[13] I. Kim, S. Park, D. J. Love, and S. Kim, "Partial channel state information unitary precoding and codebook design for MIMO broadcast systems," in *Proc. IEEE Globecom*, Washington, DC, Nov. 2007, pp. 1607–1611.
[14] Y. Kim, K. Song, R. Narasimhan, and J. M. Cioffi, "Single user random beamforming in Gaussian MIMO broadcast channels," in *Proc. IEEE ICC*, Seoul, Korea, May 2005, pp. 2695–2699.
[15] M. Arkawa, *Computational Workloads for Commonly Used Signal Processing Kernels*. Lexington, MA: Massachusetts Inst. Technol., Lincoln Lab., 2006.

A Minimum-Complexity High-Performance Channel Estimator for MIMO-OFDM Communications

Tingting Liu, Mao Wang, Yan Liang, Feng Shu, Jianxin Wang, Weixin Sheng, and Qian Chen

***Abstract*—Channel estimation is critical to the performance of a multiple-input multiple-output and orthogonal frequency-division-multiplexing (MIMO-OFDM) communication system. Current MIMO-OFDM channel estimation techniques suffer from either high complexity or poor performance. In fact, this issue has become one of the bottlenecks of a practical MIMO-OFDM system. This paper solves the problem by introducing a new technique that has minimum complexity and yet provides superior channel tracking performance close to a minimum mean square error (MMSE) estimate over a wide range of channel selectivity and noise variance. The performance was evaluated via numerical simulations.**

***Index Terms*—Channel estimation, multiple-input multiple-output and orthogonal frequency-division multiplexing (MIMO-OFDM).**

I. Introduction

Orthogonal frequency-division multiplexing (OFDM), combined with the multiple-input multiple-output (MIMO) technique, has become a core technology for next-generation wireless communication systems [1]–[3]. Channel estimation is typically one of the most crucial and high-complexity components in a MIMO-OFDM system. There are two main issues in designing channel estimators for MIMO-OFDM systems. The first issue is the pilot structure, while the second issue is the design of an estimator with low complexity, good channel tracking ability, and low pilot overhead requirement. In general, the time- and frequency-selective OFDM channel can be viewed as a 2-D, i.e., time and frequency, grid. Channel estimation is usually performed by inserting known reference symbols (pilots) onto the subcarriers in such time and frequency grid. The receiver estimates the channel by a certain criterion. Unfortunately, the optimal Wiener filter interpolation and the minimum mean square error (MMSE) estimator for such a 2-D structure are too complex for practical implementation [5]. A suite of simplified algorithms has been proposed to provide various performance and complexity tradeoffs. Some of them require either second-order channel statistics (i.e., channel covariance matrix) or moderate to high complexity [6]–[14]. Among these, the simplest channel estimation technique that does not require second-order channel statistics with moderate performance is based on polynomial interpolation. However, this type of method is typically not adaptive to noise and channel selectivity (i.e., the order of the polynomial is typically fixed), thereby causing overfit when the channel has low selectivity and/or low SNR and underfit when channel has high selectivity. It is known that MIMO-OFDM channel estimation has become the bottleneck of a *practical* MIMO-OFDM system [15], [16]. A practical low-complexity channel estimator with good tracking accuracy is thus highly desirable. In this paper, we propose a minimum complexity channel estimation technique that does not require channel statistics (i.e., covariance matrix) and provides superior channel tracking performance over a wide range of channel selectivity and noise/interference variance.

Manuscript received May 26, 2010; revised August 16, 2010 and September 14, 2010; accepted September 15, 2010. Date of publication September 30, 2010; date of current version November 12, 2010. The review of this paper was coordinated by Dr. T. Taniguchi.

The authors are with the Wireless Networking and Mobile Communications Group, School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: liutingting1026@hotmail.com; wangmao@mail.njust.edu.cn; lyliang79@yahoo.com.cn; shufeng@mail.njust.edu.cn; wangjxin@mail.njust.edu.cn; chenq@mail.njust.edu.cn).

Digital Object Identifier 10.1109/TVT.2010.2081694

II. Multiple-Input–Multiple-Output Pilots

Spatial multiplexing offers a linear increase in data rate through MIMO systems, i.e., transmitting multiple independent data streams, which are often referred to as the MIMO *layers*, within the bandwidth of operation. Spatial multiplexing gain can be realized by simply transmitting data streams via channel eigenmodes, thus maximizing the overall data rate over the MIMO system. Assume that a vector of L independent modulation symbols from L data streams is transmitted over one OFDM subcarrier of an $N_T \times N_R$ MIMO system, where N_T is the number of transmit antennas, N_R is the number of receive antennas, and $L \leq \min\{N_T, N_R\}$ [17]. The multiple transmitted data streams of the MIMO layers share the same time–frequency resource and therefore interfere with one another at the receiver. However, the modulation symbols from different MIMO layers can be decoupled from the received signals with knowledge of the MIMO layer channel (as well as the noise variance) using either the linear MMSE detector or the zero-forcing detector [17].

In a MIMO-OFDM system, data are often transmitted via a time–frequency resource block that consists of N_c contiguous OFDM subcarriers and N_s successive OFDM symbols. Hence, there are $N_c \times N_s$ modulation symbols in one resource block. A user may be assigned integer multiples of such blocks during a communication session. Each resource block is used for the simultaneous transmission of L MIMO layers, as illustrated in Fig. 1, where L MIMO layers of data are transmitted on the same time–frequency resource block through an $N_T \times N_R$ MIMO system. Let us denote $\mathbf{y}$ as the vector of received modulation symbols in one resource block at receive antenna r, i.e.,

$$\mathbf{y}^r = \sum_{l=1}^{L} \mathbf{H}^{(l,r)}\mathbf{s}^{(l,r)} + \mathbf{n}^r, \qquad 1 \leq r \leq N_R \tag{1}$$

where $\mathbf{s}^{(l)}$ is a complex vector containing N_cN_s modulation symbols from the MIMO layer l. $\mathbf{H}^{(l,r)} = \mathrm{diag}(\mathbf{h}^{(l,r)})$ and $\mathbf{h}^{(l,r)} = [h_1^{(l,r)} \quad h_2^{(l,r)} \quad \cdots \quad h_{N_cN_s}^{(l,r)}]^{\mathrm{T}}$ contain the complex-valued layer channel gains at the N_cN_s modulation symbol positions, and $\mathbf{n}$ is an $N_cN_s \times 1$ zero-mean complex noise/interference vector with variance $\sigma^2\mathbf{I}$. Note that both $\{\mathbf{h}^{(l,r)}, 1 \leq l \leq L, 1 \leq r \leq N_R\}$ and σ^2 are typically unknown and need to be estimated for MIMO receive processing.

As earlier stated, the L layers of data are transmitted on the same resource block and can only be separated if the channel of each layer at each antenna (as well as the noise variance) is known. To facilitate

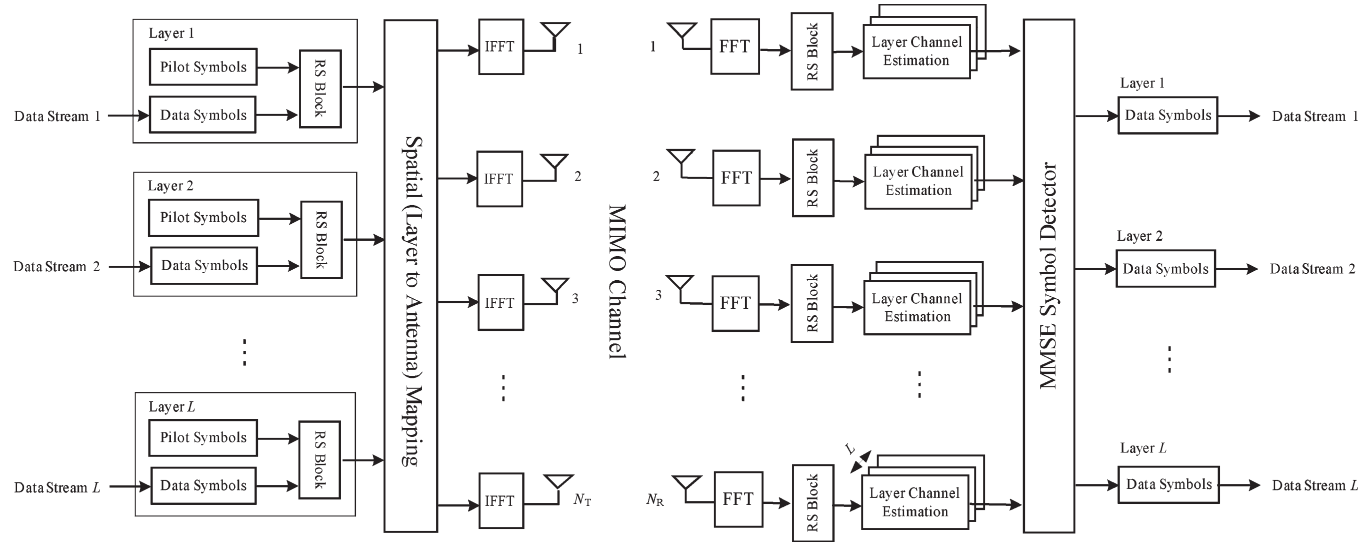

Fig. 1. MIMO communication system diagram, where L MIMO layers of data are transmitted through an $N_T \times N_R$ MIMO system.

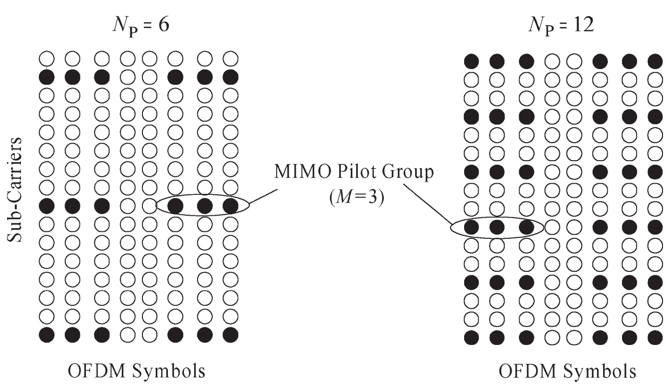

Fig. 2. MIMO pilot symbol placement in a resource block (solid: pilot symbol; hollow: data symbol), where $M = 3$, $N_P = 6$, and $N_P = 12$ for the first and the second pilot patterns, respectively, for channels with different frequency selectivities.

the estimation of the layer channel at the receiver, the transmission of layer pilots is needed, and the layer pilots themselves should be made separable at the receiver. In the implementations of Fig. 2, layer pilot symbols are inserted onto a resource block in such a way that pilot symbols are clustered into N_P groups ($N_P = 6$ and 12 for the first and second pilot patterns, respectively), each of which consists of $M = 3$ layer pilot symbols spread over time [18]. Within a group, the channel is assumed constant. The M pilot symbols of group p for layer l, i.e., $\mathbf{s}_p^{(l)}$ (an $M \times 1$ vector), are scrambled with a unique complex orthogonal sequence, i.e.,

$$\mathbf{s}_p^{(l)} = \mathbf{f}_l = \frac{1}{\sqrt{M}} \left[1 \quad e^{j\frac{2\pi l}{M}} \quad \cdots \quad e^{j\frac{2\pi l}{M}(M-1)} \right]^{\mathrm{T}}$$
$$1 \le l \le L \le M, \quad 1 \le p \le N_P \qquad (2)$$

where $\mathbf{f}_l$ is the lth column vector of an $M \times M$ discrete Fourier transform (DFT) matrix [19]. Here, we utilize the property that the columns of a DFT matrix are orthogonal, i.e., $\mathbf{f}_l^{\mathrm{H}}\mathbf{f}_m = \begin{cases} 0 & l \neq m \\ 1 & l = m \end{cases}$, providing a total of M dimensions of orthogonal subspaces spanned by these M column vectors. As such, the layer pilots are orthogonal to each other within a group and can thus be separated at the receiver. It is clear that the maximum number of layers that a pilot pattern with group size of M can support is M.

At the receiver, the lth layer pilot p at antenna r (cf., Fig. 1) is recovered by projecting the received M pilot symbols $\mathbf{y}_p^r$ in group p as

$$\begin{aligned} \mathbf{y}_p^r &= \sum_{m=1}^{L} h_p^{(m,r)} \mathbf{s}_p^{(m)} + \mathbf{n}_p^r \\ &= \sum_{m=1}^{L} h_p^{(m,r)} \mathbf{f}_m + \mathbf{n}_p^r, \qquad 1 \le p \le N_p \end{aligned} \qquad (3)$$

onto the lth dimension by multiplying the lth basis vector $\mathbf{f}_l$ with the received pilot symbol vector $\mathbf{y}_p^r$ of group p. Noting that $\mathbf{f}_l^{\mathrm{H}}\mathbf{f}_m = 0$ for $l \neq m$, we have

$$\tilde{h}_p^{(l,r)} = \mathbf{f}_l^{\mathrm{H}}\mathbf{y}_p^r = \sum_{m=1}^{L} h_p^{(m,r)} \mathbf{f}_l^{\mathrm{H}}\mathbf{f}_m + \mathbf{f}_l^{\mathrm{H}}\mathbf{n}_p = h_p^{(l,r)} + n_p^{(l,r)}$$
$$1 \le p \le N_P, \quad 1 \le l \le L, \quad 1 \le r \le N_R. \qquad (4)$$

The N_P lth layer channel gains at the rth antenna are thus separated (with noise) from the other layer and can be used for layer channel estimation.

The dimensions that are not utilized for layer pilot transmission, i.e., the $(M - L)$ dimensions, are used for noise variance estimation by extracting the interference samples from these dimensions. Since there is no pilot transmitted on any of these $(M - L)$ dimensions, the sample extracted must be the noise sample

$$n_p^{(l,r)} = \mathbf{f}_l^{\mathrm{H}}\mathbf{r}_p = \sum_{m=1}^{L} h_p^{(m,r)} \mathbf{f}_l^{\mathrm{H}}\mathbf{f}_m + \mathbf{f}_l^{\mathrm{H}}\mathbf{n}_p^r = \mathbf{f}_l^{\mathrm{H}}\mathbf{n}_p^r$$
$$1 \le p \le N_P, \quad L+1 \le l \le M, \quad 1 \le r \le N_R. \qquad (5)$$

Averaging over the $N_P(M - L)$ noise samples, the noise variance is then estimated by

$$\hat{\sigma}_r^2 = \frac{1}{N_P(M-L)} \sum_{l=L+1}^{M} \sum_{p=1}^{N_P} \left| n_p^{(l,r)} \right|^2, \qquad 1 \le r \le N_R. \qquad (6)$$

In the next section, we describe a technique that uses (4) and (6) to estimate the layer channel at each antenna in a resource block.

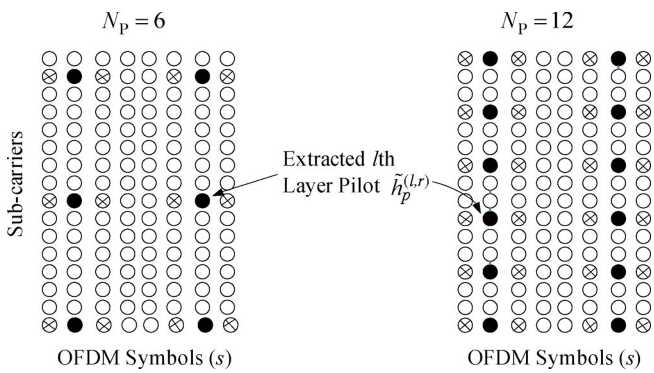

Fig. 3. Extracted pilot symbols for a single layer l $(1 \le l \le L)$ and receive antenna $r(1 \le r \le N_R)$ corresponding to the MIMO pilot patterns in Fig. 2.

III. Multiple-Input–Multiple-Output Layer Channel Estimation

The extracted layer pilots at each antenna from (4) are illustrated in Fig. 3, where $N_C = 16$, $N_S = 8$, $N_P = 6$ (Pattern 1), and $N_P = 12$ (Pattern 2). With the extracted N_P layer pilots and the associated noise variance from (6), the task is then to estimate the channel for $N_D = N_C N_S - M N_P$ layer data symbols (hollow circles) from the N_P layer pilots (solid circles) for the total of L layers and N_R receive antennas (cf., Fig. 1) by exploiting the channel time and frequency coherence within a resource block.

The time- and frequency-selective 2-D MIMO-OFDM channel $h^{(l,r)}(\mathbf{x})$ of layer l at antenna r is first modeled by

$$\psi^{(l,r)}(s,c) = \beta_0^{(l,r)} + \beta_1^{(l,r)} s + \beta_2^{(l,r)} + \sum_{p=1}^{N_P} \alpha_p^{(l,r)} \eta\left((s-s_p)^2 + (c-c_p)^2\right) \quad 1 \le l \le L, \quad 1 \le r \le N_R \tag{7}$$

where $\eta(\cdot)$ is a basis function, (s,c) denotes the symbol position in the resource block (s is the OFDM symbol index or time, c is the subcarrier index or frequency), (s_p, c_p) is the pth pilot symbol position in a resource block, and α_i and β_i are weights. That is, the channel is modeled by the linear combination of basis functions [20]. For notational simplification, we drop the layer index l and the antenna index r in the sequel wherever there is no confusion. We rewrite the foregoing equation in matrix form as

$$\psi(\mathbf{x}) = \begin{bmatrix} \boldsymbol{\eta}_{\mathbf{x}}^{\mathrm{T}} \ \boldsymbol{\chi}_{\mathbf{x}}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix} \tag{8}$$

where

$$\boldsymbol{\eta}_{\mathbf{x}} = \left[\, \eta\left(\|\mathbf{x} - \mathbf{z}_1\|^2\right) \quad \cdots \quad \eta\left(\|\mathbf{x} - \mathbf{z}_{N_P}\|^2\right) \,\right]^{\mathrm{T}} \tag{9}$$

where $\eta(r) = \begin{cases} r^2 \ln r^2, & r > 0 \\ 0, & r = 0 \end{cases}$ is the radial basis function [20], $\mathbf{x} = [s \quad c]^{\mathrm{T}}$ is the symbol position vector, and $\mathbf{z}_p = [s_p \quad c_p]^{\mathrm{T}}$ is the pth pilot symbol position vector. $\boldsymbol{\chi}_{\mathbf{x}} = [1 \quad \mathbf{x}^{\mathrm{T}}]^{\mathrm{T}}$, $\boldsymbol{\alpha} = [\alpha_1 \quad \cdots \quad \alpha_{N_P}]^{\mathrm{T}}$, and $\boldsymbol{\beta} = [\beta_0 \quad \beta_1 \quad \beta_2]^{\mathrm{T}}$ are the weight vectors that satisfy

$$\mathbf{Z}^{\mathrm{T}} \boldsymbol{\alpha} = \mathbf{0} \tag{10}$$

where $\mathbf{Z} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \mathbf{z}_1 & \mathbf{z}_2 & \cdots & \mathbf{z}_{N_P} \end{bmatrix}^{\mathrm{T}}$ is an $N_P \times 3$ matrix.

To determine the values for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, we construct a cost function consisting of a fidelity component *regularized* by a smoothness constraint

$$\zeta(\psi) = \underbrace{\frac{1}{N_P} \sum_{p=1}^{N_P} \left\| \psi(\mathbf{z}_p) - \tilde{h}(\mathbf{z}_p) \right\|^2}_{\text{Fidelity}} + \lambda \underbrace{J(\psi)}_{\text{Smoothness}} \tag{11}$$

where $\tilde{h}(\mathbf{z}_p) \triangleq \tilde{h}_p$ is the extracted channel at position $\mathbf{z}_p$, $\lambda \ge 0$ is the coefficient that governs the degrees of smoothing and is a function of the received pilot SNR, and J is a smoothness measure that provides the integral of the energy of the second-order derivative of the 2-D function ψ [21], i.e.,

$$J(\psi) = \iint \left(\frac{\partial^2 \psi}{\partial s^2}\right)^2 + 2\left(\frac{\partial^2 \psi}{\partial s \partial c}\right)^2 + \left(\frac{\partial^2 \psi}{\partial c^2}\right)^2 ds\, dc. \tag{12}$$

Equation (11) can then be rewritten as the matrix form

$$\zeta(\psi) = \frac{1}{N_P} \left\| \mathbf{Z}\boldsymbol{\beta} + (\mathbf{E} + \lambda N_P \mathbf{I})\, \boldsymbol{\alpha} - \tilde{\mathbf{h}}_P \right\|^2 + \lambda \boldsymbol{\alpha}^{\mathrm{T}} (\mathbf{E} + \lambda N_P \mathbf{I}) \boldsymbol{\alpha} \tag{13}$$

where $\tilde{\mathbf{h}}_P = [\tilde{h}_1 \quad \tilde{h}_2 \quad \cdots \quad \tilde{h}_{N_P}]^{\mathrm{T}}$ contains the extracted layer pilot vector whose elements are given by (4), and $\mathbf{E} = [\boldsymbol{\eta}_{\mathbf{z}_1} \quad \cdots \quad \boldsymbol{\eta}_{\mathbf{z}_{N_P}}]$ is an $N_P \times N_P$ matrix.

The first term of (11) or (13) represents the fidelity or closeness of the channel model ψ matched to the received (noise-corrupted) pilot symbols, and the second term reflects the smoothness or channel coherence in time and frequency. In contrast with the channel, noise samples are independent and therefore do not possess the correlation property as the channel. Therefore, the smoothness constraint helps provide extra robustness to noise in low SNR scenarios. For example, for high SNR, the received layer pilot symbols are less noisy and more reliable. The fidelity (fidelity to received pilot symbols) term should be weighted high (i.e., small λ). On the other hand, for low SNR, the received pilot symbols are heavily corrupted, and the fidelity term should then be weighted low (i.e., large λ). In general, the larger the pilot signal-to-interference-plus-noise ratio (SINR) (i.e., the more reliable the measured channel via the received pilot), the smaller the value of λ.

For a given λ, we seek the values of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ minimizing ζ as

$$\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \zeta(\psi). \tag{14}$$

It can be shown via QR decomposition of $\mathbf{Z}$ that the solution to (14) is given by a set of linear equations

$$\begin{bmatrix} \mathbf{E} + \lambda N_P \mathbf{I} & \mathbf{Z} \\ \mathbf{Z}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\boldsymbol{\beta}} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{h}}_P \\ \mathbf{0} \end{bmatrix}. \tag{15}$$

Substituting (15) into (6), we finally arrive at

$$\hat{h}(\mathbf{x}) = [\, \boldsymbol{\eta}_{\mathbf{x}}^{\mathrm{T}} \quad \boldsymbol{\chi}_{\mathbf{x}}^{\mathrm{T}} \,] \begin{bmatrix} \mathbf{E} + \lambda N_P \mathbf{I} & \mathbf{Z} \\ \mathbf{Z}^{\mathrm{T}} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{h}}_P \\ \mathbf{0} \end{bmatrix}. \tag{16}$$

That is

$$\hat{h}^{(l,r)}(\mathbf{x}) = \mathbf{c}_{\mathbf{x}}^{\mathrm{T}} \begin{bmatrix} \tilde{\mathbf{h}}_P^{(l,r)} \\ \mathbf{0} \end{bmatrix}, \qquad 1 \le l \le L, \quad 1 \le r \le N_R \tag{17}$$

which is the estimated lth layer channel for the rth antenna at symbol position $\mathbf{x}$, where

$$\mathbf{c}_{\mathbf{x}}^{\mathrm{T}} = [\,\boldsymbol{\eta}_{\mathbf{x}}^{\mathrm{T}} \quad \boldsymbol{\chi}_{\mathbf{x}}^{\mathrm{T}}\,] \begin{bmatrix} \mathbf{E} + \lambda N_P \mathbf{I} & \mathbf{Z} \\ \mathbf{Z}^{\mathrm{T}} & \mathbf{0} \end{bmatrix}^{-1}. \tag{18}$$

The layer channel estimate at symbol position $\mathbf{x}$ is thus simply a *linear* combination of the N_P layer pilots. The lth layer channel estimates at the N_D data symbol positions at the rth antenna are

$$\hat{\mathbf{h}}_D^{(l,r)} = \mathbf{C} \begin{bmatrix} \tilde{\mathbf{h}}_P^{(l,r)} \\ \mathbf{0} \end{bmatrix}, \qquad 1 \le l \le L, \quad 1 \le r \le N_R \tag{19}$$

where $\mathbf{C} = [\mathbf{c}_{\mathbf{x}_1} \quad \cdots \quad \mathbf{c}_{\mathbf{x}_{N_D}}]^{\mathrm{T}}$ is referred to as the *channel estimation matrix*.

It is clear that the channel estimation matrix $\mathbf{C}$ is a function of pilot symbol positions (which are predetermined) and λ (a function of pilot SINR). Thus, for a given pilot pattern, the channel estimation matrix $\mathbf{C}$ is simply a function of the pilot SINR. We can therefore predetermine the values of λ at various levels of pilot SINR and precalculate $\mathbf{C}$ at the respective levels of SINR. At the time of channel estimation, the layer pilot SINR is estimated using (6). The channel estimation matrix $\mathbf{C}$, corresponding to the particular SINR, is then used for layer channel estimation.

Therefore, the complexity of channel estimation for each layer data modulation symbol is $N_P + 3$ complex multiplications and additions. For the case of the first pilot pattern in Fig. 3, it is just $6 + 3$ complex multiplications and additions per layer data symbol.

IV. Simulation Results

In this section, we evaluate the performance of the proposed channel estimation algorithm via a 4×2 MIMO-OFDM simulator with resource block size of 16 subcarriers (10-kHz spacing) by 8 OFDM symbols, as shown in Fig. 2. Two layers of data streams $(L = 2)$ are transmitted. Two out of the $M = 3$ dimensions in Fig. 1 are used for layer pilot transmission, and the third dimension (no layer pilot is transmitted) is used for noise variance estimation. Different channel models with various frequency selectivities, such as Pedestrian B (low frequency selectivity), Global System for Mobile Communications (GSM) (high frequency selectivity), and Vehicular B (very high frequency selectivity) [22], [23], with fading speed from 3 to 120 km/h at a carrier frequency of 2 GHz, are used in the simulations.

We begin by determining the λ values at different levels of pilot SINRs. At a given SINR level, the λ value that provides the best simulated uncoded bit error rate (BER) under various channel models is selected, and its corresponding channel estimation matrix $\mathbf{C}$ is precalculated and stored. In general, the performance of the channel estimation is insensitive to the selection of the λ value over a wide range of SNR values. The number of channel estimation matrices needed to be stored can thus be sparse, e.g., spaced at 5 dB apart, without significant performance impact.

With the channel estimation matrices $\mathbf{C}$ determined and stored for different levels of SINRs, layer channel estimation was first performed for the GSM channel model at the fading speed of 3 km/h, assuming the first pilot pattern depicted in Fig. 2. For every resource block received, the layer pilots were extracted using (4), and the noise variance was estimated using (6). The channel estimation matrix $\mathbf{C}$ corresponding to that particular SNR was then selected from the prestored channel estimation matrices and used for calculating the layer channel gain at each layer data modulation symbol position using (19). This procedure was repeated for both layers $(L = 2)$ at both receive antennas $(N_R = 2)$. Fig. 4 shows a sample result of the estimated layer channel gain plotted against the actual layer channel at each position

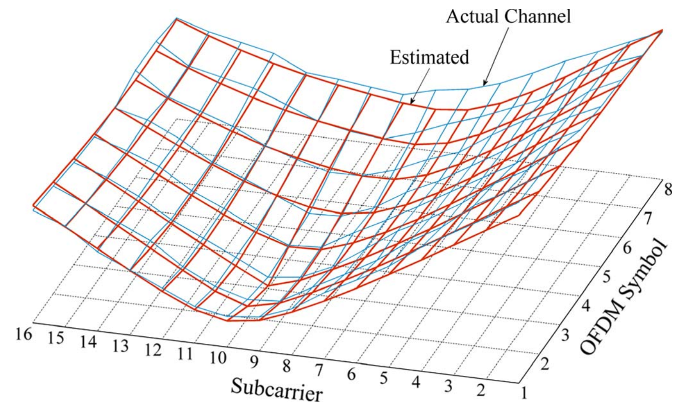

Fig. 4. Channel estimation (amplitude) sample plot for the GSM channel (SNR = 16 dB).

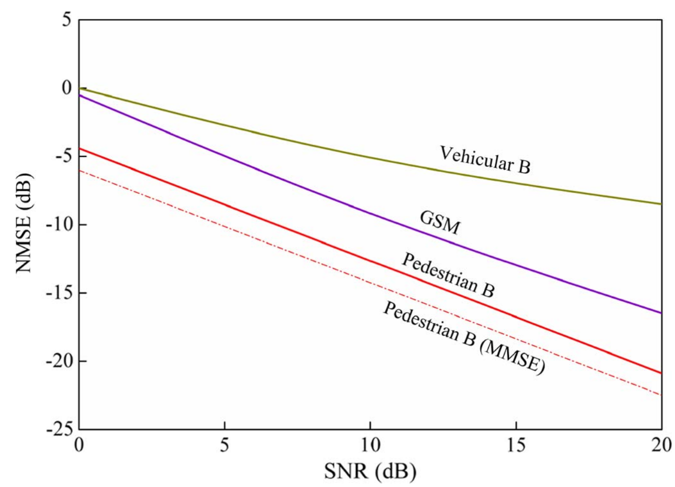

Fig. 5. NMSE for the proposed method. The performance for the MMSE-based method under the Pedestrian B model is also given. For the MMSE method, the channel covariance is assumed known at the receiver.

in a resource block. Finally, with the estimated channel gains for both layers and both antennas, together with the noise variance estimate, the MMSE detector was used to decouple the layer data symbols.

Simulations were repeated for other channels. The normalized mean square error (NMSE) of the channel estimate and the uncoded BER performance with quadrature phase-shift keying (QPSK) modulation were collected and plotted in Figs. 5 and 6, respectively, for Doppler speed of 3 km/h, together with the MMSE method for comparison. Ideal channel estimation (i.e., perfect channel knowledge) BER is also given. For Doppler speed of 120 km/h, the BER performance of the channel estimation is similar (not shown). It is seen in Fig. 6 that the BER performance from the proposed channel estimation technique follows the ideal channel estimation well for low (Pedestrian B) and high (GSM) frequency selectivity. At very high frequency selectivity (Vehicular B), the performance starts to deviate from the ideal. This deviation results from the fact that the channel estimation residual errors are no longer dominated by the pilot noise but by the severe mismatch between the channel frequency sampling rate provided by the first pilot pattern in Fig. 3 and the variation rate of the actual channel. The increase of frequency domain pilots, such as the use of the second pilot pattern in Fig. 3, provides better tracking of faster frequency variations (cf., Fig. 7) and brings down the deficit, as shown by the dotted line in Fig. 6, but at the cost of increased pilot overhead.

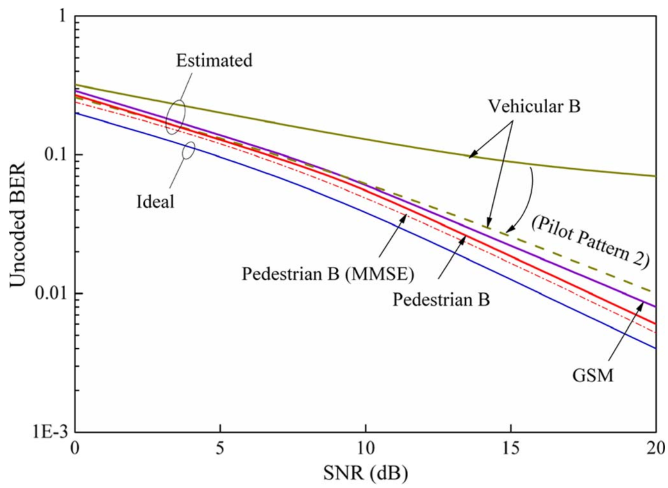

Fig. 6. Uncoded BER performance (QPSK modulation) versus SINR for Pilot Pattern 1 at Doppler speeds of 3 km/h under Pedestrian B, GSM, and Vehicular B channel models. The performance for the MMSE-based method under Pedestrian B model is also given for reference. For the MMSE method, the channel covariance is assumed known at the receiver.

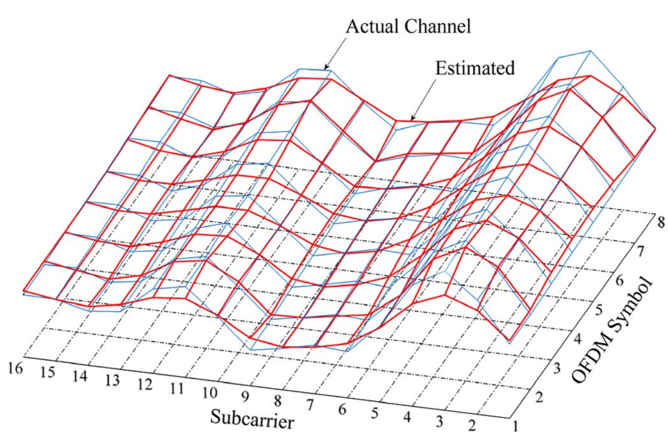

Fig. 7. Channel estimation (amplitude) sample plot for highly frequency-selective Vehicular B channel.

V. Conclusion

In this paper, a minimum-complexity high-performance channel estimation technique for MIMO-OFDM systems has been proposed. The channel estimation technique does not require knowledge of the channel statistics (i.e., channel covariance matrix). The channel estimation for each data symbol is simply a linear combination of the pilots. In particular, the number of complex multiplications and additions is equal to the number of pilots, i.e., the minimum channel estimation complexity, since any algorithm that uses less than N_P multiplications and additions would mean that some of the pilots are not used in the estimation and, thereby, would not lead to the best estimation performance. Despite its minimum complexity, this technique demonstrated its superiority (close to the high complexity MMSE method) in its ability of tracking channels with nonlinear selectivity inside a MIMO-OFDM resource block under low pilot overhead, providing channel tracking performance close to the MMSE over a wide range of channel selectivity, such as the GSM and Vehicular B channels. This performance is the direct result of the use of a cost function that is adaptive to channel noise and selectivity by incorporating a fidelity component regularized by a smoothing constraint. The solution is in a simple closed form that can be implemented with minimum complexity. Although the pilot design example used in this paper is for typical wireless mobile communication environments (Doppler speeds from 3 to 120 km/h), the proposed channel estimation technique can be, in general, used for channels with high selectivity in both frequency and time by simply using appropriate pilot patterns. The minimum complexity and superior channel tracking capability of the proposed technique enable an OFDM receiver to process high-rate MIMO data in high-selectivity channels with high performance and low cost, which has particular importance to a complexity and battery-limited mobile device.

Acknowledgment

The authors would like to thank the editor and the reviewers for their constructive suggestions.

References

[1] D. Astely, E. Dahlman, A. Furuskar, Y. Jading, M. Lindstrom, and S. Parkvall, "LTE: The evolution of mobile broadband," *IEEE Commun. Mag.*, vol. 47, no. 4, pp. 44–51, Apr. 2009.

[2] K. Lu, Y. Qian, and H.-H. Chen, "Wireless broadband access: WiMAX and beyond—A secure and service-oriented network control framework for WiMAX networks," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 124–130, May 2007.

[3] A. Greenspan, M. Klerer, J. Tomcik, R. Canchi, and J. Wilson, "IEEE 802.20: Mobile broadband wireless access for the twenty-first century," *IEEE Commun. Mag.*, vol. 46, no. 7, pp. 56–63, Jul. 2008.

[4] H. Bolcskei, "MIMO-OFDM wireless systems: Basics, perspectives, and challenges," *IEEE Wireless Commun.*, vol. 13, no. 4, pp. 31–37, Aug. 2006.

[5] O. Edfors and M. Sandell, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–938, Jul. 1998.

[6] B. Yang, K. Letaief, S. Cheng, and Z. Cao, "Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 467–479, Mar. 2001.

[7] X. Hou, S. Li, D. Liu, C. Yin, and G. Yue, "On two-dimensional adaptive channel estimation in OFDM systems," in *Proc. 60th IEEE Veh. Technol. Conf.*, Sep. 2004, vol. 1, pp. 498–502.

[8] F. Sanzi, J. Sven, and J. Speidel, "A comparative study of iterative channel estimators for mobile OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 849–859, Sep. 2003.

[9] M. Hsieh and C. Wei, "Channel estimation for OFDM systems based on comb-type pilot arrangement in frequency selective fading channels," *IEEE Trans. Consum. Electron.*, vol. 44, no. 1, pp. 217–225, Feb. 1998.

[10] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Commun.*, vol. 1, no. 1, pp. 67–75, Jan. 2002.

[11] G. Auer, "Channel estimation in two dimensions for OFDM systems with multiple transmit antennas," in *Proc. GLOBECOM*, 2003, pp. 322–326.

[12] D. Bueche, P. Corlay, M. Gazalet, and F. Coudoux, "A method for analyzing the performance of comb-type pilot-aided channel estimation in power line communications," *IEEE Trans. Consum. Electron.*, vol. 54, no. 3, pp. 1074–1081, Aug. 2008.

[13] J. Sterba and D. Kocur, "Pilot symbol aided channel estimation for OFDM system in frequency selective Rayleigh fading channel," in *Proc. 19th Int. Conf. Radio Elektronika*, Apr. 2009, pp. 77–80.

[14] J. Ylioinas and M. Juntti, "Iterative joint detection, decoding, and channel estimation in turbo-coded MIMO-OFDM," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1784–1796, Apr. 2009.

[15] A. Scherb and K. Kammeyer, "Bayesian channel estimation for doubly correlated MIMO systems," in *Proc. IEEE Workshop Smart Antennas*, Feb. 2007.

[16] P. Gupta and D. Mehra, "Simplified semi-blind channel estimation for space–time coded MIMO-OFDM systems," *Wireless Pers. Commun.*, Jun. 2010, DOI:10.1007/s11277-010-0066-9.

[17] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[18] M. Wang, A. Wang, A. Gorokhov, T. Kadus, and M. Dong, "Multi-antenna techniques for evolved 3G wireless communication networks: An overview," *J. Commun.*, vol. 4, no. 1, pp. 61–69, Feb. 2009.

[19] D. Sundararajan, *The Discrete Fourier Transform: Theory, Algorithms and Applications*. Singapore: World Scientific, 2001.
[20] M. Buhmann, *Radial Basis Functions*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[21] F. L. Bookstein, "Principle warps: Thin-plate splines and the decomposition and deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 6, pp. 567–585, Jun. 1989.
[22] ITU-R Recommendation M.1225, "Guidelines for evaluation of radio transmission technologies for IMT-2000," 1997.
[23] Specification 3GPP TS 45.005 Radio Transmission and Reception.

Power Allocation for Channel Estimation and Performance of Mismatched Decoding in Wireless Relay Networks

Duy H. N. Nguyen, *Student Member, IEEE*, and
Ha H. Nguyen, *Senior Member, IEEE*

***Abstract*—This paper is concerned with power allocation (PA) among the source and the relays of wireless relay networks to minimize the mean-square error (MSE) of the channel estimation when distributed space–time coding (DSTC) is applied. The optimal PA scheme is numerically obtained by means of geometric programming for the least-squares (LS) estimator, and a closed-form near-optimal PA scheme is also suggested. The impact of imperfect channel estimation on the error performance of DSTC is analyzed for both the LS and linear minimum MSE (LMMSE) channel estimators. It is proved that mismatched decoding of DSTC is able to achieve the same diversity order as coherent decoding of DSTC. Furthermore, when the closed-form PA obtained in the training phase is applied to the transmission phase, mismatched decoding is able to achieve a significant coding gain over the equal-PA scheme (which assigns half of the total power to the source and equally shares the other half to all the relays).**

***Index Terms*—Channel estimation, distributed space–time coding (DSTC), diversity order, geometric programming (GP), power allocation (PA).**

I. Introduction

Distributed space–time coding (DTSC) [1]–[4] has been proposed for wireless relay networks, where the relays cooperate with each other, simulate a virtual array of transmit antennas, and perform space–time coding on the source signal. By exploiting the spatial diversity provided by DSTC, it is well known that the transmission reliability of the source signal over a wireless relay network can be significantly improved. While most of the existing works on DSTC in the literature consider the relay networks with perfect channel state information (CSI) at the destination [1]–[4], only a few of them study the networks with imperfect CSI.

Mismatched decoding with imperfect channel estimation is investigated in [5] for a network with one relay. By combining with the direct source-to-destination $(S \rightarrow D)$ transmission, it is shown that the system is able to achieve a diversity order of 2. Channel estimation in a single-relay network has also been reported in [6]. Reference [7] has recently considered channel estimation and optimal training design for multiple-relay networks, where the channel variance of each link can take on any value. The optimal training at the source and the relays are presented in [7] to minimize the channel estimation mean-square error (MSE). However, [7] does not show how to optimally allocate the power among the source and the relays to further minimize the MSE, nor does it provide a diversity analysis of the mismatched decoder. These two important issues shall be examined in this paper.

Manuscript received March 25, 2010; revised July 20, 2010; accepted September 3, 2010. Date of publication September 27, 2010; date of current version November 12, 2010. The review of this paper was coordinated by Dr. C. Ling.
D. H. N. Nguyen is with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 2A7, Canada (e-mail: huu.n.nguyen@mail.mcgill.ca).
H. H. Nguyen is with the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada (e-mail: ha.nguyen@usask.ca).
Digital Object Identifier 10.1109/TVT.2010.2080296

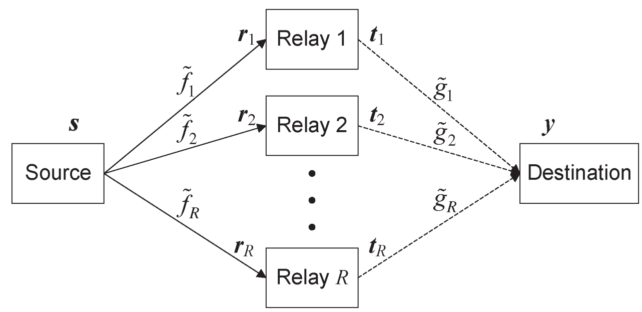

Fig. 1. Block diagram of a wireless relay network with DSTC.

This paper first studies power allocation (PA) schemes to minimize the MSE of the channel estimate, which is obtained with either the least-squares (LS) or the linear minimum MSE (LMMSE) criterion. With the LS criterion, we consider a geometric programming (GP)-based approach to find the optimal PA scheme. We also propose a closed-form PA scheme, whose performance is very close to that of the optimal scheme. Interestingly, this closed-form near-optimal scheme turns out to be same as the optimal PA scheme that has been recently proposed in [8] to maximize the signal-to-noise ratio (SNR) of the coherent DSTC at the destination under the minimum "amount of fading" constraint. With regard to the LMMSE criterion, although no optimal solution is found, the proposed PA schemes obtained under the LS criterion can be readily applied as suboptimal solutions. In Section IV, we prove that the mismatched decoder of DSTC that uses the imperfect channel estimation is able to achieve the *same* diversity order as the coherent decoder (which has the perfect channel estimation).

Notations: Superscripts $(\cdot)^{\mathcal{T}}$ and $(\cdot)^{\dagger}$ stand for transpose and complex conjugate transpose operations, respectively; $\boldsymbol{I}_M$ is an $M \times M$ identity matrix; $\mathrm{tr}(\cdot)$ denotes the trace of a square matrix; $\mathbb{E}_x[\cdot]$ indicates the expectation of random variable x; $\mathcal{CN}(0, \sigma^2)$ denotes a circularly symmetric complex Gaussian random variable with variance σ^2.

II. System Model

Consider a wireless relay network with $R+2$ nodes, as illustrated in Fig. 1. The system has one source node, one destination node, and R relay nodes. Each node is equipped with only one antenna, which is used for transmission and reception in the half-duplex mode. Assume that there is no direct link from the source to the destination, as all signals from the source are relayed to arrive at the destination. Let $\tilde{f}_i \sim \mathcal{CN}(0, \sigma^2_{F_i})$ and $\tilde{g}_i \sim \mathcal{CN}(0, \sigma^2_{G_i})$ be the channel coefficients from the source to the ith relay and from the ith relay to the destination, respectively, for $i = 1, \ldots, R$. These coefficients are assumed to be independent of each other. It is further assumed that $\tilde{f}_i$ and $\tilde{g}_i$, $i = 1, \ldots, R$ remain constant over the coherence time $T_C = 2T$, which includes both training time T and data-transmission time T. These