

Research Article

Iterative MIMO Detection and Channel Estimation Using Joint Superimposed and Pilot-Aided Training

Omar Longoria-Gandara,¹ Ramon Parra-Michel,²
Roberto Carrasco-Alvarez,³ and Eduardo Romero-Aguirre⁴

¹Department of Electronics, Systems and IT, ITESO Jesuit University, 45604 Tlaquepaque, JAL, Mexico

²Department of Electrical Engineering, CINVESTAV-IPN, 45015 Zapopan, JAL, Mexico

³Department of Electronics and Communication, CUCEI-Guadalajara University, 44430 Guadalajara, JAL, Mexico

⁴Department of Electrical and Electronic Engineering, ITSON, 85130 Ciudad Obregon, SON, Mexico

Correspondence should be addressed to Omar Longoria-Gandara; olongoria@iteso.mx

Received 10 September 2015; Revised 19 January 2016; Accepted 25 February 2016

Academic Editor: Yuh-Shyan Chen

Copyright © 2016 Omar Longoria-Gandara et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a novel iterative detection and channel estimation scheme that combines the effort of superimposed training (ST) and pilot-aided training (PAT) for multiple-input multiple-output (MIMO) flat fading channels. The proposed method, hereafter known as joint mean removal ST and PAT (MRST-PAT), implements an iterative detection and channel estimation that achieves the performance of data-dependent ST (DDST) algorithm, with the difference that the data arithmetic cyclic mean is estimated and removed from data at the receiver's end. It is demonstrated that this iterative and cooperative detection and channel estimator algorithm surpasses the effects of data detection identifiability condition that DDST has shown when higher orders of modulation are used. Theoretical performance of the MRST-PAT scheme is provided and corroborated by numerical simulations. In addition, the performance comparison between the proposed method and different MIMO channel estimation techniques is analyzed. The joint effort between ST and PAT shows that MRST-PAT is a solid candidate in communications systems for multi-amplitude constellations in Rayleigh fading channels, while achieving high-throughput data rates with manageable complexity and bit-error rate (BER) as a figure of merit.

1. Introduction

Estimation theory deals with the basic problem of inferring a set of required statistical parameters of a random experiment based on the observation of its outcome. It is assumed that it is possible to produce an effect in the experiment by means of a controlled excitation signal. This approach is normally adopted in practical communications systems where channel estimation is an essential part of standard receiver designs [1] and carried out by transmitting training symbols commonly known as pilot symbols [2]. In this case, the random experiment can be seen as an unknown system that is identified by observing how the system reacts to the applied excitation or training signal.

Traditionally, the most widely used channel estimation technique is pilot-assisted training where the pilots are

multiplexed in time or frequency. This is widely known in the literature as pilot symbol assisted modulation (PSAM) [2] and is denominated as pilot-assisted transmission (PAT) [3]. This scheme employs a nonrandom training pilot sequence known *a priori* by the transmitter and the receiver. The training pilots are periodically inserted into certain positions in the time (frequency) with the information-bearing symbols, before modulation and transmission. Using the knowledge of the training symbols and the corresponding received signal, the channel estimation block at the receiver is able to make an estimate of the channel impulse response (CIR). Conventional PAT-based channel estimation methods use pilot symbols in time-division multiplexing (TDM) schemes, thus decreasing the effective data transmission rate.

Recently, an alternative channel estimation strategy that circumvents the unwanted effect of data rate reduction has

emerged, called implicit training (IT) [4]. Two outstanding IT approaches, known as superimposed training (ST) [5, 6] and data-dependent ST (DDST) [7, 8], achieve higher effective data rates with manageable complexity [9]. These techniques are based on a training sequence added (superimposed) to the information-bearing symbols. Both schemes provide a simple (unsophisticated) channel estimation process [10, 11] and differ only in that the arithmetic mean of the transmitted data of the DDST scheme is superimposed onto the transmitted sequence. Both techniques have been successfully applied to single-input single-output (SISO), as well as multiple-input multiple-output (MIMO) systems [11–17], combined with orthogonal frequency division multiplexing (OFDM) modulations [18–20], and time-varying channels [21–23]. Likewise, there are interesting alternatives to OFDM in [24, 25], where the single-carrier modulation approach is employed, using a joint frequency domain equalization and channel estimation.

Although DDST outperforms ST [12] in terms of channel estimation error, it is worth mentioning that the data decoding under DDST is of an iterative nature, as it needs to remove the data-dependent distortion. Furthermore, DDST [8] performs similarly to TDM-based channel estimation, while saving the overhead in TDM data rate due to pilot transmission. There are, however, some drawbacks that must be taken into consideration when DDST is used. First, this technique introduces a delay in the transmitted data when it calculates the data-dependent signal; second, it assigns less transmission power to the data signal; hence, the symbol demapping operation is not suitable for higher orders of modulation due to identifiability problems, as was highlighted in [26]. This is an important issue, because recent communication standards consider this type of modulation; for example, WiMAX (IEEE 802.16e-2005 standard) uses 64-QAM digital modulation that DDST cannot support. In addition, DDST presents two more drawbacks: its performance is highly sensitive to the data block length used as well as the increased peak-power and peak-to-average ratio (PAPR) in the transmitted signal, as shown in [27].

This paper proposes an iterative detection and channel estimation structure for MIMO systems, which is capable of dealing with the data identifiability problem previously described. The novel algorithm is based on the use of joint (fusion of both techniques) implicit and explicit training for the channel estimation and symbol detection processes. Previous results of the authors related to implicit training for MIMO systems were presented in [28], which introduced the mean removal ST (MRST) technique devoted to estimate the arithmetic cycling mean of the data block signal at the receiver side (instead of at the transmitter side like DDST does). However, later on it was recognized that MRST is not able to deal with the identifiability problem, just like DDST, when the MIMO system employs high-order digital modulation schemes.

The proposal of an iterative detection and channel estimation scheme is hereafter referred to as joint mean removal ST and PAT (MRST-PAT). The method is based on a reliable preliminary channel estimate using PAT with a small number of dedicated pilot symbols. Subsequently, it uses the time-average estimator with the ST signal added to each

transmitted data block, achieving an improvement in system throughput. Thus, MRST-PAT merges the best qualities of the two techniques, achieving a better BER performance than if employed separately. In addition, the proposed method eliminates the data identifiability difficulties that DDST exhibits when higher orders of modulation are used [26].

This paper is organized as follows: Section 2 presents the space-time MIMO signal model. Section 3 summarizes the fundamental theory and performance of PAT (TDM), ST, and DDST channel estimation techniques. Section 4 details the structure of the iterative MIMO detection and channel estimation joint MRST-PAT and its theoretical performance analysis. Section 5 depicts the application of the iterative joint MRST-PAT using a \mathcal{C} 2-OSTBC coder. Section 6 considers the numerical simulation results and performance analysis for the training-based frequency-flat block-fading MIMO channel estimation using TDM, DDST, and MRST-PAT. Finally, some concluding remarks in Section 7 close this paper.

Notation. Bold letters in lower (upper) case are used to denote vector (matrices); \mathbb{C} stands for the complex number field; $(\cdot)^T$ and $(\cdot)^H$ represent transpose and conjugate transpose, respectively; \mathbf{I}_{n_T} denotes $n_T \times n_T$ identity matrix; $\|\cdot\|$ and $\|\cdot\|_F$ represent Euclidean norm and Frobenius norm, respectively; $\text{Tr}\{\cdot\}$ denotes trace of a matrix; \otimes stands for Kronecker product; and $\mathcal{C}\mathcal{N}(\mathbf{a}, \mathbf{\Sigma})$ denotes a multidimensional complex Gaussian distribution with mean \mathbf{a} and covariance matrix $\mathbf{\Sigma}$.

2. Space-Time Signal Model

This section describes the space-time signal model applicable to most existing space-time coding designs such as Vertical Bell Labs Layered Space-Time (VBLAST) [29] and the generalized schemes referred to as space-time block codes from orthogonal designs [30].

With reference to Figure 1, this paper considers a MIMO single-carrier system operating over frequency-flat quasi-static block-fading channel model, with n_T transmit and n_R receive antennas described by

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (1)$$

where $\mathbf{x} \in \mathbb{C}^{n_T}$ is the transmitted vector comprising n_T elements denoted as $[x_1, x_2, \dots, x_{n_T}]^T$ and $\mathbf{r} \in \mathbb{C}^{n_R}$ is the received vector.

We use a random matrix $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ that represents the flat fading channel in a baseband equivalent model, where each element is generated as an independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian random variable $\mathcal{C}\mathcal{N}(0, 1)$. In addition, $\mathbf{n} \in \mathbb{C}^{n_R}$ is an additive white circularly symmetric complex Gaussian noise vector, with zero mean and variance $\sigma_n^2 = N_0/2$ per real and imaginary dimension.

Let us assume the block transmission scheme, where a “block” is defined as a single transmission burst. The channel matrix \mathbf{H} is assumed to be constant for N channel uses and then changes to an independent realization for the next block [31]. Here, N denotes the block length. For any $n_T \times N$

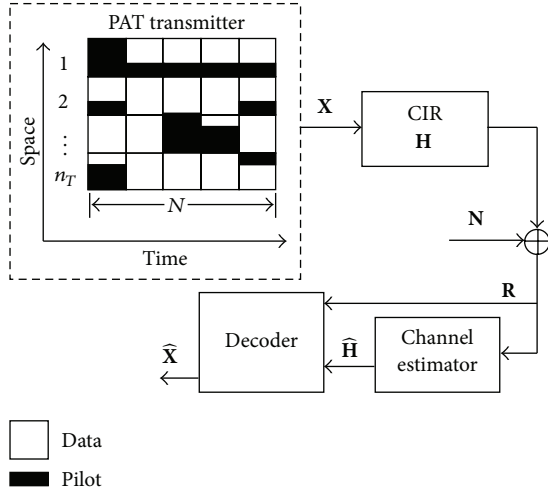


FIGURE 1: A generalized PAT transceiver.

transmitted signal matrix $\mathbf{X} = (x_{ij})$, the expression in (1) can be written as

$$\mathbf{R} = \mathbf{H}\mathbf{X} + \mathbf{N}, \quad (2)$$

where

$$\begin{aligned} \mathbf{R} &\triangleq [\mathbf{r}(1) \ \mathbf{r}(2) \ \cdots \ \mathbf{r}(N)], \\ \mathbf{X} &\triangleq [\mathbf{x}(1) \ \mathbf{x}(2) \ \cdots \ \mathbf{x}(N)], \\ \mathbf{N} &\triangleq [\mathbf{n}(1) \ \mathbf{n}(2) \ \cdots \ \mathbf{n}(N)] \end{aligned} \quad (3)$$

are the matrices of the received signals, transmitted signals, and noise, respectively.

The transmit symbol vector, prior to space-time coding, of size N_S is denoted by

$$\mathbf{s} \triangleq [s_1 \ s_2 \ \cdots \ s_{N_S}]^T, \quad (4)$$

where $s_i \in \mathbb{S}$ and \mathbb{S} denotes a normalized (average symbol energy of \mathbb{S} is assumed to be one) signal constellation to which information symbols belong.

The vector \mathbf{s} of N_S information symbols is fed into a space-time encoder, which maps (one-to-one mapping) this vector to $n_T \times N_G$ code matrix

$$\mathbf{G} = \begin{pmatrix} g_{11} & \cdots & g_{1N_G} \\ \vdots & \ddots & \vdots \\ g_{n_T 1} & \cdots & g_{n_T N_G} \end{pmatrix} \in \mathbb{C}^{n_T \times N_G} \quad (5)$$

associated with the space-time scheme in use. Therefore, the space-time code data rate, in symbols per channel use, is given by $\eta = N_S/N_G$. Additionally, the transmitted signal \mathbf{X} is formed by using $\Delta = N/N_G$ concatenated code matrices, where the block length N is a positive integer multiple of N_G .

Particularly, $n_T \times N_G$ complex matrix-valued function $\mathbf{G}(\mathbf{s})$ is called an OSTBC if it satisfies the following conditions [30]:

- (a) All entries of $\mathbf{G}(\mathbf{s})$ are linear functions of N_S complex random variables s_1, s_2, \dots, s_{N_S} and their complex conjugates.
- (b) For any arbitrary $\mathbf{s} \in \mathbb{C}^{N_S \times 1}$,

$$\mathbf{G}(\mathbf{s}) \mathbf{G}^H(\mathbf{s}) = \|\mathbf{s}\|^2 \mathbf{I}_{n_T}. \quad (6)$$

3. Channel Estimation Techniques

Consider the PAT diagram in Figure 1. During the transmission of one block of information, the data and pilot symbols, depicted by white and dark boxes, respectively, can be transmitted separately or combined in different ways. For the purpose of channel estimation, the key to unifying various schemes is to view the problem of training design as one of power allocation [3].

Conventionally, a PAT scheme where some known symbols (pilots) are time-multiplexed (TDM) with the information-bearing data is known as a conventional pilot-(CP-) based MIMO system [32]. This transmission scheme is shown in Figure 2 and analyzed in the following subsection. In this case, pilot symbols are known at the receiver and can be used for channel estimation. Since pilot symbols carry no data information, the time and the power spent on pilot symbols degrade the throughput of the system. If more frequent pilot symbols are transmitted, then a better channel estimation is achieved. However, the PAT transceiver shown in Figure 1 suggests another scheme for solving the problem, where a training sequence is superimposed onto the data. This technique, known as superimposed pilots (SIP) in [32], was introduced in [5] and has been developed extensively under different modalities such as ST in [33], implicit training in [4, 6], and DDST in [7].

It should be emphasized that, in the following calculations, \mathbf{H} is assumed to be random. The estimation process, however, will obtain estimates of the parameters of a particular realization of \mathbf{H} that corresponds to the current data block. We consider the channel estimation procedure applied to TDM, ST, DDST, and MRST-PAT schemes, starting from the least squares (LS) analysis proposed in [34].

3.1. PAT (TDM) Case. With reference to Figure 2, let us consider a PAT transceiver under TDM modality with $N = T + L$. In this scenario, T is the number of pilot symbols transmitted in a given transmission interval (block). $n_T \times T$ matrix $\mathbf{C} = (c_{ij})$ represents the training signal transmitted at the beginning of each block of information (preamble), and $n_T \times L$ matrix $\mathbf{B} = (b_{ij})$ represents the information-bearing data of length L per transmit antenna. In this scheme, the data and training symbols have the same average power defined as $P_b = E\{|b_{ij}|\}$ and $P_c = E\{|c_{ij}|\}$; that is, $P_b = P_c$. It is clear that during the training stage, the transmission power is allocated to the pilot symbols; that is, $P_c = P_x$, where $P_x = E\{|x_{ij}|\}$ denotes the average power of the transmitted symbol.

It is important to note that the TDM-based channel estimation algorithm employs only the pilot signals. It follows that the noise contribution and the received signal span during T channel uses. In this case, \mathbf{R} and \mathbf{N} have dimensions $n_R \times T$.

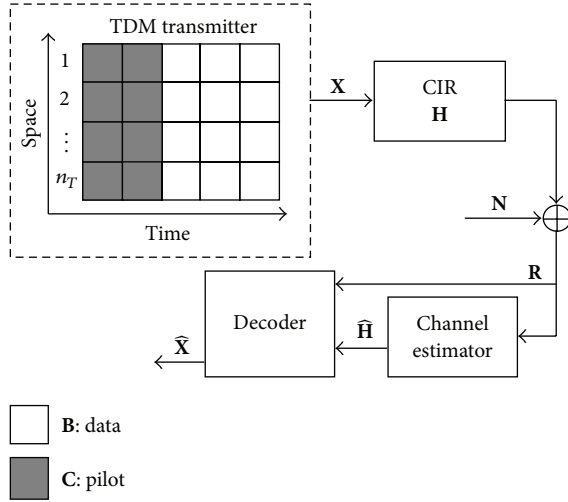


FIGURE 2: Block diagram of PAT transceiver using TDM transmitter.

Assuming that \mathbf{C} has full row-rank, knowing the received signal \mathbf{R} , and using the LS approach as is developed in [34, 35], then the realization of channel matrix \mathbf{H} can be estimated as

$$\hat{\mathbf{H}}_{\text{TDM}} = \mathbf{R}\mathbf{C}^H (\mathbf{C}\mathbf{C}^H)^{-1} = \mathbf{R}\mathbf{C}^\dagger, \quad (7)$$

where $\mathbf{C}^\dagger = \mathbf{C}^H(\mathbf{C}\mathbf{C}^H)^{-1}$ is the pseudoinverse of \mathbf{C} . Constraining the transmitted power during the training interval, the following is obtained:

$$\|\mathbf{C}\|_F^2 = n_T TP_x = \mathcal{P}. \quad (8)$$

Let us find \mathbf{C} that minimizes the channel estimation error and satisfies condition (8). This is equivalent to the following optimization problem:

$$\min_{\mathbf{C}} E \left\{ \|\mathbf{H} - \hat{\mathbf{H}}_{\text{TDM}}\|_F^2 \right\}. \quad (9)$$

Using (2) and (7), we have $\mathbf{H} - \hat{\mathbf{H}}_{\text{TDM}} = \mathbf{N}\mathbf{C}^\dagger$ and, therefore, the objective function can be written as

$$\mathcal{J}_{\text{TDM}} = E \left\{ \|\mathbf{N}\mathbf{C}^\dagger\|_F^2 \right\}. \quad (10)$$

Using $E\{\mathbf{N}^H\mathbf{N}\} = \sigma_n^2 n_R \mathbf{I}_T$ with σ_n^2 representing the receiver noise power, then (10) can be rewritten as

$$\begin{aligned} \mathcal{J}_{\text{TDM}} &= E \left\{ \|\mathbf{N}\mathbf{C}^\dagger\|_F^2 \right\} = \sigma_n^2 n_R \text{Tr} \left\{ \mathbf{C}^{\dagger H} \mathbf{C}^\dagger \right\} \\ &= \sigma_n^2 n_R \text{Tr} \left\{ (\mathbf{C}\mathbf{C}^H)^{-1} \right\}. \end{aligned} \quad (11)$$

From (11), the optimization problem shown by (9) can be equivalently reformulated as

$$\min_{\mathbf{C}} \text{Tr} \left\{ (\mathbf{C}\mathbf{C}^H)^{-1} \right\}. \quad (12)$$

The choice of \mathbf{C} must be such that $\mathbf{C}\mathbf{C}^H$ is nonsingular, which implies that $T \geq n_T$. Furthermore, any training matrix is optimal in the sense of (12) if it satisfies

$$\mathbf{C}\mathbf{C}^H = \frac{\mathcal{P}}{n_T} \mathbf{I}_{n_T}. \quad (13)$$

In other words, any matrix \mathbf{C} with orthogonal rows of the same norm $\sqrt{\mathcal{P}/n_T}$ is optimal [35].

For optimal training that satisfies (13), the LS channel estimate (7) yields

$$\hat{\mathbf{H}}_{\text{TDM}} = \left(\frac{n_T}{\mathcal{P}} \right) \mathbf{R}\mathbf{C}^H = \mathbf{H} + \left(\frac{n_T}{\mathcal{P}} \right) \mathbf{N}\mathbf{C}^H; \quad (14)$$

that is, the estimation error is $(n_T/\mathcal{P})\mathbf{N}\mathbf{C}^H$.

From (11) and (13), it follows that the channel estimation error under the MSE criterion and optimal training is given by

$$\text{MSE}_{\text{TDM}}(\hat{\mathbf{H}}) = \frac{\sigma_n^2 n_T^2 n_R}{\mathcal{P}}. \quad (15)$$

Finally, using the constraint indicated in (8), for the TDM-based estimate, the channel estimation error can be expressed as

$$\text{MSE}_{\text{TDM}}(\hat{\mathbf{H}}) = \frac{\sigma_n^2 n_T n_R}{TP_x}. \quad (16)$$

3.2. Superimposed Training Case. Let us consider Figure 1 with an ST-based channel estimation algorithm. In comparison with TDM-based channel estimation, the ST scheme has $P_b \neq 0$ throughout the transmission, and the training matrix \mathbf{C} is $n_T \times N$ matrix; that is, $T \equiv 0$. For this reason this scheme is called overlay pilot in [32]. It follows that (2) can be described as

$$\mathbf{R} = \mathbf{H}(\mathbf{B} + \mathbf{C}) + \mathbf{N}. \quad (17)$$

Accordingly, the least square (LS) channel estimate can be obtained by treating $\mathbf{H}\mathbf{B}$ as an extra additive noise term for channel estimation purposes. Even though there are an infinite number of matrices \mathbf{C} that satisfy the optimum condition of (12), the goal is to select a training matrix that makes the ST technique computationally attractive.

Under frequency-flat fading conditions, the optimal length of the training interval is n_T [36]. Hence the training matrix \mathbf{C} may be a block column matrix that can be constructed by Q concatenated replicas of \mathbf{C}_b matrix, where $Q = N/n_T$, and the block length N is chosen to be an integer multiple of the number of transmit antennas n_T . Then, the training matrix \mathbf{C} can be written as

$$\mathbf{C} = \mathbf{u} \otimes \mathbf{C}_b, \quad (18)$$

where \mathbf{u} is the unit $1 \times Q$ vector and the block training pattern \mathbf{C}_b is $n_T \times n_T$ matrix, where their rows are orthogonal to each other, and it satisfies the constraint $\mathbf{C}_b \mathbf{C}_b^H = n_T P_c \mathbf{I}_{n_T}$.

The training sequence at ℓ th transmit antenna is given by

$$c(\ell, n) = \sqrt{P_c} \exp \left(\frac{j2\pi n(\ell-1)}{n_T} \right) \quad (19)$$

with $\ell = 1, \dots, n_T$ and $n = 0, \dots, N-1$. Another way to build the training sequence is by shaping \mathbf{C}_b matrix as a circulant matrix whose first row entries are given by

$$c_b(1, \nu) = \sqrt{P_c} \left(\exp \left(\frac{j\pi}{n_T} \right) \right)^{(\nu+\alpha)}, \quad (20)$$

where $\nu = 0, \dots, n_T - 1$, with $\alpha = 2$ when n_T is even, and $\alpha = 1$ if n_T is odd [6]. The Walsh-Hadamard matrices are other attractive options.

The benefit of using any of these training patterns is that a time-domain estimator based on the synchronized averaging of the received signal \mathbf{R} can be implemented, as in the SISO case described in [12].

Let $\tilde{\mathbf{R}}$ be $n_R \times n_T$ matrix, which defines the time-average of the received signal \mathbf{R} . We assume that the channel remains constant over the entire block length of time N . Specifically, (i, j) th entry of $\tilde{\mathbf{R}}$ is given by

$$\tilde{\mathbf{R}}_{i,j} = \left(\frac{1}{Q} \right) \sum_{k=1}^Q \mathbf{R}_{i,j+n_T(k-1)}, \quad (21)$$

where $i = 1, 2, \dots, n_R$ and $j = 1, 2, \dots, n_T$.

It follows that $\tilde{\mathbf{R}}$ can be written as

$$\tilde{\mathbf{R}} = \left(\frac{1}{Q} \right) \mathbf{R}\mathbf{U}, \quad (22)$$

where \mathbf{U} is $N \times n_T$ matrix constructed by Q concatenated replicas of identity matrix \mathbf{I}_{n_T} and is given by

$$\mathbf{U} = \mathbf{u}^T \otimes \mathbf{I}_{n_T}. \quad (23)$$

From (17) and (22), $\tilde{\mathbf{R}}$ can be reformulated as

$$\begin{aligned} \tilde{\mathbf{R}} &= \mathbf{H} \left[\left(\frac{1}{Q} \right) (\mathbf{B}\mathbf{U} + \mathbf{C}\mathbf{U}) \right] + \left(\frac{1}{Q} \right) \mathbf{N}\mathbf{U} \\ &= \mathbf{H} (\tilde{\mathbf{B}} + \mathbf{C}_b) + \tilde{\mathbf{N}}, \end{aligned} \quad (24)$$

where $\tilde{\mathbf{B}}$ is $n_T \times n_T$ matrix and $\tilde{\mathbf{N}}$ is $n_R \times n_T$ matrix that represent the time-average of the data \mathbf{B} and noise \mathbf{N} , respectively, for every block transmitted. Moreover, $\mathbf{C}_b = (1/Q)\mathbf{C}\mathbf{U}$. Since the training matrix \mathbf{C}_b is known at the receiver, the channel matrix realization can be estimated as

$$\hat{\mathbf{H}}_{\text{ST}} = \tilde{\mathbf{R}}\mathbf{C}_b^{-1}. \quad (25)$$

From (24) and (25), the channel estimation error is given by

$$\hat{\mathbf{H}}_{\text{ST}} - \mathbf{H} = \mathbf{H}\tilde{\mathbf{B}}\mathbf{C}_b^{-1} + \tilde{\mathbf{N}}\mathbf{C}_b^{-1}. \quad (26)$$

Since \mathbf{C}_b is an orthogonal matrix that satisfies $\mathbf{C}_b\mathbf{C}_b^H = n_T P_c \mathbf{I}_{n_T}$ then $\mathbf{C}_b^{-1} = (1/n_T P_c)\mathbf{C}_b^H$. Therefore, using this result in (26), the channel estimate (for the ST case) can be expressed as

$$\begin{aligned} \hat{\mathbf{H}}_{\text{ST}} &= \frac{1}{n_T P_c} \left[\mathbf{H} (\tilde{\mathbf{B}} + \mathbf{C}_b) + \tilde{\mathbf{N}} \right] \mathbf{C}_b^H \\ &= \mathbf{H} + \frac{1}{n_T P_c} (\mathbf{H}\tilde{\mathbf{B}} + \tilde{\mathbf{N}}) \mathbf{C}_b^H. \end{aligned} \quad (27)$$

The MSE of the ST case can be found in [11]. It may be inferred from (27) that $\tilde{\mathbf{B}}$ represents an extra term for the channel estimate. This time-average term is exploited by the following method.

3.3. Data-Dependent ST Case. In DDST, the signal $\tilde{\mathbf{B}} = (1/Q)\mathbf{B}\mathbf{U}$ is used to define the perturbation matrix

$$\mathbf{E} = -(\mathbf{u} \otimes \tilde{\mathbf{B}}), \quad (28)$$

which will be arithmetically added to the data signal \mathbf{B} , in every transmitted block. Hence the transmitted signal is determined by $\mathbf{X} = \mathbf{B} + \mathbf{E} + \mathbf{C}$, and the corresponding received signal is given by

$$\mathbf{R} = \mathbf{H} (\mathbf{B} + \mathbf{E} + \mathbf{C}) + \mathbf{N}. \quad (29)$$

It is important to stress that the term $\tilde{\mathbf{B}}$ will be nonexistent at the receiver end, because it was removed at the transmitter side when the signal \mathbf{E} was added. Hence, using (25) and (27), the LS time-domain channel estimator based on the synchronized averaging of the received signal \mathbf{R} is given by

$$\hat{\mathbf{H}}_{\text{DDST}} = \mathbf{H} + \left(\frac{1}{n_T P_c} \right) \tilde{\mathbf{N}}\mathbf{C}_b^H. \quad (30)$$

Subsequently, the mean-square channel estimation error under optimal training with $\mathbf{C}_b\mathbf{C}_b^H = n_T P_c \mathbf{I}_{n_T}$ is given by

$$\text{MSE}_{\text{DDST}}(\hat{\mathbf{H}}) = E \left\{ \left\| \hat{\mathbf{H}}_{\text{LS}} - \mathbf{H} \right\|_F^2 \right\} = E \left\{ \left\| \tilde{\mathbf{N}}\mathbf{C}_b^{-1} \right\|_F^2 \right\}. \quad (31)$$

Using $\tilde{\mathbf{N}} = (1/Q)\mathbf{N}\mathbf{U}$, the DDST mean-square estimation error may be expressed as

$$\text{MSE}_{\text{DDST}}(\hat{\mathbf{H}}) = E \left\{ \left\| \frac{1}{Q} (\mathbf{N}\mathbf{U}) \mathbf{C}_b^{-1} \right\|_F^2 \right\}; \quad (32)$$

by replacing the inverse of \mathbf{C}_b , $\|\mathbf{A}\|_F^2 = \text{Tr}\{\mathbf{A}^H\mathbf{A}\}$ and $E\{\mathbf{N}^H\mathbf{N}\} = n_R \sigma_n^2 \mathbf{I}$, then

$$\text{MSE}_{\text{DDST}}(\hat{\mathbf{H}}) = \frac{n_R \sigma_n^2}{(n_T P_c Q)^2} \text{Tr} \left\{ \mathbf{C}_b \mathbf{U}^H \mathbf{U} \mathbf{C}_b^H \right\}, \quad (33)$$

which is finally reduced to

$$\text{MSE}_{\text{DDST}}(\hat{\mathbf{H}}) = \frac{n_R \sigma_n^2}{(n_T P_c Q)^2} Q (n_T n_T P_c) = \frac{n_T n_R \sigma_n^2}{N P_c}. \quad (34)$$

In order to have the same estimation error performance between the DDST method and the TDM-based estimate, we compare (16) and (34), and the resulting relation is given by

$$\frac{P_c}{P_x} = \frac{T}{N}. \quad (35)$$

Similar expressions for (34) and (35) are obtained using another procedure in [11].

In DDST, the data-dependent term has the purpose of removing the dependency on data; therefore, certain frequency components are removed at the transmitter and data symbols may be distorted. This data distortion becomes a problem at the receiver when its value is commensurable with the Euclidean distance between the symbols of the constellation used. It is more significant as the order of modulation

increases, at fixed transmission power. In this case, the receiver would not be able to identify the signals from the constellation. This problem was highlighted and designated the data identifiability problem in [26]. In preliminary DDST works, such as [7, 8], the data distortion is removed in the receiver using an iterative method that performs effectively for low orders of modulations. Recently in [26], another iterative removal scheme is proposed for high orders of modulations. In what follows, we will show a hybrid training method that overcomes this impairment by combining ST and PAT.

4. Iterative Joint MRST-PAT

Even though the DDST channel estimate is better (according to the LS criterion) than the ST estimate, the DDST technique presents several disadvantages compared to ST, as mentioned in the Introduction. Here, we want to stress only its main drawbacks: it cannot work with high-order constellations, and it is highly sensitive to block length. For these main reasons, it is desirable to find ST improvements that do not present these disadvantages.

The MRST scheme in [28, 37] was developed from the following two points: (i) the fact that the difference between ST (27) and DDST (30) channel estimation techniques is the factor $\mathbf{H}\tilde{\mathbf{B}}$ and (ii) the hypothesis that if we could obtain an estimate of the signal \mathbf{E} at the receiver end, we would achieve the performance of DDST in terms of channel estimate MSE. However, as more power is allocated to data signals, BER performance should increase.

Assuming this hypothesis, the signal \mathbf{E} can be approached as $\mathbf{E} \simeq -(\mathbf{u} \otimes \hat{\tilde{\mathbf{B}}})$ with $\hat{\tilde{\mathbf{B}}} = (1/Q)\hat{\mathbf{B}}\mathbf{U}$, where $\hat{\mathbf{B}}$ represents the data estimate computed at the receiver. Then, using (24) and (25), the MRST-PAT channel estimate is given by

$$\begin{aligned} \hat{\mathbf{H}}_{\text{MRST-PAT}} &= \tilde{\mathbf{R}}\mathbf{C}_b^{-1} = \left(\mathbf{H}(\tilde{\mathbf{B}} + \mathbf{C}_b) - \hat{\mathbf{H}}\hat{\tilde{\mathbf{B}}} + \tilde{\mathbf{N}} \right) \mathbf{C}_b^{-1} \\ &= \mathbf{H} + \frac{1}{n_T P_c} \left[(\mathbf{H}\tilde{\mathbf{B}} - \hat{\mathbf{H}}\hat{\tilde{\mathbf{B}}}) + \tilde{\mathbf{N}} \right] \mathbf{C}_b^H; \end{aligned} \quad (36)$$

it follows that if $\hat{\tilde{\mathbf{B}}} = \mathbf{H}\tilde{\mathbf{B}}$ then $\hat{\mathbf{H}}_{\text{MRST-PAT}} = \hat{\mathbf{H}}_{\text{DDST}}$ and, subsequently, $\text{MSE}_{\text{MRST-PAT}}(\hat{\mathbf{H}}) = \text{MSE}_{\text{DDST}}(\hat{\mathbf{H}})$.

As its name indicates, the MRST algorithm has the goal of removing the mean of the data-bearing signal at the receiver in an iterative way. The structure of this algorithm offers another benefit.

The initial channel estimation may be performed in one of several ways: employing the ST scheme, designating a few pilots according to [36] at the beginning of each block transmitted in the TDM scheme, or using the expectation maximization (EM) algorithm with iterative decoding strategies [38].

In order to make a reliable estimate ($\hat{\tilde{\mathbf{B}}}$) of $\tilde{\mathbf{B}}$, we consider the joint effort of PAT (TDM) and ST techniques. This is illustrated in Figure 3, where \mathbf{P} represents $n_T \times T$ matrix of pilot symbols. The pilots are considered to have a preliminary channel estimate and consequently an initial estimate, $\hat{\tilde{\mathbf{B}}}$, of $\tilde{\mathbf{B}}$; then the superimposed training signal is used in the

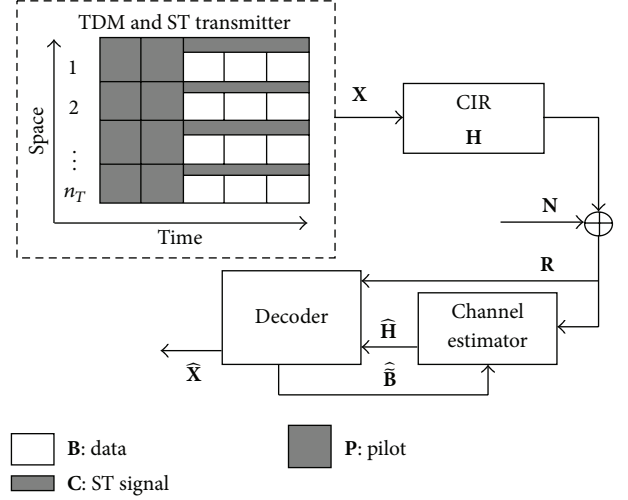


FIGURE 3: Block diagram of a transceiver with MRST-PAT.

refinement phase to improve upon the initial estimate. The following steps describe the iterative detection and channel estimation procedure of the joint MRST-PAT technique:

- (1) Set N, T and compute $L = N - T$. In this case $Q = L/n_T$.
- (2) Set the number of iterations denoted by numIterST for the ST phase.
- (3) Use (14) to obtain a preliminary channel estimate using the pilots in the TDM case.
- (4) Use the channel estimated to obtain $\hat{\tilde{\mathbf{B}}}$, which represents an estimate of the transmitted symbols.
- (5) If numIterST > 0 , compute the signal $\hat{\tilde{\mathbf{B}}}$ with $\hat{\tilde{\mathbf{B}}} = (1/Q)\hat{\mathbf{B}}\mathbf{U}$, else end.
- (6) Compute the time-average of the received signal \mathbf{R} and make $\tilde{\mathbf{R}}^{\text{old}} = \tilde{\mathbf{R}}$.
- (7) Remove $\hat{\tilde{\mathbf{B}}}$ from the received signal to obtain new $\tilde{\mathbf{R}}^{\text{new}}$ according to

$$\tilde{\mathbf{R}}^{\text{new}} = \tilde{\mathbf{R}}^{\text{old}} - \hat{\tilde{\mathbf{B}}}. \quad (37)$$

- (8) Use (36) with $\tilde{\mathbf{R}} = \tilde{\mathbf{R}}^{\text{new}}$ to obtain the channel estimate $\hat{\mathbf{H}}_{\text{MRST-PAT}}$.
- (9) Decrease the value of numIterST.
- (10) Go to step (4).

5. MRST-PAT for Alamouti Space-Time Coding

In this section, we will show the use of the proposed MRST-PAT channel estimation method with $\mathcal{E}2$ -OSTBC system (Alamouti scheme), using $n_T = 2$ and $n_R = 2$. This implementation assumes N and T to be even, so $2 \times N$

transmitted matrix is given by $\mathbf{X} = [\mathbf{G}_1 \ \mathbf{G}_2 \ \cdots \ \mathbf{G}_\Delta]$. In extended form

$$\mathbf{X} = \begin{bmatrix} s_1 & -s_2^* & s_3 & -s_4^* & \cdots & s_{N-1} & -s_N^* \\ \underbrace{s_2 & s_1^*}_{\mathbf{G}_1} & \underbrace{s_4 & s_3^*}_{\mathbf{G}_2} & \cdots & \underbrace{s_N & s_{N-1}^*}_{\mathbf{G}_\Delta} \end{bmatrix}. \quad (38)$$

The matrix's rows and columns indicate the transmit antenna and the time slot (channel use), respectively.

$2 \times T$ pilot matrix \mathbf{P} is given by

$$\mathbf{P} = \begin{bmatrix} p_1 & -p_2^* & p_3 & -p_4^* & \cdots & p_{T-1} & -p_T^* \\ p_2 & p_1^* & p_4 & p_3^* & \cdots & p_T & p_{T-1}^* \end{bmatrix}. \quad (39)$$

The entries of the pilot matrix are given by the Walsh-Hadamard bipolar spreading sequences [39]. The Walsh-Hadamard sequences of length M , with $M = 2^m$ and $m \in \mathbb{N}$, are often defined using Hadamard matrices \mathbf{W}_M [2] with

$$\mathbf{W}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad (40)$$

$$\mathbf{W}_{2M} = \begin{bmatrix} \mathbf{W}_M & \mathbf{W}_M \\ \mathbf{W}_M & -\mathbf{W}_M \end{bmatrix}.$$

The resulting matrices \mathbf{W}_M are orthogonal matrices; that is, for every M we have

$$\mathbf{W}_M \mathbf{W}_M^T = M \mathbf{I}_M, \quad (41)$$

where \mathbf{I}_M is the M -dimensional identity matrix. Additionally, the structure of \mathbf{W}_2 assures that the rows (or columns) of a Hadamard matrix are mutually orthogonal.

$2 \times L$ data-bearing matrix \mathbf{B} is given by

$$\mathbf{B} = \begin{bmatrix} b_1 & -b_2^* & b_3 & -b_4^* & \cdots & b_{L-1} & -b_L^* \\ b_2 & b_1^* & b_4 & b_3^* & \cdots & b_L & b_{L-1}^* \end{bmatrix}. \quad (42)$$

In this case, the matrix \mathbf{U} has a dimension of $L \times 2$ and therefore $\tilde{\mathbf{B}} = (1/Q)\mathbf{B}\mathbf{U}$ is a 2×2 matrix.

Assuming a frequency-flat fading scenario [36], \mathbf{C}_b matrix can be composed of two symbols c_1, c_2 and two other symbols generated through the Alamouti scheme, exploiting the property of orthogonality. In this way

$$\mathbf{C}_b = \begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix}. \quad (43)$$

Likewise, $2 \times L$ ST matrix \mathbf{C} is given by

$$\mathbf{C} = \begin{bmatrix} c_1 & -c_2^* & c_1 & -c_2^* & \cdots & c_1 & -c_2^* \\ \underbrace{c_2 & c_1^*}_{\mathbf{C}_b} & \underbrace{c_2 & c_1^*}_{\mathbf{C}_b} & \cdots & \underbrace{c_2 & c_1^*}_{\mathbf{C}_b} \end{bmatrix} \quad (44)$$

and therefore the two row vectors of matrix \mathbf{C} are orthogonal.

The initial channel estimate (step (3) of the iterative algorithm) is performed using the symbol pilot matrix \mathbf{P}

with the TDM scheme. An improved estimate of the channel matrix \mathbf{H} will be made using the ST time-domain estimator pursuant to the iterative algorithm described in Section 4.

Let us estimate the channel matrix \mathbf{H} . For the sake of simplicity, we restricted the analysis to a noise-free scenario. From (24), the time-average of the received signal is given by

$$\text{vec}\{\tilde{\mathbf{R}}\} = \begin{bmatrix} h_{11}(\tilde{b}_1 + c_1) + h_{12}(\tilde{b}_2 + c_2) \\ h_{21}(\tilde{b}_1 + c_1) + h_{22}(\tilde{b}_2 + c_2) \\ -h_{11}(\tilde{b}_2 + c_2)^* + h_{12}(\tilde{b}_1 + c_1)^* \\ -h_{21}(\tilde{b}_2 + c_2)^* + h_{22}(\tilde{b}_1 + c_1)^* \end{bmatrix}, \quad (45)$$

where $\text{vec}\{\cdot\}$ is the column-wise vectorization operator. Subsequently, the expression in (45) can be reformulated as

$$\text{vec}\{\tilde{\mathbf{R}}\} = \mathbf{A} \text{vec}\{\mathbf{H}^T\} \quad (46)$$

$$\text{with } \mathbf{A} = \begin{bmatrix} (\tilde{b}_1 + c_1) & (\tilde{b}_2 + c_2) & 0 & 0 \\ 0 & 0 & (\tilde{b}_1 + c_1) & (\tilde{b}_2 + c_2) \\ -(\tilde{b}_2 + c_2)^* & (\tilde{b}_1 + c_1)^* & 0 & 0 \\ 0 & 0 & -(\tilde{b}_2 + c_2)^* & (\tilde{b}_1 + c_1)^* \end{bmatrix}.$$

An estimate of \mathbf{A} can be used for the purpose of avoiding the computational complexity of $\widehat{\mathbf{H}}\tilde{\mathbf{B}}$ in step (7); therefore,

$$\hat{\mathbf{A}} = \begin{bmatrix} (\hat{\tilde{b}}_1 + c_1) & (\hat{\tilde{b}}_2 + c_2) & 0 & 0 \\ 0 & 0 & (\hat{\tilde{b}}_1 + c_1) & (\hat{\tilde{b}}_2 + c_2) \\ -(\hat{\tilde{b}}_2 + c_2)^* & (\hat{\tilde{b}}_1 + c_1)^* & 0 & 0 \\ 0 & 0 & -(\hat{\tilde{b}}_2 + c_2)^* & (\hat{\tilde{b}}_1 + c_1)^* \end{bmatrix}. \quad (47)$$

It is important to note that $\hat{\mathbf{A}}$ is an orthogonal matrix. Consequently, the channel estimate matrix is given by

$$\text{vec}\{\hat{\mathbf{H}}^T\} = \hat{\mathbf{A}}^{-1} \text{vec}\{\tilde{\mathbf{R}}\} \quad (48)$$

and, finally,

$$\hat{\mathbf{H}} = \left[\left(\text{vec}\{\mathbf{I}_{n_R}\}^T \otimes \mathbf{I}_{n_T} \right) \left(\mathbf{I}_{n_R} \otimes \text{vec}\{\hat{\mathbf{H}}^T\} \right) \right]^T, \quad (49)$$

where \mathbf{I}_{n_T} and \mathbf{I}_{n_R} are $n_T \times n_T$ and $n_R \times n_R$ identity matrices, respectively. Matrix $\hat{\mathbf{H}}$ represents a new MIMO channel estimate, which is used in the iterative procedure to obtain a new estimate of the arithmetic mean of the data-bearing.

The key idea of this procedure is the removal of the mean at the receiver. Instead of subtracting the mean from the time-average received matrix $\tilde{\mathbf{R}}$, this MRST-PAT implementation incorporates the estimated mean $\hat{\tilde{\mathbf{B}}}$ in the matrix $\hat{\mathbf{A}}$ to obtain an estimate of the channel.

6. Simulation Results

The following section presents the performance of the proposed channel estimation scheme in a frequency-flat fading

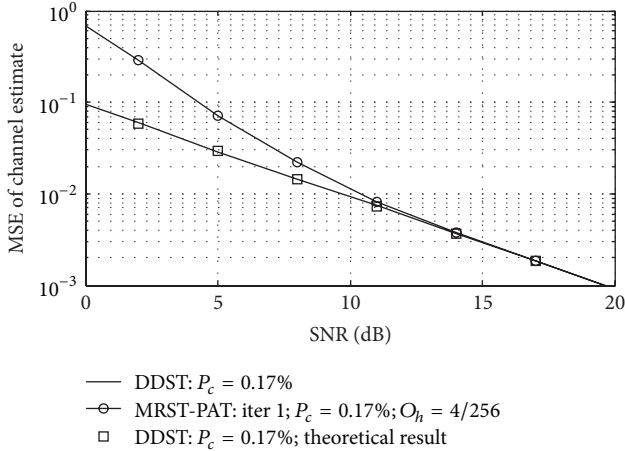


FIGURE 4: MSE of channel estimate exhibited by the $\mathcal{E}2$ -OSTBC 2×2 system employing QPSK.

scenario. A $\mathcal{E}2$ -OSTBC system with $n_T = n_R = 2$ antennas was used. This simulation scenario was also employed in [28]. The block length was fixed to $N = 256$ symbols, unless another value is indicated, and all simulations were run until 1000 block errors were found. The BER is represented as a function of the average SNR, where $N_0 = E_s n_T / \text{SNR}$ and E_s is the average symbol energy. QPSK and M -ary QAM ($M > 4$) constellations were used with Gray-coded symbols, and the transmitted power level for each antenna was normalized to one; for example, $P_x = 1$ (Watt).

To test the channel estimation techniques, we use as a figure of merit the MSE of channel estimate. Figure 4 plots this metric for DDST and MRST-PAT using QPSK. Both of these use $P_c = 17\%$ of the transmitted power P_x . This power level was chosen using the values of previous works such as [6, 15, 28].

The initial channel estimation for the MRST-PAT scheme was implemented using $T = 4$ pilots per block transmitted (remember that an Alamouti space-time coding scheme is considered); hence, the overhead associated with the pilots is $O_h = T/N = 4/256$, which is 1.56%. For the ST stage, $L = 252$ symbols are used.

Through an analysis of this performance, it can be noted that the proposed scheme approaches DDST for SNR values higher than 10 dB, as was anticipated. However, it is worth highlighting that although the MSE channel estimate in MRST-PAT is not close to DDST at low SNR levels, it is good enough to provide a close BER performance, which is the ultimate metric for system comparison.

Before considering the BER performance, let us discuss Figure 5, which shows the MSE versus power fraction for training signal, with the SNR fixed at 15 dB and QPSK signals. Three different channel estimation schemes (TDM, DDST, and MRST-PAT) were compared. If $P_c/P_x = T/N$, with $T = 4$ and $N = 256$, then the three schemes have the same MSE performance. It can be observed that the proposed MRST-PAT scheme follows DDST for all power levels of P_c . The selected SNR value is based on [11], because it allows a suitable BER.

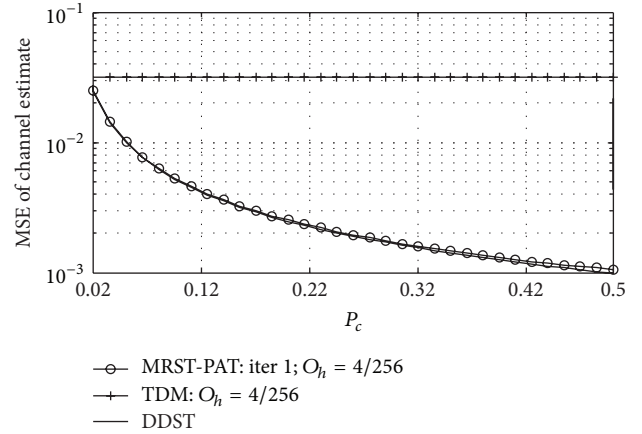


FIGURE 5: MSE of channel estimate versus power fraction for training exhibited by the $\mathcal{E}2$ -OSTBC 2×2 , at SNR = 15 dB employing QPSK.

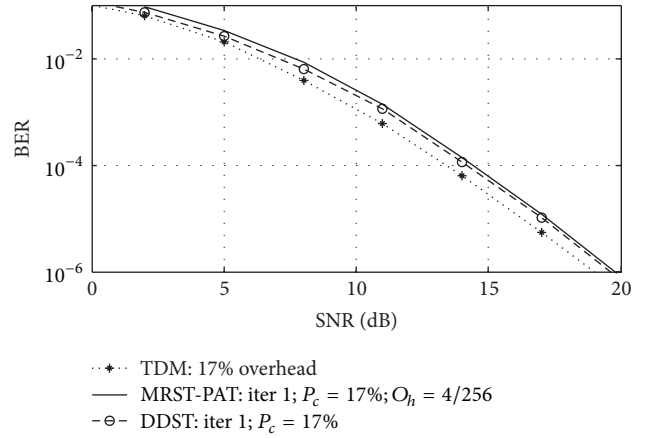


FIGURE 6: BER performance exhibited by the $\mathcal{E}2$ -OSTBC 2×2 system employing QPSK.

The BER performance of this MIMO system working with QPSK signals is shown in Figure 6. It can be seen that the TDM scheme achieves the best BER performance at the cost of 17% of data rate overhead due to pilot transmission, while MRST-PAT, with 1.56% of overhead, for example, $O_h = 4/256$, achieves very close to DDST BER performance.

In this scenario, DDST matches the relation described in (35) for the purpose of having the same MSE performance as TDM; the parameters for MRST-PAT were chosen to have the same training power P_c as DDST and 4 pilots, although under frequency-flat fading conditions the minimum length is n_T [36].

According to [26], when M -ary QAM constellations with $M \geq 16$ are used, there is evidence of DDST performance degradation due to the data identifiability problem, which is discussed in the following paragraph.

In some block realizations, the data-dependent signal distortion can have more power than the difference between two symbols; consequently, the data-bearing signal will be

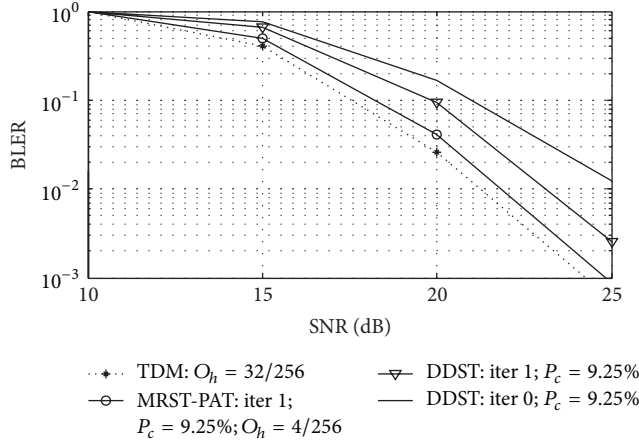


FIGURE 7: BLER performance exhibited by the $\mathcal{E}2$ -OSTBC 2×2 system employing 16 QAM.

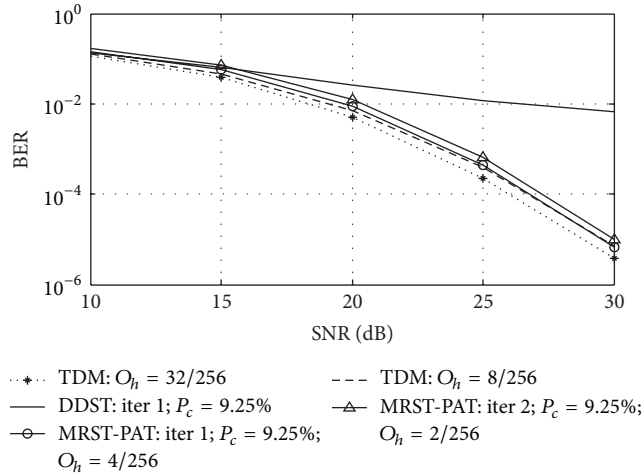


FIGURE 8: BER performance exhibited by the $\mathcal{E}2$ -OSTBC 2×2 system employing 64 QAM.

modified to such a degree that it is not possible to correctly recover it.

Figure 7 shows the block error rate (BLER) performance using 16 QAM signals. In this case, BLER performance is used because it is a relevant measurement used for 3 GPP performance testing requirements, which determine the quality of the radio link. It corresponds to the ratio of the number of blocks received with at least one symbol error to the total number of blocks transmitted.

It can be observed that at fixed BLER of 10^{-2} an advantage close to 1 dB was accomplished by MRST-PAT over DDST, both using one iteration and $P_c = 9.25\%$ of the transmitted power.

Figure 8 plots the BER performance for 64 QAM signals. Here, the advantage of the proposed MRST-PAT algorithm is more evident. An interesting analysis is that MRST-PAT ($O_h = 4/256$ and one iteration) shows the same performance as TDM with $O_h = 8/256$. MRST-PAT ($O_h = 2/256$ and two

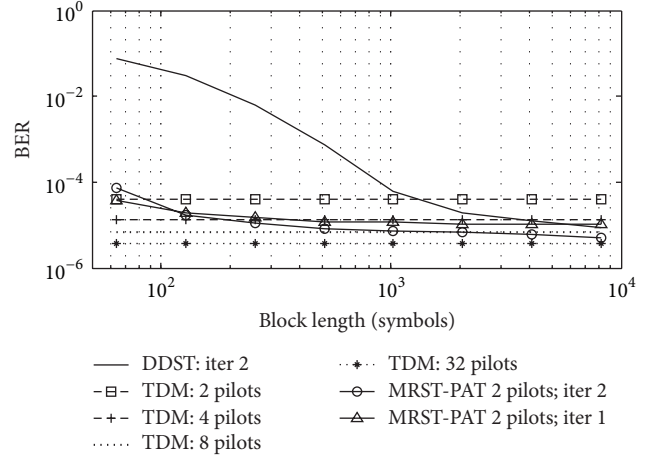


FIGURE 9: BER versus block length of $\mathcal{E}2$ -OSTBC 2×2 system at SNR = 30 dB and $P_c = 7.5\%$ for 64 QAM.

iterations) approaches TDM with $O_h = 8/256$ in 0.5 dB, for SNR levels above 20 dB.

Figure 9 plots the performance of 64 QAM $\mathcal{E}2$ -OSTBC versus the block length for the SNR fixed at 30 dB and $P_c = 7.5\%$ for DDST and MRST-PAT. It can be seen from the figure that DDST for short block length yields poor results and for large block length ($N = 8192$ symbols) approaches TDM with eight pilots, while MRST-PAT for short block length yields good results and the performance increases proportionally as N increases.

It is important to note that DDST's BER performance is highly sensitive to block length. In this scenario, we can appreciate the benefits and flexibility of the MRST-PAT structure, which exploits the presence of the known training symbols and the fact that the ST algorithm achieves a better channel estimate when the block length of the data transmitted is increased. Hence, the system user can change the parameters according to the requirements of block length or less rate loss, for example, fewer dedicated pilots or BER performance.

Finally, it is important to analyze the BER performance versus the power fraction for the training signal, with the SNR fixed at 34 dB and 256 QAM signals. This SNR value is commonly required for this order of modulation. Figure 10 shows that DDST exhibits a poor BER performance for every value of power fraction allocated to the training signal. However, the BER performance exhibited by the MRST-PAT confirms that the proposed method overcomes the identifiability problem and makes the superimposed technique an attractive choice for multiamplitude QAM constellations. The same behavior is found at other SNR levels.

For future research purposes, the proposed approach will be used with more sophisticated equalizers, such as that developed in [24, 25, 28], as well as with channel coders. The results obtained in this future research will allow determining the scenarios under which this new technique represents the best practical solution to be employed.

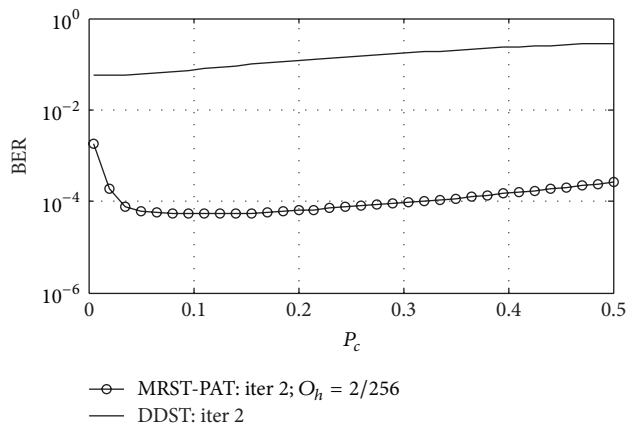


FIGURE 10: BER performance versus power fraction for training exhibited by the $\mathcal{S}2$ -OSTBC 2×2 system at SNR = 34 dB employing 256 QAM.

7. Conclusion

The iterative joint MRST-PAT MIMO detection and channel estimation technique introduced in this contribution deals effectively with DDST's identifiability problem when higher-order modulation schemes are used. This leads to communication systems that achieve near-DDST performance for low-order symbols, such as BPSK and QPSK, and attain PAT (TDM) performance for higher-order symbols, such as 16 QAM or beyond, with minimal reduction in throughput and using a computationally efficient implementation. Since MRST-PAT is an iterative joint scheme, it can take advantage of both efforts, that is, using pilots by means of explicit training (PAT) and taking the benefit of the implicit training using the ST training signal in the transmitted block. Consequently, more flexible communication systems may be built according to different requirements and constraints. Similar results are expected for MIMO frequency-selective channels.

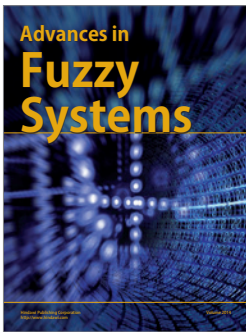
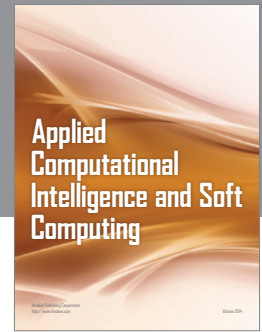
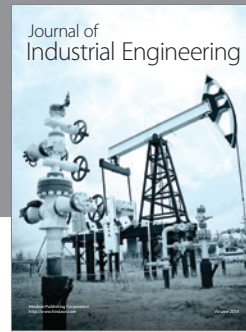
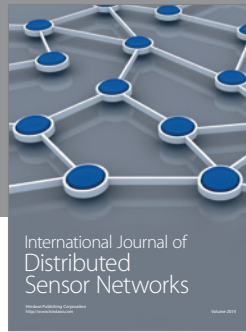
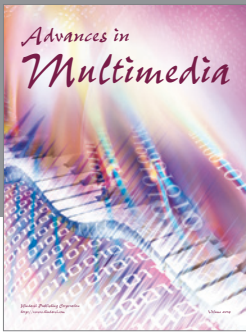
Competing Interests

The authors declare that they have no competing interests.

References

- [1] H. Arslan and G. E. Bottomley, "Channel estimation in narrow-band wireless communication systems," *Wireless Communications and Mobile Computing*, vol. 1, no. 2, pp. 201–219, 2001.
- [2] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels [mobile radio]," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 4, pp. 686–693, 1991.
- [3] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: general model, design criteria, and signal processing," *IEEE Signal Processing Magazine*, vol. 21, no. 6, pp. 12–25, 2004.
- [4] S. Haykin and K. J. R. Liu, *Handbook on Array Processing and Sensor Networks*, Wiley-Blackwell, Hoboken, NJ, USA, 2010.
- [5] B. Farhang-Boroujeny, "Pilot-based channel identification: proposal for semi-blind identification of communication channels," *Electronics Letters*, vol. 31, no. 13, pp. 1044–1046, 1995.
- [6] A. G. Orozco-Lugo, M. M. Lara, and D. C. McLernon, "Channel estimation using implicit training," *IEEE Transactions on Signal Processing*, vol. 52, no. 1, pp. 240–254, 2004.
- [7] M. Ghogho, D. McLernon, E. Alameda-Hernandez, and A. Swami, "Channel estimation and symbol detection for block transmission using data-dependent superimposed training," *IEEE Signal Processing Letters*, vol. 12, no. 3, pp. 226–229, 2005.
- [8] E. Alameda-Hernandez, D. C. McLernon, A. G. Orozco-Lugo, M. M. Lara, and M. Ghogho, "Frame/training sequence synchronization and DC-offset removal for (data-dependent) superimposed training based channel estimation," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2557–2569, 2007.
- [9] E. Romero-Aguirre, R. Carrasco-Alvarez, R. Parra-Michel, A. G. Orozco-Lugo, and A. F. Mondragón-Torres, "Full-hardware architectures for data-dependent superimposed training channel estimation," *Journal of Signal Processing Systems*, vol. 70, no. 2, pp. 105–123, 2013.
- [10] F. Martín del, R. Campo, R. Perez-Andrade, and A. G. Orozco-Lugo, "A system on a programmable chip architecture for data-dependent superimposed training channel estimation," *International Journal of Reconfigurable Computing*, vol. 2009, Article ID 912301, 10 pages, 2009.
- [11] M. Ghogho and A. Swami, "Channel estimation for MIMO systems using data-dependent superimposed training," in *Proceedings of the 42nd Annual Allerton Conference on Communication, Control and Computing*, Urbana Champaign, Ill, USA, 2004.
- [12] M. Ghogho, D. McLernon, E. Alameda-Hernandez, and A. Swami, "SISO and MIMO channel estimation and symbol detection using data-dependent superimposed training," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 3, pp. ii/461–iii/464, IEEE, March 2005.
- [13] S. M. A. Moosvi, D. C. McLernon, E. Alameda-Hernandez, A. G. Orozco-Lugo, and M. M. Lara, "A low complexity iterative channel estimation and equalisation scheme for (data-dependent) superimposed training," in *Proceedings of the 14th European Signal Processing Conference (EUSIPCO '06)*, pp. 1–5, Florence, Italy, September 2006.
- [14] X. Meng and J. K. Tugnait, "Semi-blind channel estimation and detection using superimposed training," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, vol. 4, pp. iv-417–iv-420, IEEE, Montreal, Canada, May 2004.
- [15] R. Carrasco-Alvarez, R. Parra-Michel, A. G. Orozco-Lugo, and J. K. Tugnait, "Enhanced channel estimation using superimposed training based on universal basis expansion," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1217–1222, 2009.
- [16] S. He, J. K. Tugnait, and X. Meng, "On superimposed training for MIMO channel estimation and symbol detection," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 3007–3021, 2007.
- [17] X. Meng, J. K. Tugnait, and S. He, "Iterative joint channel estimation and data detection using superimposed training: algorithms and performance analysis," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 1873–1880, 2007.
- [18] R. Dinis, N. Souto, J. Silva, A. Kumar, and A. Correia, "On the use of implicit pilots for channel estimation with OFDM modulations," in *Proceedings of the IEEE 66th Vehicular Technology Conference (VTC '07)*, pp. 1077–1081, Baltimore, Md, USA, October 2007.

- [19] N. Chen and G. T. Zhou, "Superimposed training for OFDM: a peak-to-average power ratio analysis," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2277–2287, 2006.
- [20] A. Goljahani, N. Benvenuto, S. Tomasin, and L. Vangelista, "Superimposed sequence versus pilot aided channel estimations for next generation DVB-T systems," *IEEE Transactions on Broadcasting*, vol. 55, no. 1, pp. 140–144, 2009.
- [21] H. Zhang, X. Dai, D. Li, and S. Ye, "Linearly time-varying channel estimation and symbol detection for OFDMA uplink using superimposed training," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, no. 1, Article ID 307375, 2009.
- [22] R. Carrasco-Alvarez, R. Parra-Michel, A. G. Orozco-Lugo, and J. K. Tugnait, "Time-varying channel estimation using two-dimensional channel orthogonalization and superimposed training," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4439–4443, 2012.
- [23] X. Dai, H. Zhang, and D. Li, "Linearly time-varying channel estimation for MIMO/OFDM systems using superimposed training," *IEEE Transactions on Communications*, vol. 58, no. 2, pp. 681–693, 2010.
- [24] R. Dinis, C.-T. Lam, and D. Falconer, "Joint frequency-domain equalization and channel estimation using superimposed pilots," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '08)*, pp. 447–452, IEEE, Las Vegas, Nev, USA, 2008.
- [25] N. Benvenuto, R. Dinis, D. Falconer, and S. Tomasin, "Single carrier modulation with nonlinear frequency domain equalization: an idea whose time has come—again," *Proceedings of the IEEE*, vol. 98, no. 1, pp. 69–96, 2010.
- [26] T. Whitworth, M. Ghogho, and D. C. McLernon, "Data identifiability for data-dependent superimposed training," in *Proceedings of the IEEE International Conference on Communications (ICC '07)*, pp. 2545–2550, IEEE, Glasgow, Scotland, June 2007.
- [27] T. Levanen, J. Talvitie, and M. Renfors, "Performance evaluation of time-multiplexed and data-dependent superimposed training based transmission with practical power amplifier model," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, article 49, 2012.
- [28] O. Longoria-Gandara, R. Parra-Michel, M. Bazdresch, and A. G. Orozco-Lugo, "Iterative mean removal superimposed training for SISO and MIMO channel estimation," *International Journal of Digital Multimedia Broadcasting*, vol. 2008, Article ID 535269, 9 pages, 2008.
- [29] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proceedings of the URSI International Symposium on Signals, Systems, and Electronics (ISSSE '98)*, pp. 295–300, October 1998.
- [30] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.
- [31] A. Gershman and N. Sidiropoulos, *Space-Time Processing for MIMO Communications*, John Wiley & Sons, New York, NY, USA, 2005.
- [32] M. Coldrey and P. Bohlin, "Training-based MIMO systems—part I: performance comparison," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5464–5476, 2007.
- [33] J. K. Tugnait and W. Luo, "On channel estimation using superimposed training and first-order statistics," *IEEE Communications Letters*, vol. 7, no. 9, pp. 413–415, 2003.
- [34] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1st edition, 1993.
- [35] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 884–893, 2006.
- [36] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, 2003.
- [37] O. Longoria-Gandara, R. Parra-Michel, M. Bazdresch, and A. G. Orozco-Lugo, "Iterative mean removal superimposed training for frequency selective channel estimation," in *Proceedings of the International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM '08)*, pp. 1–5, Dalian, China, October 2008.
- [38] S. Y. Ronen, S. I. Bross, S. S. Shitz, and T. M. Duman, "Iterative channel estimation and decoding in turbo coded space-time systems," *European Transactions on Telecommunications*, vol. 18, no. 7, pp. 719–734, 2007.
- [39] X. Huang, "Complementary properties of hadamard matrices," in *Proceedings of the International Conference on Communications, Circuits and Systems (ICCCAS '06)*, vol. 1, pp. 588–592, Guilin, China, June 2006.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

