Space-Time-Frequency Channel Estimation for Multiple-Antenna Orthogonal Frequency Division Multiplexing Systems

Kar Lun (Clarence) Wong

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

August 2007

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of Master of Engineering.

Your file Votre référence
ISBN: 978-0-494-38504-3
Our file Notre référence
ISBN: 978-0-494-38504-3

Abstract

We propose a linear mean square error channel estimator that exploits the joint space-time-frequency (STF) correlations of the wireless fading channel for applications in multiple-antenna orthogonal frequency division multiplexing systems. Our work generalizes existing channel estimators to the full dimensions including transmit spatial, receive spatial, time, and frequency. This allows versatile applications of our STF channel estimator to any fading environment, ranging from spatially-uncorrelated slow-varying frequency-flat channels to spatially-correlated fast-varying frequency-selective channels.

The proposed STF channel estimator reduces to a time-frequency (TF) channel estimator when no spatial correlations exist. In another perspective, the lower-dimension TF channel estimator can be viewed as an STF channel estimator with spatial correlation mismatch for space-time-frequency selective channels.

Computer simulations were performed to study the mean-square-error (MSE) behavior with different pilot parameters. We then evaluate the suitability of our STF channel estimator on a space-frequency block coded OFDM system. Bit error rate (BER) performance degradation, with respect to perfect coherent detection, is limited to less than 2 dB at a BER of 10^{-5} in the modified 3GPP fast-fading suburban macro environment. Modifications to the 3GPP channel involves reducing the base station angle spread to imitate a high transmit spatial correlation scenario to emphasize the benefit of exploiting spatial correlation in our STF channel estimator.

Sommaire

Nous proposons un estimateur de canal linéaire au sens de l'erreur quadratique qui exploite les corrélations espace-temps-fréquence (ETF) d'un canal sans fil à évanouissement aux fins d'applications de multiplexage par répartition en fréquence sur plusieurs antennes. Notre travail généralise certains estimateurs de canal existants pour utiliser toutes les dimensions incluant l'espace (transmission et réception), le temps et la fréquence. Ceci nous permet d'utiliser notre estimateur ETF dans n'importe quel environnement à évanouissement : des canaux non corrélés à changements lents et réponse fréquentielle uniforme, jusqu'aux canaux à corrélation spatiale à changement rapide et à réponse fréquentielle variable.

L'estimateur ETF que nous proposons se réduit à un estimateur temps-fréquence (TF) quand il n'y a pas de corrélation spatiale. De plus, l'estimateur TF peut être vu comme un estimateur ETF avec une désadaptation à la corrélation spatiale pour les canaux à sélection ETF. Nous avons effectué des simulations afin d'étudier l'erreur quadratique moyenne pour différents paramètre de pilotes. Nous avons ensuite évalué l'utilité de notre estimateur pour un système de multiplexage par répartition orthogonale de la fréquence utilisant un code complet espace-fréquence. Le taux d'erreur sur les bits n'était qu'à 2dB du taux associé à la détection cohérente parfaite, lorsque celle-ci mène à un taux d'erreur sur les bits de 10^{-5} pour un environnement suburbain à évanouissement rapide 3GPP modifié. La modification apportée fut la réduction de l'angle de transmission du point d'accès sans fil afin d'imiter un scénario ou la corrélation spatiale de transmission est forte, ce qui met en valeur les

avantages d'utiliser la corrélation spatiale dans notre estimateur.

Acknowledgments

First and foremost, I would like to thank my supervisor, Professor H. Leib, for his guidance, patience, and encouragement throughout this work. This work would not be possible without his leadership and tolerance.

Second, I would like to send my gratitude to my colleagues and friends for their technical and emotional support. Of particular mention are Chao-Cheng Tu, Usa Vilaipornsawai, and Marthe Kassouf for sharing their knowledge that contributed to this work, and Isabel Deslauriers for translating my Abstract into French.

Third, I am grateful for Professor H. Leib for his financial support from his Nortel Networks and NSERC grants.

Last but not least, I would like to thank my parents for their endless support towards my graduate studies at McGill University.

Contents

List of Figures

List of Tables

List of Acronyms

3GPP	Third Generation Partnership Project
ADSL	Asymmetric Digital Subscriber Line
AoA	Angle-of-Arrival
AoD	Angle-of-Departure
AS	Angle Spread
AWGN	Additive White Gaussian Noise
BER	Bit Error rate
BS	Base Station
CFR	Channel Frequency Response
CIR	Channel Impulse Response
CLT	Central Limit Theorem
CP	Cyclic Prefix
CT	Continuous-Time
DAB	Digital Audio Broadcasting
DFT	Discrete Fourier Transform
DT	Discrete-Time
DVB-T	Digital Video Broadcasting Terrestrial Television
EVD	Eigen-Value Decomposition
IDFT	Inverse Discrete Fourier Transform
ISI	Inter-Symbol-Interference

LD	Linear Dispersion
LMMSE	Linear Minimum Mean Squared-Error
LS	Least-Squares
LTV	Linear Time-Variant
MIMO	Multiple-Input Multiple-Output
ML	Maximum Likelihood
MMSE	Minimum Mean Square-Error
MRC	Maximum Ratio Combining
MS	Mobile Station
MSE	Mean Square-Error
OFDM	Orthogonal Frequency Division Multiplexing
OSTBC	Orthogonal Space-Time Block Code
P/S	Parallel-to-Serial
PACE	Pilot-symbol Assisted Channel Estimation
PDF	Probability Density Function
PSAM	pilot symbol assisted modulation
PSNR	Pilot-Signal-to-Noise Ratio
RMS	Root Mean Square
Rx	Receiver
S/P	Serial-to-Parallel
SFBC-OFDM	Space-Frequency Block Coded OFDM
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
STBC	Space-Time Block Code
STBC-OFDM	Space-Time Block Coded OFDM
STF	Space-Time-Frequency
STFBC-OFDM	Space-Time-Frequency Block Coded OFDM

STTC	Space-Time Trellis Code
TF	Time-Frequency
Tx	Transmitter
ULA	Uniform Linear Array
WCDMA	Wideband Code Division Multiple Access
WSSUS-HO	Wide Sense Stationary Uncorrelated Scattering Homogeneous Channels

List of Symbols

$\{\cdot\}^{\mathcal{F}}$	Frequency domain representation of the corresponding signal
$\{\cdot\}^{\mathcal{T}}$	Time domain representation of the corresponding signal
$\{\cdot\}_d$	Data portion of the corresponding vector
$\{\cdot\}^*$	Matrix/vector complex conjugate
$\{\cdot\}^H$	Matrix/vector complex conjugate transpose (Hermitian transpose)
$\{\cdot\}^T$	Matrix/vector transpose
$\{\cdot\}_p$	Pilot portion of the corresponding vector
$*$	Convolution operator
∂	Differential operator
$\otimes$	Kronecker product
∇	Gradient operator
$\alpha_n(t)$	Complex-valued attenuation function of the n^{th} path
$c_{N_{\mathbf{Rx}}}(\tau)$	Receive filter impulse response
$c_{N_{\mathbf{Tx}}}(\tau)$	Transmit filter impulse response
$\delta(\cdot)$	Unit impulse function
$\mathrm{diag}(\cdot)$	Diagonal matrix operator, placing the vector argument onto the diagonal of a square matrix of appropriate size

ΔD_p	Normalized pilot symbol space separation, with respect to transmit antenna element spacing
Δd_p	Pilot symbol space separation
$D_{p,\text{offset}}$	Pilot transmit space origin
$D_{p,\text{span}}$	Pilot transmit space span
d_{Rx}	Receive antenna element spacing
d_{Tx}	Transmit antenna element spacing
$D_{u,\text{offset}}$	User transmit space origin
$D_{u,\text{span}}$	User transmit space span
$E\{\cdot\}$	Statistical expectation operator
$\mathbf{e}$	Estimation error vector
$\mathbf{F}$	Normalized DFT matrix
f_c	Carrier (center) frequency
Δf	OFDM sub-carrier spacing
Δf_p	Pilot symbol frequency separation
$F_{p,\text{offset}}$	Pilot frequency origin
$F_{p,\text{span}}$	Pilot frequency span
$F_{u,\text{offset}}$	User frequency origin
$F_{u,\text{span}}$	User frequency span
$\mathbf{g}$	Discrete-time channel impulse response vector
$g_c(t,\tau)$	Overall channel impulse response (includes transmit and receive filters
$\tilde{g}_c(t,\tau)$	Discrete multipath channel impulse response
$\mathbf{h}$	Channel frequency response vector
$\hat{\mathbf{h}}$	Channel frequency response estimate
$\mathbf{I}$	Identity matrix
j	The complex number, $\sqrt{-1}$

ΔF_p	Normalized pilot symbol frequency separation, with respect to OFDM sub-carrier spacing
λ	Wavelength of the propagating electro-magnetic wave at carrier frequency
L_B	STBC block length
L_p	Pilot sequence length
$\overline{\text{MSE}}$	Average MSE per STF coordinate
$\widetilde{M}$	Number of space-time-frequency slots per user subspace
N	Number of discrete multipaths
N_f	Number of OFDM subcarriers
N_p	Number of pilot symbols per transmission frame over all transmit antenna elements
N'_p	Number of significant pilots in the low rank estimation case
ΔT_p	Normalized pilot symbol time separation, with respect to OFDM word duration (include guard time)
$T_{p,\text{offset}}$	Pilot time origin
$T_{p,\text{span}}$	Pilot time span
$\widetilde{N_p}$	Number of pilot symbols per user subspace
$N_{\mathbf{Rx}}$	Number of receive antennas
N_s	Number of source symbols in a STBC block
N_t	Number of OFDM symbols in a transmission frame
N_t	Number of OFDM words per transmission frame
$N_{\mathbf{Tx}}$	Number of transmit antennas
$\nu_{\max}$	Maximum Doppler frequency
$T_{u,\text{offset}}$	User time origin

$T_{u,\text{span}}$	User time span
$\mathbf{p}$	Pilot symbols block
$\phi_n(t)$	Phase shift due to scattering on the n^{th} path
$\mathbf{R}_{\mathbf{hy}_p}$	Cross-correlation matrix between the channel response and the received pilot symbols
$\mathbf{R}_{\mathbf{y}_p}$	Correlation matrix of the received pilot symbols
s_n	Source symbols
t	Observation time
τ	Running propagation delay variable
$\tau_{\max}$	Maximum propagation delay
$\tau_n(t)$	Propagation delay of the n^{th} path
T_c	Coherence time
T_g	Cyclic prefix time duration (or guard time) for OFDM
θ_l	Lower bound on the (azimuth) angle of departure
θ_u	Upper bound on the (azimuth) angle of departure
T_{OFDM}	OFDM word duration (including guard time)
Δt_p	Pilot symbol time separation
$\text{tr}(\cdot)$	Trace operator
T_u	Useful OFDM word duration (portion that contains source symbols, excluding guard time)
$\text{vec}(\cdot)$	Vec operator
$\mathbf{v}^i$	Filtered zero-mean additive noise vector at the i^{th} receive antenna
$v_{\max}$	Maximum velocity at mobile station
$\mathbf{v}_p$	Zero-mean additive noise vector at pilot locations

$\mathbf{W}$	Linear estimator for obtaining channel estimates from the received pilot symbols
$\mathbf{X}$	Source symbol matrix
$\mathbf{x}^j$	Coded symbol vector at the j^{th} transmit antenna
$\mathbf{X}_p$	Pilot symbol matrix
$\mathbf{y}^i$	Received symbol vector at the i^{th} receive antenna
$\mathbf{y}_p$	Received vector containing only pilot symbols

Chapter 1

Introduction

In a traditional single-antenna system, transmission rate is limited by the adverse effects of the wireless media such as multipath fading and Doppler shifts. Fortunately, Teletar [1] and Foschini [2] illustrated that higher data rates can be achieved by employing multiple antennas for transmission and/or reception. These promising capacity results sparked an enormous growth in the field of space-time coding, which combines spatial diversity and channel coding concepts, to deliver high data rate transmission. For example, numerous research has been reported in space-time block code (STBC), a subject pioneered by Alamouti [3], and in space-time trellis code (STTC), first presented by Tarokh et al. [4]. The success of Alamouti's STBC is marked by its adoption in the current wireless standards such as IEEE 802.16 (WiMAX) [5] and wideband code division multiple access (WCDMA) [6].

Most of the existing space-time coding schemes were designed for frequency-flat fading channels. However, as data rate increases, the wireless channel inevitably becomes frequency-selective. Orthogonal frequency division multiplexing (OFDM) is an effective method in handling frequency-selective fading by converting a wideband frequency-

selective channel into a set of narrowband frequency-flat subchannels [7, 8], such that complicated time-domain equalizers [9] can be replaced by simple one-tap frequency-domain equalizers per subchannel [7]. This attractive feature of OFDM made it a widely accepted technology in both wireline and wireless communication standards such as asymmetric digital subscriber line (ADSL) [10], IEEE 802.11a/g (WiFi) [11, 12], IEEE 802.16 (WiMAX) [5], digital audio broadcasting (DAB) [13], and digital video broadcasting terrestrial television (DVB-T) [14].

A natural solution to providing high data rate services in a time-varying multipath environment is the combination of space-time coding with OFDM. For example, Lee presented a space-time block coded OFDM (STBC-OFDM) system in [15], where he modified Alamouti's STBC to encode successive OFDM *vector* symbols instead of source *scalar* symbols. He showed that the two-transmitter, one-receiver, STBC-OFDM achieves the same diversity order, but incurs a 3-dB performance loss in bit error rate (BER) due to limited transmission power, when compared to the one-transmitter, two-receiver, OFDM system with maximum ratio combining (MRC) receiver diversity [15]. This behaviour is consistent with Alamouti's finding for Rayleigh flat-fading channels [3]. In addition to STBC-OFDM, Lee exploited OFDM's frequency multiplexing feature and introduced space-frequency block coded OFDM (SFBC-OFDM) in [16], where Alamouti's STBC is first applied to the source symbols, then OFDM is used to modulate these coded symbols onto the frequency dimension.

A comparative study on the BER performance of STBC-OFDM and SFBC-OFDM was reported by Lin et al. [17]. They showed that STBC-OFDM performs better in a low time-selective, high frequency-selective environment than SFBC-OFDM, whereas STBC-OFDM performs worse in a high time-selective, low frequency-selective environment than SFBC-OFDM. This behavior is expected since STBC-OFDM requires adjacent OFDM

symbols to experience similar fading and therefore STBC-OFDM suffers as the channel varies quickly with time. On the other hand, SFBC-OFDM is more sensitive to frequency variations as it requires neighboring subcarriers to encounter approximately the same frequency response. For channels with a balanced degree of time/frequency variability, the BER performance of STBC-OFDM and SFBC-OFDM closely resembles each other [17].

The above-mentioned systems [3, 4, 15–17] are categorized as *coherent* communication schemes, as the detection process requires knowledge on the propagation medium. Results in [15–17] assumes *perfect* channel knowledge is available for detection. However, in reality, the channel response must first be estimated and the channel estimates are used in place of the ideal channel coefficients for decision-making. Thus, channel estimation is compulsory when coherent detection schemes are employed.

One way of estimating the channel response is to multiplex known (pilot) symbols into the transmitted signal, and the channel response is estimated by processing the received training symbols. The estimated channel is then used in the coherent detection process. This channel estimation scheme is known as pilot-symbol assisted channel estimation (PACE), and some authors refers to the combined channel estimator and signal detection scheme as pilot symbol assisted modulation (PSAM).

Cavers was first to analyze PSAM in Rayleigh flat-fading environment [18]. A Wiener, or linear minimum mean square-error (LMMSE), filter was used to interpolate the channel response between pilot symbols. He showed that an optimized PSAM can accomplish within 3-dB of the known-channel coherent detection BER performance, even in a high (5%) normalized Doppler spread environment.

PACE was extended for frequency-selective fading channels by van de Beek et al. in 1995 [19]. In their paper, the pilot symbols are inserted as a preamble of the transmission

frame, so that the first OFDM symbol contains pilot symbols on *all* subcarriers. Their channel estimator was a frequency-domain Wiener (or LMMSE) filter, and no time-domain filtering was performed as they assumed the channel response is constant throughout one transmission frame. A few years later, Edfors et al. proposed a low-rank approximation of van de Beek's channel estimator to reduce its complexity [20].

Realizing that wireless channel can be both time- and frequency-selective, Li et al. incorporated both time and frequency channel correlations in their LMMSE channel estimator [21]. Li's channel estimator assumed a reference is available at *all times and frequencies*. Such a reference can be obtained with pilot symbols during the preamble period, and with various decision-feedback techniques during the data transmission period. Furthermore, they considered the effects of correlation mismatch in their channel estimator, and proposed a worst-case design method to obtain a robust channel estimator.

Briefly before the proposal of Li's time-frequency channel estimator [21], Hoeher proposed another 2-dimensional (time-frequency) LMMSE channel estimation that involves periodic insertion of pilot symbols in the time-frequency domain satisfying the 2-D sampling theorem [22].

When multiple antennas are employed, the major difficulty for channel estimation is that the pilot symbols from different transmit antennas interfere with each other at the receiver. In order to perform channel estimation for MIMO channels, the training symbols across all transmit antennas has to be designed properly such that the spatial interference can be suppressed at the receiver.

Li et al. was first to tackle the problem of channel estimation for MIMO-OFDM systems. In [23], they assumed the channel is constant throughout a transmission frame which includes several OFDM symbols. Therefore, they use the first OFDM symbol as a train-

ing symbol, and derived a least-squares (LS) channel estimator on the channel impulse response. A special training sequence that would simplify their channel estimator is proposed subsequently in [24]. Li's channel estimator is based on one OFDM symbol only, therefore it lacks the ability to use the time-correlation of a practical wireless channels. Jeon et al. extends the time-frequency (2-D) LMMSE channel estimation concepts from [21,22] to the MIMO-OFDM case. They provided an orthogonal pilot sequence design which allows the channel estimator to cancel the spatial interference on the pilot symbols. At the same time, Auer also proposed a time-frequency channel estimation for MIMO-OFDM systems, but his approach is based on two concatenated one-dimensional filter, first in frequency, then in time, to obtained the final channel estimates [25]. Other realizations of two-dimensional channel estimation methods are reported by Miao and Juntti [26], and Zhang et al. [27]. In [26], they made use of the space-time channel correlation to obtain a channel estimation, whereas in [27], space-frequency correlation is exploited.

1.1 Thesis Objectives

To our knowledge, we have not found a channel estimator that exploits all space, time, and frequency channel correlations. This motivates us to generalize the above-mentioned lower-dimensional channel estimators into the complete space-time-frequency framework.

Computer simulations are setup to support the effectiveness of our space-time-frequency (STF) channel estimator for SFBC-OFDM systems.

1.2 Thesis Organization and Contributions

This thesis is organized as follows. We review the operations of multiple-input multiple-output (MIMO)-OFDM systems in chapter 2. In particular, we revisit how to incorporate transmitter diversity in an OFDM system by combining OFDM with STBC. Once the system model is described, we discuss the mathematical model of wireless channels. We focus on the discussion of a triply-selective (time, frequency, and space) channels, and explore various properties of such channels. At the end, a realistic ray-based channel model as specified by the Third Generation Partnership Project (3GPP) is reviewed [28].

In chapter 3, we present our linear STF channel estimator that achieves minimum mean square-error (MMSE). We then showed that our STF channel estimator simplifies to the time-frequency (TF) channel estimator when no spatial correlations exist. Pilot design issues will be discussed also. Finally, due to the high complexity of the proposed STF channel estimation method, complexity reduction by STF partitioning is introduced.

After the channel estimation methods are presented, we include simulation results to confirm the operation of our estimation methods in chapter 4. We look into cases where STF estimation provides better overall BER performance than TF methods.

Conclusions and recommendations for future work are given in chapter 5. The simulation software used to generate the results is included in the attached CD-ROM along with the corresponding user guide in the appendix.

This thesis contains the following contributions:

1. A channel estimation scheme that exploits the *full* space-time-frequency correlation of MIMO time-varying frequency-selective channels. Furthermore, our STF channel

estimator can be applied to MIMO-OFDM systems with any arbitrary STF coding scheme. When different transmit/receive antenna pairs are spatially uncorrelated, our STF channel estimator simplifies to the same TF channel estimator in [29]. (Chapter 3)

2. A strategy to reduce complexity of our STF channel estimator by STF partitioning. (Chapter 3)

3. Computer simulation results showing that SFBC-OFDM with our proposed STF channel estimator yields better BER performance than with the TF channel estimator. (Chapter 4)

Chapter 2

System and Channel Models

This chapter provides the background for MIMO-OFDM channel estimation. Two important aspects are covered here: 1) description of SFBC-OFDM systems, and 2) description of an STF-selective wireless channel model. We will first start with the overall system model.

2.1 System Model

An SFBC-OFDM system is composed of two parts. The inner part is the OFDM system, whose purpose is to convert frequency selective wideband channels into a number of frequency flat subchannels. The outer part is an STBC system which provides spatial diversity. For an OFDM system with N_f subcarriers, a block of N_f consecutive symbols are transmitted in parallel on N_f subcarriers at the same time. When the STBC code length, L_B, is less than or equal to N_f, the overall system is an SFBC-OFDM because coding is only performed across the subcarriers of each OFDM symbol. However, when $L_B > N_f$, then the STBC codeword spans over several OFDM symbols, thus the resulting overall system

is called space-time-frequency block coded OFDM (STFBC-OFDM).

In our work, we use an SFBC-OFDM system similar to [16] for evaluating the effectiveness of our proposed channel estimator. However, the channel estimator presented in next section is independent of the code structure and can be applied to any general STF coded OFDM system.

We start with a description of STBC systems, then OFDM, and finally we will demonstrate how to combine STBC with OFDM.

2.1.1 Space-Time Block Coded System

Space-Time Block Coding (STBC) is an open-loop multiple-antenna transmission technique for wireless communications. The benefit of an open-loop transmission system is that the transmitter does not require any prior channel knowledge. On the other hand, a closed-loop multiple-antenna scheme, such as beam-forming [30], requires the transmitter to exploit certain channel knowledge in providing reliable data transmission. Although it has been shown that closed-loop systems are capable of offering higher capacity than open-loop systems [1, 2], the simplicity and robustness of open-loop systems (i.e. STBC) made them an attractive choice for current multiple-antenna transmission schemes.

There are two classes of STBC codes: 1) linear STBC, and 2) non-linear STBC. The simple structure of linear STBC, also known as linear dispersion (LD) space-time codes, is developed by Hassibi and Hochwald [31]. This structure allows efficient V-BLAST-like decoding schemes [31], which made them a preferred choice over non-linear STBC. We will, therefore, focus our presentation on linear STBC systems.

STBC System Model

The first linear STBC was introduced by Alamouti in 1998 [3], and further analyzed by Tarokh in 1999 [32]. While explicit linear STBC designs were shown in the mentioned papers, we adopt a more general framework in presenting linear STBC codes as seen in [31, 33].

Figure 2.1 shows the overall block diagram of a STBC system.

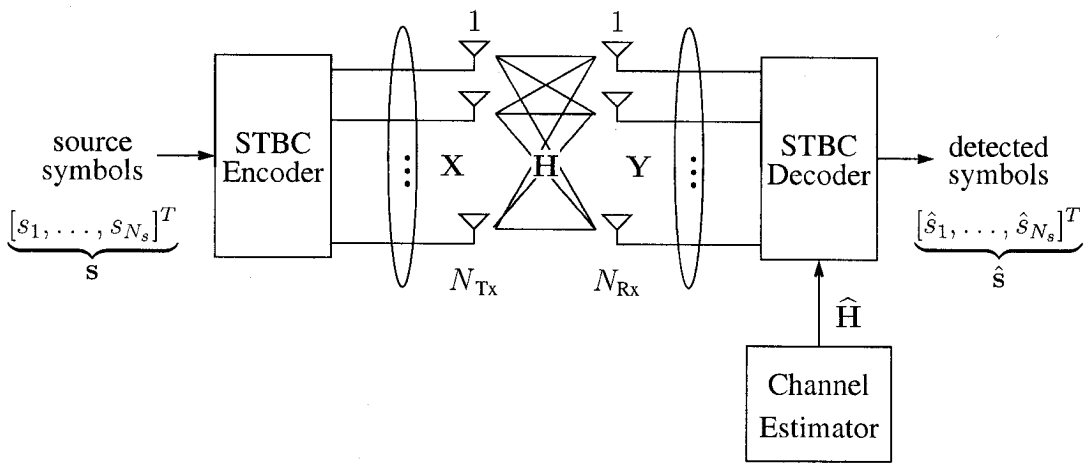


Figure 2.1 Block diagram of a STBC system.

In an STBC system with N_{Tx} transmit antennas, blocks of N_s source symbols, $\mathbf{s} = [s_1, \ldots, s_{N_s}]^T$, are encoded by the STBC encoder to produce STBC code matrices, $\mathbf{X} \in \mathbb{C}^{N_{\text{Tx}} \times L_B}$.[1] The i^{th} row of the STBC matrix constitutes the coded *data-stream* to be transmitted on the i^{th} transmit antenna, whereas the j^{th} column of the STBC matrix constitutes the coded *spatial vector* to be transmitted simultaneously on all N_{Tx} antennas at the j^{th}

[1] $\{\cdot\}^T$ denotes the vector transpose operation and $\mathbb{C}^{N \times M}$ denotes an N-by-M matrix with complex elements.

time slot.

$$\mathbf{X} = \overset{\text{Time}}{\longrightarrow} \begin{bmatrix} x_{1,1} & \cdots & x_{1,L_B} \\ \vdots & \vdots & \vdots \\ x_{N_{\text{Tx}},1} & \cdots & x_{N_{\text{Tx}},L_B} \end{bmatrix} \downarrow \text{Tx antenna}. \tag{2.1}$$

Since transmission of a STBC matrix carrying N_s source symbols requires L_B channel uses (time slots), we define the STBC encoding rate as

$$r_c = \frac{N_s}{L_B} \text{ symbols/channel use}. \tag{2.2}$$

Definition 2.1.1 (Linear STBC Encoder) *All linear STBC codes can be represented as* [2]

$$\mathbf{X} = \sum_{i=1}^{N_s} \left(s_i \, \widetilde{\mathbf{A}}_i + s_i^* \, \widetilde{\mathbf{B}}_i \right), \tag{2.3}$$

where $\mathbf{X} \in \mathbb{C}^{N_{Tx} \times L_B}$ *denotes the STBC code matrix, and* $\left\{ \widetilde{\mathbf{A}}_i, \widetilde{\mathbf{B}}_i \right\}$ *is a set of complex matrices,* $\widetilde{\mathbf{A}}_i, \widetilde{\mathbf{B}}_i \in \mathbb{C}^{N_{Tx} \times L_B}$*, that completely describes a particular STBC design [33].* ■

By representing the complex source symbols, s, in terms of its real, $s^{\Re}$, and imaginary parts, $s^{\Im}$,[3]

$$s = s^{\Re} + \jmath \, s^{\Im}, \tag{2.4}$$

[2] $\{\cdot\}^*$ denotes the complex conjugate operation.
[3] $\jmath = \sqrt{-1}$ is the complex number.

it would be useful to rewrite (2.3) as

$$\mathbf{X} = \sum_{i=1}^{N_s} \left[\left(s_i^{\Re} + \jmath\, s_i^{\Im} \right) \widetilde{\mathbf{A}}_i + \left(s_i^{\Re} - \jmath\, s_i^{\Im} \right) \widetilde{\mathbf{B}}_i \right], \tag{2.5a}$$

$$= \sum_{i=1}^{N_s} \left[s_i^{\Re} \left(\widetilde{\mathbf{A}}_i + \widetilde{\mathbf{B}}_i \right) + \jmath\, s_i^{\Im} \left(\widetilde{\mathbf{A}}_i - \widetilde{\mathbf{B}}_i \right) \right]. \tag{2.5b}$$

Definition 2.1.2 (Linear STBC Encoder - Alternate Form) *The encoding process can be re-written in terms of the real and imaginary parts of the source symbols,*

$$\mathbf{X} = \sum_{i=1}^{N_s} \left(s_i^{\Re} \mathbf{A}_i + \jmath\, s_i^{\Im} \mathbf{B}_i \right), \tag{2.6}$$

where the relationships between the matrices $\{\widetilde{\mathbf{A}}_i, \widetilde{\mathbf{B}}_i\}$ *and the matrices* $\{\mathbf{A}_i, \mathbf{B}_i\}$ *are given by*

$$\mathbf{A}_i = \widetilde{\mathbf{A}}_i + \widetilde{\mathbf{B}}_i\,, \quad \text{and} \quad \mathbf{B}_i = \widetilde{\mathbf{A}}_i - \widetilde{\mathbf{B}}_i\,. \tag{2.7}$$

■

Once the set of matrices, $\{\widetilde{\mathbf{A}}_i, \widetilde{\mathbf{B}}_i\}$ or equivalently $\{\mathbf{A}_i, \mathbf{B}_i\}$, are specified, the encoding process is executed according to (2.3) or (2.6).

Assuming that the flat-fading channel response matrix, $\mathbf{H} \in \mathbb{C}^{N_{\text{Rx}} \times N_{\text{Tx}}}$, is constant throughout the duration of the STBC code (L_B time slots), the following equation describes the input-output relationship of an STBC system.

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{V}\,, \tag{2.8}$$

where $\mathbf{Y} \in \mathbb{C}^{N_{\text{Rx}} \times L_B}$ is the received data matrix over L_B time slots, $\mathbf{H} \in \mathbb{C}^{N_{\text{Rx}} \times N_{\text{Tx}}}$ is the channel matrix, $\mathbf{X} \in \mathbb{C}^{N_{\text{Tx}} \times L_B}$ is the transmitted STBC matrix, and $\mathbf{V} \in \mathbb{C}^{N_{\text{Rx}} \times L_B}$ is the additive noise matrix.

Components of the additive noise matrix are samples from a two-dimensional complex Gaussian random process with zero-mean, and variance σ_v^2. Furthermore, the noise is assumed to be spatially and temporally white. In other words, if we stack the columns of $\mathbf{V}$, denoted as $\mathbf{v}_i$'s, by the *vec operator* [34], we get

$$\mathbf{v} = \mathrm{vec}(\mathbf{V}) = \mathrm{vec}\left([\mathbf{v}_1,\ \mathbf{v}_2,\ \cdots,\ \mathbf{v}_{L_B}]\right) = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_{L_B} \end{bmatrix}, \tag{2.9}$$

the resulting vector, $\mathbf{v} \in \mathbb{C}^{N_{\mathrm{Rx}} L_B}$, has the following distribution:

$$\mathbf{v} \sim \mathcal{CN}(\mathbf{0}_{N_{\mathrm{Rx}} L_B},\ \sigma_v^2 \mathbf{I}_{N_{\mathrm{Rx}} L_B}), \tag{2.10}$$

where $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the N-dimensional complex Gaussian distribution with mean vector $\boldsymbol{\mu} \in \mathbb{C}^N$ and covariance matrix $\boldsymbol{\Sigma} \in \mathbb{C}^{N \times N}$, $\mathbf{0}$ is the all-zero vector, and $\mathbf{I}$ is the identity matrix.

To ease our discussion of the STBC maximum likelihood (ML) decoder, it is convenient to reorder the matrix input-output relationship (2.8) into the following vector form:[4]

$$\begin{aligned} \mathbf{y} &= \mathrm{vec}(\mathbf{Y}) && \text{(2.11a)} \\ &= \mathrm{vec}(\mathbf{H}\,\mathbf{X} + \mathbf{V}) && \text{(2.11b)} \\ &= \mathrm{vec}(\mathbf{H}\,\mathbf{X}) + \mathrm{vec}(\mathbf{V}) && \text{(2.11c)} \\ &= (\mathbf{I}_{L_B} \otimes \mathbf{H})\ \mathrm{vec}(\mathbf{X}) + \mathrm{vec}(\mathbf{V}) && \text{(2.11d)} \\ &= (\mathbf{I}_{L_B} \otimes \mathbf{H})\,\mathbf{x} + \mathbf{v}, && \text{(2.11e)} \end{aligned}$$

[4] The properties of $\mathrm{vec}(\cdot)$ and $\otimes$ can be found in [34].

where $\mathrm{vec}(\cdot)$ is the vec operator, $\otimes$ is the Kronecker product [34], $\mathbf{y} \in \mathbb{C}^{N_{\mathrm{Rx}}L_B}$ is the vectorized received data, $\mathbf{x} \in \mathbb{C}^{N_{\mathrm{Tx}}L_B}$ is the vectorized transmitted data, and $\mathbf{v} \in \mathbb{C}^{N_{\mathrm{Rx}}L_B}$ is the vectorized additive noise.

When the receiver has *perfect* knowledge of $\mathbf{H}$, the conditional PDF of $\mathbf{y}$ is complex Gaussian distributed. With the noise distribution specified in (2.10), the conditional probability density function (PDF) of $\mathbf{y}$ is

$$\mathrm{p}(\mathbf{y}|\,\mathbf{x},\mathbf{H}) = \frac{1}{\pi^{N_{\mathrm{Rx}}L_B}\cdot N_{\mathrm{Rx}}L_B\,\sigma_v^2}\exp\left[\frac{-\|\mathbf{y}-(\mathbf{I}_{L_B}\otimes\mathbf{H})\,\mathbf{x}\|^2}{\sigma_v^2}\right], \tag{2.12}$$

where $\|\cdot\|$ denotes the 2-norm of a vector [34].

Definition 2.1.3 (Maximum Likelihood Decoder) *The maximum likelihood (ML) decoder decides* $\hat{\mathbf{s}}$ *was transmitted when* $\hat{\mathbf{s}}$ *maximizes the conditional PDF,* $\mathrm{p}(\mathbf{y}|\,\mathbf{x}(\mathbf{s}),\mathbf{H})$,[5]

$$\hat{\mathbf{s}} = \arg\max_{\mathbf{s}}\ \mathrm{p}(\mathbf{y}|\,\mathbf{x}(\mathbf{s}),\mathbf{H})\,. \tag{2.13}$$

■

In (2.13), we explicitly write the vectorized STBC matrix, $\mathbf{x}(\mathbf{s})$, as a function of the source symbols, $\mathbf{s}$, to illustrate the dependency. This functional notation will be dropped subsequently for simplicity.

Since the exponential function in (2.13) is a monotonically increasing function, maximizing the conditional PDF is equivalent to maximizing the exponential function argument scaled by any positive constant. Furthermore, maximizing the exponential function argument is equivalent to minimizing the *negative* of the argument. Therefore, we can convert

the maximization problem in (2.13) as

$$\begin{aligned}
\hat{\mathbf{s}} &= \arg\max_{\mathbf{s}} \, \mathrm{p}(\mathbf{y}|\,\mathbf{x}, \mathbf{H}) \\
&= \arg\min_{\mathbf{s}} \, \|\mathbf{y} - (\mathbf{I}_{L_B} \otimes \mathbf{H})\,\mathbf{x}\|^2 && \text{(2.14a)} \\
&= \arg\min_{\mathbf{s}} \, \|\operatorname{vec}(\mathbf{Y} - \mathbf{H}\,\mathbf{X})\|^2 && \text{(2.14b)} \\
&= \arg\min_{\mathbf{s}} \, \|\mathbf{Y} - \mathbf{H}\,\mathbf{X}\|^2 . && \text{(2.14c)}
\end{aligned}$$

We use (2.11) in (2.14a) to get (2.14b), and we use the property of $\|\operatorname{vec}(\mathbf{A})\| = \|\mathbf{A}\|$ to reach the final result (2.14c).

According to (2.14), the ML decoder for a general linear STBC involves an exhaustive search in the N_s-dimensional space containing all possible source vectors. Therefore, the complexity of the ML detector grows exponentially in the source vector's size, N_s, i.e. complexity grows at $\mathcal{O}(n^{N_s})$ with n being the constellation size. Fortunately, there exists a special class of STBC, namely the orthogonal STBC, which allows further simplification of (2.14) into a symbol-wise ML decoder. We now address this type of STBC in the following part.

Orthogonal STBC

An interesting class of all STBC is the orthogonal space-time block code (OSTBC) [35,36]. The orthogonality allows ML decoding in linear complexity, given that the channel is slowly varying. Because of the linear decoding complexity feature, we choose to use OSTBC in this thesis.

Definition 2.1.4 (Orthogonality of OSTBC) *Orthogonal STBC has the following properties:*

$$\mathbf{X}\,\mathbf{X}^H = \Big(\sum_{i=1}^{N_s} |s_i|^2\Big)\,\mathbf{I}_{N_{Tx}}\,, \quad \text{and} \quad \mathbf{X}^H\mathbf{X} = \Big(\sum_{i=1}^{N_s} |s_i|^2\Big)\,\mathbf{I}_{L_B}\,, \tag{2.15}$$

where $\{\cdot\}^H$ *denotes the complex conjugate transpose operation.*

■

With this orthogonal property, we now revisit on how to decouple the optimal ML decoder into a symbol-wise decoder.

Simplifying the ML metric, $\|\mathbf{Y}-\mathbf{H}\,\mathbf{X}\|^2$, with the orthogonality property of OSTBC (2.15) gives

$$\begin{aligned}
&\|\mathbf{Y}-\mathbf{H}\,\mathbf{X}\|^2 \\
&= \mathrm{tr}\left[(\mathbf{Y}-\mathbf{H}\,\mathbf{X})(\mathbf{Y}-\mathbf{H}\,\mathbf{X})^H\right] && \text{(2.16a)}\\
&= \mathrm{tr}\left(\mathbf{Y}\,\mathbf{Y}^H\right) - \mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{Y}^H\right) - \mathrm{tr}\left(\mathbf{Y}\,\mathbf{X}^H\mathbf{H}^H\right) + \mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{X}^H\mathbf{H}^H\right) && \text{(2.16b)}\\
&= \mathrm{tr}\left(\mathbf{Y}\,\mathbf{Y}^H\right) - \mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{Y}^H\right) - \mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{Y}^H\right)^H + \mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{X}^H\mathbf{H}^H\right) && \text{(2.16c)}\\
&= \mathrm{tr}\left(\mathbf{Y}\,\mathbf{Y}^H\right) - \mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{Y}^H\right) - \mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{Y}^H\right)^* + \mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{X}^H\mathbf{H}^H\right) && \text{(2.16d)}\\
&= \mathrm{tr}\left(\mathbf{Y}\,\mathbf{Y}^H\right) - 2\,\Re\left\{\mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{Y}^H\right)\right\} + \mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{X}^H\mathbf{H}^H\right) && \text{(2.16e)}\\
&= \mathrm{tr}\left(\mathbf{Y}\,\mathbf{Y}^H\right) - 2\,\Re\left\{\mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{Y}^H\right)\right\} + \Big(\sum_{i=1}^{N_s} |s_i|^2\Big)\,\mathrm{tr}\left(\mathbf{H}\,\mathbf{H}^H\right)\,, && \text{(2.16f)}
\end{aligned}$$

where in (2.16e), we use the fact that for any complex number, z, we have

$$z + z^* = 2\,\Re\{z\}\,. \tag{2.17}$$

With (2.16), we can rewrite the ML optimization problem (2.14), as

$$\begin{aligned}
\hat{\mathbf{s}} &= \arg\min_{\mathbf{s}} \|\mathbf{Y} - \mathbf{H}\,\mathbf{X}\|^2 \\
&= \arg\min_{\mathbf{s}} \left[\mathrm{tr}\left(\mathbf{Y}\,\mathbf{Y}^H\right) - 2\,\Re\left\{\mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{Y}^H\right)\right\} + \left(\sum_{i=1}^{N_s} |s_i|^2\right) \mathrm{tr}\left(\mathbf{H}\,\mathbf{H}^H\right) \right] && \text{(2.18a)} \\
&= \arg\min_{\mathbf{s}} \left[-2\,\Re\left\{\mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{Y}^H\right)\right\} + \left(\sum_{i=1}^{N_s} |s_i|^2\right) \|\mathbf{H}\|^2 \right] && \text{(2.18b)} \\
&= \arg\min_{\mathbf{s}} \left[\sum_{i=1}^{N_s} |s_i|^2 - \frac{2\,\Re\left\{\mathrm{tr}\left(\mathbf{H}\,\mathbf{X}\,\mathbf{Y}^H\right)\right\}}{\|\mathbf{H}\|^2} \right] . && \text{(2.18c)}
\end{aligned}$$

In (2.18b) and (2.18c), we use the fact that addition/subtraction of any constants and multiplication/division by any positive constants would not alter the minimization problem.

Substituting the STBC definition in the real/imaginary form (2.6) into (2.18c) yields

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \left[\sum_{i=1}^{N_s} |s_i|^2 - \frac{2\,\Re\left\{\mathrm{tr}\left(\mathbf{H}\sum_{i=1}^{N_s}\left(s_i^{\Re}\mathbf{A}_i + \jmath\, s_i^{\Im}\mathbf{B}_i\right)\mathbf{Y}^H\right)\right\}}{\|\mathbf{H}\|^2} \right] \tag{2.19a}$$

$$= \arg\min_{\mathbf{s}} \left[\sum_{i=1}^{N_s} |s_i|^2 - \frac{2\,\Re\left\{\mathrm{tr}\left(\sum_{i=1}^{N_s}\left(s_i^{\Re}\mathbf{H}\mathbf{A}_i\mathbf{Y}^H + \jmath\, s_i^{\Im}\mathbf{H}\mathbf{B}_i\mathbf{Y}^H\right)\right)\right\}}{\|\mathbf{H}\|^2} \right] \tag{2.19b}$$

$$= \arg\min_{\mathbf{s}} \sum_{i=1}^{N_s} \left[|s_i|^2 - \frac{2\,\Re\left\{\mathrm{tr}\left(s_i^{\Re}\mathbf{H}\mathbf{A}_i\mathbf{Y}^H\right) + \mathrm{tr}\left(\jmath\, s_i^{\Im}\mathbf{H}\mathbf{B}_i\mathbf{Y}^H\right)\right\}}{\|\mathbf{H}\|^2} \right] \tag{2.19c}$$

$$= \arg\min_{\mathbf{s}} \sum_{i=1}^{N_s} \left[|s_i|^2 - \frac{2\,\Re\left\{s_i^{\Re}\,\mathrm{tr}\left(\mathbf{H}\mathbf{A}_i\mathbf{Y}^H\right) + \jmath\, s_i^{\Im}\,\mathrm{tr}\left(\mathbf{H}\mathbf{B}_i\mathbf{Y}^H\right)\right\}}{\|\mathbf{H}\|^2} \right] \tag{2.19d}$$

$$= \arg\min_{\mathbf{s}} \sum_{i=1}^{N_s} \left[\left| s_i - \frac{\Re\left\{\mathrm{tr}\left(\mathbf{H}\mathbf{A}_i\mathbf{Y}^H\right)\right\} - \jmath\,\Im\left\{\mathrm{tr}\left(\mathbf{H}\mathbf{B}_i\mathbf{Y}^H\right)\right\}}{\|\mathbf{H}\|^2} \right|^2 \right.$$
$$\left. - \left| \frac{\Re\left\{\mathrm{tr}\left(\mathbf{H}\mathbf{A}_i\mathbf{Y}^H\right)\right\} - \jmath\,\Im\left\{\mathrm{tr}\left(\mathbf{H}\mathbf{B}_i\mathbf{Y}^H\right)\right\}}{\|\mathbf{H}\|^2} \right|^2 \right] \tag{2.19e}$$

$$= \arg\min_{\mathbf{s}} \sum_{i=1}^{N_s} \left[\left| s_i - \frac{\Re\left\{\mathrm{tr}\left(\mathbf{H}\mathbf{A}_i\mathbf{Y}^H\right)\right\} - \jmath\,\Im\left\{\mathrm{tr}\left(\mathbf{H}\mathbf{B}_i\mathbf{Y}^H\right)\right\}}{\|\mathbf{H}\|^2} \right|^2 \right] . \tag{2.19f}$$

The second term of (2.19e) is a constant, and can therefore be dropped from the minimization. From (2.19f), we see that the minimization involves a sum of positive terms which contains only s_i. Thus, the joint minimization over N_s symbols is equivalent to minimizing each of N_s terms independently.

Definition 2.1.5 (Decoupled, or Symbol-wise, ML Decoder) *The final decoupled ML solution becomes*

$$\hat{s}_i = \arg\min_{s_i} \left| s_i - \frac{\Re\left\{\mathrm{tr}\left(\mathbf{H}\mathbf{A}_i\mathbf{Y}^H\right)\right\} - \jmath\,\Im\left\{\mathrm{tr}\left(\mathbf{H}\mathbf{B}_i\mathbf{Y}^H\right)\right\}}{\|\mathbf{H}\|^2} \right|^2 , \text{ for } i = 1, \ldots, N_s . \tag{2.20}$$

■

OSTBC Examples

Two OSTBC's were used in this thesis – one that employs two transmit antennas, and the other employs four transmit antennas.

Two Transmit Antennas ($N_{\text{Tx}} = 2$)

We use Alamouti's scheme for our two transmit antennas system [3]. Alamouti's OSTBC code matrix is defined by

$$\mathbf{X} = \begin{bmatrix} s_1 & s_2 \\ -s_2* & s_1^* \end{bmatrix}. \tag{2.21}$$

The corresponding matrices $\widetilde{\mathbf{A}}_i$'s and $\widetilde{\mathbf{B}}_i$'s are given by

$$\widetilde{\mathbf{A}}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \widetilde{\mathbf{A}}_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \tag{2.22}$$

and

$$\widetilde{\mathbf{B}}_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \widetilde{\mathbf{B}}_2 = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}. \tag{2.23}$$

The $\mathbf{A}_i$'s and $\mathbf{B}_i$'s matrices can be found by using (2.7). Since we transmit two symbols over two time slots in Alamouti's OSTBC, the encoding rate is 1.

Four Transmit Antennas ($N_{\text{Tx}} = 4$)

We use Tirkkonen and Hottinen's scheme for our four transmit antennas system [37].

Tirkkonen and Hottinen's OSTBC code matrix is defined by

$$\mathbf{X} = \begin{bmatrix} s_1 & s_2 & s_3 & 0 \\ -s_2^* & s_1^* & 0 & -s_3 \\ -s_3^* & 0 & s_1^* & s_2 \\ 0 & s_3^* & -s_2^* & s_1 \end{bmatrix}. \tag{2.24}$$

The corresponding matrices $\widetilde{\mathbf{A}}_i$'s and $\widetilde{\mathbf{B}}_i$'s are given by

$$\widetilde{\mathbf{A}}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \widetilde{\mathbf{A}}_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \widetilde{\mathbf{A}}_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{2.25}$$

and

$$\widetilde{\mathbf{B}}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \widetilde{\mathbf{B}}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad \widetilde{\mathbf{B}}_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{2.26}$$

The $\mathbf{A}_i$'s and $\mathbf{B}_i$'s matrices can be found by using (2.7). With this OSTBC scheme, we transmit three symbols over four time slots. Therefore, the encoding rate is $3/4$.

In usual STBC operations, the allowable total transmit power is divided equally to each transmit antennas. For example, with four transmit antennas, the per-antenna transmit power should be $1/4$, assuming unit total transmit power. However, in Tirkkonen and Hottinen's OSTBC, only three of the four transmit antennas are *active* at any time instant.

Therefore, the total power is divided equally into the *three* active antennas in this case. In other words, the per-antenna transmit power is set to $1/3$. A more detailed discussion about the signal-to-noise ratio is included in Appendix A.

In this section, we reviewed the encoding and the symbol-wise ML decoding process of OSTBC. Furthermore, we revisited the code structure of two OSTBC's that are used in this thesis. Recall that STBC is only designed to work in a frequency flat fading environment. We will therefore rely on OFDM to convert frequency selective fading channels into frequency flat subchannels such that OSTBC can be applied. The operations of OFDM is the topic for the next section.

2.1.2 OFDM System

For high data-rate applications, the wireless fading channel is likely to be frequency selective. An efficient way to handle frequency selectivity is to decompose the channels into many narrowband subchannels via orthogonal frequency division multiplexing (OFDM) [7, 8], which is a bandwidth-efficient multi-carrier transmission scheme [38–40]. Figure 2.2 illustrates a block diagram of a baseband OFDM system.

At the transmitter, a block of N_f source symbols, $\{x_0^{\mathcal{F}}, \ldots, x_{N_f-1}^{\mathcal{F}}\}$, are serial-to-parallel (S/P) converted into a vector, $\mathbf{x}^{\mathcal{F}}$. The S/P converter altered the high symbol rate, $1/T_s$, at the input, to a lower OFDM symbol rate, $1/(N_f T_s)$, at the output. Each component of $\mathbf{x}^{\mathcal{F}}$ is modulated on a different subcarrier (subchannel) via the inverse discrete Fourier transform (IDFT) operation, such that

$$\mathbf{x}^{\mathcal{T}} = \mathbf{F}^H \mathbf{x}^{\mathcal{F}}, \tag{2.27}$$

where $\mathbf{F}^H$ is the unitary IDFT matrix, which is the Hermitian transpose of the unitary

discrete Fourier transform (DFT) matrix, F,

$$\mathbf{F} = \frac{1}{\sqrt{N_f}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{-j2\pi\frac{1\cdot 1}{N_f}} & \cdots & e^{-j2\pi\frac{1\cdot(N_f-1)}{N_f}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j2\pi\frac{(N_f-1)\cdot 1}{N_f}} & \cdots & e^{-j2\pi\frac{(N_f-1)\cdot(N_f-1)}{N_f}} \end{bmatrix}. \tag{2.28}$$

There are N_f subcarriers so that each symbol, $x_i^{\mathcal{F}}$ for $i = 0, \ldots, N_f$, occupies one subcarrier on its own. Because of this subcarrier modulation property, the source symbols are placed on the *frequency*-domain; thus, the $\{\cdot\}^{\mathcal{F}}$ notation is used.

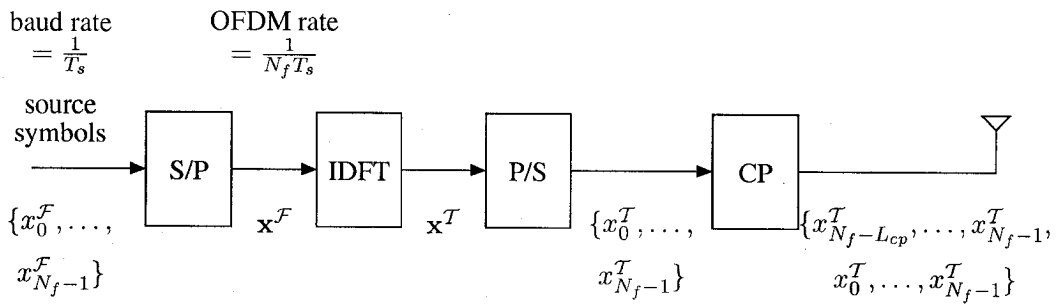
The modulated symbol vector, $\mathbf{x}^T$, is then parallel-to-serial (P/S) converted back to symbol rate of $1/T_s$. A cyclic prefix (CP) of length, L_{cp}, composed of the last L_{cp} symbols of $\mathbf{x}^T$, is appended as a prefix in the current transmission block, and later discarded at the receiver to eliminate inter-symbol-interference (ISI) from the previous transmission block. ISI is due to the different propagation delays associated with the multipath scatterers in a wireless channel [41].

The discrete-time (DT) baseband equivalent channel impulse response (CIR), g, can be obtained by sampling the continuous-time (CT) multipath CIR (including transmit and receive filters) at symbol rate $1/T_s$ [9],

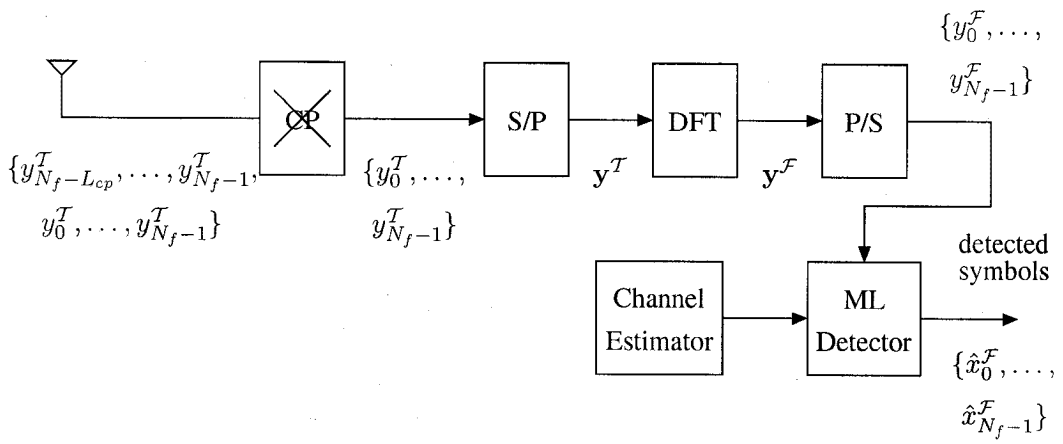
$$\mathbf{g} = [g_0, g_1, \ldots, g_{L-1}]^T, \tag{2.29}$$

where L is the length of the DT CIR.

Since the received signal is the convolution of the transmitted signal with the channel CIR and the first L_{cp} symbols of the received sequence is discarded at the receiver, ISI will



(a) Transmitter



(b) Receiver

Figure 2.2 OFDM block diagram.

be completely eliminated if $L_{cp} \geq L - 1$. This proper choice of the CP length is assumed in this work.

Assuming the channel impulse response is constant over one OFDM symbol duration and after discarding the CP, the remaining received symbols are given by

$$\begin{aligned} y_0^{\mathcal{T}} &= g_0\, x_0^{\mathcal{T}} + g_1\, x_{N_f-1}^{\mathcal{T}} + g_2\, x_{N_f-2}^{\mathcal{T}} + \ldots + g_{L-1}\, x_{N_f-L+1}^{\mathcal{T}} + v_0^{\mathcal{T}} \\ y_1^{\mathcal{T}} &= g_0\, x_1^{\mathcal{T}} + g_1\, x_0^{\mathcal{T}} + g_2\, x_{N_f-1}^{\mathcal{T}} + \ldots + g_{L-1}\, x_{N_f-L+2}^{\mathcal{T}} + v_1^{\mathcal{T}} \\ &\vdots \qquad\qquad\qquad\qquad \vdots \\ y_{N_f-1}^{\mathcal{T}} &= g_0\, x_{N_f-1}^{\mathcal{T}} + g_1\, x_{N_f-2}^{\mathcal{T}} + g_2\, x_{Nf-3}^{\mathcal{T}} + \ldots + g_{L-1}\, x_{N_f-L}^{\mathcal{T}} + v_{N_f-1}^{\mathcal{T}}\,, \end{aligned} \tag{2.30}$$

where $v^{\mathcal{T}}$ represents the time domain additive white Gaussian noise (AWGN).

From (2.30), it is clear that the use of CP converts the linear convolution into a *circular* convolution. Using matrix notation, we have

$$\begin{bmatrix} y_0^{\mathcal{T}} \\ y_1^{\mathcal{T}} \\ \vdots \\ y_{N_f-1}^{\mathcal{T}} \end{bmatrix} = \begin{bmatrix} g_0 & 0 & \cdots & \cdots & 0 & g_{L-1} & \cdots & g_1 \\ g_1 & g_0 & 0 & \cdots & \cdots & 0 & \ddots & \vdots \\ \vdots & g_1 & g_0 & 0 & \cdots & \cdots & 0 & g_{L-1} \\ g_{L-1} & \vdots & g_1 & \ddots & 0 & \cdots & \cdots & 0 \\ 0 & g_{L-1} & \vdots & \ddots & g_0 & \ddots & \cdots & \vdots \\ \vdots & 0 & g_{L-1} & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & \vdots & 0 & \ddots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & g_{L-1} & \cdots & g_1 & g_0 \end{bmatrix} \begin{bmatrix} x_0^{\mathcal{T}} \\ x_1^{\mathcal{T}} \\ \vdots \\ x_{N_f-1}^{\mathcal{T}} \end{bmatrix} + \begin{bmatrix} v_0^{\mathcal{T}} \\ v_1^{\mathcal{T}} \\ \vdots \\ v_{N_f-1}^{\mathcal{T}} \end{bmatrix} \tag{2.31a}$$

$$\mathbf{y}^{\mathcal{T}} = \mathbf{G}_{\text{circ}}\, \mathbf{x}^{\mathcal{T}} + \mathbf{v}^{\mathcal{T}}\,, \tag{2.31b}$$

where $\mathbf{G}_{\text{circ}}$ is a circulant matrix [42], and $\mathbf{v}^{\mathcal{T}}$ is the time domain AWGN with distribution

$\mathcal{CN}(\mathbf{0}, \sigma_v^2 \mathbf{I})$.

After the DFT operation at the receiver, we get

$$\begin{aligned}
\mathbf{y}^{\mathcal{F}} &= \mathbf{F}\,\mathbf{y}^{\mathcal{T}} && \text{(2.32a)}\\
&= \mathbf{F}\,\mathbf{G}_{\text{circ}}\,\mathbf{x}^{\mathcal{T}} + \mathbf{F}\mathbf{v}^{\mathcal{T}} && \text{(2.32b)}\\
&= \mathbf{F}\,\mathbf{G}_{\text{circ}}\,\mathbf{F}^H\,\mathbf{F}\,\mathbf{x}^{\mathcal{T}} + \mathbf{F}\mathbf{v}^{\mathcal{T}} && \text{(2.32c)}\\
&= \mathbf{F}\,\mathbf{G}_{\text{circ}}\,\mathbf{F}^H\,\mathbf{x}^{\mathcal{F}} + \mathbf{v}^{\mathcal{F}} && \text{(2.32d)}\\
&= \mathbf{H}_{\text{diag}}\,\mathbf{x}^{\mathcal{F}} + \mathbf{v}^{\mathcal{F}}\,, && \text{(2.32e)}
\end{aligned}$$

where $\mathbf{F}$ is the unitary DFT matrix defined by (2.28).

We use a property of unitary matrices, $\mathbf{F}^H\,\mathbf{F} = \mathbf{I}$, in reaching (2.32c). Equation (2.32e) is obtained due to that any circulant matrix can be diagonalized by the unitary DFT matrix [42]. The resulting diagonal matrix is

$$\mathbf{H}_{\text{diag}} = \text{diag}\left(h[0], h[1], \ldots, h[N_f - 1]\right)\,, \tag{2.33}$$

where $h[k]$ for $k = 0, \ldots, N_f - 1$ denotes the sampled frequency response of the fading channel given by

$$h[k] = \sum_{l=0}^{L-1} g_l\, e^{j2\pi \frac{k \cdot l}{N_f}} \quad \text{for } k = 0, \ldots, N_f - 1. \tag{2.34}$$

Note that the frequency domain additive noise, $\mathbf{v}^{\mathcal{F}}$, has the same distribution as the time domain additive noise, $\mathbf{v}^{\mathcal{T}}$, because unitary transformations (i.e. unitary DFT for our case) preserves both the Gaussian distribution as well as its variance properties. Furthermore, since $\mathbf{v}^{\mathcal{T}}$ is zero-mean, so is $\mathbf{v}^{\mathcal{F}}$. Therefore, the distribution of $\mathbf{v}^{\mathcal{F}}$ is still $\mathcal{CN}(\mathbf{0}, \sigma_v^2 \mathbf{I})$.

From (2.32) and (2.33), we see that an CP-OFDM system effectively converts a fre-

quency selective fading channel into N_f parallel frequency flat fading channels, corresponding to the N_f subcarriers. The received data, after DFT, is just a scaled version of the transmitted source data and corrupted by additive noise. In other words, we can rewrite (2.32) on a per-subcarrier basis as

$$y^{\mathcal{F}}[k] = h[k]\, x^{\mathcal{F}}[k] + v^{\mathcal{F}}[k]\,, \qquad \text{for } k = 0, \ldots, N_f - 1. \tag{2.35}$$

For the rest of this thesis, we focus on the frequency description of an OFDM system and by dropping the $\{\cdot\}^{\mathcal{F}}$ notation for simplicity. The per-subcarrier input-output relationship of an OFDM system, (2.35), is rewritten as

$$y[k] = h[k]\, x[k] + v[k]\,, \qquad \text{for } k = 0, \ldots, N_f - 1. \tag{2.36}$$

Since (2.35), or (2.36), are analogous to a flat-fading system, we may apply the STBC techniques discussed in section 2.1.1 to further enhance the traditional single-input single-output (SISO) OFDM system's performance, which is the topic of our next section.

2.1.3 SFBC-OFDM

One way to realize transmit diversity on frequency selective channels is to combine STBC with OFDM. The source symbols are first encoded with STBC in the same fashion detailed in section 2.1.1, producing N_{Tx} streams of coded symbols corresponding to the N_{Tx} transmit antennas. Each of N_{Tx} coded symbol streams are modulated via OFDM as detailed in section 2.1.2 and transmitted through the N_{Tx} transmit antennas. Because of the OFDM's IDFT operation, the coded symbol sequence is placed across the *frequency*-domain; thus, this system is an *SFBC-OFDM* system when STBC coding is applied before the serial-to-

parallel operation [15]. A STBC-OFDM system can be obtained by applying STBC coding after the serial-to-parallel converter as in [16].

Figure 2.3 shows the block diagram of the SFBC-OFDM system used to evaluate the effectiveness of our STF channel estimator.

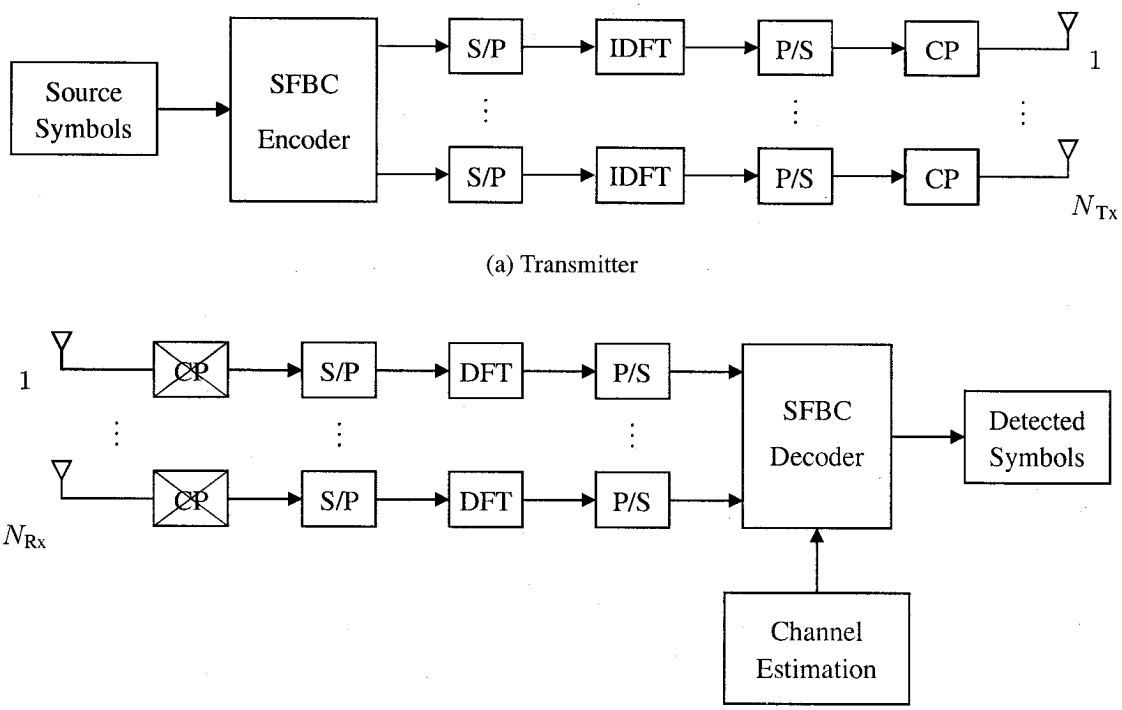


Figure 2.3 SFBC-OFDM block diagram.

An example of SFBC-OFDM with Alamouti's code is shown in Figure 2.4. The serial data symbols are first STBC coded with Alamouti's code, producing two streams of data for the two transmit antennas. Each of the data streams are then multi-carrier modulated via OFDM. From Figure 2.4, s_1 occurs on the first OFDM symbol on both antennas, but on different carriers. Redundancy (coding) is introduced on the space-frequency domain, thus an SFBC-OFDM system. Note that when the STBC code word is longer than N_f, the

STBC code word would span over several OFDM symbols, resulting in a STFBC-OFDM system.

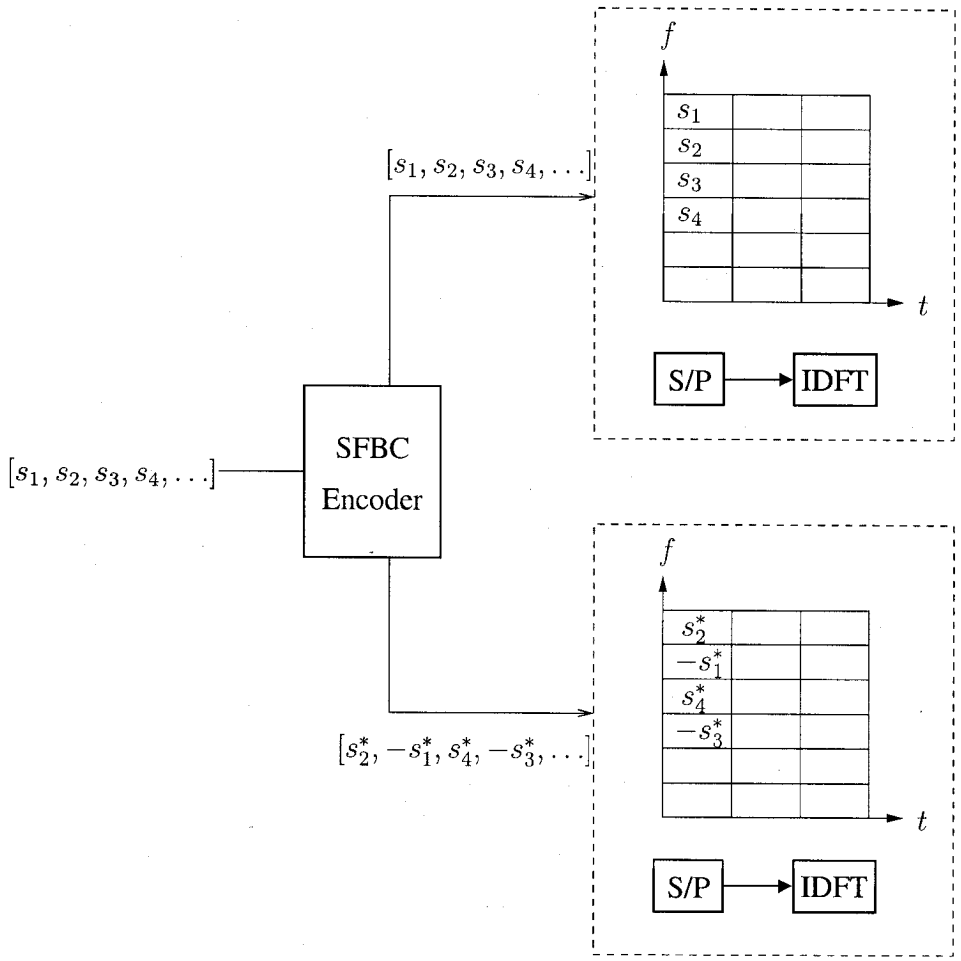


Figure 2.4 SFBC-OFDM example with Alamouti's code.

Since OFDM decomposes a frequency selective channel into N_f parallel frequency flat channels, we can extend the input-output relationship of an SISO-OFDM system, given by (2.36), to the MIMO-OFDM system on a per carrier basis:

$$\mathbf{y}[k] = \mathbf{H}[k]\,\mathbf{x}[k] + \mathbf{v}[k]\,, \qquad \text{for } k = 0, \ldots, N_f - 1. \tag{2.37}$$

In (2.37), $\mathbf{y} = [y_0, \ldots, y_{N_{\text{Rx}}}]^T$ is the received data vector over N_{Rx} receive antennas, $\mathbf{H}$ is the $N_{\text{Rx}} \times N_{\text{Tx}}$ MIMO channel matrix, $\mathbf{x} = [x_0, \ldots, x_{N_{\text{Tx}}}]^T$ is the transmitted data vector over N_{Tx} transmit antennas, $\mathbf{v} = [v_0, \ldots, v_{N_{\text{Rx}}}]^T$ is the additive noise vector at the N_{Rx} receive antennas, and finally this equation holds for each of k^{th} subcarriers in the OFDM system.

Assuming the L_B adjacent subcarriers, for each transmit antenna, has the same channel response, we can formulate the following:

$$\mathbf{Y}[m] = \Big[\mathbf{y}[m],\, \mathbf{y}[m+1],\, \cdots,\, \mathbf{y}[m+L_B-1]\Big] \tag{2.38a}$$

$$= \mathbf{H}[m] \cdot \Big[\mathbf{x}[m],\, \mathbf{x}[m+1],\, \cdots,\, \mathbf{x}[m+L_B-1]\Big] + \Big[\mathbf{v}[m],\, \mathbf{v}[m+1],\, \cdots,\, \mathbf{v}[m+L_B-1]\Big] \tag{2.38b}$$

$$= \mathbf{H}[m]\,\mathbf{X}[m] + \mathbf{V}[m]\,, \tag{2.38c}$$

where $m = \{0, L_B, \ldots, (M-1)\cdot L_B\}$ and $M = N_f/L_B \in \mathbb{I}$. It is clear that (2.38) is in the same form of (2.8), and therefore the encoding and decoding procedure is performed as discussed in section (2.1.1). The only difference is that the second dimension of $\mathbf{Y}$, $\mathbf{X}$, and $\mathbf{V}$ spans the frequency domain instead of the time domain, and we require the L_B adjacent subcarrier frequency response to be constant.

Next, we consider a space-time-frequency selective channel that we aim to estimate in this thesis.

2.2 Channel Model

This section considers the mobile wireless channels modeled as linear time-variant (LTV) systems. Furthermore, the mobile wireless channels consist of many randomly moving objects that scatter the information-bearing electromagnetic waves, causing distortions on the transmitted signal. These random distortions can be captured by modeling the mobile wireless channels as stochastic processes.

2.2.1 SISO Multipath Wireless Channel

Definition 2.2.1 (Linear Time-Varying Multipath Wireless Channel) *An LTV wireless channel composed of discrete multipaths [41] has the following form,*

$$\tilde{g}_c(t,\tau) = \sum_{n=1}^{N} \alpha_n(t)\,\delta(\tau - \tau_n)\,, \tag{2.39}$$

where $\tilde{g}_c(t,\tau)$ is the continuous-time time-varying multipath CIR observed at time t due to an impulse excitation at time $t-\tau$ [43], N is the number of discrete multipaths, $\alpha_n(t)$ is the complex-valued attenuation function of the n^{th} path, τ_n is the propagation delay of the n^{th} path, and $\delta(\cdot)$ is the unit impulse function [44].

Typically, $\tau_0(t)$ is set to 0, implying that the system is synchronized to the first component arriving at the receiver.

■

Each path might consists of many subpaths having approximately the same propagation delay, as in the case of the 3GPP channel [28]. However, each of the subpaths introduce a different phase shift, and therefore the subpaths might cause both constructive and destructive interferences. As the number of subpaths become large, then Central Limit Theorem (CLT) applies [45], and $\alpha_n(t)$ can be modeled as Gaussian random processes.

The overall channel impulse response includes the transmit and receive pulse shaping filters, $c_{\text{Tx}}(\tau)$ and $c_{\text{Rx}}(\tau)$ respectively (Figure 2.5),

$$\begin{aligned} g_c(t,\tau) &= c_{\text{Tx}}(\tau) * \tilde{g}_c(t,\tau) * c_{\text{Rx}}(\tau) \\ &= \sum_{n=1}^{N} \alpha_n(t)\, c(\tau - \tau_n)\,, \end{aligned} \tag{2.40}$$

where $c(\tau)$ is the combined Tx/Rx filter response, $*$ is the convolution operator and is defined as the following for LTV systems [46],

$$a(\tau) * b(t,\tau) = \int_{-\infty}^{\infty} a(\tau - \gamma)\, b(t,\gamma)\, d\gamma\,. \tag{2.41}$$

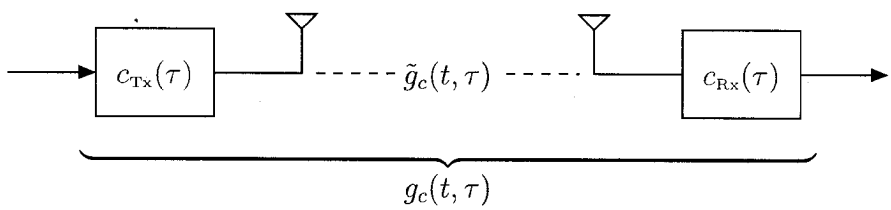


Figure 2.5 Composition of the overall channel impulse response.

Typically, the transmit and received pulse shaping filters are chosen to be a square-root raised cosine filter, which results in preserving the whiteness of the additive noise at the received filter output, when sampled at baud rate [9].

2.2.2 MIMO Multipath Wireless Channel

When multiple transmit/received antennas are employed, each pair of transmit/receive antenna would have an impulse response corresponding to (2.40). We can represent the transmit/receive antennas by their respective displacement vectors from an arbitrary spatial origin.

The LTV CIR, between the receive antenna located at $\mathbf{d}_{\text{Rx}}$ and the transmit antenna located at $\mathbf{d}_{\text{Tx}}$, is denoted as $g_c(\mathbf{d}_{\text{Rx}}, \mathbf{d}_{\text{Tx}}, t, \tau)$, where $\mathbf{d}$ is a 3-dimensional spatial vector.

It is conventional to describe OFDM systems in the frequency (carrier) domain. Therefore, it would be convenient to obtain a space-time-frequency transfer function by performing a Fourier Transform on the τ variable of $g_c(\mathbf{d}_{\text{Rx}}, \mathbf{d}_{\text{Tx}}, t, \tau)$ [47],

$$h_c(\mathbf{d}_{\text{Rx}}, \mathbf{d}_{\text{Tx}}, t, f) = \int_{-\infty}^{\infty} g_c(\mathbf{d}_{\text{Rx}}, \mathbf{d}_{\text{Tx}}, t, \tau)\, \exp(-\jmath 2\pi\tau f)\, d\tau\,. \tag{2.42}$$

Equation (2.42) represents the MIMO time-varying channel frequency response.

2.2.3 Space-Time-Frequency Selectivity

In general, a MIMO wireless channel exhibits space, time, and frequency selectivity. Durgin provided a set of transfer functions that equivalently describes the underlying MIMO time-varying channel response, $g_c(\mathbf{d}_{\text{Rx}}, \mathbf{d}_{\text{Tx}}, t, \tau)$ [47]. These transfer functions, summarized in Table 2.2.3, are helpful in understanding how each kind of selectivity relates to the physical parameters of a wireless channel.

Space Selectivity

Because of the spatial separation between multiple antennas, the channel responses of different pairs of transmit/receive antennas could potentially be different. A rich scattering environment would introduce high spatial variations due to the constructive and destructive interferences caused by a large number of scatterers. When the antenna separation is large, or when the scattering radius is large, high space-selectivity is experienced. This means that the STF transfer functions (2.42) are quite different for different pairs of $(\mathbf{d}_{\text{Rx}}, \mathbf{d}_{\text{Tx}})$.

Since the displacement (space) variable is a Fourier transform dual of the wavenum-

Table 2.1 Space-Time-Frequency Fourier Transform Relationships.

Transform Pairs	Forward Transform	Backward Transform
Displacement $\mathbf{d}$ $\longleftrightarrow$ Wavenumber $\mathbf{k}$	$\int_{-\infty}^{\infty}\{\cdot\}\exp(-\jmath\mathbf{k}^T\mathbf{d})\,d\mathbf{d}$	$\frac{1}{2\pi}\int_{-\infty}^{\infty}\{\cdot\}\exp(\jmath\mathbf{k}^T\mathbf{d})\,d\mathbf{k}$
Time t $\longleftrightarrow$ Doppler ν	$\int_{-\infty}^{\infty}\{\cdot\}\exp(-\jmath 2\pi\nu t)\,dt$	$\int_{-\infty}^{\infty}\{\cdot\}\exp(\jmath 2\pi\nu t)\,d\nu$
Frequency f $\longleftrightarrow$ Delay τ	$\int_{-\infty}^{\infty}\{\cdot\}\exp(\jmath 2\pi\tau f)\,df$	$\int_{-\infty}^{\infty}\{\cdot\}\exp(-\jmath 2\pi\tau f)\,d\tau$

ber variable, see Table (2.2.3) or [47], space selectivity exists whenever the corresponding transfer function in the wavenumber domain is not a single impulse function. The wavenumber variable is generally parameterized by the angles of propagation paths with respect to a reference direction [30]. This means that when the propagation angles are dispersed, space selectivity is experienced. We therefore refer to space-selective channels as angle-spread channels.

In a typical suburban macrocell environment, for the downlink case, most of the scatters are concentrated at the mobile station (MS) and therefore *high receive* space-selectivity is experienced. However, when the base station (BS) is located at a tall tower and the scatterers are close to the MS, the electro-magnetic waves travel through a concentrated path to the MS, leading to *low transmit* space selectivity.

Frequency Selectivity

When $N = 1$ in (2.39), the channel is a frequency flat channel because it consists of a single unit impulse function. In contrast, when $N > 1$, we have a frequency selective channel. Frequency selective channels are also known as delay-spread channels, because

of the Fourier transform relationship between frequency and delay.

Time Selectivity

We would have a time selective channel when $\alpha_n(t)$ varies with time. On the other hand, if $\alpha_n(t) = \alpha_n$ is constant over the observation period, we say the channel is quasi-static. For a cellular system, time selectivity results from the mobile unit movements. When the mobile is moving, it introduces Doppler shifts to the received signal [41]. Thus, time selective channels are also known as Doppler-spread channels.

Quantitative measures of space , time , and frequency selectivity exist if we considered the MIMO wireless channel as a random process in all three domains. Simply put, the MIMO wireless channel response is approximately constant over its coherence space, time, and frequency, with *coherence* typically defined as the inverse of the expected root mean square (RMS) values of their respective Fourier transform dual. More detailed discussions on these measures can be found in [47, 48].

2.2.4 Space-Time-Frequency Correlation Function

The STF transfer function, as in (2.42), is generally regarded as a random process in *four* dimensions: transmit spatial, receive spatial, temporal, and spectral. Our channel estimator makes use of the complete space-time-frequency correlation function defined as follows.

Definition 2.2.2 (Space-Time-Frequency Correlation Function) *The space-time-frequency correlation function of a MIMO random channel is defined as*

$$\begin{aligned} &R_c(\mathbf{d}_{\mathrm{Rx},1}, \mathbf{d}_{\mathrm{Rx},2}, \mathbf{d}_{\mathrm{Tx},1}, \mathbf{d}_{\mathrm{Tx},2}, t_1, t_2, f_1, f_2) \\ &= E\left\{h_c(\mathbf{d}_{\mathrm{Rx},1}, \mathbf{d}_{\mathrm{Tx},1}, t_1, f_1) \cdot h_c(\mathbf{d}_{\mathrm{Rx},2}, \mathbf{d}_{\mathrm{Tx},2}, t_2, f_2)^*\right\} , \end{aligned} \tag{2.43}$$

■

The design of the antenna arrays is not a primary interest of this thesis, hence, we parameterized the displacement variable with the indices of the transmit/received antenna elements, and represent the STF correlation function as

$$R_c(i_1, i_2, j_1, j_2, t_1, t_2, f_1, f_2) = E\left\{h_c(i_1, j_1, t_1, f_1) \cdot h_c(i_2, j_2, t_2, f_2)^*\right\} , \qquad (2.44)$$

where i corrsponds to the receive antenna element index, and j corresponds to the transmit antenna element index.

So far, our channel model discussion is based on continuous-time descriptions. In a digital communciation systems, there is a need to represent the continuous-time channel description with a discrete-time equivalent. By using an uniform linear array (ULA) at the transmitter and receiver, we are effective sampling the wireless channel in the spatial-domain. Thus, the above transmitter/receiver indexing already represents the spatial-sampling of the wireless channel. In the next section, we present the time and frequency sampling adopted by OFDM to result in a discrete-time space-time-frequency channel description.

2.2.5 Discrete-Time Space-Time-Frequency Selective Channel

For OFDM systems, assuming perfect time synchronization, proper cyclic prefix usage, CIR is quasi-static over one OFDM symbol duration, and tolerable spectral leakage [19], we sampled the continuous-time CIR observation time variable at OFDM symbol rate, $1/T_{\text{OFDM}}$, and the delay variable at baud rate, $1/T_s$, to yield a discrete-time CIR representation as

$$g^{i,j}[n, l] = g_c(i, j, n \cdot T_{\text{OFDM}}, l \cdot T_s) , \quad \text{for } n = 0, \ldots, N_t - 1 , \text{ and } l = 0, \ldots, L-1 . \qquad (2.45)$$

where T_{OFDM} is the OFDM symbol duration including the cyclic prefix, N_t is the number of OFDM symbols transmitted in one radio frame,[6] and L is the overall channel length.

Since the input-output relationship of an OFDM system can be specified in the frequency domain, it is useful to transform the discrete-time channel impulse response into its frequency dual.

$$h^{i,j}[n,k] = \sum_{l=0}^{L-1} g^{i,j}[n,l]\, e^{\jmath 2\pi \frac{k \cdot l}{N_f}} \quad \text{for } k = 0, \ldots, N_f - 1\,, \tag{2.46}$$

where $h^{i,j}[n,k]$ is the channel frequency response (CFR) for the n^{th} OFDM symbol at the k^{th} subcarrier linking the i^{th} receive and the j^{th} transmit antennas.

2.2.6 3GPP Channel Model

Our simulations are based on the 3GPP spatial channel model [28], which specifies the $\alpha_n(t)$'s in (2.39) for each pair of transmit/receive antennas as

$$\begin{aligned} \alpha_n^{i,j}(t) = \sqrt{\frac{A}{M}} \sum_{m=1}^{M} \big(\exp(\jmath\, k\, d_j\, \sin(\theta_{n,m,AoD})) \cdot \exp(\jmath\, k\, d_i\, \sin(\theta_{n,m,AoA})) \cdot \\ \exp(\jmath\, \Phi_{n,m}) \exp(\jmath\, k\, \|\mathbf{v}\|\, \cos((\theta_{n,m,AoA} - \theta_v)t) \cdot \big)\,, \end{aligned} \tag{2.47}$$

where A represents the large-scale fading effects [41], $k = 2\pi/\lambda$ is the magnitude of the wavenumber, d_j is the incremental transmit antenna distance from the reference transmit antenna, d_i is the incremental receive antenna distance from the reference receive antenna, $\theta_{n,m,AoD}$ is the angle-of-departure (AoD) for the m^{th} subpath of the n^{th} path, $\theta_{n,m,AoA}$ is the angle-of-arrival (AoA) for the m^{th} subpath of the n^{th} path, $\Phi_{n,m}$ is the phase shift of the m^{th} subpath of the n^{th} path, $\|\mathbf{v}\|$ is the speed of the MS, and θ_v is the direction of the MS

[6] To be defi ned in chapter3.

movement. These parameters are generated according to the 3GPP specifications [28].

Chapter 3

Pilot-Symbol Assisted MIMO OFDM Channel Estimation

This chapter presents the channel estimators for MIMO-OFDM systems. The channel is assumed to be triply selective (i.e. joint space-time-frequency selective). We took the Bayesian estimation approach and derived the LMMSE estimator based on the second-order statistics of the channel response.

In the current literature, the two most common ways to perform channel estimation are 1) pilot-symbol assisted [18, 19, 21, 22, 49, 50], and 2) blind methods [51, 52]. We chose to follow the pilot-symbol assisted method due to its ability to acquire an accurate channel estimate quickly, and independent of the coding structure employed. The fallback of pilot-symbol assisted method is the loss of bandwidth as known pilot (training) symbols are inserted over the space-time-frequency domain. Although blind methods do not incur any reduction in bandwidth, these schemes rely on exploiting the coding structure of a particular system [51, 52], which are less versatile for general coding schemes.

In a practical communication system, pilot-symbols are used for timing and frequency synchronization as well. Therefore, with proper pilot sequence designs, the pilot symbols can serve multiple purposes in a communication system. Our proposed channel estimator will make use of pilot symbols because of a guaranteed acquisition of the channel response, as well as they can serve for other practical purposes.

3.1 Space-Time-Frequency PACE

We start our discussion with reordering the input-output relationship of the MIMO-OFDM system into a simpler matrix form.

3.1.1 MIMO OFDM System Model

Figure 3.1 illustrates the block diagram of the MIMO-OFDM system. Signal definitions in Figure 3.1 are as follows: $\mathbf{x}^j$ is the transmit symbol vector, also called OFDM symbol, at the j^{th} transmit antenna, $\mathbf{y}^i$ is the receive symbol vector at the i^{th} receive antenna, $\mathbf{g}^{i,j}$ is the DT CIR vector observed at the i^{th} receiver antenna due to the j^{th} transmit antenna, N_{Tx} is the number of transmitter (Tx) antennas, N_{Rx} is the number of receiver (Rx) antennas, $\{\cdot\}_d$ denotes the data portion of the respective vector, $\{\cdot\}_p$ denotes the pilot portion of the respective vector, $\{\cdot\}^{\mathcal{F}}$ denotes the frequency domain representation of the corresponding vector, and $\{\cdot\}^{\mathcal{T}}$ denotes the time domain representation of the corresponding vector.

The CFR is obtained by taking a DFT on the CIR. That is,

$$\mathbf{h}^{i,j} = \mathbf{F}\mathbf{g}^{i,j} \tag{3.1}$$

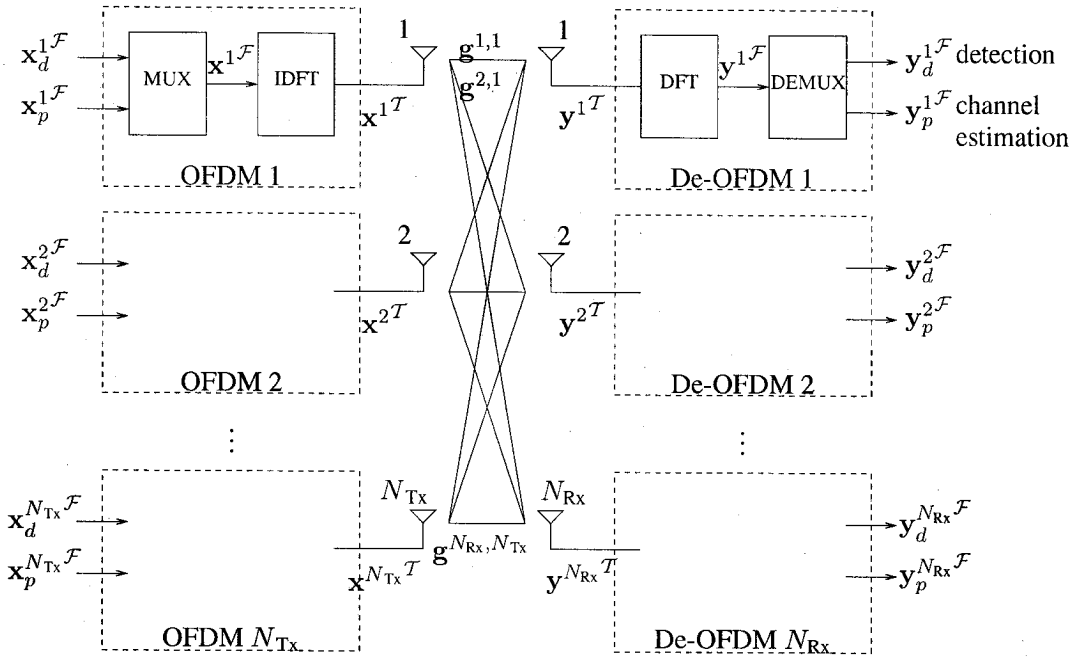


Figure 3.1 Block diagram of MIMO-OFDM system.

where $\mathbf{h}^{i,j}$ is the CFR connecting the j^{th} Tx element with the i^{th} Rx element, and $\mathbf{F}$ is the DFT matrix as defined in (2.28).

Note that the CP insertion (at Tx) and removal (at Rx) is implied but not shown in Figure 3.1.

In order to eliminate ISI, the following condition is required.

$$T_g \geq \tau_{\max}$$

where T_g denote the duration of the CP or the guard time, and $\tau_{\max}$ is the maximum propagation delay introduced by the channel.

Figure 3.2 is a graphical illustration of a transmission frame. We consider a frame transmission scheme similar to [53], where there are N_t OFDM words per frame. Each OFDM

word is composed of the useful source symbols of duration T_u (containing both data and pilot symbols) and the cyclic prefix of duration T_g (used to avoid ISI). The duration of one OFDM symbol is denoted as T_{OFDM}. In practice, the ends of the source symbols block might be filled with zeros to insert additional guard bands, in order to minimize the interference with devices operating in adjacent spectrum. In this thesis, we assume all N_f symbols are filled with either data or pilot symbols, occupying all of N_f OFDM subcarriers.

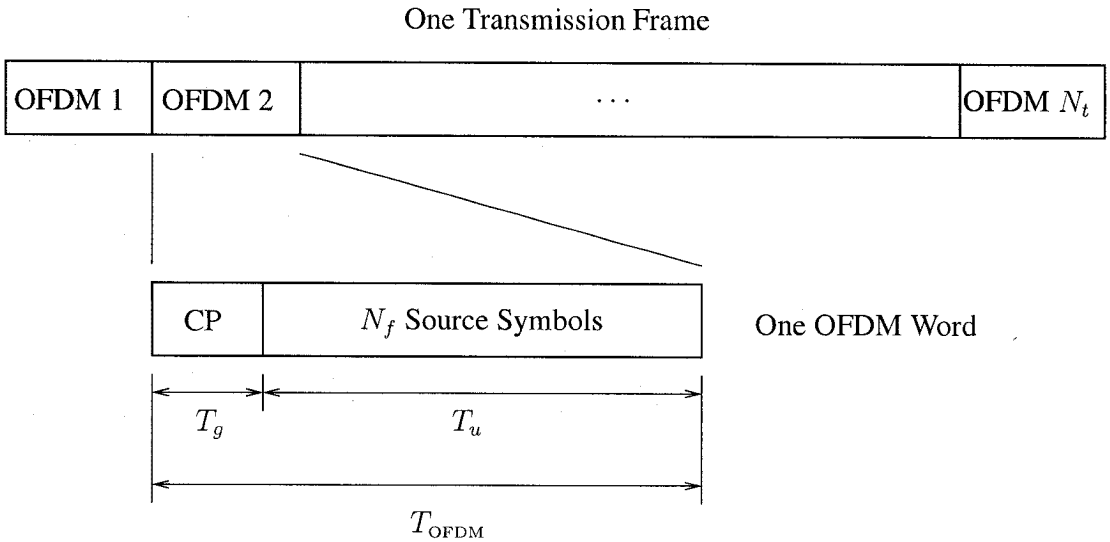


Figure 3.2 Timing information for one transmission frame.

We assumed that the channel is static over one OFDM word, but is time-varying over several OFDM words within a transmission frame. Typical duration of a frame is 2ms [53].

It is further assumed that the channel response are statistically independent from frame to frame.

With proper CP insertion and removal, we may represent the OFDM system in the frequency domain as shown in section 2.1.2.

For the rest of this manuscript, we drop the $\{\cdot\}^{\mathcal{F}}$ notation for simplicity. All signal quantities are in their frequency representation unless specified otherwise.

The received vector after the DFT operation is

$$\mathbf{y}^i(n) = \sum_{j=1}^{N_{\text{Tx}}} \mathbf{X}^j(n)\, \mathbf{h}^{i,j}(n) + \mathbf{v}^i(n), \qquad \text{for } i = 1, 2, \ldots, N_{\text{Rx}}\,, \tag{3.2}$$

where $\mathbf{y}^i(n) \in \mathbb{C}^{N_f \times 1}$ is the received frequency-domain vector at the i^{th} received antenna element at time n, $\mathbf{X}^j(n) = \text{diag}\big(\mathbf{x}_j(n)\big) \in \mathbb{C}^{N_f \times N_f}$ is the diagonal source symbol matrix at the j^{th} transmit antenna element at time n, $\mathbf{h}^{i,j}(n) \in \mathbb{C}^{N_f \times 1}$ is the channel frequency response vector linking the j^{th} transmit antenna element and the i^{th} receive antenna element at time n, $\mathbf{v}^i(n) \in \mathbb{C}^{N_f \times 1}$ is the filtered zero-mean AWGN vector at the i^{th} receive antenna element at time n, N_{Tx} is the number of transmit antennas, N_{Rx} is the number of receive antennas, and N_f is the number of OFDM subcarriers.

Collecting $\mathbf{y}^j(n)$ at N_t consecutive times gives,

$$\mathbf{y}^i = \sum_{j=1}^{N_{\text{Tx}}} \mathbf{X}^j\, \mathbf{h}^{i,j} + \mathbf{v}^i, \qquad \text{for } i = 1, 2, \ldots, N_{\text{Rx}}\,, \tag{3.3}$$

where $\mathbf{y}^i = \big[\mathbf{y}^i(n)^T,\, \mathbf{y}^i(n+1)^T,\, \ldots,\, \mathbf{y}^i(n+N_t-1)^T\big]^T \in \mathbb{C}^{N_t N_f \times 1}$ is the time stacked received vector at the i^{th} receive antenna, $\mathbf{X}^j = \text{diag}\big(\mathbf{X}^j(n),\, \mathbf{X}^j(n+1),\, \ldots,\, \mathbf{X}^j(n+N_t-1)\big) \in \mathbb{C}^{N_t N_f \times N_t N_f}$ is the diagonal source symbol matrix at N_t consecutive times, at the j^{th} transmit antenna element, $\mathbf{h}^{i,j} = \big[\mathbf{h}^{i,j}(n)^T,\, \mathbf{h}^{i,j}(n+1)^T,\, \ldots,\, \mathbf{h}^{i,j}(n+N_t-1)^T\big]^T \in \mathbb{C}^{N_t N_f \times 1}$ is the time stacked channel frequency response vector linking the j^{th} transmit antenna element and the i^{th} receive antenna element, and $\mathbf{v}^i = \big[\mathbf{v}^i(n)^T,\, \mathbf{v}^i(n+1)^T,\, \ldots,\, \mathbf{v}^i(n+N_t-1)^T\big]^T \in \mathbb{C}^{N_t N_f \times 1}$ is the time stacked filtered additive noise vector at the i^{th} receive antenna element.

Writing the summation in (3.3) in matrix form yields,

$$\mathbf{y}^i = \widetilde{\mathbf{X}}\,\mathbf{h}^i + \mathbf{v}^i, \qquad \text{for } i = 1, 2, \ldots, N_{\text{Rx}}\,, \tag{3.4}$$

where $\mathbf{y}^i$ and $\mathbf{v}^i$ are the same as in (3.3), $\widetilde{\mathbf{X}} = [\mathbf{X}^1, \mathbf{X}^2, \ldots, \mathbf{X}^{N_{\text{Tx}}}] \in \mathbb{C}^{N_t N_f \times N_{\text{Tx}} N_t N_f}$ is a horizontal cascade of N_{Tx} diagonal matrices, and $\mathbf{h}^i = \left[\mathbf{h}^{i,1^T}, \mathbf{h}^{i,2^T}, \ldots, \mathbf{h}^{i,N_{\text{Tx}}T}\right]^T \in \mathbb{C}^{N_{\text{Tx}} N_t N_f \times 1}$ is the Tx spatial time-frequency channel response.

Finally, collecting the received vectors at all the N_{Rx} receive antenna elements gives,

$$\mathbf{y} = \mathbf{X}\,\mathbf{h} + \mathbf{v}\,, \tag{3.5}$$

where $\mathbf{y} = \left[\mathbf{y}^{1^T}, \mathbf{y}^{2^T}, \ldots, \mathbf{y}^{N_{\text{Rx}}T}\right]^T \in \mathbb{C}^{N_{\text{Rx}} N_t N_f \times 1}$ is the Rx spatially stacked time-frequency received vector, $\mathbf{X} = (\mathbf{I}_{N_{\text{Rx}}} \otimes \widetilde{\mathbf{X}}) \in \mathbb{C}^{N_{\text{Rx}} N_t N_f \times N_{\text{Rx}} N_{\text{Tx}} N_t N_f}$ is a banded matrix containing the Tx spatial time-frequency source symbols,[1] $\mathbf{h} = \left[\mathbf{h}^{1^T}, \mathbf{h}^{2^T}, \ldots, \mathbf{h}^{N_{\text{Rx}}T}\right]^T \in \mathbb{C}^{N_{\text{Tx}} N_{\text{Rx}} N_t N_f \times 1}$ is the full Tx-Rx spatial time-frequency channel response vector, and $\mathbf{v} = \left[\mathbf{v}^{1^T}, \mathbf{v}^{2^T}, \ldots, \mathbf{v}^{N_{\text{Rx}}T}\right]^T \in \mathbb{C}^{N_{\text{Rx}} N_t N_f \times 1}$ is the Rx spatial time-frequency AWGN vector.

Figure 3.3 is a graphical illustration on the structure of $\mathbf{y}^T$. Note that each received OFDM symbol is a superposition of N_{Tx} transmitted OFDM symbols.

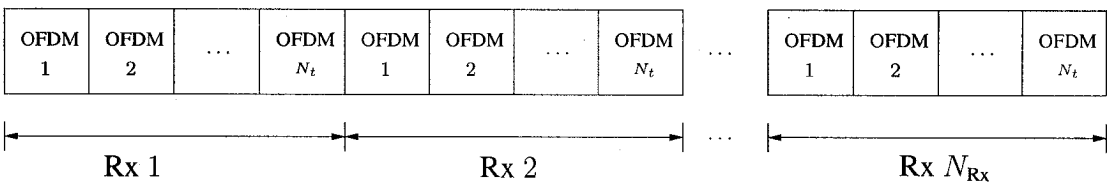


Figure 3.3 Graphical illustration on the structure of $\mathbf{y}^T$ in (3.5).

[1] $\otimes$ denotes the Kronecker product [34].

3.1.2 STF LMMSE Derivations

We assume that pilot symbols are placed uniformly throughout the space-time-frequency domains, and we need to extract the pilot components of the received vector from (3.5) to perform our channel estimation. The extraction can be done by multiplying the received vector by a pilot selection matrix, $\mathbf{Q} \in \mathfrak{B}^{N_p \times N_f N_t N_{\text{Rx}}}$, where $\mathfrak{B}$ denotes the binary number space containing the elements $\{0, 1\}$, and N_p denotes the number of pilot symbols per transmission frame. This pilot selection matrix contains 1's at the pilot locations, and 0's at the data locations. The $(N_p \times N_f N_t N_{\text{Rx}})$ pilot selection matrix is defined mathematically as

$$[\mathbf{Q}]_{i,j} = \begin{cases} 1, & (i,j) \in \Omega_p \\ 0, & \text{otherwise.} \end{cases} \tag{3.6}$$

where Ω_p is the set of pilot symbol locations in the STF domain.

Figure 3.4 is a visual illustration of the pilot and data structure on the received signal, $\mathbf{y}^T$. From this figure, we see that each row of $\mathbf{Q}$ consists of a single 1 located at successive pilot locations. For example, the n^{th} row of $\mathbf{Q}$ is a row of zeros except there is a 1 at the n^{th} pilot location.

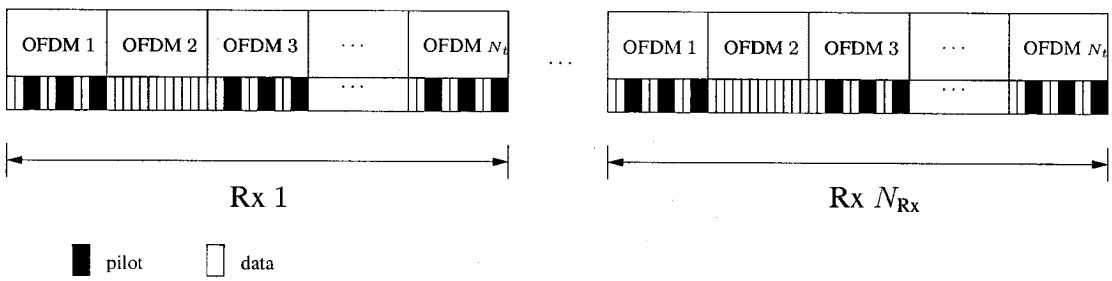


Figure 3.4 Graphical illustration on the pilot and data structure of $\mathbf{y}^T$.

Then the $(N_p \times 1)$ received pilot symbols, $\mathbf{y}_p$, are represented by

$$\begin{aligned} \mathbf{y}_p &= \mathbf{Q}\,\mathbf{y} \\ &= \mathbf{Q}\,\mathbf{X}\,\mathbf{h} + \mathbf{Q}\,\mathbf{v} \\ &= \breve{\mathbf{X}}_p\,\mathbf{h} + \mathbf{v}_p \end{aligned} \tag{3.7}$$

where $\breve{\mathbf{X}}_p \in \mathbb{C}^{N_p \times N_{\mathrm{Rx}} N_{\mathrm{Tx}} N_t N_f}$ is the pilot symbol matrix, and $\mathbf{v}_p \in \mathbb{C}^{N_p \times 1}$ is the vector denoting the additive noise samples at the pilot locations. Second equality is obtained by substituting (3.5).

Note that the structure of $\widetilde{\mathbf{X}}$ from (3.4) looks like

$$\widetilde{\mathbf{X}} = \underbrace{\left[\begin{array}{c|c|c|c} \times & \times & & \times \\ & \times & \cdots & \\ \times & & & \times \\ & \times & & \end{array}\right]}_{N_{\mathrm{Tx}} \text{ diagonal matrices}}. \tag{3.8}$$

Since $\mathbf{X} = \mathbf{I}_{N_{\mathrm{Rx}}} \otimes \widetilde{\mathbf{X}}$, removing rows of $\mathbf{X}$ is similar to removing rows of $\widetilde{\mathbf{X}}$. From (3.8), we see that whenever we remove a row of $\widetilde{\mathbf{X}}$, the remaining matrix would have N_{Tx} columns with zeros only. Hence, when we remove $N_{\mathrm{Rx}} N_t N_f - N_p$ rows of $\mathbf{X}$ in (3.7), the resulting $\breve{\mathbf{X}}_p$ contains $N_{\mathrm{Tx}}(N_{\mathrm{Rx}} N_t N_f - N_p)$ columns of zeros. In other words, when we retain only N_p rows of $\mathbf{X}$, there are only $N_{\mathrm{Tx}} N_p$ columns with non-zero entries. This means that we can further simplify (3.7) by removing those columns with zeros only, and correspondingly remove those components of $\mathbf{h}$ that would otherwise be multiplied with zeros. With this simplification, we have

$$\mathbf{y}_p = \mathbf{X}_p\,\mathbf{h}_p + \mathbf{v}_p \tag{3.9}$$

where $\mathbf{X}_p \in \mathbb{C}^{N_p \times N_{\mathrm{Tx}} N_p}$ is the reduced-dimension pilot symbol matrix, and $\mathbf{h}_p \in \mathbb{C}^{N_{\mathrm{Tx}} N_p \times 1}$

is the channel space-time-frequency response vector at the pilot locations.

We want to establish a good estimate of the channel response, $\hat{h}$, via linear processing $\mathbf{y}_p$.

$$\hat{\mathbf{h}} = \mathbf{W}\,\mathbf{y}_p \tag{3.10}$$

where the linear processor, $\mathbf{W} \in \mathbb{C}^{N_{\mathrm{Rx}}N_{\mathrm{Tx}}N_tN_f \times N_p}$ is to be chosen appropriately.

Definition 3.1.1 (Mean-Square-Error Criterion) *Define the estimation error vector as*

$$\mathbf{e} = \mathbf{h} - \hat{\mathbf{h}}. \tag{3.11}$$

The mean square-error (MSE) cost function, is computed as

$$\mathrm{MSE} = E\left\{\|\mathbf{e}\|^2\right\} \tag{3.12}$$

where $E\{\cdot\}$ is the statistical expectation operator [45].

■

We would like to choose the linear processor, $\mathbf{W}$, such that the MSE is minimized [54]. That is, we want to choose $\mathbf{W}$ according to

$$\mathbf{W}_{\mathrm{LMMSE}} = \arg\min_{\mathbf{W}} E\left\{\|\mathbf{e}\|^2\right\}. \tag{3.13}$$

A necessary condition for minimizing MSE is [30]

$$\frac{\partial\,\mathrm{MSE}}{\partial\,\mathrm{w}_{i,j}} = 0 \tag{3.14}$$

where $\mathrm{w}_{i,j}$ is the i^{th} and j^{th} component of the linear processor matrix, $\mathbf{W}$. The partial

derivative is with respect to a complex quantity and is defined as [30]

$$\frac{\partial \,\mathrm{MSE}}{\partial \,\mathrm{w}_{i,j}} = \frac{1}{2} \cdot \left(\frac{\partial \,\mathrm{MSE}}{\partial \,\Re\{\mathrm{w}_{i,j}\}} - \frac{\partial \,\mathrm{MSE}}{\partial \,\Im\{\mathrm{w}_{i,j}\}} \right). \tag{3.15}$$

The error vector can be expressed as

$$\|\mathbf{e}\|^2 = \sum_{i=1}^{N_{\mathrm{tot}}} \underbrace{|\mathrm{e}_i|^2}_{\geq 0}, \tag{3.16}$$

where $N_{\mathrm{tot}} = N_f N_t N_{\mathrm{Tx}} N_{\mathrm{Rx}}$ is the dimension of $\mathbf{e}$.

With (3.16), minimizing (3.12) is equivalent to minimizing the MSE of each component of $\mathbf{e}$ individually, i.e.

$$\arg\min_{\mathbf{W}} E\left\{\|\mathbf{e}\|^2\right\} = \left\{ \arg\min_{\mathbf{w}_i^H} E\left\{|\mathrm{e}_i|^2\right\}, \quad i = 1, \ldots, N_{\mathrm{tot}} \right\}, \tag{3.17}$$

where $\mathbf{w}_i^H$ denotes the i^{th} row of $\mathbf{W}$.

Denote the MSE per component as

$$\mathrm{MSE}_i = E\left\{|\mathrm{e}_i|^2\right\}, \qquad \text{for } i = 1, \ldots, N_{\mathrm{tot}}. \tag{3.18}$$

With

$$\mathrm{e}_i = \mathbf{h}_i - \mathbf{w}_i^H \mathbf{y}_p. \tag{3.19}$$

Following the Appendix of [30], we take the complex gradient of (3.18) with respect to

$\mathbf{w}_i$ to find a stationary point.

$$\begin{aligned}\nabla_{\mathbf{w}_i} E\left\{|\mathrm{e}_i|^2\right\} &= \nabla_{\mathbf{w}_i} E\left\{|\mathrm{h}_i|^2 - \mathrm{h}_i^* \mathbf{w}_i^H \mathbf{y}_p - \mathrm{h}_i \mathbf{y}_p^H \mathbf{w}_i + \mathbf{w}_i^H \mathbf{y}_p \mathbf{y}_p^H \mathbf{w}_i\right\} \\ &= -E\left\{\mathrm{h}_i \mathbf{y}_p\right\} + \mathbf{w}_i^H E\left\{\mathbf{y}_p \mathbf{y}_p^H\right\} \\ &= \mathbf{0}.\end{aligned} \tag{3.20}$$

The LMMSE solution is

$$\mathbf{w}_{i,\mathrm{LMMSE}}^H = E\left\{\mathrm{h}_i \mathbf{y}_p^H\right\} E\left\{\mathbf{y}_p \mathbf{y}_p^H\right\}. \tag{3.21}$$

Finally the complete linear processor, $\mathbf{W}$, is obtained by stacking it rows in the form of (3.21)

$$\begin{aligned}\mathbf{W}_{\mathrm{LMMSE}} &= \begin{bmatrix} \mathbf{w}_{1,\mathrm{LMMSE}}^H \\ \mathbf{w}_{2,\mathrm{LMMSE}}^H \\ \vdots \\ \mathbf{w}_{N_{\mathrm{tot}},\mathrm{LMMSE}}^H \end{bmatrix} \\ &= E\left\{\mathbf{h}\mathbf{y}_p^H\right\} E\left\{\mathbf{y}_p \mathbf{y}_p^H\right\} \\ &= \mathbf{R}_{\mathbf{h}\mathbf{y}_p} \mathbf{R}_{\mathbf{y}_p}^{-1}.\end{aligned} \tag{3.22}$$

where $\mathbf{R}_{\mathbf{h}\mathbf{y}_p} = E\left\{\mathbf{h}\mathbf{y}_p^H\right\}$ is the cross-correlation matrix between the channel response and the received pilot symbols, and $\mathbf{R}_{\mathbf{y}_p} = E\left\{\mathbf{y}_p \mathbf{y}_p^H\right\}$ is the auto-correlation matrix of the received pilot symbols. In the following, we simply denote the LMMSE linear processor as $\mathbf{W}$ instead of $\mathbf{W}_{\mathrm{LMMSE}}$.

With this choice of $\mathbf{W}$, the MSE is

$$\mathrm{MSE}_{\mathbf{W}} = E\left\{\|\mathbf{e}\|^2\right\} = E\left\{\|\mathbf{h} - \mathbf{R}_{\mathbf{hy}_p}\mathbf{R}_{\mathbf{y}_p}^{-1}\mathbf{y}_p\|^2\right\} = \mathrm{tr}\left(\mathbf{R}_{\mathbf{h}} - \mathbf{R}_{\mathbf{hy}_p}\mathbf{R}_{\mathbf{y}_p}^{-1}\mathbf{R}_{\mathbf{hy}_p}^{H}\right) \quad (3.23)$$

Substituting (3.9) into the correlation matrices, $\mathbf{R}_{\mathbf{hy}_p}$ and $\mathbf{R}_{\mathbf{y}_p}$, gives

$$\begin{aligned}\mathbf{R}_{\mathbf{hy}_p} &= E\left\{\mathbf{h}\mathbf{y}_p^H\right\} \\ &= E\left\{\mathbf{h}(\mathbf{X}_p\,\mathbf{h}_p + \mathbf{v}_p)^H\right\} \\ &= E\left\{\mathbf{h}\mathbf{h}_p^H\right\}\mathbf{X}_p^H \\ &= \mathbf{R}_{\mathbf{hh}_p}\mathbf{X}_p^H\,,\end{aligned} \quad (3.24)$$

and

$$\begin{aligned}\mathbf{R}_{\mathbf{y}_p} &= E\left\{\mathbf{y}_p\mathbf{y}_p^H\right\} \\ &= E\left\{(\mathbf{X}_p\,\mathbf{h}_p + \mathbf{v}_p)(\mathbf{X}_p\,\mathbf{h}_p + \mathbf{v}_p)^H\right\} \\ &= \mathbf{X}_p E\left\{\mathbf{h}_p\mathbf{h}_p^H\right\}\mathbf{X}_p^H + E\left\{\mathbf{v}_p)\mathbf{v}_p^H\right\} \\ &= \mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H + \mathbf{R}_{\mathbf{v}_p}\,.\end{aligned} \quad (3.25)$$

In deriving (3.24) and (3.25), we assumed that the channel response is uncorrelated with the zero mean additive noise.

Substituting (3.24) and (3.25) into (3.23) yields the final LMMSE estimator

$$\mathbf{W} = \mathbf{R}_{\mathbf{hh}_p}\mathbf{X}_p^H\left(\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H + \mathbf{R}_{\mathbf{v}_p}\right)^{-1}. \quad (3.26)$$

The resulting MSE is given as

$$\mathrm{MSE}_{\mathbf{W}} = \mathrm{tr}\Big(\mathbf{R}_{\mathbf{h}} - \mathbf{R}_{\mathbf{hh}_p}\mathbf{X}_p^H\left(\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H + \mathbf{R}_{\mathbf{v}_p}\right)^{-1}\mathbf{X}_p\mathbf{R}_{\mathbf{hh}_p}^H\Big). \tag{3.27}$$

Typically, the zero mean additive noise is spatially, spectrally, and temporally white, such that

$$\mathbf{R}_{\mathbf{v}_p} = \sigma_{\mathrm{v}}^2\mathbf{I}. \tag{3.28}$$

In this case, the LMMSE estimator becomes

$$\mathbf{W} = \mathbf{R}_{\mathbf{hh}_p}\mathbf{X}_p^H\left(\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H + \sigma_{\mathrm{v}}^2\mathbf{I}\right)^{-1}, \tag{3.29}$$

with an MSE of

$$\mathrm{MSE}_{\mathbf{W}} = \mathrm{tr}\Big(\mathbf{R}_{\mathbf{h}} - \mathbf{R}_{\mathbf{hh}_p}\mathbf{X}_p^H\left(\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H + \sigma_{\mathrm{v}}^2\mathbf{I}\right)^{-1}\mathbf{X}_p\mathbf{R}_{\mathbf{hh}_p}^H\Big). \tag{3.30}$$

It should be emphasized that since our STF channel estimator is based on the received pilot vectors from *all* receive antennas, it is capable of exploiting any Rx spatial correlations if it exists. In usual circumstances, however, rich scattering is experienced around the MS such that Rx spatial correlation is nearly non-existent. In certain cases where the scatters are far from the MS, our channel estimator can exploit the receive correlation and provide better channel estimates.

3.1.3 TF PACE as a Mismatched STF

We demonstrate the generality of our STF channel estimator by reducing it into the time-frequency MIMO-OFDM channel estimator reported in [29]. We will also compare the

performance of our STF channel estimator with this time-frequency channel estimator for MIMO-OFDM systems.

A time-frequency channel estimator does not exploit any spatial correlations at all. Thus, the assumption in deriving the time-frequency channel estimator in [29] is that there no Tx/Rx spatial correlation. This results in $\mathbf{R}_{\mathbf{h}_p}$ being a block diagonal matrix,

$$\mathbf{R}_{\mathbf{h}_p} = \begin{bmatrix} \mathbf{R}_{\mathbf{h}_p}^{1,1} & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \mathbf{R}_{\mathbf{h}_p}^{1,N_{\mathrm{Tx}}} & \ddots & & \vdots \\ \vdots & & \ddots & \mathbf{R}_{\mathbf{h}_p}^{2,1} & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & \mathbf{R}_{\mathbf{h}_p}^{N_{\mathrm{Rx}},N_{\mathrm{Tx}}} \end{bmatrix}, \tag{3.31}$$

where $\mathbf{R}_{\mathbf{h}_p}^{i,j} \in \mathbb{C}^{N_p/N_{\mathrm{Rx}} \times N_p/N_{\mathrm{Rx}}}$ is the time-frequency correlation matrix for the $(i,j)^{th}$ pair of receive/transmit antennas.

The pilot symbol matrix has the following structure:

$$\mathbf{X}_p = \begin{bmatrix} \begin{pmatrix} \mathbf{X}_p^1 & \cdots & \mathbf{X}_p^{N_{\mathrm{Tx}}} \end{pmatrix} & & 0 & & \cdots & & 0 \\ & 0 & & \begin{pmatrix} \mathbf{X}_p^1 & \cdots & \mathbf{X}_p^{N_{\mathrm{Tx}}} \end{pmatrix} & & \ddots & \vdots \\ \vdots & & \cdots & & & \ddots & & 0 \\ & 0 & & \cdots & & 0 & \begin{pmatrix} \mathbf{X}_p^1 & \cdots & \mathbf{X}_p^{N_{\mathrm{Tx}}} \end{pmatrix} \end{bmatrix}, \tag{3.32}$$

where $\mathbf{X}_p^j \in \mathbb{C}^{N_p/N_{\mathrm{Rx}} \times N_p/N_{\mathrm{Rx}}}$ is the time-frequency pilot symbols for the j^{th} transmit antenna.

Computing the $\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H$ term in (3.29) gives,

$$\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H = \begin{bmatrix} \sum_{j=1}^{N_{\mathrm{Tx}}} \mathbf{X}_p^j \mathbf{R}_{\mathbf{h}_p}^{1,j} \mathbf{X}_p^{jH} & 0 & \cdots & 0 \\ 0 & \sum_{j=1}^{N_{\mathrm{Tx}}} \mathbf{X}_p^j \mathbf{R}_{\mathbf{h}_p}^{2,j} \mathbf{X}_p^{jH} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sum_{j=1}^{N_{\mathrm{Tx}}} \mathbf{X}_p^j \mathbf{R}_{\mathbf{h}_p}^{N_{\mathrm{Rx}},j} \mathbf{X}_p^{jH} \end{bmatrix}. \tag{3.33}$$

Since $\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H$ is a block diagonal matrix, it is clear that

$$\begin{aligned} &(\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H + \sigma_{\mathrm{v}}^2\mathbf{I})^{-1} \\ &= \mathrm{diag}\left(\left(\sum_{j=1}^{N_{\mathrm{Tx}}} \mathbf{X}_p^j \mathbf{R}_{\mathbf{h}_p}^{1,j} \mathbf{X}_p^{jH} + \sigma_{\mathrm{v}}^2\mathbf{I} \right)^{-1}, \cdots, \left(\sum_{j=1}^{N_{\mathrm{Tx}}} \mathbf{X}_p^j \mathbf{R}_{\mathbf{h}_p}^{N_{\mathrm{Rx}},j} \mathbf{X}_p^{jH} + \sigma_{\mathrm{v}}^2\mathbf{I} \right)^{-1} \right). \end{aligned} \tag{3.34}$$

When there is no Tx/Rx spatial correlation,

$$\mathbf{R}_{\mathbf{h}\mathbf{h}_p} = \mathrm{diag}(\mathbf{R}_{\mathbf{h}\mathbf{h}_p}^{1,1}, \cdots, \mathbf{R}_{\mathbf{h}\mathbf{h}_p}^{1,N_{\mathrm{Tx}}}, \mathbf{R}_{\mathbf{h}\mathbf{h}_p}^{2,1}, \cdots, \mathbf{R}_{\mathbf{h}\mathbf{h}_p}^{N_{\mathrm{Rx}},N_{\mathrm{Tx}}}). \tag{3.35}$$

Substituting (3.32), (3.34), and (3.35) into (3.29) yields the diagonal band matrix

$$\mathbf{W}_{\mathrm{TF}} = \begin{bmatrix} \mathbf{R}_{\mathbf{hh}_p}^{1,1}\mathbf{X}_p^{1\,H}\left(\sum_{j=1}^{N_{\mathrm{Tx}}}\mathbf{X}_p^j\mathbf{R}_{\mathbf{h}_p}^{1,j}\mathbf{X}_p^{j\,H}+\sigma_{\mathrm{v}}^2\mathbf{I}\right)^{-1} & & \\ \vdots & & \\ \mathbf{R}_{\mathbf{hh}_p}^{1,N_{\mathrm{Tx}}}\mathbf{X}_p^{N_{\mathrm{Tx}}\,H}\left(\sum_{j=1}^{N_{\mathrm{Tx}}}\mathbf{X}_p^j\mathbf{R}_{\mathbf{h}_p}^{1,j}\mathbf{X}_p^{j\,H}+\sigma_{\mathrm{v}}^2\mathbf{I}\right)^{-1} & & \\ & \ddots & \\ & & \mathbf{R}_{\mathbf{hh}_p}^{N_{\mathrm{Rx}},1}\mathbf{X}_p^{1\,H}\left(\sum_{j=1}^{N_{\mathrm{Tx}}}\mathbf{X}_p^j\mathbf{R}_{\mathbf{h}_p}^{N_{\mathrm{Rx}},j}\mathbf{X}_p^{j\,H}+\sigma_{\mathrm{v}}^2\mathbf{I}\right)^{-1} \\ & & \vdots \\ & & \mathbf{R}_{\mathbf{hh}_p}^{N_{\mathrm{Rx}},N_{\mathrm{Tx}}}\mathbf{X}_p^{N_{\mathrm{Tx}}\,H}\left(\sum_{j=1}^{N_{\mathrm{Tx}}}\mathbf{X}_p^j\mathbf{R}_{\mathbf{h}_p}^{N_{\mathrm{Rx}},j}\mathbf{X}_p^{j\,H}+\sigma_{\mathrm{v}}^2\mathbf{I}\right)^{-1} \end{bmatrix}, \tag{3.36}$$

which corresponds to the time-frequency channel estimation method in [29].

To obtain a TF channel estimator from an STF estimator, we modified the channel correlation matrices to reflect the no-spatial-correlation assumption used in [29]. We can view the TF channel estimator as an STF channel estimator with correlation mismatch. Hence, our STF channel estimator cannot perform worse than the TF estimator in terms of MSE.

3.2 Pilot-Symbol Design Issues

In this section, we discuss some pilot symbol (sequence) design issues for MIMO-OFDM channel estimation. In order to obtain a good channel estimation, pilot symbols have to be placed according to the 3-dimensional sampling theorem. Since multiple antennas are used, pilot sequences has to be jointly designed for all N_{Tx} transmit antennas such that Tx spatial signatures can be separated at the receiver.

The amount of power and bandwidth allocated to the pilot symbols also plays an important role in power and bandwidth limited scenarios. We describe the system implications of these parameters in this section.

3.2.1 Pilot Placement

Pilot symbols are spaced in the STF domain for our STF channel estimator. A typical pilot symbol placement is shown in Figure 3.5. In this thesis, we focus on the rectangular layout of pilot symbol placement. Other sampling patterns, like hexagonal, might yield better efficiency in covering the entire STF space, but was not investigated here. More details in multi-dimensional sampling pattern and efficiency are discussed in [55].

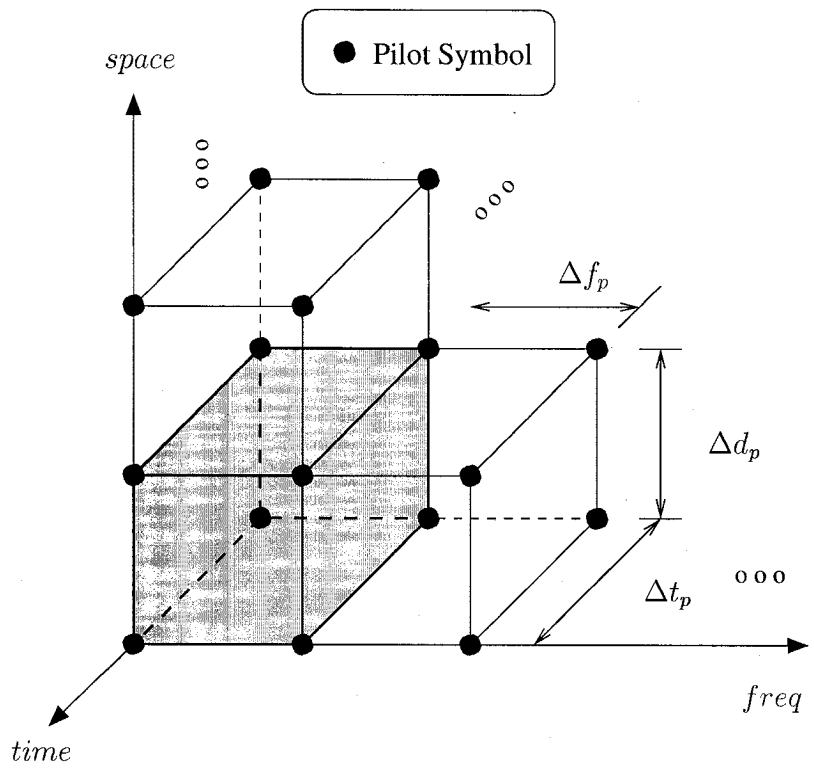


Figure 3.5 Pilot symbol placement for LMMSE channel estimation.

In Figure 3.5, the separation between pilots in the spatial, temporal, and spectral domains are Δd_p, Δt_p, and Δf_p, respectively. These separations must be chosen appropriately as small separations would introduce large training overhead into the system, yet large separations introduce inaccuracies in the channel estimates.

Wireless channels are typically limited in the Doppler frequency, propagation delay, and wavenumber (propagation direction) domains. Since these quantities are Fourier transform pairs with the observation time, ordinary frequency, and spatial position domains [30, 43], the sampling theorem applies to this random process representing the wireless channel.[2]

From the multi-dimensional sampling theorem [47], the pilot STF spacing must satisfy the following criteria to avoid aliasing:

$$\Delta d_p \leq \frac{\lambda}{|\cos(\theta_u) - \cos(\theta_l)|} \tag{3.37a}$$

$$\Delta t_p \leq \frac{1}{2\nu_{\max}} \tag{3.37b}$$

$$\Delta f_p \leq \frac{1}{\tau_{\max}} \tag{3.37c}$$

where λ is the wavelength of the propagating wave, θ_u is the upper bound on the angle of departure, θ_l is the lower bound on the angle of departure, $\nu_{\max}$ is the maximum Doppler frequency, and $\tau_{\max}$ is the maximum propagation delay.

In the worst case, the AoD ranges from $0°$ to $180°$.[3] In this case, the spatial separation between the pilots is $\Delta d_p \leq \frac{\lambda}{2}$.

Notice that for *small* AoD, the spatial sampling upper bound could be many multiples

[2] Sampling theorem of a band-limited random process was described in [56]. The multi-dimensional sampling theorem is a simple extension of this one-dimensional case. See also [55] for the multi-dimensional sampling theorem of deterministic signals.

[3] One-dimensional ULA's contains left-right ambiguity, therefore the spatial response is the same for $[0°, 180°]$ as $[-180°, 0°]$.

of λ. For example, when the MS is located at the broadside ($90°$) of the BS antenna array, and with $5°$ angle spread at the BS, the spatial sampling upper bound is 11.5λ. In this case, it is possible to train only a subset of the transmit antennas when they are spaced less than 11.5λ apart.

In the sequel, the pilots separation in space, time, and frequency are normalized with respect to the Tx antenna element spacings, d_{Tx}, to the OFDM word duration, T_{OFDM}, and to the OFDM subcarrier spacing, Δf, respectively.

$$\Delta D_p = \frac{\Delta d_p}{d_{\text{Tx}}} \leq \frac{\lambda}{|\cos(\theta_u) - \cos(\theta_l)|\, d_{\text{Tx}}} \tag{3.38a}$$

$$\Delta T_p = \frac{\Delta t_p}{T_{\text{OFDM}}} \leq \frac{1}{2\nu_{\max} T_{\text{OFDM}}} \tag{3.38b}$$

$$\Delta F_p = \frac{\Delta f_p}{\Delta f} \leq \frac{1}{\tau_{\max} \Delta f} \tag{3.38c}$$

With (3.38), the pilot blocks[4] are placed at multiples of ΔD_p, ΔT_p, and ΔF_p in the Tx space, time, and frequency domains, respectively. Figure 3.2.1 illustrates the pilot blocks placement with the time and frequency separation explicitly shown. The space separation is hidden but implied in this figure.

3.2.2 Pilot Sequence Design

When multiple transmit antennas are employed, there will be spatial interference between the pilot symbols. To counter the spatial interference, a properly designed pilot block is used as pilot symbols. In 1999, Guey et al. studied the pilot sequence design issue for flat Rayleigh fading channels with transmitter diversity [57]. They showed that the usage

[4] A sequence of pilot symbols are required for MIMO channel estimation in order to suppress the spatial interference at the receiver.

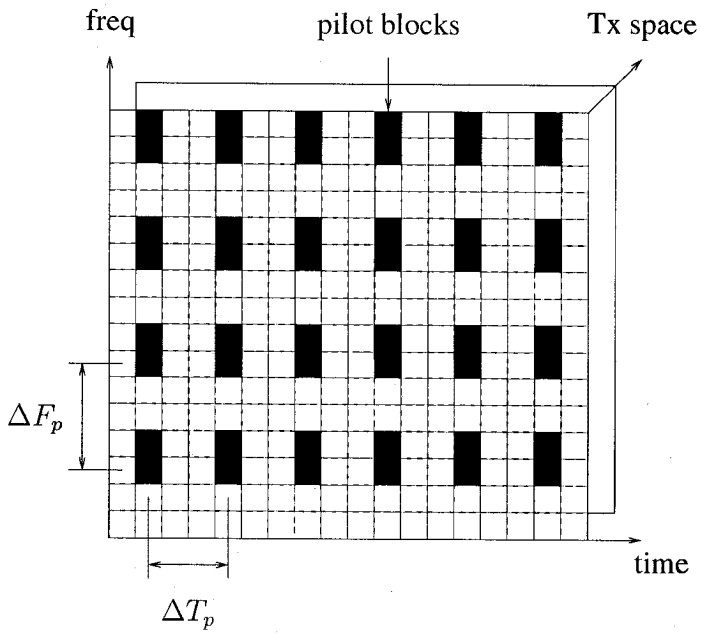


Figure 3.6 Pilot placement on discrete space-time-frequency grid.

of constant modulus orthogonal pilot blocks satisfies the following desirable properties on the channel estimation error covariance matrix, $\mathrm{Cov}(\mathbf{e}) \triangleq E\{\mathbf{e}\mathbf{e}^H\} = E\{(\mathbf{h} - \hat{\mathbf{h}})(\mathbf{h} - \hat{\mathbf{h}})^H\}$ [57]:

1. Total MSE over all channel responses, $\mathrm{tr}(\mathrm{Cov}(\mathbf{e}))$, is minimized.

2. Estimation errors are as evenly distributed throughout all channel responses as possible.

3. Cross-correlation of estimation errors, i.e. off diagonal of $\mathrm{Cov}(\mathbf{e})$, are kept as small as possible when the pilot block length, L_p, is chosen to be a multiple of the number of transmit antennas, N_{Tx}.

The columns of the unitary DFT matrix, defined in (2.28), is a practical pilot sequence which satisfies all of the above properties. We choose the DFT matrix size to be N_{Tx} in

order to maintain property 3 mentioned above and minimize the training overhead. Each column of the unitary DFT matrix would correspond to the pilot sequence used for each transmit antennas.

For MIMO-OFDM systems, we can place the pilot block over adjacent subcarriers, or over adjacent OFDM symbols on certain subcarriers. A large number of subcarriers are used in a typical OFDM system. Therefore, the channel is almost constant over several subcarriers. Also, we use an SFBC-OFDM system to evaluate the error-rate performance in the end, so we should also place the pilot sequences over adjacent subcarriers by the same reasoning in section 2.1.3.

3.2.3 Power Allocation to Pilot Symbols

Usually, the total allowable transmit power is strictly governed. For a pilot-symbol assisted communication system, we must allocate some of the transmit power to the pilot symbols such that we get a good enough channel estimate for coherent detection. Allocating too little power on the pilot symbols would produce bad channel estimates. And since we have seen that maximum likelihood decoding assumes perfect channel knowledge, it is logical to deduce that the decoding performance would be badly impaired when the decoder uses bad channel estimates for decoding.

On the other hand, allocating too much power on the pilot symbols might yield a highly accurate channel estimate. However, since the total allowable transmit power is fixed, the power for data transmission is proportionally reduced. This results in operating in a lower E_b/N_0, thus resulting in bad error-rate performance also.

It is clear that for any pilot-symbol assisted communication systems, there exists an optimal power allocation for the pilot symbols. However, in a practical setting, the percentage

of power allocated to pilots is usually fixed. For example, the OFDM feasibility proposal for 3GPP has fixed the percentage of power allocated to the pilots to be 10% [53]. In this case, we can try to optimize the overall performance by varying the distribution of the pilot power across all pilot symbols in the STF domain. Optimizing the pilot power distribution is only possible when the transmitter has some prior knowledge of the channel. If no prior channel knowledge is available at the transmitter, it is best to equally distribute the power across all pilot symbols, which is our approach in this work.

3.2.4 STF Resource Allocation to Pilot Symbols

The insertion of pilot symbols reduces the bandwidth efficiency of the communication system. Since there is one pilot sequence of length N_{Tx} within the volume spanned by the pilot space-time-frequency separation, $\Delta D_p \cdot \Delta T_p \cdot \Delta F_p$, the amount of STF resource allocated to pilot symbols is given by $N_{\text{Tx}}/(\Delta D_p \cdot \Delta T_p \cdot \Delta F_p)$. The space-time-frequency pilot separations are upper bounded by the sampling theorem stated in (3.38). Hence, a minimum amount of STF resource must be allocated to the pilot symbols in order to get a sensible channel estimate.

3.3 Estimation Noise Reduction by Low Rank Approximation via Eigenvalue Decomposition

Recall that the STF channel estimator involves a matrix inversion. In (3.29), when the additive noise variance is small and the pilot symbols are highly correlated, such as in an oversampled pilot placement, the matrix to be inverted might be rank deficient. If we blindly invert a rank deficient matrix when calculating the STF channel estimator, we are

introducing unnecessary numerical errors due to small eigenvalues. To illustrate the numerical instability, we first perform an eigen-value decomposition (EVD) on the $\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H$ term of (3.29),

$$\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^H \tag{3.39}$$

where $\boldsymbol{\Sigma}$ is a diagonal matrix with its diagonal represented by a descending ordered sequence of eigenvalues, and columns of $\mathbf{U}$ are the corresponding eigenvectors. Due to the Hermitian structure of $\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H$, all the eigenvalues are non-negative.

When the pilot symbols are highly correlated, $\mathbf{R}_{\mathbf{h}_p}$ becomes rank deficient. Therefore the eigenvalues of $\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H$ would look like

$$\lambda_1 > \lambda_2 > \ldots > \lambda_r > \lambda_{r+1} = \ldots = \lambda_{N_p} = 0\,. \tag{3.40}$$

The eigenvectors of $\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H + \sigma_{\mathrm{v}}^2\mathbf{I}$ are the same as those of $\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H$, and the eigenvalues of $\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H + \sigma_{\mathrm{v}}^2\mathbf{I}$ is

$$\lambda_i' = \lambda_i + \sigma_{\mathrm{v}}^2 \qquad \text{for } i = 1, \ldots, N_p\,. \tag{3.41}$$

When calculating the linear estimator matrix, $\mathbf{W}$, we need to invert the matrix $\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H + \sigma_{\mathrm{v}}^2\mathbf{I}$. The matrix inverse can be written as

$$\begin{aligned}\left(\mathbf{X}_p\mathbf{R}_{\mathbf{h}_p}\mathbf{X}_p^H + \sigma_{\mathrm{v}}^2\mathbf{I}\right)^{-1} &= \mathbf{U}\left(\boldsymbol{\Sigma} + \sigma_{\mathrm{v}}^2\mathbf{I}\right)^{-1}\mathbf{U}^H \\ &= \sum_{i=1}^{N_p} \frac{1}{\lambda_i + \sigma_{\mathrm{v}}^2}\mathbf{u}_i\mathbf{u}_i^H\,,\end{aligned} \tag{3.42}$$

where $\mathbf{u}_i$ denotes the i^{th} column (eigen-vector) of $\mathbf{U}$.

When the noise variance is small, we can see that for $i > r$, $1/(\lambda_i + \sigma_{\mathrm{v}}^2) = 1/\sigma_{\mathrm{v}}^2$ can become very large, thus introducing numerical errors. To avoid these numerical errors, we perform the following as the matrix inverse,

$$\left(\mathbf{X}_p \mathbf{R}_{\mathbf{h}_p} \mathbf{X}_p^H + \sigma_{\mathrm{v}}^2 \mathbf{I}\right)^{-1} \simeq \sum_{i=1}^{r'} \frac{1}{\lambda_i + \sigma_{\mathrm{v}}^2} \mathbf{u}_i \mathbf{u}_i^H . \tag{3.43}$$

When $r' = r$, we have eliminated numerical errors while preserving the optimality of our channel estimates. However, in reality, we do not have the exact value of r. In this case, we need to choose a large enough $r' < r$ to minimize the effect of truncation (approximation) error. In a high SNR setting, it is safer to *under-estimate* the value of r because of the potential large introduction of numerical errors.

3.4 Complexity Reduction by Space-Time-Frequency Partitioning

The STF LMMSE estimator described in (3.29) has high complexity due to its utilization of the complete channel correlation at each space-time-frequency point. Conceptually, the channel is only effectively correlated within a *localized* STF space. This section exploits the localized correlation among the neighbouring pilots to perform STF estimation. This approach is inspired by the development in [22].

For an appropriate STF pilot placement, we expect that the channel is only highly correlated within the region defined by neighboring pilot symbols in the STF domain (i.e. the shaded region of Figure 3.5). The entire STF space is partitioned into smaller STF subspaces corresponding to the region enclosed by the nearest $\widetilde{N}_p$ neighbouring pilots. In effect, the $(N_{\mathrm{tot}} \times N_p)$ linear estimator is broken into a set of $\mathcal{W}$, $(\widetilde{N}_{\mathrm{tot}} \times \widetilde{N}_p)$ linear esti-

mators, each corresponding to an STF subspace. Mathematically, we have

$$\widetilde{\mathbf{W}}_i = \widetilde{\mathbf{R}}_{\mathbf{hh}_{p,i}} \widetilde{\mathbf{X}}_{p,i}^H \left(\widetilde{\mathbf{X}}_{p,i} \widetilde{\mathbf{R}}_{\mathbf{h}_{p,i}} \widetilde{\mathbf{X}}_{p,i}^H + \sigma_{\mathrm{v}}^2 \mathbf{I} \right)^{-1}, \qquad i = 1, \ldots, \mathcal{W} \tag{3.44}$$

where each quantity is partitioned into the corresponding subspace, respectively.

Minimal complexity is achieved when $\widetilde{N}_p = 8L_p$, corresponding to the eight closest space-time-frequency pilot sequences to a particular space-time-frequency coordinate. Recall that each pilot sequence is of length L_p.

All $\mathcal{W}$ linear estimators become the same if the following assumptions hold true.

1. Partitioning is uniform over the STF space.
2. Pilot placement is uniform across all STF partitions.
3. Pilot signal scheme is the same for all STF partitions.
4. Channel response is shift-invariant in the STF domain, also known as wide sense stationary uncorrelated scattering homogeneous (WSSUS-HO) channels in [48].

Item 1, 2, 4 ensures that $\widetilde{\mathbf{R}}_{\mathbf{hh}_{p,i}} = \widetilde{\mathbf{R}}_{\mathbf{hh}_{p,j}}$ and $\widetilde{\mathbf{R}}_{\mathbf{h}_{p,i}} = \widetilde{\mathbf{R}}_{\mathbf{h}_{p,j}}$ for all i's and j's, and item 3 ensures that $\widetilde{\mathbf{X}}_{p,i} = \widetilde{\mathbf{X}}_{p,j}$ for all i's and j's. Substituting these equalities into (3.44) would yield the same linear processor, $\widetilde{\mathbf{W}}_i$, for all partitions.

In a multi-user environment, a particular user might only be assigned with a small group of time-frequency resources. Thus, this user would only have access to the pilots within the allocated resource, which is another practical reason for performing channel estimation on a smaller STF partition.

3.5 STF Channel Estimator Modification for OSTBC Symbol-Wise Decoding

Recall that the use of OSTBC allows a lower-complexity symbol-wise ML decoding, see Section 2.1.1. Such a symbol-wise ML decoder assumes the channel to be quasi-static over the duration of the block length. Since we applied OSTBC over adjacent subcarriers when combined with OFDM, the combined SFBC-OFDM would require the channel response to be quasi-static over a number of adjacent subcarriers equal to the block length.

However, our proposed STF channel estimator allows slight variations between adjacent subcarriers. In order to use our STF channel estimator together with the OSTBC symbol-wise ML decoder, we simply average the channel estimates over L_B (OSTBC block length) adjacent subcarriers to produce the desired channel coefficient per OSTBC block,

$$\hat{h}^{i,j}(n,k') = \frac{1}{L_B}\sum_{k=0}^{L_B-1}\hat{h}^{i,j}(n,k'+k)\,, \tag{3.45}$$

where L_B denotes the OSTBC block length, and $\hat{h}^{i,j}(n,k')$ denotes the estimated channel response for the k'^{th} OSTBC block in the frequency domain at time n for the link between the i^{th} receive antenna and the j^{th} transmit antenna.

Chapter 4

Simulation Results and Discussions

All simulations are written in C with the source code and binaries attached on the accompanying CD. Appendix B describes the program environment such that any printed results can be regenerated using the attached software.

We present our simulation results in two sections. In the first part, we analyze the MSE behavior of our channel estimator in the high SNR regime. The purpose of this section is to discover an irreducible estimation error floor caused by the implicit LMMSE interpolation between pilot symbols. In the second part, we apply our STF channel estimator to an OSTBC system to evaluate the overall error probability performance in the low-to-medium SNR regime. The purpose of this section is to evaluate the practicality of incorporating our STF channel estimator into a MIMO-OFDM system at realistic SNR values.

All simulations are based on the suburban macro environment of the 3GPP spatial channel model [28]. Table 4.1 presents the basic setup for our simulations.

Some system parameters vary for different plots, and therefore we will specify these variable system parameters prior to presenting a particular plot. The variable system parameters

Table 4.1 Default system parameters.

Parameter	Value
Channel Model	3GPP Suburban Macro
Min. number of channel realizations	10000
Min. number of OSTBC block errors	300
Modulation	QPSK
Carrier center frequency, f_c	2 GHz
Wavelength, λ	0.15 m
Number of subcarriers, N_f	256
CP length, L_{cp}	28
Baud rate, $1/T_s$	3.84×10^6 symbols/sec
Symbol duration, T_s	0.26 μs
OFDM duration, T_{OFDM}	74.0 μs
Frame duration, T_{frame}	2 ms
Number of OFDM symbols per frame N_t	27

are tabulated in Table 4.2 with possible values.

Table 4.2 Variable system parameters.

Parameter	Description	Range
d_{Tx}	Tx antenna spacing	$\mathbb{R}^+$
d_{Rx}	Rx antenna spacing	$\mathbb{R}^+$
N_{Tx}	Number of Tx antennas	$\{2, 3, 4\}$
N_{Rx}	Number of Rx antennas	$\{2, 4\}$
$\|\mathbf{v}\|$	Mobile speed	$\{3, 30, 120\}$ km/h
ν	Normalized Doppler	$\{0.04\%, 0.4\%, 1.6\%\}$

In a practical multi-user system, a particular user is only allocated with a small group of time-frequency resource. Together with the spatial freedom that the user chooses to utilize, we refer to the assigned STF resource as the *user* STF subspace. The shaded area in Figure 4.1 is an example of an user STF subspace.

From Figure 4.1, the user STF subspace is defined by two sets of numbers: the *origin* and *span* of the user STF subspace. Table 4.3 lists the parameters defining the user STF subspace.

Likewise, the pilot symbols also span a region in the STF domain. Figure 4.2 illustrates a

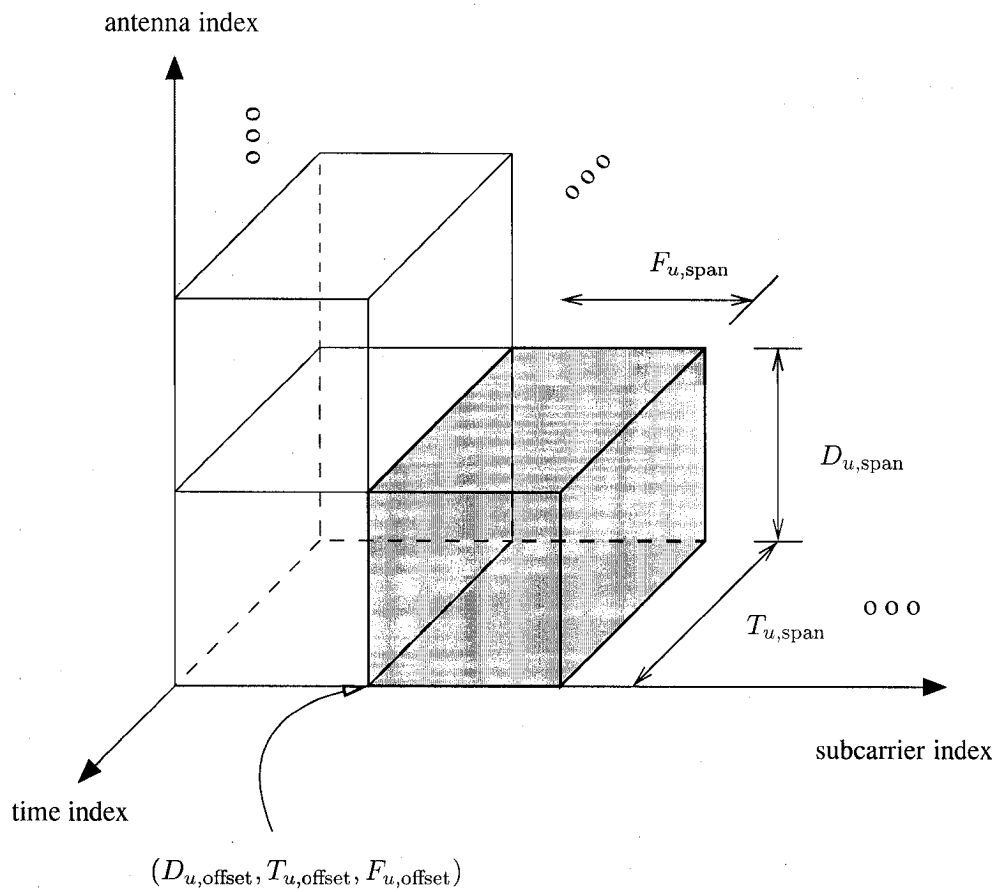


Figure 4.1 User STF subspace.

Table 4.3 User STF subspace parameters.

Parameter	Description	Range
$(D_{u,\text{offset}}, T_{u,\text{offset}}, F_{u,\text{offset}})$	Origin of the user STF subspace	$[(0,0,0),(N_{\text{Tx}}, N_t, N_f)]$
$D_{u,\text{span}}$	User Tx space span	$[0,(N_{\text{Tx}} - D_{u,\text{offset}} - 1)]$
$T_{u,\text{span}}$	User time span	$[0,(N_t - T_{u,\text{offset}} - 1)]$
$F_{u,\text{span}}$	User frequency span	$[0,(N_f - F_{u,\text{offset}} - 1)]$

typical pilot layout on the STF domain. Note that the pilot symbols may span a larger space than the user STF subspace. This is possible since pilots are typically shared in a multi-user environment. A user may choose to utilize pilots *outside* of his/her assigned STF resources to obtain better channel estimates.

From Figure 4.2, the pilot symbols are placed uniformly in the STF domain. In a MIMO system, pilot symbols sent from different Tx antennas interfere with each other at the receiver. In order to separate the pilots from different Tx antennas, Guey et al. has shown that transmission of pilot sequences, of length $L_p \geq N_{\text{Tx}}$, is required [57]. The complete set of parameters regarding the pilot arrangement is specified in Table 4.4.

Table 4.4 Pilot arrangement parameters.

Parameter	Description	Range
L_p	Pilot block (sequence) length	$[N_{\text{Tx}}, N_f]$
$(D_{p,\text{offset}}, T_{p,\text{offset}}, F_{p,\text{offset}})$	Origin of the pilot blocks placement	$[(0,0,0),(N_{\text{Tx}}, N_t, N_f - L_p)]$
$D_{p,\text{span}}$	Pilot Tx space span	$[0,(N_{\text{Tx}} - D_{p,\text{offset}} - 1)]$
$T_{p,\text{span}}$	Pilot time span	$[0,(N_t - T_{p,\text{offset}} - 1)]$
$F_{p,\text{span}}$	Pilot frequency span	$[0,(N_f - F_{p,\text{offset}} - L_p)]$
ΔD_p	Pilot Tx space separation	$[1,(N_{\text{Tx}} - 1)]$
ΔT_p	Pilot time separation	$[1,(N_t - 1)]$
ΔF_p	Pilot frequency separation	$[L_p,(N_f - L_p)]$

In all our simulations, we chose the pilot block length to be $L_p = N_{\text{Tx}}$ to minimize training overhead. Moreover, the SFBC's used here are square matrices, i.e. the frequency spread, L_B, of the STBC's is equal the number of transmit antennas N_{Tx}. This leads us to choose the pilot frequency separation, ΔF_p, to be a multiple of N_{Tx} such that the entire

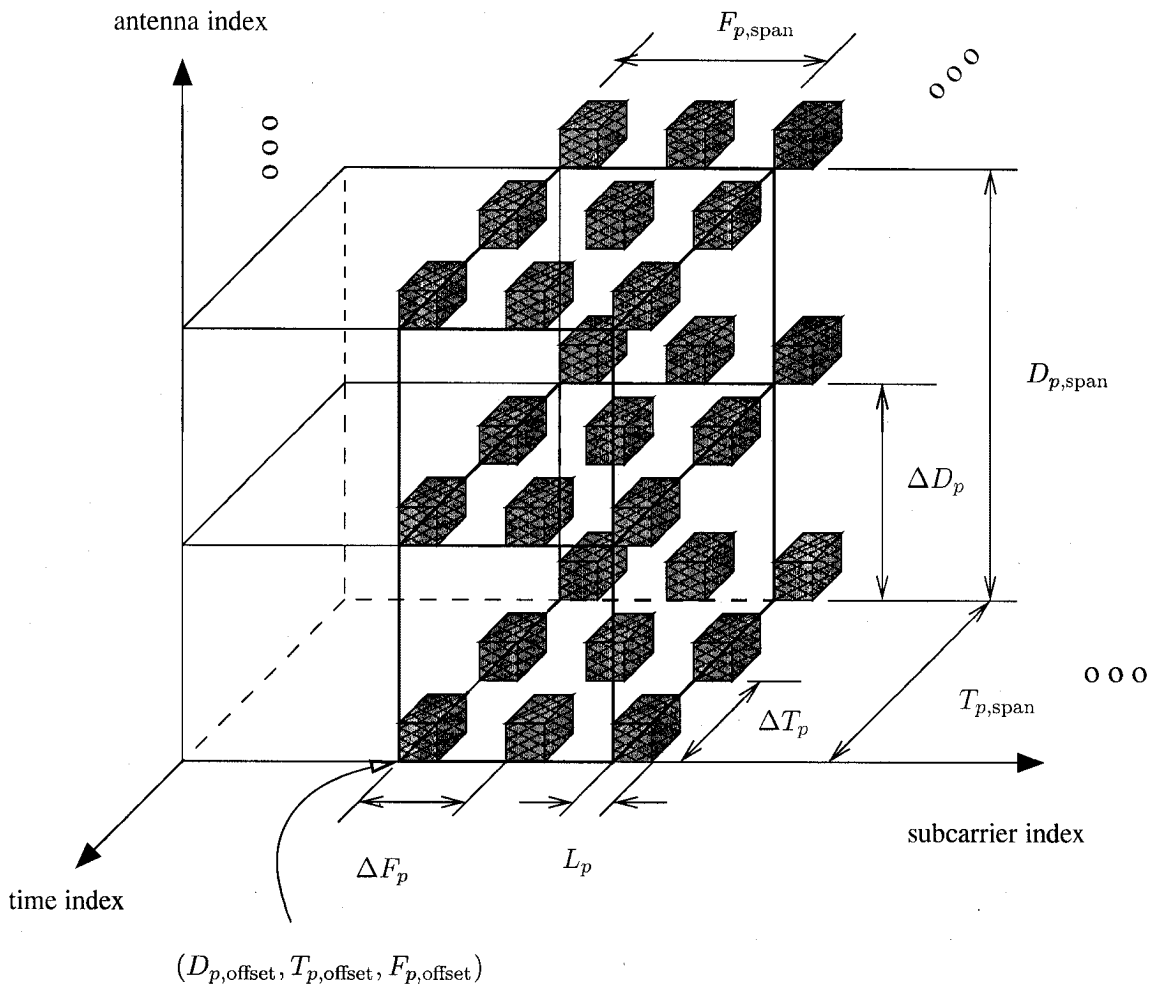


Figure 4.2 Pilot coverage.

bandwidth is filled with either coded or pilots symbols.

4.1 MSE Behavior at High SNR

We discuss the MSE behavior by varying numerous design parameters when noise is negligible.

4.1.1 MSE Distribution over the Estimated STF Subspace

In Section 3.2.2, a proper pilot sequence design based on the DFT was presented. For a system that employs N_{Tx} transmit antennas, the pilot sequence for the j^{th} transmit antennas corresponds to the j^{th} column of the unitary DFT matrix defined in (2.28). We refer to this pilot design as the DFT pilot design.

We consider a system of $N_{\text{Tx}} = 3$ and $N_{\text{Rx}} = 2$ here. The Tx antenna spacing is 0.1λ, and the Rx antenna spacing is 0.5λ. Pilots are sent from Tx antenna 1 and 3 only, which means that the channel response due to Tx antenna 2 is deduced using the spatial correlation between the Tx antennas. Only the closest eight pilot blocks (two in each space, time, and frequency domains) were used to obtain the channel estimate of the STF subspace. The user STF subspace, whose channel response is to be estimated, is chosen to be the same as the pilot subspace spanning the eight closest pilot blocks.

Table 4.5 summarizes the user STF and pilot subspace parameters for Figure 4.3 – 4.5.

A pilot-signal-to-noise ratio (PSNR) of 100 dB is considered here to imitate the noiseless case. This demonstrates the inherent error involved, due to implicit LMMSE interpolation between the pilot symbols, in the STF channel estimation method. The inherent error can be

Table 4.5 Simulation parameters for Section 4.1.1.

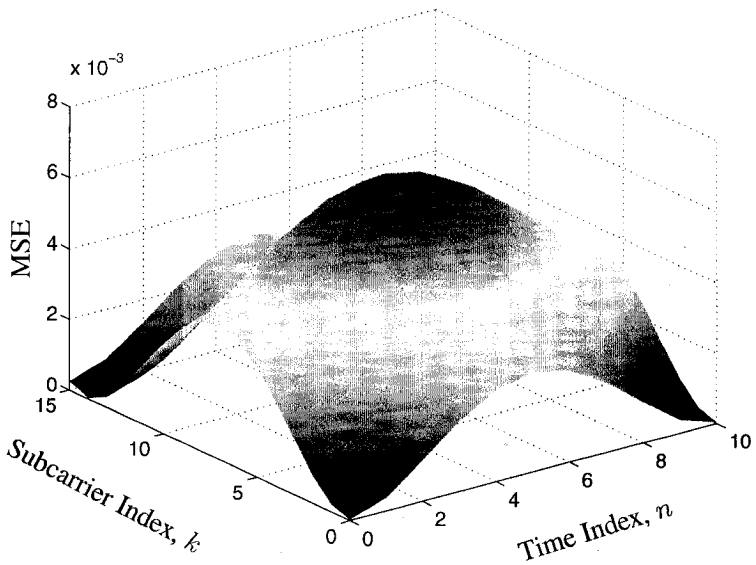
System Parameters	Values	User STF Parameters	Values	Pilot Parameters	Values
d_{Tx}	0.1λ	$D_{u,\text{offset}}$	0	$D_{p,\text{offset}}$	0
d_{Rx}	0.5λ	$T_{u,\text{offset}}$	0	$T_{p,\text{offset}}$	0
N_{Tx}	3	$F_{u,\text{offset}}$	0	$F_{p,\text{offset}}$	0
N_{Rx}	2	$D_{u,\text{span}}$	3	$D_{p,\text{span}}$	3
$\|\mathbf{v}\|$	120 km/h	$T_{u,\text{span}}$	11	$T_{p,\text{span}}$	11
ν	1.6%	$F_{u,\text{span}}$	16	$F_{p,\text{span}}$	16
				ΔD_p	2
				ΔT_p	10
				ΔF_p	14
				L_p	2

reduced as the pilot blocks become increasingly correlated, i.e. when pilot STF separation decreases.

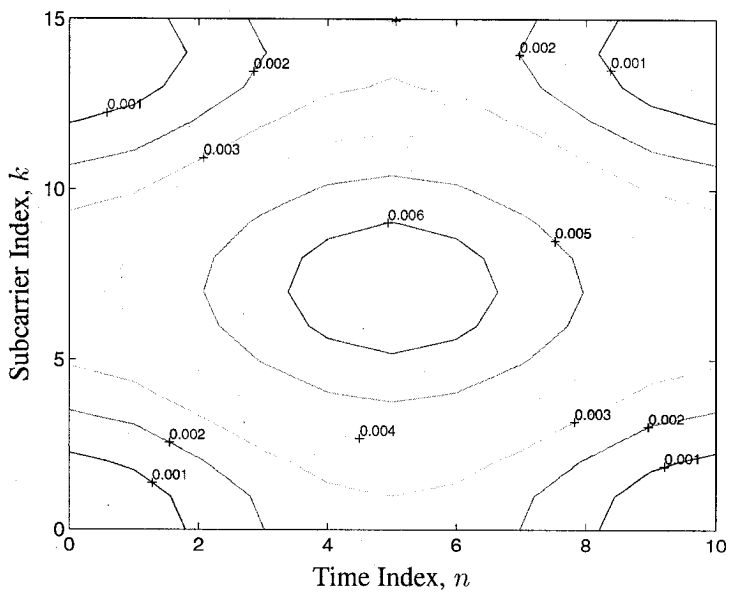
Pilot blocks are located at time/frequency index $(0,0)$, $(0,15)$, $(11,0)$, and $(11,15)$. From Figures 4.3 and 4.5, we see that the MSE is approximately zero at the pilot locations. Our channel estimator relies on the STF correlation between the pilot and data locations to deduce the overall channel response. Thus for locations close to the pilots, the achieved MSE is low. MSE gets progressively worse as the STF coordinate moves away from the pilot locations, and the worst MSE occurs at the STF coordinate that is furthest away from any pilots.

No pilot symbols were sent on Tx 2. As a result, the MSE on Tx 2 (Figures 4.4 and 4.4) is higher compared to the MSE for Tx 1 and Tx 3, even though all the Figures are similar in shape.

In following section, we showed that MSE can be controlled by reducing the pilot spacings, which is equivalent to increasing the auto-correlation of the pilot symbols.

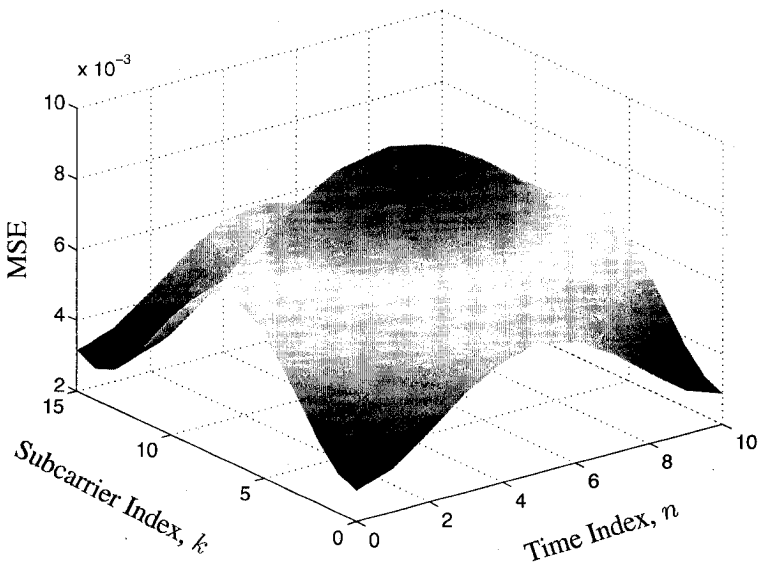


(a) TF distribution

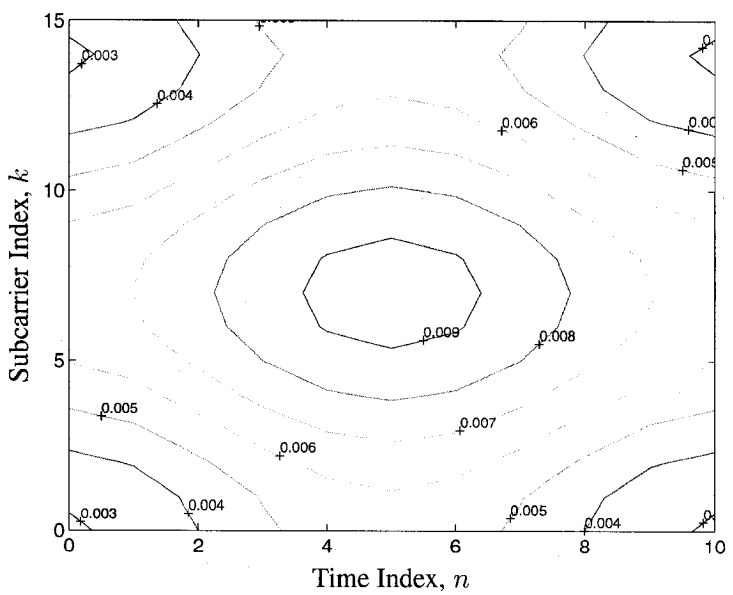


(b) contour plot

Figure 4.3 MSE of the estimated channel response due to Tx antenna 1.

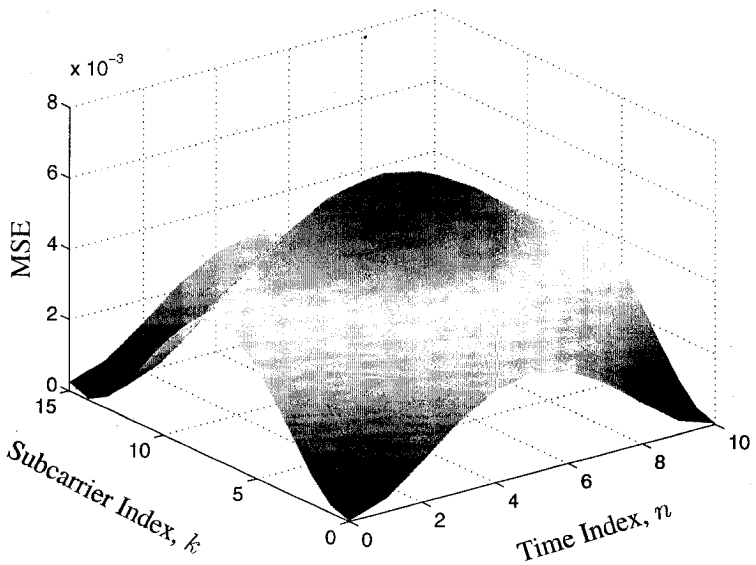


(a) TF distribution

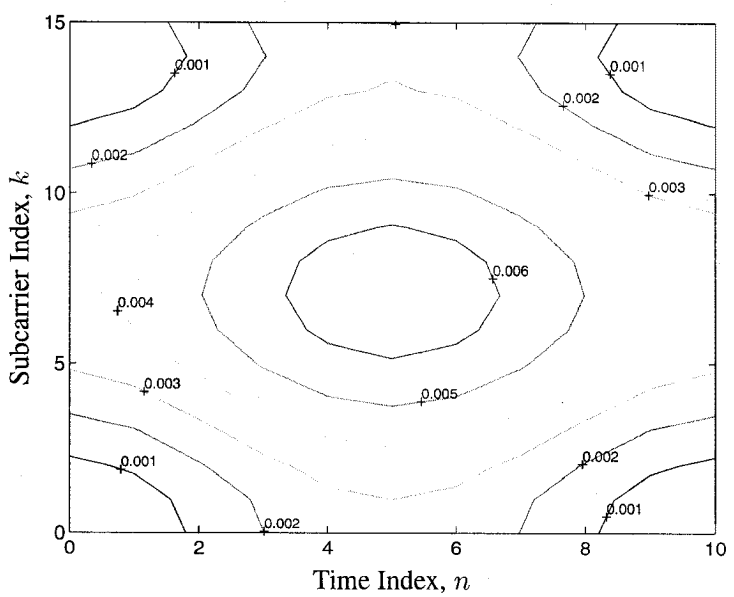


(b) contour plot

Figure 4.4 MSE of the estimated channel response due to Tx antenna 2.



(a) TF distribution



(b) contour plot

Figure 4.5 MSE of the estimated channel response due to Tx antenna 3.

4.1.2 Varying Pilot Symbol STF Separation

This section considers the effects of pilot symbol separation in each of the space, time, and frequency domains on the MSE. Only eight pilot blocks are used for channel estimation and their location defines both the pilot and the user STF subspace. We obtained the MSE behavior for two PSNR values: 30 dB and ∞. This is done to emphasize the effects of noise in the PSNR = 30 dB case versus the noiseless case.

Varying Transmit Antenna Element Separation

Simulation parameters for Figure 4.6 are summarized in Table 4.6.

Table 4.6 Simulation parameters for Figure 4.6.

System Parameters	Values	User STF Parameters	Values	Pilot Parameters	Values
d_{Tx}	varies	$D_{u,\text{offset}}$	0	$D_{p,\text{offset}}$	0
d_{Rx}	0.5λ	$T_{u,\text{offset}}$	0	$T_{p,\text{offset}}$	0
N_{Tx}	3	$F_{u,\text{offset}}$	0	$F_{p,\text{offset}}$	0
N_{Rx}	2	$D_{u,\text{span}}$	3	$D_{p,\text{span}}$	3
$\|\mathbf{v}\|$	120 km/h	$T_{u,\text{span}}$	8	$T_{p,\text{span}}$	8
ν	1.6%	$F_{u,\text{span}}$	8	$F_{p,\text{span}}$	8
				ΔD_p	2
				ΔT_p	7
				ΔF_p	6
				L_p	2

Figure 4.6 illustrates the relationship between the MSE and pilot symbol separation in the Tx spatial domain. Here, pilot blocks are sent from Tx 1 and Tx 3 only. In other words, the channel STF response due to Tx 2 are interpolated from the STF response due to Tx 1 and Tx 3. The interpolated STF response due to Tx 2 would be close to the actual channel response if Tx 2 is highly correlated with Tx 1 and Tx 3.

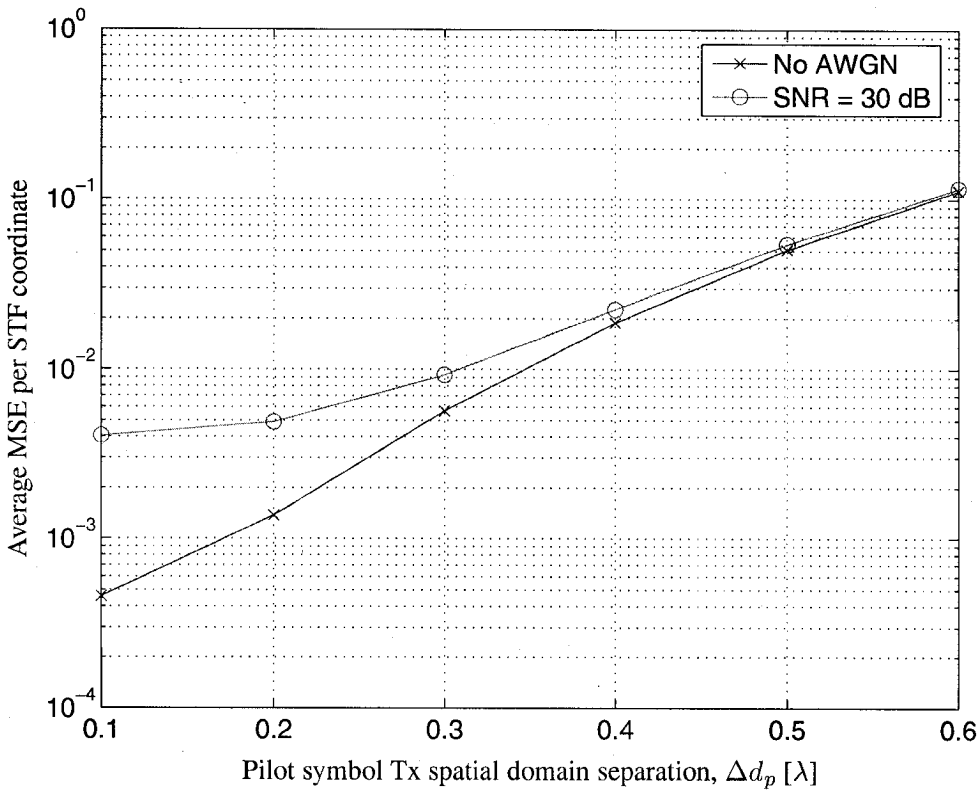


Figure 4.6 MSE vs. pilot symbol Tx spatial domain separation.

Two curves are shown in Figure 4.6: one at PSNR = 30 dB, and the other in a noiseless case. At a high spatial correlation regime (small antenna spacing), the MSE is dominated by AWGN. As the antenna spacing increases, the reduction in spatial correlation causes both MSE to get worse, and to approach each other.

Finally, it should be pointed out that for Tx spatial antenna separation of 0.6λ, we achieved a 10% MSE. A 10% MSE can cause significant BER degradation in an SFBC-OFDM system [58]. However, as most transmit diversity systems use large antenna spacing (i.e. $0.5, 4, 10\lambda$ as in [28]), it seems impractical to place pilot blocks in just a subset of the transmit antennas. This means that for transmit diversity system where Tx antenna spacing is large, we would require training on *all* Tx antennas.

Varying Time Separation

Figure 4.7 illustrates the MSE behavior as pilot time separation varies. Similar to above, as the time separation of the pilot blocks increases, the time correlation between pilot blocks decreases. This decrease in time correlation causes an increase in the MSE. Again, the MSE is dominated by AWGN in the high temporal correlation regime, and MSE is limited as time separation increases (i.e. lower temporal correlation). In other words, for a given operating PSNR, there exists a choice of pilot time separation which would yield a satisfactory MSE yet minimizing the overhead allocated for pilot symbols. For example, at PSNR = 30 dB, a pilot time separation of 14 T_{OFDM} would be an appropriate choice to maintain a 1% MSE.

Varying Frequency Separation

Figure 4.8 shows the MSE behavior as a function of frequency separation of the pilot blocks. We see that the MSE increases as frequency separation increases (i.e. spectral correlation

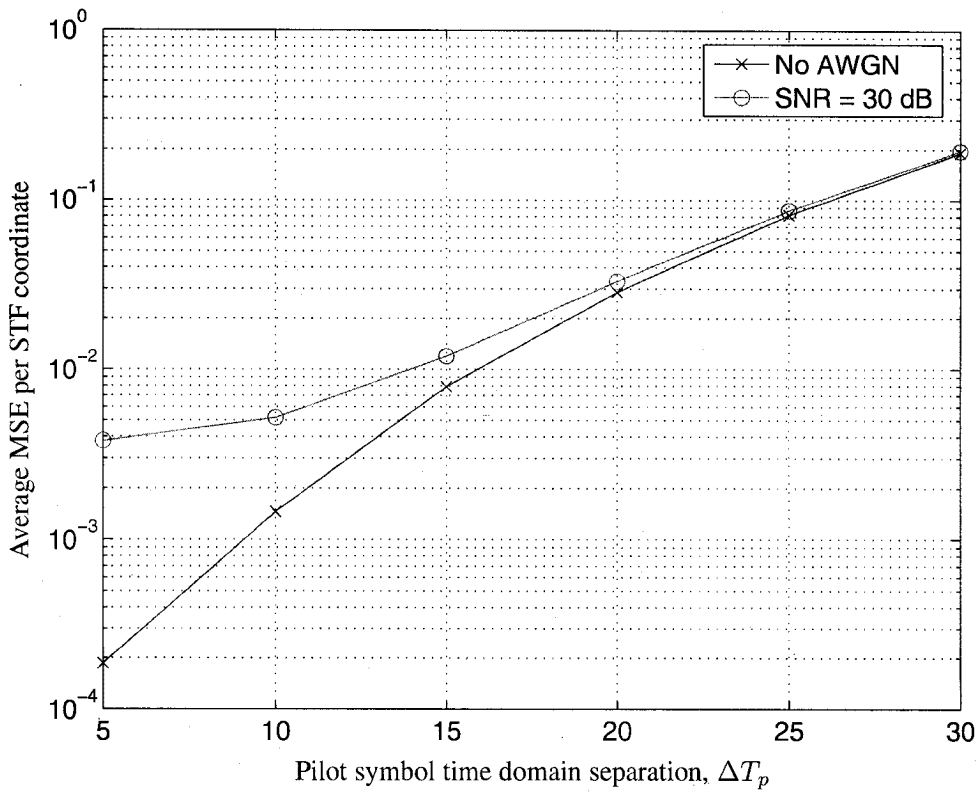


Figure 4.7 MSE vs. pilot symbol time domain separation.

decreases), and AWGN dominates MSE at high spectral correlation. At a particular operating signal-to-noise ratio (SNR), there exists an appropriate choice of pilot frequency separation to maintain a certain MSE while minimizing the pilot symbols overhead.

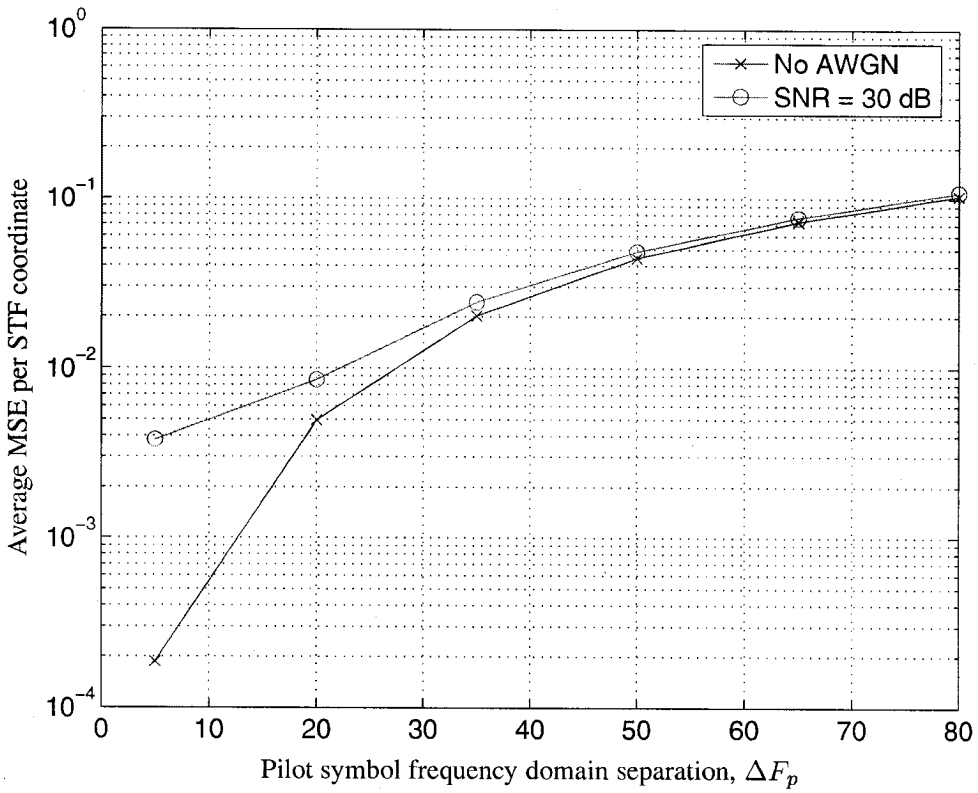


Figure 4.8 MSE vs. pilot symbol frequency domain separation.

Selection of pilot STF separation

A practical communication system is designed to tolerate a certain amount of channel estimation error. An appropriate pilot STF separation is chosen such that a given MSE specification is met. For typical fading scenarios, we can generate the above MSE curves and choose the smallest STF separation that meets the MSE specification while minimizing the

training overhead of the system.

4.1.3 Varying the size of pilot coverage

Section 4.1.2 assumes that the STF subspace coincides with the pilot coverage subspace defined by the eight closest STF pilot blocks. In this section, we look at the MSE behavior as we vary the pilot coverage while keeping the user STF subspace fixed. For a given STF subspace, the MSE can be further reduced by shrinking the pilot STF coverage to be encapsulated by the STF subspace. As a result, the optimal (minimum MSE) pilot pattern is *no longer rectangularly uniform* in the view of pilot placement over the *entire* STF space.

As before, the eight closest pilot blocks, two in each space/time/frequency domain, are used to obtain an estimation of the user STF subspace. Since the number of transmit antennas is usually small, we do not consider varying the pilot coverage in the spatial domain and focus on the MSE effects as we vary the time/frequency pilot coverage.

The user STF subspace is defined in Table 4.7 along with other simulation parameters. The origin of the pilot subspace is shifted according to its varying size such that the center of the pilot subspace coincide with the center of the user STF subspace.

Figure 4.9 illustrates the MSE behavior as we vary the time/frequency pilot separation at two PSNR's. At PSNR = 100 dB, the optimal time/frequency separation was found to be about half of the corresponding time/frequency span of the user STF subspace. At PSNR = 30 dB, the optimal time/frequency separation is 12 T_{OFDM}, and 16 subcarriers, respectively. This corresponds to $3/4$ the time span and $2/3$ the frequency span of the user STF subspace. Although the optimal time/frequency pilot separation varies with different noise power levels, we see that in both cases shown in Figure 4.9, the optimal time/frequency pilot

Table 4.7 Simulation parameters for Figure 4.9.

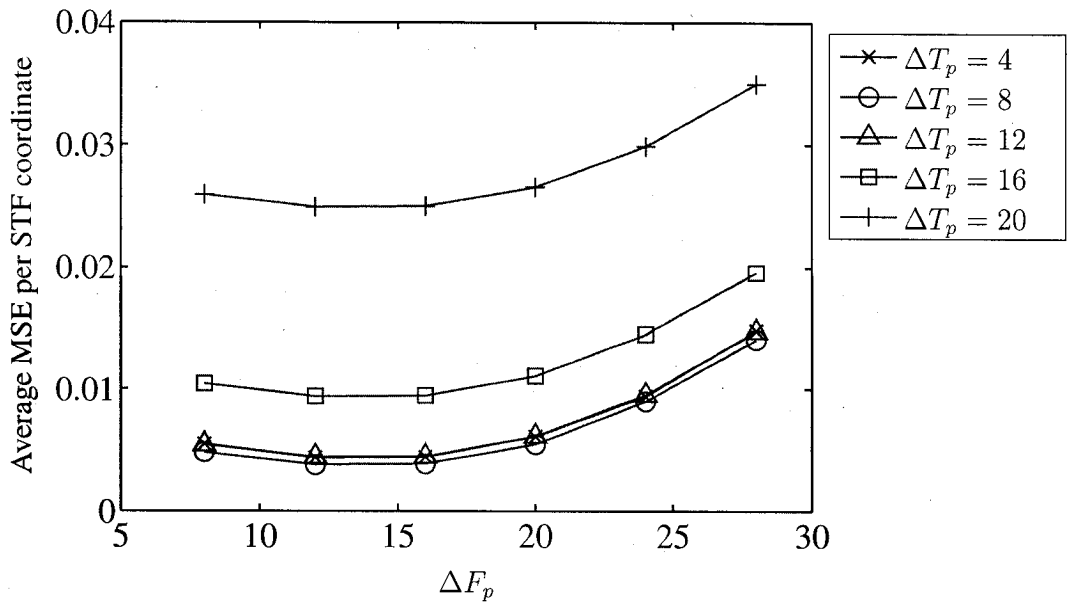
System Parameters	Values	User STF Parameters	Values	Pilot Parameters	Values
d_{Tx}	0.1λ	$D_{u,\text{offset}}$	0	$D_{p,\text{offset}}$	0
d_{Rx}	0.5λ	$T_{u,\text{offset}}$	2	$T_{p,\text{offset}}$	varies
N_{Tx}	3	$F_{u,\text{offset}}$	100	$F_{p,\text{offset}}$	varies
N_{Rx}	2	$D_{u,\text{span}}$	2	$D_{p,\text{span}}$	3
$\|\mathbf{v}\|$	120 km/h	$T_{u,\text{span}}$	16	$T_{p,\text{span}}$	varies
ν	1.6%	$F_{u,\text{span}}$	24	$F_{p,\text{span}}$	varies
				ΔD_p	2
				ΔT_p	varies
				ΔF_p	varies
				L_p	2

placement should fall *within*, rather than on the boundary of, the allocated time/frequency resource for a specific user.

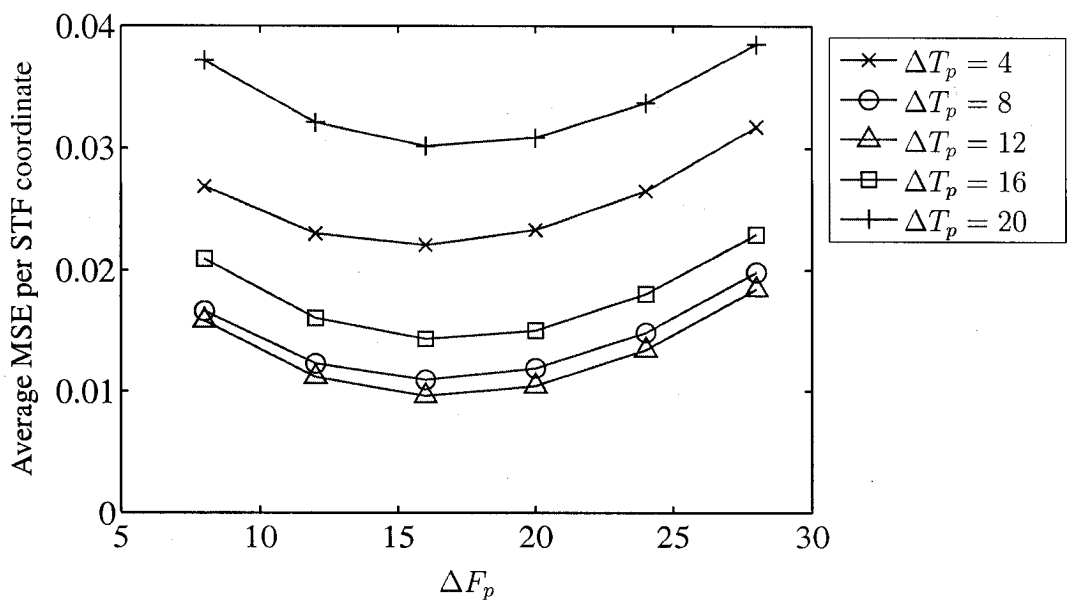
It is interesting to note that smaller MSE is achieved when the pilot STF coverage is smaller than the user STF subspace. Although, when the pilot STF becomes too small (less than half the user STF subspace in each of time and frequency domains), the MSE rises again due to the limited complexity of using only eight pilots blocks. Also note that the MSE increases quickly as the pilot STF coverage grows beyond the user STF subspace. This MSE behavior suggested that if a user is only allowed to use pilots within his/her assigned STF resources, the pilots should be surrounded by data symbols in all dimensions to achieve lower MSE.

4.1.4 Varying the number of neighbouring pilots used for PACE

In the previous sections, STF channel estimation was performed based on the eight closest pilot blocks in the STF space. We investigate the estimation improvements when more than eight pilot blocks are involved in the proposed STF PACE in this section.



(a) SNR = 100 dB



(b) SNR = 30 dB

Figure 4.9 MSE vs. various sizes of pilot coverage with fixed user STF subspace and fixed complexity.

The user STF subspace and the pilot STF separation are fixed, while we vary the number of neighbouring pilot blocks used to estimated this user STF subspace. Here the pilot blocks are uniformly placed in each of the space/time/frequency domains. We do not consider increasing the number of pilots in the Tx spatial domain because this is limited by the physical number of Tx antennas used in the system. Therefore, only variations in the number of time/frequency separated pilot blocks are considered here.

Simulation parameters for this section were tabulated in Table 4.8. Again, the origin of the pilot subspace varies as its span increases.As we increase the number of pilots, we ensure that the center of the pilot subspace coincides with the center of the user STF subspace.

Table 4.8 Simulation parameters for Section 4.1.4.

System Parameters	Values	User STF Parameters	Values	Pilot Parameters	Values
d_{Tx}	0.1λ	$D_{u,\text{offset}}$	0	$D_{p,\text{offset}}$	0
d_{Rx}	0.5λ	$T_{u,\text{offset}}$	10	$T_{p,\text{offset}}$	varies
N_{Tx}	3	$F_{u,\text{offset}}$	100	$F_{p,\text{offset}}$	varies
N_{Rx}	2	$D_{u,\text{span}}$	2	$D_{p,\text{span}}$	3
$\|\|\text{v}\|\|$	120 km/h	$T_{u,\text{span}}$	5	$T_{p,\text{span}}$	varies
ν	1.6%	$F_{u,\text{span}}$	8	$F_{p,\text{span}}$	varies
				ΔD_p	2
				ΔT_p	5
				ΔF_p	8
				L_p	2

Figure 4.10 shows how MSE changes with respect to different numbers of pilot blocks used in the STF PACE. In this figure, t_n and f_n denote the number of time- and frequency-separated pilot blocks used for PACE, respectively.

At PSNR = 100 dB, we observed that smaller MSE can be achieved when the number of time-separated pilots increases from 2 to 3. However, using more than 3 time-separated pilot blocks gives no MSE improvement. In the frequency domain, there is diminishing improvement as we use more frequency-separated pilots to aid channel estimation. Estimation

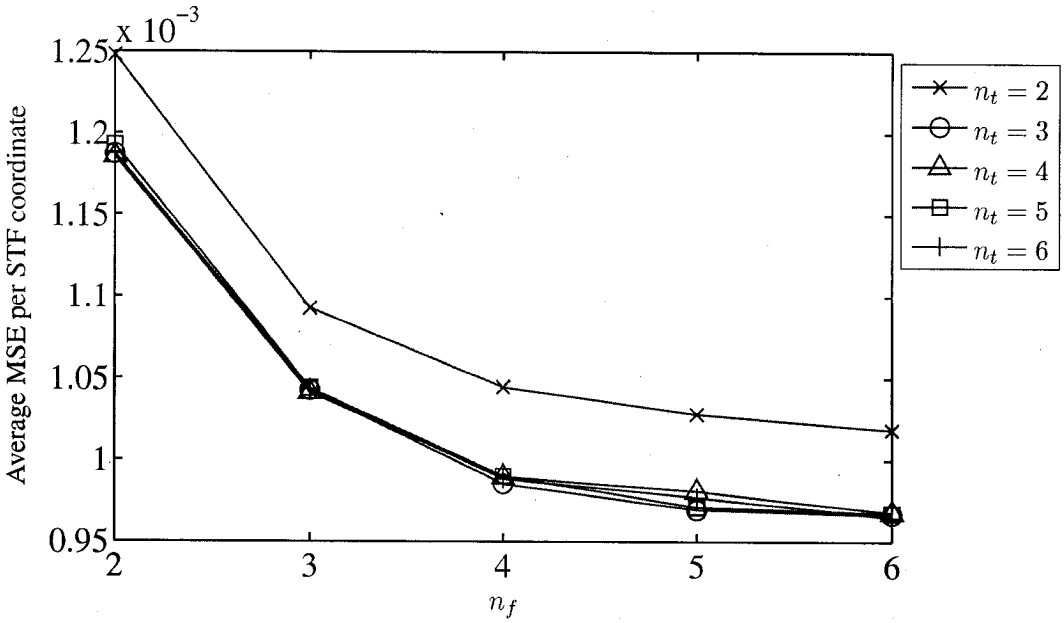
improvement becomes insignificant when more than 5 frequency-separated pilots are used.

At a more practical PSNR of 30 dB, we see that lower MSE can be achieved as we involve more time/frequency-separated pilot blocks in our STF PACE. This gives the possibility of improving the channel estimates if we are willing to increase the complexity. From Figure 4.1.4, satisfactory MSE is obtained for $n_t = 5$ and $n_f = 4$, corresponding to the pilot STF subspace spanning $(n_t - 1)\Delta T_p + 1 = 21$ time slots and $(n_f - 1)\Delta F_p + L_p = 26$ subcarriers.

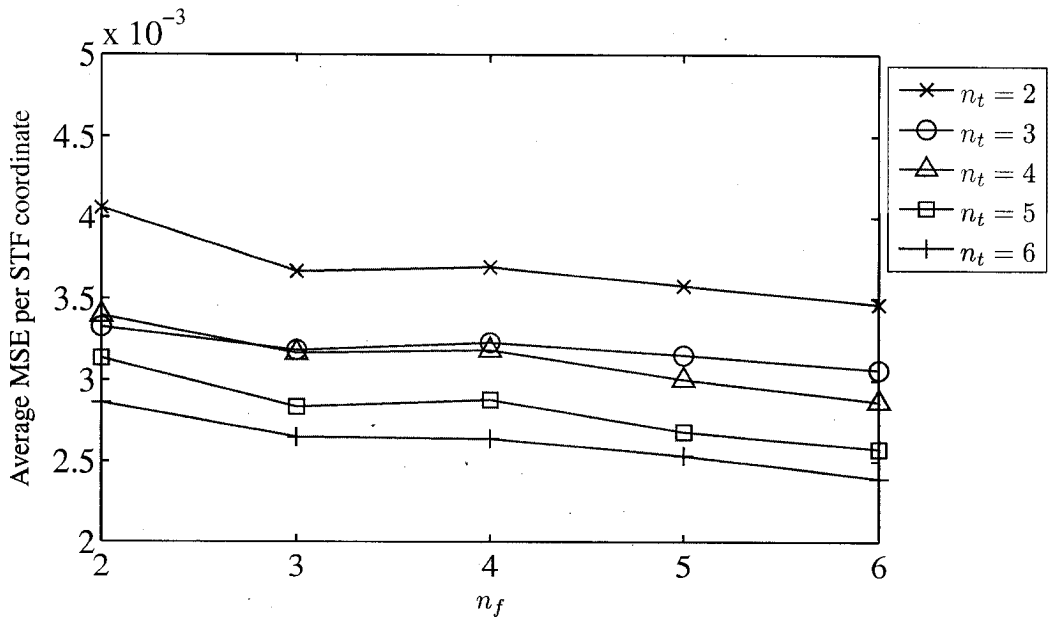
4.2 BER and MSE Performance at Low to Medium SNR

We present the BER performance curves for a MIMO-OFDM system employing four transmit, $N_{\text{Tx}} = 4$, and four receive antennas, $N_{\text{Rx}} = 4$. Tirkkonen's OSTBC code for four transmit antennas [37] is applied over the Tx spatial-frequency domain to provide an SFBC-OFDM system. Our proposed LMMSE STF channel estimator provides the channel estimates for ML symbol-wise linear decoding. The effects of the LMMSE STF channel estimator can be seen by comparing the BER performance between the ML detection with known channel and the ML detection with our LMMSE STF channel estimates. We also compare the BER performance degradation between our LMMSE STF channel estimator against the LMMSE TF channel estimator discussed in [29].

In the following results, the pilot block length is set at $L_p = 4$. The user STF subspace to be estimated starts at origin $(D_{u,\text{offset}}, T_{u,\text{offset}}, F_{u,\text{offset}}) = (0, 0, 0)$ and spans over $D_{u,\text{span}} = 4$ transmit antennas, $T_{u,\text{span}} = 3$ time slots, and $F_{u,\text{span}} = 16$ subcarriers. Also, the user STF subspaces at all N_{Rx} antennas are estimated jointly. The pilot STF subspace starts at the same origin as the user STF subspace, with the pilot subspace spanning over $D_{u,\text{span}} = 4$ transmit antennas, $T_{u,\text{span}} = 27$ time slots, and $F_{u,\text{span}} = 40$ subcarriers.



(a) SNR = 100 dB



(b) SNR = 30 dB

Figure 4.10 MSE vs. number of pilots used for PACE with fixed user STF subspace

The channel model used is the 3GPP suburban macro channel [28]: six paths with a delay spread of 0.17 μs, and a mobile speed of 120 km/h. Such a channel exhibits a moderate amount of time and frequency selectivity. To emphasize the spatial effects in our STF channel estimation, we further decreased the angle spread (AS) at the BS by reducing the mean AS from 5° to 2° and the per-path AS from 2° to 0.5°. Recall in section 3.1.3, our STF channel estimator reduces to the TF channel estimator [29] when spatial correlation is negligible. Therefore, our STF channel estimator performs best in a high Tx spatial correlation case with time and frequency selective fading.

Table 4.9 summarizes the parameters used for the rest of this section.

Table 4.9 Simulation parameters for Section 4.2.

System Parameters	Values	User STF Parameters	Values	Pilot Parameters	Values
d_{Tx}	0.5λ	$D_{u,\text{offset}}$	0	$D_{p,\text{offset}}$	0
d_{Rx}	0.5λ	$T_{u,\text{offset}}$	0	$T_{p,\text{offset}}$	0
N_{Tx}	4	$F_{u,\text{offset}}$	0	$F_{p,\text{offset}}$	0
N_{Rx}	4	$D_{u,\text{span}}$	4	$D_{p,\text{span}}$	4
$\|\|\text{v}\|\|$	varies	$T_{u,\text{span}}$	3	$T_{p,\text{span}}$	27
ν	varies	$F_{u,\text{span}}$	16	$F_{p,\text{span}}$	40
$\overline{\theta}_{\text{AoD}}$	$U(-5°, 5°)$			ΔD_p	varies
Mean AS at BS	2°			ΔT_p	varies
Per-path AS at BS	0.5°			ΔF_p	varies
				L_p	4

Two important pilot parameters should be introduced before we present the simulation results. These are the percentage of training overhead, and the percentage of transmit power allocated to the pilot blocks. By introducing these two parameters, the SNR has to be re-defined to factor in the bandwidth and power reduction due to the insertion of pilot blocks. Appendix A details the definition of SNR, represented by E_b/N_0, and the respective definitions of the percentage training overhead and the percentage of pilot power.

4.2.1 Varying Pilot Overhead

In this section, we fixed the percentage of power allocated to the pilot blocks, β, to be 0.1 (or 10%) while we study the effects of varying the percentage of pilot overhead, α. Since β is fixed in this section, the per-pilot power is varying according to different training overheads. For a fixed β, the per-pilot power increases as we decrease the training overhead and vice versa. Table 4.10 tabulates the power ratio of the pilot symbol to coded symbol for various training overhead values and the corresponding pilot time/frequency arrangments. Note that each pilot block spans over $L_p = 4$ subcarriers, so for a pilot frequency separation of ΔF_p, there will be $(\text{floor}((N_f - L_p)/\Delta F_p) + 1) \cdot L_p$ subcarriers that contains pilot symbols.

Table 4.10 Pilot parameters and power allocation for section 4.2.1.

Training overhead, α	Pilot symbol to data symbol power ratio [dB]	Pilot time separation, ΔT_p	Number of OFDM symbols containing pilots in a frame	Pilot frequency separation, ΔF_p	Number of subcarriers containing pilots in an OFDM symbol
0.05	2.93	7	4	12	88
0.11	−0.46	5	6	8	128
0.22	−3.47	5	6	4	256

From Table 4.10, we see that as we increase the training overhead, the per-pilot power is decreased with the overall pilot power being spread over more pilot symbols.

Figure 4.11 shows the MSE curves for different percentages of training overhead. As we increase the training overhead, the MSE shifts upwards. The increase in MSE is due to the reduction of per-pilot power. Some of the pilot-power is redistributed *outside* the user STF subspace, thereby lowering the effective pilot power within the relevant user STF subspace.

In contrast to slight degradations in MSE, the BER performance improves as we increase the training overhead as seen in Figure 4.12. This shows that a lower MSE does not neces-

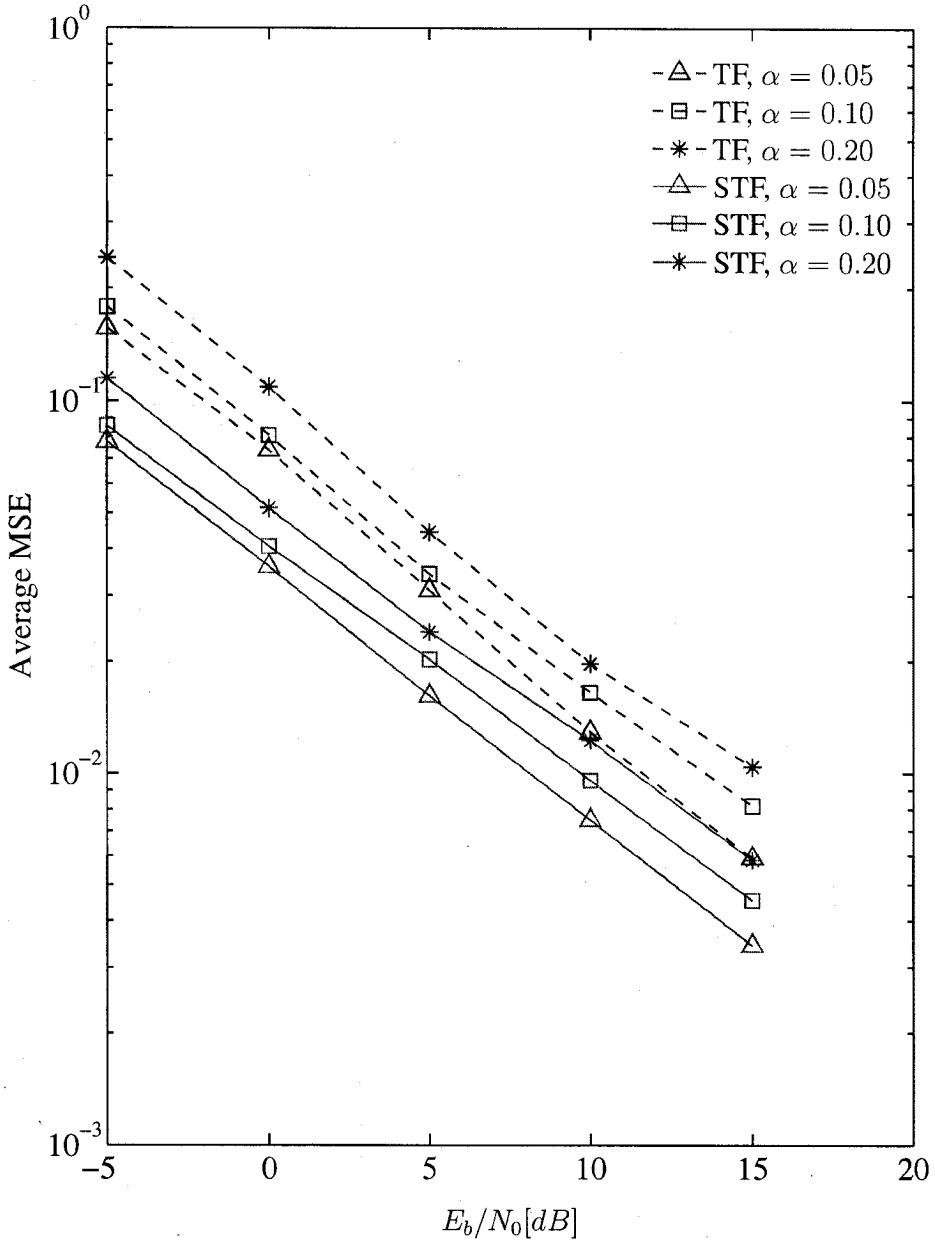


Figure 4.11 $\overline{\text{MSE}}$ performance for different training overhead.

sarily translate to a lower BER performance as the decoder's behavior is dependent on the MSE distribution. We see that the *average* MSE is slightly higher as we increase the training overhead; however, the MSE distribution is more even throughout the STF subspace as we increase the pilot density. In this case, the OSFBC decoder performs better when the MSE is evenly spread, thus giving better BER while the average MSE was slightly increased.

It should be noted that even though BER performance is similar for $\alpha = 0.10$ and $\alpha = 0.20$, the throughput for $\alpha = 0.10$ is 12.5% higher than the $\alpha = 0.20$ case. When $\alpha = 0.10$, 90% of the STF resource is data-occupied; whereas when $\alpha = 0.20$, only 80% of the STF resource is available for data transmission. Thus, throughput achieved with $\alpha = 0.10$ is $90\%/80\% = 1.125$ times the throughput achieved with $\alpha = 0.20$.

4.2.2 Varying Power Allocated to Pilots

In this section, we fixed the percentage of pilot overhead, α, to be 0.1 (or 10%) while we study the effects of varying the percentage of pilot power, β. The per-pilot power is increased by increasing the overall power allocated to the pilot symbols, β. As the pilot symbols become more powerful, we expect our STF channel estimator to give more accurate channel estimates, i.e. lower MSE. This behavior is observed in Figure 4.13.

However, an increase in the percentage of power allocated to the pilots results in a proportional decrease in the percentage of power allocated to the coded symbols. While we expect to get better channel estimates, i.e. lower MSE, as β increases, we incur a lost in power devoted to the coded symbols. Clearly, there is a trade-off between power allocation to the pilot and coded symbols, as powerful pilot symbols lead to weak coded symbols and vice versa.

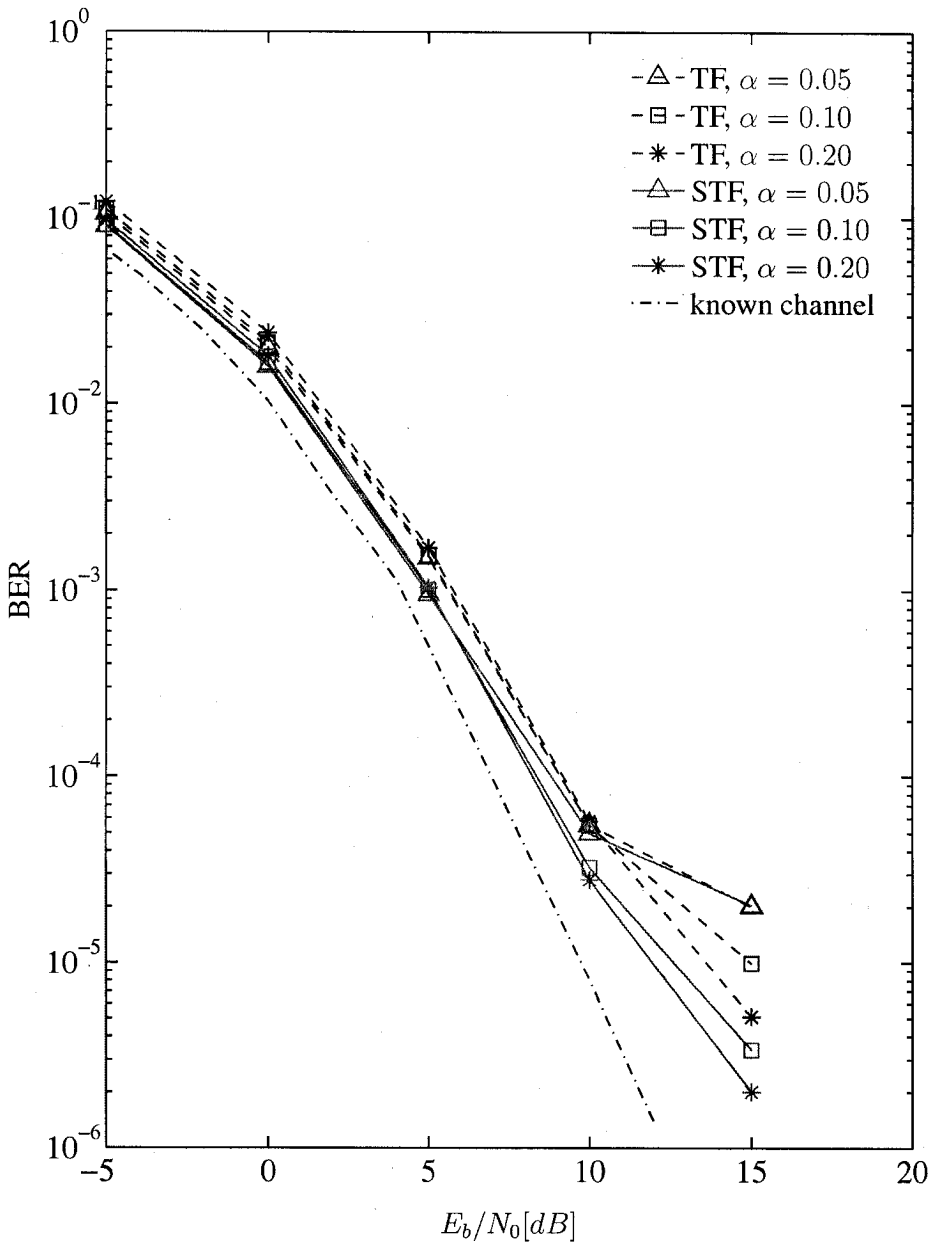


Figure 4.12 BER performance for different training overhead.

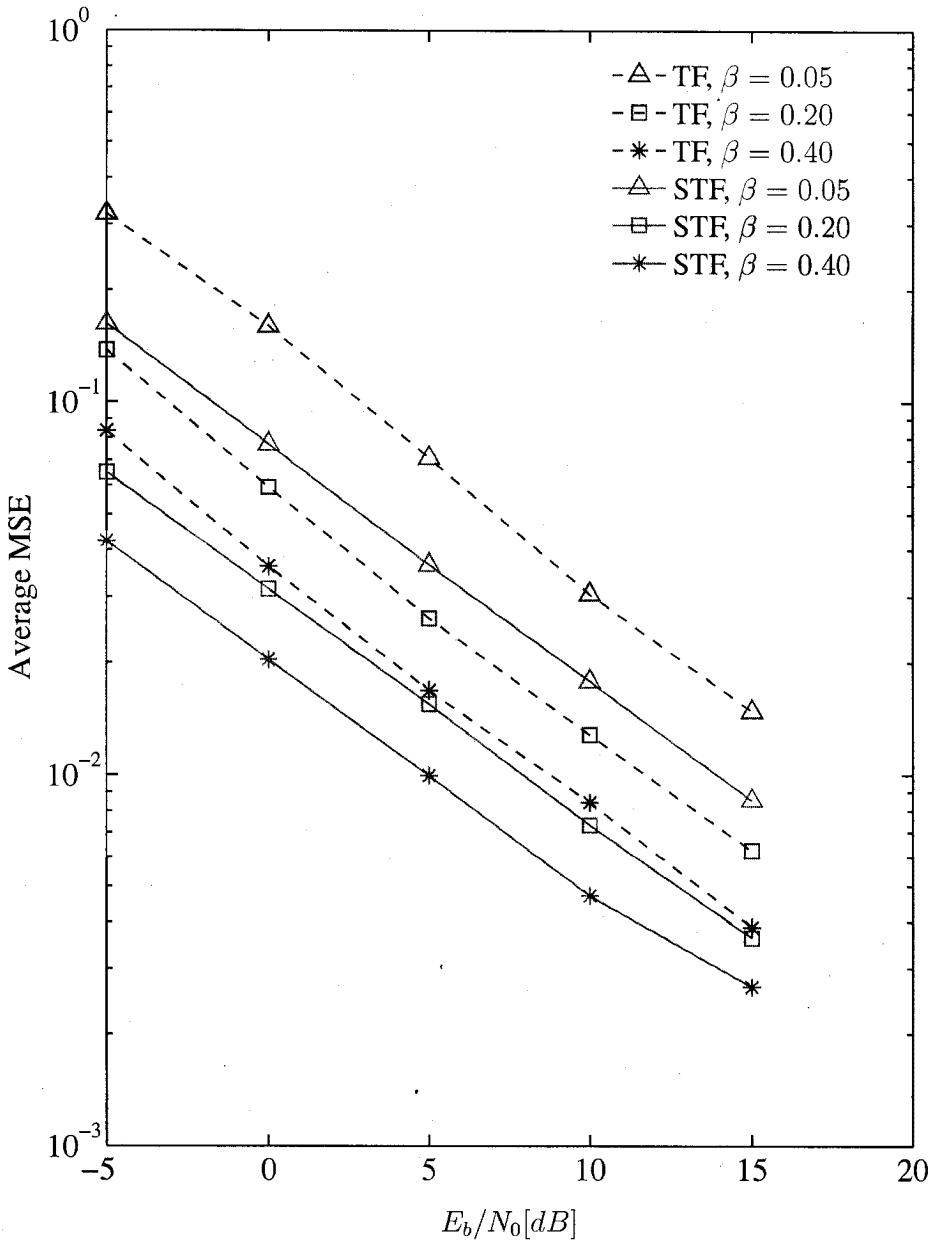


Figure 4.13 $\overline{\text{MSE}}$ performance for different pilot powers.

This trade-off is demonstrated in Figure 4.14 as we observe the higher β curves shift to the right as higher β reduces E_b/N_0 (or SNR). Higher β provides better channel estimates, of which the benefits are only realized at high E_b/N_0 values. For $\beta = 0.05$, the channel estimates are not accurate enough at high SNR which results in a BER floor at 2×10^{-5}. The BER floor lowers as β increases to 0.20 and 0.40. From Figure 4.14, we see that the most suitable choice for β is 0.2 as it gives the best overall BER performance over the range of practical SNR's.

Note that our STF channel estimator *consistently* achieves lower MSE than the TF channel estimator [29], since the TF channel estimator was shown to be a mismatched (in spatial correlation) version of our STF channel estimator in section 3.1.3. Our STF channel estimator also allows lower BER than with the TF channel estimator as shown in Figure 4.14.

4.2.3 Varying Pilot Time/Frequency Separations

Figures 4.15 and 4.16 show the BER and MSE performance for different pilot time/frequency separations for a fast fading channel. The MS speed is 120 km/h, corresponding to 1.6% normalized Doppler frequency, and both the percentage of training overheand and the percentage of pilot power is fixed at 10% (i.e. $\alpha = 0.1$ and $\beta = 0.1$). Figures 4.15 and 4.16 illustrate that proper pilot time/frequency arrangement is necessary to obtain the best BER and MSE performance. A good rule of thumb is to select ΔT_p and ΔF_p with a balanced time and frequency correlation between adjacent time and frequency separated pilots. However, the transmitter often has no knowledge on the Doppler and delay spreads of the wireless channel. This calls for a feedback scheme that would allow adaptive pilot arrangements to minimize BER.

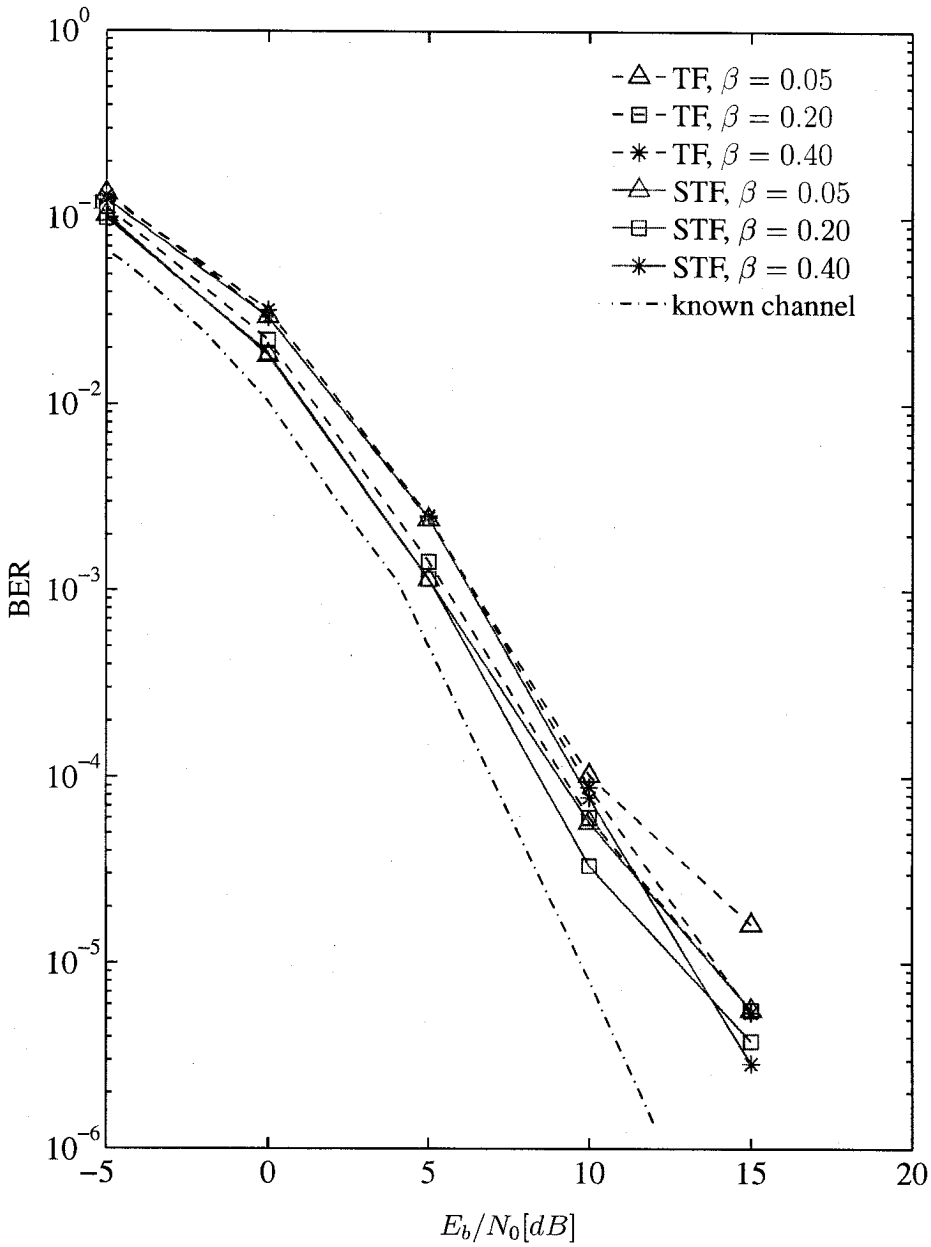


Figure 4.14 BER performance for different pilot powers at 10% training overhead.

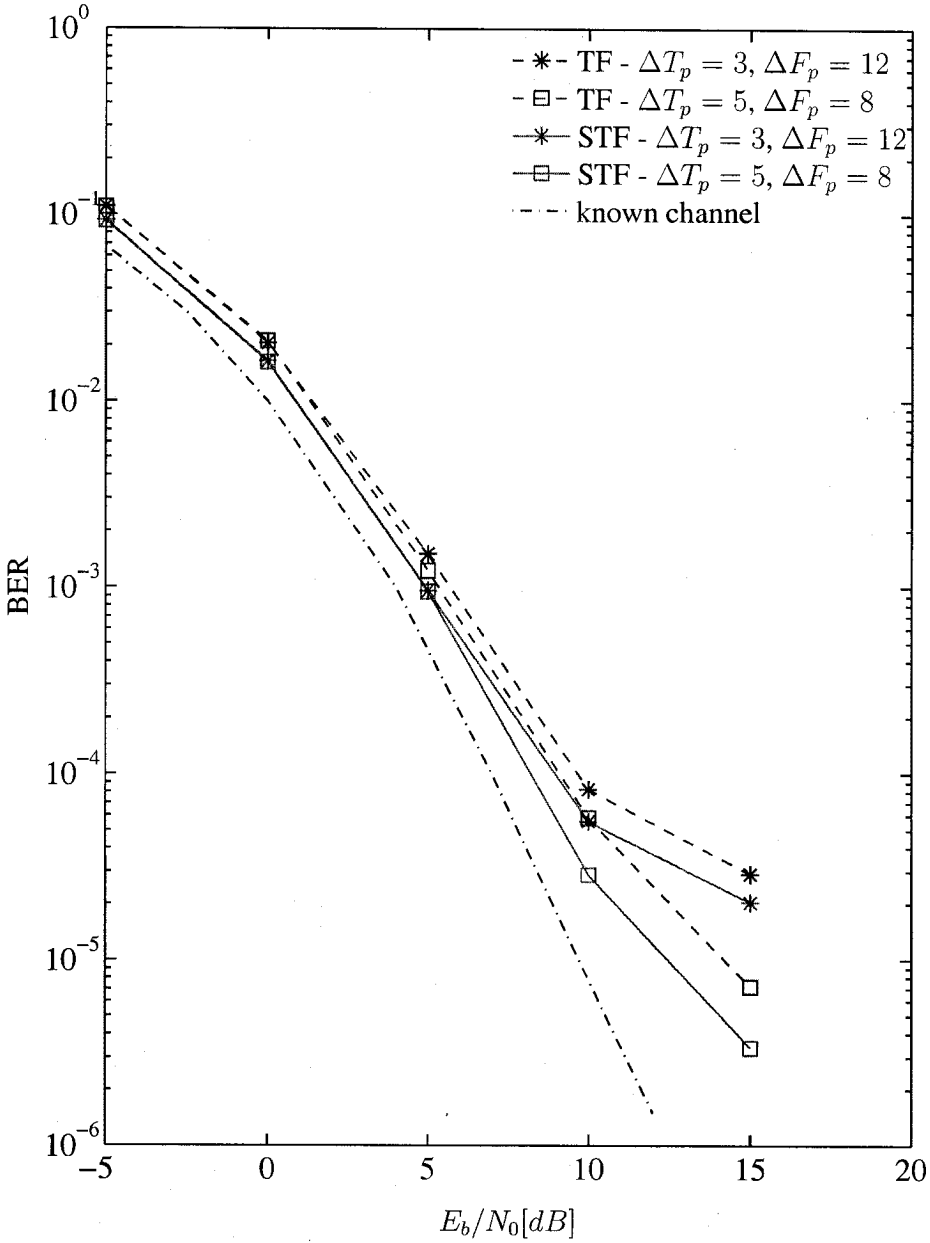


Figure 4.15 BER performance for different pilot arrangements.

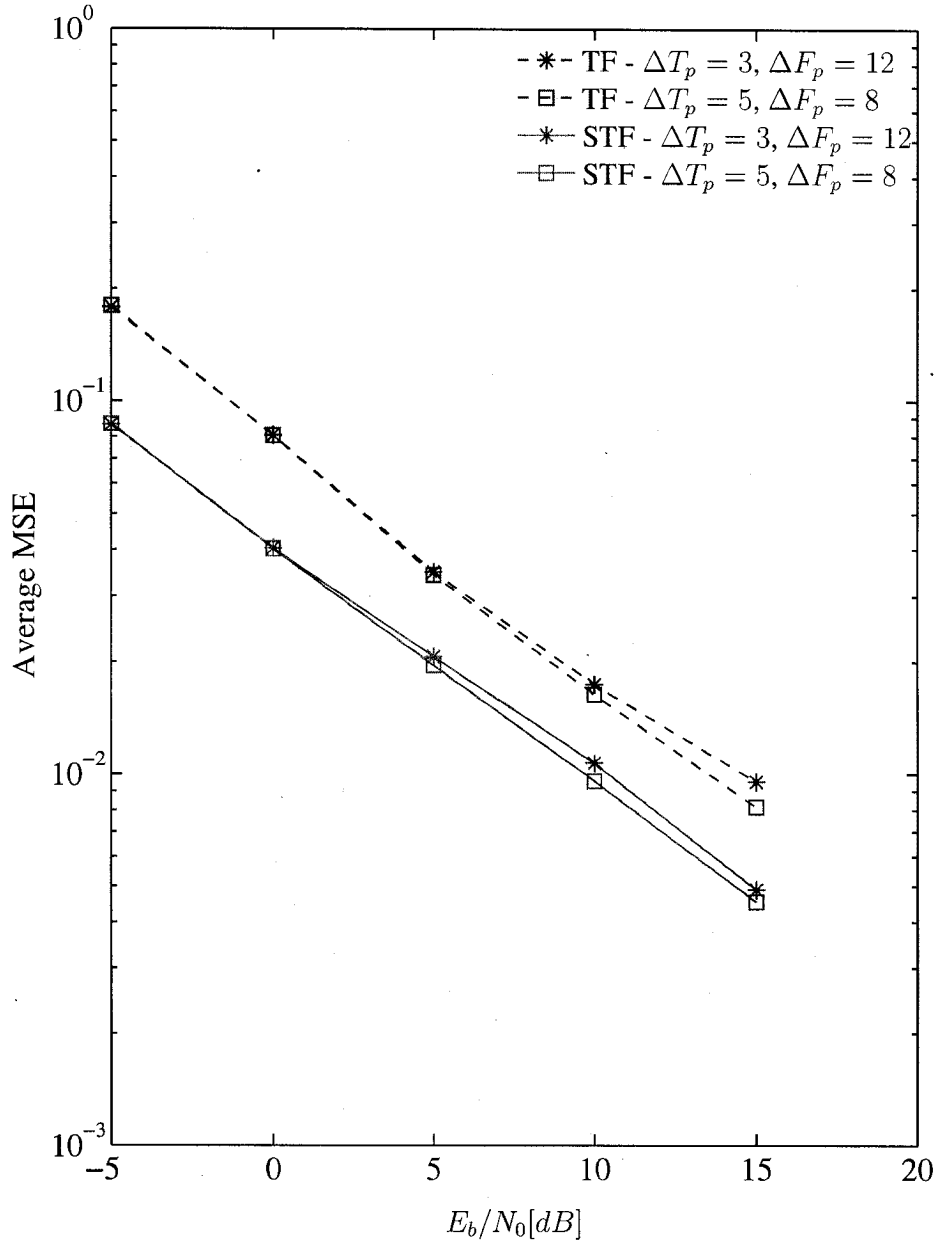


Figure 4.16 $\overline{\text{MSE}}$ performance for different pilot arrangements.

4.2.4 Varying Mobile Speed

Figure 4.17 shows the BER performance for different mobile speeds. From section 4.2.3, we see that the pilot time/frequency separation must be carefully planned to achieve the best BER possible. Table 4.11 summarizes the careful pilot time/frequency plan which yields the best possible BER given the constrains $\alpha = \beta = 0.1$.

Table 4.11 Pilot time/frequency plan for different mobile speeds.

Mobile speed, $\|\mathbf{v}\|$	Normalized Doppler, ν	Pilot time separation, ΔT_p	Pilot frequency separation, ΔF_p
3 km/h	0.04%	9	4
30 km/h	0.4%	9	4
120 km/h	1.6%	5	8

The BER curves with known channel is the same for all mobile speeds because we assumed that the channel is quasi-static over one OFDM symbol, and time-varying with the respective rate from one OFDM symbol to another. From Figure 4.17, the SFBC-OFDM with STF channel estimator can closely approach the SFBC-OFDM with perfect channel knowledge. In fact, a 2 dB performance loss is observed at BER $= 10^{-5}$ for a fast fading environment (120 km/h).

4.2.5 Rank Reduction

Lastly, we look at the employment of the rank reduction technique discussed in section 3.3 for reducing the complexity in calculating the matrix inverse in our STF channel estimator. Figures 4.18 and 4.19 show the BER and MSE impact when we employ the rank reduction technique discussed in section 3.3 in a medium fading environment ($\|\mathbf{v}\| = 30$ km/h, or 1.6% normalized Doppler, ν) at $E_b/N_0 = 10$ dB. The pilot time/frequency placement

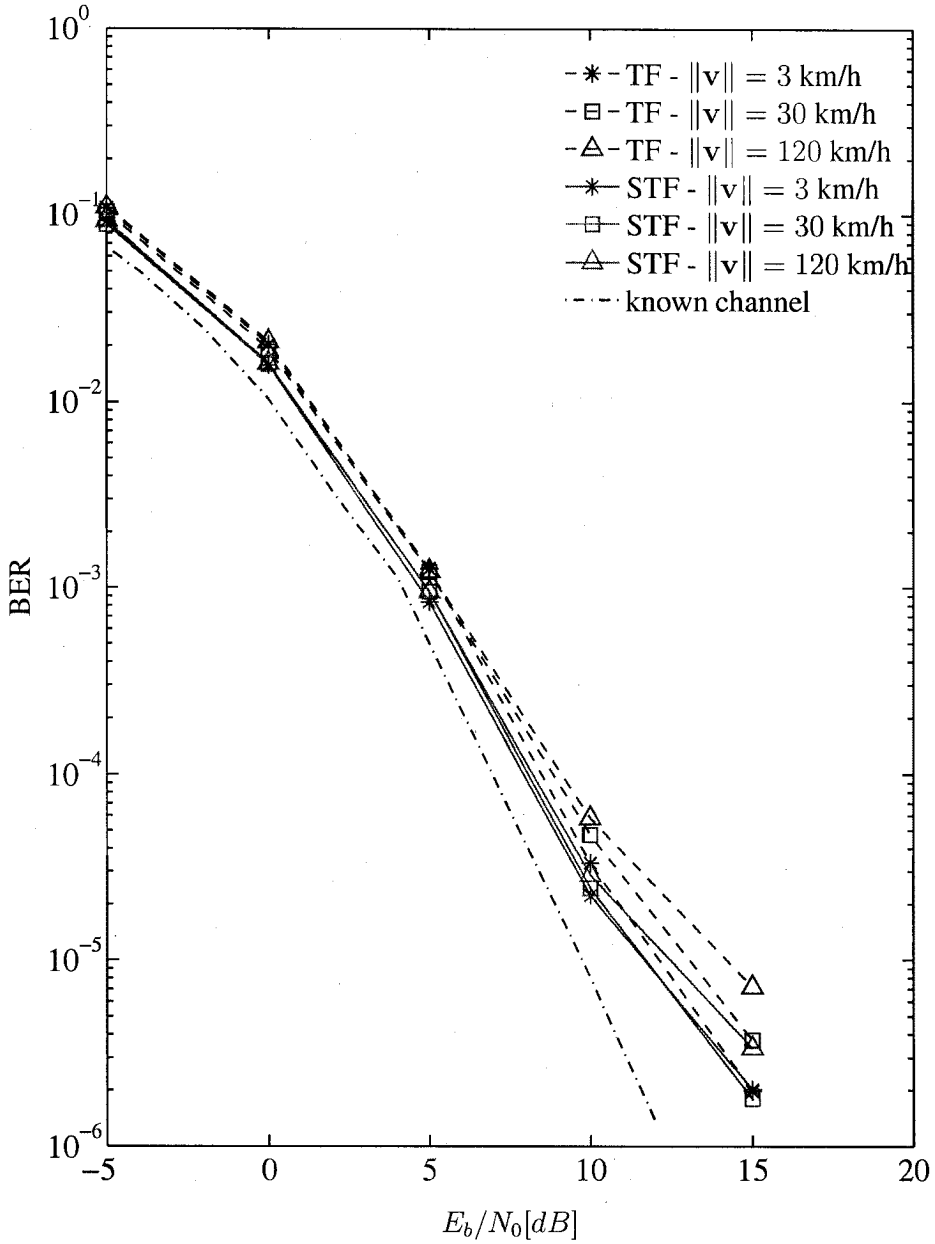


Figure 4.17 BER performance for different mobile speed.

follows the specifications from Table 4.11, and $\alpha = \beta = 0.1$ is assumed.

As shown in Figure 4.19, we get negligible improvement in MSE after rank $r > 16$ for our STF channel estimator. The MSE for the TF channel estimator remains high as we increase the rank because of its inability to exploit the channel's spatial correlation. As a result of improvement in MSE at a reasonable rank, our STF channel estimator is able to provide a much lower BER in the SFBC-OFDM system than the TF channel estimator.

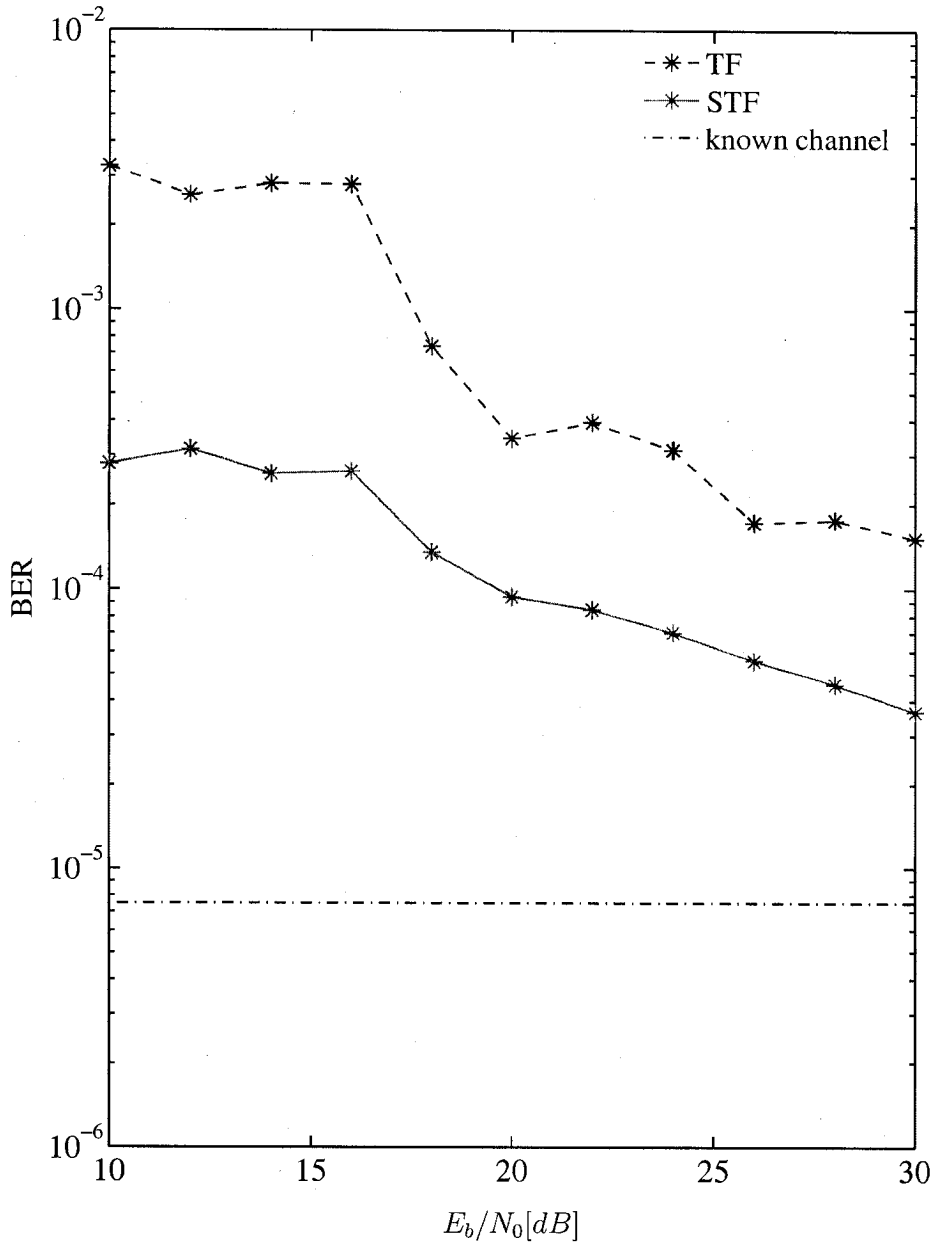


Figure 4.18 BER impact when rank reduction is employed.

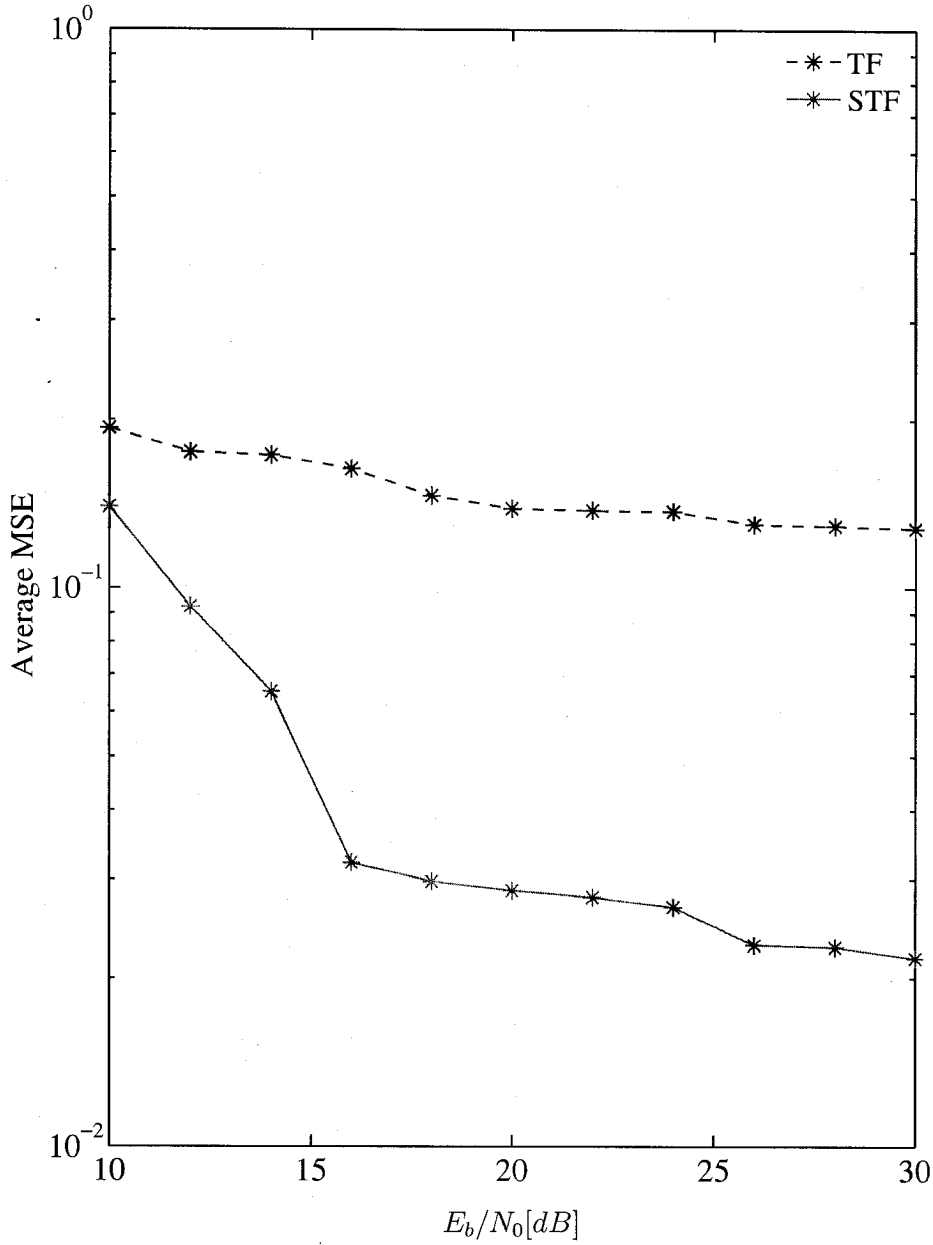


Figure 4.19 $\overline{\text{MSE}}$ impact when rank reduction is employed.

Chapter 5

Conclusions

We have presented an LMMSE STF channel estimator that exploits the space-time-frequency correlations of a MIMO wireless channel. Specifically, our STF channel estimator is suitable for any MIMO-OFDM system that allows periodic pilot symbols insertion in the space-time-frequency domain. For MIMO systems, orthogonal pilot sequences must be used in order to allow spatial interference suppression at the receivers.

Most of the existing MIMO-OFDM channel estimators exploit the correlations in only two out of the possibly four dimensions (transmit spatial, receive spatial, time, and frequency) of the MIMO wireless channel. Our STF channel estimator includes all four dimensions in the estimation process. It does so by observing the correlation among pilots that are carefully separated in all four dimensions, then formulating the LMMSE estimator based on the four-dimensional correlation function. We showed that our STF channel estimator reduces to the the TF channel estimator described in [29] when no spatial correlation exists. Our general STF framework allows various simplifications to lower dimensional estimators when certain correlation properties hold.

Complexity of our STF channel estimator can be reduced by partitioning the entire STF space into smaller STF subspaces, and estimates these subspaces with close neighboring pilot blocks. STF partition is also suitable for a multi-user system, where a single user is only allocated with a subset of the STF resources.

The effects of different pilot strategies were experimented. Careful planning of the pilot placement in the STF domain is required to achieve the lowest MSE. In a practical system, the pilot parameters such as the placement, spacing, coverage per user STF subspace, overhead, and power, must be carefully chosen such that the overall MSE, power allocation, and throughput specifications are met.

We applied our STF channel estimator to an SFBC-OFDM system employing four transmit and four receive antennas. The BER performance degradation, compared to ideal coherent detection, is limited to less than 2 dB at a BER of 10^{-5} in the 3GPP fast fading suburban macro environment. Since the SFBC-OFDM system does not exploit the time and frequency diversity offered by the wireless channel, we anticipate that our STF is even more suitable for a space-time-frequency coded system that exploits all of the space-time-frequency selectivity that exists in the wireless channel.

Appendix A

Signal-to-Noise Ratio Calculations

The performance of any communication system is evaluated at a certain signal-to-noise ratio, E_b/N_0. For a multiple antenna OFDM system, a block of k bits are first buffered at the source. The data bits are then space-time-frequency encoded into codewords of length n_c for each antenna, and the coded symbols are mapped onto the spatial OFDM time-frequency plan. In addition, pilot symbols of length n_p are inserted into the transmission block to allow channel estimation at the receiver. Figure A.1 illustrates the block diagram of a transmission frame, where a system of N_T transmit antennas is considered.

Noted that in Figure A.1, the time-frequency plan for the pilot symbols is the same for every transmit antenna. This constraint is required so that the pilots can be designed jointly to cancel interference from each other. Moreover, the coded symbols and pilot symbols cannot co-exist at the same space-time-frequency coordinate to avoid interference.

From Figure A.1, the energy per bit, E_b, is the total transmitted energy, E_T, divided by

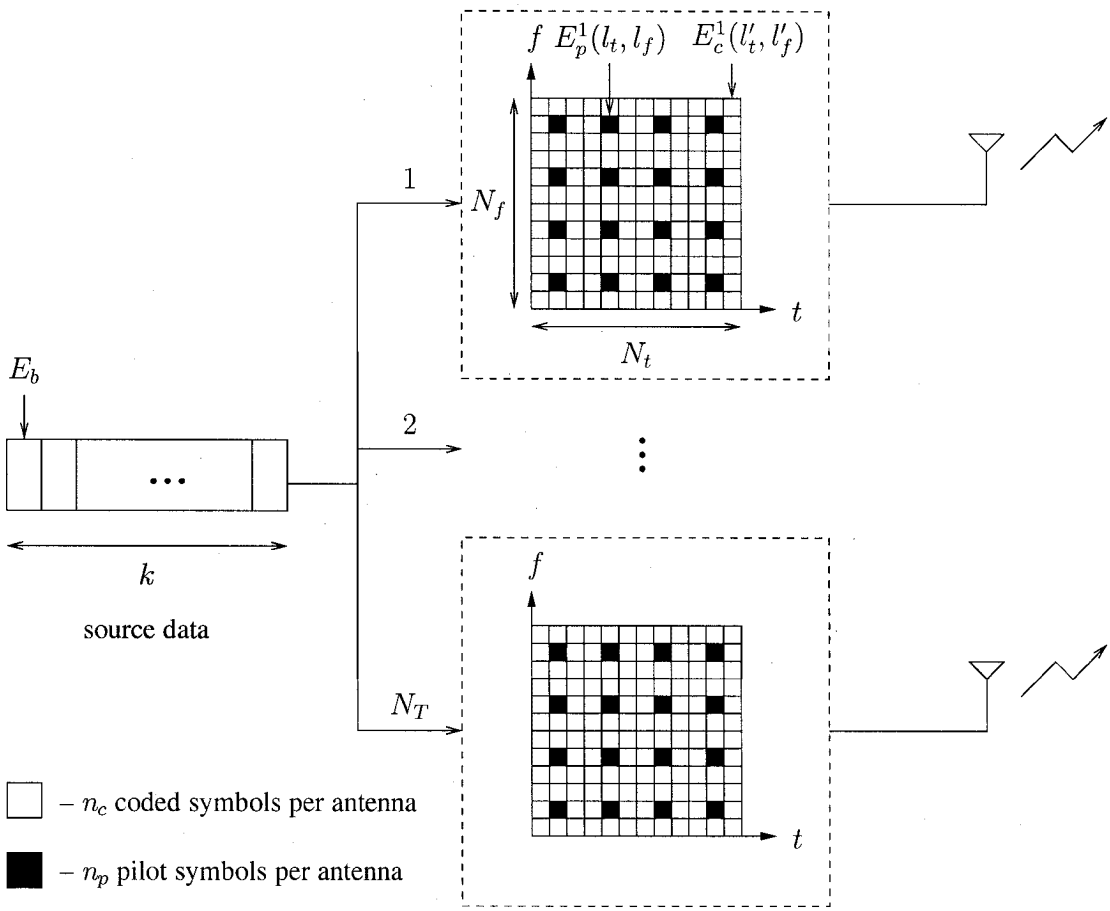


Figure A.1 A transmission frame for multiple transmit antennas systems.

the number of bits transmitted,

$$E_b = \frac{E_T}{k} . \tag{A.1}$$

The total transmitted energy is

$$E_T = \sum_{i=1}^{N_T} \sum_{l_t=1}^{N_t} \sum_{l_f=1}^{N_f} \left[E_c^i(l_t, l_f) + E_p^i(l_t, l_f) \right] , \tag{A.2}$$

where N_T is the number of transmit antennas, N_t is the number of OFDM symbols in a transmission frame, and N_f is the number of subcarriers for each OFDM symbol. The

coded symbol and pilot symbol energies are allowed to vary for different space-time-frequency locations on the transmission frame. Thus, $E_c^i(l_t, l_f)$ and $E_p^i(l_t, l_f)$ correspond to the coded symbol and pilot symbol energies, respectively, at time l_t and subcarrier l_f for the i^{th} transmit antenna.

Before deriving the relationship between E_b and the $E_c^i(l_t, l_f)$'s, we introduce two important parameters for pilot-symbol-assisted systems – namely the percentage of training overhead, and the percentage of total transmit power allocated to the pilot symbols.

Percentage of time/frequency resources allocated for pilot symbols

The percentage of training overhead, α, is given by

$$\alpha = \frac{n_p}{n_c + n_p}. \tag{A.3}$$

Solving for n_p/n_c yields,

$$\frac{n_p}{n_c} = \frac{\alpha}{1-\alpha}. \tag{A.4}$$

Percentage of total transmit power allocated for pilot symbols

The percentage of total transmit power allocated for pilot symbols, β, is

$$\beta = \frac{\sum_{i=1}^{N_T} \sum_{l_t=1}^{N_t} \sum_{l_f=1}^{N_f} E_p^i(l_t, l_f)}{E_T}. \tag{A.5}$$

Conversely, the percentage of total transmit power allocated for the coded symbols is $1-\beta$,

$$1-\beta = \frac{\sum_{i=1}^{N_T}\sum_{l_t=1}^{N_t}\sum_{l_f=1}^{N_f} E_c^i(l_t,l_f)}{E_T}. \tag{A.6}$$

Relationship between E_b and $E_c^i(l_t,l_f)$'s

Putting (A.1) and (A.6) together, we have a relationship between E_b and all the $E_c^i(l_t,l_f)$'s.

$$\begin{aligned} E_b &= \frac{1}{k}\cdot\frac{1}{1-\beta}\sum_{i=1}^{N_T}\sum_{l_t=1}^{N_t}\sum_{l_f=1}^{N_f} E_c^i(l_t,l_f) \\ &= \underbrace{\frac{N_T}{\left(\frac{k}{n_c}\right)}}_{r}\cdot\frac{1}{1-\beta}\cdot\underbrace{\frac{1}{N_T n_c}\sum_{i=1}^{N_T}\sum_{l_t=1}^{N_t}\sum_{l_f=1}^{N_f} E_c^i(l_t,l_f)}_{\overline{E_c}} \\ &= \frac{N_T}{r}\cdot\frac{1}{1-\beta}\cdot\overline{E_c}. \end{aligned} \tag{A.7}$$

The encoding rate, r, is defined as

$$r = \frac{k}{n_c} = \left(\frac{k}{n_s}\right)\left(\frac{n_s}{n_c}\right) = r_s r_c, \tag{A.8}$$

where $r_s = k/n_s$ is the number of bits per symbol in the signal constellation, and $r_c = n_s/n_c \leq 1$ is the coding rate of a given SFBC encoder. n_s is the number of source symbols embedded in a given SFBC codeword, and n_c is the number of coded symbols in a given SFBC codeword. For example, for QPSK signaling with Alamouti SFBC coding scheme, $r_s = 2$ bits/symbol and $r_c = 1$.

The average energy per coded symbol per antenna is defined as

$$\overline{E_c} = \frac{1}{N_T n_c} \sum_{i=1}^{N_T} \sum_{l_t=1}^{N_t} \sum_{l_f=1}^{N_f} E_c^i(l_t, l_f)\,, \tag{A.9}$$

because there are only n_c coded symbols spreading over the time-frequency plan for each antenna.

Example of $\overline{E_c}$ calculation

Normally, the coded symbols are selected from a constant-modulus constellation normalized to unit energy. In such a case, $\overline{E_c} = 1$. However, care must be taken when using an encoding scheme that does *not* yield coded symbols with constant amplitude.

For example, for the rate-3/4 SFBC code for four transmit antennas used in this thesis [37],

$$C = \begin{array}{c} \xrightarrow{\text{Frequency}} \\ \begin{pmatrix} z_1 & z_2 & z_3 & 0 \\ -z_2^* & z_1^* & 0 & -z_3 \\ -z_3^* & 0 & z_1^* & z_2 \\ 0 & z_3^* & -z_2^* & z_1 \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ \downarrow \text{Tx antenna}\,. \end{array} \tag{A.10}$$

each codeword per transmit antenna (columns of (A.10)) occupies four subcarriers but only three of those subcarriers actually contains coded symbols, with nothing transmitted on the remaining subcarrier. In this case, if a unit-energy constant-modulus constellation is used for the coded symbols, $\overline{E_c}$ should be set to 3/4 to account for the zeros included in the SFBC code matrix.

Setting the AWGN variance

To set the proper noise variance for computer simulations, we first divide (A.7) by N_0 to get

$$\frac{E_b}{N_0} = \frac{N_T}{r} \cdot \frac{1}{1-\beta} \cdot \frac{\overline{E_c}}{N_0}. \tag{A.11}$$

We substitute $N_0 = 2\sigma_n^2$ for the N_0 on the right hand side of (A.11) and solve for the AWGN variance, σ_n^2.

$$\sigma_n^2 = \frac{N_T}{r}\left(\frac{1}{1-\beta}\right)\frac{\overline{E_c}}{2\left(\frac{E_b}{N_0}\right)}. \tag{A.12}$$

The noise variance for a particular E_b/N_0 is set according to (A.12) in our simulations.

Appendix B

Simulation Program User Guide

Simulation programs are written in C and uses the GNU Scientific Library [59]. Simulation results seen in this thesis are generated from the attached C programs. Compilation instructions can be found in the README file on the attached CD. Table B.1 lists all the C programs used in this thesis.

Table B.1 Simulation programs list.

File Name	Description
sfbc_lmmse.c	Simulation of SFBC-OFDM system with STF channel estimator.
sfbc_lmmse_2d.c	Simulation of SFBC-OFDM system with TF channel estimator.
corr.c	Simulation of correlation matrices of the 3GPP channel.
modules.c	Communication system modules.
channel3gpp.c	3GPP channel generator.
channel3gppAS.c	Modified 3GPP channel generator (reduced base station angle spread).
gsl_custom.c	Additional numerical routines built upon GSL.
rc_coeff.dat	Raised-cosine filter coefficients.
param.txt	Simulation parameters setup file.
Makefile	Makefile for the mentioned *.c files
run_all	C-shell script for running corr.c and sfbc_lmmse.c sequentially
run_seq	C-shell script for running a series of sfbc_lmmse.c or sfbc_lmmse_2d.c with different parameters files.

Makefile's are included to compile the corresponding *.c files. Type "make" to compile. After compiling corr.c/sfbc_lmmse.c/sfbc_lmmse_2d.c, locate the executable named "run". Type "run param.txt output.m RHp.dat RcohH_Hp.dat" to execute the corresponding simulation.

Since our STF channel estimator requires the space-time-frequency channel correlations, we need to first obtain the STF correlations by simulations before using our channel estimator. Executing the "run_all output.m" C-shell script automates the sequential execution of the two programs: corr.c, then sfbc_lmmse.c or sfbc_lmmse_2d.c. This would generate a Matlab-readable output file called "output.m".

All simulation program loads simulation parameters from the "param.txt" file at the start of execution. Table B.2 lists all the simulation parameters along with their possible input values.

The user STF subspace and pilot coverage subspace is also defined in "param.txt". Please see Tables 4.3 and 4.4 for their descriptions and possible input values.

For simulations that does not require new correlation matrices, user could setup multiple "param.txt" and modify "run_seq" to sequentially executes a series of simulations.

Table B.2 Simulation parameters.

Description	Input Range/Values
Minimum number of frames simulated.	$0 \leq$ minFrames (default=10000)
Minimum number of error events.	$0 \leq$ minStopErr (default=300)
Type of error events.	stopErr_type = {0 - bit, 1 - symbol, 2 - STBC word (default), 3 - frame}
Modulation type.	modulation = PSK
Number of constellation signals.	Nsignals = {2, 4 (default)}
Number of transmit antennas	txNum = {1, 2, 4 (default)}
Number of receive antennas	rxNum = {1, 2, 4 (default)}
Frame length	TTILength = 27
Number of OFDM subcarriers	subCNum = 256
Symbol Rate [symbols/sec]	Ts = 2.6042e-7
Wavelength [m]	lambda = 0.15
Transmit antenna spacing [λ]	txSpacing = continuous (default 0.5)
Receive antenna spacing [λ]	rxSpacing = continuous (default 0.5)
Mobile Speed [m/s]	maxVelocity = {0.8333, 8.3333, 33.3 (default) }
3GPP channel model	ch3gpp = {SUBURBAN_MACRO (default), URBAN_MACRO_8, URBAN_MACRO_15}
Cyclic prefix length	cpLength = 28
Raised-cosine filter rolloff factor	rollOff = {0.0 (default), 0.22, 1.0 }
Pilot sequence design	p_mode = {1 - switch-antenna, 2 - DFT orthogonal}
Matrix inverse tolerance	$0 \leq$ tol (default 0 means use machine precision)
Matrix inverse rank	$1 \leq$ rank $\leq$ Number of pilot symbols
Signal-to-noise ratio [dB]	E_b/N_0 = continuous

Bibliography

[1] I. E. Telatar, "Capacity of multi-antenna gaussian channels," Bell Laboratories, Lucent Technologies," Technical Memorandum, Oct 1995.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, 1998.

[3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[4] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–763, Mar. 1998.

[5] *IEEE Std 802.16™-2004 IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE Std., June 2004. [Online]. Available: http://standards.ieee.org/getieee802/802.16.html

[6] *Physical Channels and Mapping of Transport Channels onto Physical Channels (FDD)*, ETSI 3GPP TR 25.211, Rev. 6.7.0 release 6, Dec. 2005. [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/25_series/25.211/25211-670.zip

[7] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency-division multiplexing using the discrete fourier transform," *IEEE Trans. Commun. Tech.*, vol. COM-19, no. 5, pp. 628–634, Oct. 1971.

[8] L. J. Cimini, Jr., "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. 33, no. 7, pp. 665–675, July 1985.

[9] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY: McGraw-Hill, 2001.

[10] *Series G: Transmission Systems and Media, Digital Systems and Networks - Asymmetric Digital Subscriber Line (ADSL) Transceivers*, ITU-T Std. G.992.1, June 1999. [Online]. Available: http://www.itu.int/rec/T-REC-G.992.1-199907-I/en

[11] *IEEE Std 802.11a™-1999 IEEE Standard for Local and Metropolitan Area Networks Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications — High-speed Physical Layer in the 5 GHz Band*, IEEE Std., June 2003. [Online]. Available: http://standards.ieee.org/getieee802/download/802.11a-1999.pdf

[12] *IEEE Std 802.11g™-2003 IEEE Standard for Local and Metropolitan Area Networks Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications — Amendment 4: Further Higher Data Rate Extension in the 2.4 GHz Band*, IEEE Std., June 2003. [Online]. Available: http://standards.ieee.org/getieee802/download/802.11g-2003.pdf

[13] *Radio Broadcasting Systems; Digital Audio Broadcasting (DAB) to Mobile, Portable and Fixed Receivers*, ETSI DAB EN 300 401, Rev. 1.4.1, May 2001. [Online]. Available: http://webapp.etsi.org/workprogram/Report_WorkItem.asp?WKI_ID=21095

[14] *Digital Video Broadcasting (DVB); Framing Structure, Channel Coding and Modulation for Digital Terrestrial Television*, ETSI DVB-T EN 300 744, Rev. 1.5.1, Nov. 2004. [Online]. Available: http://webapp.etsi.org/workprogram/Report_WorkItem.asp?WKI_ID=19899

[15] K. F. Lee and D. B. Williams, "A space-time transmitter diversity technique for frequency selective fading channels," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, Cambridge, MA, Mar 2000, pp. 149–152.

[16] ——, "A space-frequency transmitter diversity technique for OFDM systems," in *Proc. IEEE GLOBECOM '00*, vol. 3, San Francisco, CA, Nov. 2000, pp. 1473–1477.

[17] D.-B. Lin, P.-H. Chiang, and H.-J. Li, "Performance analysis of two-branch transmit diversity block-coded OFDM systems in time-varying multipath rayleigh-fading channels," *IEEE Trans. Veh. Technol.*, vol. 54, no. 1, pp. 136–148, Jan 2005.

[18] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 686–693, Nov. 1991.

[19] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Börjesson, "On channel estimation in OFDM systems," in *Proc. 45th IEEE VTC*, Chicago, IL, July 1995, pp. 815–819.

[20] O. Edfors, M. Sandell, J.-J. van de Beek, S. K. Wilson, and P. O. Börjesson, "OFDM channel estimation by singular value decomposition," in *Proc. 46th IEEE VTC*, Atlanta, GA, Apr. 1996, pp. 923–927.

[21] Y. G. Li, L. J. Cimini, Jr., and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 902–915, July 1998.

[22] P. Hoecher, S. Kaiser, and P. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by wiener filtering," in *Proc. IEEE ICASSP '97*, vol. 3, Munich, Apr. 1997, pp. 1845–1848.

[23] Y. G. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 461–471, Mar. 1999.

[24] Y. G. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 67–75, Jan. 2002.

[25] G. Auer, "Channel estimation for OFDM systems with multiple transmit antennas by filtering in time and frequency," in *Proc. 58th IEEE VTC*, vol. 2, Oct. 2003, pp. 1204–1208.

[26] H. Miao and M. J. Juntti, "Space-time MMSE channel estimation for MIMO-OFDM system with spatial correlation," in *Proc. 59th IEEE VTC*, vol. 3, May 17-19, 2004, pp. 1806–1810.

[27] H. Zhang, Y. G. Li, A. Reid, and J. Terry, "Channel estimation for MIMO OFDM in correlated fading channels," in *Proc. IEEE ICC 2005*, vol. 4, May 16-20, 2005, pp. 2626–2630.

[28] *Spacial channel model for Multiple Input Multiple Output (MIMO) simulations*, ETSI 3GPP TR 25.996, Rev. 6.1.0 release 6, Sept. 2003. [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/25_series/25.996/25996-610.zip

[29] W. G. Jeon, K. H. Park, and Y. S. Cho, "Two-dimensional MMSE channel estimation for OFDM systems with transmitter diversity," in *Proc. 54th IEEE VTC*, vol. 3, Atlantic City, NJ, Oct. 2001, pp. 1682–1685.

[30] H. L. Van Trees, *Optimum Array Processing. Part IV of Detection, Estimation, and Modulation Theory*. New York, NY: John Wiley & Sons, 2002.

[31] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1804–1824, July 2002.

[32] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: Performance results," *IEEE J. Select. Areas Commun.*, vol. 17, no. 3, pp. 451–460, Mar. 1999.

[33] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge, United Kingdom: Cambridge University Press, 2003.

[34] D. A. Harville, *Matrix Algebra from a Statistician's Perspective*. New York, NY: Springer-Verlag, 1997.

[35] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[36] ——, "Correction to 'space-time block codes from orthogonal designs'," *IEEE Trans. Inform. Theory*, vol. 46, no. 1, p. 314, Jan. 2000.

[37] O. Tirkkonen and A. Hottinen, "Square-matrix embeddable space-time block codes for complex signal constellations," *IEEE Trans. Inform. Theory*, vol. 48, no. 2, pp. 384–395, Feb. 2002.

[38] J. A. C. Bingham, "Multicarrer modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, no. 5, pp. 5–14, May 1990.

[39] J. M. Cioffi. (1991) A multicarrier primer. [Online]. Available: http://www.stanford.edu/group/cioffi/pdf/multicarrier.pdf

[40] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications," *IEEE Signal Processing Mag.*, vol. 17, no. 3, pp. 29–48, May 2000.

[41] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, Inc., 2002.

[42] R. M. Gray. (2006) Toeplitz and circulant matrices: A review. [Online]. Available: http://ee.stanford.edu/~gray/toeplitz.pdf

[43] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Comm. Syst.*, vol. 11, pp. 360–393, Dec. 1963.

[44] A. V. Oppenheim and S. H. Willsky, Alan S. with Nawab, *Signals and Systems*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, Inc., 1997.

[45] H. Stark and J. W. Woods, *Probability and Random Processes with Applications to Signal Processing*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, Inc., 2002.

[46] W. H. Tranter, K. S. Shanmugan, T. S. Rappaport, and K. L. Kosbar, *Principles of Communication Systems Simulation with Wireless Applications*. Upper Saddle River, NJ: Prentice Hall, Inc., 2004.

[47] G. D. Durgin, *Space-Time Wireless Channels*. Upper Saddle River, NJ: Prentice Hall, PTR, 2003.

[48] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, UK: Cambridge University Press, 2003.

[49] T.-A. Chen, M. P. Fitz, S. Li, and M. D. Zoltowski, "Two-dimensional space-time pilot-symbol assisted demodulation for frequency-nonselective rayleigh fading channels," *IEEE Trans. Commun.*, vol. 52, no. 6, pp. 953–963, June 2004.

[50] H. Senol, H. Cirpan, and E. Panayirci, "Pilot-aided Bayesian MMSE channel estimation for OFDM systems: Algorithm and performance analysis," in *Proc. IEEE GLOBECOM '04*, vol. 4, Nov. 2004, pp. 2361–2365.

[51] A. L. Swindlehurst and G. Leus, "Blind and semi-blind equalization for generalized space-time block codes," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2489–2498, Oct. 2002.

[52] H. Wang, Y. Lin, and B. Chen, "Data-efficient blind OFDM channel estimation using receiver diversity," *IEEE Trans. Signal Processing*, vol. 51, no. 10, pp. 2613–2623, Oct. 2003.

[53] *Feasibility Study for Orthogonal Frequency Division Multiplexing (OFDM) for UTRAN enhancement*, ETSI 3GPP TR 25.892, Rev. 6.0.0 release 6, June 2004. [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/25_series/25.892/25892-600.zip

[54] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice Hall, Inc., 1993.

[55] D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, Inc., 1984.

[56] A. Papoulis and S. U. Pillai., *Probability, random variables, and stochastic processes*, 4th ed. Boston, NY: McGraw-Hill, 2002.

[57] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, no. 4, pp. 527–537, Apr. 1999.

[58] R. Narasimha, "Performance of diversity schemes for ofdm systems with frequency offset, phase noise, and channel estimation errors," *IEEE Trans. Commun.*, vol. 50, no. 10, pp. 1561–1565, Oct. 2002.

[59] (2004, June) The gnu scientific library (gsl). [Online]. Available: http://ee.stanford.edu/~gray/toeplitz.pdf