Signal-Perturbation-Free Semi-Blind Channel Estimation for MIMO-OFDM Systems

Feng Wan

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy at

Concordia University

Montreal, Quebec, Canada

February 2009

Your file Votre référence
ISBN: 978-0-494-63419-6
Our file Notre référence
ISBN: 978-0-494-63419-6

ABSTRACT

Signal-Perturbation-Free Semi-Blind Channel Estimation for MIMO-OFDM Systems

Feng Wan, Ph.D.

Concordia University, 2009

Multiple-input multiple-output orthogonal frequency division multiplexing (MIMO-OFDM) has been considered as a strong candidate for the beyond 3G (B3G) wireless communication systems, due to its high data-rate wireless transmission performance. It is well known that the advantages promised by MIMO-OFDM systems rely on the precise knowledge of the channel state information (CSI). In real wireless environments, however, the channel condition is unknown. Therefore, channel estimation is of crucial importance in MIMO-OFDM systems.

Semi-blind channel estimation as a combination of the training-based or pilot-assisted method and the pure blind approach is considered to be a feasible solution for practical wireless systems due to its better estimation accuracy as well as spectral efficiency. In this thesis, we address the semi-blind channel estimation issue of MIMO-OFDM systems with an objective to develop very efficient channel estimation approaches.

In the first part of the dissertation, several nulling-based semi-blind approaches are presented for the estimation of time-domain MIMO-OFDM channels. By incorporating a blind constraint that is derived from MIMO linear prediction (LP) into a

training-based least-square method, a semi-blind solution for the time-domain channel estimation is first obtained. It is revealed through a perturbation analysis that the semi-blind solution is not subject to signal perturbation and therefore is superior to pure blind estimation methods. The LP-based semi-blind method is then extended for the channel estimation of MIMO-OFDM systems with pulse-shaping. By exploiting the pulse-shaping filter in the transmitter and the matched filter in the receiver, a very efficient semi-blind approach is developed for the estimation of sampling duration based multipath channels. A frequency-domain correlation matrix estimation algorithm is also presented to facilitate the computation of time-domain second-order statistics required in the LP-based method. The nulling-based semi-blind estimation issue of sparse MIMO-OFDM channels is also addressed. By disclosing and using a relationship between the positions of the most significant taps (MST) of the sparse channel and the lags of nonzero correlation matrices of the received signal, a novel estimation approach consisting of the MST detection and the sparse channel estimation, both in a semi-blind fashion, is developed. An intensive simulation study of all the proposed nulling-based methods with comparison to some existing techniques is conducted, showing a significant superiority of the new methodologies.

The second part of the dissertation is dedicated to the development of two signal-perturbation-free (SPF) semi-blind channel estimation algorithms based on a novel transmit scheme that bears partial information of the second-order statistics of the transmitted signal to receiver. It is proved that the new transmit scheme can completely cancel the signal perturbation error in the noise-free case, thereby improving largely the estimation accuracy of correlation matrix for channel estimation in noisy

conditions. It is also shown that the overhead caused by the transmission of the SPF data is negligible as compared to that of regular pilot signals. By using the proposed transmit scheme, a whitening rotation (WR)-based algorithm is first developed for frequency-domain MIMO-OFDM channel estimation. It is shown through both theoretical analysis and simulation study that the new WR-based algorithm significantly outperforms the conventional WR-based method and the nulling-based semi-blind method. By using MIMO linear prediction, the new WR-based algorithm utilizing the SPF transmit scheme is then extended for time-domain MIMO-OFDM channel estimation. Computer simulations show that the proposed signal-perturbation-free LP-based semi-blind solution performs much better than the LP semi-blind method without using the proposed transmit scheme, the LS method as well as the nulling-based semi-blind method in terms of the MSE of the channel estimate.

To my loving family!

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisors, Dr. Wei-Ping Zhu and Dr. M.N.S. Swamy, for their invaluable guidance and support throughout the span of this research. It was their advices and critiques that allowed me to convey my ideas through this thesis as well as the publications we completed together, and ultimately, led to timely accomplishment of this work.

I wish to thank Dr. M. Omair Ahmad and Dr. Wenfang Xie for their valuable comments and suggestions on my doctoral proposal and seminar, which have greatly helped shape this thesis. I would also like to thank Dr. Wu-Sheng Lu, the External Examiner, at University of Victoria, and Dr. Walaa Hamouda for their kind acceptance to evaluate my thesis and attend the thesis defense. Their comments, insights and questions are very important for further improvement and final presentation of my thesis.

My thanks also go to my friends and colleagues at the Center for Signal Processing and Communications, Dept. of Electrical and Computer Engineering, including Chao Wu, Wei Wu, Haining Zhang, Min Yang, Salman Khan, Gui Xie, Celia Shahnaz, Shaikh Fattah, Pavel Sinha, etc. I cherish every moment we have shared while working together.

I am very grateful to Concordia University and NSERC, Canada for providing me with financial support through my supervisors' research grants. Without such a support, this thesis would not have been possible.

Finally, I would like to express my love and appreciation to my family and thank them for their consistent encouragement and care. In particular, I would like to thank my wife, Peng Xia, who has been so supportive and patient during my doctoral study, and my son, David Wan, who brings a lot of joy to me.

Contents

List of Figures

List of Tables

List of Symbols

- $\mathbf{X}^T$: transpose of $\mathbf{X}$
- $\mathbf{X}^H$: complex conjugate transpose of $\mathbf{X}$
- $\mathbf{X}^\dagger$: pseudo-inverse (Moore-Penrose inverse) of $\mathbf{X}$
- Trace($\mathbf{X}$): trace of $\mathbf{X}$
- diag ($\mathbf{x}$): a diagonal matrix with its main diagonal elements denoted by $\mathbf{x}$
- $\|\mathbf{X}\|_F$: Frobenius norm of $\mathbf{X}$
- $\otimes$: Kronecker product
- $\circ$: Hadamard product
- $\phi(k)$: $\mathrm{e}^{-j2\pi\frac{k}{K}}$
- $*$: linear convolution
- $\circledast$: circular convolution
- $\mathrm{E}(x)$: expectation of x
- $\delta(\cdot)$: delta function

- vec($\mathbf{X}$): a stacking of the columns of $\mathbf{X}$ into a vector, which has the following properties: $\text{vec}\left[\mathbf{ABC}\right] = \left(\mathbf{C}^T \otimes \mathbf{A}\right) \text{vec}\left(\mathbf{B}\right)$, and $\text{vec}\left[\mathbf{AB}\right] = \left(\mathbf{I}^T \otimes \mathbf{A}\right) \text{vec}\left(\mathbf{B}\right)$ $= \left(\mathbf{B}^T \otimes \mathbf{I}\right) \text{vec}\left(\mathbf{A}\right)$.

List of Abbreviations

- AR: Autoregressive
- AWGN: Additive White Gaussian Noise
- B3G: Beyond 3rd Generation
- BER: Bit Error Rate
- CDMA: Code Division Multiple Access
- CP: Cyclic Prefix
- CRB: Cramer-Rao Bound
- CSI: Channel State Information
- DFT: Discrete Fourier Transform
- DOA: Direction of Arrival
- DOD: Direction of Departure
- ESPRIT: Estimation of Signal Parameters via Rotational invariance technique
- FFT: Fast Fourier Transform

- FIR: Finite Impulse Response
- GAIC: Generalized Akaike Information Criterion
- HOS: High Order Statistics
- IDFT: Inverse Discrete Fourier Transform
- IFFT: Inverse Fast Fourier Transform
- *i.i.d.*: Independently Identically Distributed
- ISI: Inter Symbol Interference
- LHS: Left Hand Side
- LMS: Least Mean Square
- LP: Linear Prediction
- LS: Least Square
- MAP: Maximum A Posteriori
- ML: Maximum Likelihood
- MIMO: Multiple-Input Multiple-Output
- MMSE: Minimum Mean Square Error
- MP: Matching Pursuit
- MSE: Mean Square Error

- MSL: Most Significant Lag
- MST: Most Significant Tap
- OFDM: Orthogonal Frequency Division Multiplexing
- OOK: On OFF Keying
- PAPR: Peak to Average Power Ratio
- QPSK: Quadrature Phase Shift Keying
- RHS: Right Hand Side
- SNR: Signal to Noise Ratio
- SPF: Signal-Perturbation-Free
- SVD: Signal Value Decomposition
- TOA: Time of Arrival
- TS: Training Sequence
- ULA: Uniform Linear Array
- V-BLAST: Vertical-Bell Laboratories Layered Space-Time
- WLS: Weighted Least Square
- WR: Whitening-Rotation

Chapter 1

Introduction

1.1 Background

The third-generation (3G) mobile communication technologies, which were proposed in the 1990s and are still under development today, provide users with high-data-rate mobile access. The three major radio interface standards for 3G are wide-band CDMA (WCDMA), time-division CDMA (TD-SCDMA) and CDMA2000. The data rate of 3G systems are from 144kb/s for high-mobility traffic up to 2Mb/s for very-low-mobility traffic in good conditions. Due to some intrinsic limitations with the radio interface of 3G systems, it is very difficult to extend to very high data rate such as 100Mb/s. Driven by wide-band multimedia and integrated services, therefore, future wireless networks, called B3G (beyond 3G) or 4G are anticipated to provide reliable fraudulence of very high data rates ranging from 100Mb/s to 1Gb/s for high to low-mobility applications. This so-called Long-Term Evolution (LTE) is driven by the telecommunication operator's desire to increase their profit-to-investment ratio as well as the ever increasing demand on new wide-band services. In addition to

its much wider bandwidth, B3G networks are expected to be highly intelligent and cognitive. In particular, they should provide high Quality of Service (QoS) to different grades of users and be robust to any interruptions/interferences, fast changing channel conditions and environments, and large fluctuations in traffic load.

Two key enabling technologies have been identified to meet the goals of B3G networks: multi-input multi-output (MIMO) and orthogonal frequency division multiplexing (OFDM) [1–5]. It is now well recognized that the combination of the two technologies, namely, MIMO-OFDM, would improve the spectral efficiency as well as the system performance in wireless fading channels. With multiple transmit and multiple receive antennas, MIMO systems can achieve either a diversity gain to combat signal fading (e.g. space-time block codes) or to obtain a capacity gain called spatial multiplexing (e.g. V-BLAST) [3, 4]. It means that, both high data rate and superior system performance can be achieved, without increasing the total transmission power or bandwidth, by employing MIMO transmissions. Meanwhile, the common frequency-selective problem of wireless channels caused by inter-symbol-interference (ISI) can be solved by OFDM technique without the need for complex equalization [6]. By dividing an ISI channel into a set of sub-carriers, the information symbols can be transmitted in parallel over the channel. Since the data rate per subchannel is only a small fraction of the system's data throughput, the subchanel becomes frequency-flat. Thus, the need for channel equalization can be circumvented. Additionally, OFDM offers easy scalability to different bandwidth requirement. The MIMO-OFDM technology will improve the spectral efficiency and capacity of the wireless networks, and it has therefore been considered as a strong candidate for B3G systems [7].

1.2 MIMO-OFDM Channel Estimation

The advantages promised by MIMO-OFDM systems rely on the precise knowledge of the channel state information (CSI). The detection of the data and some other signal processing tasks in MIMO-OFDM systems, such as data decoding, space-time processing, limited feedback pre-coding, source and relay power allocation etc, require full or partial knowledge of CSI. It has been proved in [3] that when the channel is Rayleigh fading and perfectly known to the receiver, the capacity of MIMO-OFDM systems grows linearly with the less of transmit and receive antennas. Therefore, channel estimation is of crucial importance in MIMO-OFDM systems.

Generally speaking, MIMO-OFDM channels can be estimated in either the frequency or the time domain [8]. In the frequency-domain channel estimation approach, the frequency-domain channel with respect to some pilot subcarriers is first estimated individually and then used to estimate the frequency-domain channel with respect to other desired subcarriers by using a frequency interpolation technique [9, 10]. In the time-domain channel estimation approach, the time-domain channel impulse response is first estimated and then utilized to form a frequency-domain channel via an FFT process. Since the frequency-domain channel estimation approach focuses on the estimation of frequency-flat MIMO channels, while the time-domain channel estimation approach focuses on the frequency-selective MIMO channels, the former is much simpler to implement. However, the frequency-domain channel estimation approach requires much larger number of OFDM symbols to acquire an accurate channel estimate.

Since in the frequency-domain channel estimation approach, the MIMO-OFDM channel with respect to each subcarrier can be described by a frequency-flat fading channel, it is necessary to briefly review some frequency flat MIMO channel estimation techniques. Generally speaking, there are three kinds of channel estimation approaches for frequency flat MIMO channel estimation. First, the training based methods [9–12], such as the least squares (LS), maximum likelihood (ML), maximum a posteriori (MAP) and maximum mean square error (MMSE) algorithms, employ known training signals to render an accurate channel estimation. In contrast to training based methods, blind channel estimation algorithms, such as those proposed in [13–16], which exploit the second-order cyclo-stationary statistics, correlative coding and other properties, have a better spectral efficiency. With the idea of both the training-based and blind algorithms, semi-blind channel estimation techniques can potentially enhance the quality of MIMO channel estimation [17–24]. With a small number of training symbols, problems such as ambiguities and mis-convergence of the blind methods can be solved. On the other hand, the use of the available data in semi-blind techniques can improve the accuracy of channel estimation.

We now give a brief review of another popular approach, namely, the time-domain channel estimation approach. Similar to the frequency-domain channel estimation approach, the time-domain MIMO-OFDM channel estimation approach can be also categorized into three classes, namely, the training-based method, the blind method and the semi-blind one as a combination of the first two methods. First, the training-based methods employ known training signals to render an accurate channel estimation [25–28]. One of the most efficient training-based methods is the LS algorithm, for

which an optimum pilot design scheme has been given in [25,29,30]. When the full or partial information of the channel correlation is known, a better channel estimation performance can be achieved via some minimum mean-square-error (MMSE) methods [26]. By using decision feedback symbols, the Takagi-Sugeno-Kang (TSK) fuzzy approach proposed in [27] can achieve a performance similar to the MMSE methods while with a low complexity. In contrast to training-based methods, blind MIMO-OFDM channel estimation algorithms, such as those proposed in [31–36], often use the second-order stationary statistics, correlative coding, or other properties to attain a better spectral efficiency. With a small number of training symbols, semi-blind channel estimation algorithms have been proposed to estimate the channel ambiguity matrix for space-time coded OFDM systems [37], the uplink of multiuser MIMO-OFDM systems [38] and MIMO-OFDM systems with non-redundant precoding [33], respectively.

It is worth pointing out that, most of the existing blind and semi-blind MIMO-OFDM channel estimation methods are based on the second-order statistics of a long vector whose size is equal to or larger than the number of subcarriers. To estimate the correlation matrix reliably, they need a large number of OFDM symbols, which is not suitable for fast time-varying channels. In addition, since the matrices involved in these algorithms are of huge size, their computational complexity is extremely high. In contrast, linear prediction-based algorithms for the channel estimation of MIMO systems are based on the second-order statistics of a short vector with a size only slightly larger than the channel length [17, 39–46]. Linear prediction has been widely used in blind MIMO channel estimation and equalization. The key

idea in this technique is to represent the received MIMO signal as a finite-order *autoregressive* (AR) series provided that the transmitted signal is uncorrelated in time [39]. Based on the AR series, a linear prediction filter can be obtained and used for second-order deconvolution. By combining the linear prediction with a higher-order statistics (HOS) or the weighted least-squares method, some blind channel estimation algorithms have been derived [39]. However, these algorithms require a large sample-size and are not robust. Alternatively, a semi-blind algorithm that uses the linear prediction as the blind constraint, in combination with some training data, has been proposed [17,42,43], yielding a better estimation performance. On the other hand, the superiority of the semi-blind method has not been theoretically justified. Moreover, the weighting factor employed to trade off the LS and the blind criteria has not been appropriately determined. As such, the resulting channel estimation performance, though better than that of the LS method, could be further improved. It should also be noted that the existing semi-blind method for MIMO channel estimation cannot be applied to MIMO-OFDM systems.

The previously introduced channel estimation algorithms are all based on the assumption that the channel is time-invariant. In practice, due to the relative motion between the transmitter and the receiver, or the motion of the scatters, doppler spread effect exists, thus always introducing a time-varying channel [47]. In many of the earlier works, a quasi-static block-fading channel model is used. The propagation channel is assumed to be constant within each time-slot, but it is changed after some time-slots. Under the quasi-static assumption, the previously discussed channel estimation algorithms can then be used for each time-invariant slot, while Kalman filter

or some adaptive algorithms can be employed to track the change of the channel matrix for different time slots [48–51]. Moreover, some channel model based algorithms have been proposed for the tracking of time-varying MIMO channels [52–58]. One approach is to characterize the channel impulse response by delayed paths. Based on the time of arrival (TOA) estimation, this approach has been shown to be able to improve the estimation of time-varying channels [54, 57, 58]. In addition, based on the beamforming model [59], the estimation performance of time-varying MIMO channels can be improved by estimating the delay and the angular modal spaces via modal analysis and by filtering the LS channel estimates as projection onto the modal spaces [53].

1.3 Motivation and Objectives of the Research

Considering that the channel estimation techniques suggested by the current mobile communication standards are all based on the pilot or training sequence, they are not suitable for future broadband services due to their poor spectral efficiency as well as possible long processing delay in the receiver. On the other hand, the blind channel estimation techniques developed thus far by the research community are not likely to be used in B3G networks because of their inferior reliability and accuracy as well as high computational complexity. Hence, it is imperative to develop efficient and completely new channel estimation techniques for future B3G systems. Semi-blind channel estimation as a combination of the training-based or pilot-assisted method and a pure blind approach is considered to be a feasible solution for practical

wireless systems owing to its better estimation accuracy as well as spectral efficiency. Therefore, the main objective of this research is to develop new semi-blind channel estimation approaches for MIMO-OFDM systems.

The first objective of the thesis is to develop new semi-blind approaches for the time-domain channel estimation of MIMO-OFDM systems. Most of the existing blind and semi-blind methods for MIMO-OFDM channel estimation are based on the second-order statistics of a long vector, whose size is equal to or larger than the number of subcarriers. As mentioned earlier, in general, these techniques are not suitable for fast time-varying channels and moreover they suffer from a high computational complexity. In contrast, a linear prediction-based semi-blind algorithm, that is based on the second-order statistics of a short vector with a size only slightly larger than the channel length, has been found to be much more efficient than the conventional LS methods for the estimation of frequency-selective MIMO channels [17,42]. In this thesis, we will extend the linear prediction-based semi-blind approach to the channel estimation of MIMO-OFDM systems, leading to a new nulling-based semi-blind channel estimation algorithm.

It is well known that the pulse-shaping filter as well as the matched filter are commonly used in digital communication systems. Perhaps for the sake of simplicity, however, many existing channel estimation methods did not take into consideration either the effect of the pulse-shaping filter in the transmitter or the matched filter in the receiver. As such, these methods have actually been developed for the estimation of the composite channel including the pulse-shaping and matched filters. Considering that both filters are known to the receiver and the only unknown part is the discrete-

time channel [60], ignoring their existence would lead to less accurate estimation results. By utilizing the information of both filters, some improved channel estimation algorithms have been obtained for OFDM systems [61,62] and CDMA systems [63,64]. Motivated by this observation, in this thesis, we will propose an improved frequency-domain nulling-based semi-blind channel estimation algorithm as well as an improved LS algorithm for pulse-shaped MIMO-OFDM systems.

In addition, we will investigate new algorithms for sparse channel estimation of MIMO-OFDM systems. By exploiting the sparse structure of the channel, some sparse channel estimation algorithms have been developed for OFDM systems [65–68] and CDMA systems [69, 70]. Most of them utilize a training sequence and follow two steps: (1) detect the position of the most significant taps (MSTs), and (2) obtain an improved channel estimate by exploiting the position of the MSTs. The common problem of the existing sparse channel estimation methods is that a large number of pilots is needed in order to render an accurate MST detection and channel estimation. To increase the spectral efficiency, the available information of users data could be applied to both the MST detection and the channel estimation. Unfortunately, very little work on blind MST detection and blind sparse channel estimation is found in the existing literature. In this thesis, we will develop, for MIMO-OFDM systems, an efficient semi-blind sparse channel estimation approach, which comprises the MST detection and sparse channel estimation both in a semi-blind fashion.

The second objective of the thesis is to perform the analysis of the proposed semi-blind channel estimation approaches. It is known that the solution of linear prediction or subspace-based methods is always perturbed by various sources, such as finite data

length and measurement noise [71–74]. Perturbation theory has been successfully applied to the analysis of subspace-based methods [72, 75–77]. In this thesis, we will employ a first-order perturbation theory to analyze the MIMO linear prediction as well as the resulting blind constraint in order to justify the superiority of our semi-blind solution over the blind algorithms. Moreover, to facilitate the calculation of the weighting factor for the semi-blind estimation solution of the proposed nulling-based semi-blind algorithm, we will derive a closed-form expression for the mean square error (MSE) of its blind estimation.

The third objective of the thesis is to develop novel signal-perturbation-free transmit schemes and the associated signal-perturbation-free channel estimation algorithms. By conducting the perturbation analysis of the subspace-based and linear prediction-based channel estimation methods, it is found that some of the blind and semi-blind existing channel estimation methods are subject to a signal perturbation error, which is caused by the finite data length in the calculation of the correlation matrix of the received signal. Since the signal perturbation error exists even in the noise free case, the channel estimation performance of those blind and semi-blind channel estimation methods, comparing with that of the training-based methods, becomes worse in the moderate to high SNR cases. In this thesis, to improve the performance of those methods in the moderate to high SNR cases, we will propose two novel transmit schemes, which can cancel the signal perturbation error at the receiver in channel estimation. Based on the new transmit schemes, we will develop two signal-perturbation-free algorithms for the estimation of the frequency-domain and the time-domain channels, respectively.

1.4 Organization of the Thesis

The following is a detailed outline of the remaining chapters of the thesis.

Chapter 2: Some preliminaries required for the development of new channel estimation techniques for MIMO-OFDM systems are introduced. First, the modelling of MIMO-OFDM systems including the transceiver, channel model and signal model is studied. Then, some of the existing algorithms for estimating the frequency-domain and the time-domain channels of MIMO-OFDM systems are briefly reviewed.

Chapter 3: A nulling-based semi-blind approach that uses a training-based least-square criterion along with a blind constraint is proposed for MIMO-OFDM channel estimation. The blind constraint is derived from the linear prediction of the received MIMO-OFDM signal and is used with a weighting factor in the semi-blind cost function. An appealing scheme for the determination of the weighting factor is also presented as a part of the proposed approach. A perturbation analysis of the proposed method is conducted to justify the superiority of the new semi-blind solution and to obtain a closed-form expression for the MSE of the blind constraint, further facilitating the calculation of the weighting factor. The proposed method is validated through computer simulation-based experimentations, showing a very high estimation accuracy of the proposed semi-blind solution in terms of the MSE of the channel estimate.

Chapter 4: A frequency-domain semi-blind channel estimation approach is proposed for the pulse-shaped MIMO-OFDM systems. First, the effect on channel estimation of pulse-shaping that is normally used in a practical communication system

is studied. By utilizing the knowledge of pulse-shaping and matched filtering, an improved semi-blind algorithm is developed for estimating sampling duration-based channels and an LS algorithm is designed for estimating upsampling duration-based channels. Moreover, in the nulling-based semi-blind approach, we propose an efficient algorithm for the computation of the time-domain correlation matrix directly from the received frequency-domain signal. A number of computer simulation-based experiments are conducted, with results confirming the effectiveness of the proposed methods.

Chapter 5: A semi-blind approach is proposed for the estimation of sparse MIMO-OFDM channels. An analysis of the second-order statistics of the signal received through a noise-free sparse channel is first conducted, revealing the sparsity and some other properties of the correlation matrices of the received signal. These properties lead to a direct relationship between the positions of the most significant taps (MST) of the sparse channel and the lags of the nonzero correlation matrices, based on which an efficient MST detection algorithm is developed via a pilot-assisted least-squares estimation. Then, by using the acquired MST information and imposing a blind constraint on the channel vector with respect to the MSTs, a semi-blind approach is presented for the estimation of the sparse channel in the least squares sense. Moreover, a signal perturbation analysis of the proposed approach is conducted, showing that the new semi-blind solution is not subject to the signal perturbation error when the sparse channel is a decimated version of a full FIR channel. Computer simulation for the estimation of various sparse channels using the proposed semi-blind approach are also undertaken.

Chapter 6: A signal-perturbation-free whitening-rotation (WR)-based approach is proposed for semi-blind estimation of frequency-domain MIMO-OFDM channels. A perturbation analysis of two subspace-based methods, namely, the whitening-rotation (WR)-based algorithm and the nulling-based algorithm, is first conducted, showing that in the noise-free case, the former is subject to a signal perturbation error, while the latter is devoid of signal perturbation error due to an ideal nulling constraint imposed on the channel matrix. This explains why the WR-based method is efficient only in the low SNR case, and concludes that the nulling-based approach is better for moderate to high SNRs. A new transmit scheme is then proposed to cancel the signal perturbation error at the receiver in order to improve the performance of the WR-based method in the case of high SNR. Moreover, a perturbation analysis of the new WR-based semi-blind method incorporating the proposed transmit scheme is conducted, leading to a novel closed-form expression for the mean square error (MSE) of the channel estimate. Computer simulations show that the proposed approach significantly outperforms the existing WR-based method as well as the nulling-based method for all SNRs.

Chapter 7: A novel signal-perturbation-free (SPF) linear prediction (LP)-based approach is presented for semi-blind estimation of time-domain MIMO-OFDM channels. A very efficient transmit scheme that can completely cancel the signal perturbation error at the receiver in the noise-free case is first proposed. Based on the new transmit structure, a signal-perturbation-free semi-blind channel estimation approach is then developed by using the MIMO linear prediction along with a whitening rotation algorithm in which the ambiguity matrix is estimated via a training-based

maximum likelihood method. Computer simulations show that the proposed signal-perturbation-free LP-based semi-blind solution significantly outperforms the LP semi-blind method without using the proposed transmit scheme as well as the LS method.

Chapter 8: This final chapter contains concluding remarks and some possible directions for future work to extend this research.

1.5 Main Contributions

The work presented in this thesis represents a number of novel contributions to the field of MIMO-OFDM channel estimation. The most significant contributions of this research are summarized as follows:

1. A nulling-based semi-blind approach is proposed for MIMO-OFDM channel estimation. Unlike most of the existing blind and semi-blind MIMO-OFDM channel estimation methods, which require a large number of OFDM symbols, the proposed approach is able to achieve a very high channel estimation accuracy by using only a few OFDM symbols while the full or partial information of the channel correlation is not needed.

2. A perturbation analysis of linear prediction-based channel estimation algorithms is conducted. It proves that the proposed nulling-based semi-blind algorithm is not subject to the signal perturbation error, and justifies why the conventional LP-based blind algorithms suffer from a poor performance in MIMO-OFDM channel estimation even in the high SNR case. Moreover, a closed-form expression is derived for the MSE of the blind constraint in the nulling-based

semi-blind channel estimation approach, further facilitating the calculation of the weighting factor.

3. Two channel estimation methods are proposed for pulse-shaped MIMO-OFDM systems. By exploiting the pulse-shaping filter available in the transmitter and the matched filter in the receiver, a semi-blind estimation algorithm and an enhanced LS algorithm are developed for the sampling- and upsampling-duration-based channels, respectively.

4. A frequency-domain correlation matrix estimation algorithm is proposed for channel estimation of MIMO-OFDM systems. By computing the time-domain correlation matrix directly from the received frequency-domain signal, an IFFT operation is avoided, that is usually required in time-domain estimation methods to convert a good frequency-domain signal to time domain when a high quality time-domain signal is not available.

5. An efficient semi-blind sparse channel estimation approach, which comprises the MST detection and sparse channel estimation both in a semi-blind fashion, is proposed for MIMO-OFDM systems. Simulation studies based on various sparse channels confirm that the proposed sparse semi-blind approach significantly outperforms the sparse LS method as well as the regular LS and semi-blind techniques.

6. A very efficient signal-perturbation-free transmit scheme is proposed to cancel the signal perturbation error at the receiver to improve the performance of the

frequency-domain channel estimation. This novel idea is also extended for the estimation of time-domain channels.

7. A new whitening-rotation (WR)-based semi-blind method using the signal-perturbation-free transmit scheme is developed for the frequency-domain channel estimation. A thorough study of the new WR-based algorithm via perturbation analysis leads to a novel closed-form expression for the mean square error (MSE) of the channel estimate.

8. A signal-perturbation-free linear prediction-based semi-blind approach is proposed for MIMO-OFDM systems, which can achieve a significantly improved channel estimation performance over some other training-based and semi-blind techniques.

Chapter 2

Preliminaries

This chapter introduces some basic concepts and principles involved in the development of channel estimation techniques for MIMO and MIMO-OFDM systems. Some state of the art literatures relating to the proposed work are also reviewed, providing necessary background materials for the development of new approaches in later chapters.

2.1 Modelling of MIMO-OFDM Systems

2.1.1 Transceiver

Fig. 2.1 shows a block diagram of a typical transmitter in a MIMO-OFDM system with the V-BLAST structure, in which there are N_T independent links, each connected to a transmit antenna and containing both pilots and information data. The m-th OFDM symbol can be written as a vector of the frequency-domain signals, namely,

$$\mathbf{X}_{i_T}(m) \triangleq [X_{i_T}(m,0), X_{i_T}(m,1),, \cdots, X_{i_T}(m,K-1)]^T.$$

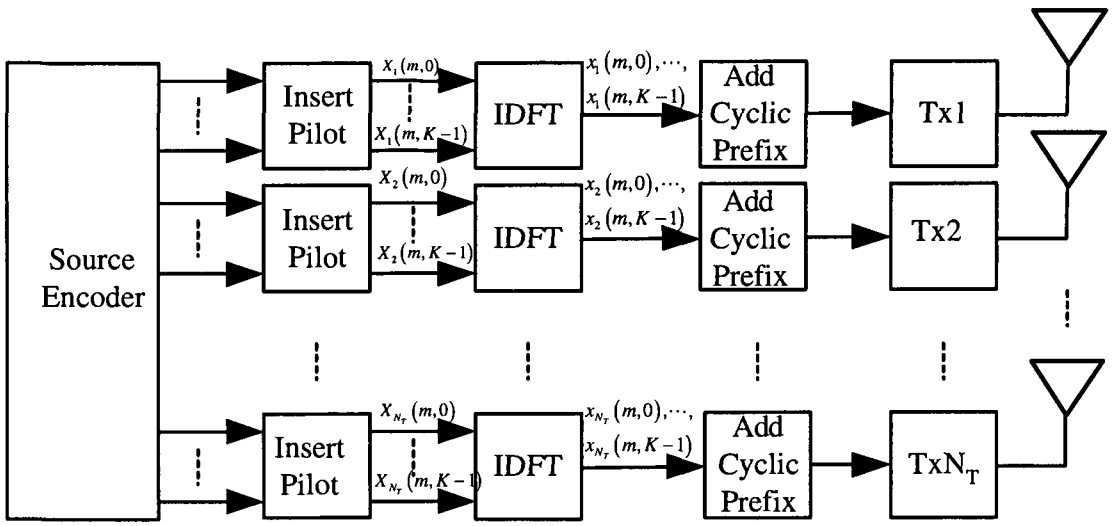

Figure 2.1: Schematic representation of an MIMO-OFDM transmitter

where K denotes the number of subcarriers. The output of the IDFT gives the time-domain OFDM signal,

$$\mathbf{x}_{i_T}(m) \triangleq [x_{i_T}(m,0), x_{i_T}(m,1), \cdots, x_{i_T}(m,K-1)]^T.$$

After adding a cyclic prefix, each OFDM signal passes through a transmit pulse-shaping filter and is then sent out by the corresponding antenna.

Fig. 2.2 shows the block diagram of the MIMO-OFDM receiver including N_R receive antennas as well as a channel estimation unit. After passing through the matched filter and removing cyclic prefix in each link, the received signal at the i_R-th link can be described as

$$\mathbf{y}_{i_R}(m) \triangleq [y_{i_R}(m,0), y_{i_R}(m,1), \cdots, y_{i_R}(m,K-1)]^T. \tag{2.1}$$

Then, the received frequency domain signal after the DFT processing is given by

$$\mathbf{Y}_{i_R}(m) \triangleq [Y_{i_R}(m,0), Y_{i_R}(m,1), \cdots, Y_{i_R}(m,K-1)]^T.$$

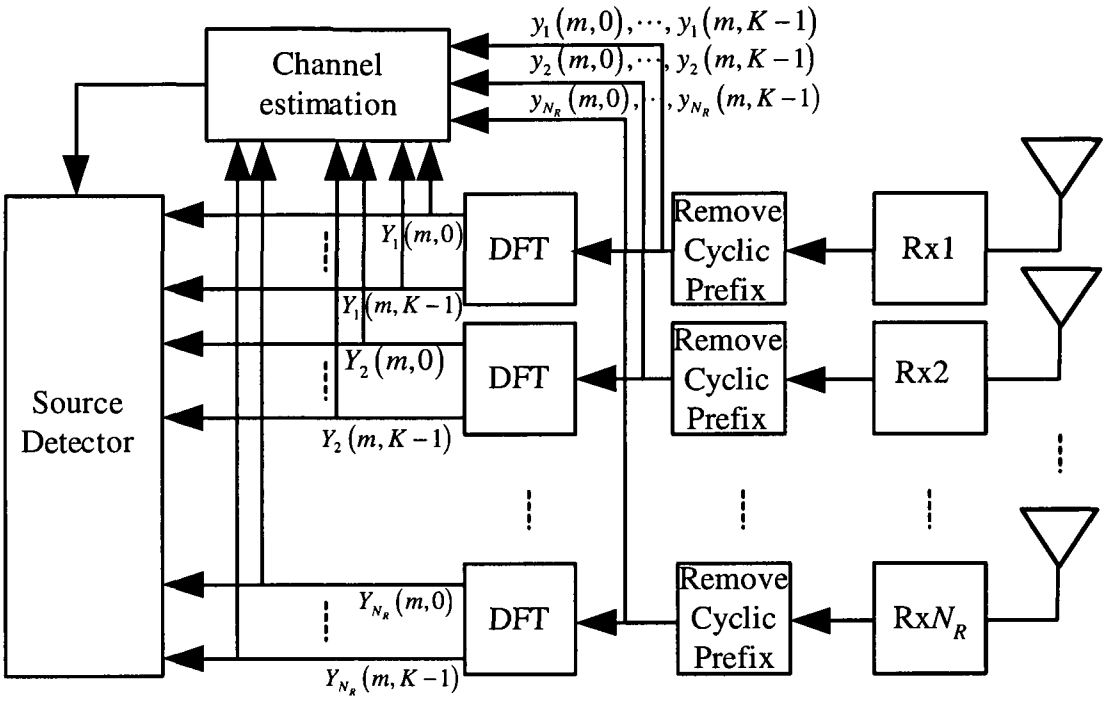

Figure 2.2: Schematic representation of MIMO-OFDM receiver

2.1.2 Channel Model

Considering that MIMO-OFDM systems are designed for broadband wireless communications, the signal bandwidth is always larger than the coherence bandwidth, implying that the channel is frequency-selective. Depending on the geometry of the antenna array and scatterers, there are two different multipath MIMO channel models (or a combination of the two), which are described below [59]:

1. Beamforming model: In this model, the elements of both the transmit and receive antenna arrays are co-located and the scatters can be considered as point sources. Each multipath channel is characterized by a direction of departure (DOD), direction of arrival (DOA), time of arrival (TOA) and a complex fading amplitude. In general, this model fits outdoor channels. For instance, it is

adopted as the MIMO model in the 3rd Generation Partnership Project (3GPP) TR 25.996.

2. Diversity model: The elements of the transmit and/or receive antenna arrays are not co-located and/or different scatters are modelled as distributed sources. This model is generally suitable for an indoor channel. In this model, the channel gain from a transmit to a receive antenna is modelled as both a spatially and temporally correlated jointly Gaussian random variable with zero mean (Rayleigh fading). This model is known to be useful for wireless LAN in IEEE P802.11 (IEEE 802.11-03).

Nevertheless, both of these channel models can be considered as a combination of L_c multi-paths, namely,

$$\mathbf{H}_c(t) = \sum_{l=0}^{L_c-1} \mathbf{\Gamma}_l \delta(t - t_l)$$

where t_l is the delay of the l-th path and $\mathbf{\Gamma}_l$ is an $N_R \times N_T$ attenuation matrix. Considering the pulse-shaping filter $g_t(t)$ in the transmitter and the matched filter $g_r(t)$ in the receiver, the composite channel can be represented by an $N_R \times N_T$ matrix $\mathbf{H}(t)$, with its (i_R, i_T)-th element as

$$h_{i_R,i_T}(t) = h_{i_R,i_T,c}(t) * g_t(t) * g_r(t) \tag{2.2}$$

where $h_{i_R,i_T,c}(t)$ is the (i_R, i_T)-th element of $\mathbf{H}_c(t)$. Most of the existing channel estimation methods focus on the composite discrete-time channel, i.e., the sampled version of the continuous-time channel response [78]. Thus, each element of the discrete-time MIMO-FIR channel is an L-tap FIR filter. Moreover, the channel is

considered constant during one OFDM symbol, even though it may change over different symbols. Therefore, for the l-th tap, the channel matrix is given by

$$\mathbf{H}(m,l) \triangleq \begin{bmatrix} h_{1,1}(m,l) & h_{1,2}(m,l) & \dots & h_{1,N_T}(m,l) \\ h_{2,1}(m,l) & h_{2,2}(m,l) & \dots & h_{2,N_T}(m,l) \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_R,1}(m,l) & h_{N_R,2}(m,l) & \dots & h_{N_R,N_T}(m,l) \end{bmatrix} \in \mathcal{C}^{N_R \times N_T}$$

where $h_{i_R,i_T}(m,l), (0 \leq l \leq L-1)$ represents the composite channel response between the i_R-th receive antenna and i_T-th transmit antenna for the l-th tap for the m-th OFDM symbol.

2.1.3 Signal Model

We first consider the signal model in the time domain. For notational simplicity, the index m of OFDM symbols can be dropped without loss of clarity. Thus, $\mathbf{H}(m,l)$, $x_{i_T}(m,n)$ and $y_{i_R}(m,n)$ are reduced to $\mathbf{H}(l)$, $x_{i_T}(n)$ and $y_{i_R}(n)$, respectively. If the length of the cyclic prefix is not less than the channel length L, the time-domain signal model for the frequency-selective fading channel can be written as

$$y_{i_R}(m,n) = \sum_{i_T=1}^{N_T} h_{i_R,i_T}(n) \circledast x_{i_T}(m,n) + v_{i_R}(m,n), m \in \{0,\cdots,g-1\} \qquad (2.3)$$

where g is the number of OFDM symbols within which the channel remains unchanged, and $v_{i_R}(m,n) \in \mathcal{C}^{N_R \times 1}$ is a spatio-temporally uncorrelated noise with zero-mean and variance σ_v^2.

Now we will address the signal model in the frequency domain [25]. Define

$$\mathbf{h}_{i_R,i_T} = [h_{i_R,i_T}(0),\cdots,h_{i_R,i_T}(L-1)]^T. \qquad (2.4)$$

After the cyclic prefix is removed, the received signal vector at the i_R-th receive antenna, which is defined in (2.1), can be rewritten as

$$\mathbf{y}_{i_R}(m) = \sum_{i_T=1}^{N_T} \bar{\mathbf{H}}_{i_R,i_T} \mathbf{F}_0^H \mathbf{X}_{i_T}(m) + \mathbf{v}_{i_R}(m), m \in \{0, \cdots, g-1\} \tag{2.5}$$

where $\bar{\mathbf{H}}_{i_R,i_T}$ is a circulant matrix with first column given by $\left[\mathbf{h}_{i_R,i_T}^T, \mathbf{0}_{1\times(K-L)}\right]^T$, the vector $\mathbf{v}_{i_R}(m)$ is the noise vector, and $\mathbf{F}_0$ is the $K \times K$ unitary DFT matrix. It can be easily found that the eigenvalue decomposition of $\bar{\mathbf{H}}_{i_R,i_T}$ leads to:

$$\bar{\mathbf{H}}_{i_R,i_T} = \mathbf{F}_0^H \mathrm{diag}\left\{\sqrt{K}\mathbf{F}_0\left[\mathbf{h}_{i_R,i_T}^T, \mathbf{0}_{1\times(K-L)}\right]^T\right\}\mathbf{F}_0.$$

Taking the DFT on both sides of (2.5), the signal model in the frequency domain can be derived as:

$$\mathbf{Y}_{i_R}(m) = \sum_{i_T=1}^{N_T} \mathrm{diag}\left\{\sqrt{K}\mathbf{F}_0\left[\mathbf{h}_{i_R,i_T}^T, \mathbf{0}_{1\times(K-L)}\right]^T\right\}\mathbf{X}_{i_T}(m) + \boldsymbol{\xi}_{i_R}(m), m \in \{0, \cdots, g-1\} \tag{2.6}$$

where $\boldsymbol{\xi}_{i_R}(m) = \mathbf{F}_0 \mathbf{v}_{i_R}(m)$.

2.2 Estimation of Frequency-Domain Channel of MIMO-OFDM Systems

With the OFDM modulation, the MIMO channel with respect to the k-th subcarrier can be described as

$$\mathbf{H}_{\mathrm{F}}(k) = \frac{1}{\sqrt{K}} \sum_{n=0}^{L-1} \mathbf{H}(n)\, \mathrm{e}^{-j2\pi(kn/K)}.$$

Due to some constraints such as the low complexity requirement for channel estimation and only a few number of subcarriers being used for the data transmis-

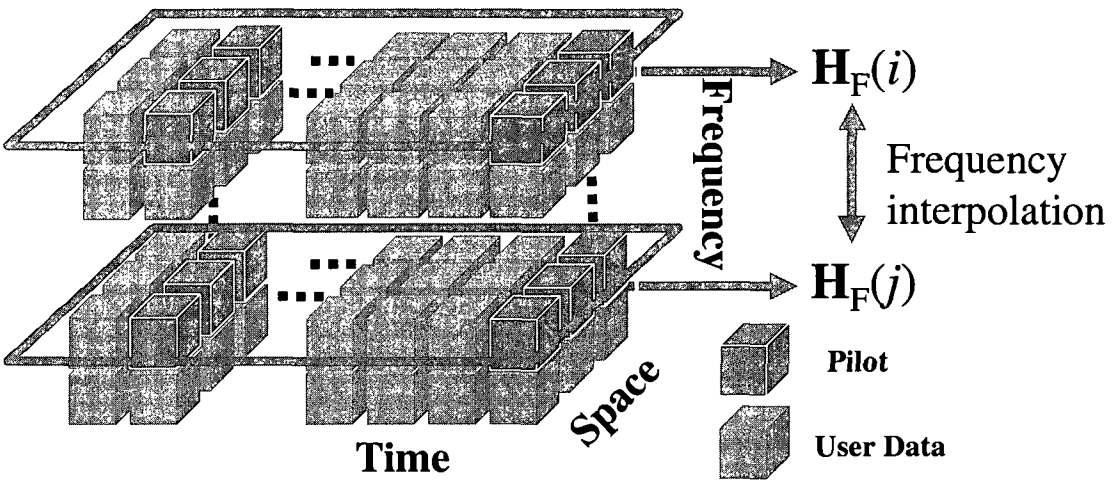

Figure 2.3: Estimation of frequency-domain channel of MIMO-OFDM systems

sion of a mobile user in OFDMA systems, the channel estimation is often implemented in the frequency-domain as shown in Fig. 2.3. In the frequency-domain approach, instead of estimating the time-domain channel $\mathbf{H}(l)$, $(l = 0, 1, \cdots, L-1)$, the frequency-domain channel $\mathbf{H}_F(k)$ with respect to certain subcarriers are first estimated individually, which are in turn used to estimate the frequency-domain channel with respect to other desired subcarriers by using an interpolation technique. In this case, for each specific subcarrier, the MIMO channel can be simply denoted as $\mathbf{H}$ by dropping the index k without loss of clarity. Thus, the received signal vector $\mathbf{y}(n)$ can be written as

$$\mathbf{y}(n) = \mathbf{H}\mathbf{x}(n) + \mathbf{v}(n). \tag{2.7}$$

In the following, both training-based and semi-blind algorithms are introduced for the estimation of the frequency-flat fading MIMO channel.

2.2.1 Training-based MIMO Channel Estimation

In general, in each data block, the first N_{P} of the N signals are used for training purpose. Defining a pilot matrix consisting of all N_{P} pilots and the corresponding received matrix, respectively, as

$$\mathbf{X}_{\mathrm{P}} \triangleq [\mathbf{x}(1), \cdots, \mathbf{x}(N_{\mathrm{P}})],$$

$$\mathbf{Y}_{\mathrm{P}} \triangleq [\mathbf{y}(1), \cdots, \mathbf{y}(N_{\mathrm{P}})],$$

a training-based optimization problem can be written as

$$\min_{\mathbf{H}} \Delta = \|\mathbf{Y}_{\mathrm{P}} - \mathbf{H}\mathbf{X}_{\mathrm{P}}\|_F^2. \tag{2.8}$$

One efficient way to solve this problem is by using an training-based LS approach, yielding an channel estimate as given by

$$\hat{\mathbf{H}}_{\mathrm{LS}} = \mathbf{Y}_{\mathrm{P}}\mathbf{X}_{\mathrm{P}}^{\dagger}. \tag{2.9}$$

Given a constraint on the transmitted training pilot's power as $\|\mathbf{X}_{\mathrm{P}}\|_F^2 = N_{\mathrm{P}}N_T\sigma_x^2$, the optimal pilot can be found by minimizing the following cost function

$$\min_{\mathbf{X}_{\mathrm{P}}} J_{LS} = \min_{\mathbf{X}_{\mathrm{P}}} \mathrm{E}\left\{\left\|\mathbf{H} - \hat{\mathbf{H}}_{\mathrm{LS}}\right\|_F^2\right\} \text{ subject to } \|\mathbf{X}_{\mathrm{P}}\|_F^2 = N_{\mathrm{P}}N_T\sigma_x^2. \tag{2.10}$$

It has been shown in [12] that the optimal pilot matrix should satisfy

$$\mathbf{X}_{\mathrm{P}}\mathbf{X}_{\mathrm{P}}^H = N_{\mathrm{P}}\sigma_x^2\mathbf{I}_{N_T}.$$

It implies that the optimal pilot matrix should consist of orthogonal vectors with the same norm, i.e., $\sqrt{N_{\mathrm{P}}}\sigma_x$. In this case, the MSE of the LS method can be derived as

$$\mathrm{MSE}_{\mathrm{LS}} \triangleq \mathrm{E}\left\{\left\|\hat{\mathbf{H}}_{\mathrm{LS}} - \mathbf{H}\right\|_F^2\right\} = \frac{N_R N_T \sigma_v^2}{N_{\mathrm{P}}\sigma_x^2}. \tag{2.11}$$

It is obvious that the channel estimation error increases when the number of the transmit or receive antennas is increased. On the other hand, the error can be reduced by increasing the transmitted power or the number of pilot slots.

If the channel correlation $\mathbf{R_H} = \mathrm{E}\left\{\mathbf{H}\mathbf{H}^H\right\}$ is known *a priori*, a channel estimate with improved accuracy can be obtained by using an MMSE estimator [12]

$$\hat{\mathbf{H}}_{\mathrm{MMSE}} = \mathbf{Y}_{\mathrm{P}}\left(\mathbf{X}_{\mathrm{P}}^H\mathbf{R_H}\mathbf{X}_{\mathrm{P}} + N_R\sigma_v^2\mathbf{I}_{N_T}\right)^{-1}\mathbf{X}_{\mathrm{P}}^H\mathbf{R_H}. \tag{2.12}$$

The performance of this estimator can be evaluated by

$$J_{\mathrm{MMSE}} \triangleq \mathrm{E}\left\{\left\|\mathbf{H} - \hat{\mathbf{H}}_{\mathrm{MMSE}}\right\|_F^2\right\} = \mathrm{Trace}\left\{\left(\mathbf{R}_{\mathbf{H}}^{-1} + N_R^{-1}\sigma_v^{-2}\mathbf{X}_{\mathrm{P}}\mathbf{X}_{\mathrm{P}}^H\right)^{-1}\right\}.$$

Given the transmit power constraint $\|\mathbf{X}_{\mathrm{P}}\|_F^2 = N_{\mathrm{P}}N_T\sigma_x^2$, the optimal pilot matrix for the MMSE method should satisfy

$$\mathbf{X}_{\mathrm{P}}\mathbf{X}_{\mathrm{P}}^H = \frac{1}{N_T}\left(N_{\mathrm{P}}N_T\sigma_x^2 + \sigma_v^2N_R\mathrm{Trace}\left\{\mathbf{R}_{\mathbf{H}}^{-1}\right\}\right)\mathbf{I}_{\mathbf{N_R}} - N_R\sigma_v^2\mathbf{R}_{\mathbf{H}}^{-1}.$$

In this case, it can be shown that the MSE of the MMSE method is given by

$$\mathrm{MSE}_{\mathrm{MMSE}} = \frac{N_T^2N_R\sigma_v^2}{N_{\mathrm{P}}N_T\sigma_x^2 + N_T\sigma_v^2\mathrm{Trace}\left\{\mathbf{R}_{\mathbf{H}}^{-1}\right\}}. \tag{2.13}$$

From (2.11) and (2.13), the gain of the MMSE method over the LS method in terms of the MSE is found to be

$$\frac{1}{\frac{\mathrm{MSE}_{\mathrm{MMSE}}}{\mathrm{MSE}_{\mathrm{LS}}}} = 1 + \frac{N_T\sigma_v^2\mathrm{Trace}\left\{\mathbf{R}_{\mathbf{H}}^{-1}\right\}}{N_{\mathrm{P}}N_T\sigma_x^2}. \tag{2.14}$$

It should be mentioned that the performance of the MMSE method largely depends on $\mathbf{R_H}$, which is however unknown in practice. Thus, an accurate estimation of $\mathbf{R_H}$ is of crucial importance for the implementation of the MMSE method. Normally,

it can be done by using the second order statistics (SOS)-based methods. These methods, however, require a huge number of samples to obtain an accurate estimate of $\mathbf{R_H}$. In the next subsection, we will introduce two efficient subspace-based semi-blind channel estimation algorithms, which require only a small number of symbols.

2.2.2 Semi-Blind MIMO Channel Estimation

First, we introduce the WR-based semi-blind MIMO channel estimation algorithm [20, 21, 23, 24]. Its idea originates from a decomposition of the channel matrix,

$$\mathbf{H} = \mathbf{W}\mathbf{Q}^H, \tag{2.15}$$

where $\mathbf{W}$ is a whitening matrix and $\mathbf{Q}$ is a unitary rotation matrix. Performing the singular value decomposition (SVD) of $\mathbf{H}$ gives

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H. \tag{2.16}$$

As $\mathbf{H}$ has a full column rank, one can partition $\mathbf{U}$ and $\mathbf{\Sigma}$ according to the signal and the noise subspaces of $\mathbf{H}$, namely,

$$\mathbf{U} = [\mathbf{U}_S, \mathbf{U}_N], \quad \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_S \\ \mathbf{0} \end{bmatrix}.$$

Thus, (2.16) can be rewritten as

$$\mathbf{H} = \mathbf{U}_S\mathbf{\Sigma}_S\mathbf{V}^H. \tag{2.17}$$

Obviously, one possible choice of $\mathbf{W}$ and $\mathbf{Q}$ can be $\mathbf{U}_S\mathbf{\Sigma}_S$ and $\mathbf{V}$, respectively. Therefore, the WR-based channel estimation method may be implemented with two steps.

The first step is to estimate the whitening matrix $\mathbf{W}$ in a blind fashion. From (2.7), we can obtain

$$\mathbf{R_Y} \triangleq \mathrm{E}\left\{\mathbf{y}\left(n\right)\mathbf{y}^H\left(n\right)\right\} - \sigma_v^2\mathbf{I}_{N_R} = \sigma_x^2\mathbf{H}\mathbf{H}^H. \tag{2.18}$$

Substituting (2.17) into (2.18) yields

$$\mathbf{R_Y} = \mathbf{U}_S\mathbf{\Sigma}_S^2\mathbf{U}_S^H. \tag{2.19}$$

Utilizing (2.19), $\hat{\mathbf{U}}_S$ and $\hat{\mathbf{\Sigma}}_S$ can be estimated from the SVD of the matrix $\hat{\mathbf{R}}_\mathbf{Y} = \frac{1}{N}\sum_{n=0}^{N-1}\mathbf{y}\left(n\right)\mathbf{y}^H\left(n\right) - \sigma_v^2\mathbf{I}_{N_R}$. Accordingly, an estimate of the whitening matrix $\mathbf{W}$ is obtained as

$$\hat{\mathbf{W}} = \hat{\mathbf{U}}_S\hat{\mathbf{\Sigma}}_S. \tag{2.20}$$

We now consider the second step, estimating the rotation matrix $\mathbf{Q}$ using the pilot signal. As $\mathbf{Q}$ is an unitary matrix, the training-based estimation problem of $\mathbf{Q}$ can be formulated as

$$\min_{\hat{\mathbf{Q}}}\left\|\mathbf{Y}_\mathrm{P} - \hat{\mathbf{W}}\hat{\mathbf{Q}}\mathbf{X}_\mathrm{P}^H\right\|_F^2, \text{ subject to } \hat{\mathbf{Q}}\hat{\mathbf{Q}}^H = \mathbf{I}. \tag{2.21}$$

An ML solution to this optimization problem has been derived in [20,21,23,24] as

$$\hat{\mathbf{Q}} = \mathbf{V}_Q\mathbf{U}_Q^H. \tag{2.22}$$

where $\mathbf{U}_Q$ and $\mathbf{V}_Q$ are obtained from an SVD of the matrix

$$\mathbf{Y}_Q \triangleq \frac{1}{N_\mathrm{P}\sigma_x^2}\hat{\mathbf{W}}^H\mathbf{Y}_\mathrm{P}\mathbf{X}_\mathrm{P}^H, \tag{2.23}$$

namely,

$$\mathbf{Y}_Q = \mathbf{U}_Q\mathbf{\Sigma}_Q\mathbf{V}_Q^H. \tag{2.24}$$

It has been proved in [21] that when the pilot matrix is optimal in the LS sense, i.e., $\mathbf{X}_{\mathrm{P}}\mathbf{X}_{\mathrm{P}}^{H} = N_{\mathrm{P}}\sigma_x^2\mathbf{I}_{N_T}$, the Cramer-Rao bound (CRB) of the channel estimate is directly proportional to ζ the number of unconstrained parameters required to describe $\mathbf{H}$, namely,

$$\mathrm{E}\left\{\left\|\hat{\mathbf{H}} - \mathbf{H}\right\|_{\mathrm{F}}^2\right\} \geq \frac{\sigma_v^2}{2N_{\mathrm{P}}\sigma_x^2}\zeta. \tag{2.25}$$

Since the unitary rotation matrix $\mathbf{Q}$ has $\zeta_Q = N_T^2$ real parameters, from (2.25), one can obtain the CRB of the WR-based semi-blind algorithm as [21, 23]

$$\mathrm{E}\left\{\left\|\hat{\mathbf{H}} - \mathbf{H}\right\|_{\mathrm{F}}^2\right\} \geq \frac{N_T^2\sigma_v^2}{2N_{\mathrm{P}}\sigma_x^2}. \tag{2.26}$$

Although the performance of the above semi-blind method is superior to that of some training-based methods in the low SNR case, it gets poor with the increase of SNR. In the higher SNR case, a nulling-based semi-blind approach may be employed [17, 18, 42, 43]. Instead of estimating the whitening matrix $\mathbf{W}$, this approach uses a subspace method to obtain an estimate of the nulling subspace of the channel matrix, $\hat{\mathbf{U}}_{\mathrm{null}}$. Then, by utilizing $\hat{\mathbf{U}}_{\mathrm{null}}$ in conjunction with a training-based LS criterion (2.8), a semi-blind cost function can be formulated as

$$\min_{\mathbf{H}} \Delta = \|\mathbf{Y}_{\mathrm{P}} - \mathbf{H}\mathbf{X}_{\mathrm{P}}\|_F^2 + \alpha\left\|\hat{\mathbf{U}}_{\mathrm{null}}^H\mathbf{H}\right\|_F^2 \tag{2.27}$$

where $\alpha > 0$ is a weighting factor. Although the solution to this minimization problem gives a better estimation performance than the WR-based semi-blind method does for high SNRs, its superiority in the high SNR case has not been theoretically justified. Moreover, the weighting factor employed to trade off the least square and the blind criteria has not been appropriately determined.

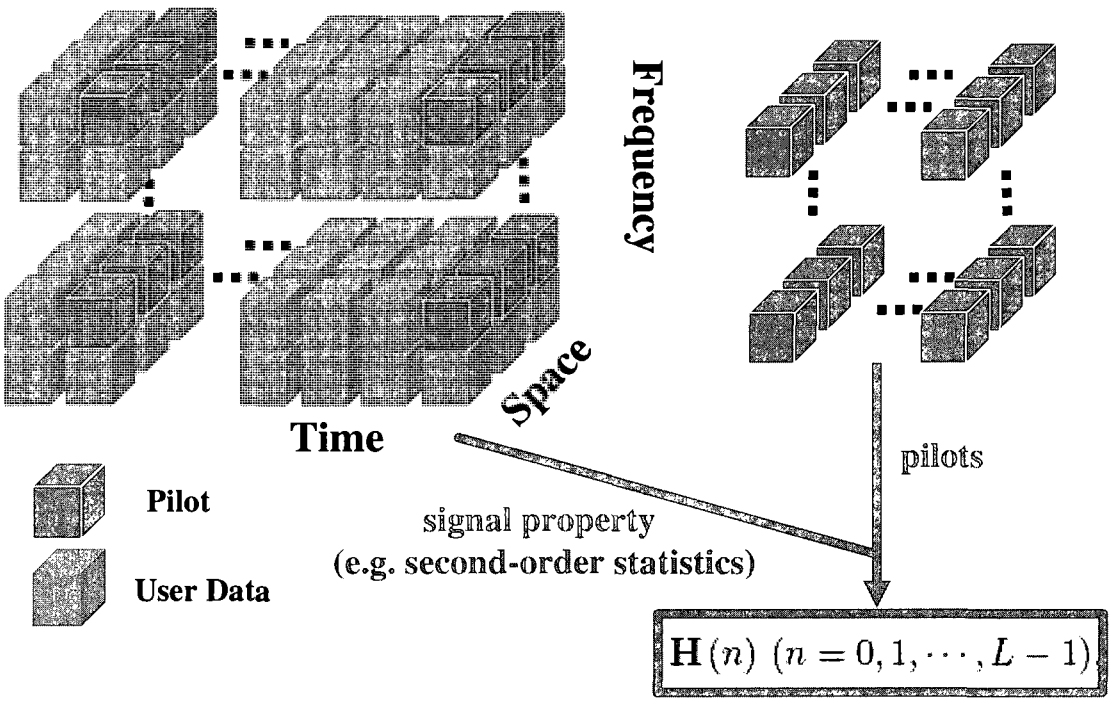

Figure 2.4: Estimation of time-domain channel of MIMO-OFDM systems

2.3 Estimation of Time-Domain Channel of MIMO-OFDM Systems

Although the algorithms of estimating frequency-domain channels have a low computational complexity, they require a large number of OFDM symbols to render an accurate channel estimation. In this section, we address the issue of estimating time-domain channels. As shown in Fig. 2.4, both pilot and data signals at all subcarriers can be employed for the estimation of the time-domain channel $\mathbf{H}(l)$, $(l = 0, 1, \cdots, L-1)$. Since the signals on all subcarriers are utilized, a good channel estimate can be expected by using only a small number of OFDM symbols in this approach.

Here, we introduce a training-based LS channel estimation algorithm for MIMO-OFDM systems [25]. Assume that the K_p sub-carriers, say from i_{pilot1} to $i_{\text{pilot}K_p}$, of each OFDM symbol carry the pilot signal. The transmitted and the received pilot vectors for each transmit and receive antenna pair can be defined as

$$\mathbf{X}_{i_T,\text{pilot}}(m) \triangleq \left[X_{i_T}(m, i_{\text{pilot1}}), \cdots, X_{i_T}\left(m, i_{\text{pilot}K_p}\right)\right]^T,$$

$$\mathbf{Y}_{i_R,\text{pilot}}(m) \triangleq \left[Y_{i_R}(m, i_{\text{pilot1}}), \cdots, Y_{i_R}\left(m, i_{\text{pilot}K_p}\right)\right]^T.$$

It should be noted that the pilot signal might not be located at the same position in different OFDM symbols. Let $\mathbf{F}_1$ be a $K \times L$ matrix formed by the first L columns of a $K \times K$ DFT matrix $\mathbf{F}_0$. For the m-th OFDM symbol, one can form a $K_P \times L$ matrix, say $\mathbf{F}(m)$, by taking only the rows of $\mathbf{F}_1$ associated with the K_p pilot sub-carriers. It can be easily verified that, for the pilot sub-carriers, (2.6) can be rewritten as

$$\mathbf{Y}_{i_R,\text{pilot}}(m) = \sum_{i_T=1}^{N_T} \mathbf{X}_{i_T,\text{pilot-diag}}(m)\, \mathbf{F}(m)\, \mathbf{h}_{i_R,i_T} + \boldsymbol{\xi}_{i_R,\text{pilot}}(m) \tag{2.28}$$

where

$$\mathbf{X}_{i_T,\text{pilot-diag}}(m) \triangleq \text{diag}\left(\mathbf{X}_{i_T,\text{pilot}}(m)\right).$$

From (2.28), the frequency-domain pilot signal received at the i_R-th receive antenna with respect to g OFDM symbols can be obtained as

$$\mathbf{Y}_{i_R,\text{pilot}} = \mathbf{A}\mathbf{h}_{i_R} + \boldsymbol{\xi}_{i_R,\text{pilot}} \tag{2.29}$$

where

$$\mathbf{Y}_{i_R,\text{pilot}} \triangleq \left[\mathbf{Y}^T_{i_R,\text{pilot}}(0), \cdots, \mathbf{Y}^T_{i_R,\text{pilot}}(g-1)\right]^T,$$

$$\mathbf{A} \triangleq \begin{bmatrix} \mathbf{X}_{1,\text{pilot-diag}}(0)\,\mathbf{F}(0) & \cdots & \mathbf{X}_{N_T,\text{pilot-diag}}(0)\,\mathbf{F}(0) \\ \vdots & \ddots & \vdots \\ \mathbf{X}_{1,\text{pilot-diag}}(g-1)\,\mathbf{F}(g-1) & \cdots & \mathbf{X}_{N_T,\text{pilot-diag}}(g-1)\,\mathbf{F}(g-1) \end{bmatrix},$$

$$\mathbf{h}_{i_R} \triangleq \left[\mathbf{h}_{i_R,1}^T, \cdots, \mathbf{h}_{i_R,N_T}^T\right]^T, \tag{2.30}$$

$$\boldsymbol{\xi}_{i_R,\text{pilot}} \triangleq \left[\boldsymbol{\xi}_{i_R,\text{pilot}}^T(0), \cdots, \boldsymbol{\xi}_{i_R,\text{pilot}}^T(g-1)\right]^T.$$

From (2.29), one can easily verify that if $\mathbf{A}$ has a full column rank of LN_T, the channel with respect to the i_R-th receive antenna can be estimated using the least squares approach as

$$\hat{\mathbf{h}}_{i_R} = \mathbf{A}^\dagger \mathbf{Y}_{i_R,\text{pilot}} = \mathbf{h}_{i_R} + \mathbf{A}^\dagger \boldsymbol{\xi}_{i_R,\text{pilot}}. \tag{2.31}$$

It has been shown in [25] that by using a small number of OFDM symbols in the least squares method, the MIMO-OFDM channel can be estimated with a high accuracy. In the next chapter, based on the LS MIMO-OFDM channel estimation algorithm, we will derive a training-based criterion and combine it with a blind criterion to formulate a semi-blind channel estimation problem, leading to an appealing channel estimation solution that is superior to the original LS approach.

2.4 Linear Prediction-based Channel Estimation Algorithms for MIMO Systems

In the previous two sections, we have introduced a few algorithms for MIMO-OFDM channel estimation. Considering that there exists some similarity between the MIMO system model and its OFDM counterpart, one may expect that the chan-

nel estimation techniques originally proposed for MIMO systems be generalized for MIMO-OFDM systems. In this section, we present two efficient MIMO channel estimation algorithms that are based on linear prediction. This linear prediction idea will be extended for MIMO-OFDM channel estimation in the next chapter. Linear prediction has been widely used in blind MIMO channel estimation and equalization [17, 39–41, 44–46], where the key idea is to represent the received MIMO signal as a finite-order *autoregressive* (AR) process under the assumption that the transmitted signals are uncorrelated in time [39]. In the following, we review in brief the linear prediction-based blind and semi-blind channel estimation techniques for MIMO systems.

2.4.1 Linear Prediction-based Blind Channel Estimation

Here, we introduce the MIMO linear prediction and its application to blind channel estimation [39]. Consider a MIMO system with N_T transmit and N_R ($> N_T$) receive antennas. The MIMO channel can be characterized by an array of L-tap FIR filters described as L matrices $\mathbf{H}(n)$ ($n = 0, 1, \cdots, L-1$) of size $N_R \times N_T$, whose (i_R, i_T)-th element $h_{i_T, i_R}(n)$ represents the channel response from the i_T-th transmit antenna to the i_R-th receive antenna. Given the transmitted signal vector

$$\mathbf{x}(n) \triangleq [x_1(n), \cdots, x_{N_T}(n)]^T, \tag{2.32}$$

the received signal vector can be written as

$$\mathbf{y}(n) \triangleq [y_1(n), \cdots, y_{N_R}(n)]^T, \tag{2.33}$$

with its element being given by

$$y_{i_R}(n) = \sum_{i_T=1}^{N_T} h_{i_R,i_T}(n) * x_{i_T}(n) + v_{i_R}(n) \tag{2.34}$$

where $*$ denotes the linear convolution and $v_{i_R}(n)$ is a spatio-temporally uncorrelated zero mean noise with variance σ_v^2.

Consider the problem of predicting $\mathbf{y}(n)$ from

$$\mathbf{y}_P(n-1) \triangleq \left[\mathbf{y}^T(n-1), \cdots, \mathbf{y}^T(n-P)\right]^T$$

for the noise-free case. The prediction error can be defined as

$$\tilde{\mathbf{y}}(n)\big|_{\mathbf{y}_P(n-1)} \triangleq \mathbf{y}(n) - \hat{\mathbf{y}}(n)\big|_{\mathbf{y}_P(n-1)} = \tilde{\mathbf{P}}_P \mathbf{y}_{P+1}(n)$$

where $\tilde{\mathbf{P}}_P = [\mathbf{I}_{N_R}, -\mathbf{P}_P]$ and $\mathbf{P}_P$ is the linear predictor matrix consisting of P $N_R \times N_R$ matrices as given by

$$\mathbf{P}_P = [\mathbf{P}_P(1), \mathbf{P}_P(2), \cdots, \mathbf{P}_P(P)].$$

Minimizing the variance of the prediction error leads to the following optimization problem [17]

$$\boldsymbol{\sigma}^2_{\tilde{\mathbf{y}},P} \triangleq \min_{\mathbf{P}_P} \mathrm{E}\left\{\tilde{\mathbf{y}}(n)\big|_{\mathbf{y}_P(n-1)}\, \tilde{\mathbf{y}}^H(n)\big|_{\mathbf{y}_P(n-1)}\right\} = \min_{\mathbf{P}_P} \tilde{\mathbf{P}}_P \mathbf{R}_{\mathbf{y}_{P+1}} \tilde{\mathbf{P}}_P^H \tag{2.35}$$

with

$$\tilde{\mathbf{P}}_P \mathbf{R}_{\mathbf{y}_{P+1}} = \begin{bmatrix} \boldsymbol{\sigma}^2_{\tilde{\mathbf{y}},P} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}. \tag{2.36}$$

Defining

$$\mathbf{R}(k = k_2 - k_1) \triangleq \mathrm{E}\left[\mathbf{y}(n-k_1)\mathbf{y}^H(n-k_2)\right],$$

$$\tilde{\mathbf{R}}_{n-1} \triangleq \mathrm{E}\left[\mathbf{y}_P(n-1)\,\mathbf{y}_P^H(n-1)\right] \tag{2.37}$$
$$= \begin{bmatrix} \mathbf{R}(0) & \mathbf{R}(1) & \cdots & \mathbf{R}(P-1) \\ \mathbf{R}(-1) & \mathbf{R}(0) & \cdots & \mathbf{R}(P-2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}(1-P) & \mathbf{R}(2-P) & \cdots & \mathbf{R}(0) \end{bmatrix},$$

$$\ddot{\mathbf{R}}_n \triangleq \mathrm{E}\left[\mathbf{y}(n)\,\mathbf{y}_P^H(n-1)\right] \tag{2.38}$$
$$= \begin{bmatrix} \mathbf{R}(1) & \mathbf{R}(2) & \cdots & \mathbf{R}(P-1) \end{bmatrix},$$

the MIMO linear predictor as a solution to the optimization problem in (2.35) can be written as [39,40,44,79]

$$\mathbf{P}_P = \ddot{\mathbf{R}}_n \tilde{\mathbf{R}}_{n-1}^{-1}. \tag{2.39}$$

It should be mentioned that, if the additive white Gaussian noise (AWGN) is considered, $\tilde{\mathbf{R}}_{n-1}$ can be calculated by

$$\tilde{\mathbf{R}}_{n-1} = \mathrm{E}\left[\mathbf{y}_P(n-1)\,\mathbf{y}_P^H(n-1)\right] - \sigma_v^2\mathbf{I}.$$

Moreover, from (2.35), the covariance matrix of the prediction error can be derived as

$$\boldsymbol{\sigma}_{\tilde{\mathbf{y}},P}^2 = \mathbf{R}(0) - \mathbf{P}_P\ddot{\mathbf{R}}_n^H. \tag{2.40}$$

Further, by defining

$$\mathbf{P}_P(z) = \mathbf{I} - \sum_{n=1}^{P} \mathbf{P}_P(n)\, z^{-n}$$
$$\mathbf{H}(z) = \sum_{n=0}^{L-1} \mathbf{H}(n)\, z^{-n},$$

it has been shown in [39, 40] that if the transmitted signals are uncorrelated and moreover, $PN_R \geq (L+P-1)N_T$, one can obtain

$$\mathbf{P}_P(z)\mathbf{H}(z) = \mathbf{H}(0), \tag{2.41}$$

$$\boldsymbol{\sigma}^2_{\tilde{\mathbf{y}},P} = \mathbf{H}(0)\mathbf{H}^H(0). \tag{2.42}$$

Based on (2.41) and (2.42), some blind algorithms have been proposed for MIMO channel estimation [39, 79, 80]. The basic idea is to first acquire an estimate of $\mathbf{H}(0)$ from that of $\boldsymbol{\sigma}^2_{\tilde{\mathbf{y}},P}$ according to (2.42), and then use (2.41) to obtain an estimate of the channel matrix $\mathbf{H}(z)$.

2.4.2 Linear Prediction-based Semi-Blind Channel Estimation

Using the above linear prediction along with training data, a semi-blind approach for MIMO channel estimation has been developed to achieve a better estimation performance [17, 42, 43]. This approach is briefly described as follows.

Denote the null column space of $\mathbf{H}(0)$ as an $N_R \times (N_R - N_T)$ matrix $\mathbf{U}_{0null}$. From (2.42), one can easily find that $\mathbf{U}_{0null}$ can be estimated from $\boldsymbol{\sigma}^2_{\tilde{\mathbf{y}},P}$. Using $\mathbf{U}_{0null}$ into (2.41) gives

$$\mathbf{U}^H_{0null}\mathbf{H}(0) = \mathbf{U}^H_{0null}\mathbf{P}_P(z)\mathbf{H}(z) = \mathbf{0}, \tag{2.43}$$

which imposes a blind constraint $\mathbf{U}^H_{0null}\mathbf{P}_P(z)$ on the channel matrix $\mathbf{H}(z)$. The above equality can then be used to derive a blind constraint $\bar{\mathbf{B}}$ for the channel vector defined as

$$\mathbf{h} \triangleq \left[\mathbf{h}_1^T(0), \cdots, \mathbf{h}_{N_T}^T(0), \cdots, \mathbf{h}_1^T(L-1), \cdots, \mathbf{h}_{N_T}^T(L-1)\right]^T$$

where $\mathbf{h}_{i_T}(l) = [h_{1,i_T}(l), \cdots, h_{N_R,i_T}(l)]^T$. By using this blind constraint in conjunction with a training based least square (LS) criterion [75], a semi-blind cost function is then formulated as

$$\min_{\hat{\mathbf{h}}} \left\{ \left\| \mathbf{Y}_{\mathrm{TS}} - \mathbf{A}_{\mathrm{TS}}\hat{\mathbf{h}} \right\|^2 + \alpha \left\| \bar{\mathbf{B}}\hat{\mathbf{h}} \right\|^2 \right\} \tag{2.44}$$

where $\mathbf{A}_{\mathrm{TS}}$ is the pilot signal matrix, $\mathbf{Y}_{\mathrm{TS}}$ the corresponding received signal vector, and $\alpha > 0$ is a weighting factor. The minimization problem (2.44) can be easily solved by treating α as a fixed constant, giving a semi-blind solution for the channel estimate $\hat{\mathbf{h}}$. It should be stressed that the estimation accuracy of the above semi-blind algorithm heavily depends on the choice of the weighting factor. This issue, however, has not been investigated. In the next chapter, we will extend the semi-blind MIMO channel estimation method for estimating MIMO-OFDM channels. We will also propose a very efficient scheme for the determination of the value of α according to the derived MSE of the least-square estimation as well as that of the blind constraint.

2.5 Conclusion

In this chapter, we have provided necessary background materials for the study of MIMO-OFDM estimation techniques in this dissertation. First, the modelling of MIMO-OFDM systems including the transceiver, channel model and signal model has been presented. Then, the existing channel estimation approaches of MIMO-OFDM systems have been described in detail. In particular, the training-based and semi-blind algorithms for the estimation of the frequency-domain channels and a training-based

LS algorithm for the estimation of the time-domain channels have been presented. At last, linear prediction-based blind and semi-blind channel estimation algorithms for MIMO systems have been discussed.

Chapter 3

A Nulling-based Semi-Blind Channel Estimation Approach

3.1 Introduction

Most of the existing blind and semi-blind MIMO-OFDM channel estimation methods are based on the second-order statistics of a long vector whose size is equal to or larger than the number of subcarriers. To estimate the correlation matrix reliably, they need a large number of OFDM symbols, which is not suitable for fast time-varying channels. In addition, since the matrices involved in these algorithms are of huge size, their computational complexity is extremely high. In contrast, a linear prediction-based semi-blind algorithm that is based on the second-order statistics of a short vector with a size only slightly larger the channel length, has been found more efficient than the conventional LS methods for the estimation of frequency-selective MIMO channels [17,42,43]. In this chapter, we will extend the linear prediction-based semi-blind approach to the channel estimation of MIMO-OFDM systems.

Linear prediction has been widely used in blind MIMO channel estimation and

equalization [17, 39–41, 44–46], and its key idea is to represent the received MIMO signal by a finite-order *autoregressive* (AR) process under the assumption that the transmitted signals are uncorrelated in time [39]. Based on the AR process, a linear prediction filter can be obtained to solve a second-order deconvolution problem for channel equalization. By combining the linear prediction with a higher-order statistics (HOS) or the weighted least-squares method, some blind channel estimation algorithms have been derived [39, 46]. However, these algorithms require a large number of signal samples and moreover, they are not robust. Medles et al. have proposed a semi-blind algorithm by incorporating a blind criterion derived from the linear prediction into a training-based LS cost function [17,42,43], leading to a closed-form expression for the estimate of the MIMO channel response. It has been shown in their papers that the semi-blind method provides a much better channel estimation performance over the pure training-based LS method. However, the superiority of the proposed semi-blind method has not been theoretically justified. Moreover, the weighting factor employed to trade off the least square and the blind criteria has not been appropriately determined. As such, the resulting channel estimation performance, though better than that of the LS method, could be further improved. It should also be noted that this semi-blind method for MIMO channel estimation cannot be directly applied to MIMO-OFDM systems, due to different signal models in the two systems. In MIMO systems, both information data and pilots are generated and transmitted in the time-domain, whereas in MIMO-OFDM systems, signals are first generated in the frequency domain and then converted to the time domain via the inverse discrete Fourier transform (IDFT). As a consequence, the training-based LS cost

function for MIMO channel estimation cannot be used for MIMO-OFDM systems. Also, the uncorrelation among the time-domain MIMO signals, which is desirable for linear prediction, may not be available in MIMO-OFDM systems. Therefore, a lot of work on the formulation of MIMO-OFDM signals has to be done first to develop a new semi-blind estimation solution.

In this chapter, we propose a channel estimation algorithm for MIMO-OFDM systems by employing the afore-mentioned semi-blind strategy. First, a new training-based LS criterion for MIMO-OFDM channel estimation is obtained through proper matrix formulations. Then, by proving that the transmitted time-domain MIMO-OFDM signals are uncorrelated, we validate the use of linear prediction for the formulation of the blind criterion. By employing the LS and the blind criteria, a semi-blind channel estimation solution is obtained. As a part of the semi-blind approach, we also propose an appealing scheme for choosing the weighting factor, which is shown to be very efficient according to our extensive computer simulations.

The second part of this chapter deals with the analysis of the MIMO linear prediction and the corresponding blind constraint. We apply the perturbation theory, which has been successfully used in analyzing the antenna array based signal processing algorithms [71–74, 76, 77], to the analysis of the proposed MIMO linear prediction based method, justifying the superiority of our semi-blind solution over the blind algorithms [39, 79, 80]. The perturbation analysis of the blind constraint also leads to a novel closed-form expression for the mean square error (MSE) of the blind estimation that is essential to the calculation of the weighting factor for the semi-blind estimation solution.

The rest of the chapter is organized as follows. Section 3.2 presents a new semi-blind approach for the estimation of MIMO-OFDM channels, including the formulation of the semi-blind estimation problem, derivation of the blind constraint from the linear prediction and development of an appealing scheme for the determination of the weighting factor. Section 3.3 conducts a perturbation analysis of the linear prediction-based semi-blind method, justifying the superior estimation performance of the semi-blind solution over a pure blind method. The analysis also yields a closed-form expression for the MSE of the blind estimation part of the proposed approach, facilitating the calculation of the weighting factor. Section 3.4 comprises a number of experimentations validating the proposed method, showing significant advantages of the semi-blind solution over the least-square method in terms of the MSE of the channel estimate. Finally, Section 3.5 highlights some of the distinct features of the proposed approach.

3.2 Proposed Nulling-based Semi-Blind Channel Estimation Algorithm for MIMO-OFDM Systems

Note that unlike the MIMO signal model in (2.34), the MIMO-OFDM signal given by (2.3) involves the circular convolution with a length K. In the following, we would like to develop a semi-blind MIMO-OFDM channel estimation approach based on (2.3) and the semi-blind criterion in (2.44) [81–83].

3.2.1 The Training based LS Criterion

In MIMO-OFDM systems, the training signal is transmitted in the frequency domain and thus, the LS criterion in (2.44) should be modified. We now derive a new LS criterion based on the approach in [25]. By defining

$$\bar{\mathbf{Y}}_{\text{pilot}} \triangleq [\mathbf{Y}_{1,\text{pilot}}, \cdots, \mathbf{Y}_{N_R,\text{pilot}}],$$

$$\mathbf{H} \triangleq [\mathbf{h}_1, \cdots, \mathbf{h}_{N_R}], \tag{3.1}$$

$$\bar{\boldsymbol{\xi}}_{\text{pilot}} \triangleq [\boldsymbol{\xi}_{1,\text{pilot}}, \cdots, \boldsymbol{\xi}_{N_R,\text{pilot}}],$$

from (2.29), one can have

$$\bar{\mathbf{Y}}_{\text{pilot}} = \mathbf{A}\mathbf{H} + \bar{\boldsymbol{\xi}}_{\text{pilot}}. \tag{3.2}$$

Further, letting

$$\mathbf{Y}_{\text{pilot}} \triangleq \text{vec}\left(\bar{\mathbf{Y}}_{\text{pilot}}\right),$$

$$\tilde{\mathbf{A}} \triangleq \mathbf{I} \otimes \mathbf{A},$$

$$\mathbf{h} \triangleq \text{vec}\left(\mathbf{H}\right),$$

from (3.2), one can obtain a new LS criterion

$$\left\| \mathbf{Y}_{\text{pilot}} - \tilde{\mathbf{A}}\hat{\mathbf{h}} \right\|^2, \tag{3.3}$$

which will be used for the training signal in the proposed semi-blind method.

3.2.2 The MIMO Linear Prediction based Blind Criterion

First of all, we show that the transmitted time-domain MIMO-OFDM signal is uncorrelated to validate the use of linear prediction for the received MIMO-OFDM

signal. Generally speaking, the frequency-domain signal in MIMO-OFDM systems before the DFT module can be considered as an independent identically distributted (i.i.d.) Gaussian process with zero mean and variance δ_x^2, implying that the frequency-domain signal at the i_T-th antenna $X_{i_T}(k)$ is uncorrelated in both time and spatial domains. Given $X_{i_{T1}}(k)$ and $X_{i_{T2}}(k)$ being the frequency-domain signals at the i_{T1}-th and i_{T2}-th antennas, respectively, the corresponding time-domain signals can be written as

$$x_{i_{T1}}(n_1) = \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} X_{i_{T1}}(k) \, \mathrm{e}^{j2\pi(kn_1/K)},$$

$$x_{i_{T2}}(n_2) = \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} X_{i_{T2}}(k) \, \mathrm{e}^{j2\pi(kn_2/K)}.$$

Consider the correlation between the two time-domain signals, i.e.

$$\mathrm{E}\left\{x_{i_{T1}}(n_1) \, x_{i_{T2}}^*(n_2)\right\} = \frac{1}{K} \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} \mathrm{E}\left\{X_{i_{T1}}(k_1) \, X_{i_{T2}}^*(k_2)\right\} \mathrm{e}^{j2\pi(k_1 n_1 - k_2 n_2)/K}. \quad (3.4)$$

Obviously

$$\mathrm{E}\left\{x_{i_{T1}}(n_1) \, x_{i_{T2}}^*(n_2)\right\} = 0 \text{ if } i_{T1} \neq i_{T2}. \quad (3.5)$$

If $i_{T1} = i_{T2}$, (3.4) reduces to

$$\mathrm{E}\left\{x_{i_{T1}}(n_1) \, x_{i_{T2}}^*(n_2)\right\} = \frac{1}{K} \sum_{k=0}^{K-1} \delta_x^2 \mathrm{e}^{j2\pi k(n_1-n_2)/K} = \begin{cases} \delta_x^2 & \text{if } (n_1 = n_2) \\ 0 & \text{if } (n_1 \neq n_2) \end{cases}. \quad (3.6)$$

It is clear from (3.5) and (3.6) that the transmitted time-domain MIMO-OFDM signal is uncorrelated.

We now use the MIMO linear prediction to obtain a blind constraint for the channel matrix. By following the linear prediction process in Section 2.4, we can obtain a time-domain representation of (2.41),

$$[\mathbf{I}, -\mathbf{P}_P] \, \mathbf{H}_D = [\mathbf{H}(0), \mathbf{0}, \cdots, \mathbf{0}] \quad (3.7)$$

where $\mathbf{P}_P$ is the linear predictor as given by (2.39) and $\mathbf{H}_D$ is a $(P+1)N_R \times (L+P)N_T$ block Toeplitz matrix with the first block row as

$$[\mathbf{H}(0), \cdots, \mathbf{H}(L-1), \mathbf{0}, \cdots, \mathbf{0}].$$

Letting

$$\mathbf{H}_F \triangleq \begin{bmatrix} \mathbf{H}(0) \\ \vdots \\ \mathbf{H}(L-1) \end{bmatrix}, \tag{3.8}$$

$$\mathbf{P}_Q \triangleq \begin{bmatrix} \mathbf{I}_{N_R} & & & \mathbf{0} \\ -\mathbf{P}_P(1) & \ddots & & \\ \vdots & \ddots & \ddots & \\ -\mathbf{P}_P(P) & \vdots & \ddots & \mathbf{I}_{N_R} \\ & \ddots & \vdots & -\mathbf{P}_P(1) \\ & & \ddots & \vdots \\ \mathbf{0} & & & -\mathbf{P}_P(P) \end{bmatrix}, \tag{3.9}$$

(3.7) can be rewritten as

$$\mathbf{P}_Q \mathbf{H}_F = \begin{bmatrix} \mathbf{H}(0) \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}. \tag{3.10}$$

Using (2.42), the null column space of $\mathbf{H}(0)$, $\mathbf{U}_{0null}$, can easily be obtained, which is then used to form

$$\mathbf{P}_\Sigma \triangleq \left(\mathbf{I}_{L+P} \otimes \mathbf{U}_{0null}^H\right) \mathbf{P}_Q. \tag{3.11}$$

From (3.10) and (3.11), we have

$$\mathbf{P}_{\Sigma}\mathbf{H}_{F} = \mathbf{0}, \tag{3.12}$$

which is equivalent to

$$(\mathbf{I} \otimes \mathbf{P}_{\Sigma})\mathrm{vec}\,(\mathbf{H}_{F}) = \mathbf{0}. \tag{3.13}$$

Noting that $\mathrm{vec}\,(\mathbf{H}_{F}) = \mathbf{E}_{P}\mathbf{h}$, where $\mathbf{E}_{P}$ is a known permutation matrix, (3.13) can be rewritten as

$$(\mathbf{I} \otimes \mathbf{P}_{\Sigma})\,\mathbf{E}_{P}\mathbf{h} = \mathbf{B}\mathbf{h} = \mathbf{0}, \tag{3.14}$$

implying that $\mathbf{B} = (\mathbf{I} \otimes \mathbf{P}_{\Sigma})\,\mathbf{E}_{P}$ is a blind constraint for the channel vector $\mathbf{h}$.

In the computation of the linear predictor $\mathbf{P}_{P}$ and the covariance matrix $\boldsymbol{\sigma}^{2}_{\tilde{\mathbf{y}},P}$, one has to estimate various correlation matrices $\tilde{\mathbf{R}}_{n-1}$, $\ddot{\mathbf{R}}_{n}$ and $\mathbf{R}\,(0)$, as discussed in Section 2.4. Considering that the circular convolution is used in MIMO-OFDM systems, a more accurate estimate of these correlation matrices can be obtained in comparision to that in MIMO systems. For example, the estimate of $\tilde{\mathbf{R}}_{n-1}$ in MIMO systems is computed as

$$\hat{\tilde{\mathbf{R}}}_{n-1} = \frac{1}{K}\sum_{n=P-1}^{K-1}\mathbf{y}_{P}\,(n)\,\mathbf{y}_{P}^{H}\,(n) - \sigma_{v}^{2}\mathbf{I} \tag{3.15}$$

where only $K - P + 1$ received signal vectors $\mathbf{y}_{P}\,(n)$ $(n = P-1, \cdots, K-1)$, are available for estimation. In MIMO-OFDM systems, however, the estimation of $\tilde{\mathbf{R}}_{n-1}$ can be modified as

$$\hat{\tilde{\mathbf{R}}}_{n-1} = \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{y}_{P}\,(n)\,\mathbf{y}_{P}^{H}\,(n) - \sigma_{v}^{2}\mathbf{I} \tag{3.16}$$

where $\mathbf{y}_{P}\,(n)$, $(n = 0, \cdots, P-2)$, can be obtained using $\mathbf{y}\,(n-i) \triangleq \mathbf{y}\,(n-i+K)$ when $n < i$ due to the circular convolution. Since more signal samples are used in the

estimate, a better linear prediction result can be expected in MIMO-OFDM systems. Note that, when multiple OFDM symbols are used, $\hat{\bar{\mathbf{R}}}_{n-1}$ can be easily calculated by averaging the results obtained from each OFDM symbol using (3.16).

3.2.3 The Semi-Blind Solution

Combining (3.3) and (3.14), a semi-blind cost function for the estimation of the channel vector $\mathbf{h}$ can be formulated as

$$\min_{\mathbf{h}} \Delta = \left\| \mathbf{Y}_{\text{pilot}} - \tilde{\mathbf{A}}\mathbf{h} \right\|^2 + \alpha \left\| \hat{\mathbf{B}}\mathbf{h} \right\|^2 \tag{3.17}$$

where $\hat{\mathbf{B}}$ is an estimate of the blind constraint. The solution to this minimization problem can be obtained by letting

$$\frac{\partial \Delta}{\partial \mathbf{h}^H} = -\tilde{\mathbf{A}}^H \left(\mathbf{Y}_{\text{pilot}} - \tilde{\mathbf{A}}\mathbf{h} \right) + \alpha \hat{\mathbf{B}}^H \hat{\mathbf{B}}\mathbf{h} = \mathbf{0},$$

which gives

$$\hat{\mathbf{h}} = \left(\tilde{\mathbf{A}}^H \tilde{\mathbf{A}} + \alpha \hat{\mathbf{B}}^H \hat{\mathbf{B}} \right)^{\dagger} \tilde{\mathbf{A}}^H \mathbf{Y}_{\text{pilot}}. \tag{3.18}$$

Clearly, the performance of the semi-blind algorithm depends on the choice of α. However, the selection of α has not yet been discussed in the existing semi-blind MIMO channel estimation methods such as those in [17, 42, 75, 84]. Here, we propose an explicit formula for the calculation of α in terms of the estimated MSE of the training based LS criterion and that of the blind part.

It has been shown in [85, 86] that the weight for a weighted least-squares (WLS) minimization problem can be determined according to the variance of the individual estimation error involved provided that the error is Gaussian distributed. A vector

version of the WLS problem can be described as

$$S = \sum_{i=1}^{N} w_i \|\mathbf{e}_i\|^2 = \sum_{i=1}^{N} w_i \|\hat{\mathbf{x}}_i - \mathbf{x}_i\|^2 \tag{3.19}$$

where $\hat{\mathbf{x}}_i$ denotes the estimate of the true vector $\mathbf{x}_i$ and w_i the weight for the i-th error term, which should be chosen as the reciprocal of the variance σ_i^2 of $\mathbf{e}_i$, i.e.

$$w_i = \frac{1}{\sigma_i^2} = \frac{1}{E\left\{\|\mathbf{e}_i\|^2\right\}}.$$

In order to obtain a closed-form expression for the weighting factor α in (3.17), let us consider the following WLS problem,

$$\min\Delta = w_{\mathrm{T}} \|\mathbf{e}_{\mathrm{T}}\|^2 + w_{\mathrm{B}} \|\mathbf{e}_{\mathrm{B}}\|^2 \tag{3.20}$$

where

$$\mathbf{e}_{\mathrm{T}} \triangleq \hat{\mathbf{h}}_{\mathrm{T}} - \mathbf{h},$$

$$\mathbf{e}_{\mathrm{B}} \triangleq \hat{\mathbf{h}}_{\mathrm{B}} - \mathbf{h},$$

with $\hat{\mathbf{h}}_{\mathrm{T}}$ being the estimate of the channel vector resulting from the training based LS criterion and $\hat{\mathbf{h}}_{\mathrm{B}}$ that from the blind estimation part. Since the error vectors $\mathbf{e}_{\mathrm{T}}$ and $\mathbf{e}_{\mathrm{B}}$ can be considered Gaussian, the coefficients w_{T} and w_{B} can be determined from the variance of $\mathbf{e}_{\mathrm{T}}$ and that of $\mathbf{e}_{\mathrm{B}}$, respectively, thereby the value of α can be estimated.

We first consider the variance of $\mathbf{e}_{\mathrm{T}}$. From [25], one can easily find that the optimal pilots are given by

$$\tilde{\mathbf{A}}^H \tilde{\mathbf{A}} = \left(gK_p\sigma_x^2\right) \mathbf{I}_{LN_RN_T}.$$

Thus, for the optimal pilots, one can obtain

$$\left\|\tilde{\mathbf{A}}\right\|_F^2 = gK_pLN_RN_T\sigma_x^2, \tag{3.21}$$

$$\mathrm{MSE_T} \triangleq \mathrm{E}\left\{\|\mathbf{e}_\mathrm{T}\|^2\right\} = \mathrm{E}\left\{\left\|\hat{\mathbf{h}}_\mathrm{T} - \mathbf{h}\right\|^2\right\} = \frac{N_R N_T L \sigma_v^2}{g K_p \sigma_x^2}. \tag{3.22}$$

Note that it is not possible to directly compute the variance of $\mathbf{e}_\mathrm{B}$, since the blind solution $\hat{\mathbf{h}}_\mathrm{B}$ is not available in the proposed method. However, we may use the following variance to replace the variance of $\mathbf{e}_\mathrm{B}$,

$$\mathrm{MSE_B} \triangleq \mathrm{E}\left\{\left\|\hat{\mathbf{B}}\left(\hat{\mathbf{h}}_\mathrm{B} - \mathbf{h}\right)\right\|^2\right\} = \mathrm{E}\left\{\left\|\hat{\mathbf{B}}\mathbf{h}\right\|^2\right\}. \tag{3.23}$$

It will be shown through simulation study in Section 3.4 that this approximation gives an appropriate choice of α in terms of the channel estimation performance.

Using (3.22) and (3.23), (3.20) can be rewritten as [85,86]

$$\min_{\mathbf{h}} \Delta = \frac{1}{\mathrm{MSE_T}}\left\|\hat{\mathbf{h}}_\mathrm{T} - \mathbf{h}\right\|^2 + \frac{1}{\mathrm{MSE_B}}\left\|\hat{\mathbf{h}}_\mathrm{B} - \mathbf{h}\right\|^2. \tag{3.24}$$

The above minimization problem can be nearly reformulated as the following least-square problem

$$\min_{\mathbf{h}} \Delta = \frac{\left\|\mathbf{Y}_\mathrm{pilot} - \tilde{\mathbf{A}}\mathbf{h}\right\|^2}{\left\|\tilde{\mathbf{A}}\right\|_F^2 \mathrm{MSE_T}} + \frac{\left\|\hat{\mathbf{B}}\mathbf{h}\right\|^2}{\left\|\hat{\mathbf{B}}\right\|_F^2 \mathrm{MSE_B}}. \tag{3.25}$$

Evidently, the solution of (3.25) is equivalent to that of (3.17) when α is chosen as

$$\alpha = \frac{\mathrm{MSE_T}}{\mathrm{MSE_B}} \frac{\left\|\tilde{\mathbf{A}}\right\|_F^2}{\left\|\hat{\mathbf{B}}\right\|_F^2}. \tag{3.26}$$

What remains in the computation of α is to determine the $\mathrm{MSE_B}$. In the next section, a perturbation analysis of the linear prediction-based blind algorithm is performed, leading to a novel closed-form expression for $\mathrm{MSE_B}$.

3.3 Perturbation Analysis of Linear Prediction-based Blind Channel Estimation

It is known that the solution of linear prediction or subspace-based methods is always perturbed by various sources, such as finite data length, measurement noise, etc [71–74]. Perturbation theory has been successfully applied to the analysis of subspace-based methods [72, 75–77]. In this section, the first-order perturbation theory is employed to analyze the MIMO linear prediction as well as the resulting blind constraint.

3.3.1 The MIMO Linear Prediction with Perturbation

Most of the LP-based blind and semi-blind methods are based on the two key equations (2.41) and (2.42). It is important to perform the perturbation analysis of these two equations in order to evaluate the linear prediction-based methods. Letting

$$\mathbf{H}_A \triangleq [\mathbf{H}(0), \mathbf{H}(1), \cdots, \mathbf{H}(L-1)],$$

$$\mathbf{x}_L(n) \triangleq \left[\mathbf{x}^T(n) \cdots \mathbf{x}^T(n-L+1)\right]^T, \text{ for } n = 0, 1, \cdots, K-1$$

where $\mathbf{x}(n) \triangleq [x_1(n), \cdots, x_{N_T}(n)]^T$ and noting that $\mathbf{x}(n) = \mathbf{x}(K+n)$ for $n < 0$, the circular convolution (2.3) can be rewritten in the matrix form as

$$\mathbf{y}(n) \triangleq [y_1(n), \cdots, y_{N_R}(n)]^T = \mathbf{H}_A \mathbf{x}_L(n) + \mathbf{v}(n). \tag{3.27}$$

Using (3.27), we can obtain

$$\begin{aligned} \mathbf{y}_P(n-1) &\triangleq \left[\mathbf{y}^T(n-1), \cdots, \mathbf{y}^T(n-P)\right]^T \\ &= \mathbf{H}_B \mathbf{x}_{P+L-1}(n-1) + \mathbf{v}_P(n-1) \end{aligned} \tag{3.28}$$

where $\mathbf{v}_P(n-1) \triangleq \left[\mathbf{v}^T(n-1)\cdots\mathbf{v}^T(n-P)\right]^T$, $\mathbf{H}_B$ is a $PN_R \times (P+L-1)N_T$ block Toeplitz matrix with the first block row given by $[\mathbf{H}(0),\cdots,\mathbf{H}(L-1),\mathbf{0},\cdots,\mathbf{0}]$. Here, we consider only the perturbation due to the finite data length in the computation of the correlation matrices.

We now derive the perturbation form of the correlation matrices involved in the linear prediction in order to disclose the perturbation in the channel constraint in (3.7), the time-domain version of (2.41). Without loss of generality, we let the variance of the signal be unity, i.e., $\sigma_x^2 = 1$. Using (3.28), the correlation matrix $\hat{\tilde{\mathbf{R}}}_{n-1}$ with such a perturbation can be written as

$$\hat{\tilde{\mathbf{R}}}_{n-1} = \hat{\mathrm{E}}\left[\mathbf{y}_P(n-1)\,\mathbf{y}_P^H(n-1)\right] - \sigma_v^2\mathbf{I} = \mathbf{H}_B\left[\mathbf{I} + \Delta\mathbf{R}_{x1}\right]\mathbf{H}_B^H + \Delta\mathbf{R}_{v1} \tag{3.29}$$

where $\Delta\mathbf{R}_{x1}$ denotes the signal perturbation matrix,

$$\Delta\mathbf{R}_{x1} = \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_{P+L-1}(n-1)\,\mathbf{x}_{P+L-1}^H(n-1) - \mathbf{I},$$

and $\Delta\mathbf{R}_{v1}$ the perturbation matrix introduced by the noise,

$$\Delta\mathbf{R}_{v1} \triangleq \mathbf{H}_B\Delta\mathbf{R}_{xv1} + \Delta\mathbf{R}_{xv1}^H\mathbf{H}_B^H + \Delta\mathbf{R}_{vv1}, \tag{3.30}$$

with

$$\Delta\mathbf{R}_{xv1} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_{P+L-1}(n-1)\,\mathbf{v}_P^H(n-1), \tag{3.31}$$

$$\Delta\mathbf{R}_{vv1} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{v}_P(n-1)\,\mathbf{v}_P^H(n-1) - \sigma_v^2\mathbf{I}.$$

Similarly, $\hat{\ddot{\mathbf{R}}}_n$ is given by

$$\hat{\ddot{\mathbf{R}}}_n = \hat{\mathrm{E}}\left[\mathbf{y}(n)\,\mathbf{y}_P^H(n-1)\right] - \sigma_v^2\left[\mathbf{I},\mathbf{0},\cdots,\mathbf{0}\right] = \mathbf{H}_A\left[\mathbf{I}_C + \Delta\mathbf{R}_{x2}\right]\mathbf{H}_B^H + \Delta\mathbf{R}_{v2} \tag{3.32}$$

where

$$\mathbf{I}_C \triangleq \begin{bmatrix} \mathbf{0}_{N_T\times(L-1)N_T} & \mathbf{0}_{N_T\times PN_T} \\ \mathbf{I}_{(L-1)N_T\times(L-1)N_T} & \mathbf{0}_{(L-1)N_T\times PN_T} \end{bmatrix},$$

$$\Delta\mathbf{R}_{x2} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_L(n)\,\mathbf{x}_{P+L-1}^H(n-1) - \mathbf{I}_C,$$

$$\Delta\mathbf{R}_{v2} \triangleq \mathbf{H}_A\Delta\mathbf{R}_{xv21} + \Delta\mathbf{R}_{xv22}^H\mathbf{H}_B^H + \Delta\mathbf{R}_{vv2},$$

with

$$\Delta\mathbf{R}_{xv21} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_L(n)\,\mathbf{v}_P^H(n-1),$$

$$\Delta\mathbf{R}_{xv22} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_{P+L-1}(n-1)\,\mathbf{v}^H(n),$$

$$\Delta\mathbf{R}_{vv2} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{v}(n)\,\mathbf{v}_P^H(n-1) - \sigma_v^2[\mathbf{I},\mathbf{0},\cdots,\mathbf{0}].$$

Based on the above expressions for the correlation matrices with perturbation, we now derive the perturbation form of the linear prediction in (3.7). To this end, we first derive the perturbation form $\hat{\mathbf{P}}_P$ of the linear predictor $\mathbf{P}_P$. Using the first-order approximation, the pseudo-inverse of $\hat{\bar{\mathbf{R}}}_{n-1}$ given by (3.29) can be obtained as

$$\hat{\bar{\mathbf{R}}}_{n-1}^{\dagger} \approx \mathbf{\Pi}_1 - \mathbf{\Pi}_2\Delta\mathbf{R}_{x1}\mathbf{\Pi}_2^H - \mathbf{\Pi}_1\Delta\mathbf{R}_{v1}\mathbf{\Pi}_1^H \tag{3.33}$$

where

$$\mathbf{\Pi}_1 \triangleq \left(\mathbf{H}_B\mathbf{H}_B^H\right)^{\dagger},$$

$$\mathbf{\Pi}_2 \triangleq \mathbf{\Pi}_1\mathbf{H}_B.$$

It should be noted that the noise subspace of $\hat{\bar{\mathbf{R}}}_{n-1}$ in (3.29) has been omitted in obtaining (3.33), and such an omission would not affect the derivation of the perturbation form of (3.7) and (2.40). Using (3.32) and (3.33), the linear predictor with

perturbation can be derived as follows

$$\hat{\mathbf{P}}_P \approx \mathbf{H}_A \left[\mathbf{I}_C + \Delta\mathbf{R}_{x2}\right] \mathbf{\Pi}_2^H - \mathbf{H}_A \mathbf{I}_C \mathbf{\Pi}_2^H \mathbf{H}_B \Delta\mathbf{R}_{x1} \mathbf{\Pi}_2^H \tag{3.34}$$

$$- \mathbf{H}_A \mathbf{I}_C \mathbf{\Pi}_2^H \Delta\mathbf{R}_{v1} \mathbf{\Pi}_1 + \Delta\mathbf{R}_{v2} \mathbf{\Pi}_1.$$

On the other hand, noting that $\mathbf{H}_D$ can be partitioned as

$$\mathbf{H}_D \triangleq \begin{bmatrix} \mathbf{H}(0) & \mathbf{H}_E \\ \mathbf{0} & \mathbf{H}_B \end{bmatrix}$$

where $\mathbf{H}_E \triangleq [\mathbf{H}(1), \cdots, \mathbf{H}(L-1), \mathbf{0}, \cdots, \mathbf{0}]$, the perturbation form of the left hand side (LHS) of (3.7) can be written as

$$\left[\mathbf{I}, -\hat{\mathbf{P}}_P\right] \mathbf{H}_D = \left[\mathbf{H}(0), \mathbf{H}_E\right] - \left[\mathbf{0}, \hat{\mathbf{P}}_P \mathbf{H}_B\right]. \tag{3.35}$$

Using (3.34) and $\mathbf{H}_E = \mathbf{H}_A \mathbf{I}_C$, and noting that $\mathbf{H}_B^H \left(\mathbf{H}_B \mathbf{H}_B^H\right)^\dagger \mathbf{H}_B = \mathbf{I}$ when $\mathbf{H}_B$ has a full column rank and $PN_R \geq (P + L - 1) N_T$, one can deduce

$$\hat{\mathbf{P}}_P \mathbf{H}_B = \mathbf{H}_E + \mathbf{\Pi}_P \tag{3.36}$$

where

$$\mathbf{\Pi}_P \triangleq \mathbf{H}_A \Delta\mathbf{R}_{x2} - \mathbf{H}_A \mathbf{I}_C \Delta\mathbf{R}_{x1} + \Delta\mathbf{R}_{v2} \mathbf{\Pi}_2 - \mathbf{H}_A \mathbf{I}_C \mathbf{\Pi}_2^H \Delta\mathbf{R}_{v1} \mathbf{\Pi}_2 \tag{3.37}$$

respresents the perturbation terms due to both signal and noise. By noting that

$$\Delta\mathbf{R}_{x2} - \mathbf{I}_C \Delta\mathbf{R}_{x1} = \frac{1}{K} \sum_{n=0}^{K-1} \begin{bmatrix} \mathbf{x}(n) \\ \mathbf{0}_{(L-1)N_T \times 1} \end{bmatrix} \mathbf{x}_{P+L-1}^H (n-1),$$

$\mathbf{\Pi}_P$ can be expressed as

$$\mathbf{\Pi}_P = \mathbf{H}(0) \Delta\mathbf{R}_{x4} + \mathbf{\Pi}'_P \tag{3.38}$$

where

$$\Delta \mathbf{R}_{x4} \triangleq \frac{1}{K} \sum_{n=0}^{K-1} \mathbf{x}(n) \mathbf{x}_{P+L-1}^{H}(n-1), \tag{3.39}$$

$$\mathbf{\Pi}_P' \triangleq \Delta \mathbf{R}_{v2} \mathbf{\Pi}_2 - \mathbf{H}_A \mathbf{I}_C \mathbf{\Pi}_2^H \Delta \mathbf{R}_{v1} \mathbf{\Pi}_2. \tag{3.40}$$

It is clear from (3.38) that the perturbation term $\mathbf{\Pi}_P$ consists of the signal perturbation $\mathbf{H}(0) \Delta \mathbf{R}_{x4}$ and the perturbation introduced by the noise, $\mathbf{\Pi}_P'$. Using (3.36) in (3.35) yields

$$\left[\mathbf{I}, -\hat{\mathbf{P}}_P\right] \mathbf{H}_D = [\mathbf{H}(0), \mathbf{0}, \cdots, \mathbf{0}] - [\mathbf{0}, \mathbf{\Pi}_P]. \tag{3.41}$$

Obviously, (3.41) is the perturbation form of (3.7).

3.3.2 The Covariance Matrix of Prediction Error

In this subsection, we derive the perturbation form of the covariance matrix in (2.42). We now turn to the derivation of the perturbed version of (2.42). The perturbed form of (2.40) can be written as

$$\hat{\boldsymbol{\sigma}}_{\tilde{\mathbf{y}},P}^2 = \hat{\mathbf{R}}(0) - \hat{\mathbf{P}}_P \hat{\ddot{\mathbf{R}}}_n^H \tag{3.42}$$

where

$$\hat{\mathbf{R}}(0) = \hat{\mathrm{E}}\left[\mathbf{y}(n) \mathbf{y}^H(n)\right] - \sigma_v^2 \mathbf{I}. \tag{3.43}$$

It is easy to show that

$$\hat{\mathbf{R}}(0) = \mathbf{H}_A \left[\mathbf{I} + \Delta \mathbf{R}_{x3}\right] \mathbf{H}_A^H + \Delta \mathbf{R}_{v3} \tag{3.44}$$

where

$$\Delta \mathbf{R}_{x3} \triangleq \frac{1}{K} \sum_{n=0}^{K-1} \mathbf{x}_L(n) \mathbf{x}_L^H(n) - \mathbf{I},$$

$$\Delta\mathbf{R}_{v3} \triangleq \mathbf{H}_A\Delta\mathbf{R}_{xv3} + \Delta\mathbf{R}_{xv3}^H\mathbf{H}_A^H + \Delta\mathbf{R}_{vv3},$$

with

$$\Delta\mathbf{R}_{xv3} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_L(n)\,\mathbf{v}^H(n),$$

$$\Delta\mathbf{R}_{vv3} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{v}(n)\mathbf{v}^H(n) - \sigma_v^2\mathbf{I}.$$

Using (3.34), the second term of the RHS of (3.42) can be expressed as

$$\begin{aligned}\hat{\mathbf{P}}_P\hat{\hat{\mathbf{R}}}_n^H &= \mathbf{H}_A\mathbf{I}_C\mathbf{I}_C^H\mathbf{H}_A^H + \mathbf{H}_A\Delta\mathbf{R}_{x2}\mathbf{I}_C^H\mathbf{H}_A^H + \mathbf{H}_A\mathbf{I}_C\Delta\mathbf{R}_{x2}^H\mathbf{H}_A^H + \mathbf{H}_A\mathbf{I}_C\mathbf{\Pi}_2^H\Delta\mathbf{R}_{v2}^H \\ &+ \Delta\mathbf{R}_{v2}\mathbf{\Pi}_2\mathbf{I}_C^H\mathbf{H}_A^H - \mathbf{H}_A\mathbf{I}_C\Delta\mathbf{R}_{x1}\mathbf{I}_C^H\mathbf{H}_A^H - \mathbf{H}_A\mathbf{I}_C\mathbf{\Pi}_2^H\Delta\mathbf{R}_{v1}\mathbf{\Pi}_2\mathbf{I}_C^H\mathbf{H}_A^H \quad (3.45)\end{aligned}$$

From

$$\mathbf{I}_C\mathbf{I}_C^H = \begin{bmatrix} \mathbf{0}_{N_T\times N_T} & \mathbf{0}_{N_T\times(L-1)N_T} \\ \mathbf{0}_{(L-1)N_T\times N_T} & \mathbf{I}_{(L-1)N_T\times(L-1)N_T} \end{bmatrix},$$

one can get

$$\mathbf{H}_A\left[\mathbf{I} - \mathbf{I}_C\mathbf{I}_C^H\right]\mathbf{H}_A^H = \mathbf{H}(0)\,\mathbf{H}^H(0). \quad (3.46)$$

Using (3.44) and (3.45) along with (3.46), the covariance matrix given by (2.42) has the following perturbation form,

$$\hat{\boldsymbol{\sigma}}_{\hat{\mathbf{y}},P}^2 = \mathbf{H}(0)\,\mathbf{H}^H(0) + \Delta\hat{\boldsymbol{\sigma}}_{\hat{\mathbf{y}},P}^2 \quad (3.47)$$

where

$$\begin{aligned}\Delta\hat{\boldsymbol{\sigma}}_{\hat{\mathbf{y}},P}^2 &\triangleq \mathbf{H}_A\Delta\mathbf{R}_{x3}\mathbf{H}_A^H - \mathbf{H}_A\Delta\mathbf{R}_{x2}\mathbf{I}_C^H\mathbf{H}_A^H - \mathbf{H}_A\mathbf{I}_C\Delta\mathbf{R}_{x2}^H\mathbf{H}_A^H + \mathbf{H}_A\mathbf{I}_C\Delta\mathbf{R}_{x1}\mathbf{I}_C^H\mathbf{H}_A^H \\ &+ \Delta\mathbf{R}_{v3} - \mathbf{H}_A\mathbf{I}_C\mathbf{\Pi}_2^H\Delta\mathbf{R}_{v2}^H - \Delta\mathbf{R}_{v2}\mathbf{\Pi}_2\mathbf{I}_C^H\mathbf{H}_A^H + \mathbf{H}_A\mathbf{I}_C\mathbf{\Pi}_2^H\Delta\mathbf{R}_{v1}\mathbf{\Pi}_2\mathbf{I}_C^H\mathbf{H}_A^H. \quad (3.48)\end{aligned}$$

We now show that the perturbation term $\Delta\hat{\boldsymbol{\sigma}}_{\hat{\mathbf{y}},P}^2$ can be split into the signal perturbation and the perturbation due to the noise. By using

$$\Delta\mathbf{R}_{x3} - \Delta\mathbf{R}_{x2}\mathbf{I}_C^H = \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_L(n)\left[\mathbf{x}^H(n), \mathbf{0}_{1\times(L-1)N_T}\right] - \left(\mathbf{I} - \mathbf{I}_C\mathbf{I}_C^H\right),$$

$$\mathbf{I}_C \Delta\mathbf{R}_{x1}\mathbf{I}_C^H - \mathbf{I}_C \Delta\mathbf{R}_{x2} = -\frac{1}{K}\sum_{n=0}^{K-1}\begin{bmatrix} \mathbf{0}(n) \\ \mathbf{x}_{L-1}(n-1) \end{bmatrix}\left[\mathbf{x}^H(n), \mathbf{0}_{1\times(L-1)N_T}\right],$$

it is easy to verify that the sum of the first four terms in the RHS of (3.48) equals $\mathbf{H}(0)\Delta\mathbf{R}_{x5}\mathbf{H}^H(0)$, where

$$\Delta\mathbf{R}_{x5} \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}(n)\mathbf{x}^H(n) - \mathbf{I}. \tag{3.49}$$

Therefore, (3.48) can be rewritten as

$$\Delta\hat{\boldsymbol{\sigma}}^2_{\tilde{\mathbf{y}},P} = \mathbf{H}(0)\Delta\mathbf{R}_{x5}\mathbf{H}^H(0) + \boldsymbol{\Xi} \tag{3.50}$$

where

$$\boldsymbol{\Xi} \triangleq \Delta\mathbf{R}_{v3} - \mathbf{H}_A\mathbf{I}_C\boldsymbol{\Pi}_2^H\Delta\mathbf{R}_{v2}^H - \Delta\mathbf{R}_{v2}\boldsymbol{\Pi}_2\mathbf{I}_C^H\mathbf{H}_A^H + \mathbf{H}_A\mathbf{I}_C\boldsymbol{\Pi}_2^H\Delta\mathbf{R}_{v1}\boldsymbol{\Pi}_2\mathbf{I}_C^H\mathbf{H}_A^H \tag{3.51}$$

represents the perturbation error caused by the noise.

Thus far, we have obtained the perturbation form of the two key equations in the linear prediction. Based on this result, we would like to point out the advantage of the proposed semi-blind method over the conventional linear prediction-based blind methods. Note that in the noise-free case, all the perturbation terms caused by the noise in the above discussions can be ignored. As such, (3.41) and (3.47) can be simplified as

$$\left[\mathbf{I}, -\hat{\mathbf{P}}_P\right]\mathbf{H}_D = \left[\mathbf{H}(0), \mathbf{0}, \cdots, \mathbf{0}\right] - \left[\mathbf{0}, \mathbf{H}(0)\Delta\mathbf{R}_{x4}\right], \tag{3.52}$$

$$\hat{\boldsymbol{\sigma}}^2_{\tilde{\mathbf{y}},P} = \mathbf{H}(0)\mathbf{H}^H(0) + \mathbf{H}(0)\Delta\mathbf{R}_{x5}\mathbf{H}^H(0). \tag{3.53}$$

In conventional blind algorithms [39, 79, 80], an estimate of $\mathbf{H}(0)$ is first acquired from $\hat{\boldsymbol{\sigma}}^2_{\tilde{\mathbf{y}},P}$, and is then used to estimate the channel matrix $\mathbf{H}_D$. This approach

would give rise to a large estimation error due to the presence of signal perturbation terms as seen from (3.52) and (3.53). In the proposed semi-blind approach, however, an ideal nulling constraint on the channel matrix $\mathbf{H}_D$ has been obtained from the ideal null space of $\mathbf{H}(0)$, which is not affected by the signal perturbation terms. Therefore, the semi-blind method using MIMO linear prediction is superior to the blind algorithms [39, 79, 80]. It should also be pointed out that in the presence of noise, although both the semi-blind and the blind methods are subject to the noise perturbation terms, the semi-blind method still outperforms the blind one, since the perturbation introduced by the noise is in general significantly smaller than the signal perturbation.

3.3.3 The MSE of Blind Channel Estimate

To derive the MSE expression, a perturbed version of (3.12) needs to be determined. By partitioning $\mathbf{\Pi}_P$ as $\mathbf{\Pi}_P = [\mathbf{\Pi}_{P,1}, \cdots, \mathbf{\Pi}_{P,P+L-1}]$, where $\mathbf{\Pi}_{P,i}$, $(i = 1, 2, \cdots, P + L - 1)$ is an $N_R \times N_T$ matrix and using (3.41), we can derive the perturbation form of (3.10),

$$\hat{\mathbf{P}}_Q \mathbf{H}_F = \left[\mathbf{H}^H(0), \mathbf{0}, \cdots, \mathbf{0}\right]^H - \mathbf{Q} \tag{3.54}$$

where

$$\mathbf{Q} \triangleq \left[\mathbf{0}, \mathbf{\Pi}_{P,1}^H, \cdots, \mathbf{\Pi}_{P,P+L-1}^H\right]^H. \tag{3.55}$$

On the other hand, using (3.47) along with the first-order approximation [77], one can get the estimate of $\mathbf{U}_{0null}$,

$$\hat{\mathbf{U}}_{0null} \approx \mathbf{U}_{0null} - \left(\mathbf{H}(0)\mathbf{H}^H(0)\right)^{\dagger} \hat{\boldsymbol{\sigma}}^2_{\tilde{\mathbf{y}},P} \mathbf{U}_{0null}.$$

Using (3.50) into the above equation gives

$$\hat{\mathbf{U}}_{0null} \approx \mathbf{U}_{0null} - \left(\mathbf{H}(0)\,\mathbf{H}^{H}(0)\right)^{\dagger}\Xi\mathbf{U}_{0null}. \tag{3.56}$$

Substituting (3.56) into (3.11) and using (3.54), we obtain

$$\hat{\mathbf{P}}_{\Sigma}\mathbf{H}_{F} = \mathbf{G}_1 - \mathbf{G}_2 \tag{3.57}$$

where

$$\mathbf{G}_1 \triangleq \left[\left(\mathbf{U}_{0null}^{H}\Xi^{H}\left(\mathbf{H}(0)\,\mathbf{H}^{H}(0)\right)^{\dagger}\mathbf{H}(0)\right)^{H}, \mathbf{0}, \cdots, \mathbf{0}\right]^{H}, \tag{3.58}$$

$$\mathbf{G}_2 \triangleq \left(\mathbf{I}_{P+L} \otimes \mathbf{U}_{0null}^{H}\right)\mathbf{Q}. \tag{3.59}$$

Clearly, (3.57) is the perturbed version of (3.12). Using (3.57) and noting that $\left\|\hat{\mathbf{B}}\mathbf{h}\right\|^2 = \left\|\hat{\mathbf{P}}_{\Sigma}\mathbf{H}_{F}\right\|_{F}^{2}$, the MSE of the blind criterion defined in (3.23) can be calculated as

$$\mathrm{MSE_B} = \mathrm{Trace}\left\{\mathrm{E}\left[\mathrm{vec}\left(\mathbf{G}_1 - \mathbf{G}_2\right)\mathrm{vec}^{H}\left(\mathbf{G}_1 - \mathbf{G}_2\right)\right]\right\}. \tag{3.60}$$

In Appendix A, we investigate the computation of $\mathrm{vec}\left(\mathbf{G}_1\right)$ and $\mathrm{vec}\left(\mathbf{G}_2\right)$ to obtain a closed-form expression for $\mathrm{MSE_B}$ in terms of the correlation of the transmitted signal, the correlation of the noise as well as the channel matrix.

3.4 Simulation Results

We consider a MIMO-OFDM system with 2 transmit and 4 receive antennas. The number of subcarriers is set to 512, the length of cyclic prefix is 10, and the length of the linear predictor is $P = 4$. In our simulation, the QPSK modulation is used and a Rayleigh channel modelled by a 3-tap MIMO-FIR filter is assumed, in which each tap corresponds to a 2×4 random matrix whose elements are i.i.d. complex

Gaussian variables with zero mean and unit variance. As the blind constraint of the proposed semi-blind algorithm is derived directly from $\mathbf{H}(0)$ as seen from (2.41) and (2.42), it is necessary to consider in the simulation study the effect of $\mathbf{H}(0)$ on the channel estimation performance. To this end, we define the following metric

$$\eta \triangleq \frac{\|\mathbf{H}(0)\|_F^2}{\sum_{n=0}^{2} \|\mathbf{H}(n)\|_F^2}$$

and conduct our investigation with respect to different ranges of η.

It is seen from (3.26) that one has to choose the value of $\mathrm{MSE_B}$ to determine the parameter α. Although a closed-form expression for $\mathrm{MSE_B}$ has been obtained in Section 3.3, it requires the true channel matrix, which is, however, unknown in practice. Therefore, we first estimate the channel matrix using the LS method and then utilize the preliminary estimate to compute the $\mathrm{MSE_B}$ and thereby the value of α. To evaluate the proposed algorithm, we also calculate the value of $\mathrm{MSE_B}$ based on the true channel matrix, which gives a reference for the choice of α. The semi-blind solution thus obtained is called the reference algorithm. The estimation performance is evaluated in terms of the MSE of the estimate of the channel matrix given by

$$\mathrm{MSE} = \frac{1}{N_{\mathrm{MC}}} \sum_{n=1}^{N_{\mathrm{MC}}} \left\| \hat{\mathbf{h}}_n - \mathbf{h}_n \right\|^2$$

where N_{MC} is the number of Monte Carlo iterations, and $\mathbf{h}_n$, $\hat{\mathbf{h}}_n$ are the true and the estimated channel vectors with respect to the n-th Monte Carlo iteration, respectively.

Experiment 1: *MSE versus weight factor α*

In the first experiment, the channel estimation performance in terms of the plot of the MSE versus α is investigated. The simulation is based on four randomly gen-

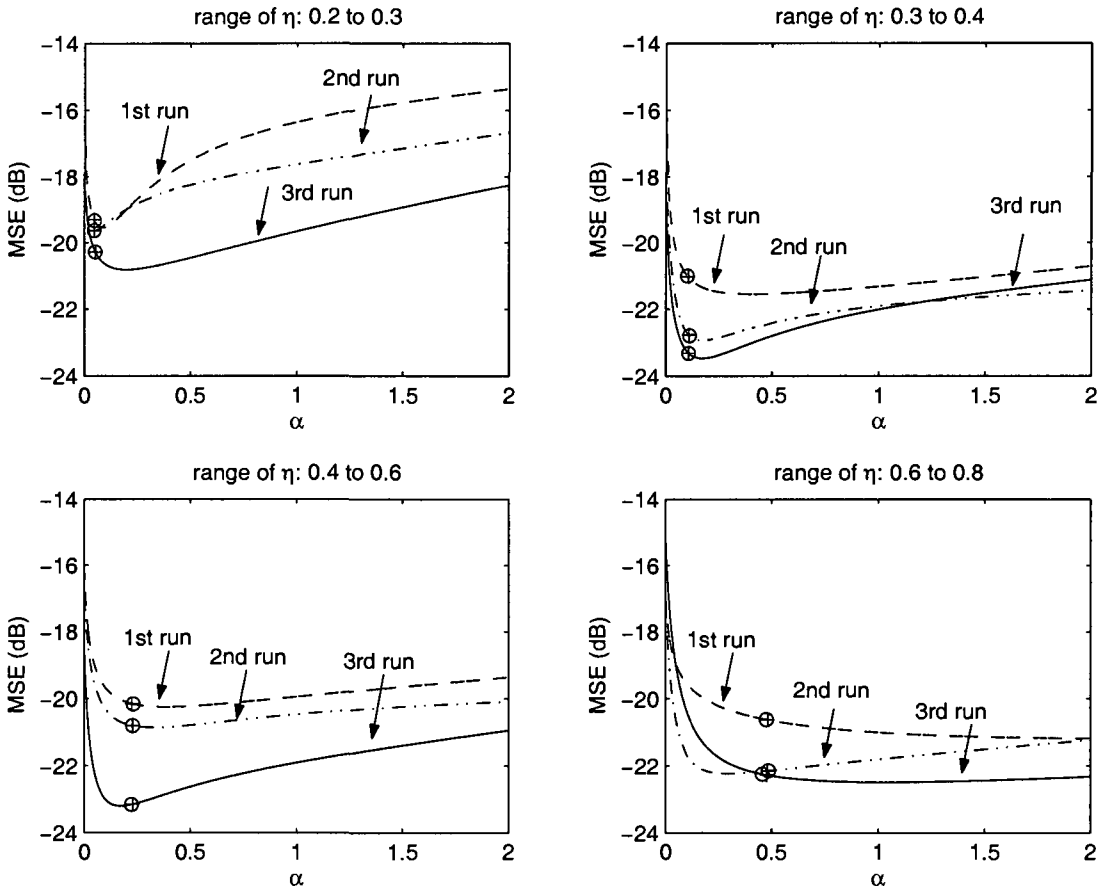

Figure 3.1: MSE versus α for different ranges of η in which the two points indicated by "+" and "o" for each run are obtained from the proposed and reference algorithms, respectively.

erated channel matrices, each having a value of η falling within the range of $[0.2, 0.3]$, $[0.3, 0.4]$, $[0.4, 0.6]$ or $[0.6, 0.8]$. For each channel matrix, three runs of the transmission of one OFDM symbol at 512 subcarriers, of which 32 are used as pilot for training purpose, are performed. Fig. 3.1 shows the MSE plots of the semi-blind algorithm with respect to different choices of α in the range of $[0, 2]$ for the four channels at an SNR of 15dB. The two points identified by "+" and "o" in each run indicates the results from the proposed and reference algorithms, respectively. It is observed that the proposed scheme for the determination of α gives a competitive MSE result. It is also seen that the use of the preliminary least-square estimate of the channel matrix, instead of the true channel characteristic, would suffice for the calculation of the value of α.

Experiment 2: *MSE versus* η

In this experiment, the channel estimation performance in terms of the MSE versus η is investigated. The simulation is undertaken based on 5000 Monte Carlo runs of the transmission of one OFDM symbol on 512 subcarriers at an SNR of 15dB. Fig. 3.2 shows the MSE plots resulting from the proposed as well as the reference algorithm along with that from the LS estimation, indicating a high consistency of the two semi-blind methods. It is noted that the proposed semi-blind algorithm significantly outperforms the LS method. It is also clear that the MSE performance of the semi-blind estimation depends on the value of η. A better estimation performance is reached when η varies from 0.3 to 0.7, which represents a typical mobile communication scenario, where the first arrived path is comparable to or stronger than other

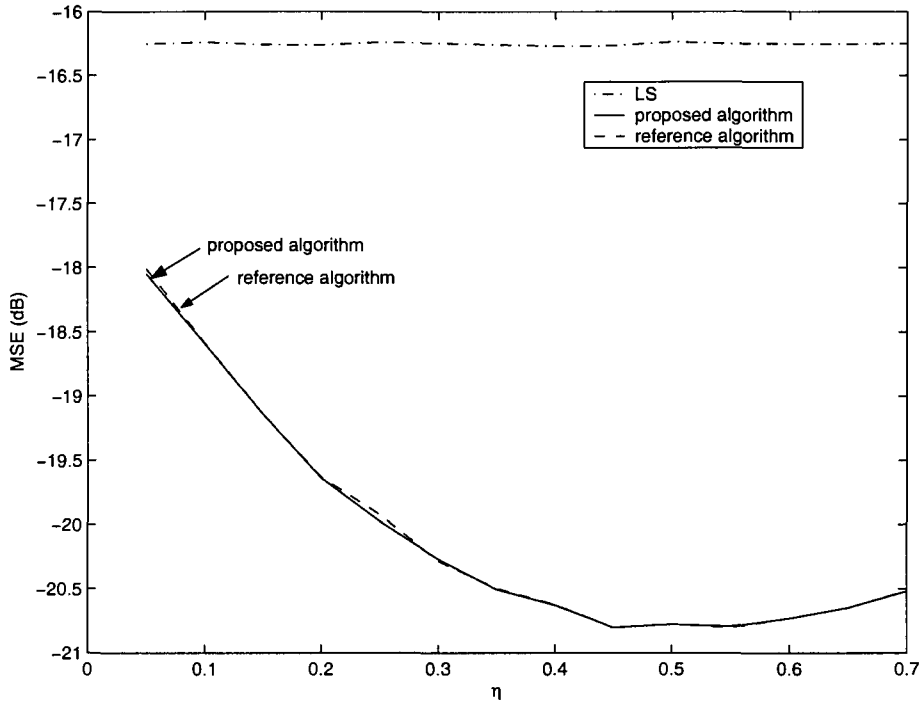

Figure 3.2: MSE versus the value of η for a 2×4 frequency-selective channel. One OFDM symbol with 32 pilot subcarriers and SNR=15 dB.

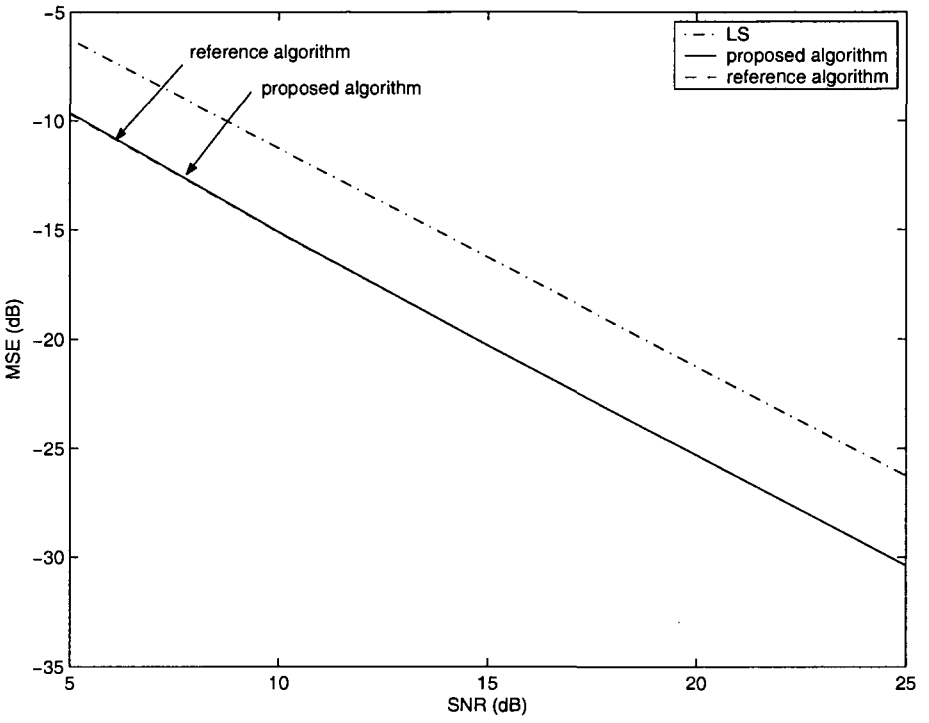

Figure 3.3: MSE versus SNR for a 2×4 frequency-selective channel. One OFDM symbol with 32 pilot subcarriers and $\eta > 0.2$.

paths [87]. From Fig. 3.2, one can find that, when η is larger than 0.55, the MSE becomes a little larger with the increase of η. We believe the possible reason is that the full column rank assumption of $\mathbf{H}_B$ for the proposed semi-blind algorithm cannot be always satisfied in the experiments, since the Rayleigh channel is generated for the Monte Carlo iterations. With the increase of the value of η, especially when η is larger than 0.55, $\mathbf{H}(1)$ and $\mathbf{H}(2)$ becomes more and more insignificant. As a result, the probability that $\mathbf{H}_B$ does not have a full column rank becomes larger, making the MSE increases slightly.

Experiment 3: *MSE versus SNR*

Now, we examine the channel estimation performance as a function of the SNR. Again, the simulation involves 5000 Monte Carlo runs of the transmission of one OFDM symbol. Fig. 3.3 shows the channel estimation results of the three methods, when $\eta > 0.2$. It is seen that the performances of the two semi-blind algorithms are very close and both can achieve a gain of nearly 3.3 dB over the LS method regardless of the level of the SNR.

Experiment 4: *MSE versus pilot length*

Here, we investigate the channel estimation performance of the proposed semi-blind algorithm versus the number of OFDM symbols as well as the number of pilot subcarriers per symbol, in comparison with that of the LS method. The number of OFDM symbols is set to be from 1 to 4, and in each case, the number of pilot subcarriers per OFDM symbol varies from 8 to 48. Note that the proposed method can easily be applied to the case of multiple OFDM symbols, where the calculation of $\mathbf{Y}_{\text{pilot}}$, $\tilde{\mathbf{A}}$ and $\hat{\mathbf{B}}$ in (3.18) should be based on the multiple OFDM symbols. Fig. 3.4 shows the MSE plots from 500 Monte Carlo iterations for an SNR of 15 dB, when $\eta > 0.2$. It is seen that for the same number of OFDM symbols, the performance of all the algorithms improved with increasing number of pilot subcarriers. Again, the performance of the proposed method is very close to that of the reference method, both being superior to the LS method by 6 dB and 4 dB when the number of pilot subcarriers is 8 and 48, respectively. It implies that the proposed semi-blind method is more advantageous for pilot signals of a shorter length. Furthermore, it is observed that the performance improvement of the proposed semi-blind method over the LS

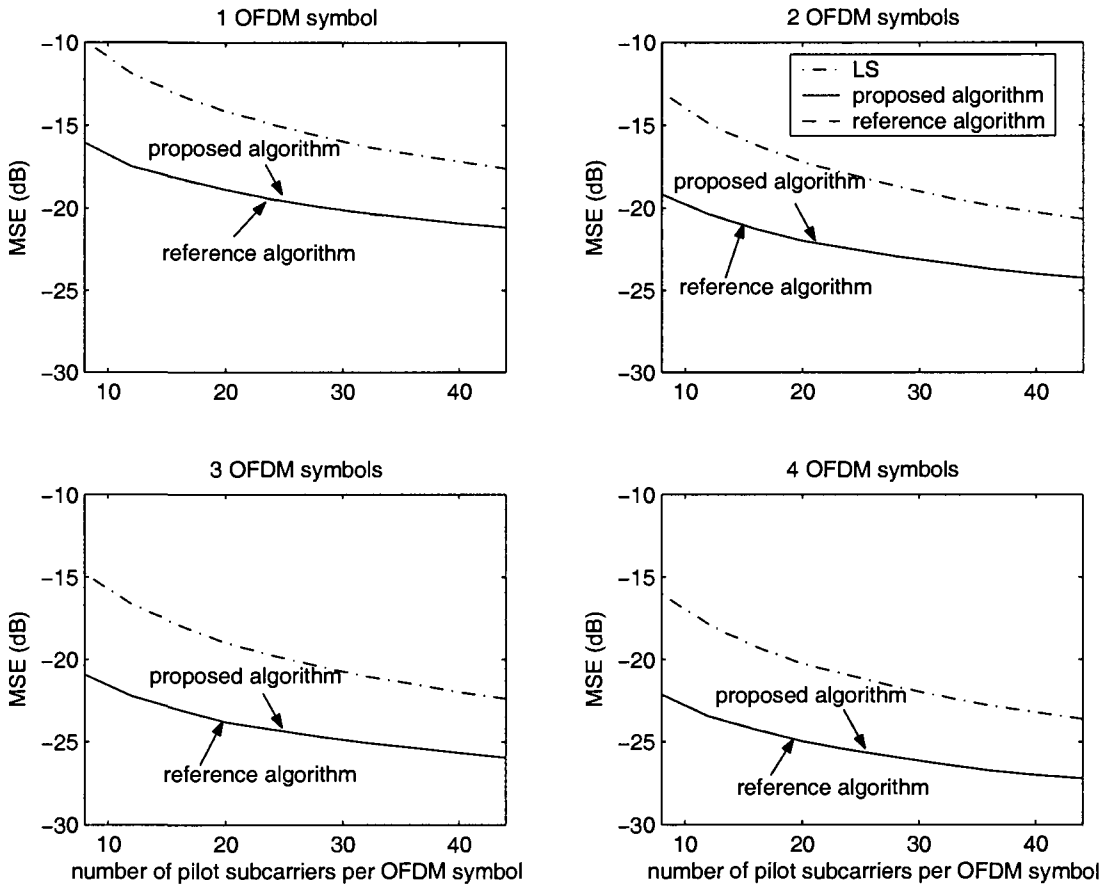

Figure 3.4: MSE versus pilot length for different numbers of OFDM symbols for a 2×4 frequency-selective channel. SNR=15 dB and $\eta > 0.2$.

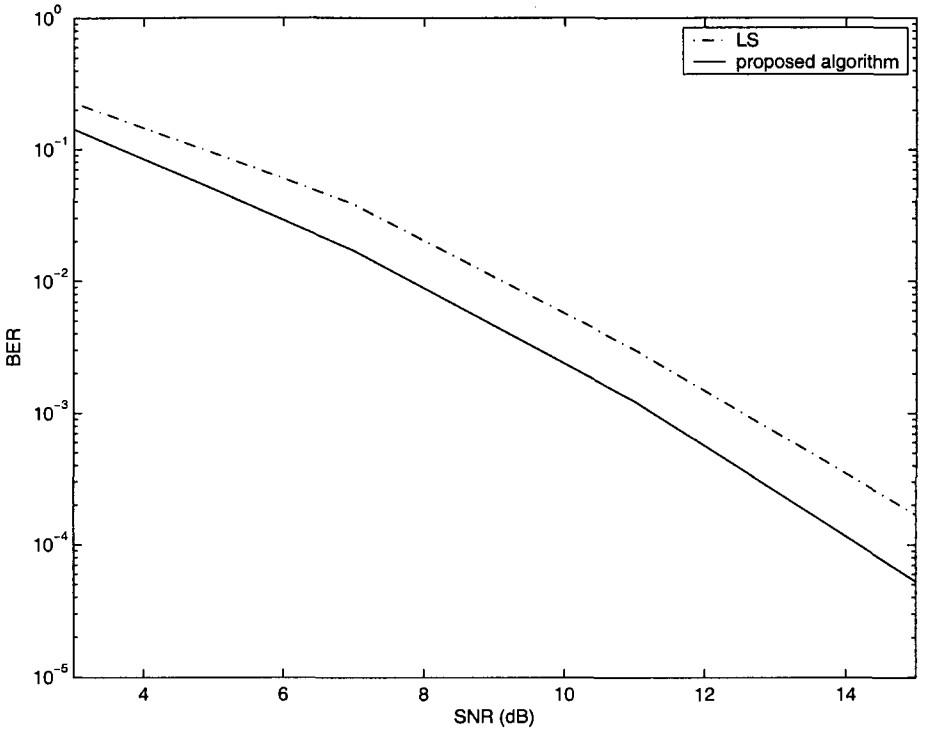

Figure 3.5: BER versus SNR for a 2×4 frequency-selective channel. 1 OFDM symbol with 8 pilot subcarriers and $\eta > 0.2$.

method remains almost the same with the increase of the number of OFDM symbols employed for channel estimation.

Experiment 5: *BER versus SNR*

In this experiment, the BER performance of the MIMO-OFDM system is investigated by using the estimated channel matrix and an ordered vertical-Bell laboratories layered space time (V-BLAST) decoder. The simulation involves 5000 Monte Carlo runs of the transmission of one OFDM symbol with 8 pilot subcarriers. Fig. 3.5 shows the BER performance versus the SNR for the LS method and the proposed semi-blind method, when $\eta > 0.2$. It is seen that the performance of the proposed

semi-blind algorithm is superior to the LS method by 2$\sim$5 dB.

3.5 Conclusion

In this chapter, a nulling-based semi-blind MIMO-OFDM channel estimation approach that incorporates a linear prediction based blind criterion into the least-square method, has been proposed. A practical yet very efficient scheme has been presented for the determination of the weighting factor in the semi-blind cost function. The perturbation analysis of the MIMO linear prediction has justified the advantage of the semi-blind method over the pure blind estimation, and led to a closed-form MSE expression for the blind criterion. The proposed method has been simulated and compared with the training based LS method, showing a significant improvement in terms of the MSE of the channel estimate.

Chapter 4

Frequency-Domain Nulling-based Semi-Blind Channel Estimation with Pulse-Shaping

4.1 Introduction

It is well known that the pulse-shaping filter as well as the matched filter are commonly used in digital communication systems. Perhaps for the sake of simplicity, however, many existing channel estimation methods did not take into consideration either the effect of the pulse-shaping filter in the transmitter or that of the matched filter in the receiver. As such, these methods have actually been developed for the estimation of the composite channel including the pulse-shaping and matched filters. Considering that both filters are known to the receiver and the only unknown part is the discrete-time channel [60], ignoring their existence would lead to less accurate estimation results. By utilizing the information of both filters, some improved channel estimation algorithms have been obtained for OFDM systems [61, 62] and CDMA systems [63,64]. Motivated by this observation, in this chapter, we will propose a new methodology for the channel estimation of pulse-shaped MIMO-OFDM systems with

an emphasis on the development of a frequency-domain linear prediction (LP)-based semi-blind method for multipath channels whose taps are in line with the sampling duration. Another contribution of this chapter is to present an improved least-square (LS) method for estimating the multipath channels whose taps occur in a fraction of the sampling duration.

In MIMO-OFDM systems, the received time-domain signal is often corrupted due to imperfections caused by some factors such as the frequency offset and the larger peak-to-average power ratio (PAPR) etc. Considering that many compensation techniques for these imperfections can only be implemented in the frequency-domain [61, 88, 89], a high quality time-domain signal may not be available. Thus, the existing time-domain MIMO channel estimation techniques, such as the LP-based semi-blind method, cannot be directly applied to MIMO-OFDM systems. To use those time-domain methods, an IFFT processor is needed in the receiver to convert the good frequency-domain signal to the desired high-quality time-domain signal, which imposes an additional computational burden and a long time delay in real-time implementation. In this chapter, therefore, we will develop a method of estimating the time-domain correlation matrix in the LP-based semi-blind method directly from the received frequency-domain signal.

The structure of this chapter is briefed as follows. Section 4.2 deals with the frequency-domain channel estimation of MIMO-OFDM systems with pulse-shaping. By utilizing the knowledge of both the pulse-shaping filter and the matched filter, two improved frequency-domain channel estimation algorithms are developed. Section 4.3 develops a frequency-domain estimation algorithm for the time-domain cor-

relation matrix used in the semi-blind channel estimation. The new development includes a derivation of the frequency-domain correlation matrix, the proof of the equivalence of the frequency-domain correlation matrix and the time-domain correlation matrix, and a direct computation of the time-domain correlation matrix from the frequency-domain version. Section 4.4 comprises a number of experimentations validating the effectiveness of the proposed method. Finally, Section 4.5 summarizes the main contributions of this chapter.

4.2 Proposed Channel Estimation with Pulse-Shaping

As the pulse-shaping filter and the matched filter are normally pre-determined in a communication system, their knowledge should be exploited to improve the channel estimation accuracy. However, many of the existing channel estimation methods [25–28, 31, 32, 83] have not yet taken into account the pulse-shaping in the transmitter and the matched filtering in the receiver. In this section, by utilizing the knowledge of pulse shaping and matched filtering, two improved frequency-domain channel estimation algorithms are proposed for pulse-shaped MIMO-OFDM systems.

4.2.1 Channel Modelling of Pulse-Shaped MIMO-OFDM Systems

A typical pulse-shaping filter in communication systems has the following raised-cosine impulse response [90]

$$g(t) = \text{sinc}\left(\frac{\pi t}{T}\right)\frac{\cos\left(\frac{\beta\pi t}{T}\right)}{1-\left(\frac{2\beta t}{T}\right)^2}$$

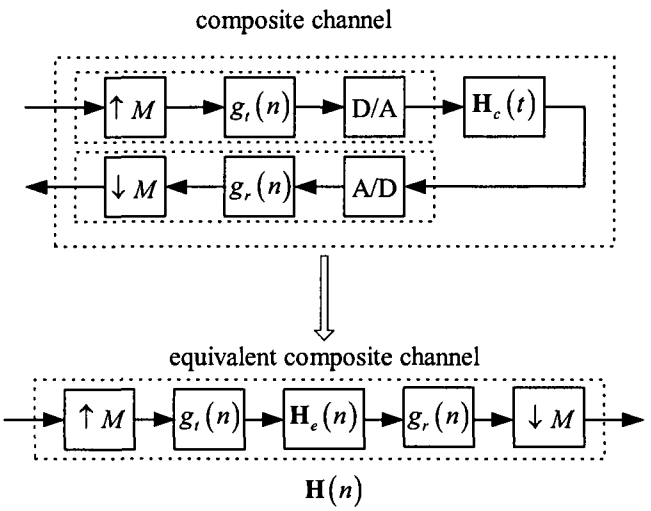

Figure 4.1: Discrete-time channel model with pulse-shaping

where β is the roll-off factor and T the symbol period. In digital communications, the pulse-shaping filter is often realized by an up-sampled raised-cosine FIR filter [91]. Thus, the composite channel model should include the pulse-shaping filter, the analog multi-path channel $\mathbf{H}_c(t)$ and the matched filter as shown in Fig. 4.1 [60–64]. In this model, an upsampling is implemented by inserting $M-1$ zeros between any two consecutive input samples prior to pulse-shaping. As such, the transmit filter $g_t(t)$ and the receive filter $g_r(t)$ in (2.2) are replaced by two root raised-cosine FIR filters $g_t(n)$ and $g_r(n)$, whose sampling period is $\frac{T}{M}$. In the upsampling domain, namely, the discrete-time domain with a sampling duration of $\frac{T}{M}$, the combination of the D/A converter, the multipath channel $\mathbf{H}_c(t)$ and the A/D converter can be represented by an equivalent discrete-time "multipath" channel $\mathbf{H}_e(n)$. In general, the channel path may not arrive at the exact sampling time, but it can be considered as an equivalent path occurring at the sampling instant synchronized to $\mathbf{H}_e\left(\frac{T}{M}\right)$, since the waveform of the D/A converter can normally be assumed as $p(t)=1,\ 0 \le t < \frac{T}{M}$. Thus, in the

case of L_d paths, the discrete-time channel $\mathbf{H}_e(n)$ can be represented by

$$\mathbf{H}_e(n) = \sum_{i=0}^{L_d-1} \mathbf{D}(i)\,\delta(n - l_i) \tag{4.1}$$

where $\mathbf{D}(i)$ and l_i are the channel matrix and the delay with respect to the i-th path. Now the composite discrete-time channel $\mathbf{H}(n)$ can be regarded as a downsampled version of the convolution of the transmit pulse-shaping filter $g_t(n)$, the discrete-time multipath channel $\mathbf{H}_e(n)$ and the received matched filter $g_r(n)$. It should be mentioned that the delay l_i can be determined prior to channel estimation. In particular, in advanced wireless networks, the time of arrivals (TOA) are often estimated at the start of communication and updated periodically. For example, the TOA estimation is conducted by using the ranging techniques for the uplink synchronization phase of the OFDMA (OFDM Access) systems [8] or for some geolocation applications [92]. On the other hand, TOA are known to be a slow fading parameter compared with the fast fading parameter (i.e. complex fading amplitude), which means that once an estimate of TOA is obtained, it can be used to estimate the fading amplitude for a relatively larger period of time. In the following, based on the knowledge of TOAs, i.e., l_i, $(i = 0, 1, \cdots, L_d - 1)$, we develop two improved channel estimation algorithms for MIMO-OFDM systems [93].

4.2.2 Semi-Blind Estimation of Sampling Duration-based Channels

Let us consider first the case of sampling duration-based channels in which the path delay is measured by a multiple of the sampling duration T, i.e., $l_i = iM$. In this case, the (i_R, i_T)-th element of the composite discrete-time channel $\mathbf{H}(n)$ is given

by

$$h_{i_R,i_T}(n) = g_0(n) * d_{i_R,i_T}(n) \tag{4.2}$$

where $d_{i_R,i_T}(n)$ is the (i_R, i_T)-th element of the discrete-time multipath channel $\mathbf{H}_e(n)$ $= \mathbf{D}(n)$, and $g_0(n) = g(Mn)$ with $g(n) = g_t(n) * g_r(n)$. Note that, in (4.2), $g_0(n) = 0,\ n \notin [0, L_g - 1]$, $d_{i_R,i_T}(n) = 0, n \notin [0, L_d - 1]$, and $L = L_g + L_d - 1$. It should be mentioned that a common assumption used in many existing algorithms is $g_0(n) = \delta(n)$, which implies $\mathbf{H}(n) = \mathbf{D}(n)$. In this sense, therefore, the pulse-shaping effect has been neglected. However, this assumption is not true in practical systems. In what follows, we will improve the semi-blind algorithm proposed in Chapter 3 by using (4.2), namely, we will estimate $\mathbf{D}(n)$, instead of the large-dimensional matrix $\mathbf{H}(n)$, with the information of $g_0(n)$. Since the number of channel parameters has been considerably decreased, the estimation performance of the new approach is expected to be much better than that of those focusing only on the estimation of the composite channel $\mathbf{H}(n)$ [94, 95].

Using (4.2), the channel link between the i_R-th receive antenna and i_T-th transmit antenna can be described as the following vector,

$$\mathbf{h}_{i_R,i_T} \triangleq [h_{i_R,i_T}(0), \cdots, h_{i_R,i_T}(L-1)]^T = \mathbf{\Lambda}\mathbf{d}_{i_R,i_T} \tag{4.3}$$

where $\mathbf{\Lambda}$ is an $L \times L_d$ circulant matrix with its first column given by $[g_0(0), \cdots, g_0(L_g), \mathbf{0}_{1\times(L-L_g)}]^T$ and $\mathbf{d}_{i_R,i_T} \triangleq [d_{i_R,i_T}(0), \cdots, d_{i_R,i_T}(L_d - 1)]^T$. From (4.3), the i_R-th partition of the composite channel vector $\mathbf{h}$ can be written as

$$\mathbf{h}_{i_R} \triangleq \left[\mathbf{h}^T_{i_R,1}, \cdots, \mathbf{h}^T_{i_R,N_T}\right]^T = (\mathbf{I}_{N_T} \otimes \mathbf{\Lambda})\,\mathbf{d}_{i_R} \tag{4.4}$$

where $\mathbf{d}_{i_R} \triangleq \left[\mathbf{d}_{i_R,1}^T, \cdots, \mathbf{d}_{i_R,i_T}^T\right]^T$. Using (4.4), one can obtain

$$\mathbf{h} = \mathbf{\Psi}\mathbf{d} \tag{4.5}$$

where $\mathbf{\Psi} \triangleq [\mathbf{I}_{N_R} \otimes (\mathbf{I}_{N_T} \otimes \mathbf{\Lambda})]$ and $\mathbf{d} \triangleq \left[\mathbf{d}_1^T, \cdots, \mathbf{d}_{N_R}^T\right]^T$. Thus, (4.5) gives a relationship between the composite channel vector and the pure multipath channel vector.

Substituting (4.5) into (3.17), a new semi-blind cost function for MIMO-OFDM channel estimation with pulse-shaping can be formulated as

$$\min_{\hat{\mathbf{d}}} \Delta = \left\| \mathbf{Y}_{\text{pilot}} - \tilde{\mathbf{A}}'\mathbf{d} \right\|_F^2 + \alpha \left\| \hat{\mathbf{B}}'\mathbf{d} \right\|_F^2 \tag{4.6}$$

where $\tilde{\mathbf{A}}' \triangleq \tilde{\mathbf{A}}\mathbf{\Psi}$ and $\hat{\mathbf{B}}' \triangleq \hat{\mathbf{B}}\mathbf{\Psi}$. Similar to (3.18), the estimate of the channel vector can be derived as

$$\hat{\mathbf{d}} = \left(\left(\tilde{\mathbf{A}}'\right)^H \tilde{\mathbf{A}}' + \alpha \left(\hat{\mathbf{B}}'\right)^H \hat{\mathbf{B}}' \right)^{\dagger} \left(\tilde{\mathbf{A}}'\right)^H \mathbf{Y}_{\text{pilot}}. \tag{4.7}$$

The above equation gives a time-domain semi-blind solution for the channel estimation of MIMO-OFDM systems with pulse-shaping. Note that the computational complexity of $\hat{\mathbf{d}}$ depends on the sizes of the matrices $\tilde{\mathbf{A}}'$ and $\hat{\mathbf{B}}'$, which are determined by the length of the pulse-shaping and matched filters as well as the length of the pure multipath channel. When the length L_g of the filter $g(n)$ is relatively small as compared to the channel length L_d, (4.7) gives an efficient channel estimate. For a large value of L_g, however, the total length $L = L_g + L_d - 1$ of the composite channel can be very large which may incur a high complexity in the computation of (4.7). In what follows, we propose a very efficient frequency-domain estimation approach regardless of the relative size of L_g.

The frequency-domain signal model between the i_R-th receive antenna and the i_T-th transmit antenna can be represented by

$$Y_{i_R}(k) = H_{i_R,i_T}(k) X_{i_T}(k), \text{for } k \in [0, K-1] \tag{4.8}$$

where $H_{i_R,i_T}(k) \triangleq \mathbf{f}_{0k}\left[\mathbf{h}_{i_R,i_T}^T, \mathbf{0}_{1\times(K-L)}\right]^T$ and $\mathbf{f}_{0k}$ is the k-th row of the $K \times K$ DFT matrix $\mathbf{F}_0$. From (4.2), we can derive

$$H_{i_R,i_T}(k) = G(k) D_{i_R,i_T}(k) \tag{4.9}$$

where

$$G(k) \triangleq \mathbf{f}_{0k}\left[g_0(0), \cdots, g_0(L_g-1), \mathbf{0}_{1\times(L-L_g)}\right]^T,$$

$$D_{i_R,i_T}(k) \triangleq \mathbf{f}_{0k}\left[\mathbf{d}_{i_R,i_T}^T, \mathbf{0}_{1\times(K-L_b)}\right]^T.$$

Using (4.9) into (4.8), we can obtain the received signal after removing the effect of the pulse-shaping and matched filters,

$$\begin{aligned} Y'_{i_R}(k) &\triangleq G^{-1}(k) Y_{i_R}(k) \\ &= D_{i_R,i_T}(k) X_{i_T}(k) + V'(k) \end{aligned} \tag{4.10}$$

where $V'(k)$ is the frequency-domain noise. Clearly, by applying the IFFT to (4.10), a time-domain version of the received signal without the effect of pulse shaping and matched filters can be obtained as

$$y'_{i_R}(n) = \sum_{i_T=1}^{N_T} d_{i_R,i_T}(n) \circledast x_{i_T}(n) + v'_{i_R}(n). \tag{4.11}$$

Interestingly, (4.11) is similar to the time-domain signal model in (2.3). Therefore, a semi-blind solution for the discrete-time multipath channel $\hat{\mathbf{d}}$ can be obtained by using

the semi-blind method discussed in Chapter 3. It should be mentioned that the above IFFT operation to convert the frequency-domain signal $Y'_{i_R}(k)$ to the time-domain signal $y'_{i_R}(n)$ imposes an additional computational burden and a long time delay for the implementation of the semi-blind algorithm. In Section 4.3, we will develop a frequency-domain method for the calculation of the time-domain correlation matrix which avoids an IFFT operation.

4.2.3 LS Estimation of Upsampling Duration-based Channels

Now, we consider the more general upsampling-duration-based channels. By using Fig. 4.1 and (4.1) and noting that the equivalent composite channel $\mathbf{H}(n)$ is a down-sampled version of the discrete-time multipath channel $\mathbf{H}_e(n)$ given by (4.1), one can express $\mathbf{H}(n)$ as

$$\mathbf{H}(n) = \sum_{i=0}^{L_d-1} \mathbf{D}(i)\, b_i(n - m_i) \tag{4.12}$$

where

$$m_i = \lfloor \frac{l_i + M - 1}{M} \rfloor, \tag{4.13}$$

$$b_i(n) = g_q(n) = g(Mn + q), \quad q = m_i M - l_i. \tag{4.14}$$

As $l_i\,(i = 0, 1, \cdots, L_d - 1)$ are known, m_i and $b_i(n)$, $(i = 0, 1, \cdots, L_d - 1)$ can be obtained by using (4.13) and (4.14), respectively. Now, we propose an enhanced LS algorithm for the estimation of $\mathbf{D}(i)$ $(i = 0, 1, \cdots, L_d - 1)$, with the information of m_i and $b_i(n)\,(i = 0, 1, \cdots, L_d - 1)$.

Assume that the K_p sub-carriers, say from i_{pilot1} to $i_{\text{pilot}K_p}$, of each OFDM symbol carry the pilot signal. The transmitted and the received pilot vectors for each transmit

and receive antenna pair can be defined as

$$\mathbf{X}_{i_T,\text{pilot}}(m) \triangleq \left[X_{i_T}(m, i_{\text{pilot}1}), \cdots, X_{i_T}(m, i_{\text{pilot}K_p})\right]^T,$$

$$\mathbf{Y}_{i_R,\text{pilot}}(m) \triangleq \left[Y_{i_R}(m, i_{\text{pilot}1}), \cdots, Y_{i_R}(m, i_{\text{pilot}K_p})\right]^T.$$

It should be noted that the pilot signal might not be located at the same position in each OFDM symbol. For the m-th OFDM symbol, one can form a $K_P \times L$ matrix, say $\mathbf{F}(m)$, by taking only the rows of $\mathbf{F}_1$ associated with the K_p pilot sub-carriers. It was shown in [25] that

$$\mathbf{Y}_{i_R,\text{pilot}}(m) = \sum_{i_T=1}^{N_T} \mathbf{X}^m_{i_T,\text{diag}} \mathbf{F}(m) \mathbf{h}_{i_R,i_T} + \boldsymbol{\xi}_{i_R,\text{pilot}}(m) \tag{4.15}$$

where $\mathbf{X}^m_{i_T,\text{diag}} \triangleq \text{diag}\left(\mathbf{X}_{i_T,\text{pilot}}(m)\right)$ and $\boldsymbol{\xi}_{i_R}(m)$ represents the frequency-domain noise corresponding to $v_{i_R}(m,n)$ in (2.3). From (4.12), one can obtain

$$\mathbf{h}_{i_R,i_T} = [\mathbf{b}_0, \mathbf{b}_1, \cdots, \mathbf{b}_{L_d-1}]\, \mathbf{d}_{i_R,i_T} \tag{4.16}$$

where $\mathbf{d}_{i_R,i_T} = [d_{i_R,i_T}(0), \cdots, d_{i_R,i_T}(L_d-1)]^T$ and $\mathbf{b}_i$ is a vector whose k-th element is given by

$$\mathbf{b}_i(k) = \begin{cases} b_i(k-m_i), & m_i \le k \le L_g + m_i - 1 \\ 0, & \text{otherwise} \end{cases}.$$

From (4.15) and (4.16), the frequency-domain pilot signal received at the i_R-th receive antenna with respect to g OFDM symbols can be obtained as

$$\begin{aligned} \mathbf{Y}_{i_R,\text{pilot}} &\triangleq \left[\mathbf{Y}^T_{i_R,\text{pilot}}(0), \cdots, \mathbf{Y}^T_{i_R,\text{pilot}}(g-1)\right]^T \\ &= \mathbf{A}_\text{d}\mathbf{d}_{i_R} + \boldsymbol{\xi}_{i_R,\text{pilot}} \end{aligned} \tag{4.17}$$

where

$$\mathbf{A}_{\mathrm{d}} \triangleq \begin{bmatrix} \mathbf{X}^{0}_{1,\mathrm{diag}}\mathbf{F}(0) & \cdots & \mathbf{X}^{0}_{N_T,\mathrm{diag}}\mathbf{F}(0) \\ \vdots & \ddots & \vdots \\ \mathbf{X}^{g-1}_{1,\mathrm{diag}}\mathbf{F}(g-1) & \cdots & \mathbf{X}^{g-1}_{N_T,\mathrm{diag}}\mathbf{F}(g-1) \end{bmatrix} \otimes \left[\mathbf{b}_0, \mathbf{b}_1, \cdots, \mathbf{b}_{L_d-1}\right],$$

$$\mathbf{d}_{i_R} \triangleq \left[\mathbf{d}^T_{i_R,1}, \cdots, \mathbf{d}^T_{i_R,N_T}\right]^T,$$

$$\boldsymbol{\xi}_{i_R,\mathrm{pilot}} \triangleq \left[\boldsymbol{\xi}^T_{i_R,\mathrm{pilot}}(0), \cdots, \boldsymbol{\xi}^T_{i_R,\mathrm{pilot}}(g-1)\right]^T.$$

By defining

$$\bar{\mathbf{Y}}_{\mathrm{pilot}} \triangleq \left[\mathbf{Y}_{1,\mathrm{pilot}}, \cdots, \mathbf{Y}_{N_R,\mathrm{pilot}}\right],$$

$$\bar{\mathbf{D}} \triangleq \left[\mathbf{d}_1, \cdots, \mathbf{d}_{N_R}\right], \tag{4.18}$$

$$\bar{\boldsymbol{\xi}}_{\mathrm{pilot}} \triangleq \left[\boldsymbol{\xi}_{1,\mathrm{pilot}}, \cdots, \boldsymbol{\xi}_{N_R,\mathrm{pilot}}\right],$$

one can have

$$\bar{\mathbf{Y}}_{\mathrm{pilot}} = \mathbf{A}_{\mathrm{d}}\bar{\mathbf{D}} + \bar{\boldsymbol{\xi}}_{\mathrm{pilot}}, \tag{4.19}$$

which gives

$$\hat{\bar{\mathbf{D}}} = \mathbf{A}^{\dagger}_{\mathrm{d}}\bar{\mathbf{Y}}_{\mathrm{pilot}}. \tag{4.20}$$

Finally, substituting $\hat{\bar{\mathbf{D}}}$ into (4.12) leads to the enhanced channel estimate $\hat{\mathbf{H}}(n)$. It will be shown in Section 4.4 that, when the number of multipath L_d is much less than the length L of the composite channel, the performance of the enhanced channel estimation approach is significantly superior to that of the original LS method. Moreover, the complexity of the enhanced LS method is much lower than that of the original version.

4.3 Frequency-Domain Estimation of Time-Domain Correlation Matrix

The core of the semi-blind MIMO-OFDM channel estimation algorithm is the computation of three matrices, i.e., $\tilde{\mathbf{R}}_{n-1}$ in (2.37), $\ddot{\mathbf{R}}_n$ in (2.38), and $\mathbf{R}(0)$ given by

$$\mathbf{R}(0) = \mathrm{E}\left\{\mathbf{y}(n)\,\mathbf{y}^H(n)\right\}. \tag{4.21}$$

As mentioned in Subsection 4.2.2, the IFFT process is required in the receiver to obtain the time-domain received signal after removing the effect of pulse shaping and matched filters for the calculation of the correlation matrices. In this section, we propose a frequency-domain estimation method for the time-domain correlation matrix directly from the frequency-domain signal without requiring the IFFT operation.

4.3.1 The Time-Domain Correlation Matrix $\hat{\mathbf{R}}_{\mathrm{T}}$ for Semi-blind Channel Estimation

We first give an unified representation of all the three matrices, $\tilde{\mathbf{R}}_{n-1}$, $\ddot{\mathbf{R}}_n$ and $\mathbf{R}(0)$. Suppose the time-domain version of the received signal after the IFFT is given by $\mathbf{y}'(n) = \left[y'_1(n), y'_2(n), \cdots, y'_{N_R}(n)\right]^T$. By letting

$$\mathbf{y}_{P+1}(n) \triangleq \left[\left(\mathbf{y}'(n)\right)^T, \left(\mathbf{y}'(n-1)\right)^T, \cdots, \left(\mathbf{y}'(n-P)\right)^T\right]^T,$$

the 2nd-order statistics in terms of the correlation matrix $\mathbf{R}_T$ of $\mathbf{y}_{P+1}$ can be estimated by

$$\hat{\mathbf{R}}_{\mathrm{T}} = \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{y}_{P+1}(n)\,\mathbf{y}_{P+1}^H(n). \tag{4.22}$$

Note that $\mathbf{y}_{P+1}(n)$ for $n = 0, 1, \cdots, P-1$, can be obtained using $\mathbf{y}(n-j) \triangleq \mathbf{y}(K+n-j)$ for $n < j$ due to the circular convolution. By defining

$$\hat{\mathbf{R}}(l) = \begin{bmatrix} \hat{R}_{1,1}(l) & \cdots & \hat{R}_{1,N_R}(l) \\ \vdots & \ddots & \vdots \\ \hat{R}_{N_R,1}(l) & \cdots & \hat{R}_{N_R,N_R}(l) \end{bmatrix}, \text{for } l = -P, \cdots, 0, \cdots, P \tag{4.23}$$

where

$$\hat{R}_{i_{R1},i_{R2}}(l) \triangleq \frac{1}{K} \sum_{n=0}^{K-1} y_{i_{R1}}(n) y^*_{i_{R2}}(n-l), \tag{4.24}$$

(4.22) can be rewritten as

$$\hat{\mathbf{R}}_{\mathrm{T}} = \begin{bmatrix} \hat{\mathbf{R}}(0) & \hat{\mathbf{R}}(1) & \cdots & \hat{\mathbf{R}}(P) \\ \hat{\mathbf{R}}(-1) & \mathbf{R}(0) & \cdots & \hat{\mathbf{R}}(P-1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{R}}(-P) & \hat{\mathbf{R}}(1-P) & \cdots & \hat{\mathbf{R}}(0) \end{bmatrix}. \tag{4.25}$$

By a simple analysis of (2.37), (2.38), (4.21) and (4.22), one can verify that

$$\begin{bmatrix} \hat{\mathbf{R}}(0) & \hat{\tilde{\mathbf{R}}}_n \\ \hat{\tilde{\mathbf{R}}}_n^H & \hat{\tilde{\mathbf{R}}}_{n-1} \end{bmatrix} = \hat{\mathbf{R}}_{\mathrm{T}}. \tag{4.26}$$

In other words, the estimates $\hat{\tilde{\mathbf{R}}}_{n-1}$, $\hat{\tilde{\mathbf{R}}}_n$ and $\hat{\mathbf{R}}(0)$ can directly be obtained from the estimate $\hat{\mathbf{R}}_{\mathrm{T}}$. In the following subsections, we develop a new algorithm for the computation of the time-domain correlation matrix for the sampling-duration-based channel estimation directly from the frequency-domain signal after removing the effect of pulse-shaping and matched filters.

4.3.2 The Frequency-Domain Correlation Matrix $\hat{\mathbf{R}}_{\mathrm{F}}$

Our derivation of the frequency-domain correlation matrix is inspired by the idea of frequency-domain equalization. In the noise-free case, the frequency-domain signal model corresponding to (2.3) can be written as

$$\begin{bmatrix} \mathbf{Y}_1(m) \\ \vdots \\ \mathbf{Y}_{N_R}(m) \end{bmatrix} = \begin{bmatrix} \operatorname{diag}(\mathbf{F}_1\mathbf{h}_{1,1}) & \cdots & \operatorname{diag}(\mathbf{F}_1\mathbf{h}_{1,N_T}) \\ \vdots & \ddots & \vdots \\ \operatorname{diag}(\mathbf{F}_1\mathbf{h}_{N_R,1}) & \cdots & \operatorname{diag}(\mathbf{F}_1\mathbf{h}_{N_R,N_T}) \end{bmatrix} \begin{bmatrix} \mathbf{X}_1(m) \\ \vdots \\ \mathbf{X}_{N_T}(m) \end{bmatrix} \tag{4.27}$$

where

$$\mathbf{h}_{i_R,i_T} \triangleq [h_{i_R,i_T}(0), \cdots, h_{i_R,i_T}(L-1)]^T \tag{4.28}$$

and $\mathbf{F}_1$ consists of the first L columns of a $K \times K$ DFT matrix $\mathbf{F}_0$. Let us consider a frequency-domain equalization for (4.27), namely,

$$\mathbf{Z}(m) = [\operatorname{diag}(\mathbf{F}_2\mathbf{w}_1), \cdots, \operatorname{diag}(\mathbf{F}_2\mathbf{w}_{N_R})] \begin{bmatrix} \mathbf{Y}_1(m) \\ \vdots \\ \mathbf{Y}_{N_R}(m) \end{bmatrix} \tag{4.29}$$

where $\mathbf{w}_i \triangleq [w_i(0), \cdots, w_i(P)]^T$, $(i = 1, \cdots, N_R)$ is a transversal equalizer of size $P+1$ with respect to the i-th receive antenna and $\mathbf{F}_2$ consists of the first $P+1$ columns of the DFT matrix. Substituting (4.27) into (4.29) gives

$$\mathbf{Z}(m) = \left[\sum_{i_R=1}^{N_R} \operatorname{diag}(\mathbf{F}_2\mathbf{w}_{i_R}) \operatorname{diag}(\mathbf{F}_1\mathbf{h}_{i_R,1}), \cdots, \sum_{i_R=1}^{N_R} \operatorname{diag}(\mathbf{F}_2\mathbf{w}_{i_R}) \operatorname{diag}(\mathbf{F}_1\mathbf{h}_{i_R,N_T}) \right] \begin{bmatrix} \mathbf{X}_1(m) \\ \vdots \\ \mathbf{X}_{N_T}(m) \end{bmatrix}. \tag{4.30}$$

Letting $\mathbf{F}_3$ be the first $L+P$ columns of the DFT matrix and

$$\mathbf{c}_{i_R,i_T} \triangleq \left[c_{i_R,i_T}(0), \cdots, c_{i_R,i_T}(L+P-1)\right]^T$$

where

$$c_{i_R,i_T}(n) = w_{i_R}(n) * h_{i_R,i_T}(n),$$

(4.30) can be rewritten as

$$\mathbf{Z}(m) =$$

$$\left[\mathrm{diag}\left[\mathbf{F}_3\left(\mathbf{c}_{1,1}+\cdots+\mathbf{c}_{N_R,1}\right)\right], \cdots, \mathrm{diag}\left[\mathbf{F}_3\left(\mathbf{c}_{1,N_T}+\cdots+\mathbf{c}_{N_R,N_T}\right)\right]\right] \begin{bmatrix} \mathbf{X}_1(m) \\ \vdots \\ \mathbf{X}_{N_T}(m) \end{bmatrix}. \tag{4.31}$$

If a specific set of equalizers $\mathbf{w}_i$, $(i=1,\cdots,N_R)$ with respect to the i_T-th transmit antenna is designed such that

$$\mathrm{diag}\left[\mathbf{F}_3\left(\mathbf{c}_{1,i_T}+\cdots+\mathbf{c}_{N_R,i_T}\right)\right] = \begin{cases} \mathbf{I} & \text{if } (i=i_T) \\ \mathbf{0} & \text{if } (i \neq i_T) \end{cases},$$

then the signal sent by this transmit antenna can be recovered, namely,

$$\mathbf{Z}(m) = \mathbf{X}_{i_T}(m).$$

Thus, if such N_T sets of equalizers are determined, all the signals from the N_T transmit antennas can be recovered, resulting in a frequency-domain equalization for MIMO-OFDM systems.

Using the frequency-domain equalization idea, we now derive a frequency-domain

correlation matrix. For the k-th subcarrier, (4.29) gives

$$\mathbf{Z}(m,k) = \left[\mathrm{diag}\left(\mathbf{F}_2(k)\,\mathbf{w}_1\right), \cdots, \mathrm{diag}\left(\mathbf{F}_2(k)\,\mathbf{w}_{N_R}\right)\right] \begin{bmatrix} Y_1(m,k) \\ \vdots \\ Y_{N_R}(m,k) \end{bmatrix} \tag{4.32}$$

where $\mathbf{F}_2(k)$ is the k-th row of the matrix $\mathbf{F}_2$. For notational simplicity, the index m of OFDM symbols is dropped from now on without loss of clarity. Thus, (4.32) can be rewritten as

$$\mathbf{Z}(k) = \left[\mathbf{w}_1^T, \cdots, \mathbf{w}_{N_R}^T\right] \mathbf{Y}'(k) \tag{4.33}$$

where

$$\mathbf{Y}'(k) \triangleq \left(\mathbf{I}_{N_R} \otimes \mathbf{F}_2^T(k)\right) \begin{bmatrix} Y_1(k) \\ \vdots \\ Y_{N_R}(k) \end{bmatrix}. \tag{4.34}$$

Note that $\mathbf{Y}'(k)$ can be regarded as an input frequency-domain signal for the equalizer $\mathbf{w}_i$, $(i = 1, \cdots, N_R)$ at the k-th subcarrier. Let us consider the autocorrelation matrix of $\mathbf{Y}'(k)$,

$$\hat{\mathbf{R}}_{\mathrm{F}} \triangleq \frac{1}{K} \sum_{k=0}^{K-1} \mathbf{Y}'(k) \left(\mathbf{Y}'(k)\right)^H, \tag{4.35}$$

which is called, for the sake of convenience, the frequency-domain correlation matrix in this work, even though it is actually an estimate. Using (4.34) into (4.35), $\hat{\mathbf{R}}_{\mathrm{F}}$ can be rewritten as

$$\hat{\mathbf{R}}_{\mathrm{F}} = \begin{bmatrix} \boldsymbol{\Delta}_{1,1} & \cdots & \boldsymbol{\Delta}_{1,N_R} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Delta}_{N_R,1} & \cdots & \boldsymbol{\Delta}_{N_R,N_R} \end{bmatrix} \tag{4.36}$$

where

$$\boldsymbol{\Delta}_{i_{R1},i_{R2}} \triangleq \frac{1}{K}\sum_{k=0}^{K-1} \mathbf{F}_2^T(k)\,\mathbf{F}_2^*(k)\,Y_{i_{R1}}(k)\,Y_{i_{R2}}^*(k)\,,\ (i_{R1},i_{R2}=1,\cdots,N_R)\,. \tag{4.37}$$

It will be shown in the next subsection that the frequency-domain correlation matrix $\hat{\mathbf{R}}_{\mathrm{F}}$ contains exactly the same information as the time-domain correlation matrix $\hat{\mathbf{R}}_{\mathrm{T}}$, yet it would significantly facilitate the computation of the second-order statistics in channel estimation by avoiding the IFFT operation converting the frequency-domain signal $Y'_{i_R}(k)$ to the time-domain signal $y'_{i_R}(n)$ in the proposed semi-blind algorithm.

4.3.3 Computation of $\hat{\mathbf{R}}_{\mathrm{T}}$ based on $\hat{\mathbf{R}}_{\mathrm{F}}$

In this subsection, we present a frequency-domain method for the computation of the time-domain correlation matrix $\hat{\mathbf{R}}_{\mathrm{T}}$ via the frequency-domain correlation matrix $\hat{\mathbf{R}}_{\mathrm{F}}$. To this end, we first reveal the relationship between the two matrices. By defining $\phi(k) = \mathrm{e}^{-j2\pi(k/K)}$, (4.37) can be rewritten as

$$\boldsymbol{\Delta}_{i_{R1},i_{R2}} = \frac{1}{K}\sum_{k=0}^{K-1} \begin{bmatrix} 1 & \phi^{-1}(k) & \phi^{-2}(k) & \cdots & \phi^{-P}(k) \\ \phi^{1}(k) & 1 & \phi^{-1}(k) & \cdots & \phi^{1-P}(k) \\ \phi^{2}(k) & \phi^{1}(k) & 1 & \cdots & \phi^{2-P}(k) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi^{P}(k) & \phi^{P-1}(k) & \phi^{P-2}(k) & \cdots & 1 \end{bmatrix} Y_{i_{R1}}(k)\,Y_{i_{R2}}^*(k)\,. \tag{4.38}$$

On the other hand, by letting

$$y_{i_{R1}}(n) \triangleq \frac{1}{\sqrt{K}}\sum_{k=0}^{K-1} Y_{i_{R1}}(k)\,\mathrm{e}^{j2\pi(kn/K)}, \tag{4.39}$$

$$y_{i_{R2}}(n) \triangleq \frac{1}{\sqrt{K}}\sum_{k=0}^{K-1} Y_{i_{R2}}(k)\,\mathrm{e}^{j2\pi(kn/K)}, \tag{4.40}$$

$\hat{R}_{i_{R1},i_{R2}}$ as defined in (4.24) can be expressed as

$$\begin{aligned}\hat{R}_{i_{R1},i_{R2}}(l) &= \frac{1}{K}\sum_{n=0}^{K-1} y_{i_{R1}}(n)\left[\frac{1}{\sqrt{K}}\sum_{k=0}^{K-1} Y^{*}_{i_{R2}}(k)\,\mathrm{e}^{-j2\pi k(n-l)/K}\right] \\ &= \frac{1}{K}\sum_{k=0}^{K-1}\left[\frac{1}{\sqrt{K}}\sum_{n=0}^{K-1} y_{i_{R1}}(n)\,\mathrm{e}^{-j2\pi kn/K}\right] Y^{*}_{i_{R2}}(k)\,\mathrm{e}^{j2\pi(kl/K)} \\ &= \frac{1}{K}\sum_{k=0}^{K-1} Y_{i_{R1}}(k)\,Y^{*}_{i_{R2}}(k)\,\phi^{-l}(k)\,. \end{aligned} \tag{4.41}$$

It is clear that (4.41) gives the expression for the elements of $\mathbf{\Delta}_{i_{R1},i_{R2}}$, i. e.,

$$\mathbf{\Delta}_{i_{R1},i_{R2}} = \begin{bmatrix} \hat{R}_{i_{R1},i_{R2}}(0) & \hat{R}_{i_{R1},i_{R2}}(1) & \hat{R}_{i_{R1},i_{R2}}(2) & \cdots & \hat{R}_{i_{R1},i_{R2}}(P) \\ \hat{R}_{i_{R1},i_{R2}}(-1) & \hat{R}_{i_{R1},i_{R2}}(0) & \hat{R}_{i_{R1},i_{R2}}(1) & \cdots & \hat{R}_{i_{R1},i_{R2}}(P-1) \\ \hat{R}_{i_{R1},i_{R2}}(-2) & \hat{R}_{i_{R1},i_{R2}}(-1) & \hat{R}_{i_{R1},i_{R2}}(0) & \cdots & \hat{R}_{i_{R1},i_{R2}}(P-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{R}_{i_{R1},i_{R2}}(-P) & \hat{R}_{i_{R1},i_{R2}}(1-P) & \hat{R}_{i_{R1},i_{R2}}(2-P) & \cdots & \hat{R}_{i_{R1},i_{R2}}(0) \end{bmatrix}. \tag{4.42}$$

Thus, $\mathbf{\Delta}_{i_{R1},i_{R2}}$ can be expressed in terms of $\hat{R}_{i_{R1},i_{R2}}(l)\,,(l=-P,-P+1\cdots,P)$ as

$$\mathbf{\Delta}_{i_{R1},i_{R2}} = \sum_{l=-P}^{P} \hat{R}_{i_{R1},i_{R2}}(l)\otimes\mathbf{I}_l \tag{4.43}$$

where $\mathbf{I}_l$ is a partial identity matrix, which has its nonzero elements 1 only in the l-th diagonal and zero elements otherwise. The sequence of the nonzero diagonals is demonstrated in Fig. 4.2. For example, $l=0$ corresponds to $\mathbf{I}_{(P+1)\times(P+1)}$, while $l=1$ gives

$$\begin{bmatrix} \mathbf{0}_{P\times 1} & \mathbf{I}_{P\times P} \\ 0 & \mathbf{0}_{1\times P} \end{bmatrix}.$$

Now we utilize the above result to derive the relationship between the frequency-domain correlation matrix $\hat{\mathbf{R}}_{\mathrm{F}}$ and the time-domain version $\hat{\mathbf{R}}_{\mathrm{T}}$. Using (4.43), (4.36)

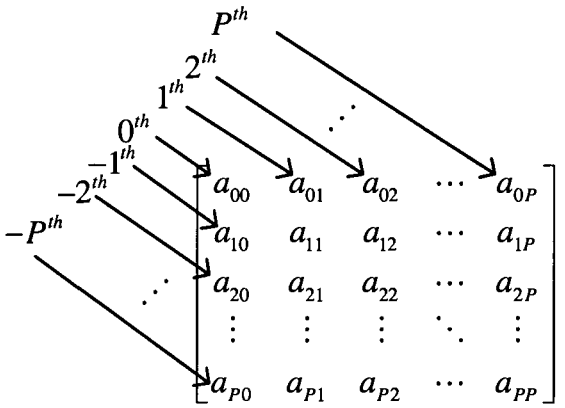

Figure 4.2: Position of the diagonal for the partial identity matrix

can be rewritten as

$$\begin{aligned}\hat{\mathbf{R}}_{\mathrm{F}} &= \begin{bmatrix} \sum_{l=-P}^{P} \hat{\mathbf{R}}_{1,1}(l) \otimes \mathbf{I}_l & \cdots & \sum_{l=-P}^{P} \hat{\mathbf{R}}_{1,N_R}(l) \otimes \mathbf{I}_l \\ \vdots & \ddots & \vdots \\ \sum_{l=-P}^{P} \hat{\mathbf{R}}_{N_R,1}(l) \otimes \mathbf{I}_l & \cdots & \sum_{l=-P}^{P} \hat{\mathbf{R}}_{N_R,N_R}(l) \otimes \mathbf{I}_l \end{bmatrix}, \\ &= \sum_{l=-P}^{P} \hat{\mathbf{R}}(l) \otimes \mathbf{I}_l. \end{aligned} \tag{4.44}$$

On the other hand, from (4.25), $\hat{\mathbf{R}}_{\mathrm{T}}$ can be expressed as

$$\hat{\mathbf{R}}_{\mathrm{T}} = \sum_{l=-P}^{P} \mathbf{I}_l \otimes \hat{\mathbf{R}}(l). \tag{4.45}$$

It is interesting to see from (4.44) and (4.45) that the frequency-domain correlation matrix $\hat{\mathbf{R}}_{\mathrm{F}}$ and the time-domain correlation matrix $\hat{\mathbf{R}}_{\mathrm{T}}$ contain exactly the same information, which is dictated by $\hat{\mathbf{R}}(l), (l = -P, -P+1, \cdots, P)$, and the only difference between the two versions is the order of the Kronecker product.

It is obvious from (4.44) that once $\hat{\mathbf{R}}_{\mathrm{F}}$ has been calculated, $\hat{\mathbf{R}}(l), (l = -P, -P+1, \cdots, P)$ can be obtained, and then $\hat{\mathbf{R}}_{\mathrm{T}}$ can be determined directly from (4.45). The new frequency-domain correlation matrix estimation algorithm can be described as

follows.

Step i) For each subcarrier, compute $\mathbf{Y}'(k)$ using (4.34) based on the received frequency-domain signal;

Step ii) Estimate the frequency-domain correlation matrix $\hat{\mathbf{R}}_{\mathrm{F}}$ using (4.35);

Step iii) Obtain $\hat{\mathbf{R}}(l)$, $(l = -P, -P+1, \cdots, P)$ from $\hat{\mathbf{R}}_{\mathrm{F}}$ in (4.44);

Step iv) Construct $\hat{\mathbf{R}}_{\mathrm{T}}$ from $\hat{\mathbf{R}}(l)$, $(l = 1-P, \cdots, P-1)$ using (4.45).

By using the above frequency-domain estimation algorithm in conjunction with the original semi-blind channel estimation approach proposed in Chapter 3, a frequency-domain semi-blind approach for estimating the sampling-duration-based channels can be obtained. Note that, since the time-domain correlation matrix has already been computed, the practical scheme of determining the weighting factor α suggested in the time-domain approach in Chapter 3 can be directly used here.

Before closing this section, we would like to evaluate the complexity of computing $\hat{\mathbf{R}}_{\mathrm{T}}$ via the proposed frequency-domain method as opposed to the time-domain estimation including the IFFT operation. The frequency-domain method requires the calculation of $\mathbf{Y}'(k)[\mathbf{Y}'(k)]^H$ and $\hat{\mathbf{R}}'_{\mathrm{F}}$, which cost $KN_R^2(P^2+3P+3)$ and $N_R^2(P+1)^2$ complex multiplications, respectively. Note that the calculation of $\hat{\mathbf{R}}_{\mathrm{F}}$ also requires $(K-1)N_R^2(P+1)^2$ complex additions. On the other hand, the time-domain estimation method mainly consists of the calculation of $\hat{\mathbf{R}}_{\mathrm{T}}$ and a K-point IFFT process. The former requires $(K+1)N_R^2(P+1)^2$ multiplications and $(K-1)N_R^2(P+1)^2$ additions, while the latter involves $N_R K\log_2(K)$ complex multiplications and additions. A detailed comparison of the two methods is given in Table 4.1. Considering that K is in general much larger than N_R and P, we have $K\log_2(K)N_R \gg N_R^2(P+2)$.

Thus, the complexity of the frequency-domain algorithm is approximately the same as that for the calculation of $\hat{\mathbf{R}}_{\mathrm{T}}$ in the time-domain algorithm excluding the IFFT. It means that, in comparison with the conventional time-domain estimation method, the IFFT processing has been avoided in the frequency-domain approach.

4.4 Simulation Results

Here we consider a MIMO-OFDM system with 2 transmit and 4 receive antennas. The number of subcarriers is set to 512, the length of cyclic prefix is 10, and the QPSK modulation is used. A square root raised cosine filter with order 16, oversampling rate 4 and rolloff factor 0.15 is used for the pulse-shaping filter and the matched filter. In our experiments, we simulate the frequency-domain semi-blind channel estimation algorithm for a sampling duration-based channel and the enhanced LS channel estimation algorithm for an upsampling duration-based channel.

4.4.1 Sampling Duration-based Channels

A Rayleigh channel modelled by a 3-tap MIMO-FIR filter is assumed, in which each tap corresponds to a 2×4 random matrix whose elements are i.i.d. complex Gaussian variables with zero mean and unit variance. The length of the linear predictor is $P = 4$. As shown in Chapter 3, the channel estimation performance is associated with $\mathbf{H}(0)$. Accordingly, we use the same metric $\eta \triangleq \frac{\|\mathbf{H}(0)\|_F^2}{\sum_{n=0}^{2} \|\mathbf{H}(n)\|_F^2}$ and conduct a simulation study with respect to different ranges of η.

In the experiments, for the purpose of comparison, the composite channel vector $\mathbf{h}$ is first estimated by the LS method. As for the estimation of the pure discrete-

Table 4.1: Complexity comparison of time-domain and frequency-domain methods for computing $\hat{\mathbf{R}}_{\mathrm{T}}$

	Frequency-Domain Algorithm	Time-Domain Algorithm
Step 1	$\mathbf{R}_A(k) =$ $\begin{bmatrix} Y_1(k) \\ \vdots \\ Y_{N_R}(k) \end{bmatrix} \left[Y_1^*(k), \cdots, Y_{N_R}^*(k)\right],$ $(k = 0, \cdots, K-1)$: KN_R^2 multiplications	a K-length IFFT processing: $K\log_2(K) N_R$ multiplications $K\log_2(K) N_R$ additions
Step 2	$\mathbf{R}_B(k) = \left[\mathbf{I}_{N_R} \otimes \mathbf{F}_2^T(k)\right] \mathbf{R}_A(k)$, $(k = 0, \cdots, K-1)$: $KN_R^2(P+1)$ multiplications	$\mathbf{R}_D(n) = \mathbf{y}_{P+1}(n)\, \mathbf{y}_{P+1}^H(n)$, $(n = 0, \cdots, K-1)$: $KN_R^2(P+1)^2$ multiplications
Step 3	$\mathbf{R}_C(k) = \mathbf{R}_B(k) \left[\mathbf{I}_{N_R} \otimes \mathbf{F}_2^*(k)\right]$, $(k = 0, \cdots, K-1)$: $KN_R^2(P+1)^2$ multiplications	$\hat{\mathbf{R}}_T = \frac{1}{K} \sum_{n=0}^{K-1} \mathbf{R}_D(n)$: $N_R^2(P+1)^2$ multiplications $(K-1) N_R^2(P+1)^2$ additions
Step 4	$\hat{\mathbf{R}}_{\mathrm{F}} = \frac{1}{K} \sum_{k=0}^{K-1} \mathbf{R}_C(k)$: $N_R^2(P+1)^2$ multiplications $(K-1) N_R^2(P+1)^2$ additions	
In Total	Multiplications: $(K+1)N_R^2(P+1)^2 + N_R^2(P+2)$ Additions: $(K-1) N_R^2(P+1)^2$	Multiplications: $K\log_2(K) N_R + (K+1) N_R^2(P+1)^2$ Additions: $K\log_2(K) N_R + (K-1) N_R^2(P+1)^2$

time channel vector $\mathbf{d}$, we consider the time-domain LS, the frequency-domain LS and the proposed frequency-domain semi-blind method, all with pulse shaping. For simplicity, we call these four methods as the basic LS, enhanced time-domain LS (TDLS), enhanced frequency-domain LS (FDLS) and enhanced semi-blind methods. Note that the LS methods can be easily obtained by setting α to zero in the proposed two semi-blind methods with pulse-shaping. The estimation performance is evaluated in terms of the MSE of the estimate of the channel matrix given by

$$\text{MSE} = \frac{1}{N_{\text{MC}}} \sum_{n=1}^{N_{\text{MC}}} \left\| \hat{\mathbf{h}}_n - \mathbf{h}_n \right\|^2 \quad \left(\text{or MSE} = \frac{1}{\text{N}_{\text{MC}}} \sum_{\text{n=1}}^{\text{N}_{\text{MC}}} \left\| \hat{\mathbf{d}}_{\text{n}} - \mathbf{d}_{\text{n}} \right\|^2 \right)$$

where N_{MC} is the number of Monte Carlo iterations, and $\mathbf{h}_n$ ($\mathbf{d}_n$) and $\hat{\mathbf{h}}_n$ ($\hat{\mathbf{d}}_n$) are the true and the estimated channel vectors with respect to the n-th Monte Carlo iteration, respectively.

Experiment A1: MSE versus η

In the first experiment, the channel estimation performance in terms of the MSE versus η is investigated. The simulation is undertaken by 1000 Monte Carlo runs of the transmission of one OFDM symbol under an SNR of 15dB at 512 subcarriers of which 30 are used as pilot for training purpose. Fig. 4.3 shows the MSE plots resulting from the proposed enhanced TDLS, FDLS and semi-blind frequency-domain methods as well as the basic LS estimation with 20, 40, 60 and 80 symbols, respectively. Obviously, the MSE performance of the proposed three enhanced methods are significantly better than that of the basic LS. In particular, the performance of the enhanced TDLS and FDLS methods with only one symbol is still better than that

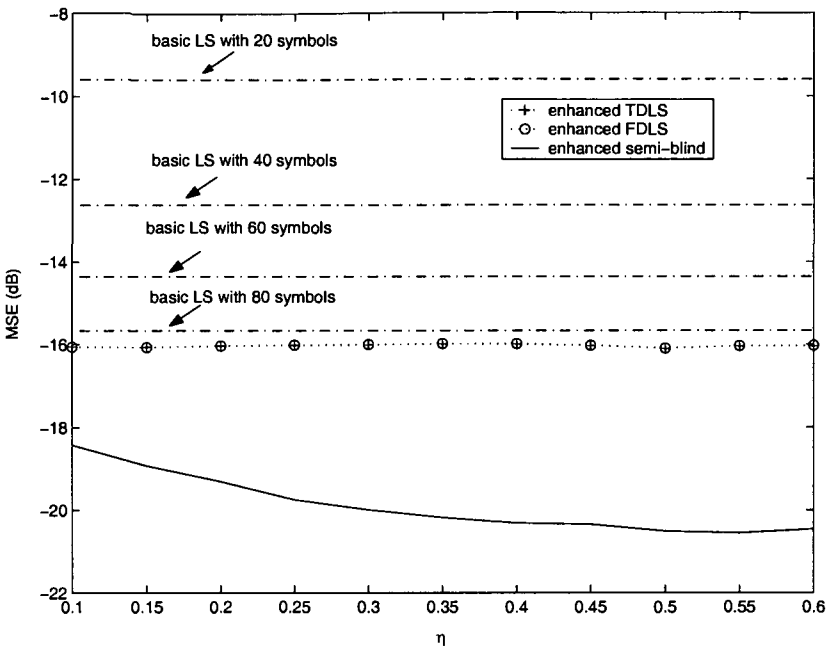

Figure 4.3: MSE versus η for a sampling duration-based channel.

of the basic LS method with 80 symbols, indicating that the new approach focusing on the pure multipath channel vector significantly outperforms other techniques for the composite channel vector irrespective of pulse-shaping. One can find a high consistency between the enhanced TDLS and the enhanced FDLS methods. Also, the semi-blind frequency-enhanced method significantly outperforms the two enhanced LS methods. In addition, the performance of the semi-blind method improves with the increasing value of η when $\eta < 0.3$, and remains almost the same when η is in the range of 0.3 to 0.6, which represents typical mobile communication scenarios where the first arrived path is comparable to or stronger than other paths.

Experiment A2: MSE versus SNR

Now we investigate the channel estimation performance versus the SNR. The sim-

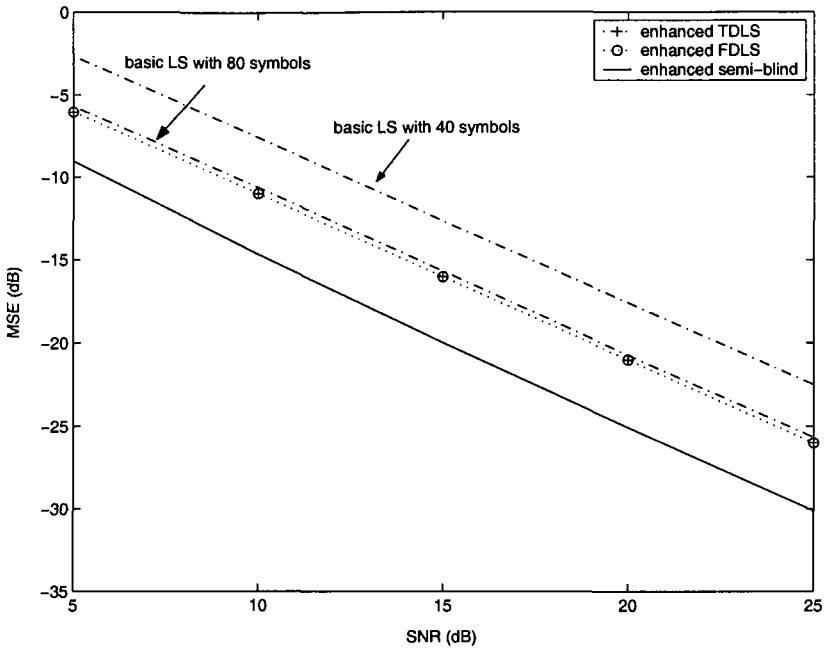

Figure 4.4: MSE versus SNR for a sampling duration-based channel

ulation involves 5000 Monte Carlo runs of the transmission of one OFDM symbol. Fig. 4.4 shows the channel estimation results of the three enhanced methods and that of the basic LS method with 40 and 80 symbols when $\eta > 0.2$. It is seen that the performances of the two LS methods are almost the same, which consistently outperform the basic LS method with 80 symbols. In addition, it is observed that the semi-blind frequency-enhanced method can achieve a nearly 3$\sim$4 dB gain over the two enhanced LS methods, when the SNR varies from 5 to 25 dB, respectively.

Experiment A3: MSE versus pilot length

Here we investigate the channel estimation performance of the semi-blind frequency-domain algorithm versus the number of pilot subcarriers per symbol, in comparison with that of the two enhanced LS methods. The number of OFDM symbols used in

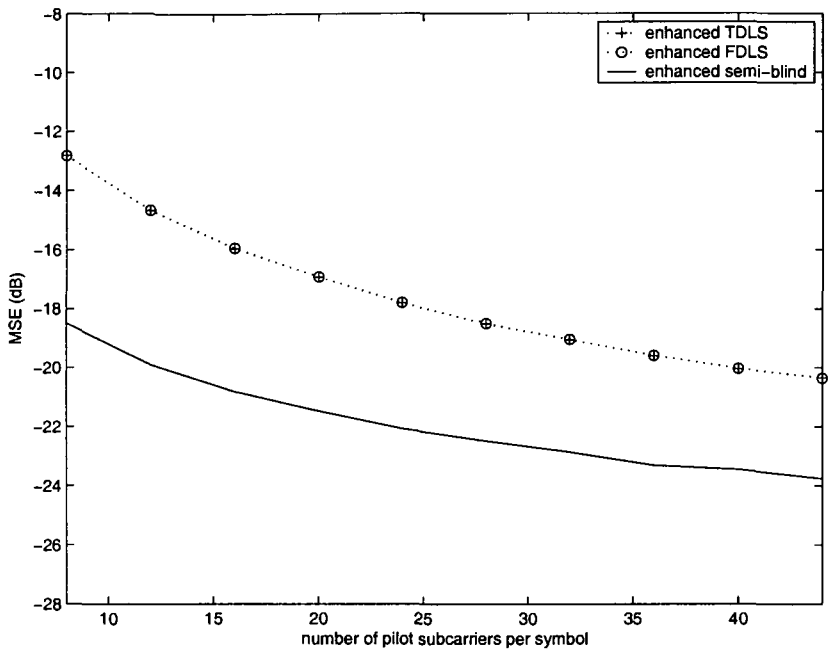

Figure 4.5: MSE versus pilot length for a sampling duration-based channel

the three methods is set to 2, and the number of pilot subcarriers per OFDM symbol varies from 8 to 48. Fig. 4.5 shows the MSE plots from 500 Monte Carlo iterations at an SNR of 15 dB when $\eta > 0.2$. It is seen that the performance of all the algorithms is improved with an increasing number of pilot subcarriers. Again, the performance of the proposed TDLS and FDLS methods are almost the same, and the semi-blind frequency-domain method is superior to both TDLS and FDLS nearly by 6 and 4dB when the number of pilot subcarriers is 8 and 48, respectively. It implies that the proposed semi-blind method is more advantageous for pilot signals of a shorter length.

Experiment A4: BER versus SNR

In this experiment, the BER performance of the MIMO-OFDM system is investigated by using the estimated channel matrix and an ordered vertical-Bell laboratories

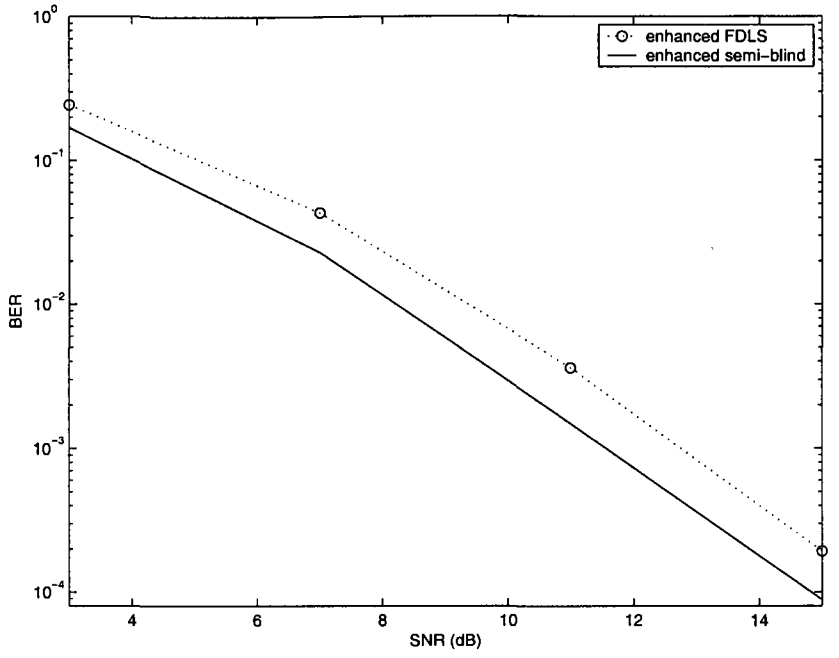

Figure 4.6: BER versus SNR for a sampling duration-based channel

layered space time (V-BLAST) decoder. The simulation involves 5000 Monte Carlo runs of the transmission of one OFDM symbol with 8 pilot subcarriers. Fig.4.6 shows the BER performance versus the SNR for the FDLS method and the semi-blind method, when $\eta > 0.2$. It is seen that the performance of the semi-blind method is superior to the LS method by 1.6~3.4 dB when the SNR varies in the range of $3 \sim 15$dB.

4.4.2 Upsampling Duration-based Channels

A 3-path Rayleigh channel is assumed, in which each path corresponds to a 2×4 random matrix whose elements are i.i.d. complex Gaussian variables with zero mean and unit variance. The delays of the three paths are set to 0, $\frac{3}{4}T$, and $2T$.

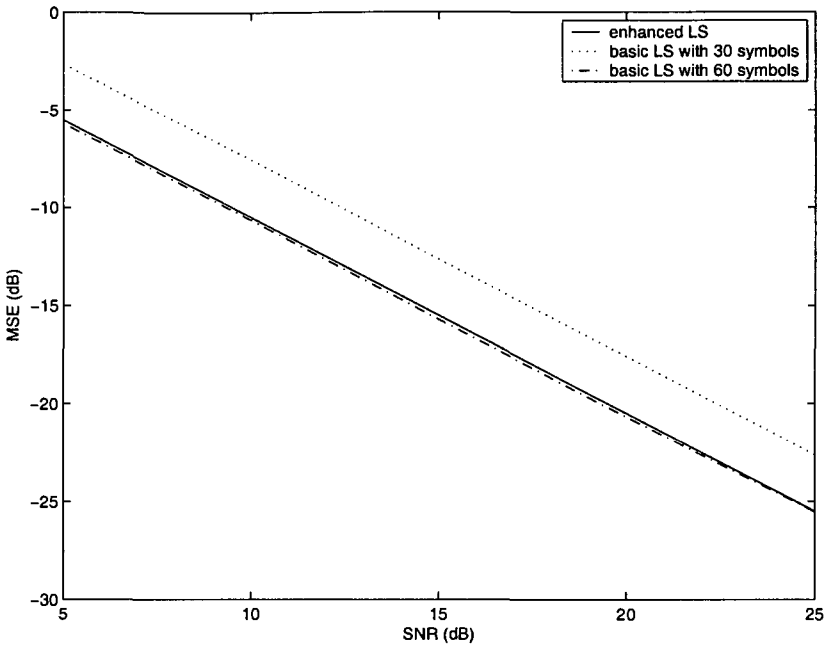

Figure 4.7: MSE versus SNR for an upsampling duration-based channel.

Experiment B1: MSE versus SNR

At first, we investigate the channel estimation performance versus the SNR. The simulation involves 5000 Monte Carlo runs of the transmission of one OFDM symbol. Fig. 4.7 shows the channel estimation result of the enhanced LS method and that of the basic LS method with 30 and 60 symbols. It is seen that the enhanced LS method with only one symbol can achieve 2.9 dB gain over the basic LS method with 30 symbols. Moreover, it is observed that 60 symbols are needed for the basic LS method to achieve the same performance as given by the enhanced LS method.

Experiment B2: MSE versus pilot length

Now we examine the channel estimation performance of the enhanced LS algorithm versus the number of pilot subcarriers per symbol. The number of OFDM

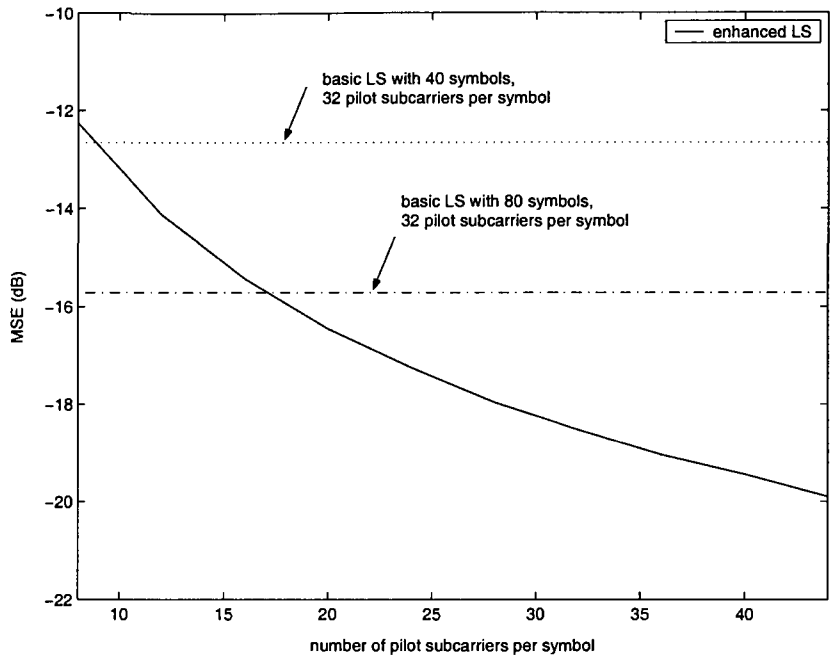

Figure 4.8: MSE versus pilot length for an upsampling duration-based channel.

symbols used is set to 2, and the number of pilot subcarriers per OFDM symbol varies from 8 to 48. Fig. 4.8 shows the MSE plots from 5000 Monte Carlo iterations at an SNR of 15 dB when $\eta > 0.2$. It is seen that the performance of the enhanced LS algorithm is improved with the increase of the number of pilot subcarriers. In particular, the enhanced LS algorithm with 18 pilots (2 symbols with 9 pilot subcarriers) can achieve a similar result of the basic LS method with 1280 pilots (40 symbols with 32 pilot subcarriers). In addition, 36 pilots in the enhanced LS method serves as 2560 pilots in the basic LS method.

4.5 Conclusions

In this chapter, the channel estimation issue of MIMO-OFDM systems with pulse shaping has been thoroughly studied. By taking into account the effect of both pulse-shaping filter in the transmitter and the matched filter in the receiver, a new channel estimation problem focusing on the estimation of the pure multipath channel has been formulated. Then, two typical multipath scenarios corresponding to the sampling- and upsampling-duration based channels have been identified, for which two frequency-domain channel estimation approaches, i.e., the semi-blind and the pilot-aided least-square methods, have been developed, respectively. To reduce the computational complexity of the semi-blind method, a frequency-domain correlation matrix estimation algorithm has also been presented. The effectiveness of the new channel estimation mthods have been confirmed by computer simulations with comparison to the basic LS method as well as its enhanced versions with pulse-shaping.

Chapter 5

Nulling-based Semi-Blind Sparse Channel Estimation

5.1 Introduction

It is well known that a wireless channel can often be modelled as a sparse channel, in which the delay spread could be very large but the number of paths is normally very small [52,65–70,96–100]. Broadly speaking, there are two kinds of approaches for the sparse channel estimation. The first one estimates the complex amplitude and the delay of each path based on a non-sampling spaced parametrical channel modelling [52,96,97]. These parameters can be determined through various means depending on the underlying wireless systems. For example, they can be obtained by employing a special spreading code in code division multiple access (CDMA) systems [96] or using the ESPRIT (the estimation of signal parameters via rotational invariance technique) with an estimated delay-subspace in OFDM systems [52, 97]. The second kind is based on the sparsity assumption of the equivalent discrete-time channel [66–70, 98–102], in which only a few taps in the long tapped delay line are considered most

significant. By exploiting the sparse structure of the channel, some improved channel estimation algorithms have been developed for OFDM systems [65–68] and CDMA systems [69, 70]. In this chapter, we are concerned with the development of a very efficient sparse channel estimation approach for MIMO-OFDM systems that requires only a few OFDM symbols. As the non-sampling spaced sparse channel estimation requires a large number of OFDM symbols for the estimation of the delay-subspace, we focus only on the sampling-spaced approach.

It should be mentioned that almost all of the sampling-spaced sparse channel estimation methods in the literature utilize a training sequence and follow two steps: (1) detect the position of the most significant taps (MSTs), which are also referred to as the nonzero taps in some of the literature; and (2) obtain an improved channel estimate by exploiting the position of the MSTs. The key to these methods is the first step, i.e., the MST detection. Several techniques based on an LS estimation have been developed for MST detection [65–67, 98]. In these methods, an unstructured channel, i.e., a channel presumably having all nonzero taps, is considered in the initial LS estimation. Using the preliminary estimate, the MST detection can simply be conducted by choosing a number of the largest amplitude channel taps [65]. To improve the detection accuracy, a generalized Akaike information criterion (GAIC) is utilized to estimate the MSTs in an iterative fashion [66], which is then simplified to a non-iterative scheme [67]. Another MST detection technique that utilizes the preliminary LS estimate in conjunction with the so called on-off-keying (OOK) detection method can be found in [98]. The matching pursuit (MP) based MST detection method has also extensively been investigated by many researchers [69, 70, 99, 100].

In the MP method, the MST is detected by sequentially selecting the column in a mixture matrix of the pilot signal that matches best the residual vector of the received signal. The detection performance of the MP method can be improved by utilizing the subspace fitting [103] or a group of OFDM symbols [104]. Another MST detection, which employs an active tap detection criterion involving the correlation between the training pilots and the received signal, has been proposed in [101, 102].

Once the MST information has been acquired in the first step, a refined channel estimate can be obtained in the second step by using the structured channel, in which only the detected MSTs are taken into account. This is usually implemented with the structured version of the training-based approaches such as the LS method [66, 67, 69, 98, 99] and the LMS method [101, 102]. It has been shown that the performance of these sparse channel estimation methods is much better than that of the unstructured channel estimation alternatives.

The common problem of the above mentioned sparse channel estimation methods is that a large number of pilots is needed in order to render an accurate MST detection and channel estimation. To increase the spectral efficiency, the available information of users data could be applied to both the MST detection and the channel estimation. Unfortunately, very little work on blind MST detection and blind sparse channel estimation is found in the existing literatures. To the best knowledge of the author, there is only one cyclic-prefix (CP)-based blind method dealing with the MST detection of the sparse channel estimation for OFDM systems [68], in which a blind estimation of the sparse channel is, however, not considered. Moreover, this detection scheme needs a large number of OFDM symbols as well as a large CP length in order

to obtain precise MST positions.

The objective of this chapter is to develop, for MIMO-OFDM systems, an efficient semi-blind sparse channel estimation approach, which comprises the MST detection and sparse channel estimation both in a semi-blind fashion. First, the analysis of the second-order statistics of the received signal passing through a sparse channel is conducted, leading to a relationship between the positions of the most significant taps and the lags of the nonzero correlation matrices of the received signal. Based on this relationship, an efficient MST detection algorithm, which requires only a few OFDM symbols and a very small number of pilots, is proposed by utilizing the lag information of the nonzero correlation matrices. Then, by using the acquired MST position, an LS criterion and a blind constraint on the sparse channel vector with respect to the most significant taps are derived to develop a semi-blind sparse channel estimation approach. In addition, a brief perturbation analysis of the proposed approach is conducted, showing that the new semi-blind sparse solution is not subject to the signal perturbation error in the M-rate sparse channel case, i.e., the sparse channel is a decimated version of the full FIR channel under consideration.

The rest of the chapter is organized as follows. Section 5.2 performs an analysis of the second-order statistics of the received signal through a sparse channel, leading to a relationship between the positions of the most significant taps and the lags of the nonzero correlation matrices of the received signal. Section 5.3 proposes a highly efficient MST detection algorithm that requires only a few OFDM symbols and a small number of pilots for the least-square based detection. Section 5.4 proposes a semi-blind approach for the estimation of sparse MIMO-OFDM channels, including

the formulation of a training-based LS criterion and a blind constraint on the sparse channel vector with respect to the most significant taps. Section 5.5 conducts a perturbation analysis of the proposed sparse semi-blind approach, justifying that a blind constraint that is not affected by signal perturbation terms can be obtained for an M-rate sparse channel. Section 5.6 comprises a number of experiments validating the proposed approach, showing the significant advantage of the sparse semi-blind solution over the sparse LS method as well as the regular LS and semi-blind techniques in terms of the MSE of the channel estimate. Finally, Section 5.7 concludes the chapter by highlighting some of the contributions presented.

5.2 Second-Order Statistics of the Received Signal through Sparse MIMO Channel

A wireless channel can very often be modelled as a sparse channel that contains many zero taps in the uniform delay line [66–70, 98–102]. In this case, the channel matrix with respect to the d-th ($d = 0, 1, \cdots, D-1$) most significant tap (namely, the nonzero tap in the noise free case) can be expressed as

$$\mathbf{Z}(d) = \mathbf{H}(l_d) \tag{5.1}$$

where l_d ($d = 0, 1, \cdots, D-1$) are integers with $0 = l_0 < l_1 < \cdots < l_{D-1}$. To distinguish from $\mathbf{H}(l)$, $\mathbf{Z}(d)$ is referred to as the effective channel matrix. In the following, we analyze the second-order statistics of the signal received through the sparse MIMO channel.

It is well known that the correlation matrix of the received signal vector $\mathbf{y}(n)$

plays a crucial role in blind or semi-blind channel estimation [36,105,106], which can be in general defined as

$$\mathbf{R}(l) \triangleq \mathrm{E}\left\{\mathbf{y}(n)\,\mathbf{y}^H(n-l)\right\},\ (l=0,1,\cdots,P). \tag{5.2}$$

Obviously, (5.2) includes the autocorrelation matrix of $\mathbf{y}(n)$ as a special case when $l=0$. In this section, we would like to express $\mathbf{R}(l)$ in terms of the effective sparse channel matrix $\mathbf{Z}(d)$, $(d=0,1,\cdots,D-1)$, in the absence of noise, and show that $\mathbf{R}(l)$ has only a few most significant lags (MSLs), i.e., most of the matrices $\mathbf{R}(l)$ with $l \in [0,P]$ are zero matrices, due to the sparse nature of the channel.

Using (2.3), (2.32), (2.33) and (5.1) into (5.2), we can obtain

$$\mathbf{R}(l) = \mathbf{Z}_A \mathbf{R}_{x,D}(l)\,\mathbf{Z}_A^H \tag{5.3}$$

where

$$\mathbf{Z}_A \triangleq \begin{bmatrix} \mathbf{Z}(0) & \mathbf{Z}(1) & \cdots & \mathbf{Z}(D-1) \end{bmatrix} \tag{5.4}$$

$$\mathbf{R}_{x,D}(l) \triangleq \mathrm{E}\left\{ \begin{bmatrix} \mathbf{x}(n) \\ \mathbf{x}(n-l_1) \\ \vdots \\ \mathbf{x}(n-l_{D-1}) \end{bmatrix} \begin{bmatrix} \mathbf{x}(n-l) \\ \mathbf{x}(n-l_1-l) \\ \vdots \\ \mathbf{x}(n-l_{D-1}-l) \end{bmatrix}^H \right\}. \tag{5.5}$$

Clearly, the structure of $\mathbf{R}(l)$ mainly depends on $\mathbf{R}_{x,D}(l)$. Due to the fact that $\mathrm{E}\left\{\mathbf{x}(n-i)\,\mathbf{x}^H(n-j)\right\} = \sigma_x^2 \delta(i-j)\,\mathbf{I}$, we can rewrite (5.5) as

$$\mathbf{R}_{x,D}(l) = \mathbf{\Lambda}_D(l) \otimes \sigma_x^2 \mathbf{I} \tag{5.6}$$

where

$$\mathbf{\Lambda}_D(l) = \begin{bmatrix} \delta(l) & \delta(l+l_1) & \cdots & \delta(l+l_{D-1}) \\ \delta(l-l_1) & \delta(l) & \cdots & \delta(l+l_{D-1}-l_1) \\ \vdots & \vdots & \ddots & \vdots \\ \delta(l-l_{D-1}) & \delta(l-l_{D-1}+l_1) & \cdots & \delta(l) \end{bmatrix}. \tag{5.7}$$

It is obvious from (5.7) that $\mathbf{\Lambda}_D(l)$ is a lower triangular matrix for $l \in [0, P]$. Moreover, the nonzero elements of $\mathbf{\Lambda}_D(l)$ occur only when $l = l_i - l_j$, $(i, j = 0, 1, \cdots, D-1;$ $i \geq j)$. Since the channel is sparse, i.e., $D \ll L \leq P$, there is only a small number of choices of l which makes $\mathbf{\Lambda}_D(l)$ a nonzero matrix. In other words, $\mathbf{\Lambda}_D(l)$ is a zero matrix for most of the values of l ranging from 0 to P. It is clear from (5.3) and (5.6) that as long as $\mathbf{\Lambda}_D(l)$ is a zero matrix, $\mathbf{R}_{x,D}(l)$ is a zero matrix and so is $\mathbf{R}(l)$. Moreover, even if $\mathbf{\Lambda}_D(l)$ is a nonzero matrix, most of its lower triangular elements are zero. Therefore, $\mathbf{\Lambda}_D(l)$ is a sparse matrix, which makes $\mathbf{R}_{x,D}(l)$ also sparse. Note that the position of the nonzero elements of $\mathbf{\Lambda}_D(l)$, and in turn that of the corresponding nonzero submatrices of $\mathbf{R}_{x,D}(l)$ depend on the values of l and l_d. We now show that since the nonzero submatrices of $\mathbf{R}_{x,D}(l)$ are simply identity matrices multiplied by σ_x^2, $\mathbf{R}(l)$ with respect to its MSLs can be expressed in terms of the effective matrix $\mathbf{Z}(d)$.

Let us consider first the simplest case when $D = 2$. In this case, there are only two nonzero effective channel matrices $\mathbf{Z}(0)$ and $\mathbf{Z}(1)$ corresponding to $\mathbf{H}(0)$ and $\mathbf{H}(l_1)$, respectively. Without loss of generality, we assume a unit signal variance, i.e., $\sigma_x^2 = 1$. From (5.3), (5.4) and (5.6), one can find that there are only two nonzero

matrices of $\mathbf{R}(l)$, i.e.,

$$\mathbf{R}(0) = \begin{bmatrix} \mathbf{Z}(0) & \mathbf{Z}(1) \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{Z}^H(0) \\ \mathbf{Z}^H(1) \end{bmatrix} = \mathbf{Z}(0)\mathbf{Z}^H(0) + \mathbf{Z}(1)\mathbf{Z}^H(1), \quad (5.8)$$

$$\mathbf{R}(l_1) = \begin{bmatrix} \mathbf{Z}(0) & \mathbf{Z}(1) \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{Z}^H(0) \\ \mathbf{Z}^H(1) \end{bmatrix} = \mathbf{Z}(1)\mathbf{Z}^H(0), \quad (5.9)$$

which means that the MSL position of $\mathbf{R}(l)$ is $l = 0, l_1$.

When $D = 3$, $\mathbf{R}_{x,3}(l)$ has different sparse structures depending on the relationship between l_1 and l_2, which leads $\mathbf{R}(l)$ to have different expressions. Using(5.3), (5.4) and (5.6), one can obtain the following results.

Case D3.1: if $l_2 = 2l_1$,

$$\mathbf{R}(l) = \begin{cases} \sum_{i=0}^{2} \mathbf{Z}(i)\mathbf{Z}^H(i) & \text{if } l = 0 \\ \mathbf{Z}(1)\mathbf{Z}^H(0) + \mathbf{Z}(2)\mathbf{Z}^H(1) & \text{if } l = l_1 \\ \mathbf{Z}(2)\mathbf{Z}^H(0) & \text{if } l = l_2 \\ \mathbf{0} & \text{otherwise} \end{cases}; \quad (5.10)$$

Case D3.2: if $l_2 \neq 2l_1$,

$$\mathbf{R}(l) = \begin{cases} \sum_{i=0}^{2} \mathbf{Z}(i)\mathbf{Z}^H(i) & \text{if } l = 0 \\ \mathbf{Z}(1)\mathbf{Z}^H(0) & \text{if } l = l_1 \\ \mathbf{Z}(2)\mathbf{Z}^H(1) & \text{if } l = l_2 - l_1 \\ \mathbf{Z}(2)\mathbf{Z}^H(0) & \text{if } l = l_2 \\ \mathbf{0} & \text{otherwise} \end{cases}. \quad (5.11)$$

It is seen that the main difference between the two cases *D3.1* and *D3.2* lies in the number of MSLs in addition to the expression of $\mathbf{R}(l)$. The two common MSLs in the

two cases are $\mathbf{R}(0)$ and $\mathbf{R}(l_2)$. Other MSLs depend on the relationship between l_1 and l_2. If l_1 is not identical to $l_2 - l_1$, for any value of l, $\delta(l - l_1)$ and $\delta(l - l_2 + l_1)$ in (5.7) cannot be unity at the same time, leading to two different MSLs, i.e., $\mathbf{R}(l_1) = \mathbf{Z}(1)\mathbf{Z}^H(0)$ and $\mathbf{R}(l_2 - l_1) = \mathbf{Z}(2)\mathbf{Z}^H(1)$. However, if l_1 equals $l_2 - l_1$, both $\delta(l - l_1)$ and $\delta(l - l_2 + l_1)$ could be unity simultaneously. Then, the two presumable MSLs overlap, yielding the only MSL, i.e., $\mathbf{R}(l = l_1 = l_2 - l_1) = \mathbf{Z}(1)\mathbf{Z}^H(0) + \mathbf{Z}(2)\mathbf{Z}^H(1)$.

In the case of $D = 4$, one can easily obtain $\mathbf{R}(l)$ with respect to the first and the last MSLs as

$$\mathbf{R}(0) = \sum_{i=0}^{3} \mathbf{Z}(i)\mathbf{Z}^H(i), \tag{5.12}$$

$$\mathbf{R}(l_3) = \mathbf{Z}(3)\mathbf{Z}^H(0), \tag{5.13}$$

which is similar to the cases of $D = 2$ and $D = 3$. All other MSLs can be determined from the relationship amongst l_1, l_2, $l_2 - l_1$, $l_3 - l_1$ and $l_3 - l_2$. As some of these values could be identical as discussed for the case of $D = 3$, the number of the MSLs of $\mathbf{R}(l)$ can be different. As such, the position of the MSLs and the expression of $\mathbf{R}(l)$ are dependent on the values of l_1, l_2, l_3 and their differences. Using (5.3), (5.4), (5.6) and (5.7), one can have a total of 8 possible cases as summarized in Table 5.1, where $\mathbf{R}_{i,j} \triangleq \mathbf{Z}(i)\mathbf{Z}^H(j)$. Note that the correlation matrices with respect to the first and the last MSLs, which are common to all the 8 cases, are not included in Table 5.1 for notational simplicity. Evidently, *Case D4.8* indicates a possibility of having a maximum of 7 MSLs of $\mathbf{R}(l)$ including the first and the last ones. When some of the values of l_1, l_2, $l_2 - l_1$, $l_3 - l_1$ and $l_3 - l_2$ happen to be identical, the number of the MSLs would be reduced. For example, *Case D4.2* corresponds to a possibility that

Table 5.1: Expression of correlation matrix $\mathbf{R}(l)$ for $D = 4$

Case D4.1:	l	l_1	l_2			
$l_1 : l_2 : l_3 = 1 : 2 : 3$	$\mathbf{R}(l)$	$\mathbf{R}_{1,0} + \mathbf{R}_{2,1} + \mathbf{R}_{3,2}$	$\mathbf{R}_{2,0} + \mathbf{R}_{3,1}$			
Case D4.2:	l	l_1	l_2	$l_3 - l_1$		
$l_1 : l_2 : l_3 = 1 : 2 : 4$	$\mathbf{R}(l)$	$\mathbf{R}_{1,0} + \mathbf{R}_{2,1}$	$\mathbf{R}_{2,0} + \mathbf{R}_{3,2}$	$\mathbf{R}_{3,1}$		
Case D4.3:	l	l_1	l_2	$l_2 - l_1$		
$l_1 : l_2 : l_3 = 2 : 3 : 4$	$\mathbf{R}(l)$	$\mathbf{R}_{1,0} + \mathbf{R}_{3,1}$	$\mathbf{R}_{2,0}$	$\mathbf{R}_{2,1} + \mathbf{R}_{3,2}$		
Case D4.4: $l_3 = l_1 + l_2$	l	l_1	l_2	$l_2 - l_1$		
$l_2 \neq 2l_1$, $l_1 : l_2 \neq 2 : 3$	$\mathbf{R}(l)$	$\mathbf{R}_{1,0} + \mathbf{R}_{3,2}$	$\mathbf{R}_{2,0} + \mathbf{R}_{3,1}$	$\mathbf{R}_{2,1}$		
Case D4.5: $l_3 \neq l_1 + l_2$	l	l_1	l_2	$l_3 - l_1$	$l_3 - l_2$	
$l_2 = 2l_1$, $l_3 \neq 2l_2$	$\mathbf{R}(l)$	$\mathbf{R}_{1,0} + \mathbf{R}_{2,1}$	$\mathbf{R}_{2,0}$	$\mathbf{R}_{3,1}$	$\mathbf{R}_{3,2}$	
Case D4.6: $l_3 \neq l_1 + l_2$	l	l_1	l_2	$l_2 - l_1$	$l_3 - l_1$	
$l_2 \neq 2l_1$, $2l_2 = l_3 + l_1$	$\mathbf{R}(l)$	$\mathbf{R}_{1,0}$	$\mathbf{R}_{2,0}$	$\mathbf{R}_{2,1} + \mathbf{R}_{3,2}$	$\mathbf{R}_{3,1}$	
Case D4.7: $l_1 : l_3 = 1 : 2$	l	l_1	l_2	$l_2 - l_1$	$l_3 - l_2$	
$l_1 : l_2 \neq 2 : 3$	$\mathbf{R}(l)$	$\mathbf{R}_{1,0} + \mathbf{R}_{3,1}$	$\mathbf{R}_{2,0}$	$\mathbf{R}_{2,1}$	$\mathbf{R}_{3,2}$	
Case D4.8: $l_3 \neq l_1 + l_2$, $l_2 \neq 2l_1$	l	l_1	l_2	$l_2 - l_1$	$l_3 - l_1$	$l_3 - l_2$
$l_1 : l_2 \neq 2 : 3$, $2l_2 \neq l_3 + l_1$	$\mathbf{R}(l)$	$\mathbf{R}_{1,0}$	$\mathbf{R}_{2,0}$	$\mathbf{R}_{2,1}$	$\mathbf{R}_{3,1}$	$\mathbf{R}_{3,2}$

the MSLs at l_1 and $l_2 - l_1$ are overlapped as well as the MSLs at l_2 and $l_3 - l_2$ are overlapped.

We have obtained in the above the expression of the correlation matrices $\mathbf{R}(l)$, $(l = 0, 1, \cdots, l_{D-1})$ in terms of the effective channel matrix $\mathbf{Z}(d)$, $(d = 0, 1, \cdots, D-1)$ for the cases with respect to $D = 2, 3, 4$. When $D > 4$, a table similar to Table 5.1 can easily be designed by a simple computer programming. Noting that $l_0 = 0$, all the potential MSLs can be written as $\epsilon_{ij} = l_i - l_j \, (i > j)$, each corresponding to $\mathbf{R}_{i,j} = \mathbf{R}(l = l_i - l_j)$. If any two of ϵ_{ij}'s are identical, the corresponding two potential MSLs are overlapped, and thus the corresponding two $\mathbf{R}_{i,j}$'s are combined.

As seen in the next section, the structure of $\mathbf{R}(l)$ can be exploited to detect the MSTs of the sparse channel. In Section 5.4, we will use the expression of $\mathbf{R}(l)$ at its MSLs to derive a blind constraint for the semi-blind estimation of the sparse channel.

5.3 Detection of Most Significant Taps (MSTs)

The MST detection is usually conducted by utilizing some training sequences. The available MST detection methods include the GAIC-, OOK- and MP-based methods [66, 67, 69, 70, 98–100, 103, 104]. These methods in general require a large number of pilots to render an accurate MST detection. In addition to training-based methods, a CP-assisted blind approach was developed for OFDM systems in [68]. This method still needs a large number of OFDM symbols plus a large CP length in order to obtain a satisfactory detection result. We now propose an efficient MST detection algorithm based on the structure of $\mathbf{R}(l)$ as analyzed in the previous section. The idea is to determine the MST position based on the MSL position of $\mathbf{R}(l)$ [107, 108]. It will be shown through simulation studies that, by using very few OFDM symbols plus only a small number of pilots, our new MST detection method has a high detection accuracy.

Prior to developing the new method, we first introduce the estimated version $\hat{\mathbf{R}}(l)$ of $\mathbf{R}(l)$. For the sake of simplicity, only one OFDM symbol is considered. The correlation matrix of $\mathbf{y}(n)$ can be estimated as

$$\hat{\mathbf{R}}(l) = \frac{1}{K} \sum_{n=0}^{K-1} \mathbf{y}(n) \mathbf{y}^H(n-l) \tag{5.14}$$

where $\mathbf{y}(n) = \mathbf{y}(K+n)$ for $n < 0$. By letting

$$\mathbf{x}_D(n) \triangleq \left[\mathbf{x}^T(n-l_0), \mathbf{x}^T(n-l_1), \cdots, \mathbf{x}^T(n-l_{D-1})\right]^T, \ (n = 0, 1, \cdots, K-1)$$

where $\mathbf{x}(n) = \mathbf{x}(K+n)$ for $n < 0$, the received signal vector $\mathbf{y}(n)$ as described by the circular convolution (2.3) in the noisy case can be rewritten in the matrix form as

$$\mathbf{y}(n) = \mathbf{Z}_A \mathbf{x}_D(n) + \mathbf{v}(n) \tag{5.15}$$

where $\mathbf{v}(n) \triangleq [v_1(n), v_1(n), \cdots, v_{N_R}(n)]^T$. Substituting (5.15) into (5.14) and using (5.3) yield

$$\hat{\mathbf{R}}(l) = \mathbf{R}(l) + \mathbf{Z}_A \Delta\mathbf{R}_{x,D}(l) \mathbf{Z}_A^H + \mathbf{R}_v \tag{5.16}$$

where

$$\Delta\mathbf{R}_{x,D}(l) \triangleq \frac{1}{K} \sum_{n=0}^{K-1} \mathbf{x}_D(n) \mathbf{x}_D^H(n-l) - \mathbf{R}_{x,D}(l) \tag{5.17}$$

represents the signal perturbation and $\mathbf{R}_v$ is the perturbation error introduced by the noise.

Obviously, the first term in the right hand side (RHS) of (5.16) is an ideal correlation matrix of the received signal vector $\mathbf{y}(n)$ without the signal perturbation as well as the noise corruption, while the second and the third terms are the errors introduced by the signal and noise perturbations, respectively. Note that in the noisy case, $\hat{\mathbf{R}}(l)$ is in general a nonzero matrix even if $\mathbf{R}(l) = 0$. However, the MSLs can be detected from the norm of $\hat{\mathbf{R}}(l)$, since the first term in the RHS of (5.16), $\mathbf{R}(l)$, would be dominant provided that the signal-to-noise ratio (SNR) is not extremely low. Accordingly, the first step of our method is to detect the MSLs of $\hat{\mathbf{R}}(l)$ by

comparing its norm with a threshold, which is defined in this thesis as

$$\eta = \frac{K_e}{P_L + 1} \sum_{l=0}^{P_L} \| \hat{\mathbf{R}}(l) \| \tag{5.18}$$

where P_L is a predetermined length and the coefficient K_e is used to adjust the average norm of $\mathbf{R}(l)$. Once the MSLs of $\hat{\mathbf{R}}(l)$ are acquired, their position information can be utilized to determine the MST position l_d of the sparse channel via the MSL of $\mathbf{R}(l)$ obtained in the previous subsection. The second step of the new MST detection algorithm can be described as follows.

Assume a total of W MSLs of $\hat{\mathbf{R}}(l)$ have been detected, whose positions are denoted as m_i with $m_0 < m_1 < \cdots < m_{W-1}$. We now determine the possible l_d's according to the structures of $\mathbf{R}(l)$ with respect to different values of D and W. Although we have in general $D \leq W$, it is clear that the positions of the first and the last MSTs are simply given by

$$l_0 = m_0 = 0, \tag{5.19}$$

$$l_{D-1} = m_{W-1}. \tag{5.20}$$

Obviously, (5.19) and (5.20) give the only two nonzero taps if $W = 2$. In the following, we determine l_d $(d = 1, 2, \cdots, D-2)$ from m_w $(w = 1, 2, \cdots, W-2)$ for the cases of $W > 2$.

If $W = 3$, we have only the case *D3.1* according to (5.10). Then, there is only one additional tap l_d to be determined, which is readily given by

$$l_1 = m_1 = \frac{m_2}{2}.$$

If $W = 4$, we should have case *D3.2* from (5.11) or *D4.1* in Table 5.1. Then, one can have the following three possible solutions,

Case W4.1: when $D = 3$ with $l_2 > 2l_1$, we have $m_1 = l_1$ and $m_2 = l_2 - l_1$, which gives $l_1 = m_1$ and $l_2 = m_1 + m_2 = m_3$;

Case W4.2: when $D = 3$ with $l_2 < 2l_1$, we obtain $m_1 = l_2 - l_1$ and $m_2 = l_1$, which yields $l_1 = m_2$ and $l_2 = m_1 + m_2 = m_3$;

Case W4.3: when $D = 4$ with $l_1 : l_2 : l_3 = 1 : 2 : 3$, we readily have $l_1 = m_1$ and $l_2 = m_2$.

In order to make a decision among the three choices, we propose a sparse LS method as described below.

By letting l'_d $(d = 0, 1, \cdots, D' - 1)$ be the potential MSTs of the sparse channel matrix and using (5.1), (2.28) can be rewritten as

$$\mathbf{Y}_{i_R,\text{pilot}}(m) = \sum_{i_T=1}^{N_T} \mathbf{X}_{i_T,\text{pilot-diag}}(m)\, \mathbf{F}'(m)\, \mathbf{z}'_{i_R,i_T} + \boldsymbol{\xi}_{i_R,\text{pilot}}(m) \tag{5.21}$$

where $\mathbf{z}'_{i_R,i_T} = \left[z'_{i_R,i_T}(0), \cdots, z'_{i_R,i_T}(D'-1)\right]^T$ and $\mathbf{F}'(m)$ is a $K_P \times D'$ matrix, whose d-th column is the l_d-th column of $\mathbf{F}(m)$, $(d = 0, 1, \cdots, D' - 1)$. By further letting

$$\mathbf{z}'_{i_R} \triangleq \left[\left(\mathbf{z}'_{i_R,1}\right)^T, \cdots, \left(\mathbf{z}'_{i_R,N_T}\right)^T\right]^T, \tag{5.22}$$

$$\mathbf{Z}' \triangleq \left[\mathbf{z}'_1, \cdots, \mathbf{z}'_{N_R}\right], \tag{5.23}$$

(3.2) can be rewritten as

$$\bar{\mathbf{Y}}_{\text{pilot}} = \mathbf{A}'\mathbf{Z}' + \bar{\boldsymbol{\xi}}_{\text{pilot}} \tag{5.24}$$

where

$$\mathbf{A}' \triangleq \begin{bmatrix} \mathbf{X}_{1,\text{pilot-diag}}(0)\,\mathbf{F}'(0) & \cdots & \mathbf{X}_{N_T,\text{pilot-diag}}(0)\,\mathbf{F}'(0) \\ \vdots & \ddots & \vdots \\ \mathbf{X}_{1,\text{pilot-diag}}(g-1)\,\mathbf{F}'(g-1) & \cdots & \mathbf{X}_{N_T,\text{pilot-diag}}(g-1)\,\mathbf{F}'(g-1) \end{bmatrix}. \tag{5.25}$$

From (5.24), an LS estimate of the sparse channel with respect to the potential MSTs, l'_d $(d = 0, 1, \cdots, D' - 1)$ can be obtained as

$$\hat{\mathbf{Z}}' = (\mathbf{A}')^{\dagger} \bar{\mathbf{Y}}_{\text{pilot}}. \tag{5.26}$$

Using the idea of joint channel estimation and zero detection proposed in [98], a cost function for the new MST detection method can be established as

$$J\left(l'_0, \cdots, l'_{D'-1}\right) = \left\| \bar{\mathbf{Y}}_{\text{pilot}} - \mathbf{A}'\hat{\mathbf{Z}}' \right\|_F^2. \tag{5.27}$$

Evidently, the cost J reaches its minimum when l'_d $(d = 0, 1, \cdots, D' - 1)$ gives true MSTs.

In the case of $W = 4$, our MST detection only needs to calculate a few costs with respect to Cases *W4.1*, *W4.2* and *W4.3*, denoted as $J_{4,1}$, $J_{4,2}$ and $J_{4,3}$. The complete scheme is shown in Fig.5.1, which gives three possible MST detection results, namely,

Case W4.1: $l_1 = m_1$; *Case W4.2:* $l_1 = m_2$ and *Case W4.3:* $l_1 = m_1, l_2 = m_2$.

It should be stressed that the detection scheme in Fig.5.1 is already the improved version of a straightforward implementation. First, instead of computing all the three costs $J_{4,1}$, $J_{4,2}$ and $J_{4,3}$ and making a detection among the three values, a test on the three MSLs m_1, m_2 and m_3 has been performed to possibly reduce the three candidates to two, which could not only save the calculation of one cost but also enhance the detection performance by excluding the false MST candidate *Case W4.3* from comparison. Second, when the test on the three candidates is passed, we have found that $J_{4,3}$ could be very close to $J_{4,1}$ ($J_{4,2}$) if *Case W4.1* (*Case W4.2*) is true, indicating a possibility of mis-detecting *Case W4.1* (*Case W4.2*) as *Case W4.3* at a low level of SNR. On the other hand, when *Case W4.3* is true, $J_{4,3}$ is found to be

much smaller than $J_{4,1}$ and $J_{4,2}$. In the improved detection scheme, therefore, we have employed a scaling coefficient c to distinguish *W4.3* from *W4.1* and *W4.2*. The value of c should be chosen according to the level of SNR. For example, for a moderate SNR, we have found $c = 0.8 \sim 0.9$ is a proper choice.

Similarly, for $W = 5$, one can have the following possible situations:

Case W5.1: $m_1 = l_1$, $m_2 = l_2$ and $m_3 = l_3 - l_1$, when $D = 4$ and $l_1 : l_2 : l_3 = 1 : 2 : 4$;

Case W5.2: $m_1 = l_2 - l_1$, $m_2 = l_1$ and $m_3 = l_2$, when $D = 4$ and $l_1 : l_2 : l_3 = 2 : 3 : 4$;

Case W5.3: $m_1 = l_1$, $m_2 = l_2 - l_1$ and $m_3 = l_2$, when $D = 4$, $l_3 = l_1 + l_2$ and $l_2 > 2l_1$;

Case W5.4: $m_1 = l_2 - l_1$, $m_2 = l_1$ and $m_3 = l_2$, when $D = 4$, $l_3 = l_1 + l_2$ and $l_2 < 2l_1$;

Case W5.5: $m_1 = l_1$, $m_2 = l_2$ and $m_3 = l_3$, when $D = 5$ and $l_1 : l_2 : l_3 : l_4 = 1 : 2 : 3 : 4$.

The corresponding MST detection scheme is described in Fig.5.2, which gives the following five detection results.

Case W5.1: $l_1 = m_1, l_2 = m_2$; *Case W5.2:* $l_1 = m_2, l_2 = m_3$; *Case W5.3:* $l_1 = m_1, l_2 = m_3$;

Case W5.4: $l_1 = m_2, l_2 = m_3$; and *Case W5.5:* $l_1 = m_1, l_2 = m_2, l_3 = m_3$.

Clearly, the key to the above detection scheme is to design a table listing all the possible MST candidates corresponding to the same value of W using the expression of $\mathbf{R}(l)$ obtained in the previous subsection. This table can easily be constructed

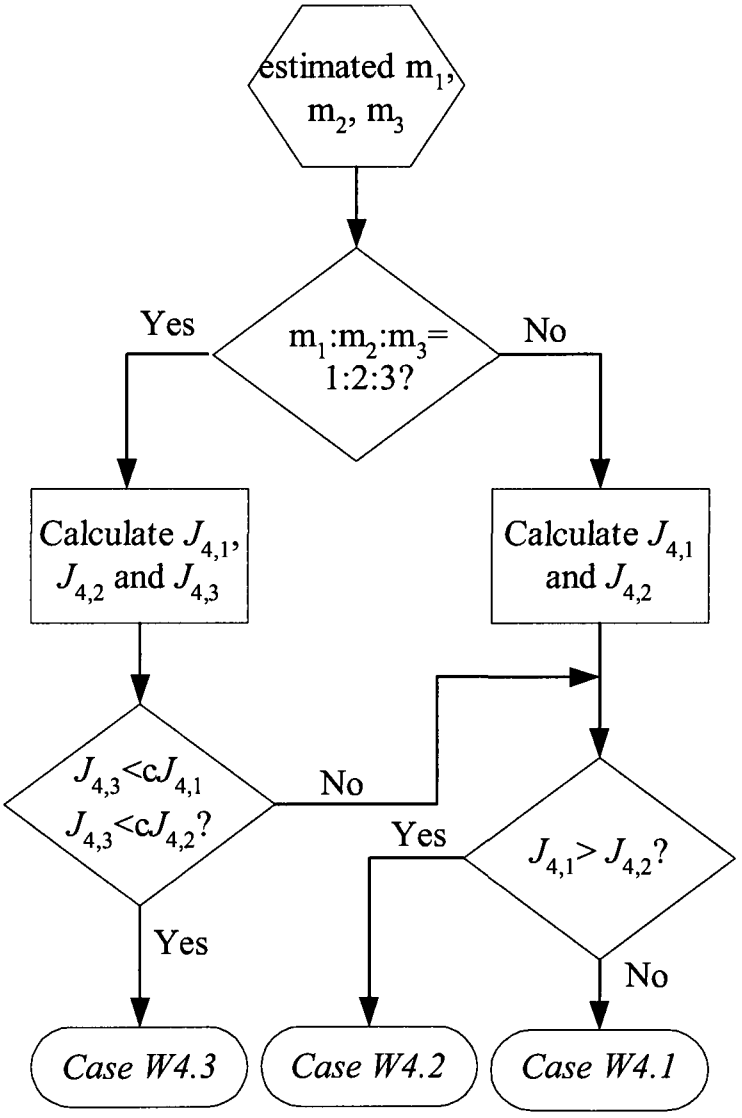

Figure 5.1: MST detection scheme for $W = 4$

even if W has a large value, since the expression of $\mathbf{R}(l)$ can simply be derived by computer programming as stated before. In comparison with many of the MST detection methods available in the literature, which normally require to search all the possible MST candidates over the entire channel length L, the proposed detection method reduces the search range to a very small number of MST candidates, by exploiting the MSL information of $\hat{\mathbf{R}}(l)$. More importantly, the conventional methods need a large number of pilots or OFDM symbols to conduct a training-based or CP-assisted MST detection. In contrast, our technique uses only a few OFDM symbols with a small number of pilots to achieve an accurate detection.

5.4 Semi-Blind Estimation of the Sparse Channel

In this section, we extend our previously developed nulling-based MIMO-OFDM semi-blind approach for the sparse channel estimation. It will be shown that, by using the available user data, the sparse semi-blind approach significantly outperforms the training-based sparse method.

The key idea of the sparse semi-blind method is to derive a blind constraint for the effective channel matrix $\mathbf{Z}(d)$ that is to be incorporated in the training-based LS cost function [109]. As shown in Chapters 3 and 4, the blind constraint in the semi-blind cost function is derived from the correlation matrices of the received signal, i.e., $\tilde{\mathbf{R}}_{n-1}$

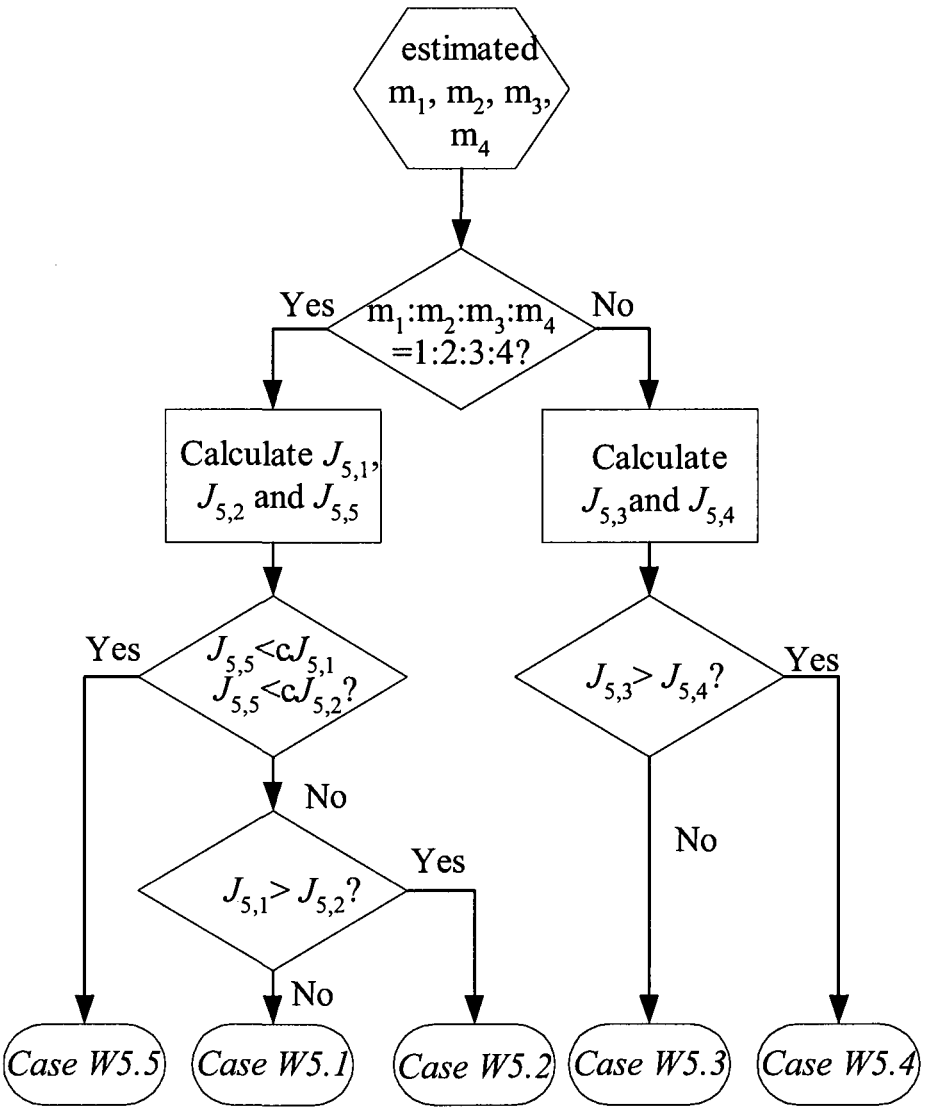

Figure 5.2: MST detection scheme for $W = 5$

in (2.37) and $\ddot{\mathbf{R}}_n$ in (2.38), namely,

$$\tilde{\mathbf{R}}_{n-1} = \begin{bmatrix} \mathbf{R}(0) & \mathbf{R}(1) & \cdots & \mathbf{R}(P-1) \\ \mathbf{R}(-1) & \mathbf{R}(0) & \cdots & \mathbf{R}(P-2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}(1-P) & \mathbf{R}(2-P) & \cdots & \mathbf{R}(0) \end{bmatrix}, \tag{5.28}$$

$$\ddot{\mathbf{R}}_n = \begin{bmatrix} \mathbf{R}(1) & \mathbf{R}(2) & \cdots & \mathbf{R}(P) \end{bmatrix}. \tag{5.29}$$

Theorem 5.1: In the absence of noise, $\mathbf{R}(l)$ can be expressed in terms of $\mathbf{H}(l)$ as

$$\mathbf{R}(l) = \begin{cases} \sum_{i=l}^{L-1} \mathbf{H}(i)\,\mathbf{H}^H(i-l), & l = 0, 1, \cdots, L-1 \\ \mathbf{0}, & l > L-1 \end{cases}. \tag{5.30}$$

The proof of this theorem is given in Appendix C.

It should be noted that (5.30) is the corner stone of the linear prediction-based blind or semi-blind channel estimation method. In a manner similar to obtaining (5.30), one can construct the matrices

$$\mathbf{R}_{\mathbf{Z}}(d) = \sum_{i=d}^{D-1} \mathbf{Z}(i)\,\mathbf{Z}^H(i-d) = \sum_{i=d}^{D-1} \mathbf{R}_{i,i-d}, \quad (d = 0, 1, \cdots, D-1)\ . \tag{5.31}$$

Letting

$$\tilde{\mathbf{R}}_{\mathbf{Z},n-1} \triangleq \begin{bmatrix} \mathbf{R}_{\mathbf{Z}}(0) & \mathbf{R}_{\mathbf{Z}}(1) & \cdots & \mathbf{R}_{\mathbf{Z}}(D-2) \\ \mathbf{R}_{\mathbf{Z}}(-1) & \mathbf{R}_{\mathbf{Z}}(0) & \cdots & \mathbf{R}_{\mathbf{Z}}(D-3) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{\mathbf{Z}}(2-D) & \mathbf{R}_{\mathbf{Z}}(3-D) & \cdots & \mathbf{R}_{\mathbf{Z}}(0) \end{bmatrix}, \tag{5.32}$$

$$\ddot{\mathbf{R}}_{\mathbf{Z},n} \triangleq \begin{bmatrix} \mathbf{R}_{\mathbf{Z}}(1) & \mathbf{R}_{\mathbf{Z}}(2) & \cdots & \mathbf{R}_{\mathbf{Z}}(D-1) \end{bmatrix}, \tag{5.33}$$

a constraint on $\mathbf{Z}(d)$ $(d = 0, 1, \cdots, D-1)$ can be obtained by using the linear prediction method shown in Section 3.2. Therefore, once the matrices $\mathbf{R}_{\mathbf{Z}}(d)$ $(d = 0, 1, \cdots, D-1)$ are available, a semi-blind estimation problem for the effective channel matrix $\mathbf{Z}(d)$ $(d = 0, 1, \cdots, D-1)$ can be formulated.

Assuming that the MSTs have been correctly estimated, we now discuss the estimation of $\mathbf{R}_{\mathbf{Z}}(d)$ $(d = 0, 1, \cdots, D-1)$ with respect to different values of D by utilizing the analysis results of Section 5.2. Our objective here is to express $\hat{\mathbf{R}}_Z(d)$ in terms of $\hat{\mathbf{R}}(l)$. For the simplest case when $D = 2$, it is clear from (5.8) and (5.9) that

$$\hat{\mathbf{R}}_{\mathbf{Z}}(0) = \hat{\mathbf{R}}(0), \tag{5.34}$$

$$\hat{\mathbf{R}}_{\mathbf{Z}}(1) = \hat{\mathbf{R}}(l_1). \tag{5.35}$$

When $D = 3$, utilizing (5.10) and (5.11), one can estimate $\mathbf{R}_{\mathbf{Z}}(d)$ $(d = 0, 1, 2)$ as follows:

$$\hat{\mathbf{R}}_{\mathbf{Z}}(0) = \hat{\mathbf{R}}(0), \tag{5.36}$$

$$\hat{\mathbf{R}}_{\mathbf{Z}}(1) = \hat{\mathbf{R}}(l_1) + \hat{\mathbf{R}}(l_2 - l_1)\left[1 - \delta(l_2 - 2l_1)\right], \tag{5.37}$$

$$\hat{\mathbf{R}}_{\mathbf{Z}}(2) = \hat{\mathbf{R}}(l_2). \tag{5.38}$$

In a similar manner, when $D = 4$, from (5.12) and (5.13), we have

$$\hat{\mathbf{R}}_{\mathbf{Z}}(0) = \hat{\mathbf{R}}(0), \tag{5.39}$$

$$\hat{\mathbf{R}}_{\mathbf{Z}}(3) = \hat{\mathbf{R}}(l_3). \tag{5.40}$$

Using Table 5.1, the estimates $\hat{\mathbf{R}}_{\mathbf{Z}}(1) = \hat{\mathbf{R}}_{1,0} + \hat{\mathbf{R}}_{2,1} + \hat{\mathbf{R}}_{3,2}$ and $\hat{\mathbf{R}}_{\mathbf{Z}}(2) = \hat{\mathbf{R}}_{2,0} + \hat{\mathbf{R}}_{3,1}$ in terms of $\hat{\mathbf{R}}(l)$ with respect to different cases of $D = 4$ can be attained as shown

Table 5.2: Estimation of correlation matrix $\mathbf{R}_{\mathbf{Z}}(d)$ for different cases of $D = 4$

	$\hat{\mathbf{R}}_{\mathbf{Z}}(0)$	$\hat{\mathbf{R}}_{\mathbf{Z}}(1)$	$\hat{\mathbf{R}}_{\mathbf{Z}}(2)$	$\hat{\mathbf{R}}_{\mathbf{Z}}(3)$
Case D4.1:	$\hat{\mathbf{R}}(0)$	$\hat{\mathbf{R}}(l_1)$	$\hat{\mathbf{R}}(l_2)$	$\hat{\mathbf{R}}(l_3)$
Case D4.2:	$\hat{\mathbf{R}}(0)$	$\hat{\mathbf{R}}(l_1) + \hat{\mathbf{R}}(l_2) - \hat{\mathbf{R}}_{2,0}$	$\hat{\mathbf{R}}_{2,0} + \hat{\mathbf{R}}(l_3 - l_1)$	$\hat{\mathbf{R}}(l_3)$
Case D4.3:	$\hat{\mathbf{R}}(0)$	$\hat{\mathbf{R}}_{1,0} + \hat{\mathbf{R}}(l_2 - l_1)$	$\hat{\mathbf{R}}(l_1) - \hat{\mathbf{R}}_{1,0} + \hat{\mathbf{R}}(l_2)$	$\hat{\mathbf{R}}(l_3)$
Case D4.4:	$\hat{\mathbf{R}}(0)$	$\hat{\mathbf{R}}(l_1) + \hat{\mathbf{R}}(l_2 - l_1)$	$\hat{\mathbf{R}}(l_2)$	$\hat{\mathbf{R}}(l_3)$
Case D4.5:	$\hat{\mathbf{R}}(0)$	$\hat{\mathbf{R}}(l_1) + \hat{\mathbf{R}}(l_3 - l_2)$	$\hat{\mathbf{R}}(l_2) + \hat{\mathbf{R}}(l_3 - l_1)$	$\hat{\mathbf{R}}(l_3)$
Case D4.6:	$\hat{\mathbf{R}}(0)$	$\hat{\mathbf{R}}(l_1) + \hat{\mathbf{R}}(l_2 - l_1)$	$\hat{\mathbf{R}}(l_2) + \hat{\mathbf{R}}(l_3 - l_1)$	$\hat{\mathbf{R}}(l_3)$
Case D4.7:	$\hat{\mathbf{R}}(0)$	$\hat{\mathbf{R}}_{1,0} + \hat{\mathbf{R}}(l_2 - l_1) + \hat{\mathbf{R}}(l_3 - l_2)$	$\hat{\mathbf{R}}(l_1) - \hat{\mathbf{R}}_{1,0} + \hat{\mathbf{R}}(l_2)$	$\hat{\mathbf{R}}(l_3)$
Case D4.8:	$\hat{\mathbf{R}}(0)$	$\hat{\mathbf{R}}(l_1) + \hat{\mathbf{R}}(l_2 - l_1) + \hat{\mathbf{R}}(l_3 - l_2)$	$\hat{\mathbf{R}}(l_2) + \hat{\mathbf{R}}(l_3 - l_1)$	$\hat{\mathbf{R}}(l_3)$

in Table 5.2. Note that for Cases *D4.2*, *D4.3* and *D4.7* in Table 5.2, besides $\hat{\mathbf{R}}(l)$, $\hat{\mathbf{R}}_{1,0}$ or $\hat{\mathbf{R}}_{2,0}$ is required. In our method, both $\hat{\mathbf{R}}_{1,0}$ and $\hat{\mathbf{R}}_{2,0}$ can be calculated based on the estimate of $\mathbf{Z}(d)$ using the sparse LS method. The above discussion can be easily extended to the case of a larger value of D.

Once $\hat{\mathbf{R}}_{\mathbf{Z}}(d)$ $(d = 0, 1, \cdots, D-1)$ are obtained, $\hat{\tilde{\mathbf{R}}}_{\mathbf{Z},n-1}$ and $\hat{\tilde{\mathbf{R}}}_{\mathbf{Z},n}$ can be constructed by using (5.32) and (5.33); these can be further used to estimate a blind constraint $\hat{\mathbf{B}}_{\mathbf{Z}}$ on the effective channel vector $\mathbf{z} \triangleq \mathrm{vec}(\mathbf{Z})$, where $\mathbf{Z}$ is formed by the true MSTs obtained from (5.23). Then, a semi-blind sparse channel estimation problem similar to (3.17) can be formulated as

$$\min_{\hat{\mathbf{z}}} \Delta = \left\| \mathbf{Y}_{\text{pilot}} - \tilde{\mathbf{A}}_{\mathbf{Z}} \hat{\mathbf{z}} \right\|_F^2 + \alpha \left\| \hat{\mathbf{B}}_{\mathbf{Z}} \hat{\mathbf{z}} \right\|_F^2 \tag{5.41}$$

where $\tilde{\mathbf{A}}_{\mathbf{Z}} = \mathbf{I} \otimes \mathbf{A}'$, with $\mathbf{A}'$ being given by (5.25) and corresponding to the true

MSTs. The solution to this optimization problem is given by

$$\hat{\mathbf{z}} = \left[\tilde{\mathbf{A}}_{\mathbf{Z}}^{H}\tilde{\mathbf{A}}_{\mathbf{Z}} + \alpha\hat{\mathbf{B}}_{\mathbf{Z}}^{H}\hat{\mathbf{B}}_{\mathbf{Z}}\right]^{\dagger}\tilde{\mathbf{A}}_{\mathbf{Z}}^{H}\mathbf{Y}_{\text{pilot}}. \tag{5.42}$$

The proposed semi-blind sparse channel estimation approach can be summarized as follows:

1. MSLs are first estimated by comparing the norm of the correlation matrices with the threshold calculated from (5.18).

2. The estimated MSLs are utilized, together with a small number of pilots, for the estimation of MSTs.

3. Based on the estimated MSTs, the channel coefficients are estimated in a semi-blind fashion as given by (5.42).

It will be shown in Section 5.6 that the performance of the proposed sparse channel estimation approach is significantly superior to that of the original semi-blind method. Clearly, the complexity of the sparse estimation method is much lower than that of the original version for non-sparse channels due to $D \ll L$. Typically, the computational complexity of the original semi-blind algorithm is in the order of $O(L^3)$, while that for the proposed sparse approach is only $O(D^3)$.

5.5 A Case Study for M-Rate Sparse Channel Estimation

It is known that the linear prediction- or subspace-based channel estimation methods are always perturbed by various sources, such as finite data length and measurement noise [71, 72]. In our previous chapters, we have successfully applied the per-

turbation theory to the analysis of MIMO and MIMO-OFDM channel estimation. In this section, we evaluate the performance of the proposed sparse channel estimation approach from the perspective of the perturbation analysis [110].

We consider only the perturbation due to the finite data length in the computation of correlation matrices. In the noise-free case, (5.16) can be rewritten as

$$\hat{\mathbf{R}}(l) = \mathbf{Z}_A \mathbf{R}_{x,D}(l) \mathbf{Z}_A^H + \mathbf{Z}_A \Delta\mathbf{R}_{x,D}(l) \mathbf{Z}_A^H \tag{5.43}$$

where $\Delta\mathbf{R}_{x,D}(l)$ is given by (5.17), which represents an error term introduced by the signal perturbation. Based on (5.43), we would like to investigate the signal perturbation error of the blind constraint $\hat{\mathbf{B}}_{\mathbf{Z}}$ in the proposed semi-blind approach. To this end, we need to derive the perturbation term $\Delta\mathbf{R}_{\mathbf{Z}}(d)$ of $\mathbf{R}_{\mathbf{Z}}(d)$ by employing $\Delta\mathbf{R}_{x,D}(l)$. Let us consider first the case of an "M-rate" sparse channel case, i.e., the sparse channel is a decimated version of the full-length FIR channel by a factor of M,

$$\mathbf{Z}(d) = \mathbf{H}(dM). \tag{5.44}$$

In this case, (5.17) is still applicable except that $\mathbf{x}_D(n)$ is replaced by

$$\mathbf{x}_D(n) = \left[\begin{array}{cccc} \mathbf{x}^H(n) & \mathbf{x}^H(n-M) & \cdots & \mathbf{x}^H(n-(D-1)M) \end{array} \right]^H. \tag{5.45}$$

On the other hand, from (5.30), (5.31) and (5.44), we have

$$\mathbf{R}_{\mathbf{Z}}(d) = \mathbf{R}(dM), \text{ for } d = 0, 1, \cdots, D-1. \tag{5.46}$$

Using (5.43) and (5.46), we obtain

$$\Delta\mathbf{R}_{\mathbf{Z}}(d) = \Delta\mathbf{R}(dM) = \mathbf{Z}_A \Delta\mathbf{R}_{x,D}(dM) \mathbf{Z}_A^H. \tag{5.47}$$

By defining $\mathbf{x}_M(n) \triangleq \mathbf{x}(nM)$ and using (5.3), (5.17) and (5.45), (5.47) can be rewritten as

$$\Delta\mathbf{R}_{\mathbf{Z}}(d) = \mathbf{Z}_A \left[\frac{1}{K} \sum_{n=0}^{K-1} \mathbf{x}_{MD}(n)\, \mathbf{x}_{MD}^H(n-d) \right] \mathbf{Z}_A^H - \mathbf{R}_{\mathbf{Z}}(d) \tag{5.48}$$

where

$$\mathbf{x}_{MD}(n) = \left[\begin{array}{cccc} \mathbf{x}_M^H(n) & \mathbf{x}_M^H(n-1) & \cdots & \mathbf{x}_M^H(n-D+1) \end{array} \right]^H. \tag{5.49}$$

Interestingly, (5.48) and (5.49) indicate that $\Delta\mathbf{R}_{\mathbf{Z}}(d)$ can be viewed as the perturbation term of the correlation matrix of the received signal through a regular FIR channel characterized by $\mathbf{Z}(l)$, $(l = 0, 1, \cdots, D-1)$ with respect to the transmitted signal $\mathbf{x}_M(n)$. Thus, the perturbation analysis of the regular channel estimation method in Section 3.3 can be directly applied to the M-rate sparse channel case. As a result, an ideal nulling constraint on the effective channel vector $\mathbf{z}$ is obtained, leading to a signal-perturbation-free sparse channel estimation. Moreover, the practical scheme of computing the weighting factor α suggested in Chapter 3 can also be used.

In a non-M-rate sparse channel case, since the expression of $\Delta\mathbf{R}_{\mathbf{Z}}(d)$ as given in (5.48) and (5.49) is not available, an ideal nulling constraint on the channel vector $\mathbf{z}$ cannot be obtained, even in the absence of noise. Therefore, the non-M-rate sparse channel estimation is in general subject to signal perturbation error. Recall that to determine the weighting factor α, the $\mathrm{MSE_B}$ needs to be calculated, which, as shown in Chapter 3, can easily be implemented online by using a closed-form expression. In the non-M-rate sparse channel estimation, it is difficult to obtain a closed-form expression for $\mathrm{MSE_B}$ due to the existence of the signal perturbation error. Considering that the $\mathrm{MSE_B}$ varies only slightly with the change of the channel, however, one can

estimate the MSE_B using a training signal, and then compute the weighting factor α off-line for the semi-blind sparse channel estimation.

5.6 Simulation Results

We consider a MIMO-OFDM system with 2 transmit and 4 receive antennas. The number of subcarriers is set to 1024, the length of cyclic prefix is 30, and the length of the linear predictor in the semi-blind algorithm is $P = L$. In our simulation, the QPSK modulation is used and a sparse Rayleigh channel modelled by a 3- or 4-nonzero tap MIMO-FIR filter is assumed, in which each tap corresponds to a 4×2 random matrix whose elements are i.i.d. complex Gaussian variables with zero mean and unit variance. For the non-M-rate sparse channel, the value of MSE_B is calculated off-line based on the true channel matrix by 500 Monte Carlo iterations. The constant K_e in (5.18) is set to 0.8.

For the purpose of comparison, the channel vector $\mathbf{h}$ is first estimated by the original LS and semi-blind methods. As for the estimation of the effective channel vector $\mathbf{z}$, we consider the proposed sparse LS method and sparse semi-blind method both with the MST detection. To evaluate the proposed sparse methods, an ideal sparse LS method and an ideal semi-blind method with the knowledge of the true MST information are also simulated. For easy citation, we call these four methods as the sparse LS, sparse semi-blind, ideal sparse LS and ideal sparse semi-blind methods. The estimation performance is evaluated in terms of the MSE of the estimate of the

channel matrix as given by

$$\text{MSE} = \frac{1}{N_{\text{MC}}} \sum_{n=1}^{N_{\text{MC}}} \left\| \hat{\mathbf{h}}_n - \mathbf{h}_n \right\|^2$$

where N_{MC} is the number of Monte Carlo iterations and $\mathbf{h}_n$ and $\hat{\mathbf{h}}_n$ are the true and the estimated channel vectors with respect to the n-th Monte Carlo iteration, respectively.

A simulation study with comparison is conducted for each of the following sparse channel models:

Channel A (Case D3.2 & Case W4.1): A 3-tap sparse channel with $l_0 = 0$, $l_1 = 4$ and $l_2 = 11$;

Channel B (Case D4.4 & Case W5.3): A 4-tap sparse channel with $l_0 = 0$, $l_1 = 6$, $l_2 = 14$ and $l_3 = 20$;

Channel C (Case D4.2 & Case W5.1): A 4-tap sparse channel with $l_0 = 0$, $l_1 = 5$, $l_2 = 10$ and $l_3 = 20$;

Channel D (Case D4.1 & Case W4.3): A 4-tap sparse channel with $l_0 = 0$, $l_1 = 4$, $l_2 = 8$ and $l_3 = 12$;

Channel E (Case D3.2 & Case W4.1): A 3-tap sparse channel with $l_0 = 0$, $l_1 = 4$ and $l_2 = 12$.

Experiment 1: *Channel A*

In the first experiment, the channel estimation performance in terms of the MSE as a function of the SNR is investigated for *Channel A*. The simulation involves 20000 Monte Carlo runs of the transmission of one OFDM symbol at 1024 subcarriers, of

Table 5.3: Error probability of the MST detection versus different SNR for *Channels A, B and C*

SNR(dB)	*Channel A*	*Channel B*	*Channel C*
5	2.7×10^{-3}	1.24×10^{-2}	1.47×10^{-2}
9	4×10^{-4}	1.9×10^{-3}	6×10^{-3}
13	3×10^{-4}	2×10^{-4}	4.8×10^{-3}
17	10^{-4}	2×10^{-4}	3.6×10^{-3}
21	10^{-4}	2×10^{-4}	3.8×10^{-3}

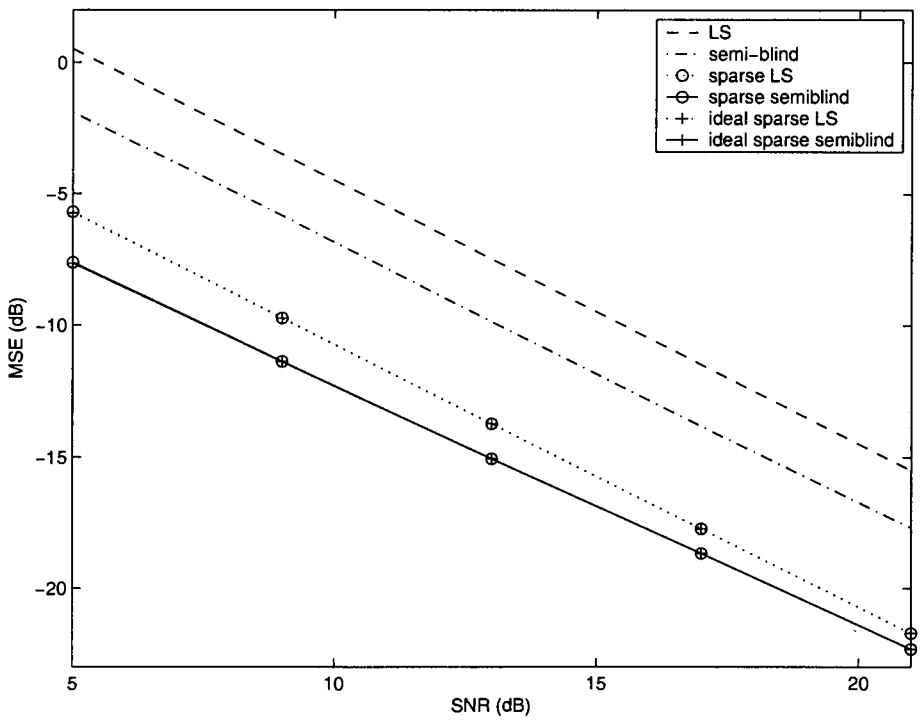

Figure 5.3: MSE versus SNR of proposed sparse LS and semi-blind algorithms for *Channel A*. Also shown for reference is the MSE of the original LS and semi-blind methods. One OFDM symbol with 30 pilot subcarriers.

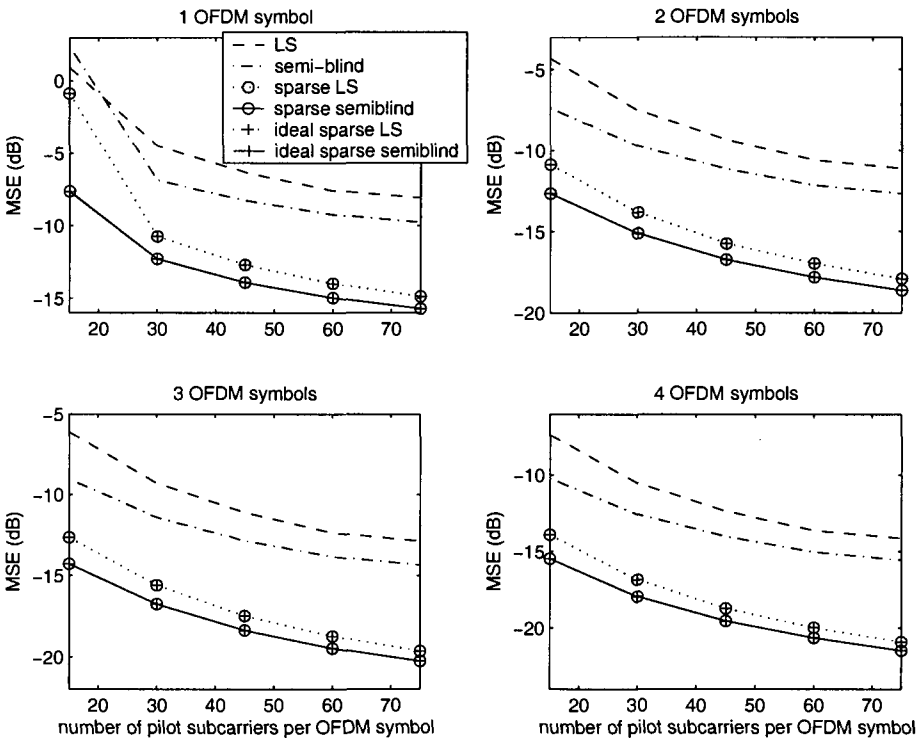

Figure 5.4: MSE versus pilot length for different numbers of OFDM symbols for *Channel A*. SNR=10 dB.

which 30 are used as pilot for training purpose. The error probability of the MST detection versus SNR is shown in the second column of Table 5.3. It is clear that a high performance of the MST detection is achieved despite only one OFDM symbol used. In particular, when SNR $\geq$ 17dB, the error probability is 10^{-4}. Fig. 5.3 shows the MSE plots from the proposed as well as the reference sparse methods along with two original methods. It is seen that the sparse LS method and the sparse semi-blind method are highly consistent with their ideal versions, respectively. It is also noted that the proposed two sparse methods significantly outperform the two original methods. Specifically, the sparse semi-blind method is superior to the original semi-blind method nearly by 5.7 dB and 4.6 dB when SNR is 5 dB and 21 dB, respectively. Moreover, the performance of the sparse semi-blind method is superior to the sparse LS method by 2 dB and 0.6 dB at the two SNR levels, respectively. It implies that the sparse semi-blind method is more advantageous for a lower SNR.

We now investigate the channel estimation performance of the proposed sparse methods versus the number of OFDM symbols as well as the number of pilot subcarriers per symbol. The number of OFDM symbols used is set to be from 1 to 4, and in each of the four cases, the number of pilot subcarriers per OFDM symbol varies from 15 to 75. Fig. 5.4 shows the MSE plots from 2000 Monte Carlo iterations for an SNR of 10 dB. It is seen that for the same number of OFDM symbols, the performance of all the algorithms is improved with increasing number of pilot subcarriers. Again, the two sparse LS methods have almost the same performance similar to the two sparse semi-blind methods. One can find that when only one OFDM symbol with 15 pilots is used, the sparse semi-blind method can achieve a gain of 6.8 $\sim$ 8.6 dB over the other

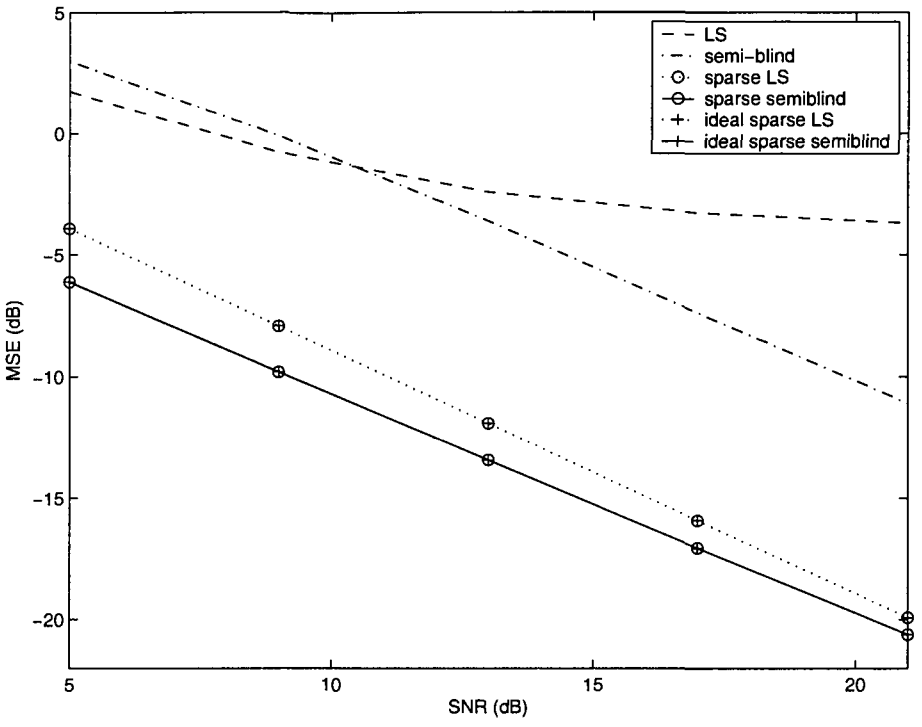

Figure 5.5: MSE versus SNR for *Channel B*. One OFDM symbol with 30 pilot sub-carriers.

methods. When the total number of pilots is not less than 30, it is observed that the performance improvement of the sparse semi-blind method over the sparse LS method remains almost the same with the increase of the number of OFDM symbols employed for channel estimation. Obviously, the proposed sparse semi-blind method is more advantageous for pilot signals of a shorter length.

Experiment 2: *Channel B*

Now, we examine the channel estimation performance as a function of the SNR for *Channel B*. The simulation involves 10000 Monte Carlo runs of the transmission of one OFDM symbol, of which 30 are used as pilot for training purpose. The error

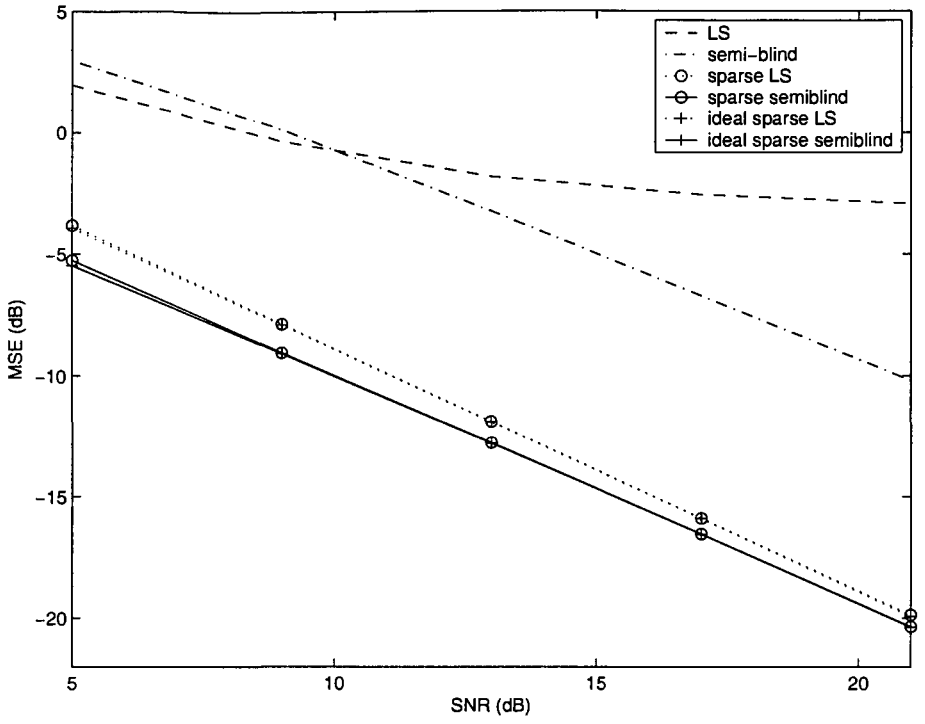

Figure 5.6: MSE versus SNR for *Channel C*. One OFDM symbol with 30 pilot sub-carriers.

probability of the MST detection versus SNR is also shown in Table 5.3, showing that the MST detection error is still very small when the channel length L and the number D of MSTs become 20 and 4, respectively. Fig. 5.5 shows the channel estimation results of the six methods. Again, the sparse LS method and the sparse semi-blind method each perform equally well to their respective ideal versions. It is seen that the sparse semi-blind method can achieve a gain of 9.1 $\sim$ 9.8 dB over the original semi-blind method, meaning that the sparse channel estimation approach is more advantageous for a larger channel length.

Experiment 3: *Channel C*

Using the same condition as in *Experiment* 2, the channel estimation result and the error probability of the MST detection versus SNR are shown in Fig. 5.6 and Table 5.3, respectively. It is seen that although the error probability of the MST detection is a little larger than that of the previous channel models, the performance of the sparse LS method and that of the sparse semi-blind method are still consistent with their ideal cases. This is due to the fact that the mis-detection of the 4-tap sparse channel with $l_0 = 0$, $l_1 = 5$, $l_2 = 10$ and $l_3 = 20$ as a 5-tap sparse channel with $l_0 = 0$, $l_1 = 5$, $l_2 = 10$, $l_3 = 15$ and $l_4 = 20$, still leads to a very good estimation performance. Also, one can find that the sparse semi-blind method can achieve a $0.5 \sim 1.4$ dB gain over the sparse LS method, confirming the effectiveness of calculating $\hat{\mathbf{R}}_{2,0}$ by using the sparse LS channel estimate.

Experiment 4: *Channel D*

Here, we investigate the estimation performance of the proposed method for an M-rate sparse channel. Fig. 5.7 shows the channel estimation performance versus the SNR obtained using the same condition as in *Experiment* 2. Again, the performances of the sparse LS and semi-blind methods are consistent with that of the ideal ones, which are significantly better than that of the original LS and semi-blind estimation methods. Furthermore, it is observed that the performance improvement of the sparse semi-blind method over the sparse LS method remains almost the same with the increase of SNR. This is different from the previous non-M-rate sparse channel examples, where the performance improvement decreases with the increase of SNR. This phenominon confirms our analysis result in Section 5.5, namely, the sparse semi-

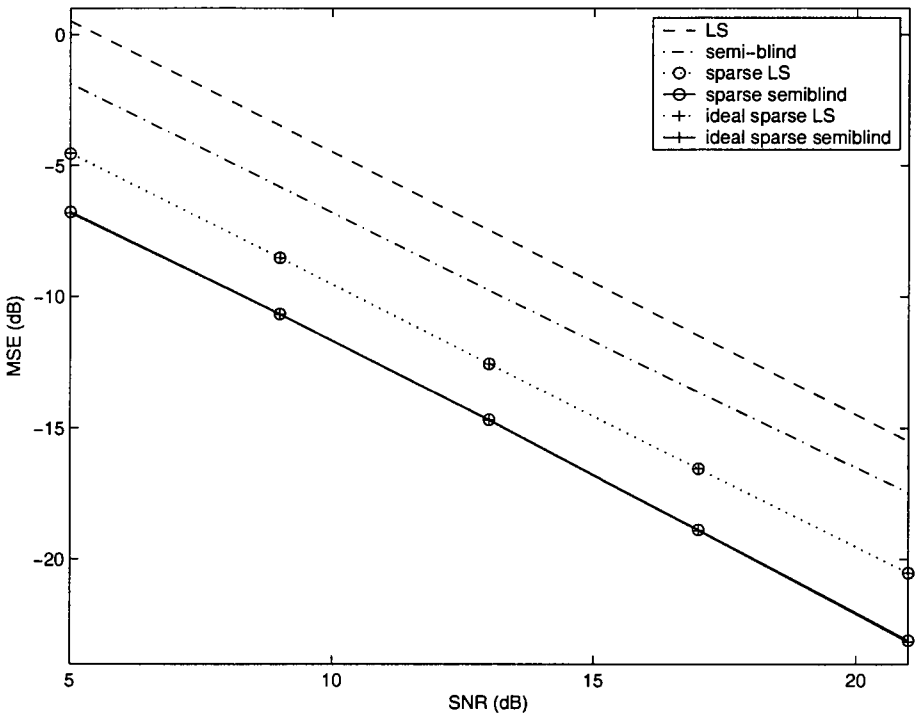

Figure 5.7: MSE versus SNR for *Channel D*. One OFDM symbol with 30 pilot subcarriers.

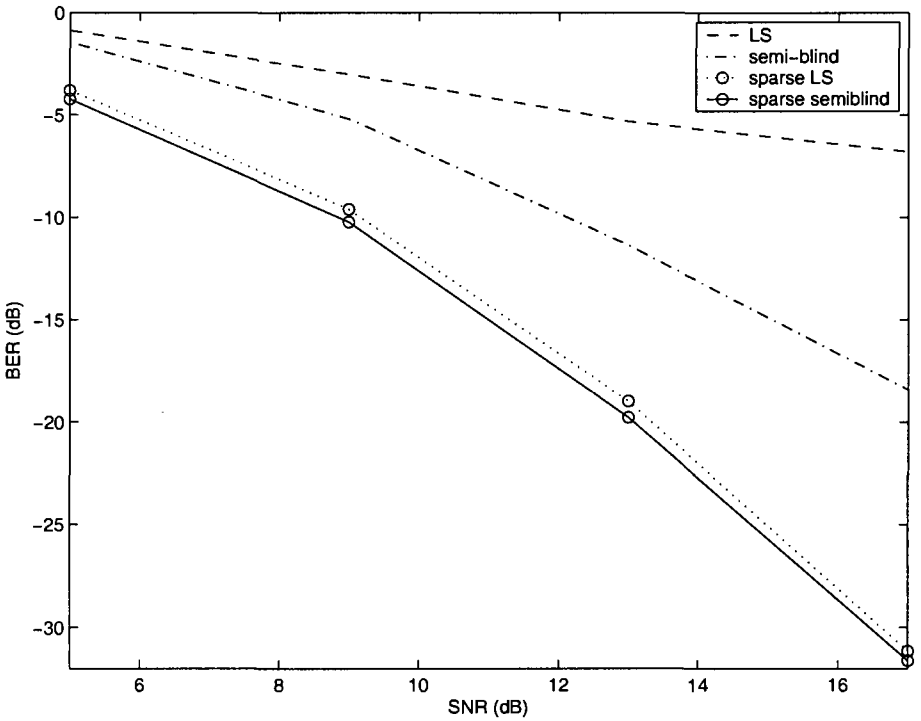

Figure 5.8: BER versus SNR for *Channel D*. One OFDM symbol with 30 pilot subcarriers.

blind method is not subject to the signal perturbation error in the M-rate sparse channel case.

Now, the BER performance of the MIMO-OFDM system is investigated by using the estimated channel matrix and an ordered vertical-Bell laboratories layered space time (V-BLAST) decoder. The simulation involves 4000 Monte Carlo runs of the transmission of one OFDM symbol with 20 pilot subcarriers. Fig. 5.8 shows the BER performance versus the SNR of the proposed sparse LS and semi-blind methods as well as their original versions. Clearly, the performance improvement of the proposed sparse semi-blind method gets more prominent compared to the original one with the increase of SNR. For example, when SNR is increased to 17 dB from 5 dB, the performance gain of the proposed sparse semi-blind method over the original one is boosted to 13.2 dB from 2.8dB. Moreover, it is seen that the performance of the sparse semi-blind method is superior to that of the sparse LS method by 0.42$\sim$0.8 dB.

Experiment 5: *Channel E*

With the same simulation condition as in *Experiment* 2, Fig.5.9 shows the channel estimation results of the previously mentioned six methods together with two other reference sparse methods: the sparse LS and semi-blind methods using *Channel D* as the MST detection result. It is seen that the sparse LS method consistently outperforms the same method with *Channel D* by about 1.1 $\sim$ 1.4 dB MSE. One can also observe that when SNR $<$ 14.4 dB, the sparse semi-blind method can achieve a better performance than the same method using *Channel D*, while the performance of the former becomes worse than that of the latter when SNR $>$ 14.4 dB. This

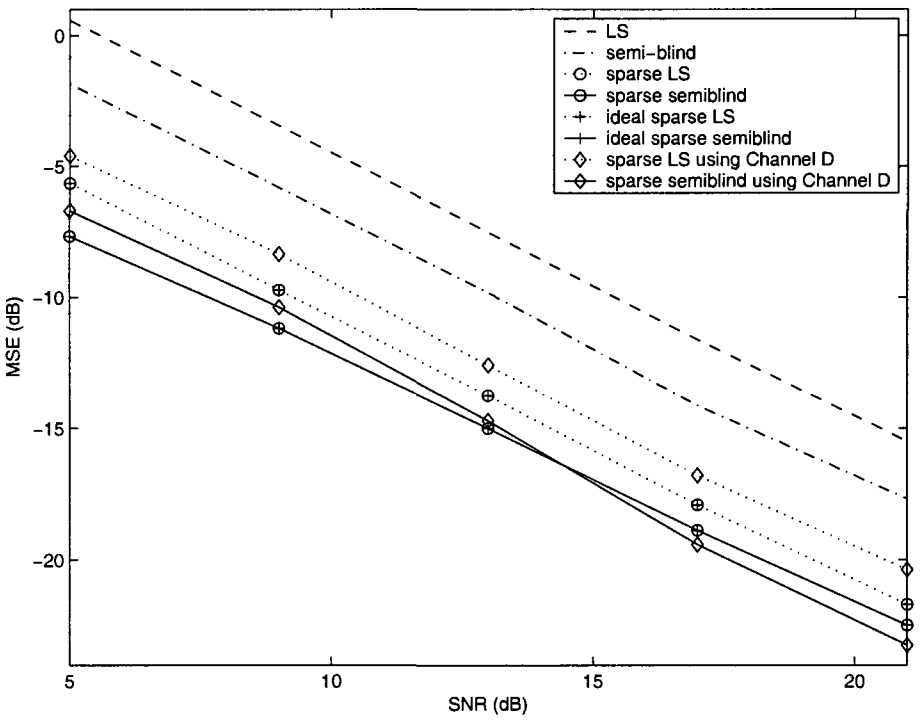

Figure 5.9: MSE versus SNR for *Channel E*. One OFDM symbol with 30 pilot subcarriers.

interesting outcome suggests that if a non-M-rate sparse channel can be converted to a M-rate one by adding several zero taps, then using the MST detection result from the M-rate channel may yield a much better channel estimation performance in the case of large SNRs.

5.7 Conclusion

In this chapter, a nulling-based semi-blind sparse channel estimation approach has been proposed for MIMO-OFDM systems. The relation between the most significant taps (MST) of the sparse channel and the most significant lags (MSL) of the correlation matrices of the received signal has first been disclosed. This relation has then been exploited to develop a highly efficient MST detection algorithm that requires only a few OFDM symbols and a small number of pilots for a least-square based detection. By employing the acquired MST information, a semi-blind approach incorporating a training-based LS criterion and a blind constraint on the sparse channel vector consisting of only a few MSTs has been proposed. As the new approach does not require estimating all the channel taps, it has saved a large amount of computations compared to a regular channel estimation method. It has also been shown via a perturbation analysis that the proposed semi-blind solution is not subject to the signal perturbation error when the sparse channel is a decimated version of a full channel. Computer simulations based on various sparse channels have confirmed that the proposed sparse semi-blind approach significantly outperforms the sparse LS method as well as the regular LS and nulling-based semi-blind techniques.

Chapter 6

Signal-Perturbation-Free Whitening Rotation based Estimation of Frequency-domain Channels

6.1 Introduction

As discussed in Section 2.2, the channel estimation is sometimes implemented by using a frequency-domain approach due to some constraints such as the low complexity requirement and only a small number of subcarriers used for the data transmission of a mobile user in OFDMA systems. In this case, channel estimation algorithms for the frequency-flat fading MIMO channel can be directly applied to the estimation of each frequency-domain channel $\mathbf{H}_{\mathrm{F}}(k)$. Usually, the frequency flat fading MIMO channel can be estimated by using the training-based algorithms [12] or the blind channel estimation algorithms like those proposed in [13–16, 45]. As a combination of the training-based and blind algorithms, a semi-blind estimation technique can potentially enhance the quality of MIMO channel estimation [17, 18, 75]. With a small number of training symbols, problems such as ambiguities and mis-convergence of the blind methods can be solved by semi-blind techniques. On the other hand, the use of

the available information data yields an improved accuracy of the channel estimation.

More recently, a whitening-rotation (WR)-based semi-blind algorithm has been proposed for frequency-flat MIMO channel estimation [17, 20, 21, 23, 24]. The idea of this algorithm was first briefly presented in [17] as a simplified version of a semi-blind algorithm for the estimation of frequency-selective MIMO channels. It was then fully disclosed in [20, 21, 23], in which a training-based constrained ML method was developed for the estimation of an ambiguity matrix for general MIMO systems. This method was later further extended for MIMO systems with maximum ratio transmission in [24]. The WR-based semi-blind algorithm consists of two steps: (1) estimation of a whitening matrix utilizing information data; and (2) estimation of a unitary rotation matrix using pilots. The Cramer-Rao bound (CRB) of this semi-blind technique shows that it can achieve a better channel estimation performance than the conventional LS method, when the number of receive antennas is greater than or equal to the number of transmit antennas. However, this method is found to be efficient only in the case of low SNRs.

A nulling-based semi-blind MIMO channel estimation approach, which can achieve a better performance in moderate to high SNR cases, was developed in [18, 42]. Instead of estimating the whitening matrix, this method uses the information data to obtain a blind constraint for the channel matrix, which is then combined with a training-based LS cost function so as to produce a semi-blind solution for the MIMO channel response. This method can be considered as a modified LS solution involving a weighted blind constraint. It has been shown in [42] that the semi-blind method provides a better channel estimation performance over the pure training-based LS

method. However, this superiority has not been theoretically proved, and the weighting factor employed to trade off the least square and the blind criteria has not well been determined.

In this chapter, we first apply the perturbation theory [71–77,111,112] to the analysis of the above mentioned two semi-blind algorithms, showing that in the noise-free case the blind part of the WR-based method is subject to a signal perturbation error, whereas the nulling-based method gives an ideal nulling constraint on the channel matrix, thus avoiding the signal perturbation error. Our analysis concludes that the nulling-based method is superior to the whitening-rotation-based method in the moderate to high SNR cases. We then derive a closed-form expression for the mean square error (MSE) of the blind estimation to facilitate the calculation of the weighting factor in the nulling-based method.

In the second part of this chapter, we propose a novel signal-perturbation-free transmit scheme to improve the performance of the WR-based method in the moderate to high SNR cases. By utilizing the eigenvalue decomposition (EVD) of the transmit signal perturbation matrix, a very efficient transmit scheme is designed for the elimination of the signal perturbation error in the receiver, leading to a signal-perturbation-free WR-based semi-blind approach. It is shown that the new approach provides a much better performance than the original WR-based method as well as the nulling-based method does for all SNR cases.

The third part of this chapter deals with the analysis of the new WR-based semi-blind method with the proposed transmit scheme. The first-order perturbation theory is employed to analyze the first step of the WR-based method, leading to an expres-

sion of the perturbation term of the whitening matrix. For the second step of the WR-based method, a perturbation analysis of the SVD of a square matrix involving only the signal subspace is conducted, yielding an closed-form expression for the perturbation term of the rotation matrix. By utilizing the perturbation terms of the whitening matrix and the rotation matrix, a closed-form expression for the MSE of the new WR-based method with the proposed transmit scheme is then derived.

The rest of the chapter is organized as follows. Section 6.2 conducts a perturbation analysis of two subspace-based semi-blind approaches, justifying why the nulling-based method is superior to the WR-based method in the moderate to high SNR case. The analysis also yields a closed-form expression for the MSE of the blind estimation part of the nulling-based approach. Section 6.3 presents a novel signal-perturbation-free transmit scheme for the WR-based method that can eliminate the signal perturbation error at the receiver. Section 6.4 provides a perturbation analysis of the new WR-based approach with the proposed transmit scheme and the derivation of a closed-form expression for the MSE of the proposed method. Section 6.5 comprises a number of experiments verifying the significant advantages provided by the new WR-based approach over the regular WR-based method as well as the nulling-based method. Finally, Section 6.6 highlights some of the distinct features of the proposed approach.

6.2 Perturbation Analysis of Subspace-based Blind Channel Estimation

It is known that the solution of subspace based methods is always perturbed by various sources, such as the finite data length and the measurement noise [71–74]. The perturbation theory has been employed for the analysis of subspace based methods [72, 76, 77, 111, 112]. In this section, the first-order perturbation analysis is used to evaluate the performance of the two subspace-based channel estimation methods [113].

6.2.1 Analysis of the WR-based Method

Here, we consider only the perturbation due to the finite data length in the computation of the correlation matrices. Our objective is to show that the whitening matrix would be perturbed even in the absence of noise. Using (2.7), the autocorrelation matrix of the received signal, $\hat{\mathbf{R}}_{\mathbf{Y}}$, with such a perturbation can be written as [83]

$$\hat{\mathbf{R}}_{\mathbf{Y}} = \hat{\mathrm{E}}\left[\mathbf{y}\left(n\right)\mathbf{y}^{H}\left(n\right)\right] - \sigma_v^2\mathbf{I} = \mathbf{H}\left[\mathbf{I} + \Delta\mathbf{R}_x\right]\mathbf{H}^{H} + \Delta\mathbf{R}_v \tag{6.1}$$

where $\Delta\mathbf{R}_x$ denotes the signal perturbation matrix

$$\Delta\mathbf{R}_x \triangleq \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}\left(n\right)\mathbf{x}^{H}\left(n\right) - \sigma_x^2\mathbf{I}, \tag{6.2}$$

and $\Delta\mathbf{R}_v$ the perturbation matrix introduced by the noise

$$\Delta\mathbf{R}_v \triangleq \mathbf{H}\Delta\mathbf{R}_{xv} + \Delta\mathbf{R}_{xv}^{H}\mathbf{H}^{H} + \Delta\mathbf{R}_{vv} \tag{6.3}$$

with

$$\Delta\mathbf{R}_{xv} \triangleq \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}\left(n\right)\mathbf{v}^{H}\left(n\right), \tag{6.4}$$

$$\Delta \mathbf{R}_{vv} \triangleq \frac{1}{N} \sum_{n=1}^{N} \mathbf{v}(n) \mathbf{v}^H(n) - \sigma_v^2 \mathbf{I}.$$

In the noise-free case, all the perturbation terms introduced by the noise would disappear. Then, (6.1) reduces to

$$\hat{\mathbf{R}}_{\mathbf{Y}} = \mathbf{H} [\mathbf{I} + \Delta \mathbf{R}_x] \mathbf{H}^H. \tag{6.5}$$

Based on the above perturbation forms, we can derive the estimate of $\mathbf{W}$. Using (2.16) into (6.5), one can get

$$\hat{\mathbf{R}}_{\mathbf{Y}} = \mathbf{U} \mathbf{I}_A \mathbf{T} \mathbf{I}_A^H \mathbf{U}^H, \tag{6.6}$$

$$\mathbf{I}_A \triangleq \begin{bmatrix} \mathbf{I}_{N_T \times N_T} \\ \mathbf{0}_{(N_R - N_T) \times N_T} \end{bmatrix}, \tag{6.7}$$

$$\mathbf{T} \triangleq \mathbf{\Sigma}_S^2 + \mathbf{\Sigma}_S \mathbf{V}^H \Delta \mathbf{R}_x \mathbf{V} \mathbf{\Sigma}_S \tag{6.8}$$

where $\mathbf{\Sigma}_S$ is a diagonal matrix satisfying $\mathbf{\Sigma} = \mathbf{I}_A \mathbf{\Sigma}_S$. Since $\mathbf{T}$ is Hermitian and the signal perturbation matrix $\Delta \mathbf{R}_x$ has a small norm, the singular value of $\mathbf{T}$ would be different from but close to $\mathbf{\Sigma}_S$. Thus, the SVD of $\mathbf{T}$ can be written as

$$\mathbf{T} = \mathbf{\Pi}_S (\mathbf{\Sigma}_S + \Delta \mathbf{\Sigma}_S)^2 \mathbf{\Pi}_S^H \tag{6.9}$$

where $\mathbf{\Pi}_S$ is a unitary matrix and $\Delta \mathbf{\Sigma}_S$ is considered as the perturbation error of $\mathbf{\Sigma}_S$. Substituting (6.9) into (6.6), comparing it with (6.5) and noting that $\mathbf{H} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}$, one can obtain an estimate of the whitening matrix $\mathbf{W}$ as

$$\hat{\mathbf{W}} = \hat{\mathbf{U}} \hat{\mathbf{\Sigma}} = \mathbf{U} \mathbf{I}_A \mathbf{\Pi}_S (\mathbf{\Sigma}_S + \Delta \mathbf{\Sigma}_S). \tag{6.10}$$

From (6.10), one can find that, even in the noise-free case, $\hat{\mathbf{W}}$ consists of two perturbation terms $\mathbf{\Pi}_S$ and $\Delta \mathbf{\Sigma}_S$, which are dictated by the signal perturbation

matrix $\Delta\mathbf{R}_x$. This could explain why the performance of the WR-based method is very poor in the moderate to high SNR cases.

6.2.2 Analysis of the Nulling-Based Method

In the nulling-based method proposed in [18,42], a nulling constraint $\hat{\mathbf{U}}_{\text{null}}$ on the channel matrix $\mathbf{H}$ is obtained from the SVD of $\hat{\mathbf{R}}_{\mathbf{Y}}$. It can be verified from (6.5) that in the absence of noise, the nulling constraint $\hat{\mathbf{U}}_{\text{null}}$ is ideal, namely,

$$\hat{\mathbf{U}}_{\text{null}}^H \mathbf{H} = \mathbf{0}.$$

This implies that in the noise-free case the blind constraint of the nulling-based method is perfectly satisfied without being affected by the signal perturbation terms, and therefore, the nulling-based method is superior to the WR-based method. In the following, we derive a closed-form expression for the MSE of the blind estimate in the nulling-based method for the noisy case.

The MSE of the blind estimate is defined as

$$\text{MSE}_{\text{B}} \triangleq \text{E}\left\{ \left\| \hat{\mathbf{U}}_{\text{null}}^H \mathbf{H} \right\|_F^2 \right\}. \tag{6.11}$$

Using (6.1) and (6.3) along with the first-order approximation [77], one can get an estimate of $\mathbf{U}_{null}$:

$$\hat{\mathbf{U}}_{\text{null}} \approx \mathbf{U}_{\text{null}} - \left(\mathbf{H}\mathbf{H}^H\right)^{\dagger} \mathbf{H} \Delta\mathbf{R}_{xv} \mathbf{U}_{\text{null}}. \tag{6.12}$$

For the medium to high SNR, the perturbation matrix $\Delta\mathbf{R}_{vv}$ of the noise autocorrelation can be neglected. Thus, substituting (6.12) into (6.11), and noting that $\mathbf{H}^H \left(\mathbf{H}\mathbf{H}\right)^{\dagger} \mathbf{H} = \mathbf{I}$ when $\mathbf{H}$ has a full column rank and $N_R > N_T$, one can obtain

$$\text{MSE}_{\text{B}} = \text{E}\left\{ \left\| \hat{\mathbf{U}}_{\text{null}}^H \Delta\mathbf{R}_{xv}^H \right\|_F^2 \right\}. \tag{6.13}$$

Due to the fact that $\|\mathbf{A}\|_F^2 = \mathrm{vec}(\mathbf{A})\,\mathrm{vec}^H(\mathbf{A})$, $\mathrm{MSE_B}$ can be calculated by

$$\mathrm{MSE_B} = \mathrm{Trace}[(\mathbf{I} \otimes \mathbf{U}_{\mathrm{null}}^H)\,\mathbf{R}_{\Delta xv}\,(\mathbf{I} \otimes \mathbf{U}_{\mathrm{null}})] \tag{6.14}$$

where

$$\mathbf{R}_{\Delta xv} \triangleq \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv}^H\right)\right\}.$$

Theorem 6.1: $\mathbf{R}_{\Delta xv}$ can be calculated in terms of the signal and noise variances σ_x^2 and σ_v^2, namely,

$$\mathbf{R}_{\Delta xv} = \frac{1}{N}\sigma_x^2\sigma_v^2\mathbf{I}_{N_R N_T}. \tag{6.15}$$

The proof of the theorem is given in Appendix D.

Substituting (6.15) into (6.14) yields

$$\mathrm{MSE_B} = \frac{1}{N}\sigma_x^2\sigma_v^2\mathrm{Trace}\left[\mathbf{I}_{N_T} \otimes \left(\mathbf{U}_{\mathrm{null}}^H\mathbf{U}_{\mathrm{null}}\right)\right] = \frac{1}{N}\sigma_x^2\sigma_v^2 N_T\left(N_R - N_T\right). \tag{6.16}$$

We now utilize the above closed-form expression of $\mathrm{MSE_B}$ to calculate the weighting factor in the nulling-based method. The weight factor α can be chosen as [81,83]

$$\alpha = \frac{\mathrm{MSE_T}}{\mathrm{MSE_B}}\frac{\|\mathbf{X}_P\|_F^2}{\left\|\hat{\mathbf{U}}_{\mathrm{null}}\right\|_F^2}, \tag{6.17}$$

where $\mathrm{MSE_T}$ represents the MSE of the training-based LS estimation of the channel vector. For orthogonal training pilots, it can be easily verified that

$$\|\mathbf{X}_{\mathrm{P}}\|_F^2 = KN_T\sigma_x^2, \tag{6.18}$$

$$\mathrm{MSE_T} \triangleq \mathrm{E}\left\{\left\|\hat{\mathbf{H}}_{\mathrm{T}} - \mathbf{H}\right\|_F^2\right\} = \frac{N_R N_T\sigma_v^2}{K\sigma_x^2}, \tag{6.19}$$

$$\left\|\hat{\mathbf{U}}_{\mathrm{null}}\right\|_F^2 = N_R - N_T. \tag{6.20}$$

Thus, by substituting (6.16), (6.18), (6.19) and (6.20) into (6.17), the weighting factor α of the nulling-based semi-blind method can be calculated as

$$\alpha = \frac{NN_TN_R}{(N_R - N_T)^2 \sigma_x^2}. \tag{6.21}$$

What remains in the nulling-based method is to solve the minimization problem in (2.27). This can be easily done by calculating the derivative of the cost function in (2.27) with respect to the channel vector, $\mathbf{h} \triangleq \mathrm{vec}(\mathbf{H})$, namely,

$$\frac{\partial \Delta}{\partial \mathbf{h}^H} = -\left(\mathbf{X}_{\mathrm{P}}^* \otimes \mathbf{I}_{N_R}\right)\left[\mathrm{vec}\left(\mathbf{Y}_{\mathrm{P}}\right) - \left(\mathbf{X}_{\mathrm{P}}^T \otimes \mathbf{I}_{N_R}\right)\mathbf{h}\right] + \alpha\left(\mathbf{I}_{N_T} \otimes \hat{\mathbf{U}}_{\mathrm{null}}\right)\left[\left(\mathbf{I}_{N_T} \otimes \hat{\mathbf{U}}_{\mathrm{null}}^H\right)\mathbf{h}\right]. \tag{6.22}$$

Letting $\frac{\partial \Delta}{\partial \mathbf{h}^H}$ be zero gives a closed-form solution for the channel vector,

$$\hat{\mathbf{h}} = \left\{\left[\left(\mathbf{X}_{\mathrm{P}}^* \mathbf{X}_{\mathrm{P}}^T\right) \otimes \mathbf{I}_{N_R}\right] + \alpha\left[\mathbf{I}_{N_T} \otimes \left(\hat{\mathbf{U}}_{\mathrm{null}} \hat{\mathbf{U}}_{\mathrm{null}}^H\right)\right]\right\}^{\dagger} \left(\mathbf{X}_{\mathrm{P}}^* \otimes \mathbf{I}_{N_R}\right) \mathrm{vec}\left(\mathbf{Y}_{\mathrm{P}}\right). \tag{6.23}$$

As will be seen from the simulation results in Section 6.5, the nulling-based solution given by (6.23) outperforms the WR-based method in the case of medium to high SNRs.

6.3 Proposed Signal-Perturbation-Free Transmit Scheme

The idea of the new signal-perturbation-free transmit scheme is to send information of the signal perturbation matrix $\Delta\mathbf{R}_x$ to the receiver. The received version of this information will be then exploited to cancel the signal perturbation error [114, 115]. The novel idea begins with the eigenvalue decomposition (EVD) of $\Delta\mathbf{R}_x$, which is re-defined as the scaled version of (6.2) for notational convenience,

$$\Delta\mathbf{R}_A = \sum_{n=1}^{N} \mathbf{x}(n)\,\mathbf{x}^H(n) - N\sigma_x^2\mathbf{I}. \tag{6.24}$$

As $\Delta\mathbf{R}_A$ is a Hermitian matrix, it can be verified that the EVD of $\Delta\mathbf{R}_A$ can be written as [116]

$$\Delta\mathbf{R}_A = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{N_T}] \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{N_T} \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^H \\ \mathbf{u}_2^H \\ \vdots \\ \mathbf{u}_{N_T}^H \end{bmatrix} \tag{6.25}$$

where $\sigma_i, (i = 1, 2, \cdots, N_T)$, are the eigenvalues in the range of $(-\infty, \infty)$, and $\mathbf{u}_i$, $(i = 1, 2, \cdots, N_T)$, are the corresponding eigenvectors. Using (6.25), one can separate $\Delta\mathbf{R}_A$ into two parts as

$$\Delta\mathbf{R}_A = \Delta\mathbf{R}_{\text{pos}} - \Delta\mathbf{R}_{\text{neg}} \tag{6.26}$$

where

$$\Delta\mathbf{R}_{\text{pos}} = \sum_{i=1}^{L_{\text{pos}}} \sigma_{\text{pos},i} \mathbf{u}_{\text{pos},i} \mathbf{u}_{\text{pos},i}^H, \tag{6.27}$$

$$\Delta\mathbf{R}_{\text{neg}} = \sum_{i=1}^{L_{\text{neg}}} (-\sigma_{\text{neg},i}) \, \mathbf{u}_{\text{neg},i} \mathbf{u}_{\text{neg},i}^H. \tag{6.28}$$

Here, $\sigma_{\text{pos},i}$ represents a positive eigenvalue, $\mathbf{u}_{\text{pos},i}$ is the eigenvector associated with $\sigma_{\text{pos},i}$, and the L_{pos} is the total number of positive eigenvalues. Likewise, $\sigma_{\text{neg},i}$, $\mathbf{u}_{\text{neg},i}$ and L_{neg} refer to similar quantities with respect to the negative eigenvalues. In what follows, we will derive a further decomposed form of $\Delta\mathbf{R}_A$ based on (6.26) such that the information of $\Delta\mathbf{R}_{\text{pos}}$ and $\Delta\mathbf{R}_{\text{neg}}$ can easily be transmitted to the receiver, which will then be employed to cancel the signal perturbation error.

By letting

$$\Delta\mathbf{R}_{\text{pos}} = \eta \mathbf{X}_{\text{pos}} \mathbf{X}_{\text{pos}}^H, \tag{6.29}$$

$$\Delta\mathbf{R}_{\text{neg}} = \eta \mathbf{X}_{\text{neg}} \mathbf{X}_{\text{neg}}^H, \tag{6.30}$$

(6.26) can be rewritten as

$$\Delta \mathbf{R}_A = \eta \left(\mathbf{X}_{\text{pos}} \mathbf{X}_{\text{pos}}^H - \mathbf{X}_{\text{neg}} \mathbf{X}_{\text{neg}}^H \right) \tag{6.31}$$

where η is a scaling factor, and $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ are two matrices containing the information of $\Delta \mathbf{R}_{\text{pos}}$ and that of $\Delta \mathbf{R}_{\text{neg}}$, respectively. We now first show that as long as $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ are transmitted to the receiver, the signal perturbation error can completely be cancelled. We will then show that $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ can easily be constructed using the singular values and the singular vectors of $\Delta \mathbf{R}_A$. As shown later, the size of $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ can be made comparable to the dimension of $\Delta \mathbf{R}_A$, namely, the number of the transmit antennas N_T. Therefore, the spectral resources used for transmitting $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ is negligible as compared to that of the user data.

Letting $\mathbf{Y}_{\text{pos}}$ and $\mathbf{Y}_{\text{neg}}$ be the received signals corresponding to $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$, respectively, namely,

$$\mathbf{Y}_{\text{pos}} = \mathbf{H}\mathbf{X}_{\text{pos}} + \mathbf{V}_{\text{pos}}, \tag{6.32}$$

$$\mathbf{Y}_{\text{neg}} = \mathbf{H}\mathbf{X}_{\text{neg}} + \mathbf{V}_{\text{neg}} \tag{6.33}$$

where $\mathbf{V}_{\text{pos}}$ and $\mathbf{V}_{\text{neg}}$ are the corresponding noise matrices, the received version of the signal perturbation matrix can be defined as

$$\Delta \hat{\mathbf{R}}_{\mathbf{Y}} = \frac{\eta}{N} \left[\left(\mathbf{Y}_{\text{pos}} \mathbf{Y}_{\text{pos}}^H - \mathbf{Y}_{\text{neg}} \mathbf{Y}_{\text{neg}}^H \right) - \left(N_{\text{pos}} - N_{\text{neg}} \right) \sigma_v^2 \mathbf{I} \right] \tag{6.34}$$

where N_{pos} and N_{neg} denote the number of columns of $\mathbf{Y}_{\text{pos}}$ and that of $\mathbf{Y}_{\text{neg}}$, respectively. Using (6.32) and (6.33) into (6.34) and noting that $\Delta \mathbf{R}_A = N \Delta \mathbf{R}_x$, we obtain

$$\Delta \hat{\mathbf{R}}_{\mathbf{Y}} = \mathbf{H} \Delta \mathbf{R}_x \mathbf{H}^H + \Delta \mathbf{R}_{vp} \tag{6.35}$$

where $\Delta\mathbf{R}_{vp}$ represents a perturbation term introduced by the noise. By utilizing (6.1) and (6.35), the received correlation matrix without the signal perturbation error can be obtained from

$$\hat{\mathbf{R}}'_{\mathbf{Y}} = \hat{\mathbf{R}}_{\mathbf{Y}} - \Delta\hat{\mathbf{R}}_{\mathbf{Y}} = \mathbf{H}\mathbf{H}^H + \Delta\mathbf{R}'_v \tag{6.36}$$

where

$$\Delta\mathbf{R}'_v = \Delta\mathbf{R}_v - \Delta\mathbf{R}_{vp}. \tag{6.37}$$

As a result, the signal perturbation error has been completely eliminated through the transmission of $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$. What remains to be done in the proposed scheme is to determine the matrices $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ from the singular values σ_i and the singular vectors $\mathbf{u}_i$.

Note that the total power of N_T transmit antennas in each time slot can be written as $\sigma_{\text{int}} \triangleq N_T\sigma_x^2$. It is found from a large number of simulation experiments that the value of $\sigma_{\text{pos},i}$ is much larger than σ_{int}. To transmit $\sigma_{\text{pos},i}$ with a small number of slots, it is first divided by the scaling factor η and then is split into $N_{\text{pos},i}$ terms of σ_{int} plus one fractional term as

$$\frac{\sigma_{\text{pos},i}}{\eta} = (N_{\text{pos},i}\sigma_{\text{int}} + \sigma_{\text{pos-frac},i}) \tag{6.38}$$

where

$$N_{\text{pos},i} = \lfloor \frac{\sigma_{\text{pos},i}}{\eta\sigma_{\text{int}}} \rfloor, \tag{6.39}$$

$$\sigma_{\text{pos-frac},i} = \frac{\sigma_{\text{pos},i}}{\eta} - N_{\text{pos},i}\sigma_{\text{int}}. \tag{6.40}$$

Letting

$$\mathbf{x}_{\text{pos-int},i} = \sqrt{\sigma_{\text{int}}}\mathbf{u}_{\text{pos},i},$$

$$\mathbf{x}_{\text{pos-frac},i} = \sqrt{\sigma_{\text{pos-frac},i}}\mathbf{u}_{\text{pos},i},$$

one can construct an $N_T \times (N_{\text{pos},i} + 1)$ matrix $\mathbf{X}_{\text{pos},i}$ for the i-th singular value by stacking $N_{\text{pos},i}$ consecutive vectors $\mathbf{x}_{\text{pos-int},i}$ and one vector $\mathbf{x}_{\text{pos-frac},i}$, which satisfies

$$\eta \mathbf{X}_{\text{pos},i}\mathbf{X}_{\text{pos},i}^H = \sigma_{\text{pos},i}\mathbf{u}_{\text{pos},i}\mathbf{u}_{\text{pos},i}^H.$$

Thus, the complete $\mathbf{X}_{\text{pos}}$ can be formed as

$$\mathbf{X}_{\text{pos}} = \left[\mathbf{X}_{\text{pos},1}, \mathbf{X}_{\text{pos},2}, \cdots, \mathbf{X}_{\text{pos},L_{\text{pos}}}\right]. \tag{6.41}$$

Obviously, the number of columns of $\mathbf{X}_{\text{pos}}$ is given by $N_{\text{pos}} = L_{\text{pos}} + \sum_{i=1}^{L_{\text{pos}}} N_{\text{pos},i}$.

In a similar manner, $\mathbf{X}_{\text{neg}}$ can be constructed as follows

$$\mathbf{X}_{\text{neg}} = \left[\mathbf{X}_{\text{neg},1}, \mathbf{X}_{\text{neg},2}, \cdots, \mathbf{X}_{\text{neg},L_{\text{neg}}}\right] \tag{6.42}$$

where $\mathbf{X}_{\text{neg},i}$ consists of $N_{\text{neg},i} = \lfloor \frac{\sigma_{\text{neg},i}}{\eta\sigma_{\text{int}}} \rfloor$ consecutive vectors $\mathbf{x}_{\text{neg-int},i}$ and one vector $\mathbf{x}_{\text{neg-frac},i}$ as given by

$$\mathbf{x}_{\text{neg-int},i} = \sqrt{\sigma_{\text{int}}}\mathbf{u}_{\text{neg},i}, \tag{6.43}$$

$$\mathbf{x}_{\text{neg-frac},i} = \sqrt{\sigma_{\text{neg-frac},i}}\mathbf{u}_{\text{neg},i}, \tag{6.44}$$

$$\sigma_{\text{neg-frac},i} = \frac{\sigma_{\text{neg},i}}{\eta} - N_{\text{neg},i}\sigma_{\text{int}}. \tag{6.45}$$

Note that the number of columns of $\mathbf{X}_{\text{neg}}$ is

$$N_{\text{neg}} = L_{\text{neg}} + \sum_{i=1}^{L_{\text{neg}}} N_{\text{neg},i}. \tag{6.46}$$

From the above discussion, a new transmit structure, which consists of conventional pilots, user's data and the additional data $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$, can be obtained as shown in Fig.6.1. It is now clear that the total column size of $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ is inversely

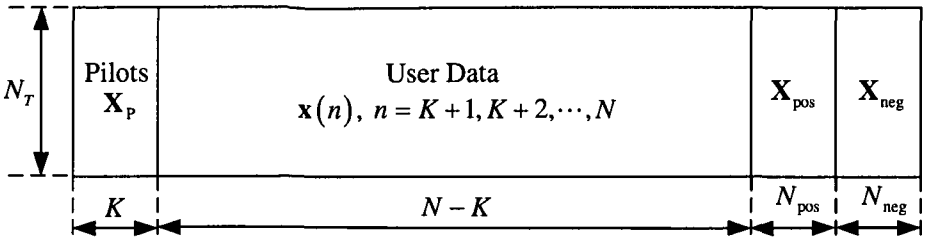

Figure 6.1: Signal-perturbation-free transmit structure for flat-fading channel

proportional to the scaling factor η. It can be shown that when η is sufficiently large, as less as N_T slots can be used for the transmission of $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$. In general, the choice of η should depend on the number of the transmit antennas as well as on the length of the user data. Our extensive simulations show that $\eta = 16$ is a proper choice to achieve a very good channel estimate for a 4×8 MIMO system, in which case, the transmission of $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ requires only 9 slots when the user data length is about 1000.

In summary, the scheme developed above gives a signal-perturbation-free estimate of the whitening matrix in the noise-free case as seen from (6.36), leading to an ideal WR-based method. It should be pointed out that in the presence of noise, although the WR-based method with the new transmit scheme is subject to the noise perturbation, the proposed method still outperforms the WR-based method, since the perturbation introduced by the noise is, in general, significantly smaller than the signal perturbation. It should also be mentioned that the proposed signal-perturbation-free transmit scheme is also very useful for other correlation-based methods for MIMO channel estimation.

6.4 Analysis of New WR-based Method with Proposed Transmit Scheme

In this section, the first-order perturbation theory is employed to analyze the proposed WR-based method, in which the channel estimate can be denoted as $\hat{\mathbf{H}} = \hat{\mathbf{W}}\hat{\mathbf{Q}}^H$. We first derive the perturbation errors of the whitening matrix $\mathbf{W}$ and the rotation matrix $\mathbf{Q}$ due to the noise, and then reveal a closed-form expression for the MSE of the WR-based method with the proposed transmit scheme.

6.4.1 The Perturbation Error of the Whitening Matrix W

First of all, from (6.36), the ideal correlation matrix $\mathbf{R_Y}$ without both the signal and the noise perturbation terms can be written as

$$\mathbf{R_Y} \triangleq \mathbf{H}\mathbf{H}^H = [\mathbf{U}_S, \mathbf{U}_N] \begin{bmatrix} \mathbf{\Sigma}_S^2 & \\ & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_S^H \\ \mathbf{U}_N^H \end{bmatrix}. \tag{6.47}$$

In obtaining (6.47), we have employed the SVD of $\mathbf{H}$ as given by

$$\mathbf{H} = [\mathbf{U}_S, \mathbf{U}_N] \begin{bmatrix} \mathbf{\Sigma}_S \\ \mathbf{0} \end{bmatrix} \mathbf{V}_S^H \tag{6.48}$$

where $\mathbf{V}_S = \mathbf{V}$ under the assumption of the full-column rank of $\mathbf{H}$.

In the case of noise perturbation but without the signal perturbation error, the autocorrelation matrix of the received signal can be written as

$$\hat{\mathbf{R}}'_{\mathbf{Y}} = \hat{\mathbf{H}}\hat{\mathbf{H}}^H \tag{6.49}$$

where $\hat{\mathbf{H}}$ denotes the estimate of the channel matrix containing the noise perturbation

error. Let the SVD of $\hat{\mathbf{H}}$ be given by

$$\hat{\mathbf{H}} = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\hat{\mathbf{V}}^H. \tag{6.50}$$

As $\mathbf{H}$ has a full column rank, one can partition $\hat{\mathbf{U}}$ and $\hat{\mathbf{\Sigma}}$ according to the signal and the noise subspaces of $\hat{\mathbf{H}}$ as

$$\hat{\mathbf{U}} = [\hat{\mathbf{U}}_S, \hat{\mathbf{U}}_N], \ \hat{\mathbf{\Sigma}} = \begin{bmatrix} \hat{\mathbf{\Sigma}}_S \\ \mathbf{0} \end{bmatrix}.$$

From (6.49) and (6.50) along with the above partitioned form, we have

$$\hat{\mathbf{R}}'_{\mathbf{Y}} = [\hat{\mathbf{U}}_S, \hat{\mathbf{U}}_N] \begin{bmatrix} \hat{\mathbf{\Sigma}}_S^2 & \\ & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{U}}_S^H \\ \hat{\mathbf{U}}_N^H \end{bmatrix}. \tag{6.51}$$

Clearly, (6.51) gives an SVD of $\hat{\mathbf{R}}'_{\mathbf{Y}}$, implying that $\hat{\mathbf{U}}_S$ and $\hat{\mathbf{\Sigma}}_S$ can be obtained from the SVD of $\hat{\mathbf{R}}'_{\mathbf{Y}}$. Using (6.49) and (6.51) and noting that $\hat{\mathbf{H}} = \hat{\mathbf{W}}\hat{\mathbf{Q}}^H$, we can obtain the perturbed version of $\mathbf{W}$,

$$\hat{\mathbf{W}} = \hat{\mathbf{U}}_S\hat{\mathbf{\Sigma}}_S. \tag{6.52}$$

By defining $\Delta\mathbf{U}_S \triangleq \hat{\mathbf{U}}_S - \mathbf{U}_S$ and $\Delta\mathbf{\Sigma}_S \triangleq \hat{\mathbf{\Sigma}}_S - \mathbf{\Sigma}_S$, the perturbation error of $\mathbf{W}$ can be expressed as

$$\begin{aligned} \Delta\mathbf{W} &= \hat{\mathbf{W}} - \mathbf{W} = \hat{\mathbf{U}}_S\hat{\mathbf{\Sigma}}_S - \mathbf{U}_S\mathbf{\Sigma}_S, \\ &\approx \Delta\mathbf{U}_S\mathbf{\Sigma}_S + \mathbf{U}_S\Delta\mathbf{\Sigma}_S \end{aligned} \tag{6.53}$$

where the terms involving the second-order perturbation error have been neglected. In the following, we simplify the computation of $\Delta\mathbf{W}$ by investigating $\Delta\mathbf{U}_S$ and $\Delta\mathbf{\Sigma}_S$.

Using the first-order approximation [77], it can be shown that

$$\Delta\mathbf{U}_S \approx \mathbf{U}_N\mathbf{U}_N^H\Delta\mathbf{R}'_v\mathbf{U}_S\mathbf{\Sigma}_S^{-2}. \tag{6.54}$$

From (6.37), one can easily verify that, when $N \gg N_{\text{plus}}+N_{\text{neg}}$, $\Delta\mathbf{R}_{vp}$ can be neglected and $\Delta\mathbf{R}'_v$ approximates to $\Delta\mathbf{R}_v$. Using this result into (6.3) and neglecting the perturbation matrix $\Delta\mathbf{R}_{vv}$ of the noise autocorrelation for a medium to high SNR, we have

$$\Delta\mathbf{R}'_v \approx \mathbf{H}\Delta\mathbf{R}_{xv} + \Delta\mathbf{R}_{xv}^H\mathbf{H}^H. \tag{6.55}$$

Substituting (6.55) into (6.54) and utilizing (6.48), one can obtain

$$\Delta\mathbf{U}_S = \mathbf{U}_N\mathbf{U}_N^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S\boldsymbol{\Sigma}_S^{-1}. \tag{6.56}$$

On the other hand, using (6.36), (6.51) and the first-order approximation [72], one can show that

$$(\boldsymbol{\Sigma}_S + \Delta\boldsymbol{\Sigma}_S)^2 - \boldsymbol{\Sigma}_S^2 \approx \mathbf{U}_S^H\Delta\mathbf{R}'_v\mathbf{U}_S. \tag{6.57}$$

Using (6.48) and (6.55), (6.57) can be rewritten as

$$(\boldsymbol{\Sigma}_S + \Delta\boldsymbol{\Sigma}_S)^2 - \boldsymbol{\Sigma}_S^2 = \boldsymbol{\Sigma}_S\mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{U}_S + \mathbf{U}_S^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S\boldsymbol{\Sigma}_S^H, \tag{6.58}$$

which, under the assumption that the second-order perturbation error terms can be ignored, leads to

$$\Delta\boldsymbol{\Sigma}_S \approx \mathbf{U}_S^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S. \tag{6.59}$$

Using (6.56) and (6.59) into (6.53), $\Delta\mathbf{W}$ can be rewritten as

$$\Delta\mathbf{W} \approx \mathbf{U}_N\mathbf{U}_N^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S + \mathbf{U}_S\mathbf{U}_S^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S. \tag{6.60}$$

Recalling that $\mathbf{U}_N\mathbf{U}_N^H + \mathbf{U}_S\mathbf{U}_S^H = \mathbf{I}$, (6.60) reduces to

$$\Delta\mathbf{W} \approx \Delta\mathbf{R}_{xv}^H\mathbf{V}_S. \tag{6.61}$$

As will be shown in the following subsections, (6.61) can be used to derive the MSE of the proposed signal-perturbation-free WR-based approach.

6.4.2 The Perturbation Error of the Rotation Matrix Q

Prior to the derivation of the perturbation error of $\mathbf{Q}$, we first conduct an analysis for the training-based estimation of $\mathbf{Q}$ given in [23]. In the noise-free case, the rotation matrix $\mathbf{Q}$ can be calculated from

$$\mathbf{Q} = \mathbf{V}_Q \mathbf{U}_Q^H \tag{6.62}$$

where $\mathbf{U}_Q$ and $\mathbf{V}_Q$ are obtained from an SVD of the matrix

$$\mathbf{Y}_Q \triangleq \frac{1}{K\sigma_x^2} \mathbf{W}^H \mathbf{Y}_P \mathbf{X}_P^H, \tag{6.63}$$

namely,

$$\mathbf{Y}_Q = \mathbf{U}_Q \mathbf{\Sigma}_Q \mathbf{V}_Q^H. \tag{6.64}$$

By noting that $\mathbf{W} = \mathbf{U}_S \mathbf{\Sigma}_S$ and

$$\frac{1}{K\sigma_x^2} \mathbf{Y}_P \mathbf{X}_P^H = \mathbf{H}, \tag{6.65}$$

(6.63) can be rewritten as

$$\mathbf{Y}_Q = \mathbf{\Sigma}_S^2 \mathbf{V}_S^H. \tag{6.66}$$

As such, one realization of the SVD of $\mathbf{Y}_Q$ is given by

$$\mathbf{U}_Q = \mathbf{I}, \mathbf{\Sigma}_Q = \mathbf{\Sigma}_S^2, \mathbf{V}_Q = \mathbf{V}_S. \tag{6.67}$$

Using (6.67) into (6.62) yields

$$\mathbf{Q} = \mathbf{V}_S. \tag{6.68}$$

The above discussion indicates that the method in [23] gives an ideal rotation matrix $\mathbf{V}_S$ in the noise-free case. In the following, we derive the expression of the perturbation error $\Delta\mathbf{Q}$ of $\mathbf{Q}$ in the presence of noise.

In the noisy case, (6.63) should be modified to

$$\hat{\mathbf{Y}}_Q = \frac{1}{K\sigma_x^2}\hat{\mathbf{W}}^H \mathbf{Y}_P \mathbf{X}_P^H \tag{6.69}$$

where $\hat{\mathbf{W}} = \mathbf{W} + \Delta\mathbf{W}$ is the perturbed version of $\mathbf{W}$, which has been discussed in the previous subsection. Noting that $\sigma_x^2 = 1$, one can easily verify that

$$\frac{1}{K}\mathbf{Y}_P \mathbf{X}_P^H = \mathbf{H} + \Delta\mathbf{R}_{xv,P}^H \tag{6.70}$$

where

$$\Delta\mathbf{R}_{xv,P} \triangleq \frac{1}{K}\sum_{n=1}^{K} \mathbf{x}(n)\,\mathbf{v}^H(n)\,. \tag{6.71}$$

Using (6.61) and (6.70), (6.69) can be expressed as

$$\begin{aligned} \hat{\mathbf{Y}}_Q &\approx \mathbf{W}^H\mathbf{H} + \Delta\mathbf{W}^H\mathbf{H} + \mathbf{W}^H\Delta\mathbf{R}_{xv,P}^H, \\ &= \boldsymbol{\Sigma}_S^2\mathbf{V}_S^H + \mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{H} + \mathbf{W}^H\Delta\mathbf{R}_{xv,P}^H, \end{aligned} \tag{6.72}$$

which can be rewritten, utilizing (6.66), as

$$\hat{\mathbf{Y}}_Q = \mathbf{Y}_Q + \Delta\mathbf{Y}_Q \tag{6.73}$$

where

$$\Delta\mathbf{Y}_Q = \mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{H} + \mathbf{W}^H\Delta\mathbf{R}_{xv,P}^H. \tag{6.74}$$

Our next goal is to disclose an expression for $\Delta\mathbf{Q}$ in terms of the perturbation error $\Delta\mathbf{Y}_Q$ of $\mathbf{Y}_Q$ that is caused by noise. Let us consider the SVD of $\hat{\mathbf{Y}}_Q$:

$$\hat{\mathbf{Y}}_Q = \hat{\mathbf{U}}_Q\hat{\boldsymbol{\Sigma}}_Q\hat{\mathbf{V}}_Q^H = [(\mathbf{U}_Q + \Delta\mathbf{U}_Q)\,\mathbf{P}]\,(\boldsymbol{\Sigma}_Q + \Delta\boldsymbol{\Sigma}_Q)\,[(\mathbf{V}_Q + \Delta\mathbf{V}_Q)\,\mathbf{P}]^H \tag{6.75}$$

where $\mathbf{P}$ is a diagonal unitary matrix used to represent a general form of the SVD, since the SVD of $\hat{\mathbf{Y}}_Q$ is not unique. By utilizing $\hat{\mathbf{U}}_Q = (\mathbf{U}_Q + \Delta\mathbf{U}_Q)\,\mathbf{P}$ and $\hat{\mathbf{V}}_Q =$

$(\mathbf{V}_Q + \Delta\mathbf{V}_Q)\,\mathbf{P}$, we have

$$\begin{aligned}\Delta\mathbf{Q} &\triangleq \hat{\mathbf{Q}} - \mathbf{Q} = \hat{\mathbf{V}}_Q\hat{\mathbf{U}}_Q^H - \mathbf{V}_Q\mathbf{U}_Q^H \\ &\approx \Delta\mathbf{V}_Q\mathbf{U}_Q^H + \mathbf{V}_Q\Delta\mathbf{U}_Q^H. \end{aligned} \tag{6.76}$$

Using (6.67), (6.76) reduces to

$$\Delta\mathbf{Q} = \Delta\mathbf{V}_Q + \mathbf{V}_S\Delta\mathbf{U}_Q^H. \tag{6.77}$$

In general, if the noise subspace of $\mathbf{Y}_Q$ exists, the expressions for $\Delta\mathbf{V}_Q$ and $\Delta\mathbf{U}_Q$ can be easily derived in terms of the perturbation error $\Delta\mathbf{Y}_Q$ by utilizing the analysis results of the existing perturbation theory such as that in [71–77, 111, 112]. However, these results are not applicable to our case, since only the signal subspace exists in $\mathbf{Y}_Q$. To overcome this difficulty, we propose the following theorem to derive an expression for $\Delta\mathbf{Q}$ in terms of the perturbation error $\Delta\mathbf{Y}_Q$.

Theorem 6.2: Given an SVD of a full rank $M \times M$ matrix $\mathbf{Z}$, $\mathbf{Z} = \mathbf{U}_Z\mathbf{\Sigma}_Z\mathbf{V}_Z^H$, where $\mathbf{\Sigma}_Z = \mathrm{diag}\,(\sigma_{z_1}, \sigma_{z_2}, \cdots, \sigma_{z_M})$, for a small perturbed term $\Delta\mathbf{Z}$, the SVD of $\hat{\mathbf{Z}} = \mathbf{Z} + \Delta\mathbf{Z}$ is defined as

$$\hat{\mathbf{Z}} = \hat{\mathbf{U}}_Z\hat{\mathbf{\Sigma}}_Z\hat{\mathbf{V}}_Z^H = \left[\left(\mathbf{U}_Z + \Delta\mathbf{U}_Z\right)\mathbf{P}_Z\right]\left(\mathbf{\Sigma}_Z + \Delta\mathbf{\Sigma}_Z\right)\left[\left(\mathbf{V}_Z + \Delta\mathbf{V}_Z\right)\mathbf{P}_Z\right]^H \tag{6.78}$$

where $\mathbf{P}_Z$ is a diagonal unitary matrix whose role is the same as $\mathbf{P}$ in (6.75). Then, it can be shown that

$$\mathbf{\Omega} \triangleq \Delta\mathbf{U}_Z^H\mathbf{U}_Z + \mathbf{V}_Z^H\Delta\mathbf{V}_Z = \mathbf{\Gamma}_Z \circ \left(\mathbf{V}_Z^H\Delta\mathbf{Z}^H\mathbf{U}_Z - \mathbf{U}_Z^H\Delta\mathbf{Z}\mathbf{V}_Z\right) \tag{6.79}$$

where

$$\mathbf{\Gamma}_Z = \begin{bmatrix} \frac{1}{2\sigma_{z_1}} & \frac{1}{\sigma_{z_1}+\sigma_{z_2}} & \cdots & \frac{1}{\sigma_{z_1}+\sigma_{z_M}} \\ \frac{1}{\sigma_{z_2}+\sigma_{z_1}} & \frac{1}{2\sigma_{z_2}} & \cdots & \frac{1}{\sigma_{z_2}+\sigma_{z_M}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sigma_{z_M}+\sigma_{z_1}} & \frac{1}{\sigma_{z_M}+\sigma_{z_2}} & \cdots & \frac{1}{2\sigma_{z_M}} \end{bmatrix}. \tag{6.80}$$

The proof of the theorem is given in Appendix E.

In order to use *Theorem* 6.2, we first rewrite (6.77) as

$$\Delta\mathbf{Q} = \mathbf{V}_S \left(\Delta\mathbf{U}_Q^H\mathbf{I} + \mathbf{V}_S^H\Delta\mathbf{V}_Q\right). \tag{6.81}$$

With replacements $\Delta\mathbf{U}_Z = \Delta\mathbf{U}_Q$, $\mathbf{U}_Z = \mathbf{I}$, $\mathbf{V}_Z = \mathbf{V}_S$, $\Delta\mathbf{V}_Z = \Delta\mathbf{V}_Q$, $\mathbf{\Sigma}_Z = \mathbf{\Sigma}_S^2$ and $\Delta\mathbf{Z} = \Delta\mathbf{Y}_Q$ in (6.79), we immediately obtain

$$\Delta\mathbf{Q} = \mathbf{V}_S \left(\mathbf{\Gamma}_Q \circ \mathbf{\Pi}\right) \tag{6.82}$$

where

$$\mathbf{\Gamma}_Q = \begin{bmatrix} \frac{1}{2\sigma_{S_1}^2} & \frac{1}{\sigma_{S_1}^2+\sigma_{S_2}^2} & \cdots & \frac{1}{\sigma_{S_1}^2+\sigma_{S_{N_T}}^2} \\ \frac{1}{\sigma_{S_2}^2+\sigma_{S_1}^2} & \frac{1}{2\sigma_{S_2}^2} & \cdots & \frac{1}{\sigma_{S_2}^2+\sigma_{S_{N_T}}^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sigma_{S_{N_T}}^2+\sigma_{S_1}^2} & \frac{1}{\sigma_{S_{N_T}}^2+\sigma_{S_2}^2} & \cdots & \frac{1}{2\sigma_{S_{N_T}}^2} \end{bmatrix}, \tag{6.83}$$

$$\mathbf{\Pi} = \mathbf{V}_S^H\Delta\mathbf{Y}_Q^H - \Delta\mathbf{Y}_Q\mathbf{V}_S. \tag{6.84}$$

Note that σ_{S_i}, $(i = 1, 2, \cdots, N_T)$ is the diagonal element of $\mathbf{\Sigma}_S$. Substituting (6.74) into (6.84), we obtain

$$\mathbf{\Pi} = \mathbf{\Sigma}_S\mathbf{U}_S^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S + \mathbf{V}_S^H\Delta\mathbf{R}_{xv,P}\mathbf{U}_S\mathbf{\Sigma}_S - \mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{U}_S\mathbf{\Sigma}_S - \mathbf{\Sigma}_S\mathbf{U}_S^H\Delta\mathbf{R}_{xv,P}^H\mathbf{V}_S. \tag{6.85}$$

As a result, the perturbation error of $\mathbf{Q}$ has been expressed in terms of the perturbation errors, $\Delta\mathbf{R}_{xv}$ and $\Delta\mathbf{R}_{xv,P}$. In the next subsection, we will derive the MSE of the proposed signal-perturbation-free WR-based approach by utilizing the expression of $\Delta\mathbf{W}$ and $\Delta\mathbf{Q}$ given by (6.61) and (6.82), respectively.

6.4.3 MSE of the Proposed Signal-Perturbation-Free WR-based Method

The estimation error of the channel matrix $\mathbf{H}$ due to the perturbation error can be written as

$$\Delta\mathbf{H} \triangleq \hat{\mathbf{H}} - \mathbf{H} = \hat{\mathbf{W}}\hat{\mathbf{Q}}^H - \mathbf{W}\mathbf{Q}^H \approx \Delta\mathbf{W}\mathbf{Q}^H + \mathbf{W}\Delta\mathbf{Q}^H. \tag{6.86}$$

Substituting (6.61) and (6.82) into (6.86) gives

$$\Delta\mathbf{H} = \Delta\mathbf{R}_{xv}^H + \mathbf{U}_S\mathbf{\Sigma}_S\left(\mathbf{\Gamma}_Q \circ \mathbf{\Pi}^H\right)\mathbf{V}_S^H. \tag{6.87}$$

Thereby, we have

$$\mathrm{vec}\left(\Delta\mathbf{H}\right) = \mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right) + \left[\mathbf{V}_S^* \otimes \left(\mathbf{U}_S\mathbf{\Sigma}_S\right)\right]\left[\mathrm{vec}\left(\mathbf{\Gamma}_Q\right) \circ \mathrm{vec}\left(\mathbf{\Pi}^H\right)\right]. \tag{6.88}$$

Then, the MSE of the channel estimate can be calculated as

$$\mathrm{MSE}_{\mathrm{WR}} = \mathrm{Trace}\left\{\mathrm{E}\left[\mathrm{vec}\left(\Delta\mathbf{H}\right)\mathrm{vec}^H\left(\Delta\mathbf{H}\right)\right]\right\} = G_0 + G_1 + G_2 + G_2^* \tag{6.89}$$

where

$$G_0 = \frac{N_R N_T}{N}\sigma_v^2, \tag{6.90}$$

$$\begin{aligned} G_1 \;=\; & \mathrm{Trace}\left\{\mathrm{E}\left\{\left[\mathbf{V}_S^* \otimes \left(\mathbf{U}_S\mathbf{\Sigma}_S\right)\right]\left\{\left[\mathrm{vec}\left(\mathbf{\Gamma}_Q\right)\mathrm{vec}^H\left(\mathbf{\Gamma}_Q\right)\right]\right.\right.\right. \\ & \left.\left.\left. \circ \left[\mathrm{vec}\left(\mathbf{\Pi}^H\right)\mathrm{vec}^H\left(\mathbf{\Pi}^H\right)\right]\right\}\left[\mathbf{V}_S^* \otimes \left(\mathbf{U}_S\mathbf{\Sigma}_S\right)\right]^H\right\}\right\}, \end{aligned} \tag{6.91}$$

$$G_2 = \text{Trace}\left\{\text{E}\left\{\text{vec}\left(\Delta\mathbf{R}_{xv}^H\right)\left[\text{vec}\left(\mathbf{\Gamma}_Q\right) \circ \text{vec}\left(\mathbf{\Pi}^H\right)\right]^H \left[\mathbf{V}_S^* \otimes \left(\mathbf{U}_S\mathbf{\Sigma}_S\right)\right]^H\right\}\right\}. \quad (6.92)$$

In obtaining G_0 in (6.90), (D-3) has been used. For the computation of G_1 and G_2, we have the following theorem.

Theorem 6.3: G_1 and G_2 defined by (6.91) and (6.92), respectively, have the following closed-form expression.

$$G_1 = \frac{N_T^2\left(N-K\right)\sigma_v^2}{2KN}, \quad (6.93)$$

$$G_2 = 0. \quad (6.94)$$

The proof of the theorem is given in Appendix F. Using (6.90), (6.93) and (6.94) into (6.89) yields

$$\text{MSE}_{\text{WR}} = \frac{N_R N_T \sigma_v^2}{N} + \frac{N_T^2\left(N-K\right)\sigma_v^2}{2KN}. \quad (6.95)$$

The above result gives the MSE of the proposed signal-perturbation-free WR-based method in terms of the system configuration and the noise variance, when the signal energy σ_x^2 is normalized to unity. As will be shown in the next section, the theoretical value of the MSE given by (6.95) is highly consistent with the simulation result. It is to be noted that when $\Delta\mathbf{W} = \mathbf{0}$, the WR-based method turns to the ideal case, where the whitening matrix is obtained directly from the true channel matrix. In this case, the second term on the RHS of (6.72) vanishes due to $\Delta\mathbf{W} = \mathbf{0}$, and one can prove that the above analysis would lead to the following MSE expression

$$\text{MSE}_{\text{WR,ideal}} = \frac{N_T^2\sigma_v^2}{2K}. \quad (6.96)$$

It is of interest to note that (6.96) is the same as the Cramer-Rao bound of the WR-based method derived in [23].

6.5 Simulation Results

We consider a MIMO system with 4 transmit and 8 receive antennas, in which the QPSK modulation is used and a Rayleigh channel, whose elements are i.i.d. complex Gaussian variables with zero mean and unit variance, is assumed. Here, the orthogonal pilots are generated by using the scheme proposed in [25]. It is noted that a joint optimization approach has been proposed in [23] to improve the estimation of the whitening matrix. Although this approach can improve the estimation accuracy of the whitening matrix slightly, the complexity of this method is extremely high since it involves many iterations in the computation of the "fminunc" function in MATLAB, making its implementation very difficult for real-world applications. Thus, this approach is not considered in our experiments. Instead, an ideal WR-based method, which obtains the whitening matrix directly from the true channel matrix, is simulated in our experiments for comparison.

Experiment 1: *MSE versus SNR*

In the first experiment, the channel estimation performance in terms of the MSE versus the SNR is investigated. The simulation is undertaken based on 20000 Monte Carlo runs of the transmission of one data frame with 1000 slots out of which 100 are used as pilots. Fig. 6.2 shows the MSE plots of the LS method, the ideal WR-based method, the WR-based method, the nulling-based method and the proposed method with $\eta = 1, 4, 9, 16$, respectively. It is seen that the MSE of the proposed method is closest to that of the ideal WR-based method in comparison to the other methods

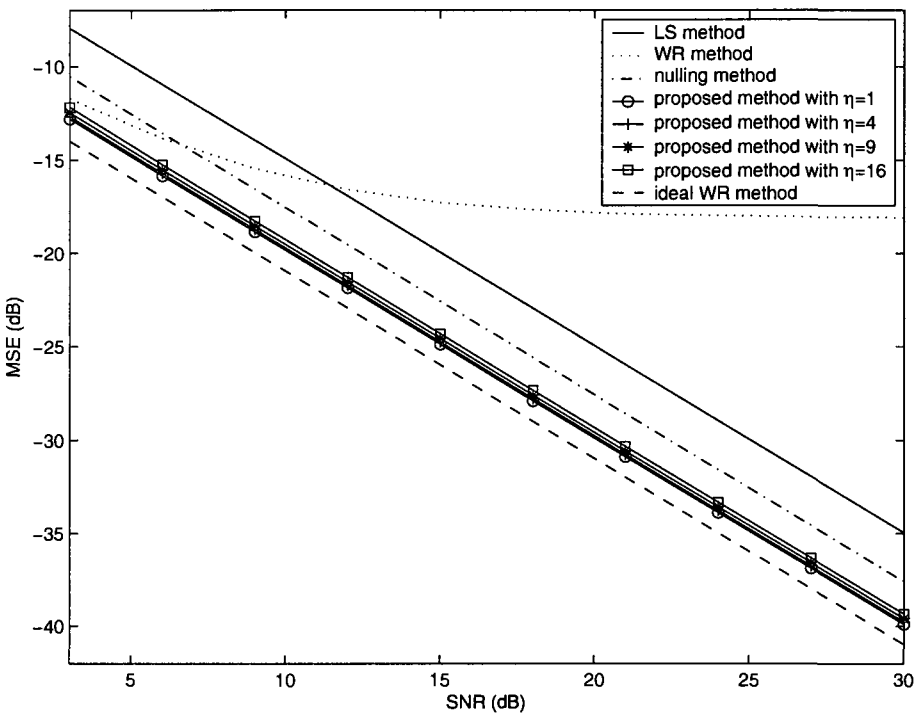

Figure 6.2: MSE versus SNR of the proposed WR method for a 4×8 frequency-flat channel. Also shown for reference is MSE of the LS, the nulling and the ideal WR methods. Pilot length is 100.

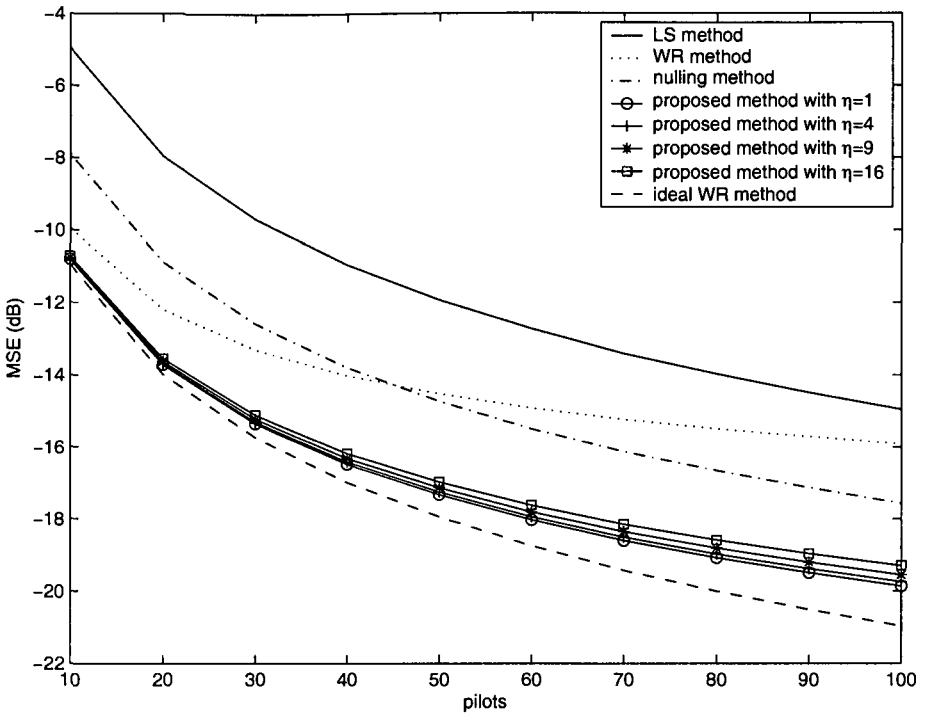

Figure 6.3: MSE versus pilot length for a 4×8 frequency-flat fading channel. SNR=10 dB.

irrespective of the choice of η. Interestingly, the different values of η only make a little difference on the MSE result. However, the number of slots for the transmission of $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ depends largely on the value of η. We have found that $\eta = 16$ is a very good choice for the proposed method, since it requires only 9 slots for $\mathbf{X}_{\text{pos}}$ and $\mathbf{X}_{\text{neg}}$ while significantly outperforming the WR-based, the nulling-based and the LS methods at all SNR levels.

Experiment 2: *MSE versus pilot length*

Here, we investigate the channel estimation performance versus the pilot length. Fig. 6.3 shows the MSE plots of the channel estimate from 20000 Monte Carlo runs

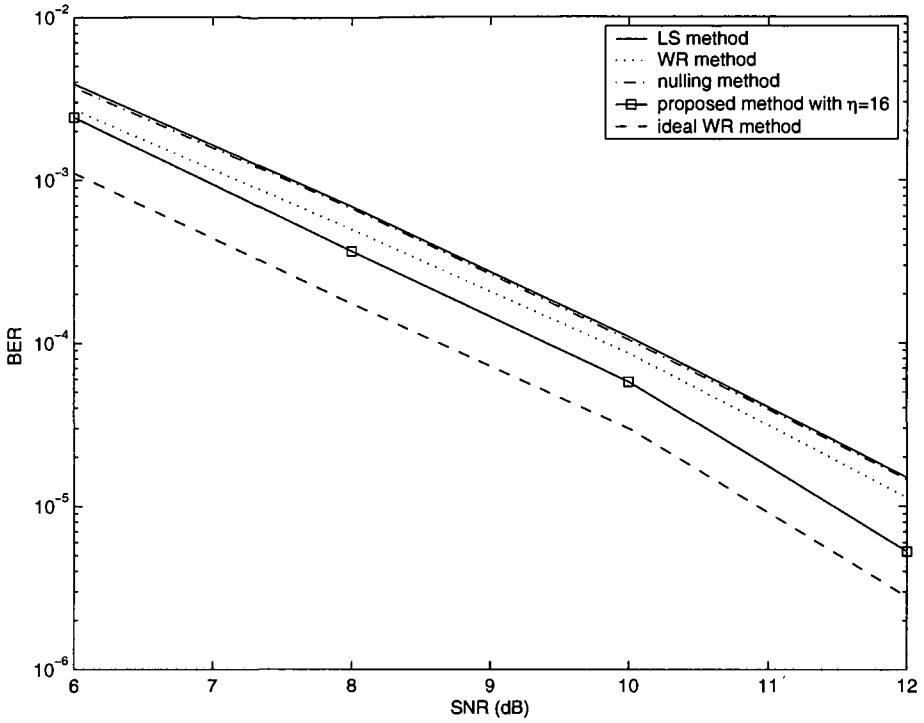

Figure 6.4: BER versus SNR for a 4×8 frequency-flat fading channel. Pilot length is 10.

of the transmission of one data frame of 1000 slots for an SNR of 10 dB, indicating a high estimation consistency of the proposed method with different values of η. Clearly, the improvement in the performance of the proposed method becomes more prominent compared to the WR-based method with the increase of the pilot length. For example, when the pilot length is increased to 100 from 30, the performance gain of the proposed method over the WR-based method is boosted to 3.4 dB from 2dB.

Experiment 3: *BER versus SNR*

Now, the BER performance is investigated by using the estimated channel matrix and an ordered vertical-Bell laboratories layered space time (V-BLAST) decoder.

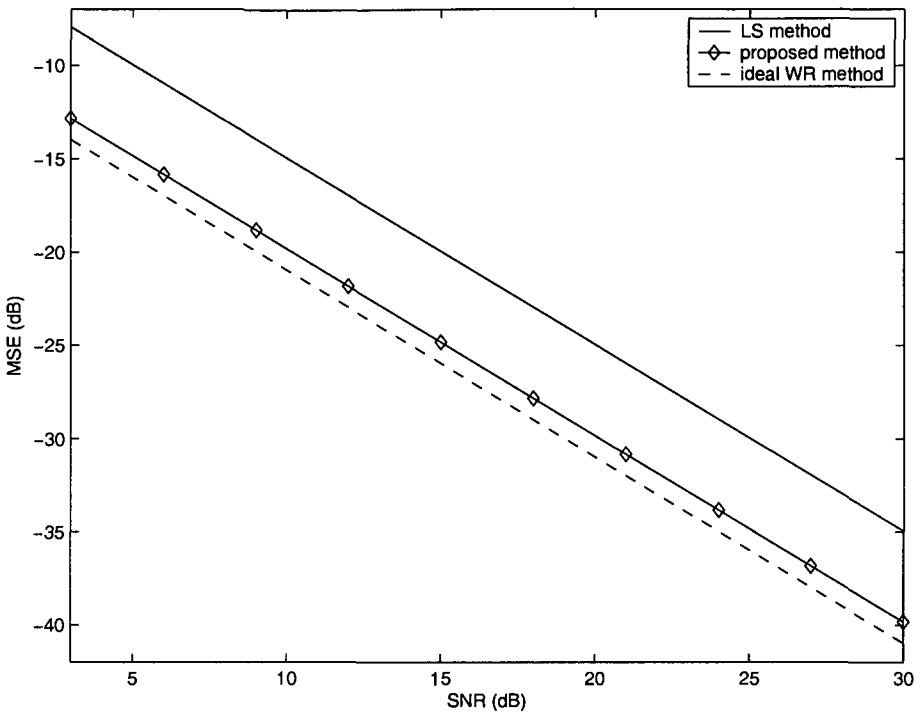

Figure 6.5: Theoretical MSE versus SNR for a 4×8 frequency-flat channel.

The simulation involves 50000 Monte Carlo runs of the transmission of one data frame with 100 slots out of which 10 are used as pilots. Fig. 6.4 shows the BER performance versus the SNR. It is seen that the BER performance of the proposed method with $\eta = 16$ becomes much better than that of the WR-based method with the increase of SNR. In particular, when SNR is 12 dB, the BER of the proposed method is about 3.3 dB over that of the WR-based method.

Experiment 4: *The theoretical value of MSE*

In this experiment, we compare the proposed method with the ideal WR-based method as well as the LS method in terms of their theoretical MSE expressions. Using the same conditions as in *Experiment* 1 and *Experiment* 2, the three MSE

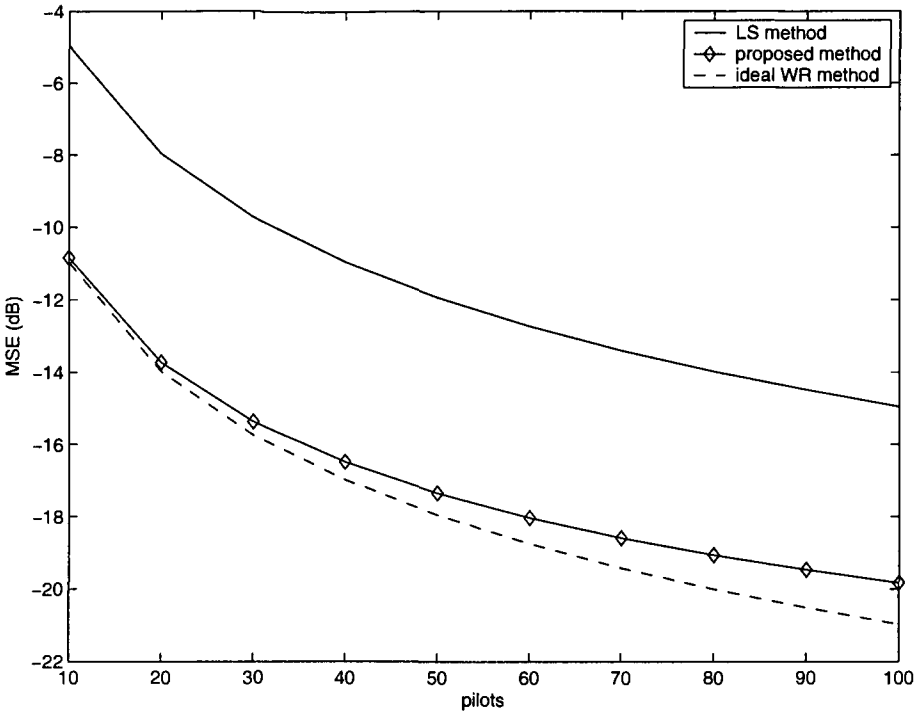

Figure 6.6: Theoretical MSE versus pilot length for a 4×8 frequency-flat channel.

plots calculated from (6.95), (6.96) and (6.11) are shown in Fig. 6.5 and Fig. 6.6, respectively. By comparing Fig. 6.5 to Fig. 6.2 and Fig. 6.6 to Fig. 6.3, one can see that the theoretical MSE values are consistent with the simulation results, confirming the high accuracy of the derivation of the MSE expressions in Section 6.4.

6.6 Conclusion

In this chapter, a new signal-perturbation-free WR-based semi-blind approach has been proposed for the estimation of frequency-domain channel. A perturbation analysis of two subspace-based semi-blind MIMO channel estimation algorithms has been performed, justifying that the WR-based method is efficient only in the low SNR case,

and the nulling-based method is a better choice when the SNR is moderate or high. To improve the performance of the WR-based method in the high SNR case, a new transmit structure, which contains user specific data bearing the information of the signal perturbation matrix, has been proposed for the cancellation of the signal perturbation error at the receiver. Furthermore, based on a novel perturbation analysis of the SVD of the square matrix without the noise subspace, a closed-form expression for the MSE of the new WR-based method integrating the proposed transmit scheme is derived. Simulation results have confirmed that, by using a small number of additional slots conveying the information of the autocorrelation matrix of the transmitted signal, a significant improvement both in the MSE of the channel estimate and in the BER of the data transmission can be achieved over the WR-based method as well as the nulling-based method for all SNRs.

Chapter 7

Signal-Perturbation-Free Linear Prediction based Estimation of Time-domain Channels

7.1 Introduction

By using the perturbation theory [71,72,77], our previous study on the LP-based blind channel estimation given in Chapter 3 has shown that some conventional LP-based blind algorithms such as those in [46,79,80] are subject to a signal perturbation error due to the finite data length effect in the calculation of the correlation matrix of the received signal. It means that these algorithms would suffer from a poor performance in the MIMO-OFDM channel estimation if the number of the OFDM symbols is not large enough. In contrast, the semi-blind algorithm proposed in Chapter 3 imposes an ideal nulling constraint on the channel matrix in the absence of noise and therefore, gives a better channel estimation performance. In Chapter 6, we have proposed a signal-perturbation-free (SPF) transmit scheme for the estimation of frequency-domain channels, based on which the signal perturbation error can be cancelled at the receiver, leading to a signal-perturbation-free semi-blind algo-

rithm for the estimation of the frequency-domain channels. In this chapter, we will extend the idea of the signal perturbation error cancellation to the estimation of time-domain channels. By developing a new signal-perturbation-free transmit scheme, we will achieve an efficient semi-blind solution for the estimation of time-domain MIMO-OFDM channel.

The rest of the chapter is organized as follows. Section 7.2 presents a novel signal-perturbation-free transmit scheme for the LP-based blind estimation of time-domain channels. We will first discuss the signal perturbation error in the LP-based blind estimation and show the signal perturbation error cancellation principle of the proposed scheme, and then develop a detailed transmit structure that can be used in the receiver to cancel the signal perturbation error in the estimated correlation matrix of the received signal. In Section 7.3, based on the proposed transmit scheme, a signal-perturbation-free semi-blind algorithm is developed for the estimation of the time-domain channel using the MIMO linear prediction in conjunction with a whitening rotation (WR) algorithm. The new approach is considered semi-blind in the sense that the ambiguity matrix in the WR algorithm is estimated by means of a training-based maximum likelihood method. Section 7.4 conducts a simulation study comprising a number of computer simulation based experiments to validate the proposed signal-perturbation-free semi-blind algorithm and show its significant advantages over some of the existing channel estimation techniques. Finally, Section 7.5 concludes the chapter by highlighting some of the contributions presented.

7.2 A Novel Signal-Perturbation-Free Transmit Scheme

In this section, the perturbation form of the correlation matrix of the transmitted signal used in linear prediction is first revisited, showing that the existence of perturbation error in the estimated correlation matrix of the received signal would degrade the performance of the linear prediction-based estimation methods. To improve the channel estimation performance of the LP-based blind algorithms, a signal-perturbation-free transmit scheme is proposed for the cancellation of the signal perturbation error [117, 118].

7.2.1 Revisit to Signal Perturbation in Linear-Prediction

It is clear from Section 2.4 that the estimation of the correlation matrix

$$\mathbf{R}_y(l) \triangleq \mathrm{E}\left\{\mathbf{y}(n)\,\mathbf{y}^H(n-l)\right\}, \text{ for } l = 0, 1, \cdots, P \tag{7.1}$$

of the received signal at the receiver plays a key role in the linear prediction-based blind method. For the sake of simplicity, only one OFDM symbol with K subcarriers is considered here. Letting

$$\mathbf{H}_A \triangleq [\mathbf{H}(0), \mathbf{H}(1), \cdots, \mathbf{H}(L-1)]$$

$$\mathbf{x}(n) \triangleq [x_1(n), \cdots, x_{N_T}(n)]^T$$

$$\mathbf{x}_L(n) \triangleq \left[\mathbf{x}^T(n) \cdots \mathbf{x}^T(n-L+1)\right]^T, \text{ for } n = 0, 1, \cdots, K-1$$

where $\mathbf{x}(n) = \mathbf{x}(K+n)$ for $n < 0$, the circular convolution (2.3) can be rewritten in the matrix form as

$$\mathbf{y}(n) = \mathbf{H}_A \mathbf{x}_L(n) + \mathbf{v}(n). \tag{7.2}$$

Using (7.2), the estimate of $\mathbf{R}_y(l)$ can be expressed as

$$\hat{\mathbf{R}}_y(l) = \hat{\mathrm{E}}\left[\mathbf{y}(n)\,\mathbf{y}^H(n-l)\right] - \delta_v^2\delta(l)\,\mathbf{I}_{N_R} = \mathbf{H}_A\hat{\mathbf{R}}_{x,L}(l)\,\mathbf{H}_A^H + \Delta\mathbf{R}_v(l) \tag{7.3}$$

where

$$\hat{\mathbf{R}}_{x,L}(l) \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}_L(n)\,\mathbf{x}_L^H(n-l) \tag{7.4}$$

and $\Delta\mathbf{R}_v(l)$ is the perturbation error introduced by the noise. Clearly, $\hat{\mathbf{R}}_{x,L}(l)$ can be rewritten as

$$\hat{\mathbf{R}}_{x,L}(l) = \begin{bmatrix} \hat{\mathbf{R}}_x(l) & \hat{\mathbf{R}}_x(l+1) & \cdots & \hat{\mathbf{R}}_x(l+L-1) \\ \hat{\mathbf{R}}_x(l-1) & \hat{\mathbf{R}}_x(l) & \cdots & \hat{\mathbf{R}}_x(l+L-2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{R}}_x(l-L+1) & \hat{\mathbf{R}}_x(l-L+2) & \cdots & \hat{\mathbf{R}}_x(l) \end{bmatrix} \tag{7.5}$$

where

$$\hat{\mathbf{R}}_x(l) \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}(n)\,\mathbf{x}^H(n-l)\,. \tag{7.6}$$

It has been proved in Chapter 3 that, when the transmitted frequency-domain signal is considered as an i.i.d. Gaussian process with zero mean and unit variance, the transmitted time-domain MIMO-OFDM signal is uncorrelated, i.e., $\mathbf{R}_x(l) \triangleq \mathrm{E}\left\{\mathbf{x}(n)\,\mathbf{x}^H(n-l)\right\} = \delta(l)\,\mathbf{I}_{N_T}$. Thus, (7.6) can be rewritten as

$$\hat{\mathbf{R}}_x(l) = \mathbf{R}_x(l) + \Delta\mathbf{R}_x(l) \tag{7.7}$$

where $\Delta\mathbf{R}_x(l)$ is the perturbation term of $\mathbf{R}_x(l)$ as given by

$$\Delta\mathbf{R}_x(l) \triangleq \frac{1}{K}\sum_{n=0}^{K-1}\mathbf{x}(n)\,\mathbf{x}^H(n-l) - \delta(l)\,\mathbf{I}_{N_T}. \tag{7.8}$$

Note that, when multiple OFDM symbols are used, $\hat{\mathbf{R}}_x(l)$ as well as $\hat{\mathbf{R}}_{x,L}(l)$ can be easily calculated by averaging the results obtained from each OFDM symbol.

It is clear from the above analysis that the existence of the signal perturbation terms $\Delta\mathbf{R}_x(l)$ would in general introduce the perturbation error to the linear prediction-based estimation methods. By conducting a perturbation analysis, it has been shown in [83] that, even in the noise-free case, the conventional blind algorithms such as those proposed in [39, 79, 80] are subject to a signal perturbation error, while the semi-blind algorithm is free of signal perturbation error since it gives an ideal nulling constraint on the channel matrix. It should also be pointed out that in the presence of noise, although both the semi-blind and the blind methods are subject to the noise perturbation terms, the semi-blind method still outperforms the blind one, since the perturbation introduced by the noise is in general significantly smaller than the signal perturbation. In order to improve the performance of the conventional blind approach, in the next section, we will propose an efficient signal-perturbation-free transmit scheme to cancel the signal perturbation error in MIMO-OFDM channel estimation. Our idea is to send information of the signal perturbation matrix $\Delta\mathbf{R}_x(l)$ to the receiver. The received version of this information will be then exploited to cancel the signal perturbation error. The idea of cancelling the signal perturbation error is first presented in the following subsection.

7.2.2 The Principle of Signal Perturbation Cancellation

The new idea begins with the computation of the signal perturbation matrix $\Delta\mathbf{R}_x(l)$ at the transmitter using the frequency-domain data. Define the transmitted frequency-domain signal vector at the k-th subcarrier as

$$\mathbf{X}(k) \triangleq [X_1(k), \cdots, X_{N_T}(k)]^T. \tag{7.9}$$

It is proved in Appendix G that the estimated correlation matrix of the time-domain signal $\mathbf{x}(n)$, $\hat{\mathbf{R}}_x(l)$, can be calculated as

$$\hat{\mathbf{R}}_x(l) = \frac{1}{K}\sum_{k=0}^{K-1}\mathbf{X}(k)\,\mathbf{X}^H(k)\,\phi^{-l}(k) \tag{7.10}$$

which, using (7.7), leads to

$$\Delta\mathbf{R}_x(l) = \frac{1}{K}\sum_{k=0}^{K-1}\mathbf{X}(k)\,\mathbf{X}^H(k)\,\phi^{-l}(k) - \delta(l)\,\mathbf{I}_{N_T}. \tag{7.11}$$

Interestingly, $\hat{\mathbf{R}}_x(l)$ given by (7.10) can be regarded as the IDFT of $\mathbf{X}(k)\,\mathbf{X}^H(k)$. Accordingly, the estimate of the correlation matrix of the received signal $\mathbf{y}(n)$, $\hat{\mathbf{R}}_y(l)$, can be represented as an IDFT of $\mathbf{Y}(k)\,\mathbf{Y}^H(k)$, where $\mathbf{Y}(k)$ is the frequency-domain version of $\mathbf{y}(n)$. In order to develop a new transmit scheme to cancel the perturbation error $\Delta\mathbf{R}_y(l) = \hat{\mathbf{R}}_y(l) - \mathbf{R}_y(l)$ due to $\Delta\mathbf{R}_x(l)$, we would like to express $\Delta\mathbf{R}_x(l)$ as an IDFT of a set of data $\mathbf{T}(k)$. From (7.1), (7.3) and (7.5), one can see that the range of l for $\Delta\mathbf{R}_x(l)$ is given by

$$-L_1 \le l \le L_2,\ (L_1 = L-1, L_2 = P+L-1). \tag{7.12}$$

As such, the size of the IDFT should be at least $K_T = L_1 + L_2 + 1$. Considering that the total number of subcarriers usually satisfies $K \gg L_1 + L_2 + 1$, it would be sufficient and convenient to choose $K_T = \frac{K}{M} \ge L_1 + L_2 + 1$ where M is the largest possible integer such that K_T is the smallest possible power of two. By using the periodicity property of IDFT, we have

$$\Delta\mathbf{R}_x(l) = \begin{cases} \sum_{k=0}^{K_T-1}\mathbf{T}(k)\,e^{j2\pi kl/K_T},\ (l = 0, 1, \cdots L_2) \\ \sum_{k=0}^{K_T-1}\mathbf{T}(k)\,e^{j2\pi k(K_T+l)/K_T},\ (l = -L_1, 1-L_1, \cdots -1) \end{cases}. \tag{7.13}$$

Clearly, (7.13) gives a K_T-size IDFT of $\mathbf{T}(k)$ with a gain of K_T, which means $\mathbf{T}(k)$ can easily be obtained from a K_T-size DFT of $\Delta\mathbf{R}_x(l)$ with a gain of $1/K_T$, namely,

$$\mathbf{T}(k) = \frac{1}{K_T}\left[\sum_{l=0}^{L_2} \Delta\mathbf{R}_x(l)\, \mathrm{e}^{-j2\pi kl/K_T} + \sum_{l=K_T-L_1}^{K_T-1} \Delta\mathbf{R}_x(l-K_T)\, \mathrm{e}^{-j2\pi kl/K_T}\right] \quad \text{for} k = 0, 1, \cdots, K_T - 1. \tag{7.14}$$

In what follows, we show that as long as the $N_T \times N_T$ matrix $\mathbf{T}(k)$ can be factorized into

$$\mathbf{T}(k) = \mathbf{T}_{\mathrm{L}}(k)\, \mathbf{T}_{\mathrm{R}}^{H}(k) \tag{7.15}$$

and $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}^{H}(k)$ are transmitted to the receiver, the signal perturbation error in $\hat{\mathbf{R}}_y(l)$ can be eliminated in the absence of noise.

First of all, we reveal that only a small number of subcarriers are needed to transmit $\mathbf{T}(k)$. Noting that $\mathrm{e}^{j2\pi k(l+K_T)/K_T} = \mathrm{e}^{j2\pi kl/K_T} = \mathrm{e}^{j2\pi kMl/K}$, (7.13) can be rewritten as

$$\Delta\mathbf{R}_x(l) = \sum_{k=0}^{K_T-1} \mathbf{T}(k)\, \phi^{-Ml}(k), \text{ for } l = -L_1, -L_1+1, \cdots L_2. \tag{7.16}$$

Interestingly, (7.16) corresponds to an M-rate decimated version of the K-size DFT of $\mathbf{T}'(k)$, where $\mathbf{T}'(k)$ is an up-sampled version of $\mathbf{T}(k)$ by a factor of M, i.e., it can easily be obtained by inserting $M-1$ zero matrices following each $\mathbf{T}(k)$. This observation gives us an idea that $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ need be transmitted over the kM-th subcarriers only ($k = 0, 1, \cdots, K_T - 1$), which is very advantageous to practical applications, since only a small number of subcarriers are required for the transmission of $\mathbf{T}(k)$.

In order to exploit the received version of $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ for the cancellation of signal perturbation error $\Delta\mathbf{R}_y(l)$, let us first express the estimated correlation matrix

of the received signal in terms of the channel matrix and the estimate of the correlation matrix of the transmitted signal. Noting that the received frequency-domain signal at the k-th subcarrier $\mathbf{Y}(k)$ can be written as

$$\mathbf{Y}(k) = \mathbf{H}_{\mathrm{F}}(k)\,\mathbf{X}(k) \tag{7.17}$$

where $\mathbf{H}_{\mathrm{F}}(k)$, $(k = 0, 1, \cdots, K-1)$ is the frequency-domain channel matrix defined by

$$\mathbf{H}_{\mathrm{F}}(k) \triangleq \sum_{l=0}^{L-1} \mathbf{H}(l)\,\phi^{l}(k), \ (k = 0, 1, \cdots, K-1). \tag{7.18}$$

In a manner similar to the derivation of (7.10) as shown in Appendix G, one obtains

$$\hat{\mathbf{R}}_{y}(l) = \frac{1}{K}\sum_{k=0}^{K-1} \mathbf{Y}(k)\,\mathbf{Y}^{H}(k)\,\phi^{-l}(k), \tag{7.19}$$

which, using (7.17) and (7.18), leads to

$$\hat{\mathbf{R}}_{y}(l) = \sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1} \mathbf{H}(l_1)\,\frac{1}{K}\sum_{k=0}^{K-1}\left[\mathbf{X}(k)\,\mathbf{X}^{H}(k)\,\phi^{-(l-l_1+l_2)}(k)\right]\mathbf{H}^{H}(l_2). \tag{7.20}$$

Further by using (7.10) into (7.20), we have

$$\hat{\mathbf{R}}_{y}(l) = \sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1} \mathbf{H}(l_1)\,\hat{\mathbf{R}}_{x}(l-l_1+l_2)\,\mathbf{H}^{H}(l_2). \tag{7.21}$$

On the other hand, by using the received noise-free version of the user specific data $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$, denoted as

$$\mathbf{Y}_{\mathrm{TL}}(k) = \mathbf{H}_{\mathrm{F}}(kM)\,\mathbf{T}_{\mathrm{L}}(k), \tag{7.22}$$

$$\mathbf{Y}_{\mathrm{TR}}(k) = \mathbf{H}_{\mathrm{F}}(kM)\,\mathbf{T}_{\mathrm{R}}(k), \tag{7.23}$$

we can construct matrices $\mathbf{R}_{\mathrm{YT}}(l)$, $(l = 0, 1, \cdots, L-1)$ as

$$\mathbf{R}_{\mathrm{YT}}(l) = \sum_{k=0}^{K_T-1} \mathbf{Y}_{\mathrm{TL}}(k)\,\mathbf{Y}_{\mathrm{TR}}^{H}(k)\,\phi^{-Ml}(k). \tag{7.24}$$

Using (7.18), (7.22) and (7.23) in (7.24) yields

$$\mathbf{R}_{\mathrm{YT}}(l) = \sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1}\mathbf{H}(l_1)\sum_{k=0}^{K_T-1}\left[\mathbf{T}_{\mathrm{L}}(k)\,\mathbf{T}_{\mathrm{R}}^{H}(k)\,\phi^{-M(l-l_1+l_2)}(k)\right]\mathbf{H}^{H}(l_2), \tag{7.25}$$

which, by using (7.15) and (7.16), can be rewritten as

$$\mathbf{R}_{\mathrm{YT}}(l) = \sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1}\mathbf{H}(l_1)\,\Delta\mathbf{R}_x(l-l_1+l_2)\,\mathbf{H}^{H}(l_2). \tag{7.26}$$

From (7.21) and (7.26), and noting that $\Delta\mathbf{R}_x(l) = \hat{\mathbf{R}}_x(l) - \mathbf{R}_x(l)$ and $\Delta\mathbf{R}_y(l) = \hat{\mathbf{R}}_y(l) - \mathbf{R}_y(l)$, it is now clear that $\mathbf{R}_{\mathrm{YT}}(l)$ gives exactly the signal perturbation error $\Delta\mathbf{R}_y(l)$. Therefore, the correlation matrix of the received signal without the signal perturbation error can be calculated by

$$\hat{\mathbf{R}}'_y(l) = \hat{\mathbf{R}}_y(l) - \mathbf{R}_{\mathrm{YT}}(l) \tag{7.27}$$

$$= \sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1}\mathbf{H}(l_1)\,\mathbf{R}_x(l-l_1+l_2)\,\mathbf{H}^{H}(l_2) = \mathbf{R}_y(l). \tag{7.28}$$

The above discussion shows that via the transmission of $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$, the signal perturbation error in the receiver has been completely eliminated in the noise free case. From (7.28), it is interesting to note that, under the assumption of $\mathbf{R}_x(l) = \delta(l)\mathbf{I}_{N_T}$, the ideal correlation matrix of the received signal, $\mathbf{R}_y(l)$, can be expressed in terms of the channel matrices $\mathbf{H}(l)$, i.e.,

$$\mathbf{R}_y(l) = \sum_{i=l}^{L-1}\mathbf{H}(i)\,\mathbf{H}^{H}(i-l),\ (l = 0, 1, \cdots, L-1). \tag{7.29}$$

It should be mentioned that, in the noisy case, $\hat{\mathbf{R}}'_y(l)$ calculated using (7.27) would be, in general, different from $\mathbf{R}_y(l)$ obtained from (7.28) or (7.29) due to the existence of both signal and noise perturbation errors in $\hat{\mathbf{R}}_y(l)$. However, it is consistently found through a computer simulation study that $\left\|\hat{\mathbf{R}}'_y(l) - \mathbf{R}_y(l)\right\|_F^2 \ll \left\|\hat{\mathbf{R}}_y(l) - \mathbf{R}_y(l)\right\|_F^2$

as seen from Fig. 7.2 in the simulation section. This is because the noise perturbation is much smaller than the signal perturbation whereas the latter has been completely cancelled in $\hat{\mathbf{R}}'_y(l)$ by using (7.27). This means that the estimation accuracy of the correlation matrix of the received noisy signal has been significantly improved using $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$. In the following subsection, we will propose a detailed structure for $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ for the implementation of the new transmit scheme.

7.2.3 The Signal-Perturbation-Free Transmit Structure

Our idea is to obtain $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ by using the singular value decomposition (SVD) technique. Performing the SVD on $\mathbf{T}(k)$ gives

$$\mathbf{T}(k) = \mathbf{U}_T(k)\,\mathbf{\Sigma}_T(k)\,\mathbf{V}_T^H(k) \tag{7.30}$$

where

$$\mathbf{U}_T(k) = [\mathbf{u}_{T,1}(k), \mathbf{u}_{T,2}(k), \cdots, \mathbf{u}_{T,N_T}(k)],$$

$$\mathbf{V}_T(k) = [\mathbf{v}_{T,1}(k), \mathbf{v}_{T,2}(k), \cdots, \mathbf{v}_{T,N_T}(k)],$$

and $\mathbf{\Sigma}_T(k)$ is a diagonal matrix conposed of the singular values $\sigma_{T,i}(k)$, $(i = 1, 2, \cdots, N_T)$ of $\mathbf{T}(k)$. We now construct the matrices $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ using the singular values $\sigma_{T,i}(k)$ and the singular vectors $\mathbf{u}_{T,i}(k)$ and $\mathbf{v}_{T,i}(k)$.

Note that the total power of N_T transmit antennas at one subcarrier in each OFDM symbol can be written as $\delta_{\mathrm{int}} \triangleq N_T$. Clearly, the power required to transmit $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ depends on $\sigma_{T,i}(k)$. It is found from extensive computer simulations that the value of $\sigma_{T,i}(k)$ is much smaller than δ_{int}. In order to ensure a reliable transmission of the SPF data in noisy conditions, we should allow the use of one

or multiple OFDM symbols at the kM-th subcarrier to carry an amplified version of $\sigma_{T,i}(k)$. To this end, we split $\sigma_{T,i}(k)$ into $M_{\mathrm{T},i}(k) \geq 0$ times of δ_{int} plus one fractional term $\delta_{\mathrm{T},i}(k)$ as

$$\frac{\sigma_{T,i}(k)}{\eta} = M_{\mathrm{T},i}(k)\,\delta_{\mathrm{int}} + \delta_{\mathrm{T},i}(k) \tag{7.31}$$

where

$$M_{\mathrm{T},i}(k) = \lfloor \frac{\sigma_{T,i}(k)}{\eta\delta_{\mathrm{int}}} \rfloor, \tag{7.32}$$

$$\delta_{\mathrm{T},i}(k) = \frac{\sigma_{T,i}(k)}{\eta} - M_{\mathrm{T},i}(k)\,\delta_{\mathrm{int}}. \tag{7.33}$$

By letting

$$\mathbf{T}_{\mathrm{L-int},i}(k) = \sqrt{\delta_{\mathrm{int}}}\mathbf{u}_{T,i}(k),$$

$$\mathbf{T}_{\mathrm{L-frac},i}(k) = \sqrt{\delta_{\mathrm{T},i}(k)}\mathbf{u}_{T,i}(k),$$

one can construct an $N_T \times [M_{\mathrm{T},i}(k)+1]$ matrix $\mathbf{T}_{\mathrm{L},i}(k)$ for the i-th singular value by stacking $M_{\mathrm{T},i}(k)$ consecutive vectors $\mathbf{T}_{\mathrm{L-int},i}(k)$ and one vector $\mathbf{T}_{\mathrm{L-frac},i}(k)$. In a similar manner, by using the right singular vector $\mathbf{v}_{T,i}(k)$, an $N_T \times [M_{\mathrm{T},i}(k)+1]$ matrix $\mathbf{T}_{\mathrm{R},i}(k)$ that satisfies

$$\mathbf{T}_{\mathrm{L},i}(k)\,\mathbf{T}_{\mathrm{R},i}^{H}(k) = \frac{\sigma_{T,i}(k)}{\eta}\mathbf{u}_{T,i}(k)\,\mathbf{v}_{T,i}^{H}(k)$$

can be constructed. Thus, the complete $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ can be formed as

$$\mathbf{T}_{\mathrm{L}}(k) = [\mathbf{T}_{\mathrm{L},1}(k), \mathbf{T}_{\mathrm{L},2}(k), \cdots, \mathbf{T}_{\mathrm{L},N_T}(k)], \tag{7.34}$$

$$\mathbf{T}_{\mathrm{R}}(k) = [\mathbf{T}_{\mathrm{R},1}(k), \mathbf{T}_{\mathrm{R},2}(k), \cdots, \mathbf{T}_{\mathrm{R},N_T}(k)]. \tag{7.35}$$

Obviously, each of $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ has a total of $N_T + \sum_{i=1}^{N_T} M_{\mathrm{T},i}(k)$ columns.

From the above discussion, a new transmit structure, which consists of user's data, pilots and the signal-perturbation-free (SPF) data $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$, can be

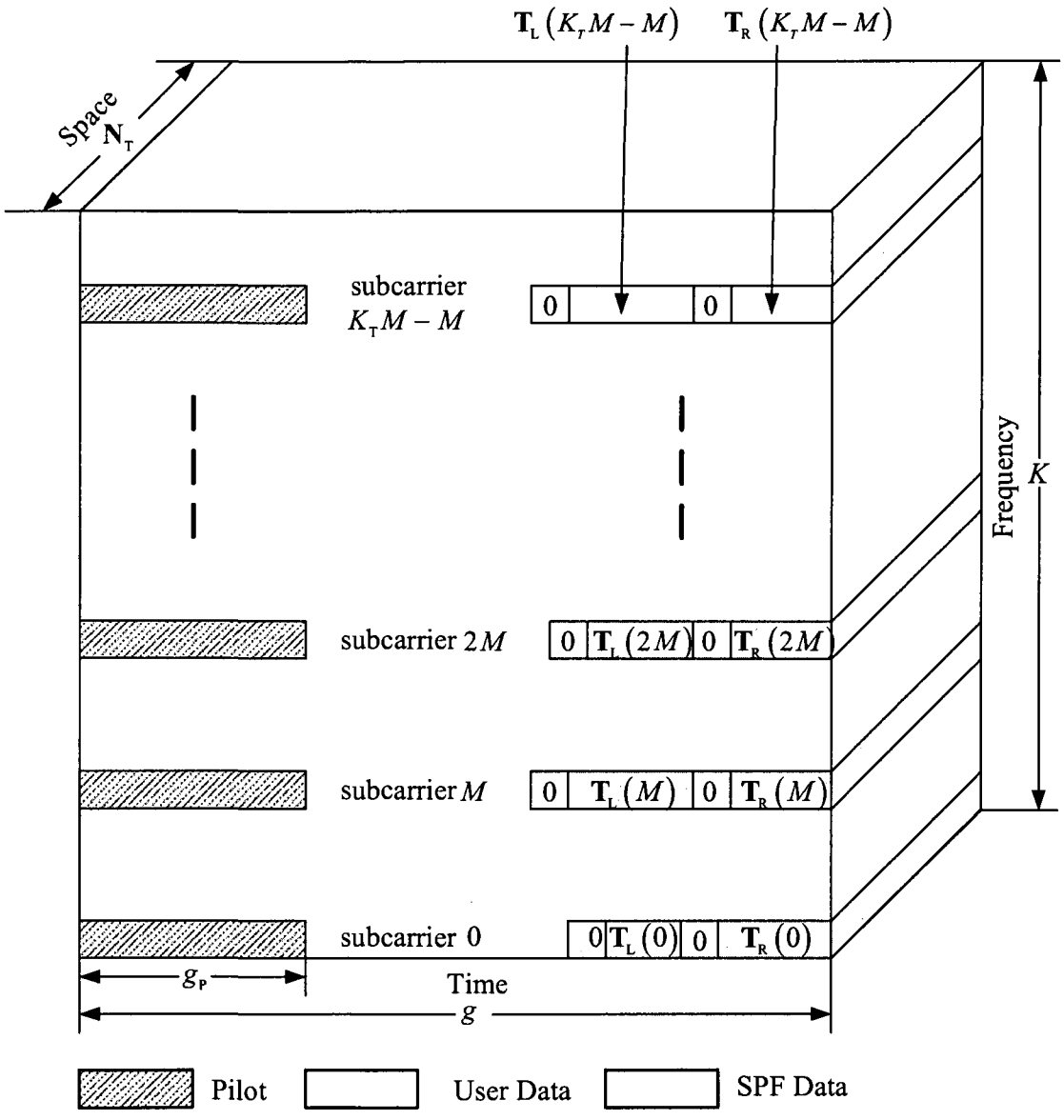

Figure 7.1: Signal-perturbation-free transmit structure for MIMO-OFDM systems

designed as shown in Fig.7.1. Here the pilots allocated in the first g_p OFDM symbols are transmitted only at the subcarriers $k = 0, M, \cdots, (K_T - 1) M$ as normally specified in OFDM systems. As such, the SPF data $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ are presumably transmitted at the same subcarriers as the pilot. Note that the pilots will be utilized to estimate the ambiguity matrix of the blind channel estimation, as will be explained in the next section. The zero symbol, namely, the symbol with zero amplitude, proceeding the SPF data is used to identify $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$. It is now clear that the total column size of $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ is inversely proportional to the scaling factor η. One can easily find a value of η such that as low as $2N_T$ symbols are required for the transmission of $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$. In general, the choice of η should depend on the number of the transmit antennas as well as the length of the user data. Our simulations show that, for a 2×4 MIMO-OFDM system, with a properly chosen value of η, on the average, up to 8 OFDM symbols per SPF subcarrier are required for the transmission of SPF data to sufficiently suppress the signal perturbation error. This overhead is negligible as compared to the pilot budget required by OFDM systems. It should be mentioned that the construction of $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ is rather simple since they are readily obtained by stacking a scaled version of the left and right singular vectors of $N_T \times N_T$ matrix $\mathbf{T}(k)$. Moreover, $\mathbf{T}(k)$ can easily be computed by a K_T-size DFT of $\Delta \mathbf{R}_x(l)$ where K_T is a small number, for example, $K_T = 8$ in our study. It should be noted that the SPF data $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ suffers from the channel noise during its transmission just as the pilot and user data do. However, this noise effect has been significantly reduced by using multiple OFDM symbols to bear properly scaled singular vectors of $\mathbf{T}(k)$. As will be seen from Fig.7.2 in Section

7.4, the proposed signal perturbation cancellation scheme is indeed very efficient at moderate to high SNR levels.

7.3 Proposed Signal-Perturbation-Free LP-based Semi-Blind Algorithm

7.3.1 Signal-Perturbation-Free LP-based Semi-Blind Channel Estimation

By employing the above signal-perturbation-free transmit scheme, the correlation matrix of the received signal after the signal perturbation cancellation, $\hat{\mathbf{R}}'_y(l)$, can be obtained from (7.27). Then, by employing (2.37) and (2.38), the estimate of $\tilde{\mathbf{R}}_{n-1}$ and $\ddot{\mathbf{R}}_n$ is obtained, which is to be used to compute the MIMO linear predictor (2.39). In the following, we will derive a blind channel estimate subject to an ambiguity matrix by utilizing the the properties of linear prediction in (2.41) and (2.42). The ambiguity matrix will then be estimated using the pilot-assisted maximum likelihood method.

We first derive the expression for $\mathbf{H}(i)$, $(i = 1, 2, \cdots, L-1)$ in terms of $\mathbf{H}(0)$ in the absence of noise. Letting the predictor length $P = L - 1$ and substituting (3.8) and (3.9) into (3.10), one can obtain

$$-\mathbf{P}_P(1)\mathbf{H}(0) + \mathbf{H}(1) = \mathbf{0}, \tag{7.36}$$

$$-\mathbf{P}_P(2)\mathbf{H}(0) - \mathbf{P}_P(1)\mathbf{H}(1) + \mathbf{H}(2) = \mathbf{0}, \tag{7.37}$$

$$\vdots$$

$$-\mathbf{P}_P(L-1)\mathbf{H}(0) - \mathbf{P}_P(L-2)\mathbf{H}(1) - \cdots$$

$$-\mathbf{P}_P(1)\mathbf{H}(L-2) + \mathbf{H}(L-1) = \mathbf{0}. \tag{7.38}$$

From (7.36)-(7.38), one can derive

$$\mathbf{H}(i) = \mathbf{P}_R(i)\,\mathbf{H}(0)\,,\ (i = 1, 2, \cdots, L-1) \tag{7.39}$$

where $\mathbf{P}_R(i)$ in (7.39) can be iteratively calculated by using (7.36)-(7.38), namely, $\mathbf{P}_R(1) = \mathbf{P}_P(1)\,,\ \mathbf{P}_R(2) = \mathbf{P}_P(2) + \mathbf{P}_P(1)\,\mathbf{P}_P(1),\ \mathbf{P}_R(3) = \mathbf{P}_P(3) + \mathbf{P}_P(2)\,\mathbf{P}_P(1) + \mathbf{P}_P(1)\,\mathbf{P}_P(2) + \mathbf{P}_P(1)\,\mathbf{P}_P(1)\,\mathbf{P}_P(1)$, etc.. Eq. (7.39) indicates that once the channel matrix $\mathbf{H}(0)$ is available, the matrices at other taps $\mathbf{H}(i)$ $(i = 1, 2, \cdots, L-1)$, can be obtained. Therefore, the key to the estimation of $\mathbf{H}(i)$ is to obtain $\mathbf{H}(0)$.

We now use the whitening rotation (WR) algorithm [21, 23] to determine $\mathbf{H}(0)$. The idea of WR method starts from the decomposition of $\mathbf{H}(0)$ as

$$\mathbf{H}(0) = \mathbf{W}_0\mathbf{Q}_0^H \tag{7.40}$$

where $\mathbf{W}_0$ is a whitening matrix and $\mathbf{Q}_0$ is a unitary rotation matrix. By using the singular value decomposition (SVD) of $\mathbf{H}(0)$,

$$\mathbf{H}(0) = \mathbf{U}_0\mathbf{\Sigma}_0\mathbf{V}_0^H, \tag{7.41}$$

one can see that a possible choice of $\mathbf{W}_0$ and $\mathbf{Q}_0$ is $\mathbf{U}_0\mathbf{\Sigma}_0$ and $\mathbf{V}_0$. In [21, 23], $\mathbf{W}_0 = \mathbf{U}_0\mathbf{\Sigma}_0$ was estimated by a subspace-based method and $\mathbf{Q}_0$ was obtained via a training-based ML method. However, this method was proposed for the frequency-flat MIMO channels, and it is not applicable to our case. We now estimate $\mathbf{W}_0$ using the MIMO linear prediction. Substituting (7.41) into (2.42) gives

$$\boldsymbol{\delta}_{\tilde{\mathbf{y}},P}^2 = \mathbf{U}_0\mathbf{\Sigma}_0\mathbf{\Sigma}_0^H\mathbf{U}_0^H, \tag{7.42}$$

implying that $\mathbf{W}_0 = \mathbf{U}_0\mathbf{\Sigma}_0$ can easily be estimated from the SVD of the prediction error $\boldsymbol{\delta}_{\tilde{\mathbf{y}},P}^2$, provided that the MIMO linear-prediction has been performed as discussed

in Chapter 3. Using the estimate of $\mathbf{W}_0$ along with (7.39) and (7.40), we have thus far estimated in a blind fashion the channel matrix $\mathbf{H}(i)$, $(i = 1, 2, \cdots, L-1)$ subject to an unitary ambiguity matrix $\mathbf{Q}_0$.

We now propose a new ML method for the estimation of $\mathbf{Q}_0$ by using the training pilots available in the kM-th subcarriers $(k = 0, 1, \cdots, K_T - 1)$ during the first g_p OFDM symbols as shown in Fig.7.1. By stacking a pilot matrix from all g_p symbols at the kM-th subcarrier as

$$\mathbf{X}_P(k) = [\mathbf{X}(kM, 1), \mathbf{X}(kM, 2), \cdots, \mathbf{X}(kM, g_p)], \tag{7.43}$$

a received version of $\mathbf{X}_P(k)$ can be expressed as

$$\mathbf{Y}_P(k) = \mathbf{H}_{\mathrm{F}}(kM)\,\mathbf{X}_P(k). \tag{7.44}$$

On the other hand, using (7.39) in (7.18) gives

$$\mathbf{H}_{\mathrm{F}}(k) = \mathbf{P}_{R,F}(k)\,\mathbf{H}(0) \tag{7.45}$$

where

$$\mathbf{P}_{R,F}(k) \triangleq \sum_{l=0}^{L-1} \mathbf{P}_R(i)\,\phi^l(k). \tag{7.46}$$

By defining

$$\mathbf{Y}_Q(k) \triangleq \mathbf{W}_0^H \mathbf{P}_{R,F}^H(k)\,\mathbf{Y}_P(k)\,\mathbf{X}_P^H(k), \tag{7.47}$$

we have proved in Appendix H that the rotation matrix $\mathbf{Q}_0$ can be calculated as

$$\mathbf{Q}_0 = \mathbf{V}_Q \mathbf{U}_Q^H \tag{7.48}$$

where $\mathbf{U}_Q$ and $\mathbf{V}_Q$ are obtained from an SVD of the matrix

$$\mathbf{Y}_Q \triangleq \sum_{k=1}^{K_T} \mathbf{Y}_Q(k), \tag{7.49}$$

namely,

$$\mathbf{Y}_Q = \mathbf{U}_Q \boldsymbol{\Sigma}_Q \mathbf{V}_Q^H. \tag{7.50}$$

Having estimated $\mathbf{W}_0$ and $\mathbf{Q}_0$, the channel matrix $\mathbf{H}(l)$, $(l = 0, 1, \cdots, L-1)$ can easily be obtained from (7.39) and (7.40).

7.3.2 Identification of the Ill Condition

It should be mentioned that, in the noisy case the blind estimation part of the proposed channel estimation algorithm suffers from the common ill-conditioning problem as many other LP-based blind methods like the one proposed in [46]. As mentioned in [46], the ill condition arises from the pseudo-inverse of the matrix $\hat{\tilde{\mathbf{R}}}_{n-1}$ in (2.37). Performing the SVD of $\hat{\tilde{\mathbf{R}}}_{n-1}$ gives

$$\tilde{\mathbf{R}}_{n-1} = \mathbf{U}_{Rn} \boldsymbol{\Sigma}_{Rn} \mathbf{V}_{Rn}^H. \tag{7.51}$$

Based on (7.51), the pseudo-inverse of the matrix $\hat{\tilde{\mathbf{R}}}_{n-1}$ can be obtained from

$$\tilde{\mathbf{R}}_{n-1}^{\dagger} = \mathbf{V}_{Rn} \boldsymbol{\Sigma}_{Rn}^{-1} \mathbf{U}_{Rn}^H. \tag{7.52}$$

Occasionally, $\boldsymbol{\Sigma}_{Rn}$ is quite ill-conditioned. To solve this problem, a regularized pseudo-inverse

$$\tilde{\mathbf{R}}_{n-1}^{\dagger} = \mathbf{V}_{Rn} \left(\boldsymbol{\Sigma}_{Rn} + \sigma_{Rn} \mathbf{I} \right)^{-1} \mathbf{U}_{Rn}^H \tag{7.53}$$

with $\sigma_{Rn} > 0$, was used in [46]. However, it has been found through a large amount of computer simulations that in ill-conditioned cases, the performance of the LP-based methods using the regularized scheme is still much worse than the traditional LS method, leading to a poor overall channel estimation performance. Therefore,

we suggest replacing the LP-based channel estimation solution with our previously developed nulling-based semi-blind algorithm in Chapter 3, once the ill-condition case is identified. Thus, what remains is to identify the ill conditions of the proposed signal-perturbation-free LP-based method [119].

Here, we first give an LS-based criterion to identify the ill condition.

1. Calculate the cost of the LS estimate:

$$\Delta_{\mathrm{LS}} = \left\| \mathbf{Y}_{\mathrm{pilot}} - \tilde{\mathbf{A}}\hat{\mathbf{h}}_{\mathrm{LS}} \right\|_F^2 \tag{7.54}$$

   where $\hat{\mathbf{h}}_{\mathrm{LS}}$ is the LS channel estimate resulting from the method in [25].

2. Calculate the cost of the SPF semi-blind estimate:

$$\Delta_{\mathrm{SB}} = \left\| \mathbf{Y}_{\mathrm{pilot}} - \tilde{\mathbf{A}}\hat{\mathbf{h}}_{\mathrm{SB}} \right\|_F^2 \tag{7.55}$$

   where $\hat{\mathbf{h}}_{\mathrm{SB}}$ is the semi-blind channel estimate, which can be obtained in the previous subsection.

3. Define a coefficient ρ_1 to evaluate the deviation of Δ_{SB} from Δ_{LS} as

$$\rho_1 = \frac{|\Delta_{\mathrm{SB}} - \Delta_{\mathrm{LS}}|}{\Delta_{\mathrm{LS}}}. \tag{7.56}$$

4. If ρ_1 is larger than a predetermined threshold τ_1, the SPF semi-blind channel estimate is considered as ill conditioned.

As the above criterion may not cover all the ill-conditioned cases, we suggest another testing to further improve the reliability of the ill condition identification.

1. Obtain a difference matrix of $\boldsymbol{\delta}^2_{\tilde{\mathbf{y}},P}$ between the LP-method and the LS method as

$$\boldsymbol{\Upsilon} = \hat{\boldsymbol{\delta}}^2_{\tilde{\mathbf{y}},P} - \hat{\mathbf{H}}_{\mathrm{LS}}(0)\,\hat{\mathbf{H}}^H_{\mathrm{LS}}(0)\,. \tag{7.57}$$

where $\hat{\boldsymbol{\delta}}^2_{\tilde{\mathbf{y}},P}$ is an estimate of $\boldsymbol{\delta}^2_{\tilde{\mathbf{y}},P}$ given in (2.40), $\hat{\mathbf{H}}_{\mathrm{LS}}(0)$ is the first tap of the LS channel estimate.

2. Perform the eigenvalue decomposition on $\boldsymbol{\Upsilon}$, giving the largest and the second largest eigenvalues as $\sigma_{\Upsilon 1}$ and $\sigma_{\Upsilon 2}$.

3. Define a coefficient ρ_2 as

$$\rho_2 = \sigma_{\Upsilon 1}/\sigma_{\Upsilon 2}. \tag{7.58}$$

4. If ρ_2 is larger than a predetermined threshold τ_2, the SPF semi-blind channel estimate is considered as ill-conditioned.

We will show in the next section that by using both the above ill-condition test rules, the proposed semi-blind algorithm performs very well.

7.4 Simulation Results

We consider a MIMO-OFDM system with 2 transmit and 4 receive antennas. The number of subcarriers is set to 512, the length of cyclic prefix is 10. In this simulation, the QPSK modulation is used and an SUI-3 type MIMO channel is considered. In particular, the channel is modelled as a 3-tap MIMO-FIR filter, in which each tap corresponds to a 2×4 random matrix whose elements are i.i.d. complex Gaussian variables with zero mean and an equal variance. Moreover, the channel has an

exponentially decaying profile, giving 0 dB, -5 dB and -10 dB powers for the first, second and third taps, respectively. In our simulation, M and K_{T} are set to be 64 and 8, respectively, suggesting that the pilots and SPF data are transmitted only in the subcarriers indexed by $64 \times k$, $(k = 0, 1, \cdots, 7)$, which, for the convenience, are referred to as the SPF subcarriers.

For the purpose of comparison, the channel vector $\mathbf{h}$ is estimated by the proposed SPF LP-based semi-blind algorithm, the LP-based semi-blind algorithm without using the SPF transmit scheme, the LS and the nulling-based semi-blind methods. For easy citation, we call these four methods as the SPF LP semi-blind, LP semi-blind, LS, and nulling semi-blind methods. In the SPF LP semi-blind and LP semi-blind algorithms, the ill condition is identified if $\rho_1 > 0.5$ or $\rho_2 > 5$ or $\rho_1 \times \rho_2 > 0.2$. The estimation performance is evaluated in terms of the MSE of the estimate of the channel matrix given by

$$\mathrm{MSE} = \frac{1}{N_{\mathrm{MC}}} \sum_{n=1}^{N_{\mathrm{MC}}} \left\| \hat{\mathbf{h}}_n - \mathbf{h}_n \right\|^2 \tag{7.59}$$

where N_{MC} is the number of Monte Carlo iterations, and $\mathbf{h}_n$ and $\hat{\mathbf{h}}_n$ are the true and the estimated channel vectors with respect to the n-th Monte Carlo iteration, respectively.

Experiment 1: *MSE versus SNR*

In the first experiment, the channel estimation performance in terms of the MSE versus the SNR is investigated. The simulation involves 2000 Monte Carlo runs of the transmission of 60 OFDM symbols with pilot length $g_{\mathrm{p}} = 20$. Here, the scaling factor in the SPF scheme is set to be $\eta = 9.5 \times 10^{-4}$, which corresponds, on the average, to

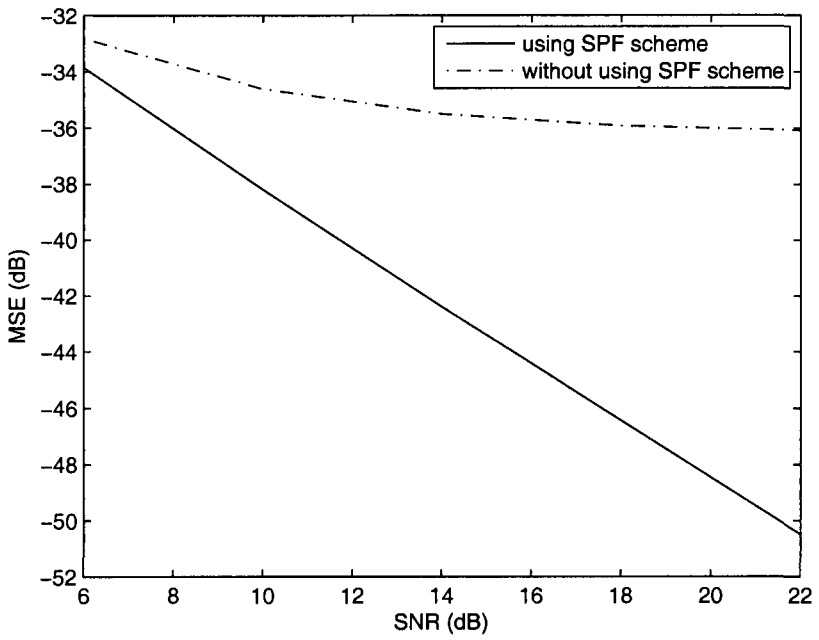

Figure 7.2: MSE of the estimated correlation matrix of the received signal versus SNR

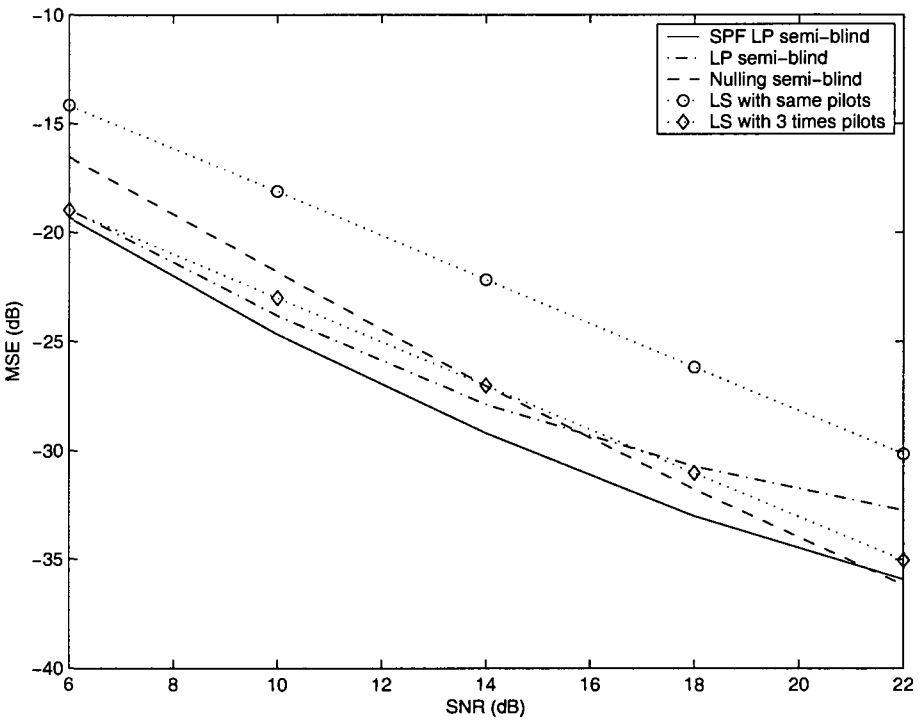

Figure 7.3: MSE of the time-domain channel estimate versus SNR

5.8 OFDM symbol per SPF subcarrier used for the transmission of $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$. First of all, Fig. 7.2 shows the MSE of the estimated correlation matrix of the received signal resulting from N_{MC} Monte Carlo iterations, which is defined in a manner similar to (7.59) by using the norm of the error correlation matrix. Clearly, the conventional correlation matrix estimation without using the proposed SPF cancellation scheme achieves very little gain in the MSE with increasing the SNR level. In contrast, by using the SPF cancellation scheme, the MSE of the estimated correlation matrix has been significantly improved, which is linearly propertional to the increase of the SNR. Fig. 7.3 shows the channel estimation results of the SPF LP semi-blind, LP semi-blind, nulling semi-blind methods as well as the LS method with 20 pilot symbols. Moreover, the result from the LS method using 60 pilot symbols which is three times the pilot length of other methods is also provided for comparison. It is seen that the SPF LP semi-blind algorithm consistently outperforms the nulling semi-blind method and the LS method. Also, one can find that the performance gain of the SPF LP semi-blind algorithm over the LP semi-blind algorithm becomes larger with increasing SNR value. In particular, the MSE is improved by 3.2 dB when the SNR is 22 dB.

Experiment 2: *The effect of scaling factor η on the channel estimation performance*

In this experiment, the channel estimation performance in terms of MSE for different scaling factors η is investigated. Using the same condition as in *Experiment 1*, the simulation is undertaken based on 2000 Monte Carlo runs for the scaling factor η given by $\eta_1 = 1.9 \times 10^{-3}$, $\eta_2 = 9.5 \times 10^{-4}$ and $\eta_3 = 4.7 \times 10^{-4}$, respectively. Table 7.1

scaling factor	$\eta_1 = 1.9 \times 10^{-3}$	$\eta_2 = 9.5 \times 10^{-4}$	$\eta_3 = 4.7 \times 10^{-4}$
average number of SPF symbols	4.1	5.8	9.7

Table 7.1: The average number of SPF symbols needed for the transmission of $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ vs value of η

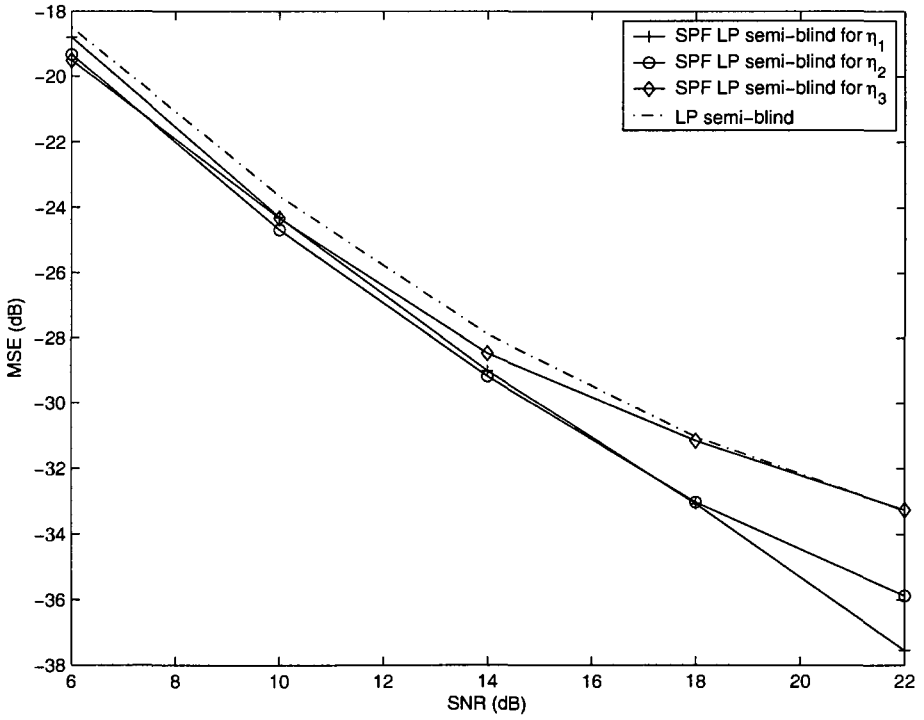

Figure 7.4: The effect of the scaling factor η on the MSE of channel estimate versus SNR

OFDM block size	1st singular value	2st singular value
40	2.8×10^{-3}	1.7×10^{-3}
60	2.2×10^{-3}	1.4×10^{-3}
80	1.9×10^{-3}	1.2×10^{-3}
100	1.7×10^{-3}	1.1×10^{-3}
120	1.6×10^{-3}	1.0×10^{-3}

Table 7.2: The two average singular values for five OFDM block sizes with 2 transmit antennas

shows the average number of OFDM symbols per SPF subcarrier used for $\mathbf{T}_{\mathrm{L}}(k)$ and $\mathbf{T}_{\mathrm{R}}(k)$ for three values of η. Clearly, more OFDM symbols are needed for a smaller value of η. Fig. 7.4 shows the MSE plots of the LP semi-blind algorithm and the SPF LP semi-blind algorithm for the three η values. It is seen that the performance of the SPF LP semi-blind algorithm for all three cases is better than that of the LP semi-blind algorithm. Moreover, the performance of the SPF LP semi-blind algorithm corresponding to each value of η depends on the SNR range. For example, η_2 gives the best MSE when the SNR is between 10 and 18 dB.

Experiment 3: *The effect of OFDM block size on the channel estimation performance*

Now, we examine the effect of the OFDM block size on channel estimation performance. In order to have a fair comparison, we first need to determine a proper value

OFDM block size	The normalized 1st singular value	The normalized 2nd singular value	The value of η
40	1.227	1.223	1.164×10^{-3}
60	1	1	9.5×10^{-4}
80	0.861	0.869	8.216×10^{-4}
100	0.765	0.774	7.313×10^{-4}
120	0.702	0.711	6.712×10^{-4}

Table 7.3: The normalized average singular values and the selected η values for five OFDM block sizes

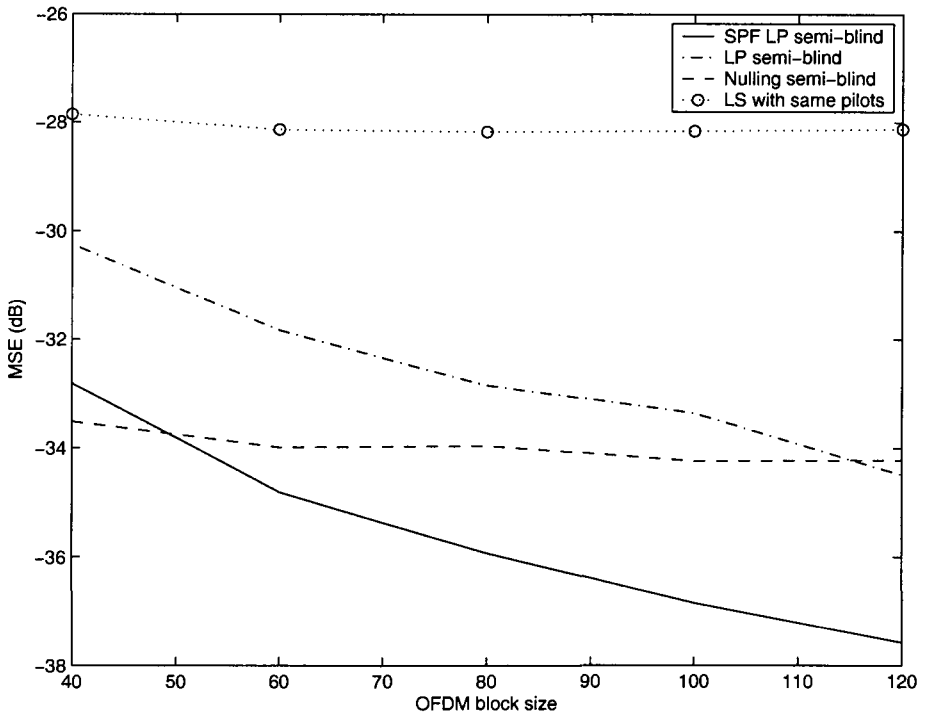

Figure 7.5: MSE versus the OFDM block size for a fixed number of pilots

of η for a different OFDM block size such that the same average number of OFDM symbols is used for the transmission of the SPF data. Table 7.2 shows the two averaged singular values $\sigma_{T,i}(k)$, $(i = 1, 2)$ which are obtained from 1000 Monte Carlo runs at an SNR of 15 dB for the OFDM block size g = 40, 60, 80, 100, 120. From Table 7.2, we now normalize the 1st average singular value of the five cases according to the 1st singular value corresponding to $g = 60$, as shown in the 2nd columns of Table 7.3. A similar operation is applied to the 2nd averaged singular values, giving the normalized singular values in the third column of Table 7.3. It is interesting to see that, in each of the five cases, the two normalized averaged singular values are very close. Here, we would like to set $\eta = 9.5 \times 10^{-4}$ for the case of $g = 60$. Considering that η is proportional to the singular values $\sigma_{T,i}(k)$ as seen from (7.31), one can calculate the η value for the other four cases by multiplying $\eta = 9.5 \times 10^{-4}$ with the average of the corresponding normalized 1st and 2nd singular values as given in the last column of Table 7.3. It has been verified through simulations that the average number of OFDM symbols for the SPF data is then 5.8 for all the five cases.

Fig. 7.5 shows the channel estimation results of the SPF LP semi-blind and the LP semi-blind algorithms as a function of the OFDM block size based on 5000 Monte Carlo runs for the pilot length $g_{\mathrm{p}} = 20$. For comparison, the performance of the LS method and the nulling semi-blind method with the same pilot length is also included. It is seen that the performance of both the SPF LP semi-blind and the LP semi-blind algorithms get much better with increasing OFDM block size. Although the LP semi-blind method can achieve a much better performance than the LS method, in general it does not perform as well as the nulling semi-blind method. However, the SPF LP

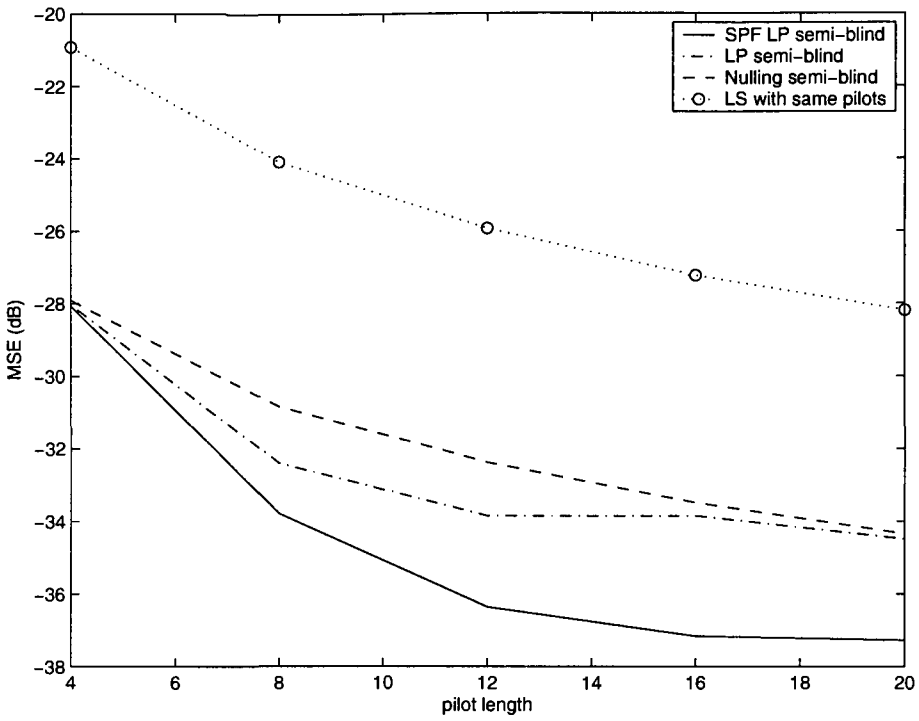

Figure 7.6: MSE versus pilot length for a block of 120 OFDM symbols

semi-blind algorithm outperforms the nulling semi-blind method when the OFDM block size is larger than 50, and becomes more advantageous as the OFDM block size increases. In particular, when $g = 120$, the SPF LP semi-blind algorithm gains about 3.3 dB over the nulling semi-blind method.

Experiment 4: *MSE versus pilot length*

Here, we investigate the channel estimation performance of the proposed algorithm versus the pilot length. Fig. 7.6 shows the MSE plots from 500 Monte Carlo iterations for $g = 120$ at an SNR of 15 dB. It is seen that the performance of all the algorithms is improved with increasing pilot length except for the LP semi-blind method. Again, the SPF LP semi-blind algorithm outperforms the other three methods for all pilot

lengths. In particular, when the pilot length is 20, the gain of the SPF LP semi-blind algorithm over the LS, the LP semi-blind and the nulling semi-blind algorithms is 9 dB, 2.8 dB and 2.9 dB, respectively.

7.5 Conclusions

We had previously shown in Chapter 3 that the LP-based blind channel estimation method is subject to signal perturbation error and therefore gives a poor estimation performance in the moderate to high SNR levels. To improve its performance, in this chapter, we have developed a new scheme that transmits user specific data bearing the information of the correlation matrix of the information signal to cancel the signal perturbation error at the receiver. The new transmit structure has then been used to develop a semi-blind LP-based approach for the estimation of time-domain channels. In the new approach, the whitening rotation algorithm is used along with linear prediction to obtain a blind channel estimate subject to an ambiguity matrix, and the ambiguity matrix is then estimated via a pilot-assisted maximum likelihood method. It has been shown that the new semi-blind MIMO-OFDM channel estimation solution is devoid of any signal perturbation error in the noise-free case and is capable of efficiently suppressing the signal perturbation error in the noisy case. Simulation results have confirmed that, by using a small number of additional slots for the transmission of the user specific SPF data, the new semi-blind approach can achieve a significantly improved channel estimation performance as compared to the LP-based semi-blind method without using the signal-perturbation-free transmit scheme, the

LS method as well as the nulling-based semi-blind technique.

Chapter 8

Conclusion and Future Work

8.1 Concluding Remarks

In this dissertation, the semi-blind channel estimation issue of MIMO-OFDM systems has been thoroughly studied. Based on the MIMO linear prediction technique, a nulling-based semi-blind channel estimation approach has been first proposed for the estimation of the time-domain MIMO-OFDM channels. It has been then extended to the channel estimation of pulse-shaped MIMO-OFDM systems and further applied to the estimation of the sparse MIMO-OFDM channels. By conducting a perturbation analysis, the proposed semi-blind channel estimation approaches have been shown to significantly outperform some other blind and semi-blind methods recently reported in open literature, in the case of moderate to high SNR levels. In order to improve the performance of those methods for high SNRs, two novel signal-perturbation-free transmit schemes have also been proposed for the estimation of frequency-domain channels and that of time-domain channels, leading to two very efficient signal-perturbation-free channel estimation techniques.

The first part of the thesis has been devoted to the development of nulling-based semi-blind estimation approaches for the time-domain MIMO-OFDM channels. By using a training-based least-square method in conjunction with a blind constraint on the channel vector, that is derived from MIMO linear prediction, a nulling-based semi-blind MIMO-OFDM channel estimation approach has been developed. As a part of the new approach, a practical yet very efficient scheme has been proposed for the determination of the weighting factor in the semi-blind cost function. The perturbation analysis of the MIMO linear prediction has justified the advantage of the semi-blind method over the pure blind estimation, and led to a closed-form expression for the MSE of the blind constraint. This study can be regarded as a significant extension of the existing linear prediction-based MIMO channel estimation.

The proposed nulling-based semi-blind approach has then been extended to the channel estimation of pulse-shaped MIMO-OFDM systems. By exploiting the pulse-shaping filter available in the transmitter and the matched filter in the receiver, a very efficient semi-blind estimation algorithm and an enhanced LS algorithm have been developed for sampling- and upsampling-duration-based channels, respectively. To reduce the computational complexity of the semi-blind method, a frequency-domain estimation algorithm for the time-domain correlation matrix has also been presented. The significance of this work lies in the study of, for the first time, the pure wireless channel estimation of MIMO-OFDM systems with pulse shaping, a practical yet very complicated issue, which results in a significantly improved channel estimation performance.

The proposed nulling-based semi-blind estimation idea has also been applied to

the estimation of sparse MIMO-OFDM channels. It is revealed for the first time that the most significant taps (MST) of the sparse channel can be expressed in terms of the most significant lags (MSL) of the correlation matrices of the received signal. This finding has helped develop a highly efficient MST detection algorithm that requires only a few OFDM symbols and a small number of pilots for a least-square based detection. By employing the detected MST information, a semi-blind approach incorporating a training-based LS criterion and a blind constraint on the sparse channel vector consisting of only a few MSTs has been proposed. As the new approach does not require estimating all the channel taps, it has saved a large amount of computations compared to a regular channel estimation method. It has also been shown via a perturbation analysis that the proposed semi-blind solution is not subject to the signal perturbation error when the sparse channel is a decimated version of a full wireless fading channel. Computer simulations based on various sparse channels have confirmed that the proposed sparse semi-blind approach outperforms much better than the sparse LS method as well as the regular LS and semi-blind techniques.

The second part of the dissertation has been focused on the development of completely new transmit schemes for the cancellation of signal perturbation errors in the the estimation of frequency-domain as well as time-domain MIMO-OFDM channels. The novelty of this idea lies in the transmission of limited user-dependent data bearing partial information of the correlation matrix of the transmitted signal that can be exploited to enhance the estimation accuracy of the second-order statistics of the received signal at the receiver, leading to a signal-perturbation-free channel estimation algorithm.

Firstly, a new signal-perturbation-free semi-blind approach has been proposed for the estimation of the frequency-domain channel of MIMO-OFDM systems. A perturbation analysis of two subspace-based semi-blind frequency-flat MIMO channel estimation algorithms has been performed, justifying that the existing WR-based method is efficient only in the low SNR case, and the nulling-based method is a better choice when the SNR is moderate or high. To improve the performance of the WR-based method in the high SNR case, a new transmit structure, which contains known data bearing the information of the signal perturbation matrix, has been proposed for the cancellation of the signal perturbation error at the receiver. Furthermore, in the computation of second order statistics of the received signal via an in-depth perturbation analysis, a closed-form expression for the MSE of the new WR-based method integrating the proposed transmit scheme is derived. Simulation results have confirmed that, by using a small number of additional slots bearing the information of the autocorrelation matrix of the transmitted signal, a significant improvement in terms of both the MSE of the channel estimate and the BER of the data transmission can be achieved over the WR-based method as well as the nulling-based method for all SNR cases.

Secondly, a signal-perturbation-free linear prediction-based semi-blind approach has been proposed for the estimation of time-domain channels. We have developed a new scheme that transmits user specific data bearing the information of the correlation matrix of the information signal to the receiver for the cancellation of signal perturbation error at the receiver. The new transmit structure has then been used to develop a semi-blind LP-based channel estimation approach. In the new approach, a

whitening rotation-based algorithm is used along with linear prediction to obtain a blind channel estimate subject to an ambiguity matrix, and the ambiguity matrix is then estimated via a pilot-assisted maximum likelihood method. It has been shown that the new semi-blind MIMO-OFDM channel estimation solution is devoid of any signal perturbation error in the noise-free case and is capable of efficiently suppressing the signal perturbation error in the noisy case. Simulation results have confirmed that, by using a small number of additional slots for the transmission of the user specific data, the new semi-blind approach can achieve a significantly improved channel estimation performance as compared to the LP-based semi-blind method without using the signal-perturbation-free transmit scheme as well as the LS method.

8.2 Suggestions for Future Investigation

Considering that a number of original ideas have been proposed in this thesis and the related channel estimation algorithms have been developed only over a four-year span, several aspects of the proposed techniques are worth well further studying. The following lists a number of directions for future research.

- As shown in Chap. 4, the time of arrivals (TOA) are needed for the channel estimation of pulse-shaped MIMO-OFDM systems. Normally, they can be estimated by using some training-based techniques. For example, the TOA estimation was conducted by using the ranging techniques for the uplink synchronization of the OFDMA (OFDM Access) systems [8] or for some geolocation applications [92]. The TOA can be also estimated by using beamforming meth-

ods such as the MUSIC and ESPRIT [52, 54, 56–58]. In general, the TOA are considered sparse in the upsampling domain. Therefore, our semi-blind MST detection algorithm proposed for sampling-spaced sparse channels can be extended to the estimation of TOA. Accordingly, the proposed sparse semi-blind algorithm for the sampling-spaced sparse channel estimation can be extended for the estimation of upsampling duration-based channels.

- The proposed frequency-domain estimation algorithm for the correlation matrix could be generalized to develop other second-order statistics-based blind or semi-blind algorithms of MIMO systems for MIMO-OFDM systems.

- The signal-perturbation-free transmit scheme proposed for the MIMO-OFDM channel estimation can be extended to improve other signal processing algorithms for multiple antenna systems. Whenever such algorithms are based on second-order statistics, they are subject to signal perturbation error. Then, a similar signal-perturbation-free transmit scheme could be developed to cancel the signal perturbation at the receiver.

- The proposed semi-blind channel estimation approaches could be extended for the estimation of time-varying MIMO-OFDM channels in urban areas. In an urban scenario, the channel can be regarded as a combination of a large number of multipaths, which can be modelled as Rayleigh or Rician process. In this case, the maximum Doppler frequency can be estimated using some of the existing methods such as those in [120, 121], which can then be utilized to determine the AR model parameters for the Kalman filter-based channel tracking [48, 122].

- The proposed semi-blind approaches could also be extended for the estimation of time-varying MIMO-OFDM channels in suburban areas. In the suburban environment, the MIMO channel can be modelled as a combination of a few multi-paths, each of which may be characterized by a DOA, DOD, TOA, doppler frequency and complex fading amplitude. The proposed semi-blind channel estimation approaches can be extended for the joint estimation of the multi-path parameters including the Doppler frequency, by combining our proposed virtual array processing techniques [123,124] with some of the existing training-based or blind parameter estimation algorithms such as those in [125–127].

References

[1] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj, "A fourth-generation MIMO-OFDM broadband wireless system: design, performance, and field trial results", *IEEE Communications Magazine*, vol. 40, no. 9, pp. 143–149, 2002.

[2] H. Yang, "A road to future broadband wireless access: MIMO-OFDM-based air interface", *IEEE Communications Magazine*, vol. 43, no. 1, pp. 53–60, 2005.

[3] G. Stuber, J. R. Barry, S. W. Mclaughlin, Y. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications", *Proceedings of the IEEE*, vol. 92, no. 2, pp. 271–294, 2004.

[4] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An overview of MIMO communications-a key to gigabit wireless", *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198–218, 2004.

[5] Y. Jung, J. Kim, S. Lee, H. Yoon, and Kim. J., "Design and implementation of MIMO-OFDM baseband processor for high-speed wireless lans", *IEEE Trans. on Circuits and Systems II: Express Briefs*, vol. 54, no. 7, pp. 631–635, July 2007.

[6] T. Keller and L. Hanzo, "Adaptive multicarrier modulation: a convenient framework for time-frequency processing in wireless communications", *Proceedings of the IEEE*, vol. 88, no. 5, pp. 611–640, 2000.

[7] M. Steer, "Beyond 3G", *IEEE Microwave Magazine*, vol. 8, no. 1, pp. 76–82, 2007.

[8] J. G. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX-understanding Broadband Wireless Networking*, Prentice Hall, Englewood Cliffs, NJ, 2007.

[9] M.K. Ozdemir and H. Arslan, "Toward real-time adaptive low-rank LMMSE channel estimation of MIMO-OFDM systems", *IEEE Trans. on Wireless Communications*, vol. 5, no. 10, pp. 2675–2678, Oct. 2006.

[10] S. J. Lee, "On the training of MIMO-OFDM channels with least square channel estimation and linear interpolation", *IEEE Communications Letters*, vol. 12, no. 2, pp. 100–102, February 2008.

[11] D. Samardzija and N. Mandayam, "Pilot-assisted estimation of MIMO fading channel response and achievable data rates", *IEEE Trans. on Signal Processing*, vol. 51, no. 11, pp. 2882–2890, Nov 2003.

[12] M. Biguesh and A. B. Gershman, "MIMO channel estimation: optimal training and tradeoffs between estimation techniques", in *Proc. IEEE International Conference on Communications*, 2004, vol. 5, pp. 2658–2662.

[13] V. Barroso and J. Xavier, "Blind identification of MIMO channels: a closed form solution based on second order statistics", in *Thirty-Third Asilomar Conference on Signals, Systems, and Computers*, 1999, vol. 1, pp. 70–74.

[14] J. K. Tugnait, "Blind estimation and equalization of MIMO channels via multidelay whitening", *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 8, pp. 1507–1519, 2001.

[15] Z. Ding and L. Qiu, "Blind MIMO channel identification from second order statistics using rank deficient channel convolution matrix", *IEEE Trans. on Signal Processing*, vol. 51, no. 2, pp. 535–544, 2003.

[16] Z. Ding and D. B. Ward, "Subspace approach to blind and semi-blind channel estimation for space-time block codes", *IEEE Trans. on Wireless Communications*, vol. 4, no. 2, pp. 357–362, 2005.

[17] A. Medles, D. T. M. Slock, and E. De Carvalho, "Linear prediction based semi-blind estimation of MIMO FIR channels", in *Proc. IEEE Third Workshop on Signal Processing Advances in Wireless Communications*, 2001, pp. 58–61.

[18] A. Medles and D. T. M. Slock, "Augmenting the training sequence part in semiblind estimation for MIMO channels", in *Thirty-Seventh Asilomar Conference on Signals, Systems and Computers*, 2003, vol. 2, pp. 1825–1829.

[19] Y. Zeng and T. Ng, "A semi-blind channel estimation method for multiuser multiantenna OFDM systems", *IEEE Trans. on Signal Processing*, vol. 52, no. 5, pp. 1419–1429, 2004.

[20] A. K. Jagannatham and B. D. Rao, "Constrained ML algorithms for semi-blind MIMO channel estimation", in *IEEE Global Telecommunications Conference*, 2004, vol. 4, pp. 2475–2479.

[21] A. K. Jagannatham and B. D. Rao, "Cramer-Rao lower bound for constrained complex parameters", *IEEE Signal Processing Letters*, vol. 11, no. 11, pp. 875–878, 2004.

[22] Y. H. Zeng, S. D. Ma, and T. S. Ng, "Semi-blind estimation of channels and symbols for asynchronous MIMO systems", *IEE Proceedings-Communications*, vol. 152, no. 6, pp. 883–889, 2005.

[23] A. K. Jagannatham and B. D. Rao, "Whitening-rotation-based semi-blind MIMO channel estimation", *IEEE Trans. on Signal Processing*, vol. 54, no. 3, pp. 861–869, 2006.

[24] C.R. Murthy, A. K. Jagannatham, and B. D. Rao, "Training-based and semiblind channel estimation for MIMO systems with maximum ratio transmission", *IEEE Trans. on Signal Processing*, vol. 54, no. 7, pp. 2546–2558, 2006.

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

[26] X. Ma, L. Yang, and G. B. Giannakis, "Optimal training for MIMO frequency-selective fading channels", *IEEE Trans. on Wireless Communications*, vol. 4, no. 2, pp. 453–466, 2005.

[27] J. Zhang, Z. He, and Y. Huang, "Tsk fuzzy approach to channel estimation for MIMO-OFDM systems", *IEEE Signal Processing Letters*, vol. 14, no. 6, pp. 381–384, 2007.

[28] R. Chen, H. Zhang, Y. Xu, and H. Luo, "On MM-type channel estimation for MIMO OFDM systems", *IEEE Trans. on Wireless Communications*, vol. 6, no. 3, pp. 1046–1055, 2007.

[29] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation", *IEEE Trans. on Wireless Communications*, vol. 5, no. 5, pp. 1158–1168, May 2006.

[30] H. Minn, N. Al-Dhahir, and Y. Li, "Optimal training signals for MIMO OFDM channel estimation in the presence of frequency offset and phase noise", *IEEE Trans. on Communications*, vol. 54, no. 10, pp. 1754–1759, Oct. 2006.

[31] C. Shin, R. W. Heath, and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems", *IEEE Trans. on Vehicular Technology*, vol. 56, no. 2, pp. 670–685, 2007.

[32] F. Gao and A. Nallanathan, "Blind channel estimation for MIMO OFDM systems via nonredundant linear precoding", *IEEE Trans. on Signal Processing*, vol. 55, no. 2, pp. 784–789, 2007.

[33] C. Shin, R. W. Heath, and E. J. Powers, "Non-redundant precoding-based blind and semi-blind channel estimation for MIMO block transmission with a cyclic prefix", *IEEE Trans. on Signal Processing*, vol. 56, no. 6, pp. 2509–2523, June 2008.

[34] F. Gao, Y. Zeng, A. Nallanathan, and T. Ng, "Robust subspace blind channel estimation for cyclic prefixed MIMO ODFM systems: algorithm, identifiability and performance analysis", *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 2, pp. 378–388, February 2008.

[35] F. Gao and A. Nallanathan, "Resolving multidimensional ambiguity in blind channel estimation of MIMO-FIR systems via block precoding", *IEEE Trans. on Vehicular Technology*, vol. 57, no. 1, pp. 11–21, Jan. 2008.

[36] Y.-S. Chen and C.-A. Lin, "Blind-channel identification for MIMO single-carrier zero-padding block-transmission systems", *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 55, no. 6, pp. 1571–1579, July 2008.

[37] Y. Zeng, W. H. Lam, and T. S. Ng, "Semiblind channel estimation and equalization for MIMO space-time coded OFDM", *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 53, no. 2, pp. 463–474, 2006.

[38] Y. Zeng, A.R. Leyman, and T. Ng, "Joint semiblind frequency offset and channel estimation for multiuser MIMO-OFDM uplink", *IEEE Trans. on Communications*, vol. 55, no. 12, pp. 2270–2278, Dec. 2007.

[39] A. Gorokhov and P. Loubaton, "Blind identification of MIMO-FIR systems: A generalized linear prediction approach", *Signal Processing*, vol. 73 (1-2), pp. 105–124, 1999.

[40] Y. Inouye and R. Liu, "A system-theoretic foundation for blind equalization of an FIR MIMO channel system", *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 49, no. 4, pp. 425–436, 2002.

[41] R. Lopez-Valcarce and S. Dasgupta, "Blind channel equalization with colored sources based on second-order statistics: a linear prediction approach", *IEEE Trans. on Signal Processing*, vol. 49, no. 9, pp. 2050–2059, 2001.

[42] A. Medles and D. T. M. Slock, "Semiblind channel estimation for MIMO spatial multiplexing systems", in *Proc. IEEE Vehicular Technology Conference*, 2001, vol. 2, pp. 1240–1244.

[43] A. Medles and D. T. M. Slock, "Linear precoding for spatial multiplexing MIMO systems: blind channel estimation aspects", in *Proc. IEEE International Conference on Communications*, 2002, vol. 1, pp. 401–405.

[44] C. B. Papadias and D. T. M. Slock, "Fractionally spaced equalization of linear polyphase channels and related blind techniques based on multichannel linear prediction", *IEEE Trans. on Signal Processing*, vol. 47, no. 3, pp. 641–654, 1999.

[45] J. K. Tugnait, "On linear predictors for MIMO channels and related blind identification and equalization", *IEEE Signal Processing Letters*, vol. 5, no. 11, pp. 289–291, 1998.

[46] J. K. Tugnait and B. Huang, "Multistep linear predictors-based blind identification and equalization of multiple-input multiple-output channels", *IEEE Trans. on Signal Processing*, vol. 48, no. 1, pp. 26–38, 2000.

[47] S. Haykin, *Modern Wireless Communications*, Prentice-Hall, 2005.

[48] K. J. Kim, T. Reid, and R. A. Iltis, "A sequential Monte-Carlo Kalman filter based delay and channel estimation method in the MIMO-OFDM system", in *Proc. IEEE Vehicular Technology Conference*, 2004, vol. 1, pp. 573–577.

[49] K. J. Kim and T. Reid, "Multiple hypothesis channel estimation for the MIMO-OFDM system", in *Proc. IEEE Global Telecommunications Conference*, 2004, vol. 4, pp. 2674–2678.

[50] T. Roman, M. Enescu, and V. Koivunen, "Time-domain method for tracking dispersive channels in MIMO OFDM systems", in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2003, vol. 4, pp. 393–396.

[51] D. Schafhuber, G. Matz, and F. Hlawatsch, "Kalman tracking of time-varying channels in wireless MIMO-OFDM systems", in *Proc. the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers*, 2003, vol. 2, pp. 1261–1265.

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

[53] M. Cicerone, O. Simeone, N. Geng, and U. Spagnolini, "Modal analysis/filtering to estimate time-varying MIMO-OFDM channels", in *ITG Workshop on Smart Antennas*, 2004, pp. 35–40.

[54] H. Miao and M. J. Juntti, "Space-time channel estimation and performance analysis for wireless MIMO-OFDM systems with spatial correlation", *IEEE Trans. on Vehicular Technology*, vol. 54, no. 6, pp. 2003–2016, 2005.

[55] M. Hua and J. Zhu, "A blind uplink OFDM synchronization algorithm based on cyclostationarity", in *Proc. IEEE Vehicular Technology Conference*, 2005, vol. 2, pp. 1002–1006.

[56] M. Oziewicz, "On application of MUSIC algorithm to time delay estimation in OFDM channels", *IEEE Trans. on Broadcasting*, vol. 51, no. 2, pp. 249–255, 2005.

[57] T. A. Thomas and F. W. Vook, "Broadband MIMO-OFDM channel estimation via near maximum likelihood time of arrival estimation", in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2002, vol. 3, pp. 2569–2572.

[58] Z. J. Wang, H. Zhu, and K. J. R. Liu, "A MIMO-OFDM channel estimation approach using time of arrivals", *IEEE Trans. on Wireless Communications*, vol. 4, no. 3, pp. 1207–1213, 2005.

[59] O. Simeone and U. Spagnolini, "Channel estimation for block-fading frequency-selective rayleigh MIMO channels: performance limits", in *Proc. IEEE 59th Vehicular Technology Conference*, 2004, vol. 2, pp. 856–860.

[60] Z. Ding, "Multipath channel identification based on partial system information", *IEEE Trans. on Signal Processing*, vol. 45, no. 1, pp. 235–240, 1997.

[61] N. J. Baas and D. P. Taylor, "Pulse shaping for wireless communication over time- or frequency-selective channels", *IEEE Trans. on Communications*, vol. 52, no. 9, pp. 1477–1479, 2004.

[62] H. Bolcskei, P. Duhamel, and R. Hleiss, "A subspace-based approach to blind channel identification in pulse shaping OFDM/OQAM systems", *IEEE Trans. on Signal Processing*, vol. 49, no. 7, pp. 1594–1598, 2001.

[63] M. J. Borran, P. Varshney, H. Vilpponen, and P. D. Papadimitriou, "Channel estimation and signal detection for multi-carrier CDMA systems with pulse-shaping filter", in *Proc. IEEE International Conference on Communications*, 2003, vol. 5, pp. 3457–3461.

[64] M. Y. A. Gaffar, A. D. Broadhurst, and F. Takawira, "Improved subspace-based channel estimation algorithm for DS-CDMA systems exploiting pulse-shaping information", *IEE Proceedings-Communications*, vol. 152, no. 5, pp. 533–540, 2005.

[65] H. Minn and V. K. Bhargava, "An investigation into time-domain approach for OFDM channel estimation", *IEEE Trans. on Broadcasting*, vol. 46, no. 4, pp. 240–248, 2000.

[66] M.R. Raghavendra and K. Giridhar, "Improving channel estimation in OFDM systems for sparse multipath channels", *IEEE Signal Processing Letters*, vol. 12, no. 1, pp. 52–55, 2005.

[67] J. K. Hwang, R. L. Chung, M. F. Tsai, and J. H. Deng, "Highly efficient sparse multipath cfhannel estimator with chu-sequence preamble for frequency-domain MIMO

DFE receiver", *IEICE Trans. on Communications*, vol. E90B, no. 8, pp. 2103–2110, 2007.

[68] C. R. N. Athaudage and A. D. S. Jayalath, "Delay-spread estimation using cyclic-prefix in wireless OFDM systems", *IEE Proceedings-Communications*, vol. 151, no. 6, pp. 559–566, 2004.

[69] S. Kim and R. A. Iltis, "A matching-pursuit/GSIC-based algorithm for DS-CDMA sparse-channel estimation", *IEEE Signal Processing Letters*, vol. 11, no. 1, pp. 12–15, 2004.

[70] D. K. Borah, "Estimation of frequency-selective CDMA channels with large possible delay and Doppler spreads", *IEEE Trans. on Vehicular Technology*, vol. 55, no. 4, pp. 1126–1136, 2006.

[71] B. Friedlander and A. J. Weiss, "On the second-order statistics of the eigenvectors of sample covariance matrices", *IEEE Trans. on Signal Processing*, vol. 46, no. 11, pp. 3136–3139, 1998.

[72] Z. Xu, "Perturbation analysis for subspace decomposition with applications in subspace-based algorithms", *IEEE Trans. on Signal Processing*, vol. 50, no. 11, pp. 2820–2830, 2002.

[73] Z. Xu, "On the second-order statistics of the weighted sample covariance matrix", *IEEE Trans. on Signal Processing*, vol. 51, no. 2, pp. 527–534, 2003.

[74] Z. Xu, "Statistical performance of a data-based covariance estimator", *IEEE Trans. on Vehicular Technology*, vol. 53, no. 3, pp. 939–943, 2004.

[75] V. Buchoux, O. Cappe, E. Moulines, and A. Gorokhov, "On the performance of semi-blind subspace-based channel estimation", *IEEE Trans. on Signal Processing*, vol. 48, no. 6, pp. 1750–1759, 2000.

[76] F. Li and R. J. Vaccaro, "Analysis of min-norm and MUSIC with arbitrary array geometry", *IEEE Trans. on Aerospace and Electronic Systems*, vol. 26, no. 6, pp. 976–985, 1990.

[77] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: unification, simplification, and observations", *IEEE Trans. on Aerospace and Electronic Systems*, vol. 29, no. 4, pp. 1170–1184, 1993.

[78] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, 3 edition, 1995.

[79] T.W.S. Chow, B. Wang, and K.T. Ng, "Linear prediction based multipath channel identification algorithm", *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 50, no. 6, pp. 769–774, 2003.

[80] S. Houcke and P. Loubaton, "A weighted linear prediction approach for the blind CDMA forward link channel estimation", in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2000, vol. 5, pp. 2937–2940.

[81] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "Linear prediction based semi-blind channel estimation for MIMO-OFDM system", in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, 2007, pp. 3239–3242.

[82] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A closed-form semi-blind solution to MIMO-OFDM channel estimation", in *Proc. IEEE International Conference on Neural Networks and Signal Processing*, 2008, pp. 125–128. (Best Paper Award).

[83] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A semi-blind channel estimation approach for MIMO-OFDM systems", *IEEE Trans. on Signal Processing*, vol. 56, no. 7, pp. 2821–2834, 2008.

[84] J. Choi, "Equalization and semi-blind channel estimation for space-time block coded signals over a frequency-selective fading channel", *IEEE Trans. on Signal Processing*, vol. 52, no. 3, pp. 774–785, 2004.

[85] C. Moreno, "A least-squares-based method for determining the ratio between two measured quantities", *Measurement Science and Technology*, vol. 7, pp. 137–141, 1996.

[86] J. Valyon and G. Horvath, "Extended least squares LS-SVM", *International Journal of Computational Intelligence*, vol. 3, no. 3, pp. 1304–2386, 2006.

[87] P. Tsai, H. Kang, and T. Chiueh, "Joint weighted least-squares estimation of carrier-frequency offset and timing offset for OFDM systems over multipath fading channels", *IEEE Trans. on Vehicular Technology*, vol. 54, no. 1, pp. 211–223, 2005.

[88] M. Sandell, D. McNamara, and S. Parker, "Analysis of frequency-offset tracking in MIMO OFDM systems", *IEEE Trans. on Communications*, vol. 54, no. 8, pp. 1481–1489, 2006.

[89] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission", *IEEE Wireless Communications*, vol. 12, no. 2, pp. 56–65, 2005.

[90] W.-P. Zhu, Y. Yan, M. O. Ahmad, and M. N. S. Swamy, "A feedforward symbol timing recovery technique using two samples per symbol", *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 52, no. 11, pp. 2490–2500, 2005.

[91] W. P. Zhu, M. O. Ahmad, and M. N. S. Swamy, "ASIC implementation architecture for pulse shaping FIR filters in 3G mobile communications", in *Proc. IEEE International Symposium on Circuits and Systems*, 2002, vol. 1, pp. 433–436.

[92] F. Gustafsson and F. Gunnarsson, "Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements", *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 41–53, 2005.

[93] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "Frequency-domain channel estimation of MIMO-OFDM systems with pulse-shaping", *submitted to IEEE Trans. Wireless Communications.*

[94] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "Semi-blind channel estimation of MIMO-OFDM systems with pulse shaping", in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, 2008, pp. 125–128.

[95] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A frequency-domain correlation matrix estimation algorithm for MIMO-OFDM channel estimation", in *Proc. IEEE Vehicular Technology Conference (VTC 2008-Fall)*, 2008.

[96] E. Aktas and U. Mitra, "Single-user sparse channel acquisition in multiuser DS-CDMA systems", *IEEE Trans. on Communications*, vol. 51, no. 4, pp. 682–693, 2003.

[97] M. R. Raghavendra, E. Lior, S. Bhashyam, and K. Giridhar, "Parametric channel estimation for pseudo-random tile-allocation in uplink OFDMA", *IEEE Trans. on Signal Processing*, vol. 55, no. 11, pp. 5370–5381, 2007.

[98] C. Carbonelli, S. Vedantam, and U. Mitra, "Sparse channel estimation with zero tap detection", *IEEE Trans. on Wireless Communications*, vol. 6, no. 5, pp. 1743–1763, 2007.

[99] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization", *IEEE Trans. on Communications*, vol. 50, no. 3, pp. 374–377, 2002.

[100] Y. Liu and D. K. Borah, "Estimation of fading channels with large possible delay spreads", *Electronics Letters*, vol. 39, no. 1, pp. 130–131, 2003.

[101] J. Homer, I. Mareels, R. R. Bitmead, B. Wahlberg, and A. Gustafsson, "LMS estimation via structural detection", *IEEE Trans. on Signal Processing*, vol. 46, no. 10, pp. 2651–2663, 1998.

[102] J. Homer, I. Mareels, and C. Hoang, "Enhanced detection-guided NLMS estimation of sparse FIR-modeled signal channels", *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 53, no. 8, pp. 1783–1791, 2006.

[103] C. Wu and D. W. Lin, "Sparse channel estimation for OFDM transmission based on representative subspace fitting", in *Proc. IEEE Vehicular Technology Conference*, 2005, vol. 1, pp. 495–499.

[104] C. Wu and D. W. Lin, "A group matching pursuit algorithm for sparse channel estimation for OFDM transmission", in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, 2006, vol. 4, pp. 429–432.

[105] C.-A. Lin and Y.-S. Chen, "Blind identification of MIMO channels using optimal periodic precoding", *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 54, no. 4, pp. 901–911, April 2007.

[106] K. Kohno, Y. Inouye, and M. Kawamoto, "A matrix pseudo-inversion lemma for positive semidefinite Hermitian matrices and its application to adaptive blind deconvolution of MIMO systems", *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 55, no. 1, pp. 424–435, Feb. 2008.

[107] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A semi-blind algorithm for most significant tap detection in channel estimation of OFDM systems", *accepted for its presentation at IEEE International Symposium on Circuits and Systems (ISCAS) to be held in May 2009, Taipei.*

[108] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A semi-blind sparse channel estimation algorithm for MIMO-OFDM systems", *submitted to IEEE Trans. Signal Processing.*

[109] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "Semi-blind most significant tap detection for sparse channel estimation of OFDM systems", *submitted to IEEE Trans. on Circuits and Systems I: Regular Papers (under 2nd round of review).*

[110] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "Perturbation analysis of sparse MIMO-OFDM channel estimation algorithm", *submitted to IEEE International Midwest Symposium on Circuits and Systems to be held in August 2009, Cancun, Mexico.*

[111] F. Li and R. J. Vaccaro, "Analytical performance prediction of subspace-based algorithms for DOA estimation", in *SVD and Signal Processing, II: Algorithms, Analysis and Applications*, R. J. Vaccaro, Ed., pp. 243–260. Elsevier, 1991.

[112] G. W. Stewart, "Perturbation theory for the singular value decomposition", in *SVD and Signal Processing, II: Algorithms, Analysis and Applications*, R. J. Vaccaro, Ed., pp. 99–109. Elsevier, 1991.

[113] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "Perturbation analysis of subspace-based semi-blind MIMO channel estimation approaches", in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, 2008, pp. 129–132.

[114] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A signal perturbation free transmit scheme for MIMO channel estimation", in *Proc. IEEE Vehicular Technology Conference (VTC 2008-Fall)*, 2008.

[115] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A signal perturbation free whitening-rotation-based semi-blind approach for MIMO channel estimation", *accepted with minor revision for publication as Regular Paper in IEEE Trans. on Signal Processing.*

[116] K.-B. Yu, "Recursive updating the eigenvalue decomposition of a covariance matrix", *IEEE Trans. on Signal Processing*, vol. 39, no. 5, pp. 1136–1145, May 1991.

[117] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "An enhanced scheme for second-order-statistics estimation in MIMO-OFDM systems", *accepted for its presentation at IEEE International Symposium on Circuits and Systems (ISCAS) to be held in May 2009, Taipei.*

[118] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A signal-perturbation-free linear prediction-based semi-blind algorithm for MIMO-OFDM channel estimation", *submitted to IEEE Trans. on Signal Processing.*

[119] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "Ill condition identification for linear prediction-based MIMO-OFDM channel estimation", *submitted to IEEE International Midwest Symposium on Circuits and Systems to be held in August 2009, Cancun, Mexico.*

[120] K. E. Baddour and N. C. Beaulieu, "Robust Doppler spread estimation in nonisotropic fading channels", *IEEE Trans. on Wireless Communications*, vol. 4, no. 6, pp. 2677–2682, 2005.

[121] P. Y. Chen and H. J. Li, "An iterative algorithm for Doppler spread estimation in LOS environments", *IEEE Trans. on Wireless Communications*, vol. 5, no. 6, pp. 1223–1228, 2006.

[122] M.K. Tsatsanis and Z. Xu, "Pilot symbol assisted modulation in frequency selective fading wireless channels", *IEEE Trans. on Signal Processing*, vol. 48, no. 8, pp. 2353–2365, 2000.

[123] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A spatial extrapolation based super-resolution DOA estimation algorithm", in *Proc. IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2007, pp. 1157–1160.

[124] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A spatial extrapolation based blind DOA estimation approach for closely spaced sources", *to appear as Regular Paper in IEEE Trans. on Aerospace and Electronic Systems.*

[125] A. J. van der Veen, M. C. Vanderveen, and A. Paulraj, "Joint angle and delay estimation using shift-invariance techniques", *IEEE Trans. on Signal Processing*, vol. 46, no. 2, pp. 405–418, 1998.

[126] Y. Wang, J. Chen, and W. Fang, "TST-MUSIC for joint DOA-delay estimation", *IEEE Trans. on Signal Processing*, vol. 49, no. 4, pp. 721–729, 2001.

[127] A. N. Lemma, A. J. van der Veen, and E. F. Deprettere, "Analysis of joint angle-frequency estimation using ESPRIT", *IEEE Trans. on Signal Processing*, vol. 51, no. 5, pp. 1264–1283, 2003.

Appendix A

Derivation of $\mathrm{MSE_B}$ in Nulling-based Semi-Blind Estimation

We formulate first $\mathrm{vec}(\mathbf{G}_1)$. Defining

$$\mathbf{\Phi}_1 \triangleq \begin{bmatrix} \mathbf{I}_{(N_R-N_T)N_T\times(N_R-N_T)N_T} \\ \mathbf{0}_{(N_R-N_T)(P+L-1)N_T\times(N_R-N_T)N_T} \end{bmatrix},$$

and using (3.58), we can derive

$$\mathrm{vec}(\mathbf{G}_1) = \mathbf{\Phi}_1 \mathrm{vec}\left\{\mathbf{U}_{0null}^H \mathbf{\Xi}^H \left(\mathbf{H}(0)\mathbf{H}^H(0)\right)^\dagger \mathbf{H}(0)\right\} = \mathbf{\Omega}_1 \mathrm{vec}\left(\mathbf{\Xi}^H\right) \tag{A-1}$$

where

$$\mathbf{\Omega}_1 \triangleq \mathbf{\Phi}_1 \left\{\left[\mathbf{H}^T(0)\left(\mathbf{H}(0)\mathbf{H}^H(0)\right)^\dagger\right] \otimes \mathbf{U}_{0null}^H\right\}, \tag{A-2}$$

$$\mathrm{vec}\left(\mathbf{\Xi}^H\right) = \mathrm{vec}\left(\Delta\mathbf{R}_{v3}^H\right) - \mathbf{\Psi}_1 \mathrm{vec}\left(\Delta\mathbf{R}_{v2}\right) - \mathbf{\Psi}_2 \mathrm{vec}\left(\Delta\mathbf{R}_{v2}^H\right) + \mathbf{\Psi}_3 \mathrm{vec}\left(\Delta\mathbf{R}_{v1}^H\right), \tag{A-3}$$

with

$$\mathbf{\Psi}_1 \triangleq \left[\left(\mathbf{\Pi}_2 \mathbf{I}_C^H \mathbf{H}_A^H\right)^T \otimes \mathbf{I}_{N_R}\right],$$

$$\mathbf{\Psi}_2 \triangleq \left[\mathbf{I}_{N_R}^T \otimes \left(\mathbf{H}_A \mathbf{I}_C \mathbf{\Pi}_2^H\right)\right],$$

$$\mathbf{\Psi}_3 \triangleq \left[\left(\mathbf{\Pi}_2 \mathbf{I}_C^H \mathbf{H}_A^H\right)^T \otimes \left(\mathbf{H}_A \mathbf{I}_C \mathbf{\Pi}_2^H\right)\right].$$

It is clear from (3.59) that, to obtain vec $(\mathbf{G}_2)$, one needs to first compute vec $(\mathbf{Q})$. To this end, we define a transformation matrix $\mathbf{\Phi}_2 = \left[\mathbf{0}, \mathbf{\Phi}_{2,1}^H, \cdots, \mathbf{\Phi}_{2,P+L-1}^H\right]^H$ such that its submatrices, being of the same size and satisfying $[\mathbf{\Phi}_{2,1}, \cdots, \mathbf{\Phi}_{2,P+L-1}] = \mathbf{I}_{(P+L-1)N_T}$. Thus (3.55) can be rewritten as

$$\mathbf{Q} = (\mathbf{I}_{P+L} \otimes \mathbf{\Pi}_P)\mathbf{\Phi}_2. \tag{A-4}$$

From which we have

$$\text{vec}\,(\mathbf{Q}) = \left(\mathbf{\Phi}_2^T \otimes \mathbf{I}_{(P+L)N_R}\right)\text{vec}(\mathbf{I}_{P+L} \otimes \mathbf{\Pi}_P). \tag{A-5}$$

We now simplify vec $(\mathbf{Q})$ by introducing another transformation matrix $\mathbf{\Phi}_3$. Our idea is to construct $\mathbf{\Phi}_3$ such that

$$\text{vec}(\mathbf{I}_{P+L} \otimes \mathbf{\Pi}_P) = \mathbf{\Phi}_3 \text{vec}\,(\mathbf{\Pi}_P)\,,$$

which leads vec $(\mathbf{Q})$ to

$$\text{vec}\,(\mathbf{Q}) = \left(\mathbf{\Phi}_2^T \otimes \mathbf{I}_{(P+L)N_R}\right)\mathbf{\Phi}_3 \text{vec}\,(\mathbf{\Pi}_P)\,. \tag{A-6}$$

To obtain such a $\mathbf{\Phi}_3$, we define a $\left[(P+L)^2\,(P+L-1)\,N_R N_T\right]$-dimensional vector, denoted as

$$\mathbf{i}_{F2} = \text{vec}(\mathbf{I}_{P+L} \otimes \mathbf{I}_{F1})$$

where $\mathbf{I}_{F1}$ is an $N_R \times (P+L-1)\,N_T$ matrix whose elements are all unity. The matrix $\mathbf{\Phi}_3$ of size $\left[(P+L)^2\,(P+L-1)\,N_R N_T\right] \times \left[(P+L-1)\,N_R N_T\right]$ can then be constructed according to the following rules:

(i) If the n-th element of vector $\mathbf{i}_{F2}$ is zero, the n-th row of $\mathbf{\Phi}_3$ is set to a zero vector,

(ii) If the n-th element of vector $\mathbf{i}_{F2}$ is 1, then only one element in the n-th row of $\mathbf{\Phi}_3$ is 1 and all the other elements are zero. The position of the element 1 shifts to the right by one entry with respect to the previous row with an element of 1. If the element 1 in the previous reference row happens at the last column, the next element 1 returns to the first column. Note that the first element in the first row of $\mathbf{\Phi}_3$ is always 1. As an example, considering $P+L=3$, $N_R=1$ and $N_T=1$, we have

$$\mathbf{i}_{F2} = \text{vec}(\mathbf{I}_3 \otimes [1,1]) = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}^T.$$

Then, the corresponding $\mathbf{\Phi}_3$ is readily given by

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T.$$

Using (3.59) and (A-6) gives

$$\text{vec}\left(\mathbf{G}_2\right) = \mathbf{\Omega}_2 \text{vec}\left(\mathbf{\Pi}'_P\right) \tag{A-7}$$

where

$$\mathbf{\Omega}_2 \triangleq \left[\mathbf{I}_{(P+L)N_R} \otimes \left(\mathbf{I}_{P+L} \otimes \mathbf{U}_{0null}^H\right)\right] \left(\mathbf{\Phi}_2^T \otimes \mathbf{I}_{(P+L)N_R}\right) \mathbf{\Phi}_3, \tag{A-8}$$

$$\text{vec}\left(\mathbf{\Pi}'_P\right) = \mathbf{\Psi}_4 \text{vec}\left(\Delta\mathbf{R}_{v2}\right) - \mathbf{\Psi}_5 \text{vec}\left(\Delta\mathbf{R}_{v1}\right), \tag{A-9}$$

with

$$\mathbf{\Psi}_4 \triangleq \mathbf{\Pi}_2^T \otimes \mathbf{I}_{N_R},$$

$$\mathbf{\Psi}_5 \triangleq \mathbf{\Pi}_2^T \otimes \left(\mathbf{H}_A \mathbf{I}_C \mathbf{\Pi}_2^H\right).$$

Note that $\mathbf{\Pi}_P$ has been replaced by $\mathbf{\Pi}'_P$ in (A-7), since the signal perturbation term has been eliminated.

Finally, by using (A-1) and (A-7) in (3.60) and noting that $\Delta\mathbf{R}_{v1} = \Delta\mathbf{R}_{v1}^H$ and $\Delta\mathbf{R}_{v3} = \Delta\mathbf{R}_{v3}^H$, the MSE of the blind criterion can be derived as

$$\mathrm{MSE_B} = \mathrm{Trace}\left\{\boldsymbol{\Omega}_1\left(\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_2^H\right)\boldsymbol{\Omega}_1^H + \boldsymbol{\Omega}_2\boldsymbol{\Gamma}_3\boldsymbol{\Omega}_2^H - \boldsymbol{\Omega}_1\boldsymbol{\Gamma}_4\boldsymbol{\Omega}_2^H - \boldsymbol{\Omega}_2\boldsymbol{\Gamma}_4^H\boldsymbol{\Omega}_1^H\right\} \tag{A-10}$$

where

$$\boldsymbol{\Gamma}_1 = \mathbf{R}_{\Delta v4} + \boldsymbol{\Psi}_1\mathbf{R}_{\Delta v2}\boldsymbol{\Psi}_1^H + \boldsymbol{\Psi}_2\mathbf{R}_{\Delta v3}\boldsymbol{\Psi}_2^H + \boldsymbol{\Psi}_3\mathbf{R}_{\Delta v1}\boldsymbol{\Psi}_3^H,$$

$$\boldsymbol{\Gamma}_2 = -\mathbf{R}_{\Delta v9}^H\boldsymbol{\Psi}_1^H - \mathbf{R}_{\Delta v10}^H\boldsymbol{\Psi}_2^H + \mathbf{R}_{\Delta v7}^H\boldsymbol{\Psi}_3^H + \boldsymbol{\Psi}_1\mathbf{R}_{\Delta v8}\boldsymbol{\Psi}_2^H - \boldsymbol{\Psi}_1\mathbf{R}_{\Delta v5}^H\boldsymbol{\Psi}_3^H - \boldsymbol{\Psi}_2\mathbf{R}_{\Delta v6}^H\boldsymbol{\Psi}_3^H,$$

$$\boldsymbol{\Gamma}_3 = \boldsymbol{\Psi}_4\mathbf{R}_{\Delta v2}\boldsymbol{\Psi}_4^H - \boldsymbol{\Psi}_4\mathbf{R}_{\Delta v5}^H\boldsymbol{\Psi}_5^H - \boldsymbol{\Psi}_5\mathbf{R}_{\Delta v5}\boldsymbol{\Psi}_4^H + \boldsymbol{\Psi}_5\mathbf{R}_{\Delta v1}\boldsymbol{\Psi}_5^H,$$

$$\boldsymbol{\Gamma}_4 = \mathbf{R}_{\Delta v9}^H\boldsymbol{\Psi}_4^H - \mathbf{R}_{\Delta v7}^H\boldsymbol{\Psi}_5^H - \boldsymbol{\Psi}_1\mathbf{R}_{\Delta v2}\boldsymbol{\Psi}_4^H + \boldsymbol{\Psi}_1\mathbf{R}_{\Delta v5}^H\boldsymbol{\Psi}_5^H - \boldsymbol{\Psi}_2\mathbf{R}_{\Delta v8}^H\boldsymbol{\Psi}_4^H + \boldsymbol{\Psi}_2\mathbf{R}_{\Delta v6}^H\boldsymbol{\Psi}_5^H +$$

$$\boldsymbol{\Psi}_3\mathbf{R}_{\Delta v5}\boldsymbol{\Psi}_4^H - \boldsymbol{\Psi}_3\mathbf{R}_{\Delta v1}\boldsymbol{\Psi}_5^H,$$

with $\mathbf{R}_{\Delta vi}$, $(i = 1, \cdots, 10)$ being given by

$$\mathbf{R}_{\Delta v1} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{v1}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{v1}\right)\right\},$$

$$\mathbf{R}_{\Delta v2} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{v2}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{v2}\right)\right\},$$

$$\mathbf{R}_{\Delta v3} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{v2}^H\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{v2}^H\right)\right\},$$

$$\mathbf{R}_{\Delta v4} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{v3}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{v3}\right)\right\},$$

$$\mathbf{R}_{\Delta v5} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{v1}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{v2}\right)\right\},$$

$$\mathbf{R}_{\Delta v6} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{v1}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{v2}^H\right)\right\},$$

$$\mathbf{R}_{\Delta v7} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{v1}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{v3}\right)\right\},$$

$$\mathbf{R}_{\Delta v8} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{v2}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{v2}^H\right)\right\},$$

$$\mathbf{R}_{\Delta v9} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{v2}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{v3}\right)\right\},$$

$$\mathbf{R}_{\Delta v10} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{v2}^H\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{v3}\right)\right\}.$$

The closed-form expression of $\mathbf{R}_{\Delta vi}, (i = 1, \cdots, 10)$ is derived in Appendix B.

Appendix B

Derivation of $\mathbf{R}_{\Delta vi}, (i = 1, 2, \cdots, 10)$

We now derive the closed-form expression of $\mathbf{R}_{\Delta vi}, (i = 1, \cdots, 10)$. First of all, we consider $\mathbf{R}_{\Delta v1}$. For a medium to high SNR, the perturbation matrix of the noise autocorrelation matrix can be neglected. Thus, from (3.30), one can get

$$\mathrm{vec}\left(\Delta\mathbf{R}_{v1}\right) \approx \left(\mathbf{I}_{PN_R} \otimes \mathbf{H}_B\right) \mathrm{vec}\left(\Delta\mathbf{R}_{xv1}\right) + \left(\mathbf{H}_B^* \otimes \mathbf{I}_{PN_R}\right) \mathrm{vec}\left(\Delta\mathbf{R}_{xv1}^H\right).$$

Since the noise element $v_{i_R}(n)$ is circularly symmetric with equal variance $\delta_v^2/2$ for the real and imaginary part, one can verify that $\mathrm{E}\left\{v_{i_R}(n)\, v_{i_R}(n)\right\} = \frac{1}{2}\delta_v^2 - \frac{1}{2}\delta_v^2 = 0$ and $\mathrm{E}\left\{v_{i_R}(n)\, v_{i_R}^*(n)\right\} = \frac{1}{2}\delta_v^2 + \frac{1}{2}\delta_v^2 = \delta_v^2$. In addition, it can be easily verified that $\mathrm{E}\left[\mathrm{vec}\left(\Delta\mathbf{R}_{xv1}\right)\ \mathrm{vec}^H\left(\Delta\mathbf{R}_{xv1}^H\right)\right] = \mathbf{0}$. Therefore, $\mathbf{R}_{\Delta v1}$ can be calculated by

$$\begin{aligned}\mathbf{R}_{\Delta v1} \approx\ & \left(\mathbf{I}_{PN_R} \otimes \mathbf{H}_B\right) \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv1}\right) \mathrm{vec}^H\left(\Delta\mathbf{R}_{xv1}\right)\right\} \left(\mathbf{I}_{PN_R} \otimes \mathbf{H}_B^H\right) \\ & + \left(\mathbf{H}_B^* \otimes \mathbf{I}_{PN_R}\right) \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv1}^H\right) \mathrm{vec}^H\left(\Delta\mathbf{R}_{xv1}^H\right)\right\} \left(\mathbf{H}_B^T \otimes \mathbf{I}_{PN_R}\right). \quad \text{(B-1)}\end{aligned}$$

To this end, the following two correlation matrices are to be determined

$$\boldsymbol{\Upsilon}_1 \triangleq \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv1}\right) \mathrm{vec}^H\left(\Delta\mathbf{R}_{xv1}\right)\right\}, \quad \text{(B-2)}$$

$$\boldsymbol{\Upsilon}_1' \triangleq \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv1}^H\right) \mathrm{vec}^H\left(\Delta\mathbf{R}_{xv1}^H\right)\right\}. \quad \text{(B-3)}$$

To compute $\boldsymbol{\Upsilon}_1$, one can obtain from (3.31),

$$\mathrm{vec}\left(\Delta\mathbf{R}_{xv1}\right) = \frac{1}{K}\sum_{n=0}^{K-1}\left(\mathbf{X}_{\mathbf{I}1}\left(n\right)\mathbf{v}_P\left(n-1\right)\right) \tag{B-4}$$

where $\mathbf{X}_{\mathbf{I}1}\left(n\right) \triangleq \mathbf{I}_{PN_R} \otimes \mathbf{x}_{P+L-1}\left(n-1\right)$. Substituting (B-4) into (B-2) gives

$$\boldsymbol{\Upsilon}_1 \triangleq \frac{1}{K^2}\mathrm{E}\left\{\sum_{n_1=0}^{K-1}\sum_{n_2=0}^{K-1}\mathbf{X}_{\mathbf{I}1}\left(n_1\right)\mathbf{R}_{\mathbf{IV}1}\left(n_1-n_2\right)\mathbf{X}_{\mathbf{I}1}^H\left(n_2\right)\right\} \tag{B-5}$$

where $\mathbf{R}_{\mathbf{IV}1}\left(n_1-n_2\right) \triangleq \mathrm{E}\left(\mathbf{v}_P\left(n_1-1\right)\mathbf{v}_P^H\left(n_2-1\right)\right)$. It can be easily shown that $\mathbf{R}_{\mathbf{IV}1}\left(n_1-n_2\right)$ is a partial identity matrix. Further, it can be proved that

$$\mathbf{X}_{\mathbf{I}1}\left(n_1\right)\mathbf{R}_{\mathbf{IV}1}\left(n_1-n_2\right)\mathbf{X}_{\mathbf{I}1}^H\left(n_2\right) = \mathbf{R}_{\mathbf{IV}1}\left(n_1-n_2\right)\otimes\mathbf{R}_{\mathbf{IX}1}\left(n_1-n_2\right) \tag{B-6}$$

where $\mathbf{R}_{\mathbf{IX}1}\left(n_1-n_2\right) \triangleq \mathrm{E}\left(\mathbf{x}_{P+L-1}\left(n_1-1\right)\mathbf{x}_{P+L-1}^H\left(n_2-1\right)\right)$, which is also a partial identity matrix. Thus, substituting (B-6) into (B-5) gives

$$\boldsymbol{\Upsilon}_1 = \frac{1}{K}\sum_{i=1-P}^{P-1}\mathbf{R}_{\mathbf{IV}1}\left(i\right)\otimes\mathbf{R}_{\mathbf{IX}1}\left(i\right). \tag{B-7}$$

Similarly, one can derive

$$\boldsymbol{\Upsilon}_1' = \frac{1}{K}\sum_{i=1-P}^{P-1}\mathbf{R}_{\mathbf{IX}1}\left(i\right)\otimes\mathbf{R}_{\mathbf{IV}1}\left(i\right). \tag{B-8}$$

Thus, by using (B-7) and (B-8) into (B-1), one can obtain

$$\mathbf{R}_{\Delta v1} \approx \left(\mathbf{I}_{PN_R}\otimes\mathbf{H}_B\right)\boldsymbol{\Upsilon}_1\left(\mathbf{I}_{PN_R}\otimes\mathbf{H}_B^H\right) + \left(\mathbf{H}_B^*\otimes\mathbf{I}_{PN_R}\right)\boldsymbol{\Upsilon}_1'\left(\mathbf{H}_B^T\otimes\mathbf{I}_{PN_R}\right). \tag{B-9}$$

As a result, $\mathbf{R}_{\Delta v1}$ can now be explicitly computed in terms of $\boldsymbol{\Upsilon}_1$, $\boldsymbol{\Upsilon}_1'$ and the channel matrix $\mathbf{H}_B$.

In a similar manner, one can derive the expressions for $\mathbf{R}_{\Delta vi}, (i = 2, 3, \cdots, 10)$.

$$\mathbf{R}_{\Delta v2} \approx \left(\mathbf{I}_{PN_R}\otimes\mathbf{H}_A\right)\boldsymbol{\Upsilon}_2\left(\mathbf{I}_{PN_R}\otimes\mathbf{H}_A^H\right) + \left(\mathbf{H}_B^*\otimes\mathbf{I}_{N_R}\right)\boldsymbol{\Upsilon}_2'\left(\mathbf{H}_B^T\otimes\mathbf{I}_{N_R}\right),$$

$$\mathbf{R}_{\Delta v3} \approx \left(\mathbf{H}_A^* \otimes \mathbf{I}_{PN_R}\right) \boldsymbol{\Upsilon}_3 \left(\mathbf{H}_A^T \otimes \mathbf{I}_{PN_R}\right) + \left(\mathbf{I}_{N_R} \otimes \mathbf{H}_B\right) \boldsymbol{\Upsilon}_3' \left(\mathbf{I}_{N_R} \otimes \mathbf{H}_B^H\right),$$

$$\mathbf{R}_{\Delta v4} \approx \left(\mathbf{I}_{N_R} \otimes \mathbf{H}_A\right) \boldsymbol{\Upsilon}_4 \left(\mathbf{I}_{N_R} \otimes \mathbf{H}_A^H\right) + \left(\mathbf{H}_A^* \otimes \mathbf{I}_{N_R}\right) \boldsymbol{\Upsilon}_4 \left(\mathbf{H}_A^T \otimes \mathbf{I}_{N_R}\right),$$

$$\mathbf{R}_{\Delta v5} \approx \left(\mathbf{I}_{PN_R} \otimes \mathbf{H}_B\right) \boldsymbol{\Upsilon}_5 \left(\mathbf{I}_{PN_R} \otimes \mathbf{H}_A^H\right) + \left(\mathbf{H}_B^* \otimes \mathbf{I}_{PN_R}\right) \boldsymbol{\Upsilon}_5' \left(\mathbf{H}_B^T \otimes \mathbf{I}_{N_R}\right),$$

$$\mathbf{R}_{\Delta v6} \approx \left(\mathbf{I}_{PN_R} \otimes \mathbf{H}_B\right) \boldsymbol{\Upsilon}_6 \left(\mathbf{I}_{N_R} \otimes \mathbf{H}_B^H\right) + \left(\mathbf{H}_B^* \otimes \mathbf{I}_{PN_R}\right) \boldsymbol{\Upsilon}_6' \left(\mathbf{H}_A^T \otimes \mathbf{I}_{PN_R}\right),$$

$$\mathbf{R}_{\Delta v7} \approx \left(\mathbf{I}_{PN_R} \otimes \mathbf{H}_B\right) \boldsymbol{\Upsilon}_7 \left(\mathbf{I}_{N_R} \otimes \mathbf{H}_A^H\right) + \left(\mathbf{H}_B^* \otimes \mathbf{I}_{PN_R}\right) \boldsymbol{\Upsilon}_7' \left(\mathbf{H}_A^T \otimes \mathbf{I}_{N_R}\right),$$

$$\mathbf{R}_{\Delta v8} \approx \left(\mathbf{I}_{PN_R} \otimes \mathbf{H}_A\right) \boldsymbol{\Upsilon}_8 \left(\mathbf{I}_{N_R} \otimes \mathbf{H}_B^H\right) + \left(\mathbf{H}_B^* \otimes \mathbf{I}_{N_R}\right) \boldsymbol{\Upsilon}_8' \left(\mathbf{H}_A^T \otimes \mathbf{I}_{PN_R}\right),$$

$$\mathbf{R}_{\Delta v9} \approx \left(\mathbf{I}_{PN_R} \otimes \mathbf{H}_A\right) \boldsymbol{\Upsilon}_9 \left(\mathbf{I}_{N_R} \otimes \mathbf{H}_A^H\right) + \left(\mathbf{H}_B^* \otimes \mathbf{I}_{N_R}\right) \boldsymbol{\Upsilon}_9' \left(\mathbf{H}_A^T \otimes \mathbf{I}_{N_R}\right),$$

$$\mathbf{R}_{\Delta v10} \approx \left(\mathbf{H}_A^* \otimes \mathbf{I}_{PN_R}\right) \boldsymbol{\Upsilon}_{10} \left(\mathbf{H}_A^T \otimes \mathbf{I}_{N_R}\right) + \left(\mathbf{I}_{N_R} \otimes \mathbf{H}_B\right) \boldsymbol{\Upsilon}_{10}' \left(\mathbf{I}_{N_R} \otimes \mathbf{H}_A^H\right)$$

where

$$\boldsymbol{\Upsilon}_2 = \frac{1}{K} \sum_{i=1-L}^{L-1} \mathbf{R}_{\mathbf{IV}1}(i) \otimes \mathbf{R}_{\mathbf{IX}2}(i),$$

$$\boldsymbol{\Upsilon}_2' = \frac{\delta_x^2 \delta_v^2}{K} \mathbf{I}_{N_R N_T (P+L-1)},$$

$$\boldsymbol{\Upsilon}_3 = \frac{1}{K} \sum_{i=1-L}^{L-1} \mathbf{R}_{\mathbf{IX}2}(i) \otimes \mathbf{R}_{\mathbf{IV}1}(i),$$

$$\boldsymbol{\Upsilon}_3' = \frac{\delta_x^2 \delta_v^2}{K} \mathbf{I}_{N_R N_T (P+L-1)},$$

$$\boldsymbol{\Upsilon}_4 = \frac{\delta_x^2 \delta_v^2}{K} \mathbf{I}_{N_R N_T L},$$

$$\boldsymbol{\Upsilon}_5 = \frac{1}{K} \sum_{i=1-P}^{P-1} \mathbf{R}_{\mathbf{IV}1}(i) \otimes \mathbf{R}_{\mathbf{IX}3}(i),$$

$$\boldsymbol{\Upsilon}_5' = \frac{1}{K} \sum_{i=1}^{P} \mathbf{R}_{\mathbf{IX}1}(i) \otimes \mathbf{R}_{\mathbf{IV}2}(i),$$

$$\boldsymbol{\Upsilon}_6 = \frac{1}{K} \sum_{i=1}^{P} \mathbf{R}_{\mathbf{IV}2}(i) \otimes \mathbf{R}_{\mathbf{IX}1}(i),$$

$$\boldsymbol{\Upsilon}_6' = \frac{1}{K} \sum_{i=1-P}^{P-1} \mathbf{R}_{\mathbf{IX}3}(i) \otimes (i)\, \mathbf{R}_{\mathbf{IV}1}(i),$$

$$\boldsymbol{\Upsilon}_7 = \frac{1}{K} \sum_{i=1}^{P} \mathbf{R}_{\mathbf{IV}2}(i) \otimes \mathbf{R}_{\mathbf{IX}3}(i),$$

$$\boldsymbol{\Upsilon}_7' = \frac{1}{K} \sum_{i=1}^{P} \mathbf{R}_{\mathbf{IX}3}(i) \otimes \mathbf{R}_{\mathbf{IV}2}(i),$$

$$\boldsymbol{\Upsilon}_8 = \frac{1}{K} \sum_{i=1}^{L-2} \mathbf{R}_{\mathbf{IV}2}(i) \otimes \mathbf{R}_{\mathbf{IX}4}(i),$$

$$\boldsymbol{\Upsilon}_8' = \frac{1}{K} \sum_{i=1-L}^{-1} \mathbf{R}_{\mathbf{IX}3}(i) \otimes \mathbf{R}_{\mathbf{IV}3}(i),$$

$$\Upsilon_9 = \frac{1}{K}\sum_{i=1}^{L-1} \mathbf{R_{IV2}}(i) \otimes \mathbf{R_{IX2}}(i),$$

$$\Upsilon'_9 = \frac{\delta_v^2}{K}\mathbf{R_{IX3}}(0) \otimes \mathbf{I}_{N_R},$$

$$\Upsilon_{10} = \frac{1}{K}\sum_{i=1}^{L-1} \mathbf{R_{IX2}}(i) \otimes \mathbf{R_{IV2}}(i),$$

$$\Upsilon'_{10} = \frac{\delta_v^2}{K}\mathbf{I}_{N_R} \otimes \mathbf{R_{IX3}}(0),$$

with

$$\mathbf{R_{IX2}}(i = n_1 - n_2) = \mathrm{E}\left(\mathbf{x}_L(n_1)\,\mathbf{x}_L^H(n_2)\right),$$

$$\mathbf{R_{IX3}}(i = n_1 - n_2) = \mathrm{E}\left(\mathbf{x}_{P+L-1}(n_1 - 1)\,\mathbf{x}_L^H(n_2)\right),$$

$$\mathbf{R_{IX4}}(i = n_1 - n_2) = \mathrm{E}\left(\mathbf{x}_L(n_1)\,\mathbf{x}_{P+L-1}^H(n_2 - 1)\right),$$

$$\mathbf{R_{IV2}}(i = n_1 - n_2) = \mathrm{E}\left(\mathbf{v}_P(n_1 - 1)\,\mathbf{v}^H(n_2)\right),$$

$$\mathbf{R_{IV3}}(i = n_1 - n_2) = \mathrm{E}\left(\mathbf{v}(n_1)\,\mathbf{v}_P^H(n_2 - 1)\right).$$

Appendix C

Proof of *Theorem* 5.1

Letting

$$\mathbf{H}_A \triangleq [\mathbf{H}(0), \mathbf{H}(1), \cdots, \mathbf{H}(L-1)]$$

$$\mathbf{x}_L(n) \triangleq \left[\mathbf{x}^T(n) \cdots \mathbf{x}^T(n-L+1)\right]^T, (n = 0, 1, \cdots, K-1)$$

where $\mathbf{x}(n) = \mathbf{x}(K+n)$ for $n < 0$, the circular convolution (2.3) in the noise-free case can be rewritten in the matrix form as

$$\mathbf{y}(n) = \mathbf{H}_A \mathbf{x}_L(n). \tag{C-1}$$

Substituting (C-1) into (5.2) yields

$$\mathbf{R}(l) = \mathbf{H}_A \mathbf{R}_{x,L}(l) \mathbf{H}_A^H \tag{C-2}$$

where

$$\mathbf{R}_{x,L}(l) = \mathrm{E}\left\{\mathbf{x}_L(n) \mathbf{x}_L^H(n-l)\right\}. \tag{C-3}$$

When $l = 0$, one can easily verify that

$$\mathbf{R}_{x,L}(0) = \mathbf{I}_{N_T L}. \tag{C-4}$$

Using (C-4) into (C-2) gives

$$\mathbf{R}(0) = \sum_{i=0}^{L-1} \mathbf{H}(i)\,\mathbf{H}^H(i)\,. \tag{C-5}$$

When $l = 1$, (C-3) reduces to

$$\mathbf{R}_{x,L}(1) = \begin{bmatrix} \mathbf{0}_{N_T \times (L-1)N_T} & \mathbf{0}_{N_T \times N_T} \\ \mathbf{I}_{(L-1)N_T \times (L-1)N_T} & \mathbf{0}_{(L-1)N_T \times N_T} \end{bmatrix}, \tag{C-6}$$

leading to

$$\mathbf{R}(1) = \sum_{i=1}^{L-1} \mathbf{H}(i)\,\mathbf{H}^H(i-1)\,. \tag{C-7}$$

In a similar manner, one can derive

$$\mathbf{R}(l) = \sum_{i=l}^{L-1} \mathbf{H}(i)\,\mathbf{H}^H(i-l)\,, \text{ if } l = 2, \cdots, L-1. \tag{C-8}$$

Obviously, when $l > L-1$, one gets $\mathbf{R}(l) = \mathbf{0}$ due to the fact that $\mathbf{R}_{x,L}(l) = \mathbf{0}$. Thus, (5.30) is obtained.

Appendix D

Proof of *Theorem* 6.1

Here, we simplify the calculation of $\mathbf{R}_{\Delta xv} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^{H}\right)\mathrm{vec}^{H}\left(\Delta\mathbf{R}_{xv}^{H}\right)\right\}$ by using the statistical property of the signal as well as the noise. Using (6.4), one can get

$$\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^{H}\right) = \frac{1}{N}\sum_{n=1}^{N}\left\{\left[\mathbf{x}^{*}\left(n\right)\otimes\mathbf{I}_{N_R}\right]\mathrm{vec}\left[\mathbf{v}\left(n\right)\right]\right\}, \tag{D-1}$$

which leads to

$$\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^{H}\right)\mathrm{vec}^{H}\left(\Delta\mathbf{R}_{xv}^{H}\right)\right\} = \frac{1}{N^2}\mathrm{E}\left\{\sum_{n_1=1}^{N}\sum_{n_2=1}^{N}\left[\mathbf{x}^{*}\left(n_1\right)\otimes\mathbf{I}_{N_R}\right]\right.$$
$$\left.\mathrm{E}\left[\mathbf{v}\left(n_1\right)\mathbf{v}^{H}\left(n_2\right)\right]\left[\mathbf{x}^{T}\left(n_2\right)\otimes\mathbf{I}_{N_R}\right]\right\}. \tag{D-2}$$

Noting that $\mathrm{E}\left[\mathbf{v}\left(n_1\right)\mathbf{v}^{H}\left(n_2\right)\right] = \delta\left(n_1-n_2\right)\delta_v^2\mathbf{I}$ and $\mathrm{E}\left[\mathbf{x}\left(n\right)\mathbf{x}^{H}\left(n\right)\right] = \delta_x^2\mathbf{I}$, we can obtain

$$\begin{aligned}\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^{H}\right)\mathrm{vec}^{H}\left(\Delta\mathbf{R}_{xv}^{H}\right)\right\} &= \frac{\delta_v^2}{N}\left\{\mathrm{E}\left[\frac{1}{N}\sum_{n=1}^{N}\mathbf{x}^{*}\left(n\right)\mathbf{x}^{T}\left(n\right)\right]\otimes\mathbf{I}_{N_R}\right\}\\ &= \frac{1}{N}\delta_x^2\delta_v^2\mathbf{I}_{N_RN_T}.\end{aligned} \tag{D-3}$$

Appendix E

Proof of *Theorem* 6.2

Since $\mathbf{Z}$ is a full rank square matrix, we have $\sigma_{z_1} \geq \sigma_{z_2} \geq \cdots \geq \sigma_{z_M} > 0$. By letting $\mathbf{\Psi}_1 \triangleq \Delta\mathbf{U}_Z^H\mathbf{U}_Z$ and $\mathbf{\Psi}_2 \triangleq \mathbf{V}_Z^H\Delta\mathbf{V}_Z$, $\mathbf{\Omega}$ can be rewritten as

$$\mathbf{\Omega} = \mathbf{\Psi}_1 + \mathbf{\Psi}_2. \tag{E-1}$$

In what follows, we compute the elements of $\mathbf{\Omega}$, i.e., those of $\mathbf{\Psi}_1$ and $\mathbf{\Psi}_2$. Let us consider the non-diagonal elements first. Noting that $\hat{\mathbf{U}}_Z^H\hat{\mathbf{U}}_Z \approx \mathbf{U}_Z^H\mathbf{U}_Z + \mathbf{\Psi}_1^H + \mathbf{\Psi}_1$ and $\hat{\mathbf{U}}_Z^H\hat{\mathbf{U}}_Z = \mathbf{U}_Z^H\mathbf{U}_Z = \mathbf{I}$, one can get

$$\mathbf{\Psi}_1 + \mathbf{\Psi}_1^H = \mathbf{0}, \tag{E-2}$$

implying that $\mathbf{\Psi}_1$ is a skew-Hermitian matrix. Using (6.78) and neglecting the second-order error terms, one can prove that

$$\mathbf{U}_Z^H\left(\hat{\mathbf{Z}}\hat{\mathbf{Z}}^H - \mathbf{Z}\mathbf{Z}^H\right)\mathbf{U}_Z \approx \mathbf{\Psi}_1^H\mathbf{\Sigma}_Z^2 + \mathbf{\Sigma}_Z^2\mathbf{\Psi}_1 + 2\mathbf{\Sigma}_Z\Delta\mathbf{\Sigma}_Z. \tag{E-3}$$

Using the SVD of $\mathbf{Z}$ and that of $\hat{\mathbf{Z}} = \mathbf{Z} + \Delta\mathbf{Z}$ into (E-3) gives

$$\mathbf{\Psi}_1^H\mathbf{\Sigma}_Z^2 + \mathbf{\Sigma}_Z^2\mathbf{\Psi}_1 \approx \left(\mathbf{U}_Z^H\Delta\mathbf{Z}\mathbf{V}_Z - \Delta\mathbf{\Sigma}_Z\right)\mathbf{\Sigma}_Z + \mathbf{\Sigma}_Z\left(\mathbf{U}_Z^H\Delta\mathbf{Z}\mathbf{V}_Z - \Delta\mathbf{\Sigma}_Z\right)^H. \tag{E-4}$$

By pre- and post-multiplying both sides of (E-4) by $\mathbf{\Sigma}_Z^{-1}$, one can get

$$\mathbf{\Sigma}_Z^{-1}\mathbf{\Psi}_1^H\mathbf{\Sigma}_Z + \mathbf{\Sigma}_Z\mathbf{\Psi}_1\mathbf{\Sigma}_Z^{-1} = \mathbf{\Sigma}_Z^{-1}\mathbf{A} + \mathbf{A}^H\mathbf{\Sigma}_Z^{-1} \tag{E-5}$$

where

$$\mathbf{A} \triangleq \mathbf{U}_Z^H\Delta\mathbf{Z}\mathbf{V}_Z - \Delta\mathbf{\Sigma}_Z. \tag{E-6}$$

Using (E-2) as well as (E-5), one can obtain

$$\mathbf{\Psi}_1(i,j) = \frac{\sigma_{z_i}^{-1}\mathbf{A}(i,j) + \sigma_{z_j}^{-1}\mathbf{A}^*(j,i)}{\sigma_{z_i}\sigma_{z_j}^{-1} - \sigma_{z_j}\sigma_{z_i}^{-1}}, (i \neq j). \tag{E-7}$$

We now consider the non-diagonal elements of $\mathbf{\Psi}_2$. Utilizing $\hat{\mathbf{V}}_Z^H\hat{\mathbf{V}}_Z \approx \mathbf{V}_Z^H\mathbf{V}_Z + \mathbf{\Psi}_2^H + \mathbf{\Psi}_2$ and noting that $\hat{\mathbf{V}}_Z^H\hat{\mathbf{V}}_Z = \mathbf{V}_Z^H\mathbf{V}_Z = \mathbf{I}$, one can show that $\mathbf{\Psi}_2$ is also skew-Hermitian. In order to determine the non-diagonal elements of $\mathbf{\Psi}_2$, we compute from (6.78)

$$\mathbf{V}_Z^H\left(\hat{\mathbf{Z}}^H\hat{\mathbf{Z}} - \mathbf{Z}^H\mathbf{Z}\right)\mathbf{V}_Z \approx \mathbf{\Psi}_2\mathbf{\Sigma}_Z^2 + \mathbf{\Sigma}_Z^2\mathbf{\Psi}_2^H + 2\mathbf{\Sigma}_Z\Delta\mathbf{\Sigma}_Z. \tag{E-8}$$

In a manner similar to the derivation of (E-5), one can obtain

$$\mathbf{\Sigma}_Z^{-1}\mathbf{\Psi}_2\mathbf{\Sigma}_Z + \mathbf{\Sigma}_Z\mathbf{\Psi}_2^H\mathbf{\Sigma}_Z^{-1} = \mathbf{A}\mathbf{\Sigma}_Z^{-1} + \mathbf{\Sigma}_Z^{-1}\mathbf{A}^H, \tag{E-9}$$

which leads the non-diagonal elements of $\mathbf{\Psi}_2$ to

$$\mathbf{\Psi}_2(i,j) = \frac{\sigma_{z_j}^{-1}\mathbf{A}(i,j) + \sigma_{z_i}^{-1}\mathbf{A}^*(j,i)}{\sigma_{z_j}\sigma_{z_i}^{-1} - \sigma_{z_i}\sigma_{z_j}^{-1}}, (i \neq j). \tag{E-10}$$

Using (E-7) and (E-10) into (E-1), one can eventually obtain the non-diagonal elements of $\mathbf{\Omega}$ as

$$\mathbf{\Omega}(i,j) = \frac{\mathbf{A}^*(j,i) - \mathbf{A}(i,j)}{\sigma_{z_i} + \sigma_{z_j}}, (i \neq j). \tag{E-11}$$

We now determine the diagonal elements of $\mathbf{\Omega}$. From (6.78), one can have

$$\Delta\mathbf{Z} - \mathbf{U}_Z\Delta\mathbf{\Sigma}_Z\mathbf{V}_Z^H \approx \mathbf{U}_Z\mathbf{\Sigma}_Z\Delta\mathbf{V}_Z^H + \Delta\mathbf{U}_Z\mathbf{\Sigma}_Z\mathbf{V}_Z^H. \tag{E-12}$$

Premultiplying $\mathbf{U}_Z^H$ and postmultiplying $\mathbf{V}_Z$ on both sides of (E-12) give

$$\mathbf{U}_Z^H \Delta \mathbf{Z} \mathbf{V}_Z - \Delta \mathbf{\Sigma}_Z = \mathbf{\Sigma}_Z \Delta \mathbf{V}_Z^H \mathbf{V}_Z + \mathbf{U}_Z^H \Delta \mathbf{U}_Z \mathbf{\Sigma}_Z. \tag{E-13}$$

Using (E-6) and (E-13) and noting that $\mathbf{\Psi}_1 = \Delta \mathbf{U}_Z^H \mathbf{U}_Z$ and $\mathbf{\Psi}_2 = \mathbf{V}_Z^H \Delta \mathbf{V}_Z$, one can verify

$$\mathbf{\Psi}_2 \mathbf{\Sigma}_Z + \mathbf{\Sigma}_Z \mathbf{\Psi}_1 = \mathbf{A}^H. \tag{E-14}$$

Recall that $\mathbf{\Sigma}_Z$ is a real diagonal matrix and both $\mathbf{\Psi}_1$ and $\mathbf{\Psi}_2$ are skew-Hermitian which implies that their diagonal elements are imaginary. That is to say the diagonal elements of $\mathbf{\Psi}_2 \mathbf{\Sigma}_Z + \mathbf{\Sigma}_Z \mathbf{\Psi}_1$ are imaginary. As a result, we have

$$\operatorname{diag}\left(\mathbf{\Psi}_2 \mathbf{\Sigma}_Z + \mathbf{\Sigma}_Z \mathbf{\Psi}_1\right) = \operatorname{diag}\left(\frac{\mathbf{A}^H - \mathbf{A}}{2}\right). \tag{E-15}$$

Using (E-1) and (E-15), the diagonal elements of $\mathbf{\Omega}$ are readily given by

$$\mathbf{\Omega}(i,i) = \frac{\mathbf{A}^*(i,i) - \mathbf{A}(i,i)}{2\sigma_{z_i}}. \tag{E-16}$$

From (E-6), (E-11) and (E-16), the theorem is proved.

Appendix F

Proof of *Theorem* 6.3

Letting

$$\boldsymbol{\Upsilon} \triangleq \mathrm{E}\left\{\left[\mathrm{vec}\left(\boldsymbol{\Gamma}_Q\right)\mathrm{vec}^H\left(\boldsymbol{\Gamma}_Q\right)\right] \circ \left[\mathrm{vec}\left(\boldsymbol{\Pi}^H\right)\mathrm{vec}^H\left(\boldsymbol{\Pi}^H\right)\right]\right\}, \tag{F-1}$$

G_1 given by (6.91) can be rewritten as

$$G_1 = \mathrm{Trace}\left\{\left[\mathbf{V}_S^* \otimes \left(\mathbf{U}_S\boldsymbol{\Sigma}_S\right)\right]\boldsymbol{\Upsilon}\left[\mathbf{V}_S^* \otimes \left(\mathbf{U}_S\boldsymbol{\Sigma}_S\right)\right]^H\right\}. \tag{F-2}$$

In order to determine G_1, we first compute $\mathrm{E}\left[\mathrm{vec}\left(\boldsymbol{\Pi}^H\right)\mathrm{vec}^H\left(\boldsymbol{\Pi}^H\right)\right]$. From (6.85), one can get

$$\boldsymbol{\Pi}^H = \mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{U}_S\boldsymbol{\Sigma}_S + \boldsymbol{\Sigma}_S\mathbf{U}_S^H\Delta\mathbf{R}_{xv,P}^H\mathbf{V}_S - \boldsymbol{\Sigma}_S\mathbf{U}_S^H\Delta\mathbf{R}_{xv}^H\mathbf{V}_S - \mathbf{V}_S^H\Delta\mathbf{R}_{xv,P}\mathbf{U}_S\boldsymbol{\Sigma}_S. \tag{F-3}$$

Obviously, the computation of $\mathrm{E}\left[\mathrm{vec}\left(\boldsymbol{\Pi}^H\right)\mathrm{vec}^H\left(\boldsymbol{\Pi}^H\right)\right]$ involves a total of 16 terms. It can be shown that computing these terms requires the auto-correlation as well as the cross-correlation matrices of $\mathrm{vec}\left(\Delta\mathbf{R}_{xv}\right)$, $\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right)$, $\mathrm{vec}\left(\Delta\mathbf{R}_{xv,P}\right)$ and $\mathrm{vec}\left(\Delta\mathbf{R}_{xv,P}^H\right)$. Following a manner similar to Appendix D, one can derive

$$\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv}\right)\right\} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv}^H\right)\right\} = \frac{\delta_v^2}{N}\mathbf{I}_{N_RN_T}, \tag{F-4}$$

$$\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv,P}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv,P}\right)\right\} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv,P}^H\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv,P}^H\right)\right\} = \frac{\delta_v^2}{K}\mathbf{I}_{N_R N_T}, \tag{F-5}$$

$$\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv,P}\right)\right\} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv,P}^H\right)\right\} = \frac{\delta_v^2}{N}\mathbf{I}_{N_R N_T}, \tag{F-6}$$

$$\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv,P}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv,P}^H\right)\right\} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv}^H\right)\right\} = \mathbf{0}, \tag{F-7}$$

$$\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv,P}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv}^H\right)\right\} = \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}\right)\mathrm{vec}^H\left(\Delta\mathbf{R}_{xv,P}^H\right)\right\} = \mathbf{0}. \tag{F-8}$$

Then, the first term in the computation of $\mathrm{E}\left[\mathrm{vec}\left(\mathbf{\Pi}^H\right)\mathrm{vec}^H\left(\mathbf{\Pi}^H\right)\right]$ can be calculated using (F-4) as

$$\begin{aligned}&\mathrm{E}\left[\mathrm{vec}\left(\mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{U}_S\mathbf{\Sigma}_S\right)\mathrm{vec}^H\left(\mathbf{V}_S^H\Delta\mathbf{R}_{xv}\mathbf{U}_S\mathbf{\Sigma}_S\right)\right]\\ &= \frac{\delta_v^2}{N}\left[\left(\mathbf{\Sigma}_S\mathbf{U}_S^T\right)\otimes\mathbf{V}_S^H\right]\left[\left(\mathbf{U}_S^*\mathbf{\Sigma}_S\right)\otimes\mathbf{V}_S\right] = \frac{\delta_v^2}{N}\left(\mathbf{\Sigma}_S^2\otimes\mathbf{I}_{N_T}\right).\end{aligned} \tag{F-9}$$

In a similar manner, all the other terms can be computed using (F-4) to (F-8). By adding all the 16 terms together, we obtain

$$\mathrm{E}\left[\mathrm{vec}\left(\mathbf{\Pi}^H\right)\mathrm{vec}^H\left(\mathbf{\Pi}^H\right)\right] = \frac{\left(N-K\right)\delta_v^2}{KN}\left(\mathbf{\Sigma}_S^2\otimes\mathbf{I}_{N_T} + \mathbf{I}_{N_T}\otimes\mathbf{\Sigma}_S^2\right). \tag{F-10}$$

Utilizing (6.83) and (F-10), one can easily verify that $\mathbf{\Upsilon}$ is a diagonal matrix with diagonal elements being given by

$$\boldsymbol{\rho}\left(l = i + N_T\left(j-1\right)\right) = \frac{\left(N-K\right)\delta_v^2}{KN}\frac{1}{\sigma_{S_i}^2+\sigma_{S_j}^2}, (i,j = 1,2,\cdots,N_T). \tag{F-11}$$

Letting $\mathbf{u}_{S_i}$ and $\mathbf{v}_{S_j}$ be the i-th and j-th column vectors of $\mathbf{U}_S$ and $\mathbf{V}_S$, respectively, one can find the l-th column vector of $\mathbf{V}_S^*\otimes\left(\mathbf{U}_S\mathbf{\Sigma}_S\right)$ as $\sigma_{S_i}\mathbf{v}_{S_j}^*\otimes\mathbf{u}_{S_i}$. Therefore, (F-2) can be expressed as

$$G_1 = \frac{\left(N-K\right)\delta_v^2}{KN}\sum_{j=1}^{N_T}\sum_{i=1}^{N_T}\frac{\sigma_{S_i}^2}{\sigma_{S_i}^2+\sigma_{S_j}^2}\mathrm{Trace}\left[\left(\mathbf{v}_{S_j}^*\otimes\mathbf{u}_{S_i}\right)\left(\mathbf{v}_{S_j}^*\otimes\mathbf{u}_{S_i}\right)^H\right]. \tag{F-12}$$

Since Trace $\left[\left(\mathbf{v}_{S_j}^{\star} \otimes \mathbf{u}_{S_i}\right)\left(\mathbf{v}_{S_j}^{\star} \otimes \mathbf{u}_{S_i}\right)^H\right] = \left\|\mathbf{v}_{S_j}^{\star} \otimes \mathbf{u}_{S_i}\right\|_F^2 = 1$, from (F-12), the value of G_1 can be found to be

$$G_1 = \frac{(N-K)\,\delta_v^2}{KN} \sum_{j=1}^{N_T} \sum_{i=1}^{N_T} \frac{\sigma_{S_i}^2}{\sigma_{S_i}^2 + \sigma_{S_j}^2} = \frac{N_T^2\,(N-K)\,\delta_v^2}{2KN}. \tag{F-13}$$

Now we determine the value of G_2. The first two terms of the RHS of (6.92) can be rewritten as

$$\begin{aligned} &\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right)\left[\mathrm{vec}\left(\mathbf{\Gamma}_Q\right) \circ \mathrm{vec}\left(\mathbf{\Pi}^H\right)\right]^H\right\} \\ &= \left[\mathbf{1} \cdot \mathrm{vec}\left(\mathbf{\Gamma}_Q\right)^H\right] \circ \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right) \mathrm{vec}^H\left(\mathbf{\Pi}^H\right)\right\}. \end{aligned} \tag{F-14}$$

Using (F-3), (F-7) and (F-8), we have

$$\begin{aligned} \mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right) \mathrm{vec}^H\left(\mathbf{\Pi}^H\right)\right\} &= \mathrm{E}\left[\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right) \mathrm{vec}^H\left(\Delta\mathbf{R}_{xv,P}^H\right)\right]\left[\mathbf{V}_S^T \otimes \left(\mathbf{\Sigma}_S\mathbf{U}_S^H\right)\right]^H \\ &\quad - \mathrm{E}\left[\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right) \mathrm{vec}^H\left(\Delta\mathbf{R}_{xv}^H\right)\right]\left[\mathbf{V}_S^T \otimes \left(\mathbf{\Sigma}_S\mathbf{U}_S^H\right)\right]^H. \end{aligned} \tag{F-15}$$

Using (F-4) and (F-6) into (F-15) gives

$$\mathrm{E}\left\{\mathrm{vec}\left(\Delta\mathbf{R}_{xv}^H\right) \mathrm{vec}^H\left(\mathbf{\Pi}^H\right)\right\} = \frac{\delta_v^2}{N}\left[\mathbf{V}_S^T \otimes \left(\mathbf{\Sigma}_S\mathbf{U}_S^H\right) - \mathbf{V}_S^T \otimes \left(\mathbf{\Sigma}_S\mathbf{U}_S^H\right)\right]^H = \mathbf{0}. \tag{F-16}$$

Therefore, from (6.92), (F-14) and (F-16), we have $G_2 = 0$.

Appendix G

Frequency-Domain Estimation of Correlation Matrix $\hat{\mathbf{R}}_x(l)$

From (7.6) and $\mathbf{x}(n) \triangleq [x_1(n), \cdots, x_{N_T}(n)]^T$, we have

$$\hat{\mathbf{R}}_x(l) = \begin{bmatrix} \hat{R}_{1,1}(l) & \hat{R}_{1,2}(l) & \cdots & \hat{R}_{1,N_T}(l) \\ \hat{R}_{2,1}(l) & \hat{R}_{2,2}(l) & \cdots & \hat{R}_{2,N_T}(l) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{R}_{N_T,1}(l) & \hat{R}_{N_T,2}(l) & \cdots & \hat{R}_{N_T,N_T}(l) \end{bmatrix} \tag{G-1}$$

where

$$\hat{\mathbf{R}}_{i_{T1},i_{T2}}(l) = \frac{1}{K} \sum_{n=0}^{K-1} x_{i_{T1}}(n)\, x_{i_{T2}}^H(n-l). \tag{G-2}$$

Note that $x_{i_{T2}}(n)$ for $n < l$ can be obtained using $x_{i_{T2}}(n-l) \triangleq x_{i_{T2}}(K+n-l)$ due to the circular convolution. Utilizing

$$x_{i_{T1}}(n) = \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} X_{i_{T1}}(k)\, e^{j2\pi(kn/K)}, \tag{G-3}$$

$$x_{i_{T2}}(n) = \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} X_{i_{T2}}(k)\, e^{j2\pi(kn/K)}, \tag{G-4}$$

$\hat{R}_{i_{T1},i_{T2}}$ as defined in (G-2) can be written as

$$\begin{aligned}\hat{R}_{i_{T1},i_{T2}}(l) &= \frac{1}{K}\sum_{n=0}^{K-1} x_{i_{T1}}(n)\left[\frac{1}{\sqrt{K}}\sum_{k=0}^{K-1} X^*_{i_{T2}}(k)\,\mathrm{e}^{-j2\pi k(n-l)/K}\right] \\ &= \frac{1}{K}\sum_{k=0}^{K-1}\left[\frac{1}{\sqrt{K}}\sum_{n=0}^{K-1} x_{i_{T1}}(n)\,\mathrm{e}^{-j2\pi kn/K}\right] X^*_{i_{T2}}(k)\,\mathrm{e}^{j2\pi(kl/K)} \\ &= \frac{1}{K}\sum_{k=0}^{K-1} X_{i_{T1}}(k)\,X^*_{i_{T2}}(k)\,\phi^{-l}(k). \end{aligned} \tag{G-5}$$

Substituting (G-5) into (G-1) yields

$$\begin{aligned}\hat{\mathbf{R}}_x(l) &= \frac{1}{K}\sum_{k=0}^{K-1}\begin{bmatrix} X_1(k)X_1^*(k) & X_1(k)X_2^*(k) & \cdots & X_1(k)X_{N_T}^*(k) \\ X_2(k)X_1^*(k) & X_2(k)X_2^*(k) & \cdots & X_2(k)X_{N_T}^*(k) \\ \vdots & \vdots & \ddots & \vdots \\ X_{N_T}(k)X_1^*(k) & X_{N_T}(k)X_2^*(k) & \cdots & X_{N_T}(k)X_{N_T}^*(k) \end{bmatrix}\phi^{-l}(k), \\ &= \frac{1}{K}\sum_{k=0}^{K-1}\mathbf{X}(k)\,\mathbf{X}^H(k)\,\phi^{-l}(k). \end{aligned} \tag{G-6}$$

Appendix H

Estimation of Rotation Matrix $\mathbf{Q}_0$

Assuming orthogonal training pilots, namely, $\mathbf{X}_P(k)\mathbf{X}_P^H(k) = g_k\mathbf{I}_{N_T}$, from (7.44), one can obtain

$$\mathbf{Y}_P(k)\mathbf{X}_P^H(k) = g_k\mathbf{H}_{\mathrm{F}}(kM). \tag{H-1}$$

Using (7.45) in (H-1) and then substituting (H-1) into (7.47), we have

$$\mathbf{Y}_Q(k) = g_k\mathbf{W}_0^H\mathbf{P}_{R,F}^H(k)\mathbf{P}_{R,F}(k)\mathbf{H}(0). \tag{H-2}$$

Using (7.40) into (H-2) gives

$$\mathbf{Y}_Q(k) = g_k\left[\mathbf{P}_{R,F}(k)\mathbf{W}_0\right]^H\left[\mathbf{P}_{R,F}(k)\mathbf{W}_0\right]\mathbf{Q}_0^H. \tag{H-3}$$

Define

$$\mathbf{Y}_Q \triangleq \sum_{k=1}^{K_T}\mathbf{Y}_Q(k). \tag{H-4}$$

Substituting (H-3) into (H-4) yields

$$\mathbf{Y}_Q = \mathbf{\Theta}\mathbf{Q}_0^H \tag{H-5}$$

where

$$\mathbf{\Theta} \triangleq \sum_{k=1}^{K_T}\left\{g_k\left[\mathbf{P}_{R,F}(k)\mathbf{W}_0\right]^H\left[\mathbf{P}_{R,F}(k)\mathbf{W}_0\right]\right\}. \tag{H-6}$$

Clearly, $\boldsymbol{\Theta}$ is an $N_T \times N_T$ Hermitian matrix, whose SVD can be written as

$$\boldsymbol{\Theta} = \mathbf{U}_\Theta \boldsymbol{\Sigma}_\Theta \mathbf{U}_\Theta^H. \tag{H-7}$$

Utilizing (H-5) and (H-7) and noting that $\mathbf{Q}_0$ is an unitary matrix, the SVD of $\mathbf{Y}_Q$ can be written as

$$\mathbf{Y}_Q = \mathbf{U}_Q \boldsymbol{\Sigma}_Q \mathbf{V}_Q^H = [\mathbf{U}_\Theta \boldsymbol{\Omega}] \, \boldsymbol{\Sigma}_\Theta \, [\mathbf{Q}_0 \mathbf{U}_\Theta \boldsymbol{\Omega}]^H \tag{H-8}$$

where $\boldsymbol{\Omega}$ is a diagonal unitary matrix used to represent a general form of the SVD of $\mathbf{Y}_Q$, since, given a matrix, its SVD is not unique. Consequently, from (H-8), we have

$$\mathbf{V}_Q \mathbf{U}_Q^H = \mathbf{Q}_0 \mathbf{U}_\Theta \boldsymbol{\Omega} \, [\mathbf{U}_\Theta \boldsymbol{\Omega}]^H = \mathbf{Q}_0. \tag{H-9}$$