# We are IntechOpen,
# the world's leading publisher of
# Open Access books
# Built by scientists, for scientists

## 6,900
Open access books available

## 186,000
International authors and editors

## 200M
Downloads

## 154
Countries delivered to

Our authors are among the

## TOP 1%
most cited scientists

## 12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
**BOOK CITATION INDEX**
INDEXED

**WEB OF SCIENCE**™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

## Interested in publishing with us?
## Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

# Superimposed Training-Aided Channel Estimation for Multiple Input Multiple Output-Orthogonal Frequency Division Multiplexing Systems over High-Mobility Environment

Han Zhang[1], Xianhua Dai[2], Daru Pan[1] and Shan Gao[1]
*[1]School of Physics and Telecommunications Engineering,*
*South China Normal University, Guangzhou*
*[2]School of Information Science and Technology,*
*SUN Yat-sen University, Guangzhou*
*China*

## 1. Introduction

The combination of multiple-input multiple-output (MIMO) antennas and orthogonal frequency-division multiplexing (OFDM) can achieve a lower error rate and/or enable high-capacity wireless communication systems by flexibly exploiting diversity gain and/or the spatial multiplexing gains. Such systems, however, rely upon the knowledge of propagation channels. In many mobile communication systems, transmission is impaired by both delay and Doppler spreads [1]-[7]. In such cases, explicit incorporation of the time-varying characteristics of mobile wireless channel is called for.

The coefficients of a linearly time-varying (LTV) channel can be usually modeled as uncorrelated stationary random processes which are assumed to be low-pass, Gaussian, with zero mean (Rayleigh fading) or non-zero mean (Rician fading) depending on whether line-of-sight propagation is absent or present [1][6]. Recently, the basis expansion models, i.e. the truncated discrete Fourier basis (DFT) models, polynomial models and discrete prolate Spheroidal sequence models, have gained special attentions, especially for the situation that channel is caused by a few strong reflectors and path delays exhibit variations due to the kinematics of the mobiles [1]-[2] [5]-[6] [16] [25]-[28].

In conventional pilot-aided channel estimation approaches, MIMO channels can be effectively estimated by utilizing the time-division multiplexed (TDM) and (or) frequency-division multiplexed (FDM) training sequences [5]-[7] [20]-[23] [25]. Although the channel estimates are in general reliable, extra bandwidth or time slot is required for transmitting known pilots. In recent years, an alternative approach, referred to as superimposed training (ST), has been extensively studied in [8]-[19] [26]-[28]. In the idea of ST, additional periodic training sequences are arithmetically added to information sequence in time- or frequency-

domain. The advantage of the scheme is that there is no loss in information rate, and thus enables higher bandwidth efficiency. However, some useful power must inevitably be allocated to the pilots, and thus resulting in information signal-to-noise ratio (SNR) reduction. Meanwhile, the information sequences are viewed as interference to channel estimation since pilot symbols are superimposed at a low power to the information sequences at the transmitter. The existing ST-based channel estimations are mainly restricted to the case where the channel is linearly time-invariant (LTI), where the channel transfer function can be estimated by using first-order statistics [8]-[13] [17]-[18]. In the latest contributions, J. K. Tugnait [16] extended the conventional ST to time-varying environment where the LTV channels are modeled by complex exponential bases. For the issue of training power allocation, the optimal pilot power has been investigated by [24] for different taps of low-pass filter, and then, the optimization of ST power allocation for LTI channel is mathematically analyzed based on equalizer design [15] [19].

In this paper, a new ST-based channel estimator is proposed for OFDM/MIMO systems over LTV multipath fading channels. The main contributions are twofold. First, the LTV channel coefficients modeled by the truncated discrete Fourier bases (DFB), unlike the existing approaches [1]-[2] [5]-[6] [16], cover multiple OFDM symbols. Then, a two-step channel estimation approach is adopted for LTV channel estimation. Furthermore, a closed-form expression of the estimation variance is derived, which provides a guideline for designing the superimposed pilot symbols. We demonstrate by analytical analysis that the estimation variance, unlike that of conventional ST-based schemes [8]-[19], approaches to a fixed lower-bound as the training length increases. Second, for wireless communication systems with a limited transmission power, unlike [10] where the issue of ST power allocation is derived by optimizing the SNR for equalizer design, we provide an optimal solution of ST power allocation with a different point of view by maximizing the lower-bound of channel capacity. Comparatively, the training power allocation scheme [10] can be otherwise considered as a special case compared with the proposed approach. In simulations presented in this paper, we compare the results of our approach with that of the FDM training approaches [5] as latter serves as a "benchmark" in related works. It is shown that the proposed algorithm outperforms that of FDM training, and yields higher transmission efficiency.

The rest of the paper is organized as follows. Section II presents the system and channel models. In Section III, we estimate the LTV channel coefficients with the proposed two-step channel estimation approach. In Section IV, we derive the closed-form expression of the channel estimation variances. Section V determines the optimal ratio of the ST power to the total transmission power by maximizing the lower-bound of channel capacity. Section VI reports on some simulation experiments in order to test the validity of theoretic results, and we conclude the paper with Section VII.

*Notation:* The letter $t$ represents the time-domain variable and $k$ is the frequency-domain variable. Bold letters denote the matrices and column-vectors, and the superscripts $[\bullet]^T$ and $[\bullet]^H$ represent the transpose and conjugate transpose operations, respectively. $[\bullet]_{k,t}$ denotes the $(k, t)$ element of the specified matrix.

## 2. System and channel model

### 2.1 System model

Consider an MIMO/OFDM system of $N$ transmitters or mobile users and a receive array of $M$ receive antennas with perfect synchronization. At transmit terminals, an inverse fast Fourier transform (IFFT) is used as a modulator. The modulated outputs are given by

$$\mathbf{X}_n(i) = \left[ x_n(i,0), \cdots x_n(i,t), \cdots x_n(i,B-1) \right]^T = \mathbf{F}^{-1}\mathbf{S}_n(i) \qquad n = 1, \cdots N \tag{1}$$

where $B$ is OFDM symbol-size, $\mathbf{S}_n(i) = [s_n(i,0), \cdots s_n(i,k), \cdots s_n(i,B-1)]^T$ is the $i$ th transmitted symbol of the $n$ th transmit antenna. $\mathbf{F}^{-1}$ is the IFFT matrix with $[\mathbf{F}^{-1}]_{k,t} = e^{j2\pi kt/B}$ and $j^2 = -1$. Then, $\mathbf{X}_n(i)$ is concatenated by a cyclic-prefix (CP) of length $\bar{L}$, propagating through the respective channels. At receiver, the received signals of $m$ th receive antenna, discarding CP and stacking the received signals $y^{(m)}(i,t)$ $t = 0, \cdots B-1$, can be written in a vector-form as

$$\mathbf{Y}^{(m)}(i) = \left[ y^{(m)}(i,0), \cdots y^{(m)}(i,t), \cdots y^{(m)}(i,B-1) \right]^T \qquad m = 1, \cdots M \tag{2}$$

and the received signals $y^{(m)}(i,t)$ in (2) is given by

$$y^{(m)}(i,t) = \sum_{n=1}^{N} \mathbf{X}_n(i) \otimes \mathbf{h}_n^{(m)}(i) + v^{(m)}(i,t)$$

$$= \sum_{n=1}^{N} \sum_{l=0}^{L-1} h_{n,l}^{(m)}(i,t) x_n(i,t-l) + v^{(m)}(i,t) \qquad t = 0, \cdots B-1 \tag{3}$$

where $\mathbf{h}_n^{(m)}(i) = \left[ h_{n,0}^{(m)}(i,t), \cdots h_{n,L-1}^{(m)}(i,t), 0_{1 \times B-L} \right]^T$ is the impulse response vector of the propagating channel from the $n$ th transmit to the $m$ th receive antenna. The channel coefficients $h_{n,l}^{(m)}(i,t)$, $l = 0, \cdots L-1$ is the functions of time variable $t$ which will be defined by (6). The notation $\otimes$ represents the cyclic convolution and $v^{(m)}(i,t)$ is the additive Gaussian noise.

At receiver, an FFT operation is performed on the vector (2), and the demodulated outputs can be written as

$$\mathbf{U}^{(m)}(i) = \left[ u^{(m)}(i,0), \cdots u^{(m)}(i,k), \cdots u^{(m)}(i,B-1) \right]^T = \mathbf{F}\mathbf{Y}^{(m)}(i) \qquad m = 1, \cdots M. \tag{4}$$

From (3) and the duality of time and frequency, the FFT demodulated signals in (4) can be written as

$$u^{(m)}(i,k) = FFT \left\{ \sum_{n=1}^{N} \sum_{l=0}^{L-1} h_{n,l}^{(m)}(i,t) x_n(i,t-l) + v^{(m)}(i,t) \right\}$$

$$= \sum_{n=1}^{N} \sum_{l=0}^{L-1} FFT \left\{ h_{n,l}^{(m)}(i,t) \right\} \otimes FFT \left\{ x_n(i,t) \right\} + \bar{v}^{(m)}(i,k) \tag{5}$$

where $FFT\{\bullet\}$ represents the FFT vector of the specified function and $\overline{v}^{(m)}(i,k)$ is the frequency-domain noise. Compared with the FFT demodulated signals of OFDM systems with LTI channels, the convolution in (5) between the information sequences and the FFT vectors of time-varying channel coefficients may introduce inter-carrier interference (ICI).

## 2.2 Channel model

As mentioned in [1], the coefficients of the time- and frequency-selective channel can be modeled as Fourier basis expansions. Thereafter, this model was intensively investigated and applied in block transmission, channel estimation and equalization (e.g. [2][5]-[6][16]). In this paper, we extend the block-by-block process [2][5]-[6][16] to the case where multiple OFDM symbols are utilized. Consider a time interval or segment $\{t:(\ell-1)\Omega \le t \le \ell\Omega\}$, the channel coefficients in (3) can be approximated by truncated discrete Fourier bases (DFB) within the segment as

$$h_{n,l}^{(m)}(i,t) \approx \sum_{q=0}^{Q} h_{n,l,q}^{(m)} e^{-j2\pi(q-Q/2)t/\Omega}$$

$$= \sum_{q=0}^{Q} h_{n,l,q}^{(m)} \eta_q(t) \quad t=(\ell-1)\Omega,\cdots\ell\Omega , \ \ell=1,2,\cdots$$

(6)

where $h_{n,l,q}^{(m)}$ is a constant coefficient, $Q$ represents the basis expansion order that is generally defined as $Q \ge 2f_d\Omega/f_s$ [1], $\Omega > B$ is the segment length and $\ell$ is the segment index. Unlike [1]-[2] [5]-[6] [16], the approximation frame $\Omega$ covers multiple OFDM symbols, denoted by $i=1,\cdots I$ , where $I=\Omega/B'$ and $B'=B+\overline{L}$ . Since the proposed two-step channel estimation as will be shown in Section III is adopted within one frame, we omit the segment index $\ell$ for simplicity.

## 3. ST-based channel estimation

In this section, we propose a ST-based two-step approach for LTV channel estimation. In ST-based approaches [8]-[19], the pilot symbols are superimposed (arithmetically added) to the information sequences as

$$s_n(i,k)=c_n(i,k)+p_n(i,k) \quad k=0,\cdots B-1$$

(7)

where $c_n(i,k)$ and $p_n(i,k)$ are the information and pilot sequence, respectively. Compared with the FDM/TDM training aided methods [20]-[22], ST requires no additional bandwidth (or time-slot) for transmitting known pilots, and thus offers a higher data rate.

### 3.1 ST-based channel estimation over one OFDM symbol

For LTV environment where the channel coefficient $h_{n,l}^{(m)}(t)$ is a function of time variable $t$, the vectors $FFT\{h_{n,l}^{(m)}(t)\}$ in (5) cannot be approximated as a $\delta$-sequences and, the FFT demodulated signals at the sub-carrier $k$ of the $i$ th symbol is given by

$$u^{(m)}(i,k) = \sum_{n=1}^{N}\sum_{l=0}^{L-1} FFT\left\{h_{n,l}^{(m)}(i,t)\right\} \otimes P_n(i) + \overline{v}^{\,\prime(m)}(i,k)$$

$$\approx \sum_{n=1}^{N}\sum_{l=0}^{L-1} FFT\left\{\sum_{q=0}^{Q} h_{n,l,q}^{(m)}\eta_q(t)\right\} \otimes P_n(i) + \overline{v}^{\,\prime(m)}(i,k) \qquad (8)$$

$$= \sum_{n=1}^{N}\sum_{q=0}^{Q} H_{n,q}^{(m)}(i,k)\eta_q(t_i)W_q(i,k) \otimes P_n(i) + \overline{v}^{\,\prime(m)}(i,k)$$

where $\overline{v}^{\,\prime(m)}(i,k) = \sum_{n=1}^{N}\sum_{l=0}^{L-1} FFT\left\{h_{n,l}^{(m)}(t)\right\} \otimes C_n(i) + \overline{v}^{(m)}(i,k)$ and $t_i = (i-1)B' + B/2$. $W_q(i,k)$ with $k = 0,\cdots B-1$ is the FFT vector of the complex exponential function (CEF) and, can be written as

$$W_q(i,0) = \left[w_q(i,0),\cdots w_q(i,k),\cdots w_q(i,B-1)\right]^T$$

$$= F\left[\eta_q(t_i - B/2)/\eta_q(t_i),\cdots\eta_q(t_i + B/2 - 1)/\eta_q(t_i)\right]^T. \qquad (9)$$

Notice that $W_q(i,k)$ is a cyclic-shifted vector of $W_q(i,0)$ with a shifting length $k$. On the other hand, ICI introduced by the cyclic convolution $W_q(i,k) \otimes S_n(i)$ depends explicitly on $\eta_q(t)$, $t = 0,\cdots B-1$. When $q$ is not large, the complex exponential functions in (9) are slowly time-varying over an OFDM symbol-duration and, thereby, the principal power or major-lobe of the FFT vector $W_q(i,0)$ may concentrate on its two ends (low frequency tones) with indexes $0,\cdots T$ and $B-T,\cdots B-1$. Using the major-lobe to approximate the CEF vectors $W_q(i,0)$ $q = 0,1,\cdots Q$, we have

$$W_q(i,0) \approx \left[w_q(i,0),\cdots w_q(i,T),0,\cdots\cdots 0, w_q(i,B-T),\cdots w_q(i,B-1)\right]^T \quad q = 0,1,\cdots Q \qquad (10)$$

where $T$ is a positive integer.

In general, the FFT vector of the function $\eta_q(t)$ in (10) may have a great side-lobe that results in a great error. For improving the approximation performance, an intuitional idea is to apply a window function to the received signals in order to reduce the side-lobe leakage. The windowed vector of received signals in (3) of the $i$ th symbol is

$$\overline{y}^{(m)}(i,t) = \sum_{n=1}^{N} h_n^{(m)}(i,t)\psi_B(t)x_n(t-l) + v^{(m)}(t)\psi_B(t) \qquad t = 0,\cdots B-1 \qquad (11)$$

where $\psi_B(t)$ is a time-domain windowing function with a length $B$. Performing the FFT demodulated operation on the windowed sequences in (10), the demodulated signals, by (10), can be written by

$$u^{(m)}(i,k) \approx \sum_{n=1}^{N}\sum_{l=0}^{L-1} FFT\left\{\sum_{q=0}^{Q} h_{n,l,q}^{(m)}\eta_q(t)\psi_B(t)\right\} \otimes P_n(i) + \overline{v}^{\,\prime(m)}(i,k)$$

$$= \sum_{n=1}^{N}\sum_{q=0}^{Q} H_{n,q}^{(m)}(i,k)\eta_q(t_i)\overline{W}_q(i,k) \otimes P_n(i) + \overline{v}^{\,\prime(m)}(i,k) \qquad (12)$$

where $\overline{W}_q(i,k)$ is the CEF vector with the windowing function $\psi_B(t)$ as

$$
\begin{aligned}
\overline{W}_q(i,k) &= \left[ \overline{w}_q(i,0), \cdots \overline{w}_q(i,k), \cdots \overline{w}_q(i,B-1) \right]^T \\
&= \mathbf{F} \left[ \psi_B(t)\eta_q(t_i - B/2)/\eta_q(t_i), \cdots \psi_B(t)\eta_q(t_i + B/2 - 1)/\eta_q(t_i) \right]^T \\
&\approx \left[ \overline{w}_q(i,0), \cdots \overline{w}_q(i,T), 0, \cdots 0, \overline{w}_q(i,B-T), \cdots \overline{w}_q(i,B-1) \right]^T .
\end{aligned}
\tag{13}
$$

Compared with (10), the approximation of windowing based vector has a much smaller side-lobe with the same index $T$. The experiment studies show that by using a *Kaiser* function [5], the approximation in (13) of $T = 2$ may capture almost 99% power of $FFT\left[ \psi_B(t)\eta_q(t_i - B/2 : t_i + B/2 - 1)/\eta_q(t_i) \right]^T$ for truncated DFBs when $q < B/10$. Substituting (13) into (12), the FFT demodulated outputs can be approximated by

$$
\begin{aligned}
u^{(m)}(i,k) = \sum_{n=1}^{N}\sum_{q=0}^{Q} H_{n,q}^{(m)}(i,k)\eta_q(t_i) \Bigg[ & \sum_{k'=0}^{T} \overline{w}_q(i,k')p_n(i,k-k') + \\
& \sum_{k'=T}^{1} \overline{w}_q(i,B-k')p_n(i,k+k') \Bigg] + \overline{v}'^{(m)}(i,k).
\end{aligned}
\tag{14}
$$

The first term of (14) illustrates that $2T + 1$ tones, i.e. $p_n(i,k-k'), \cdots p_n(i,k+k')$ should be jointly designed for estimating $H_{n,q}^{(m)}(i,k)$. We refer to such $2T + 1$ consecutive pilot tones as a pilot cluster for differentiating from the isolated tones utilized in the LTI channel estimation [19] [22]-[23]. Denote $k_1, \cdots k_\Gamma$ as the pilot cluster indexes located at the $\tau$ th pilot symbol and, $p_n(k_\tau - T), \cdots p_n(k_\tau + T)$ as the pilot sequences at the pilot cluster $k_\tau$. Since the ST does not entail additional bandwidth, two adjacent pilot-clusters, i.e. $k_\tau$ and $k_{\tau+1}$ can be placed closed together. The pilot tone distribution is shown in Fig. 1.
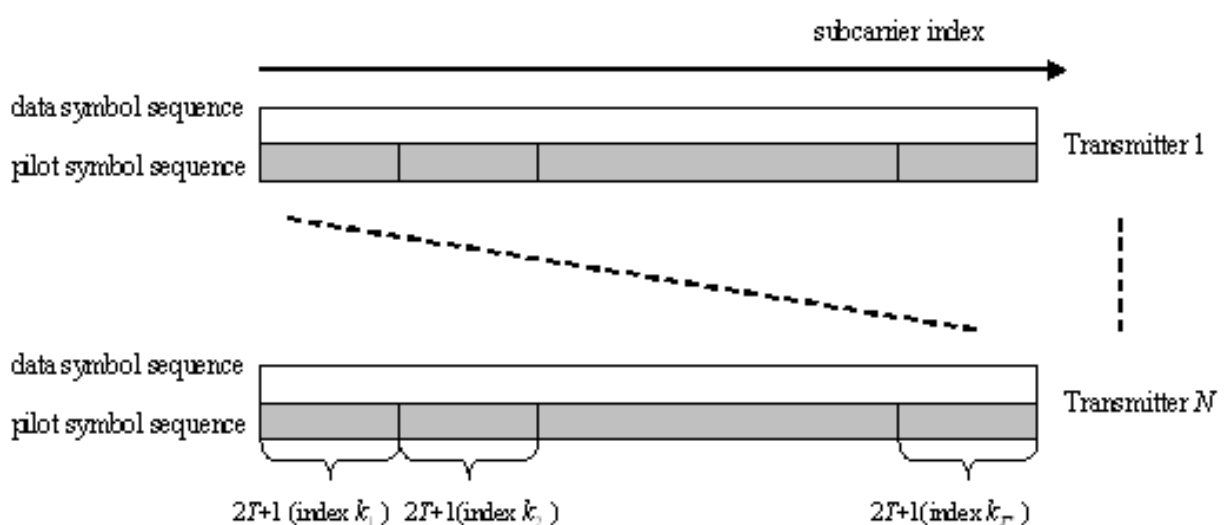


Fig. 1. A typical pilot tone distribution. $2T + 1$ consecutive tones are grouped together as one pilot cluster. All pilot clusters are uniformly distributed in frequency domain with each adjacent pilot cluster being closed together.

Then, we focus on ST design. From (14), when the training sequence at each pilot cluster is designed as either a constant modulus sequence, i.e.

$$p_n(i,k_\tau) = p_n(i,k_\tau \pm k') \quad \tau = 1,\cdots \Gamma, \ k' = 1,\cdots T \tag{15}$$

or a $\delta$ sequence, i.e.

$$p_n(i,k_\tau \pm k') = \begin{cases} p_n(i,k_\tau), k' = 0 \\ 0, \quad otherwise \end{cases} \quad \tau = 1,\cdots \Gamma, \ k' = 1,\cdots T. \tag{16}$$

Accordingly, the FFT demodulated outputs at pilot cluster $k_\tau$ can be approximated as

$$u^{(m)}(i,k_\tau) \approx \sum_{n=1}^{N} \sum_{q=0}^{Q} H_{n,q}^{(m)}(i,k_\tau) \eta_q(t_i) g_q(i,k_\tau) p_n(i,k_\tau) + \overline{v}^{\,\prime(m)}(i,k_\tau)$$

$$= \sum_{n=1}^{N} H_n^{\,\prime(m)}(i,k_\tau) p_n(i,k_\tau) + \overline{v}^{\,\prime(m)}(i,k_\tau) \tag{17}$$

where $g_q(i,k_\tau) = \sum_{k'=0}^{T} \overline{w}_q(i,k_\tau + k') + \sum_{k'=T}^{1} \overline{w}_q(i,k_\tau + B - k')$ if the training sequence takes value from (15) and $g_q(i,k_\tau) = \overline{w}_q(i,k_\tau)$ for (16), which are all known, respectively. The channel transfer functions $H_n^{\,\prime(m)}(i,k_\tau)$ is given by

$$H_n^{\,\prime(m)}(i,k_\tau) = \sum_{q=0}^{Q} H_{n,l,q}^{(m)}(i,k_\tau) \eta_q(t_i) g_q(i,k_\tau)$$

$$= \sum_{l=0}^{L-1} \sum_{q=0}^{Q} h_{n,l,q}^{(m)} \eta_q(t_i) g_q(i,k_\tau) e^{-j2\pi k_\tau l/B} \approx \sum_{l=0}^{L-1} \hbar_{n,l}^{(m)}(t_i) e^{-j2\pi k_\tau l/B}. \tag{18}$$

From (17)-(18), we note that $H_n^{\,\prime(m)}(i,k_\tau), \tau = 1,\cdots \Gamma$ is in fact a LTI system transfer function of which the coefficients are the mid-values of the LTV channel at the $i$ th OFDM symbol interval. As a result, the LTV channel estimation can be approximately reduced into that of the LTI channel [22] and [23] by simply designing the ST sequences as (15) or (16).

Let $\mathbf{H}^{(m)}(i) = \left[ \hbar_{1,0}^{(m)}(t_i),\cdots \hbar_{1,L-1}^{(m)}(t_i),\cdots \cdots \hbar_{N,0}^{(m)}(t_i),\cdots \hbar_{N,L-1}^{(m)}(t_i) \right]^T$ be the channel coefficient vector associated with the $i$ th OFDM symbol and stack the FFT demodulated signals at pilot clusters of the $i$ th OFDM symbol to form a vector

$$\mathbf{U}^{(m)}(i,k_1:k_\Gamma) = \left[ u^{(m)}(i,k_1),\cdots u^{(m)}(i,k_\tau),\cdots u^{(m)}(i,k_\Gamma) \right]^T. \tag{19}$$

The received signals at pilot clusters can be thus written as

$$\mathbf{U}^{(m)}(i,k_1:k_\Gamma) = \underbrace{\mathbf{A}(i)\mathbf{H}^{(m)}(i)}_{\text{desired signal for channel estimation}} + \underbrace{\mathbf{\Xi}^{(m)}(i,k_1:k_\Gamma)}_{\text{information interference on channel estimation}}$$

$$+ \ \overline{\mathbf{V}}^{(m)}(i,k_1:k_\Gamma) \tag{20}$$

where $\overline{\mathbf{V}}^{(m)}(i,k_1:k_\Gamma)$ is the noise vector in frequency-domain, $\Xi^{(m)}(i,k_1:k_\Gamma)=\left[\Xi^{(m)}(i,k_1),\cdots\Xi^{(m)}(i,k_\Gamma)\right]^T$ is the interference vector produced by the information sequences with $\Xi^{(m)}(i,k_\tau)=\sum_{n=1}^{N}H_n^{\prime(m)}(i,k_\tau)c_n(i,k_\tau)$, $\mathbf{A}(i)=$

$\left[\mathbf{A}(1,0),\cdots\mathbf{A}(1,L-1)\cdots\mathbf{A}(n,l)\cdots\mathbf{A}(N,0),\cdots\mathbf{A}(N,L-1)\right]$ is a $\Gamma\times NL$ matrix with the column-vectors

$$A(n,l)=\left[p_n(i,k_1)e^{-j2\pi k_1 l/B},\cdots p_n(i,k_\tau)e^{-j2\pi k_\tau l/B},\cdots p_n(i,k_\Gamma)e^{-j2\pi k_\Gamma l/B}\right]^T. \tag{21}$$

Since the matrix $\mathbf{A}(i)$ is known, when $\Gamma\geq NL$, the matrix $\mathbf{A}(i)$ is of full column rank, and the channel coefficient vectors can be thus estimated by

$$\hat{\mathbf{H}}^{(m)}(i)=\mathbf{A}^\dagger\mathbf{U}^{(m)}(i,k_1:k_\Gamma)$$
$$=\mathbf{H}^{(m)}(i)+\mathbf{A}(i)^\dagger\Xi^{(m)}(i,k_1:k_\Gamma)+\mathbf{A}(i)^\dagger\overline{\mathbf{V}}^{(m)}(i,k_1:k_\Gamma)\ m=1,\cdots M,i=1,\cdots I \tag{22}$$

where the superscript '$\dagger$' is the pseudo-inverse operation, and the hat '$\wedge$' indicates the estimation. From (22), the mainly computational effort is directly proportional to the unknown parameter number $NL$.

Using the specifically designed ST sequences in (15) and (or) (16), the problem of LTV channel estimation for MIMO/OFDM systems can be reduced into that of LTI channel. From (20) and (22), however, we notice that the interference vector due to information sequence can hardly be neglected since the power of data symbol is much larger than the pilot power. For conventional ST based schemes stated in [8]-[13] [17]-[18], first-order statistics are employed to suppress the information sequence interference over multiple training periods in the case that the channel is LTI during the record length. Such arithmetical average process, however, is no longer feasible to the channel assumed in this paper where the channel coefficients are linearly time-variant between consecutive OFDM symbols.

## 3.2 Channel estimation over multiple OFDM symbols

In this sub-section, a weighted average approach is developed to suppress the abovementioned information sequence interference over multiple OFDM symbols, and thus overcoming the shortcoming of the existing ST-based approach in estimating the time-variant channels.

By (22), the LTV channel coefficients can be obtained following the relationship $h_{n,l}^{(m)}(t_i)=\sum_{q=0}^{Q}h_{n,l,q}^{(m)}\eta_q(t_i)$. Taking the LTV channel coefficient estimation of each OFDM symbol $\hat{h}_{n,l}^{(m)}(t_i)\ i=1,\cdots I$ by (22) as a temporal result, and form a vector as $\hat{\mathbf{h}}_{n,l}^{(m)}=\left[\ \hat{h}_{n,l}^{(m)}(t_1),\cdots\hat{h}_{n,l}^{(m)}(t_I)\ \right]^T$, we thus have

$$\hat{\mathbf{h}}_{n,l}^{(m)} = \boldsymbol{\eta}\hat{\mathbf{h}}_{n,l,q}^{(m)}$$

$$= \begin{bmatrix} e^{j2\pi(0-Q/2)t_1/\Omega} & \cdots & e^{j2\pi(Q-Q/2)t_1/\Omega} \\ \vdots & \ddots & \vdots \\ e^{j2\pi(0-Q/2)t_I/\Omega} & \cdots & e^{j2\pi(Q-Q/2)t_I/\Omega} \end{bmatrix} \begin{bmatrix} \hat{h}_{n,l,0}^{(m)} \\ \vdots \\ \hat{h}_{n,l,Q}^{(m)} \end{bmatrix} \quad n = 1, \cdots N, \, l = 0, \cdots L-1 \tag{23}$$

where $\hat{\mathbf{h}}_{n,l,q}^{(m)} = \left[ \hat{h}_{n,l,0}^{(m)}, \cdots, \hat{h}_{n,l,q}^{(m)} \cdots \hat{h}_{n,l,Q}^{(m)} \right]^T$ is estimation of the complex exponential coefficients vector modeling the LTV channel, $\boldsymbol{\eta}$ is a $I \times (Q+1)$ matrix with $[\boldsymbol{\eta}]_{q,i} = e^{j2\pi(q-Q/2)t_i/\Omega}$. Thus, when $I \geq Q+1$, the matrix $\boldsymbol{\eta}$ is of full column rank, and the basis expansion model coefficients can be computed by

$$\hat{\mathbf{h}}_{n,l,q}^{(m)} = \boldsymbol{\eta}^\dagger \hat{\mathbf{h}}_{n,l}^{(m)} \quad n = 1, \cdots N, \, l = 0, \cdots L-1 . \tag{24}$$

Substituting $t_i = (i-1)B' + B/2$ into the matrix $\boldsymbol{\eta}$, we obtain the pseudo-inverse matrix as

$$\left[ \boldsymbol{\eta}^\dagger \right]_{i,q} = e^{-j2\pi(q-Q/2)((i-1)B'+B/2)/\Omega} \big/ I . \tag{25}$$

By (23)-(25), the modeling coefficients (6) can be computed by

$$\hat{h}_{n,l,q}^{(m)} = \sum_{i=1}^{I} e^{-j2\pi(q-Q/2)((i-1)B'+B/2)/\Omega} \hat{h}_{n,l}^{(m)}(i) \big/ I . \tag{26}$$

In fact, (26) is estimated over multiple OFDM symbols with a weighted average function of $e^{-j2\pi(q-Q/2)t_i/\Omega} \big/ I$.

  Compared with the conventional ST strategies, the proposed channel estimation is composed of two steps: First, with specially designed ST signals in (15) and (16), channel estimation can be reduced into that of LTI channel, and we are allowed to estimate the channel coefficients during each OFDM symbol as temporal results. Second, the temporal channel estimates are further enhanced over multiple OFDM symbols by using a weighted average procedure. That is, not only the target OFDM symbol, but also the OFDM symbols over the whole frame are invoked for channel estimation. Similar to the first-order statistics of LTI case [8]-[13] [17]-[18], it is thus anticipated that the weighted average estimation may also exhibit a considerable performance improvement for the LTV channels over a long frame $\Omega$.

## 4. Channel estimation analysis

In this section, we analyze the performance of the channel estimator proposed in Section III and derive a closed-form expression of the channel estimation variance which can be, in turn, used for ST power allocation. Before going further, we make the following assumptions:

(H1) The information sequence $\{c_n(i,k)\}$ is zero-mean, finite-alphabet, i.i.d., and equi-powered with the power $\sigma_c^2$.

(H2) The additive noise $\{v^{(m)}(i,t)\}$ is white, uncorrelated with $\{c_n(i,k)\}$, with $E\left[\left|v^{(m)}(i,t)\right|^2\right] = \sigma_v^2$.

(H3) The LTV channel coefficients $\mathbf{h}_{n,l}^{(m)}$ are complex Gaussian variables, and statistically independent for different values of $n$ and $l$.

From (22)-(26), the mean square error (MSE) of channel estimation is given by

$$MSE^{(m)} = E\left\{ \sum_{n=1}^{N}\sum_{l=0}^{L-1}\left\| h_{n,l}^{(m)}(i,t) - \sum_{q=0}^{Q}\hat{h}_{n,l,q}^{(m)}\eta_q(t)\right\|^2 \right\}$$

$$= E\left\{ \sum_{n=1}^{N}\sum_{l=0}^{L-1}\left\| h_{n,l}^{(m)}(i,t) - \sum_{q=0}^{Q}h_{n,l,q}^{(m)}\eta_q(t) + \sum_{q=0}^{Q}h_{n,l,q}^{(m)}\eta_q(t) - \sum_{q=0}^{Q}\hat{h}_{n,l,q}^{(m)}\eta_q(t)\right\|^2 \right\} \qquad (27)$$

where $\| \ \|$ is the Euclidean norm. In (27), the first error term $\sum_{n=1}^{N}\sum_{l=0}^{L-1}\left[ h_{n,l}^{(m)}(i,t) - \sum_{q=0}^{Q}h_{n,l,q}^{(m)}\eta_q(t)\right]$ is caused by the orthonormal basis expansion model in (6), which is referred to as the channel modeling error. The second error term $\sum_{n=1}^{N}\sum_{l=0}^{L-1}\sum_{q=0}^{Q}\left[ h_{n,l,q}^{(m)}\eta_q(t) - \hat{h}_{n,l,q}^{(m)}\eta_q(t)\right]$ is duo to the information interference to channel estimation (22) and additive noise. Explicitly, two error signals are mutually independent. Herein, we do not elaborate the topic of channel modeling error and focus on channel estimation error, which is mainly produced by the interference of information sequence. By (H2), the MSE of the estimation in one OFDM symbol can be written as

$$MSE^{(m)}(i) \stackrel{def}{=} \frac{1}{(Q+1)NL} E\left\{ \sum_{n=1}^{N}\sum_{l=0}^{L-1}\sum_{q=0}^{Q}\left\| h_{n,l,q}^{(m)}\eta_q(t) - \hat{h}_{n,l,q}^{(m)}\eta_q(t)\right\|^2 \right\}$$

$$= \underbrace{\frac{1}{(Q+1)NL}\text{tr}\left\{ \mathbf{A}(i)^{\dagger} E\left\{ \Xi^{(m)}(i,k_1:k_\Gamma)\left(\Xi^{(m)}(i,k_1:k_\Gamma)\right)^{H}\right\}\left(\mathbf{A}(i)^{\dagger}\right)^{H}\right\}}_{\text{estimation variance due to information sequence interference}} \qquad (28)$$

$$+ \underbrace{\frac{1}{(Q+1)NL}\text{tr}\left\{ \mathbf{A}(i)^{\dagger} E\left\{ \bar{V}^{(m)}(i,k_1:k_\Gamma)\left(\bar{V}^{(m)}(i,k_1:k_\Gamma)\right)^{H}\right\}\left(\mathbf{A}(i)^{\dagger}\right)^{H}\right\}}_{\text{estimation variance due to additive noise}}.$$

For zero-mean white noise, we have

$$E\left\{ \bar{V}^{(m)}(i,k_1:k_\Gamma)\left(\bar{V}^{(m)}(i,k_1:k_\Gamma)\right)^{H}\right\} = \sigma_v^2 I_\Gamma . \qquad (29)$$

Invoking the assumption (H1), information sequence interference $\Xi^{(m)}(i,k_\tau) = \sum_{n=1}^{N}H_n'^{(m)}(i,k_\tau)c_n(i,k_\tau), \tau = 1,\cdots\Gamma$ is approximately Gaussian distributed for a

large $\Gamma$. Therefore, the channel estimation variance due to information sequence interference can be obtained as

$$E\left\{\Xi^{(m)}(i,k_1:k_\Gamma)\left(\Xi^{(m)}(i,k_1:k_\Gamma)\right)^H\right\} = \frac{\sigma_c^2}{\Gamma^2}\sum_{\tau=0}^{\Gamma-1}\sum_{n=1}^{N}\sum_{l=0}^{L-1}\left|H'^{(m)}_n(i,k_\tau)\right|^2 = \frac{\rho_c}{\Gamma}\sum_{\tau=0}^{\Gamma-1}\sum_{n=1}^{N}\sum_{l=0}^{L-1}\left|\hbar_{n,l}^{(m)}e^{-2\pi k_\tau l/B}\right|^2. \quad (30)$$

Substituting (29) and (30) into (28), we have

$$MSE^{(m)}(i) \overset{def}{=} \frac{1}{(Q+1)NL}\left(\sigma_v^2 + \frac{\sigma_c^2}{\Gamma}\sum_{\tau=0}^{\Gamma-1}\sum_{n=1}^{N}\sum_{l=0}^{L-1}\left|\hbar_{n,l}^{(m)}e^{-2\pi k_\tau l/B}\right|^2\right)\mathrm{tr}\left[\left(\mathbf{A}(i)\right)^H\mathbf{A}(i)\right]^{-1}. \quad (31)$$

Apparently, the channel estimation performance depends crucially on the matrix $\mathbf{A}(i)$. The optimal estimation or minimum MSE (MMSE) estimation may require $\mathbf{A}(i)^H\mathbf{A}(i) = \Phi\mathbf{I}$ where $\Phi$ is a constant. From (21), we adopt the training sequence as $p_n(i,k) = \sigma_p, k = 0,\cdots B-1$ as (15), the above MMSE condition can be well satisfied. We thus have

$$\mathrm{tr}\left[\left(\mathbf{A}(i)\right)^H\mathbf{A}(i)\right] = \Gamma\sigma_p^2\mathbf{I}_{NL(Q+1)}. \quad (32)$$

Substituting (32) into (31), the MSE of channel estimation over one OFDM symbol can be derived as

$$MSE^{(m)}(i) = \frac{\sigma_c^2}{\Gamma^2\sigma_p^2}\sum_{\tau=0}^{\Gamma-1}\sum_{n=1}^{N}\sum_{l=0}^{L-1}\left|\hbar_{n,l}^{(m)}e^{-2\pi k_\tau l/B}\right|^2 + \frac{\sigma_v^2}{\Gamma\sigma_p^2}. \quad (33)$$

It is seen that the first term of (33) is the estimation variance due to information interference, and depends upon the channel transfer functions. We thus define the normalized variance as

$$NMSE^{(m)}(i) = \frac{\sigma_c^2}{\Gamma^2\sigma_p^2}\sum_{\tau=0}^{\Gamma-1}\sum_{n=1}^{N}\sum_{l=0}^{L-1}\left|\hbar_{n,l}^{(m)}e^{-2\pi k_\tau l/B}\right|^2 \Bigg/ \left|\bar{\hbar}^{(m)}(i)\right|^2 \quad (34)$$

where $\left|\bar{\hbar}^{(m)}(i)\right|^2 = \sum_{\tau=0}^{\Gamma-1}\sum_{n=1}^{N}\sum_{l=0}^{L-1}\left|\hbar_{n,l}^{(m)}e^{-2\pi k_\tau l/B}\right|^2 \Big/ NL\Gamma$. Following the definition of (34), we obtain the normalized variance as

$$NMSE^{(m)}(i) = \frac{\sigma_c^2}{\Gamma^2\sigma_p^2}\sum_{\tau=0}^{\Gamma-1}\sum_{n=1}^{N}\sum_{l=0}^{L-1}\left|\hbar_{n,l}^{(m)}e^{-2\pi k_\tau l/B}\right|^2 \Bigg/ \left|\bar{\hbar}^{(m)}(i)\right|^2 = \frac{NL}{\Gamma}\frac{\sigma_c^2}{\sigma_p^2}. \quad (35)$$

From (35), we can find that the estimation variance due to the information interference is directly proportional to the information-to-pilot power ratio $\sigma_c^2/\sigma_p^2$, thereby resulting in an inaccurate solution for the general case that $\sigma_c^2 >> \sigma_p^2$.

Then, we analyze the channel estimation performance of the weighted average approach over multiple OFDM symbols (the whole frame $\Omega$). Define the vectors $\mathrm{A} = \left[\mathbf{A}(1),...\mathbf{A}(I)\right]^T$,

$\Xi^{(m)} = \left[\Xi^{(m)}(1, k_1 : k_\Gamma), \cdots \Xi^{(m)}(I, k_1 : k_\Gamma)\right]^T$ and $\quad V^{(m)} = \left[\bar{V}^{(m)}(1, k_1 : k_\Gamma), \ldots \bar{V}^{(m)}(I, k_1 : k_\Gamma)\right]^T$,

the MSE of the weighted average channel estimator over multiple OFDM symbols is given by

$$MSE^{(m)} = \frac{1}{(Q+1)NL} \mathrm{tr}\left\{\boldsymbol{\eta}^\dagger E\left\{\left\|\mathbf{A}^\dagger \Xi^{(m)} + \mathbf{A}^\dagger V^{(m)}\right\|^2\right\}\left(\boldsymbol{\eta}^\dagger\right)^H\right\} = \frac{1}{I}\sum_{i=1}^{I} MSE^{(m)}(i)\, \mathrm{tr}\left\{\boldsymbol{\eta}^\dagger\left(\boldsymbol{\eta}^\dagger\right)^H\right\}. \tag{36}$$

Note that the column vectors of the matrix $\boldsymbol{\eta}$ in (23) are in fact the FFT vectors of a $I \times I$ matrix, we thus have $\boldsymbol{\eta}^H\boldsymbol{\eta} = I I_{(Q+1)}$ and $\mathrm{tr}\left[\boldsymbol{\eta}^H\boldsymbol{\eta}\right]^{-1} = (Q+1)/I$. Substituting (33) into (36), the MSE of channel estimation over multiple OFDM symbols is given by

$$MSE^{(m)} = \frac{(Q+1)\sigma_c^2}{I\Gamma^2\sigma_p^2}\sum_{\tau=0}^{\Gamma-1}\sum_{n=1}^{N}\sum_{l=0}^{L-1}\left|\hbar_{n,l}^{(m)} e^{-2\pi k_\tau l/B}\right|^2 + \frac{(Q+1)\sigma_v^2}{I\Gamma\sigma_p^2} \tag{37}$$

In (37), the second term is caused by information sequence interference, which may become the dominant component of the channel estimation variance for the general case of $\sigma_c^2 \gg \sigma_p^2$, especially for large SNRs. Therefore, we solely consider information sequence effect. Similar to (34)-(35), we derive the normalized variance due to information interference by removing the channel gain as

$$NMSE^{(m)} = \frac{(Q+1)\sigma_c^2}{I\Gamma^2\sigma_p^2}\sum_{\tau=0}^{\Gamma-1}\sum_{n=1}^{N}\sum_{l=0}^{L-1}\left|\hbar_{n,l}^{(m)} e^{-2\pi k_\tau l/B}\right|^2 \bigg/ \left|\bar{h}^{(m)}\right|^2 \tag{38}$$

where $\left|\bar{h}^{(m)}\right|^2 = \sum_{i=1}^{I}\left|\bar{h}^{(m)}(i)\right|^2 \big/ I$. It follows that

$$NMSE^{(m)} = \frac{\sigma_c^2}{\sigma_p^2}\frac{NL(Q+1)}{I\Gamma} \approx \frac{\sigma_c^2}{\sigma_p^2}\frac{NL(Q+1)}{\Omega}\frac{B}{\Gamma} = \frac{\sigma_c^2}{\sigma_p^2}\frac{NL(Q+1)}{\theta\Omega} \tag{39}$$

where $\theta = \Gamma/B$ is the training ratio of one OFDM symbol. For conventional ST-based LTI schemes where isolated pilots are exploited for channel estimation [8]-[13] [17]-[18], we have $\theta = 1$. However, for estimating the LTV channels addressed in this paper, $\Gamma$ pilot clusters, instead of isolated pilot tones, are exploited. Thus, the corresponding training ratio yields $\theta \leq 1/(2T+1)$. From (39), the normalized variance is directly proportional to the information-pilot power ratio $\sigma_c^2/\sigma_p^2$, the training ratio $\theta$ and the ratio of unknown parameter number $NL(Q+1)$ over the frame length $\Omega$.

Compared with the variances of channel estimation over one OFDM symbol as in (33)-(35), the estimation variances of the weighted average estimator(37)-(39) is significantly reduced owing to the fact that $I/(Q+1) \gg 1$. Theoretically, the weighted average operation can be considered as an effective approach in estimating LTV channel, where the information sequence interference can be effectively suppressed over multiple OFDM symbols. As stated

in conventional ST-based LTI schemes [8]-[13] [17]-[18], channel estimation performance can be improved along with the increment of the recorded frame length $\Omega$, i.e. the estimation variance approaches to zero as $\Omega \to \infty$. This can be easily comprehended that larger frame length $\Omega$ means more observation samples, and hence lowers the MSE level. From the LTV channel model (6), however, we note that as the frame length $\Omega$ is increased, the corresponding truncated DFB requires a larger order $Q$ to model the LTV channel (maintain a tight channel model), and the least order should be satisfied $Q/2 \geq f_d \Omega / f_s$, where $f_d$ and $f_s$ are the Doppler frequency and sampling rate, respectively. Consequently, as the frame length $\Omega$ increases, the LTV channel estimation variance (39) approaches to a fixed lower-bound associate with the system Doppler frequency as well as the information to pilot power ratio. This is quite different from the existing ST-based channel estimation approaches [8]-[19].

According to the theoretic analysis in (37)-(39), the proposed two-step LTV channel estimator achieves a significant improvement over multiple OFDM symbols compared with that of block-by-block process (33)-(35). However, as the frame length $\Omega$ is increased, the estimation variances approach to a fixed lower-bound. Further enhancement of the channel estimation should resort to increasing the ST power $\sigma_p^2$. For wireless communication systems with a limited transmission power, however, an increased ST power allocation reduces the data power $\sigma_c^2$, leading to SER degradation. Accordingly, in the analysis presented in the next section, the ratio of ST power allocation is determined by maximizing the lower bound of the average channel capacity.

## 5. Analysis of ST power allocation and system capacity

In this section, we consider the issue of ST power allocation where the lower bound of the average channel capacity is maximized and then mathematically derived for the proposed two-step channel estimator.

Define the ST power allocation factor

$$\beta = \frac{E\left[\left|p_n(k)\right|^2\right]}{E\left[\left|p_n(k)\right|^2\right] + E\left[\left|c_n(k)\right|^2\right]} = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_c^2}. \tag{40}$$

For a fixed SNR or transmitted power budget, higher $\beta$ implies smaller effective SNR at the receiver due to decreased power in the information sequence but higher channel estimation accuracy. Having removed ST sequence, we obtain the received signals in a vector-form as

$$\overline{U}^{(m)}(i) = \left[\overline{u}^{(m)}(i,0), \cdots \overline{u}^{(m)}(i,k), \cdots \overline{u}^{(m)}(i,B-1)\right]^T$$

$$= \underbrace{\sum_{n=1}^{N} \hat{H}_n'^{(m)}(i) C_n(i)}_{\text{desired signals: } = \lambda^{(m)}(i)} + \underbrace{\sum_{n=1}^{N} \Delta\hat{H}_n'^{(m)}(i)\left[P_n(i) + C_n(i)\right]}_{\text{interference to information signal recovery: } = \mu^{(m)}(i)} + \overline{V}^{(m)}(i) \tag{41}$$

with the received signals $\overline{u}^{(m)}(i,k), k=0, \cdots B-1$ in (41) as

$$\bar{u}^{(m)}(i,k) = \lambda^{(m)}(i,k) + \mu^{(m)}(i,k) + \bar{v}^{(m)}(i,k)$$

$$= \sum_{n=1}^{N} \hat{H}_n'^{(m)}(i,k) c_n(i,k) + \sum_{n=1}^{N} \Delta\hat{H}_n'^{(m)}(i,k) \left[ p_n(i,k) + c_n(i,k) \right] + \bar{v}^{(m)}(i,k) \qquad (42)$$

where $\Delta\hat{H}_n'^{(m)}(i,k) = \sum_{n=1}^{N} \left[ H_n'^{(m)}(i,k) - \hat{H}_n'^{(m)}(i,k) \right]$ is the estimation error due to information interference as well as additive noise. Using the proposed two-step estimator (23)-(26), the channel estimation variance can be smoothed over multiple OFDM symbols, and approaches to a small fixed lower bound. The estimated vector $\hat{H}_n'^{(m)}(i)$ as well as the error vector $\Delta\hat{H}_n'^{(m)}(i)$, therefore, can be thus approximated to be of the similar characteristics of distribution as that of $H_n'^{(m)}(i)$. Consequently, following the assumption (H1)-(H3), the interference vector $\mathbf{\mu}^{(m)}(i)$ is approximately white for a large symbol-size *B*, and independent of the noise vector $\bar{V}^{(m)}(i)$. Similar to the procedure of (29)-(30), the covariance matrix of $\mathbf{\mu}^{(m)}(i)$ and $\lambda^{(m)}(i)$ can be obtained as

$$\text{var}\left(\lambda^{(m)}(i)\right) = E\left\{ \left(\lambda^{(m)}(i)\right)^H \lambda^{(m)}(i) \right\} = \sigma_{\hat{H}}^2 \sigma_p^2 I \qquad (43)$$

$$\text{var}\left(\mathbf{\mu}^{(m)}(i)\right) = E\left\{ \left(\mathbf{\mu}^{(m)}(i)\right)^H \mathbf{\mu}^{(m)}(i) \right\} = \sigma_{\Delta\hat{H}}^2 \left( \sigma_p^2 + \sigma_c^2 \right) I \qquad (44)$$

where $\sigma_{\hat{H}}^2 = \left| \bar{h}^{(m)}(i) \right|^2 = \sum_{k=0}^{B-1} \sum_{n=1}^{N} \sum_{l=0}^{L-1} \left| \hat{h}_{n,l}^{(m)} e^{-2\pi kl/B} \right|^2 \Big/ NLB$ ,

and $\sigma_{\Delta\hat{H}}^2 = \sum_{k=0}^{B-1} \sum_{n=1}^{N} \sum_{l=0}^{L-1} \left| \Delta\hat{h}_{n,l}^{(m)} e^{-2\pi kl/B} \right|^2 \Big/ NLB$ . Since the ST power allocation factor is derived within each isolated OFDM symbol, we neglect the symbol-index *i* for simplicity. A lower bound on the OFDM channel capacity with channel estimation error has been derived in [20]-[21] for uniform pilot distribution. Such expression can readily be extended to issue of ST where the pilots are spread over the whole frequency band. Therefore, the lower bound of the average channel capacity for an ST-based OFDM system can be obtained by summing over all the subcarriers, i.e.,

$$C^{(m)} \geq \bar{C}^{(m)} = \frac{1}{B} \sum_{k=0}^{B-1} E\left\{ \log\left[ 1 + \frac{\sigma_c^2}{\left(\sigma_p^2 + \sigma_c^2\right) \sigma_{\Delta\hat{H}}^2 \Big/ \sigma_{\hat{H}}^2 + \sigma_v^2 \Big/ \sigma_{\hat{H}}^2} \right] \right\} \qquad (45)$$

For the sake of simplicity, we assume the transmission power satisfies that $\sigma_p^2 + \sigma_c^2 = 1$. By (40), we thus have $\sigma_p^2 = \beta$ and $\sigma_c^2 = 1 - \beta$. Considering that the normalized MSE of the proposed two-step channel estimator is sufficiently small and approaches to a fixed lower bound (37)-(39), it allows us to make the approximation of $\sigma_{\Delta\hat{H}}^2 \Big/ \sigma_{\hat{H}}^2 \approx \sigma_{\Delta\hat{H}}^2 \Big/ \sigma_{H}^2 = NMSE^{(m)}$ . As a result, $C^{(m)}$ in (45) can be approximated as

$$C^{(m)} \approx \frac{1}{B}\sum_{k=0}^{B-1} E\left\{ \log\left[ 1 + \frac{\sigma_c^2}{\sigma_{\Delta\hat{H}}^2/\sigma_H^2 + \sigma_v^2/\sigma_H^2} \right] \right\}$$

$$= \frac{1}{B}\sum_{k=0}^{B-1} E\left\{ \log\left[ 1 + \frac{1-\beta}{(1-\beta)(Q+1)NL/\beta I\Gamma + (Q+1)\sigma_v^2/\beta I\Gamma\sigma_H^2 + \sigma_v^2/\sigma_H^2} \right] \right\} \quad (46)$$

$$= \log\left( 1 + \frac{(1-\beta)\beta I\Gamma}{\beta[ I\Gamma/\Re_{\mathrm{SNR}} - (Q+1)NL ] + (Q+1)NL(1/\Re_{\mathrm{SNR}}+1)} \right)$$

where $\Re_{\mathrm{SNR}} = \sigma_H^2\left(\sigma_p^2 + \sigma_c^2\right)\big/\sigma_v^2 = \sigma_H^2/\sigma_v^2$. In fact, the averaged channel capacity of (46) is a log-function of $\beta$, which is a monotonically increasing function. Therefore, the lower-bound of $C^{(m)}$ with respect to $\beta$ can be achieved by maximizing the following function

$$\Upsilon^{(m)}(\beta) = \frac{\beta(1-\beta)}{\alpha_1\beta + \alpha_2} = \frac{(1-\beta)\beta}{\beta[ 1/\Re_{\mathrm{SNR}} - (Q+1)NL/I\Gamma ] + (Q+1)NL(1/\Re_{\mathrm{SNR}}+1)/I\Gamma} . \quad (47)$$

where

$$\alpha_1 = 1/\Re_{\mathrm{SNR}} - (Q+1)NL/I\Gamma , \quad \alpha_2 = (Q+1)NL(1/\Re_{\mathrm{SNR}}+1)\big/I\Gamma . \quad (48)$$

Setting the first derivation of $\Upsilon^{(m)}(\beta)$ with respect to $\beta$ to zero, we obtain (after some manipulations) a quadratic equation in $\beta$, i.e.

$$\beta^2 + \frac{2\alpha_2}{\alpha_1}\beta - \frac{\alpha_2}{\alpha_1} = 0 . \quad (49)$$

Consequently, the global maximum of $\Upsilon^{(m)}(\beta)$ can be obtained when

$$\beta = \frac{\sqrt{\left(1/\Re_{\mathrm{SNR}}+1\right)\left(I\Gamma/NL(Q+1)\Re_{\mathrm{SNR}}+1/\Re_{\mathrm{SNR}}\right)} - (Q+1)NL(1/\Re_{\mathrm{SNR}}+1)}{I\Gamma/\Re_{\mathrm{SNR}} - (Q+1)NL} . \quad (50)$$

As will be shown in simulations, an increase in the training power allocation factor $\beta$ does not necessarily improve the overall system performance since a larger $\beta$ implies a better channel estimation while substantially scarifying the effective received signal SNR at the same time.

## 6. Simulations

We assume the MIMO/OFDM system with $N = 2$ and $M = 4$. The symbol-size is $B = 1024$ and the transmitted data $s_n(i,k)$ is 8-PSK signals with symbol rate $f_s = 10^7$/second. Before transmission, the transmitted data are coded by $1/2$ convolutional coding and block interleaving over one OFDM symbol. The channel is assumed to be $L = 10$ taps and, the

coefficients $h_{n,l}^{(m)}(t)$ are generated as low-pass, Gaussian and zero mean random processes and uncorrelated for different values of $n$ and $l$. The multi-path intensity profile is chosen to be $\phi(l) = \exp(-l/10)$ $l = 0, \cdots L-1$. The Doppler spectra are $\Psi(f) = (\pi\sqrt{(f_n)^2 - f})$ for $f \leq f_n$, where $f_n$ is the Doppler frequency of the $n$th user, otherwise, $\Psi(f) = 0$. CP-length is chosen to be 32 to avoid inter-symbol interferences. The additive noise is a Gaussian and white random process with a zero mean.

*Test Case 1. Channel Estimation*

We run simulations with the Doppler frequency $f_n = 300\text{Hz}$ that corresponds to the maximum mobility speed of 162 km/h as the users operate at carrier frequency of 2GHz. In order to model the LTV channel, the frame is designed as $\Omega = B' \times 128 = (B + \text{CP-length}) \times 128 = 135168$, i.e. each frame consists of 128 OFDM symbols. During the frame, the channel variation is $f_n\Omega/f_s = 4.1$. Over the frame $\Omega$, we utilize truncated DFB of order $Q = 10 > 2f_d\Omega/f_s$ to model the LTV channel coefficients. In order to estimate the MIMO/OFDM channels, the superimposed pilots are designed according to (15) with the pilot power $\sigma_p^2 = 0.2\sigma_c^2$. Fig.2 depicts the LTV channel coefficient estimation over the frame $\Omega$. It is clearly observed that although the channel coefficient is accurately estimated during the centre part of the frame, the outmost samples over the whole frame still exhibit errors. A possible explanation is that as the Fourier basis expansions in (6) are truncated, and an effect similar to the Gibbs phenomenon, together with spectral leakages, will lead to some errors at the beginning and the end of the frame. This may be a common problem for the proceeding literature [1]-[2] [5]-[6] [16] that employing basis expansions to model the LTV channels. To solve the problem, the frames are designed to be partially overlapped, e.g. the frames are designed as $(\ell-1)\Omega - \Psi B' \leq t \leq \ell\Omega$, $\ell = 2,3,\cdots$, where $\Psi$ is a positive integer. By the frame-overlap, the channel at the beginning and the end of one frame can be modeled and estimated from the neighboring frames.

To further evaluate the new channel estimator, we use the mean square errors to measure the channel estimation performance by

$$MSE_n^{(m)} = \sum_{i=1}^{\Omega/B'} MSE_n^{(m)}(i)/(\Omega/B') =$$

$$\frac{B'}{\Omega}\sum_{i=1}^{\Omega/B'} E\left\{\sum_{t=0}^{B-1}\sum_{l=0}^{L-1}\left|h_{n,l}^{(m)}(i,t) - \sum_{q=0}^{Q}\hat{h}_{n,l,q}^{(m)}\text{e}^{j2\pi(q-Q/2)t/\Omega}\right|^2 \bigg/ BL\left|h_{n,l}^{(m)}(i,t)\right|^2\right\} \tag{51}$$

where $\hat{h}_{n,l,q}^{(m)}$ is the channel coefficient estimation.

We firstly test the two-step channel estimator under the different pilot powers and different channel coefficient numbers to verify the channel estimation variance analysis. The LTV channel is the same as that in Fig.2. As shown in Fig.3, the MSE of the channel estimation approach are almost independent of the additive noises, especially as SNR>5dB. This is consistent with the channel estimation analysis (38)-(39) where the additive noise has been

greatly suppressed by the weighted average procedure. Thus, the estimation errors depend mainly on the information- pilot power ratio as well as the system unknowns $NL$. This is rather different from the FDM training based schemes [20]-[23].
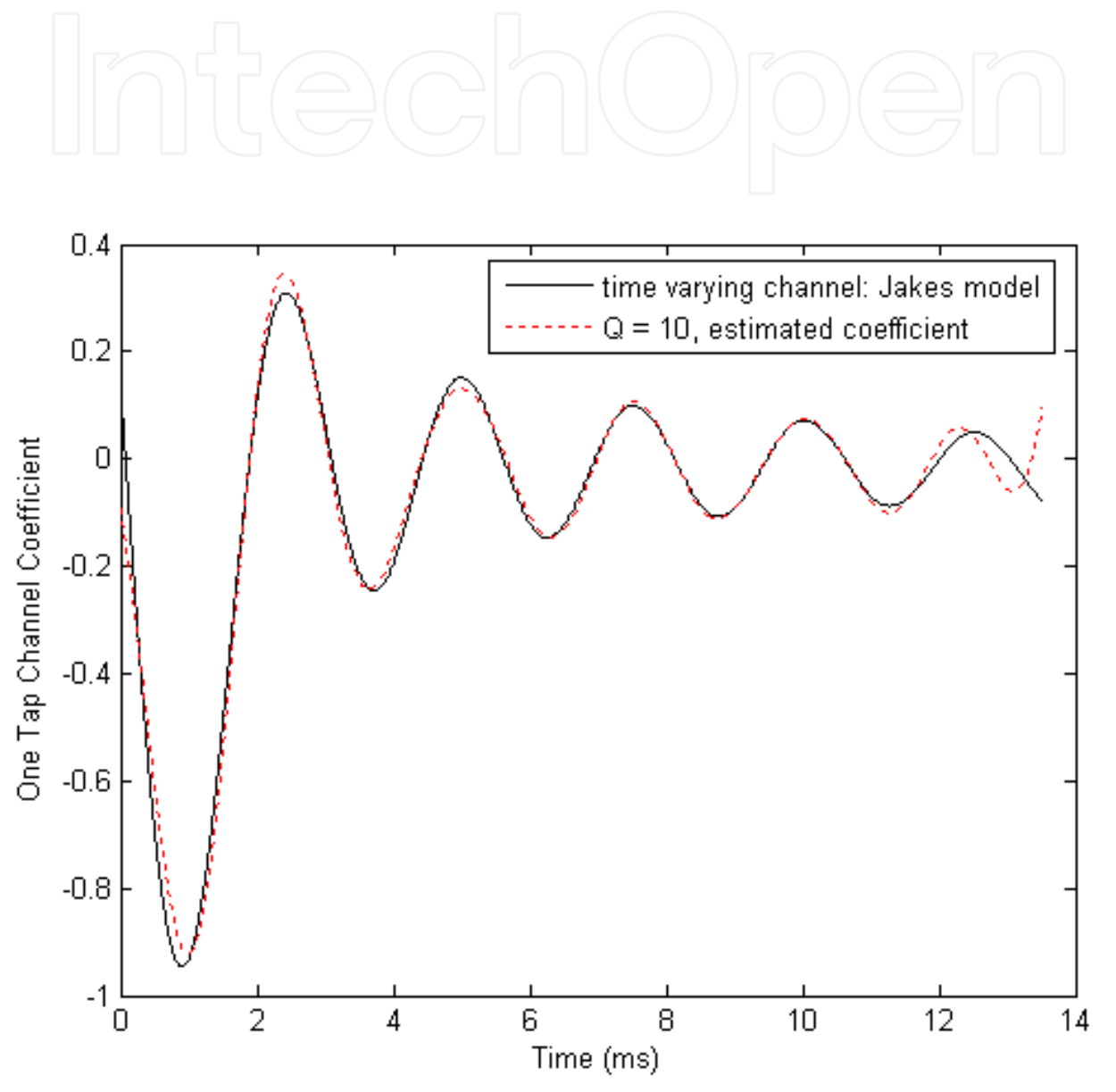


Fig. 2. One tap coefficient of the LTV channel and the estimation over the frame $\Omega = 135168/10^7 \approx 13.52$ ms.
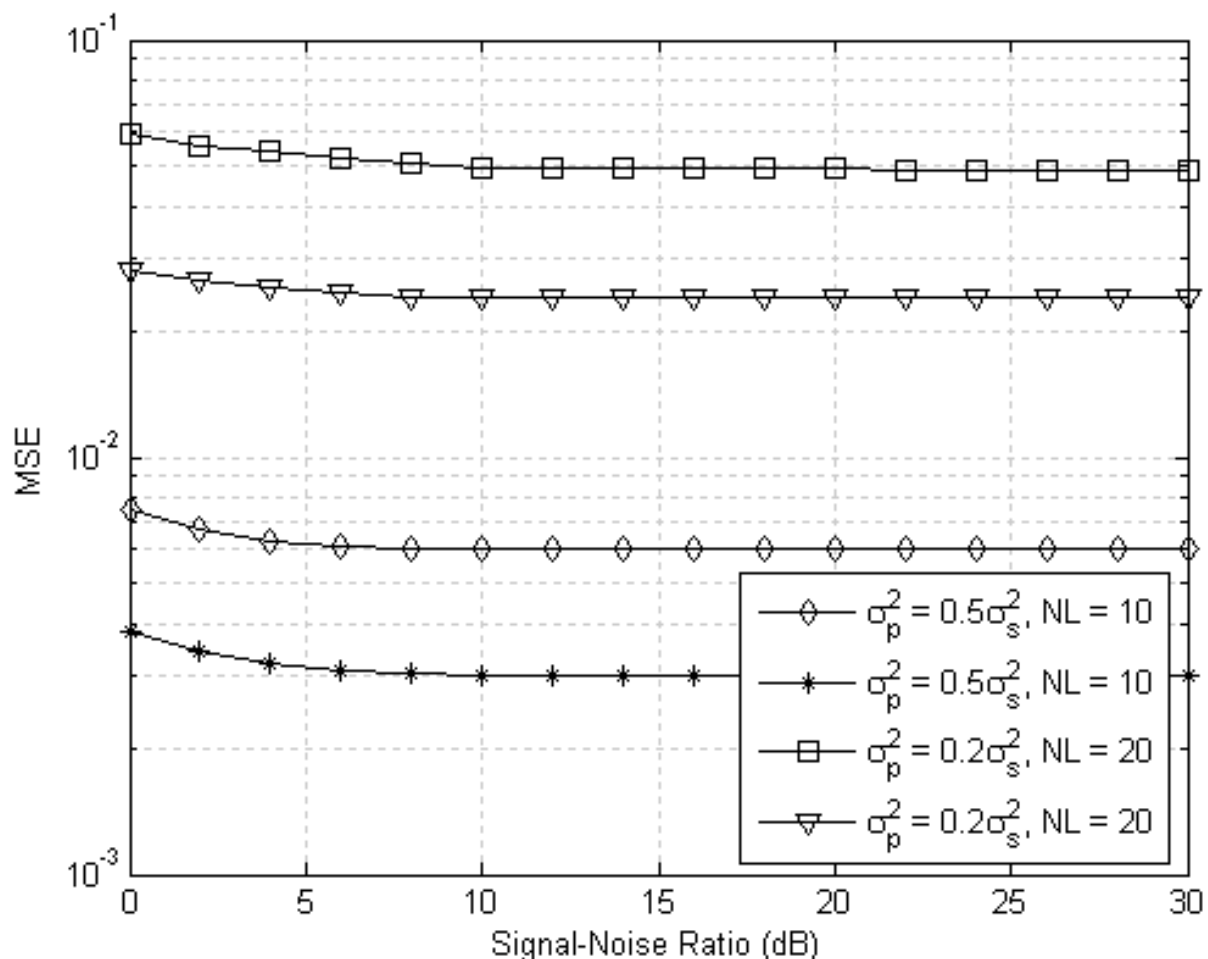
Fig. 3. MSE of the weighted average estimation versus SNR for the LTV channel
of $f_n = 300$Hz and $\Omega = 13.52$ ms under the different pilot powers and different unknown
parameters.

We then compare the proposed two-step channel estimation scheme with the conventional
ST-based methods [8]-[13] [17]-[18] under different Doppler frequencies. In the conventional
ST scheme, the LTV channel is firstly estimated from the LTI assumption at each OFDM
symbol, and then all the estimations from the frame $\Omega$ are averaged to confront the
information sequence interferences. It shows clearly in Fig. 4 that for the LTI channel
of $f_n = 0$Hz, both the conventional ST and the weighted average estimator exhibit the similar
performance. In addition, the estimation performance can be improved with the increment
of the frame or average length. However, when the channel involved in simulations is time-
varying, the channel estimation performance of the conventional ST-based schemes is
degraded extensively. The simulation reveals the shortcoming of the conventional ST in
estimating the LTV channels. On the contrary, the MSE level is reduced by the weighted
average process (23)-(26) for the LTV channels of $f_n = 100$Hz, 300Hz with $T = 2$ (one pilot
cluster is composed of $2T + 1 = 5$ pilots). We also observe that the MSE approaches to a
constant as the increment of the frame length, i.e. the lower-bound that associated with the
given Doppler frequency.
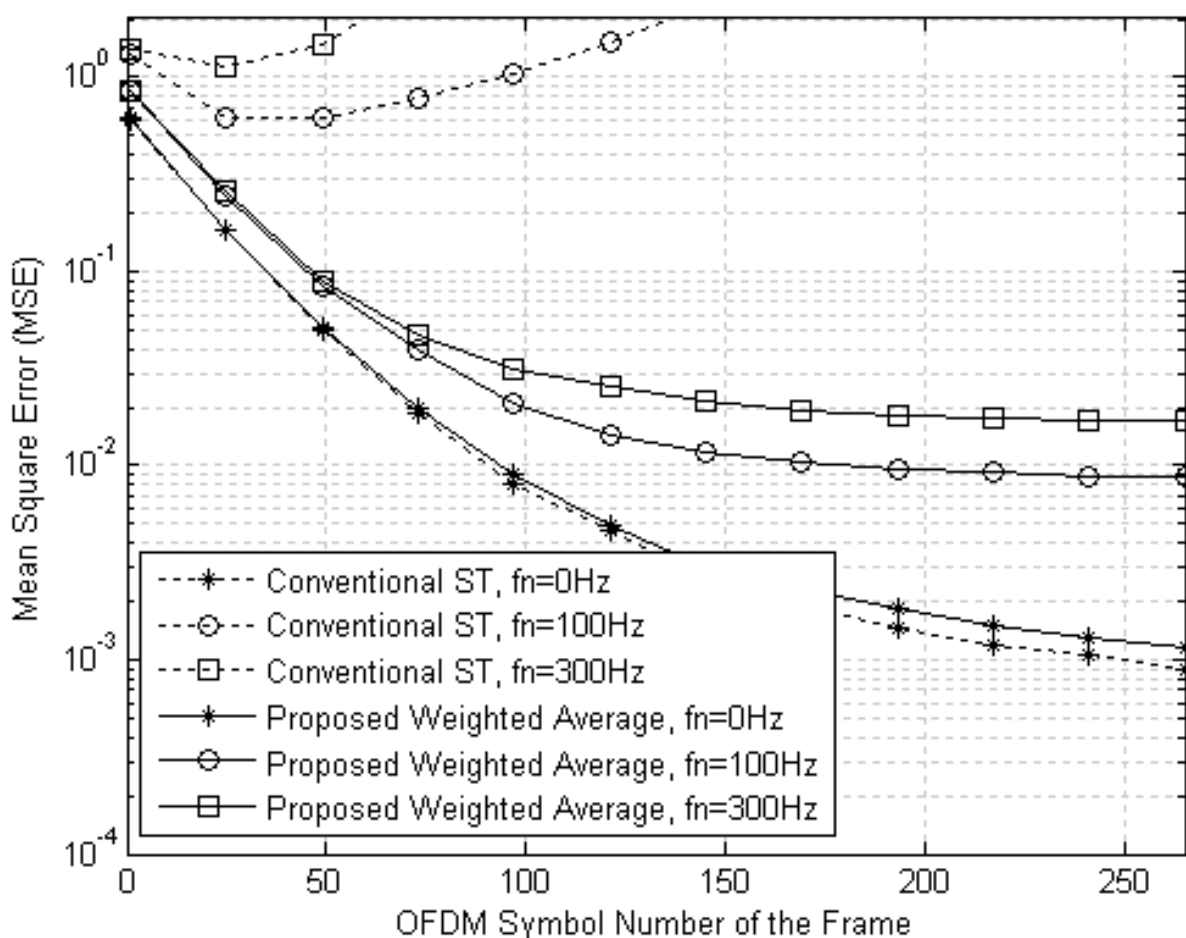
Fig. 4. MSE versus frame or average length under the different Doppler frequencies of the
LTV channel with $\sigma_p^2 = 0.25\sigma_c^2$, SNR = 20dB.

From Fig. 4, we observe that channel estimation performance would be degraded as the
increment of mobile users' speed (or corresponding system Doppler shift). To further
enhance the channel estimation performance of the systems with a limited pilot power while
suffering from a high Doppler shift, an iterative decision feedback (DF) approach can be
adopted at the receiver. Explicitly, the iterative method can be considered as a twofold
process. First, the information sequences are recovered by a hard detector [5] based on the
LTV channel estimation in Section III. Second, the recovered data symbols are removed from
the received signals to cancel the information sequence interference and, thus to enhance the
channel estimation performance. Fig. 5 depicts the performance between the weighted
average scheme and the iterative DF estimator in terms of channel MSE. For a fairness of
comparison, we also simulate the MSE of the FDM training-based channel estimator [5] as
latter serves as a "benchmark" in related works. For estimating the MIMO/OFDM channels,
$\Gamma = 40$ pilot clusters with $\Gamma(2T+1) = 200$ known pilot symbols which are subject to the

proposed pilot specifications in (15) are used in one OFDM symbol. That is, approximately 10% total bandwidth is assigned for pilot tones. Comparatively, as shown in Fig.5, the iterative DF estimation exhibits a more significant improvement than that of weighted average estimation, and outperforms the FDM channel estimator [5] by using a small pilot power of $\sigma_p^2 = 0.25\sigma_c^2$, which conforms that the information sequence interferences can be effectively cancelled by iterative DF procedure. Moreover, it should be noted that since the superimposed pilots are spread over the entire band, the proposed ST-based channel estimator is also feasible to estimate the channel with a very long delay spread, i.e. cluster-based channel.
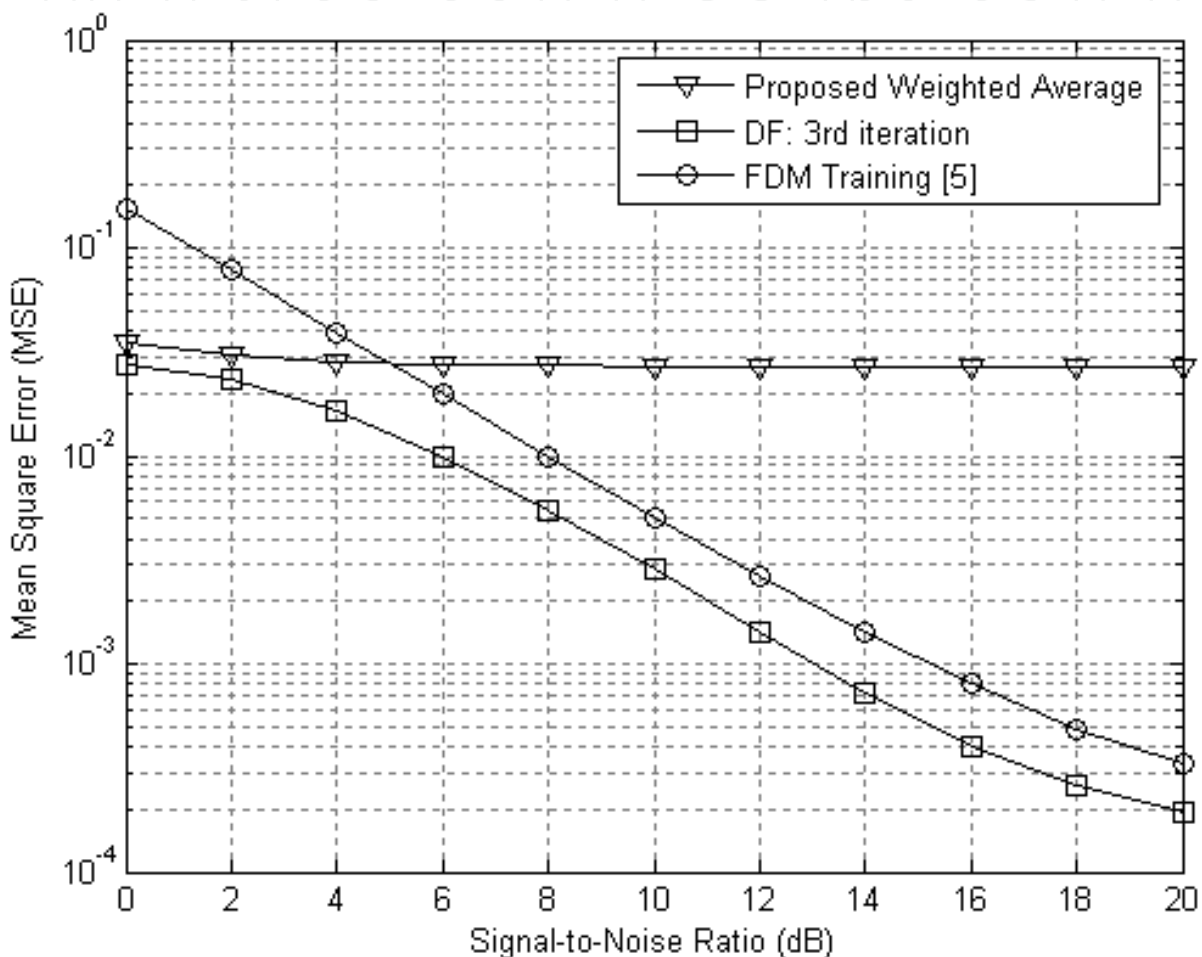


Fig. 5. MSE versus SNR for different estimators for the LTV channel of $f_n = 300\,\mathrm{Hz}$, $NL = 20$.

To further validate the effectiveness of the DF scheme, we also provide the channel estimation MSE of the DF method versus the iteration numbers under SNR = 15dB. Fig. 6 shows that the iterative DF method is feasible for a wide range of system Doppler spreads. Obviously, the enhancement of the iterative DF is at the cost of an increment in computational complexity that is directly proportional to the iteration number. However, as is shown in Fig. 6 that the iterative DF approach converges to the steady-state performance by only a few iterations, the overall computational complexity will be acceptable for many wireless communication systems.
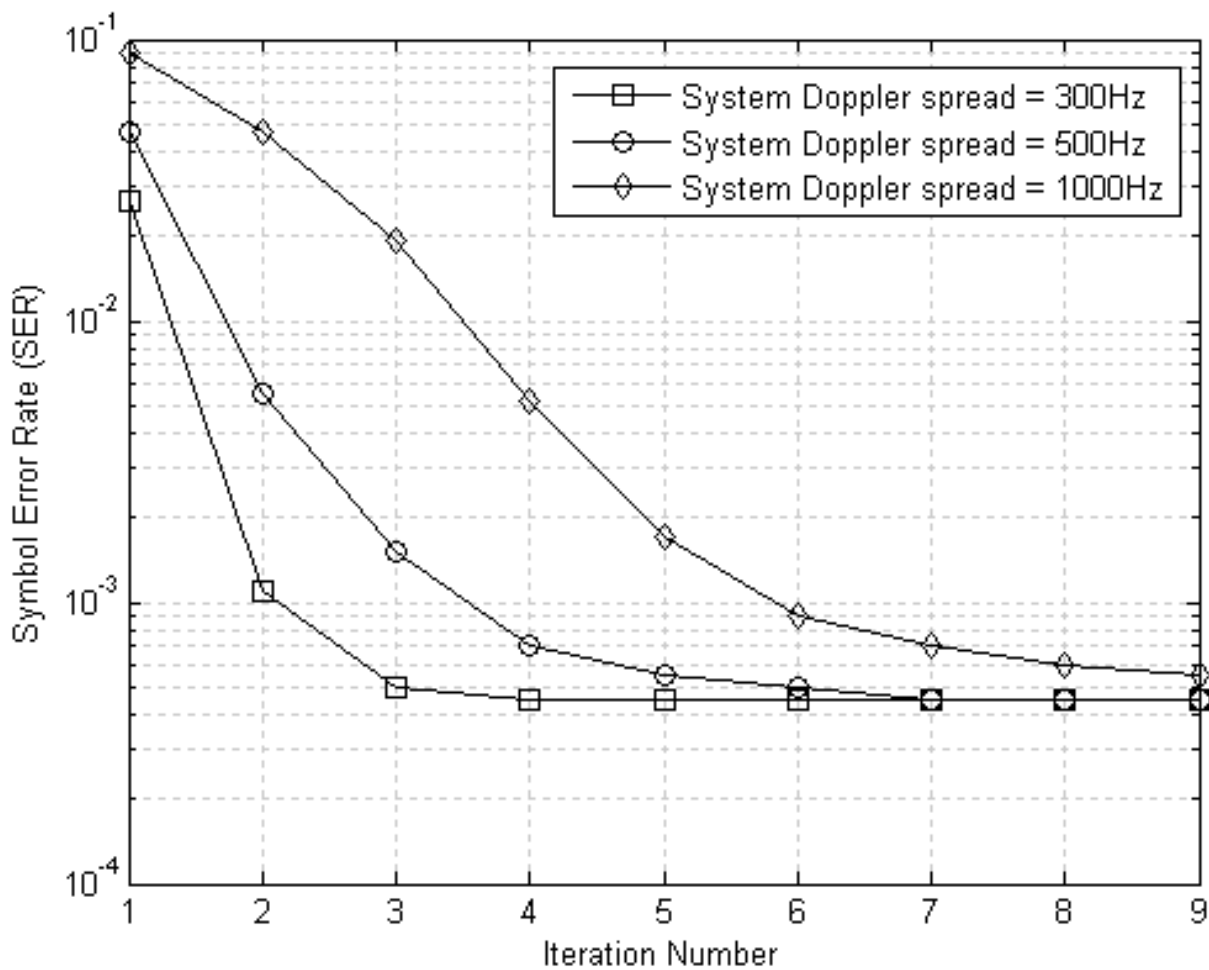
Fig. 6. MSE of the iterative DF channel estimation versus iterations for the LTV channel of different system Doppler spreads when SNR=15dB.

*Test Case 2. Training Power Allocation*

As aforementioned, for wireless communication systems with a limited transmission power, some useful power must inevitably be allocated to the superimposed pilots, and thus resulting in the received signal SNR reduction. Herein, we carry out several experiments to assess the effect of ST power allocation factor on the lower-bound of the average channel capacity for different SNRs.

Fig. 7 shows the effect of different value of training power allocation factor $\beta$ on the lower bound of the average channel capacity for received signal SNR = 10 and 20 dB, respectively. It is seen that the average channel capacity decreases with the increment of $\beta$. It reveals that although higher $\beta$ implies that higher fraction of transmitted power is allocated to training leading to more accurate channel estimates, the received signal SNR is substantially decreased, resulting in potential decrement of the average channel capacity. In addition, we further simulate the approximated $\beta$ in order to test the validity of theoretic results in (50). It can be seen that the approximation of $\beta$ is almost consistent with that of the actual results.
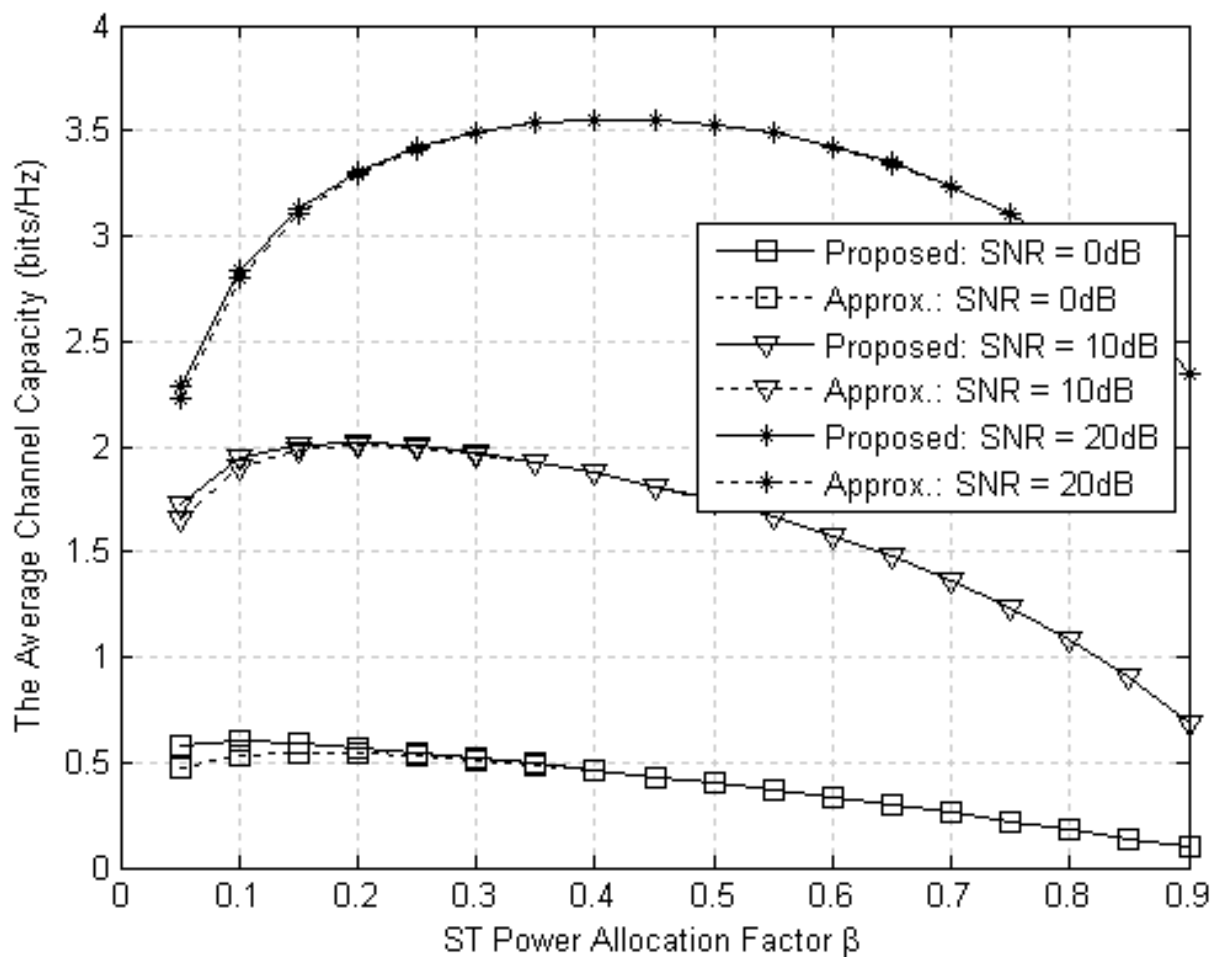
Fig. 7. Lower bound of the average channel capacity versus different values of ST power allocation factor $\beta$ under SNR = 0dB, 10dB and 20dB, respectively.

Fig. 8 shows the plots of the optimal value of training power allocation factor $\beta$ versus received SNR for different frame length. It is observed that the increment of SNR leading to a corresponding increase in the optimal ST power allocation factor. This can be easily comprehended that according to (41), the effective interference is composed of two factors, i.e. the bias of channel estimation and the additive noise. That is, for large SNRs, higher $\beta$ is required to improve the channel estimation performance, thus leading to a reduction of the effective interference. Conversely, when SNR is small, improving the channel estimation accuracy has a small effect in reducing the effective interference. On the other hand, we notice that $\beta$ decreases as the frame length increases but approximately unvaried when $\Omega$ is sufficiently large, i.e. $\beta$ is almost unchanged when $I \geq 192$. This result arises because we have theoretically analyzed in Section III that the estimation variance approaches to a fixed lower bound that can be only improved by increasing ST power allocation when the frame length is large enough. Therefore, the power allocated to the training sequence can be reduced with no loss in channel estimation performance when the frame length is increased, but finally approaches to a fixed lower bound associate with the channel estimation variance when $\Omega$ is sufficiently large. This is somewhat different from those presented in [10].
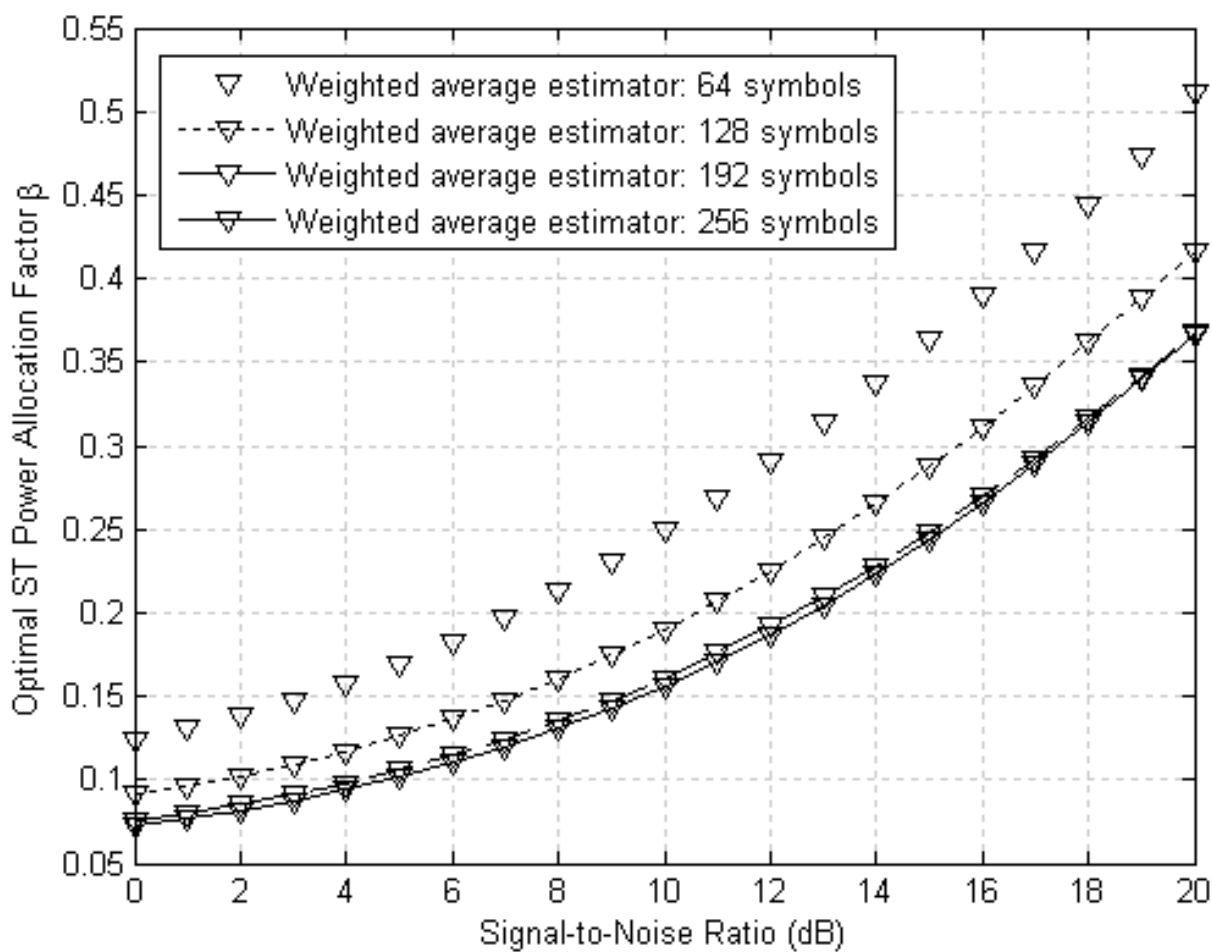
Fig. 8. Optimal ST power allocation factor of the proposed weighted average channel estimator versus SNR for different frame lengths.

## 7. Conclusions

In this paper, we have a developed superimposed training-aided LTV channel estimation approach for MIMO/OFDM systems. The LTV channel coefficients were firstly modeled by truncated DFB, and then estimated by using a two-step approach over multiple OFDM symbols. We also present a performance analysis of the proposed estimation approach and derive closed-form expressions for the channel estimation variances. It is shown that the estimation variances, unlike the conventional ST, approach to a fixed lower-bound that can only be reduced by increasing the pilot power. Using the developed channel estimation variance expression, we analyzed the system capacity and optimize the training power allocation by maximizing the lower bound of the average channel capacity for systems with a limited power. Compared with the existing FDM training based schemes, the new estimator does not entail a loss of rate while yields a better estimation performance, and thus enables a higher efficiency.

## 8. Ackownledgment

## 9. Reference

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

[2] X. L. Ma, G. B. Giannakis and B. Lu "Block differential encoding for rapidly fading channels," *IEEE Trans. Commun.*, vol. 52, no. 3, pp.416-425, March 2004.

[3] H.-C. Wu, "Analysis and characterization of intercarrier and interblock interferences for wireless mobile OFDM systems," *IEEE Trans. Broadcasting*, vol. 52, no. 2, pp. 203–210, Jun. 2006.

[4] X. Dai "Adaptive blind source separation of multiple-input multiple-output linearly time-varying FIR system," *IEE Proc.-Vis. Image Signal Process.*, Vol. 151, No. 4, pp.279-286, August 2004.

[5] Z. Tang, R. C. Cannizzaro, G. Leus and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems," *IEEE Trans. Signal Process*. Vol. 55, no. 5, pp. 2226-2238, May 2007.

[6] X. Dai, "Optimal estimation of linearly time-varying MIMO/OFDM channels modeled by a complex exponential basis expansion," *IET Commun.*, vol. 1, no. 5, pp. 945-953, 2007.

[7] H. C. Wu and Y. Wu, "Distributive pilot arrangement based on modified M-sequences for OFDM intercarrier interference estimation," IEEE Trans. Wireless Commun., vol. 5, no. 6, pp. 1605-1609, May 2007.

[8] G. T. Zhou, M. Viberg, and T. McKelvey, "A first-order statistical method for channel estimation," *IEEE Signal Processing Letters*, vol. 10, no. 3, pp.57-60, March 2003.

[9] J. K. Tugnait and W. Luo "On channel estimation using superimposed training and first-order statistics," *IEEE Comm. Letters*, vol. 7, no. 9, pp. 413-416, Sept. 2003.

[10] J. K. Tugnait and X.H. Meng "On superimposed training for channel Estimation: performance analysis, training power allocation, and frame synchronization," *IEEE Trans. Signal Process.*, vol. 54, no.2, pp.752-765, Feb. 2006.

[11] M. Ghogho, D. McLernon, E. A. Hernandez, and A. Swami "Channel estimation and symbol detection for block transmission using data-dependent superimposed training," *IEEE Signal Process. Letters*, vol. 12, no. 3, pp. 226-229, March 2005.

[12] W. C. Huang, C. P. Li and H. J. Li, "On the power allocation and system capacity of OFDM systems using superimposed training schemes," *IEEE Trans. Veh. Technol.*, vo. 58, no. 4, pp. 1731-1740, May, 2009.

[13] Q. Y. Zhu and Z.Q. Liu, "Optimal pilot superimposition for zero-padded block transmissions," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2194-2201, Aug. 2006.

[14] H. Zhang, X. Dai, "Time-varying Channel Estimation and Symbol Detection for OFDM Systems using Superimposed Training," *IEE Electronics Letter*, vol. 43, no. 22, Oct. 2007.

[15] S. M. A. Moosvi, D. C. McLernon, A. G. Orozco-Lugo, M. M. Lara, and M. Ghogho, "Carrier frequency offset estimation using data dependent superimposed training," *IEEE Commun. Letter*, vol. 12, no. 3, pp. 179-181, Mar. 2008.

[16] R. C.-Alvarez, R. P. -Michel, A. G. O. -Lugo, and J. K. Tugnait, "Enhanced channel estimation using superimposed training based on universal basis expansion," *IEEE Trans. Signal Process*. vol. 57, no. 3, Mar. 2009.

[17] F. Mazzenga, "Channel estimation and equalization for M-QAM transmission with a hidden pilot Sequence," *IEEE Trans. Broadcasting*, vol. 46, no. 2, pp. 170-176, June 2000.

[18] A. Goljahani, N. Benvenuto, S. Tomasin and L. Vangelista, "Superimposed sequence versus pilot aided channel estimations for next generation DVB-T systems," *IEEE Trans. Broadcasting*, vol. 55, no. 1, pp. 140-144, Mar. 2009.

[19] C. -P. Li and W. Hu, "Super-imposed training scheme for timing and frequency synchronization for OFDM systems," *IEEE trans. Broadcasting*, vol. 53, no. 2, pp. 574-583, Jun. 2007.

[20] A. Y. Panah, R. G. Vaughan and R. W. Heath, Jr. , "Optimizing pilot locations using feedback in OFDM systems," *IEEE trans. Vehicular Tech.* , vol. 58, no. 6, pp. 2803-2814, July 2009.

[21] S. Ohno and G.. B. Giannakis, "Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency- selective block Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 2138-2145, Sep. 2004.

[22] I. Barhumi, G. Leus and M. Moonen "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp.1615-1624, June 2003.

[23] Y. Xiaoyan, W. Jiaqing, Y. Luxi and H. Zhenya, "Doubly selective fading channel estimation in MIMO OFDM systems," *SCIENCE CHINA Information Sciences*, vol. 48, no. 6, pp. 795-807, 2005.

[24] J. Z. Wang and J. Chen, "Performance of wideband CDMA with complex spreading and imperfect channel estimation," *IEEE Journal on Selected Areas in Commun.*, vol. 19, no. 1, pp. 152-163, Jan. 2001.

[25] J. Z. Wang and L. B. Milstein, "CDMA overlay situations for microcellular mobile communications," *IEEE Trans. on Commun.*, vol. 43, no. 2/3/4, pp. 603-614, Feb/March/April 1995.

[26] X. Dai, H. Zhang and D. Li, 'Linearly time-varying channel estimation for MIMO-OFDM systems using superimposed training,' *IEEE trans. Commun.*, vol. 58, no. 2, pp. 681-693, Feb. 2011.

[27] H. Zhang, et. al., 'Linearly time-varying channel estimation and symbol detection for OFDMA uplink using superimposed training,' *EURASIP J. Wireless Commun. Networking*, vol. 2009.

[28] H. Zhang, X. Dai and D. Li, 'Time-varying channel estimation for MIMO/OFDM systems using superimposed training,' *in Proc. IEEE WiCOM* 08, pp. 1-6, 2008.
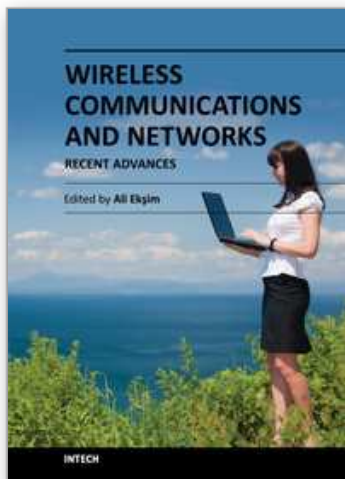
This book will provide a comprehensive technical guide covering fundamentals, recent advances and open issues in wireless communications and networks to the readers. The objective of the book is to serve as a valuable reference for students, educators, scientists, faculty members, researchers, engineers and research strategists in these rapidly evolving fields and to encourage them to actively explore these broad, exciting and rapidly evolving research areas.

# INTECH
open science | open minds