$. An increase in $K$ translates into a corresponding decrease in $\beta$ but an increase in $\gamma$ and vice versa. This explains why the channel estimation MSE may not always be reduced by an increase in $K$ as shown by the numerical example for PM in [2] and later on in our simulations for various FPMs. There is an undeniable trade-off between $\beta$ and $\gamma$. Given a fixed SNR, channel order $L$ and channel correlation $R_r$, the channel estimation MSE can be minimized by balancing both $\beta$ and $\gamma$ to yield an optimal model order $K$, which is present for each FPM. If channel correlation matrices are available, then the derivation of the optimal $K$ becomes possible by numerically evaluating (16). However, since channel correlation matrices are in general unavailable, it is difficult to decide the optimal $K$. We shall investigate the relation of $K$ with the normalized Doppler frequency $f_D$ by numerical simulations.

### 5. Performance Evaluation

We first compare the Bessel model, OBEM, PM and Slepian model in terms of the channel estimation MSE under the assumption of perfect timing and carrier synchronization, i.e., we do not take time-selectivity due to timing and carrier mismatch into account. Then, all these models will be applied in an uncoded OFDM system to test their respective BER performances. We set the block size $N = 64$. The doubly-selective channel is composed of $L + 1$ component channel taps which follows Rayleigh’s and/or Rice’s envelope with non-uniformed power profile.

Three channel models of random delay spreads [12] are used, the characteristics of each is listed in Table 1. The sampling period is taken to be equivalent to the BPSK symbol duration, $T_s = 1/R_s$, where $R_s = 14.4 \cdot 10^3$. The carrier frequency is assigned as 5 GHz. In our simulations, for OBEM, the value of $K$ is chosen such that the Doppler spread requirement $K = 2P/Nf_D$ is satisfied as in [6]. We vary the oversampling factor $P$ as $P = 1, \ldots, 4$. Note that OBEM with $P = 1$ corresponds to the conventional EBEM.

![Fig. 3](image) Monte-Carlo results of the channel estimation MSE for varying $K$ w.r.t. $f_D$ for channel model 1 at SNR = 15 dB.

For PM, Slepian model and the Bessel model, the best $K$ is selected based on simulations. The results are averaged over 200 channel realizations. Since the bandwidth and transmission rate are fixed, we refer $E_b/N_0$ as SNR. FPM coefficients, i.e., entries of $g_i^{[m]}$ are obtained from LS channel estimation with the help of training vector (or detected symbol vector) $u$ known at the receiver.

In order to see the relation of the number $K + 1$ of kernel functions with normalized Doppler frequency $f_D$, we perform Monte-Carlo simulations of the channel estimation MSE with our Bessel model for varying $K$ w.r.t. $f_D$ in the range 0 to 0.033 for channel model 1. The results are depicted in Fig. 3 for SNR = 15 dB and in Fig. 4 for SNR = 25 dB. We observe that at low $f_D$ of below 0.001, the block-constant fading model (or $K = 0$) gives the best performance. In fact, the channel estimation MSE perfor-
formance deteriorates with increasing $K$ at low $\bar{f}_D$. Conversely, at high $\bar{f}_D$, the channel estimation MSE for large $K$ is better than that for small $K$. This can be explained using (16). $\gamma$ is a term which is proportional to only $K$ when $L$ is fixed. At low $\bar{f}_D$, $\beta$ is small and the channel estimation MSE is mainly due to $\gamma$. This accounts for the larger channel estimation MSE for large $K$ at low $\bar{f}_D$. At high $\bar{f}_D$, $\beta$ becomes significant compared to $\gamma$ and the increase in the magnitude of $\beta$ for small $K$ is steeper than that for large $K$. Thus, the channel estimation MSE for small $K$ outstrips that for large $K$ at high $\bar{f}_D$. This is analogous to the bias-variance tradeoff in [5]. Indeed, at low $\bar{f}_D$, the simple block-constant model is sufficient for channel modeling and other FPMs are deemed unnecessary. This justifies the need for higher order FPMs only at higher $\bar{f}_D$. In our subsequent simulations, we will only consider high range of $\bar{f}_D$.

Since the optimal order $K$ for PM, Slepian model and Bessel model cannot be derived analytically, we find them by simulations. For each channel model, under the condition of a fixed $\bar{f}_D = 0.0183$ (mobile terminal’s velocity, $v = 57$ km/h) and SNR = 15 dB, we simulate the channel estimation MSE of the PM, Slepian model and Bessel model for varying $K$. The results are summarized in Table 2. Henceforth, the respective optimal $K$s will be used for the next few simulations at a fixed $\bar{f}_D = 0.0183$ and optimal $K$s found for other $\bar{f}_D$ will not be shown here due to space-constraint.

To verify the effectiveness of the Bessel model in channel estimation, we perform simulations using channel model 1, 2 and 3. At the normalized Doppler frequency $\bar{f}_D = 0.0183$ and using the optimal order $K$ for each FPM, we simulate the channel estimation MSE for PM, Bessel model, OBEM with $P = 1, 2, 3$ and 4 and Slepian model over the SNR range 0 dB to 30 dB. The results for channel model 1, 2 and 3 are plotted in Figs. 5, 6 and 7 respectively. In Fig. 5, because all the taps are Rayleigh-faded in channel model 1, the Bessel model based on the Bessel function is the optimal model which explains why it gives the best performance. In Figs. 6 and 7, Rice taps are present and the Bessel model is not always optimal. However, the Bessel model still gives the lowest channel estimation MSE compared to all the other FPMs. This validates the usefulness and precision of the Bessel model in describing the doubly-selective channels.

To see the robustness of the Bessel model, we test two Bessel models: i) One is produced based on a fixed $\bar{f}_D = 0.0236$ (Bessel fixed); ii) the other is generated by normalized Doppler frequency with an error $+0.001$ from the true $\bar{f}_D$ (Bessel error). At the optimal order $K$ for each FPM and SNR = 15 dB, we simulate the channel estimation MSE w.r.t. $\bar{f}_D$ for the range 0.016 to 0.033 using channel model 1, 2 and 3 and plot the respective results in Figs. 8, 9 and 10. Evidently, OBEM $P = 2$ gives better performance

![Fig. 4](image_url) Monte-Carlo results of the channel estimation MSE for varying $K$ w.r.t. $\bar{f}_D$ for channel model 1 at SNR = 25 dB.

<table>
<thead>
<tr>
<th>Channel model</th>
<th>PM</th>
<th>Slepian</th>
<th>Bessel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

**Table 2** Optimal order $K$ at SNR = 15 dB and $\bar{f}_D = 0.0183$.

![Fig. 5](image_url) Channel estimation MSE w.r.t. SNR for channel model 1 at $\bar{f}_D = 0.0183$.

![Fig. 6](image_url) Channel estimation MSE w.r.t. SNR for channel model 2 at $\bar{f}_D = 0.0183$. 
Fig. 7 Channel estimation MSE w.r.t. SNR for channel model 3 at $f_D = 0.0183$.

Fig. 8 Channel estimation MSE w.r.t. normalized Doppler frequency for channel model 1 at SNR = 15 dB.

Fig. 9 Channel estimation MSE w.r.t. normalized Doppler frequency for channel model 2 at SNR = 15 dB.

Fig. 10 Channel estimation MSE w.r.t. normalized Doppler frequency for channel model 3 at SNR = 15 dB.

Fig. 11 Channel estimation MSE w.r.t. normalized Doppler frequency error for channel model 1, 2 and 3 at SNR = 15 dB.

than PM, OBEM $P = 3$ and 4 over the calibrated range of $f_D$. Also, the average performance of OBEM $P = 1$ is worse than Slepian model and our Bessel model due to the fact that there is Gibbs’ effect and spectral leakage. Furthermore, even though the Bessel model with a fixed $f_D = 0.0236$ is not optimal except around $f_D = 0.0236$ for channel model 1, it still promises good channel modeling performance over the calibrated range of $f_D$, surpassing OBEM $P = 2, 3$ and 4. But at higher $f_D$, the Slepian model and OBEM $P = 1$ outperforms it. Then, we can also see that on the average Bessel model gives the best overall performance over the calibrated range of $f_D$ in channel model 1, 2 and 3. This demonstrates the robustness of our Bessel model. Even when Bessel model is fixed at a certain $f_D$, it is nonetheless a better candidate than either the OBEM, Slepian model and PM in approximating the channel over a wide range of $f_D$.

To see how the Bessel model fares in the presence of error in $f_D$, we also simulate the channel estimation MSE w.r.t. the deviation in $f_D$ of the Bessel model for channel model 1, 2 and 3. We do this by fixing channel model 1, 2 and 3 at $f_D = 0.0236 (v = 73$ km/h) and varying $f_D$ of the Bessel model. We set the model order $K$ as 4. The results for SNR = 15 dB and SNR = 25 dB are plotted in Fig. 11 and Fig. 12 respectively. Both figures show that for channel
model 1, the channel estimation MSE loss is symmetrical about $f_D = 0.0236$. But for channel model 2 and 3, the point of symmetry is higher at $f_D = 0.026$. The Bessel model does not give the best performance at $f_D = 0.0236$ for channel model 2 and 3 because of the presence of Rice taps.

Next, we sent BPSK symbols over channel model 1, 2 and 3 in an OFDM system with 64 subcarriers. To see the effects of modeling error on BER performance, we apply the Bessel model, PM, Slepian and OBEM with $P = 2$ set at their optimal order $K$ to describe channel model 1, 2 and 3 at $f_D = 0.0318$ ($v = 96$ km/h). Based on (7), $\tilde{h}$ for each FPM are generated and used in the effective ICI-suppressing Viterbi-type ML equalizer [8] which is employed for equalization of the received samples. Figures 13, 14 and 15 depict the simulation results conducted over the SNR range 0 dB to 15 dB for channel model 1, 2 and 3. As a benchmark, the BER performance using the true channel (where perfect CSI is available) for equalization, is also shown. We can see that albeit all the models suffer from BER degradation compared to the true channel, the Bessel model accords the closest performance to the true channel out of all the models, with the Slepian model coming a close second. This reinforces the fact that the Bessel model indeed describes the channel best, as reflected in the low channel estimation MSE in Figs. 8, 9 and 10 thus accounting for the better BER performance.

6. Conclusions

We have done a comprehensive comparative study for various finite parameter models for time-varying multipath channels through discussions and numerical simulations. In addition, we have also developed our Bessel model. Using the channel estimation MSE as the main criterion for comparison, and under reasonable settings over an ensemble of channels, numerical simulations reveal that our Bessel model is comparable or more superior than the EBEM, OBEM, Slepian model or PM. Even when the Bessel model uses kernels based on a Bessel function obtained at a fixed maximum normalized Doppler frequency, its performance is still comparable to the other models at a wide range of normalized Doppler frequency. Only at higher normal-
ized Doppler frequency is its performance surpassed by the EBEM and Slepian model. Besides that, the Bessel model demonstrates a better BER performance than the PM, OBEM and Slepian model when it is applied in an uncoded OFDM system. We have also concluded that the block-constant fading model actually suffices to model the channel at very low normalized Doppler frequency.

References


Kok Ann Donny Teo is awarded the Japanese Monbuk-Kagakusho Scholarship and the Public Service Commission (PSC) Singapore Government Scholarship in 1999 to study in Japan, where he graduated with B.Eng. degree (Good Honors) in 2004 and M.Eng. degree in 2006 from Hiroshima University, Japan. His research interests lie in the areas of channel estimation, equalization, diversity techniques, and broadband wireless communication systems.

Shuichi Ohno received the B.E., M.E. and Dr. Eng. degrees in applied mathematics and physics from Kyoto University, in 1990, 1992 and 1995, respectively. From 1995 to 1999 he was a research associate in the Department of Mathematics and Computer Science at Shimane University, Shimane, Japan. He joined the Department of Artificial Complex Systems Engineering, Hiroshima University, Hiroshima, Japan, in April 2002, where he is currently an Associate Professor. His current interests are in the areas of digital communications, signal processing for communications and adaptive signal processing. Dr. Ohno is a member of the IEEE and the Institute of Systems, Control, and Information Engineers in Japan.