行政院國家科學委員會專題研究計畫 成果報告

具頻域等化機制之單載波空時區塊碼之盲蔽式多通道判別：以非多餘傳送前置編碼為基礎的研究(第 2 年)
研究成果報告(完整版)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 95-2221-E-009-047-MY2
執 行 期 間 ：96 年 08 月 01 日至 97 年 07 月 31 日
執 行 單 位 ：國立交通大學電信工程學系(所)

計 畫 主 持 人：李大嵩
共 同 主 持 人：吳卓諭
計畫參與人員：碩士班研究生-兼任助理人員：李思漢
碩士班研究生-兼任助理人員：宋志晟
博士班研究生-兼任助理人員：林光敏
博士班研究生-兼任助理人員：黃崇榮

報 告 附 件 ：出席國際會議研究心得報告及發表論文

處 理 方 式 ：本計畫可公開查詢

中 華 民 國 97 年 10 月 31 日

行政院國家科學委員會補助專題研究計畫 ■ 成 果 報 告
□期中進度報告

具頻域等化機制之單載波空時區塊碼之盲蔽式多通道判別：以非多餘傳送前置編碼為基礎的研究

計畫類別：■ 個別型計畫　　□ 整合型計畫
計畫編號：NSC 95－2221－E－009－047－MY2
執行期間：　96 年 8 月 1 日至 97 年 7 月 31 日

計畫主持人：李大嵩 教授
共同主持人：吳卓諭 助理教授
計畫參與人員：林光敏、黃崇榮、李思漢、宋志晟

成果報告類型(依經費核定清單規定繳交)：□精簡報告　■完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
■出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢
□涉及專利或其他智慧財產權，□一年□二年後可公開查詢

執行單位：

中　華　民　國　97　年　10　月　29　日

摘要

本計畫基於非冗長對角先期編碼和獨立同分佈（i.i.d.）訊源的假設，提出以單載波頻域等化器為基礎的空時區塊碼系統的盲蔽式通道估測法。此方法開發了共軛交叉相關介於兩個時間區塊信號的先期編碼所產生的線性信號結構和循環行列式的通道矩陣特性，並且在通道雜訊為循環高斯而接收機資料統計可以完整的得到時可以產生精確解。

此通道估測的公式化建立在重組共軛交叉相關矩陣的線性方程式集合以及通道脈衝響應，使之成為一個具有區塊循環循環區塊（block-circulant with circulant-block (BCCB)）的特殊結構。這樣允許了一個簡單的僅視先期編碼參數而定的可辨識條件，也提供了一個自然而有效的最佳先期編碼器的設計架構來改善當不完全資料估測發生時的解答正確性。

我們從明確和統計的觀點考慮兩種資料不匹配的模型，並且提出相關的設計準則。最佳化的問題以利用 BCCB 系統矩陣的特性公式化並加以分析求解。所提出的最佳化先期編碼器目的在於有明確誤差擾動時解的強健度最佳化和當資料不匹配以白雜訊模型化時的均方誤差最小化。配對的誤差機率分析用來探討等化器的性能，而數值分析的例子便展示了提出方法的性能。

關鍵詞：盲蔽式通道估測， 具有循環區塊的區塊循環矩陣，循環矩陣，多輸入單輸出，非冗長先期編碼器，單載波頻域等化器，空時區塊碼，傳送多樣性

Abstract

Relying on non-redundant diagonal precoding and i.i.d. source assumption, this paper proposes a blind channel estimation scheme for single-carrier frequency-domain equalization based space-time block coded systems. The proposed method exploits the precoding-induced linear signal structure in the conjugate cross-correlation between the two temporal block received signals as well as the circulant channel matrix property, and can yield exact solutions whenever the channel noise is circularly Gaussian and the receive data statistic is perfectly obtained. The channel estimation formulation builds on rearranging the set of linear equations relating the entries of conjugate cross-correlation matrix and products of channel impulse responses into one with a distinctive block-circulant with circulant-block (BCCB) structure. This allows a simple identifiability condition depending on precoder parameters alone, and also provides a natural yet effective optimal precoder design framework for improving solution accuracy when imperfect data estimation occurs. We consider two models of data mismatch, from both deterministic and statistical points of view, and propose the associated design criteria. The optimization problems are formulated to take advantage of the BCCB system matrix property and are solved analytically. The proposed optimal precoder aims to optimize solution robustness against deterministic error perturbation and also minimize the mean square error when the data mismatch is modeled as a white noise. Pair-wise error probability analysis is conducted for investigating the equalization performance. Numerical examples are used to illustrate the performance of the proposed method.

Keywords: Blind channel estimation, block-circulant matrix with circulant blocks (BCCB), circulant matrix, multiple input single output (MISO), nonredundant precoders, single-carrier frequency-domain equalization, space–time block code (STBC), transmit diversity.

Contents

Abstract	I
Chapter 1 Introduction	1
Chapter 2 System Model and Basic Assumptions	6
Chapter 3 Blind Channel Estimation	9
Chapter 4 Identifiability and Product Unknowns Computation	17
Chapter 5 Optimal Precoder Design	21
Chapter 6 Equalization Aspect	31
Chapter 7 Simulation Results	33
Chapter 8 Conclusions	45
Appendix	47
References	55

Chapter 1

Introduction

A. Overview

Space-time block code (STBC) is a widely-known transmit diversity technique for combating channel fading in modern wireless communications [22]. Most of the existing proposals are devised for the flat- fading channel environment, e.g., the Alamouti's scheme [1] and the related generalization by Tarokh et. al [34], among others. When the propagation channels are subject to frequency-selective fading, one popular STBC technique is via time-reversal block-wise encoding, either combined with OFDM mechanism [27], [40], or resorting to time-domain equalizer [26], for removing the channel distortion. The multi-carrier related solutions, although simplifying receiver implementations, would incur high peak-to-average power ratio (PAPR) and is sensitive to carrier frequency offset. The scheme with time-domain equalization, on the other hand, can provide additional multipath diversity at the expense of decoding complexity. To avoid the drawbacks of the multi-carrier strategy and to also maintain low receiver complexity, an alternative single-carrier frequency-domain equalization (FDE) based STBC was proposed in [2]. The aforementioned STBC's capable of mitigating dispersive channels can be cast into a general code formulation [39]; comparisons pf the achievable performances and implementation costs can be found in [3].

To realize the diversity benefit of STBC, the channel state information must be

known at the receiver to coherently combine the multiple temporal received signals for decoding[a]. Since STBC potentially entails low spectral efficiency and training based channel estimation further consumes bandwidth resource, blind approaches then become appealing candidate solutions. There has been extensive literature on blind multi-input multi-output (MIMO) channel estimation [14], [16]. However, only a few studies are tailored for STBC systems, typically through a multi-input single-output (MISO) channel link. Under flat-fading assumption, several schemes were put forth for orthogonal STBC [4], [7], [32], and for a general linear code family [33]. For time-reversal STBC over frequency-selective channels, the work [5] focused on codes with time-domain equalization [26]. Through linear symbol precoding, blind schemes for OFDM-based STBC were shown in [27] and [40]. The method [27] resorts to zero-padding for removing inter-block interference, and is applicable only for constant-modulus sources and channel pairs without common zeros; the one in [40], instead, uses cyclic prefix (CP) as guard interval and leverages redundant precoding to relieve the source and channel-zero constraints imposed in [27]. For FDE-STBC, training based channel estimation is recently considered in [11]. It is known that single-carrier FDE systems fall within the class of precoded OFDM, with FFT matrix as precoder [25]. In view of this fact, the method in [40] for OFDM scenario also provides an immediate blind solution for FDE-STBC: one just chooses FFT precoding matrix to convert the multi-carrier transmission into a single-carrier scheme and then inserts certain redundancy into the symbol streams to facilitate channel identification. The price to be paid for this approach, however, would be the

[a] Differential STBC does not require channel information for decoding but incurs a 3-dB penalty in SNR [22, Sec. 9.6].

loss in the effective data rate.

B. Paper Contributions

This paper proposes a blind channel estimation scheme for FDE-STBC systems in a two transmit antennas and single receive antenna environment. The proposed approach relies on *non-redundant* diagonal precoding (hence preserving the baud rate), assumes i.i.d. source statistics (irrespective of constellation modulus), and does not impose constraints on sub-channel zero locations. It exploits the precoding-induced linear signal structure in the time-domain conjugate cross-correlation between the two temporal receive branches, as well as the circulant channel matrix property after CP is discarded. Specifically, we show that the set of linear equations relating the entries of conjugate cross-correlation matrix and *products* of channel impulse responses can be rearranged into one with a block-circulant with circulant-block (BCCB) structure. The products of channel taps are first obtained by solving this linear equation set; the channel pair is then simultaneously identified, up to a 2×2 complex matrix ambiguity, as the dominant left singular vectors of an associated rank-two matrix. A similar "bilinear" estimation strategy has also been adopted in [13], [21], [24], [37]. In our formulation, a natural sufficient condition for unique channel recovery is the non-singularity of certain BCCB matrix with precoder coefficients as its entries. Channel identifiability is thus free from any priori assumptions on sub-channel characteristics, and is shown to be fulfilled by almost all choices of precoders. As long as the channel noise is circularly Gaussian and the received data statistic is perfectly obtained, the resultant channel estimate is exact. In the presence of finite-sample estimation error, the proposed channel estimation framework allows a natural precoder design formulation for improving solution

robustness. We consider two models of data mismatch, one as an unknown deterministic perturbation while the other statistically as a white noise, and propose the associated optimal precoder design criteria, aiming for minimizing the worst-case solution sensitivity to perturbation and mean-square errors, respectively. Both optimization problems are further formulated to take advantage of the BCCB system matrix property and are then analytically solved; the resultant solutions are shown to be the same two-level form precoder. Pair-wise error probability (PEP) analysis is conducted to investigate the equalization performance of the proposed optimal solution and characterize the associated design trade-off. It is noted that blind channel estimation via non-redundant diagonal precoding has been considered in the single channel case [31], [10], [24], [37]; the related generalizations to MIMO single- and multi-carrier spatial multiplexing systems can be found in [8] and [9]. The rest of this paper is organized as follows. Section II briefly describes the system model and the underlying assumptions. Section III presents the proposed method; the associated key features are investigated in Section IV. Section V addresses the optimal precoder design against imperfect data estimation. Section VI examines the equalization performance through PEP analysis. Section VII contains the simulation results. Finally, Section VIII is the conclusion.

Notation List: Let $\mathbb{R}^{m\times n}$ and $\mathbb{C}^{m\times n}$ be respectively the sets of $m\times n$ real and complex matrices. Denote by $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$, respectively the transpose, complex conjugate, and Hermitian operations. The symbols $\mathbf{I}_m$ and $\mathbf{0}_m$ denote the $m\times m$ identity and zero matrices; $\mathbf{0}_{m\times n}$ is the $m\times n$ zero matrix. The notation $\otimes$ stands for the Kronecker product [19, p-242]. For $\mathbf{X}\in\mathbb{C}^{m\times n}$ with $\mathbf{x}_j\in\mathbb{C}^m$ being the *j*th column, define $vec(\mathbf{X}) \coloneqq [\mathbf{x}_1^T \ \cdots \ \mathbf{x}_n^T]^T \in \mathbb{C}^{mn}$. For $\mathbf{x}\in\mathbb{C}^m$, let

$diag\{\mathbf{x}\}$ be the $m \times m$ diagonal matrix with the elements of $\mathbf{x}$ on the main diagonal. The notation Ey stands for the expected value of the random variable y, and $j \coloneqq \sqrt{-1}$. We denote by $\mathbf{F} \in \mathbb{C}^{N \times N}$ the FFT matrix with the kl-th entry $[\mathbf{F}]_{k,l} \coloneqq 1/\sqrt{N} \cdot \omega^{-(k-1)(l-1)}$, where $\omega \coloneqq \exp(j2\pi/N)$, $1 \leq k,l \leq N$. We denote by $\|\cdot\|$ the two-norm of a vector and by $\kappa(\mathbf{X})$ the condition number of matrix $\mathbf{X}$.

Chapter 2

System Model and Basic Assumptions

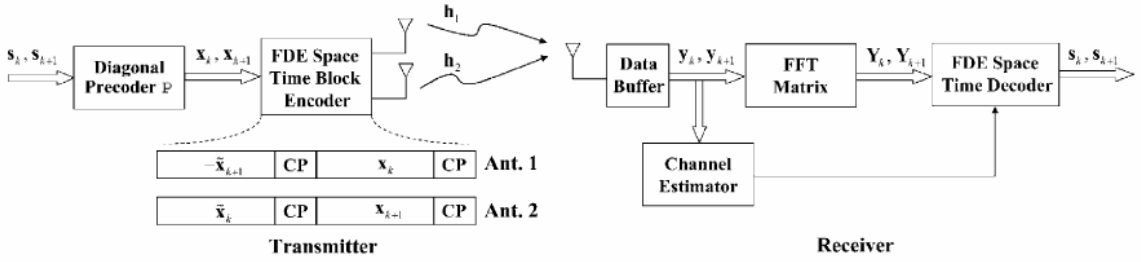

Fig. 1. FDE-STBC system diagram.

We consider the discrete-time baseband model of an FDE-STBC system [2] over frequency-selective channels as shown in Figure 1. Let $\mathbf{s}_k$ and $\mathbf{s}_{k+1}$ be two N-dimensional symbol blocks to be transmitted. Priori to the STBC encoder, each symbol block is precoded by an $N \times N$ diagonal matrix

$$\mathbf{P} := diag\left\{\left[p(0)\cdots p(N-1)\right]^T\right\}, \tag{2.1}$$

with $p(n) \in \mathbb{R}$, to obtain

$$\mathbf{x}_l = \mathbf{P}\mathbf{s}_l, \text{ for } l = k,\ k+1, \tag{2.2}$$

which are then spatially and temporally coded according to [2] for transmit diversity as well as for mitigating the multipath channel distortion. For $1 \le i \le 2$, let $h_i(n)$ be the impulse response of the channel between the ith transmit antenna and the receive antenna. In terms of block signals, the input-output relations in time-domain are described as [2]

$$\mathbf{y}_k = \mathbf{G}_1\mathbf{x}_k + \mathbf{G}_2\mathbf{x}_{k+1} + \mathbf{v}_k, \tag{2.3}$$

and

$$\mathbf{y}_{k+1} = \mathbf{G}_1\tilde{\mathbf{x}}_{k+1} - \mathbf{G}_2\tilde{\mathbf{x}}_k + \mathbf{v}_{k+1}, \tag{2.4}$$

where, for $l = k,\ k+1$, $\mathbf{y}_l$ and $\mathbf{v}_l$ are the received signal (upon CP removal) and noise, $\tilde{\mathbf{x}}_l$ is the time-reversed and element-wise conjugated version associated with $\mathbf{x}_l$, that is,

$$\tilde{\mathbf{x}}_l(n) = \mathbf{x}_l^*((-n)_N),\ \ 0 \le n \le N-1, \tag{2.5}$$

and $\mathbf{G}_i \in \mathbb{C}^{N\times N}$ is circulant with[b]

$$\mathbf{g}_i \coloneqq \begin{bmatrix} h_i(0) & \cdots & h_i(L) & 0 & \cdots & 0 \end{bmatrix}^T \in \mathbb{C}^N, \tag{2.6}$$

as the first column, $1 \le i \le 2$. Since $\mathbf{G}_i$ is circulant, we have $\mathbf{G}_i = \mathbf{F}^H\mathbf{D}_i\mathbf{F}$, where $\mathbf{D}_i$ is diagonal with $\left[\mathbf{D}_i\right]_{mm} = \sum_{n=0}^{L} h_i(n)\omega^{-(m-1)n}$, $1 \le m \le N$. Let us define $\mathbf{Y}_l \coloneqq \mathbf{F}\mathbf{y}_l$, $\mathbf{X}_l \coloneqq \mathbf{F}\mathbf{x}_l$, and $\mathbf{V}_l \coloneqq \mathbf{F}\mathbf{v}_l$, for $l = k,\ k+1$. Then the frequency-domain representation associated with (2.3) and (2.4) can be expressed in a compact vector-matrix form as [2]

$$\begin{bmatrix} \mathbf{Y}_k \\ \mathbf{Y}_{k+1}^* \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{D}_1 & \mathbf{D}_2 \\ -\mathbf{D}_2^* & \mathbf{D}_1^* \end{bmatrix}}_{\coloneqq\mathbf{D}} \begin{bmatrix} \mathbf{X}_k \\ \mathbf{X}_{k+1} \end{bmatrix} + \begin{bmatrix} \mathbf{V}_k \\ \mathbf{V}_{k+1}^* \end{bmatrix}; \tag{2.7}$$

through space-time matched filtering using the effective channel matrix $\mathbf{D}$ we get

$$\mathbf{D}^H \begin{bmatrix} \mathbf{Y}_k \\ \mathbf{Y}_{k+1}^* \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{D}} & \mathbf{0}_N \\ \mathbf{0}_N & \tilde{\mathbf{D}} \end{bmatrix} \begin{bmatrix} \mathbf{X}_k \\ \mathbf{X}_{k+1} \end{bmatrix} + \mathbf{D}^H \begin{bmatrix} \mathbf{V}_k \\ \mathbf{V}_{k+1}^* \end{bmatrix}, \tag{2.8}$$

---

[b] Without loss of generality we may take L as a common channel order, or simply an associated upper bound, which can be determined as the maximum among the two.

where $\tilde{\mathbf{D}} := \mathbf{D}_1^H \mathbf{D}_1 + \mathbf{D}_2^H \mathbf{D}_2 \in \mathbb{R}^{N \times N}$ is diagonal with $\left[\tilde{\mathbf{D}}\right]_{ii} = \left|\left[\mathbf{D}_1\right]_{ii}\right|^2 + \left|\left[\mathbf{D}_2\right]_{ii}\right|^2$, $1 \le i \le N$: this asserts that two-fold transmit diversity is achieved in the frequency domain. To recover the source signals, per-tone frequency-domain equalizer [2], [15] can be designed based on (2.8), as long as a channel estimate is available at the receiver. Based on the time-domain signal model (2.3) and (2.4), this paper proposes a blind channel estimation scheme by using the second-order statistics of the received signal and discusses an optimal design of the precoder $p(n)$ for improving channel estimation accuracy. The following assumptions are made in the sequel.

a) The source sequence $s(n)$ is independently identically distributed (i.i.d.) with zero mean and $Es(k)s(l)^* = \delta(k-l)$, where $\delta(\cdot)$ is the Kronecker delta function.

b) The noise $v(n)$ is white circular Gaussian with zero mean, variance σ_v^2, and is independent of the source sequence $s(n)$.

Chapter 3

Blind Channel Estimation

To introduce the proposed method, we first assume that all the signal statistics can be perfectly obtained; the case with imperfect data estimation will be treated later. To obtain the channel matrix $\mathbf{D}$, one may focus on direct estimation of the *N* tones of the channel frequency response. Since the block length *N* could be large, this strategy would involve considerable computational efforts. Hence, we propose to instead estimate the time-domain channel impulse response $h_i(n)$, for $0 \le n \le L$ and $1 \le i \le 2$; the gains of the associated frequency tones can then be obtained by using FFT operations.

A. Identification Equations

The proposed approach exploits the imbedded linear signal structure in the time-domain conjugate cross-correlation matrix of the two received signals $\mathbf{y}_k$ and $\mathbf{y}_{k+1}$ as well as the circulant property of the channel matrix $\mathbf{G}_i$. To proceed, let us first define the matrix

$$\mathbf{\Gamma} := \begin{bmatrix} 1 & 0 & \cdots & & 0 \\ 0 & \vdots & \cdots & & 1 \\ \vdots & 0 & \cdots & 1 & \vdots \\ 0 & 1 & \cdots & & 0 \end{bmatrix} \in \mathbb{R}^{N \times N}. \tag{3.1}$$

Then, from (2.5), it is easy to see

$$\tilde{\mathbf{x}}_k = \mathbf{\Gamma}\mathbf{x}_k^* \quad \text{and} \quad \tilde{\mathbf{x}}_{k+1} = \mathbf{\Gamma}\mathbf{x}_{k+1}^*. \tag{3.2}$$

With (2.2) and (3.2), the signal models (2.3) and (2.4) then become

$$\mathbf{y}_k = \mathbf{G}_1\mathbf{P}\mathbf{s}_k + \mathbf{G}_2\mathbf{P}\mathbf{s}_{k+1} + \mathbf{v}_k\,, \tag{3.3}$$

and

$$\mathbf{y}_{k+1} = \mathbf{G}_1\mathbf{\Gamma}\mathbf{P}\mathbf{s}_{k+1}^* - \mathbf{G}_2\mathbf{\Gamma}\mathbf{P}\mathbf{s}_k^* + \mathbf{v}_{k+1}\,. \tag{3.4}$$

From (3.3), (3.4), and by assumptions *a)* and *b)*, it is easy to check

$$\mathbf{R}_{\mathbf{y}}(1) \coloneqq E\mathbf{y}_k\mathbf{y}_{k+1}^T = \mathbf{G}_2\mathbf{P}^2\mathbf{\Gamma}\mathbf{G}_1^T - \mathbf{G}_1\mathbf{P}^2\mathbf{\Gamma}\mathbf{G}_2^T + \left\{\mathbf{G}_1\mathbf{P}E\mathbf{s}_k\mathbf{v}_{k+1}^T + \mathbf{G}_2\mathbf{P}E\mathbf{s}_{k+1}\mathbf{v}_{k+1}^T + E\mathbf{v}_k\mathbf{v}_{k+1}^T\right\} \tag{3.5}$$

.

Since the noise $v(n)$ is circular, we have $E\mathbf{v}_k\mathbf{v}_{k+1}^T = \mathbf{0}_N$. Also, we assume that both the real and imaginary components of the noise process are independent of those of the source sequence $s(n)$: this thus implies $E\mathbf{s}_k\mathbf{v}_{k+1}^T = E\mathbf{s}_{k+1}\mathbf{v}_{k+1}^T = \mathbf{0}_N$. Under these conditions, the noise contributions to the conjugate cross-correlation matrix $\mathbf{R}_{\mathbf{y}}(1)$ in (3.5) become a zero matrix, leading to

$$\mathbf{R}_{\mathbf{y}}(1) = \mathbf{G}_2\mathbf{P}^2\mathbf{\Gamma}\mathbf{G}_1^T - \mathbf{G}_1\mathbf{P}^2\mathbf{\Gamma}\mathbf{G}_2^T\,. \tag{3.6}$$

For a given $\mathbf{R}_{\mathbf{y}}(1)$, the matrix equation (3.6) defines a set of N^2 scalar equations nonlinear in the unknowns $h_i(0), \cdots, h_i(L), 1 \le i \le 2$, but is linear with respect to *product* channel coefficients $h_i(k)h_{\bar{i}}(l),\ 1 \le i,\bar{i} \le 2$. As a result, in lieu of directly solving for $h_i(0), \cdots, h_i(L)$, we propose to exploit the imbedded linear structure in $\mathbf{R}_{\mathbf{y}}(1)$ for channel estimation. This will be done by further taking into account the circulant property of the channel matrices $\mathbf{G}_i$'s.

Specifically, define the following permutation matrix

$$\mathbf{J} := \begin{bmatrix} \mathbf{0}_{1\times(N-1)} & 1 \\ \mathbf{I}_{N-1} & \mathbf{0}_{(N-1)\times 1} \end{bmatrix} \in \mathbb{R}^{N\times N}. \tag{3.7}$$

Since $\mathbf{G}_i$ is circulant, it can be expressed in terms of its first column (cf. (2.6)) as

$$\mathbf{G}_i = \begin{bmatrix} \mathbf{g}_i & \mathbf{J}\mathbf{g}_i & \cdots & \mathbf{J}^{N-2}\mathbf{g}_i & \mathbf{J}^{N-1}\mathbf{g}_i \end{bmatrix}, \ 1 \le i \le 2. \tag{3.8}$$

By definitions of $\mathbf{P}$ and $\mathbf{\Gamma}$ (see (2.1) and (3.1)) and from (3.8), it follows

$$\begin{aligned} &\mathbf{G}_2\mathbf{P}^2\mathbf{\Gamma}\mathbf{G}_1^T = \left(\mathbf{G}_2\mathbf{P}^2\right)\cdot\left(\mathbf{\Gamma}\mathbf{G}_1^T\right) \\ &= \underbrace{\begin{bmatrix} p(0)^2\,\mathbf{g}_2 & p(1)^2\mathbf{J}\mathbf{g}_2 & \cdots & p(N-1)^2\mathbf{J}^{N-1}\mathbf{g}_2 \end{bmatrix}}_{=\mathbf{G}_2\mathbf{P}^2}\cdot\underbrace{\begin{bmatrix} \mathbf{g}_1 & \mathbf{J}^{N-1}\mathbf{g}_1 & \cdots & \mathbf{J}\mathbf{g}_1 \end{bmatrix}^T}_{=\mathbf{\Gamma}\mathbf{G}_1^T} = \sum_{n=0}^{N-1} p(n)^2\mathbf{J}^n\mathbf{g}_2\mathbf{g}_1^T\left(\mathbf{J}^T\right)^{N-n}. \end{aligned} \tag{3.9}$$

Similarly, we have

$$\mathbf{G}_1\mathbf{P}^2\mathbf{\Gamma}\mathbf{G}_2^T = \sum_{n=0}^{N-1} p(n)^2\mathbf{J}^n\mathbf{g}_1\mathbf{g}_2^T\left(\mathbf{J}^T\right)^{N-n}. \tag{3.10}$$

Combining (3.9) and (3.10), $\mathbf{R}_{\mathbf{y}}(1)$ in (3.6) becomes

$$\mathbf{R}_{\mathbf{y}}(1) = \sum_{n=0}^{N-1} p(n)^2\mathbf{J}^n\tilde{\mathbf{G}}\left(\mathbf{J}^T\right)^{N-n}, \tag{3.11}$$

where[c]

$$\tilde{\mathbf{G}} := \mathbf{g}_2\mathbf{g}_1^T - \mathbf{g}_1\mathbf{g}_2^T = \begin{bmatrix} \mathbf{g}_2 & \mathbf{g}_1 \end{bmatrix}\cdot\begin{bmatrix} \mathbf{g}_1^T \\ -\mathbf{g}_2^T \end{bmatrix} \in \mathbb{C}^{N\times N}. \tag{3.12}$$

With $\mathbf{g}_i$ given in (2.6), the matrix $\tilde{\mathbf{G}}$ is seen to contain the product channel impulse responses of the form $h_2(k)h_1(l) - h_1(k)h_2(l)$, $0 \le k,l \le L$, which are to be determined from (3.11). Toward a tractable procedure for computing $\tilde{\mathbf{G}}$, we observe

---

[c] We assume that the two channel impulse response vectors are linearly independent, for otherwise $\tilde{\mathbf{G}}$ is identically a zero matrix; this assumption holds whenever the environment is with sufficiently rich scattering.

that $\mathbf{R}_{\mathbf{y}}(1)$ in (3.11) is a weighted sum of N matrices of the form $\mathbf{J}^n\tilde{\mathbf{G}}\left(\mathbf{J}^T\right)^{N-n}$, in which the unknown $\tilde{\mathbf{G}}$ are pre, and post, multiplied by the known matrices $\mathbf{J}^n$ and $\left(\mathbf{J}^T\right)^{N-n}$. Based on this structural property, we can further rearrange (3.11) into a standard linear equation form. This is done via the next lemma.

Lemma 3.1 [19, p-255]: The matrix equation $\sum_{k=1}^{K}\mathbf{A}_k\mathbf{X}\mathbf{B}_k = \mathbf{C}$ can be equivalently expressed as $\left[\sum_{k=1}^{K}\mathbf{B}_k^T \otimes \mathbf{A}_k\right]vec(\mathbf{X}) = vec(\mathbf{C})$. □

Based on Lemma 3.1, we can immediately rewrite (3.11) as

$$\left[\sum_{n=0}^{N-1} p(n)^2\mathbf{J}^{N-n} \otimes \mathbf{J}^n\right]vec(\tilde{\mathbf{G}}) = vec(\mathbf{R}_{\mathbf{y}}(1)). \tag{3.13}$$

By definitions of the Kronecker product and $\mathbf{J}$ in (3.7), equation (3.13) turns out to be

$$\underbrace{\begin{bmatrix} p(0)^2\mathbf{I}_N & p(1)^2\mathbf{J} & \cdots & p(N-2)^2\mathbf{J}^{N-2} & p(N-1)^2\mathbf{J}^{N-1} \\ p(N-1)^2\mathbf{J}^{N-1} & p(0)^2\mathbf{I}_N & \cdots & p(N-3)^2\mathbf{J}^{N-3} & p(N-2)^2\mathbf{J}^{N-2} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ p(2)^2\mathbf{J}^2 & p(3)^2\mathbf{J}^3 & \cdots & p(0)^2\mathbf{I}_N & p(1)^2\mathbf{J} \\ p(1)^2\mathbf{J} & p(2)^2\mathbf{J}^2 & \cdots & p(N-1)^2\mathbf{J}^{N-1} & p(0)^2\mathbf{I}_N \end{bmatrix}}_{:=\mathbf{Q}} vec(\tilde{\mathbf{G}}) = vec(\mathbf{R}_{\mathbf{y}}(1)) \tag{3.14}$$

The $N^2 \times N^2$ real-valued matrix $\mathbf{Q}$ defined in (3.14), which is characterized by the N circulant matrices $\left\{p(0)^2\mathbf{I}_N, p(1)^2\mathbf{J}, \cdots, p(N-1)^2\mathbf{J}^{N-1}\right\}$ on the top row block, is block circulant with circulant blocks (BCCB) [12, p-184]. Equation (3.14) forms the basis of the proposed approach.

B. Identification of Channel Impulse Response

Assume that $vec\left(\tilde{\mathbf{G}}\right)$, and hence the matrix $\tilde{\mathbf{G}}$, can be uniquely recovered from the linear equation (3.14); the uniqueness condition and the computational issue will be investigated in the next section. We then collect the product unknowns $h_2(k)h_1(l) - h_1(k)h_2(l)$, $0 \le k,l \le L$, to form the following $(L+1)\times(L+1)$ matrix

$$\mathbf{H} \coloneqq \left[\mathbf{H}_{k,l}\right]_{0\le k,l\le L}, \text{ where } \mathbf{H}_{k,l} = h_2(k)h_1(l) - h_1(k)h_2(l). \tag{3.15}$$

Observe that the matrix $\mathbf{H}$ is of rank two, and can be factorized as

$$\mathbf{H} = \mathbf{h}_2\mathbf{h}_1^T - \mathbf{h}_1\mathbf{h}_2^T = \begin{bmatrix}\mathbf{h}_1 & \mathbf{h}_2\end{bmatrix}\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\mathbf{h}_1^T\\ \mathbf{h}_2^T\end{bmatrix}, \tag{3.16}$$

where

$$\mathbf{h}_i \coloneqq \begin{bmatrix}h_i(0) & h_i(1) & \cdots & h_i(L)\end{bmatrix}^T \in \mathbb{C}^{L+1},\ 1 \le i \le 2, \tag{3.17}$$

is the desired channel impulse response vectors. Based on (3.16), the channels can thus be identified, up to a 2×2 complex matrix of the form

$$\mathbf{U} = \begin{bmatrix}a & b\\ c & d\end{bmatrix}, \text{ with } ad - bc = 1, \tag{3.18}$$

by computing the two dominant left singular vectors associated with $\mathbf{H}$; the inherent matrix ambiguity must satisfy (3.18) since, for any vector pair of the form $\bar{\mathbf{h}} = \begin{bmatrix}\mathbf{h}_1 & \mathbf{h}_2\end{bmatrix}\mathbf{U}$ with $\mathbf{U} \in \mathbb{C}^{2\times 2}$, we have

$$\bar{\mathbf{h}}\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\bar{\mathbf{h}}^T = \mathbf{h}_2\mathbf{h}_1^T - \mathbf{h}_1\mathbf{h}_2^T \tag{3.19}$$

whenever $\mathbf{U}$ verifies (3.18). We note that a similar matrix outer-product based

approach for blind channel estimation is also adopted in [13], [21], [24], [37].

C. On Ambiguity Removal

The matrix ambiguity (3.18) can be resolved through insertion of additional pilot symbols. To see this, let $\begin{bmatrix}\bar{\mathbf{h}}_1 & \bar{\mathbf{h}}_2\end{bmatrix}$ be a dominant left singular vector pair associated with the rank-two matrix $\mathbf{H}$ defined in (3.15). Then we have $\begin{bmatrix}\bar{\mathbf{h}}_1 & \bar{\mathbf{h}}_2\end{bmatrix} = \begin{bmatrix}\mathbf{h}_1 & \mathbf{h}_2\end{bmatrix}\mathbf{U}$, with $\mathbf{U} \in \mathbb{C}^{2\times 2}$ fulfilling (3.18); this implies

$$\begin{bmatrix}\mathbf{h}_1 & \mathbf{h}_2\end{bmatrix} = \begin{bmatrix}\bar{\mathbf{h}}_1 & \bar{\mathbf{h}}_2\end{bmatrix}\mathbf{U}^{-1} = \begin{bmatrix}\bar{\mathbf{h}}_1 & \bar{\mathbf{h}}_2\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix}^{-1} = \begin{bmatrix}\bar{\mathbf{h}}_1 & \bar{\mathbf{h}}_2\end{bmatrix}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}. \tag{3.20}$$

Since both $\mathbf{G}_1$ and $\mathbf{G}_2$ are circulant, the first output branch (2.3), at some $k = k_0$, can be alternatively expressed as

$$\mathbf{y}_{k_0} = \mathbf{C}_{k_0}\mathbf{g}_1 + \mathbf{C}_{k_0+1}\mathbf{g}_2 + \mathbf{v}_{k_0}, \tag{3.21}$$

where $\mathbf{g}_i$ $(i = 1, 2)$ is the zero-padded channel impulse response as in (2.6), and $\mathbf{C}_l \in \mathbb{C}^{N\times N}$ is circulant with the precoded symbol vector $\mathbf{x}_l$ as the first column, $l = k_0, k_0 + 1$. Let us write

$$\mathbf{g}_i = \begin{bmatrix}\mathbf{h}_i^T & \mathbf{0}_{1\times(N-L-1)}\end{bmatrix}^T, \ 1 \le i \le 2, \tag{3.22}$$

where $\mathbf{h}_i$ is the desired channel impulse response vector defined in (3.17). With (3.22), equation (3.21) is then reduced to

$$\mathbf{y}_{k_0} = \tilde{\mathbf{C}}_{k_0}\mathbf{h}_1 + \tilde{\mathbf{C}}_{k_0+1}\mathbf{h}_2 + \mathbf{v}_{k_0}, \tag{3.23}$$

where $\tilde{\mathbf{C}}_l \in \mathbb{C}^{N\times(L+1)}$ contains the first $L+1$ columns of $\mathbf{C}_l$. With (3.20), we can

write (3.23) in terms of the scalar ambiguities as

$$\begin{aligned}\mathbf{y}_{k_0} &= \tilde{\mathbf{C}}_{k_0}\left(d\bar{\mathbf{h}}_1 - c\bar{\mathbf{h}}_2\right) + \tilde{\mathbf{C}}_{k_0+1}\left(-b\bar{\mathbf{h}}_1 + a\bar{\mathbf{h}}_2\right) + \mathbf{v}_{k_0} \\ &= \underbrace{\left[\tilde{\mathbf{C}}_{k_0}\bar{\mathbf{h}}_1 \quad -\tilde{\mathbf{C}}_{k_0}\bar{\mathbf{h}}_2 \quad -\tilde{\mathbf{C}}_{k_0+1}\bar{\mathbf{h}}_1 \quad \tilde{\mathbf{C}}_{k_0+1}\bar{\mathbf{h}}_2\right]}_{:=\mathbf{T}}\left[d \quad c \quad b \quad a\right]^T + \mathbf{v}_{k_0}.\end{aligned} \tag{3.24}$$

It is noted that, subject to the constraint $ad - bc = 1$, there are only three independent unknowns in (3.24). One can just solve for, say (b,c,d), from (3.24) and then determine a via the nonlinear equation $a = (1+bc)/d$; this, however, would be more prone to error propagation. Hence we propose to instead compute (a,b,c,d) all at once from (3.24). Toward this end, pilot symbols should be appropriately inserted to produce *at least* four training components in $\mathbf{y}_{k_0}$. We observe that each column of $\mathbf{T}$ in (3.24) is a linear combination of $L+1$ circularly shifted symbol vector $\mathbf{x}_l$ for some $l \in \{k_0, k_0+1\}$. The cyclicity structural constraint implies at least $L+4$ pilot symbols are needed in both $\mathbf{x}_l$. One plausible placement, in particular, is to insert four (and L, respectively) consecutive pilots at the head (and tail) of $\mathbf{x}_l$, $l = k_0, k_0+1$; in this way, the first four components in $\mathbf{y}_{k_0}$, denoted by $\mathbf{y}_t$, then act as training data and the scalar unknowns are estimated via

$$\left[\hat{d} \quad \hat{c} \quad \hat{b} \quad \hat{a}\right]^T = \mathbf{T}_t^{-1}\mathbf{y}_t, \tag{3.25}$$

where $\mathbf{T}_t \in \mathbb{C}^{4\times 4}$ contains the first four rows of $\mathbf{T}$.Hence, even though the proposed blind method reduces the number of unknown channel parameters from $2L+2$ to three, no less than $2L+8$ pilot symbols are nonetheless required for ambiguity removal. This is due to the non-redundant precoding based channel

TABLE I
ALGORITHM SUMMARY

1) Estimate the conjugate cross-correlation matrix $R_y(1)$ via

$$\hat{R}_y(1) := \frac{1}{K}\sum_{i=1}^{K} y(i)y(i+1)^T$$

where K is the number of temporal received symbol block pairs.

2) Form the matrix equation (3.14), and select a precoder $p(n)$ with which the matrix Q is nonsingular.

3) Compute the product channel coefficients based on (4.3) or (4.8).

4) Form the rank-two matrix H as in (3.15). The channel impulse response pair is then computed, up to a 2×2 matrix ambiguity, as the two dominant left singular vectors associated with H.

5) Remove the matrix ambiguity by inserting extra training symbols as suggested in (3.25).

estimation formulation as well as the circulant signal structure (the proposed channel estimation procedures are outlined in Table I).

Chapter 4
Identifiability and Product Unknowns Computation

This section first specifies the channel identifiability condition, and then introduces two methods for computing the product channel coefficients. The presented results also lay the foundation for further investigating the optimal precoder design problem.

A. Channel Identifiability

From the previous discussions, it is easy to see that the channel can be identified if $vec\left(\tilde{\mathbf{G}}\right)$ is uniquely determined from (3.14): this is true if the matrix $\mathbf{Q}$ is nonsingular. By exploiting the BCCB property of $\mathbf{Q}$, the following theorem explicitly shows the associated eigenvalues, and in turn specifies the condition for $\mathbf{Q}$ to be nonsingular. Roughly speaking, if we define the vector

$$\mathbf{p} \coloneqq [p(0)^2 \quad p(1)^2 \quad \cdots \quad p(N-1)^2]^T \in \mathbb{R}^N, \tag{4.1}$$

then the N^2 eigenvalues of $\mathbf{Q}$ are completely determined by the N eigenvalues associated with the $N \times N$ circulant matrix with $\mathbf{p}^T$ as the first row (the proof of theorem is given in Appendix A).

Theorem 4.1: Let $\mathbf{F}$ be the $N \times N$ FFT matrix; also, associated with the vector $\mathbf{p}$ in (4.1) we define the polynomial

$$\mathbf{p}(z) \coloneqq p(0)^2 + p(1)^2 z^{-1} + \cdots + p(N-1)^2 z^{-(N-1)}. \tag{4.2}$$

Then the N^2 eigenvalues of the matrix $\mathbf{Q}$ defined in (3.14) are given by the N replicas of the N-tuple $\left\{\mathbf{p}(1), \mathbf{p}(\omega), \cdots, \mathbf{p}(\omega^{N-1})\right\}$.

□

Theorem 4.1 shows that channel identifiability is guaranteed whenever $\mathbf{p}(\omega^n) \neq 0$ for all $0 \leq n \leq N-1$; this condition is quite mild and can hold for almost all choices of $p(n)$. We should note that the significance of Theorem 4.1 is far above just characterizing a sufficient condition for unique channel recovery. It moreover specifies the eigenvalues associated with the matrix $\mathbf{Q}$: this result will be exploited for selecting $p(n)$ to improve the reliability of channel estimate against the finite-sample estimation error (see Section V).

B. Computation of $vec\left(\tilde{\mathbf{G}}\right)$

A crucial step for implementing the proposed channel estimation scheme is the computation of the product channel coefficient vector $vec\left(\tilde{\mathbf{G}}\right)$ based on (3.14). In what follows we propose two methods for fulfilling this task.

i) Direct Matrix Inversion: An immediate approach to solving (3.14) is through direct matrix inversion so that

$$vec\left(\tilde{\mathbf{G}}\right) = \mathbf{Q}^{-1} vec\left(\mathbf{R}_{\mathbf{y}}(1)\right). \tag{4.3}$$

Observe from (3.14) that $\mathbf{Q}$ is BCCB and is characterized by the particular set of circulant matrices $\left\{p(0)^2\mathbf{I}_N, p(1)^2\mathbf{J}, \cdots, p(N-1)^2\mathbf{J}^{N-1}\right\}$. This appealing structure allows for a potentially low-complexity implementation via FFT operations. In

Appendix B we derive a simple closed-form expression of $\mathbf{Q}^{-1}$ based on which this figure of merit is justified.

ii) Solution via Zero Entry Removal in (3.14): It is noted from (2.6) that, for $1 \le i \le 2$, the vector $\mathbf{g}_i$ contains $L+1$ channel impulse response $h_i(n)$, $0 \le n \le L$, followed by $N-L-1$ trailing zeros. As a result, the $N^2 \times N^2$ matrix $\tilde{\mathbf{G}}(= \mathbf{g}_2\mathbf{g}_1^T - \mathbf{g}_1\mathbf{g}_2^T)$, and hence the associated vectorized representation $vec(\tilde{\mathbf{G}})$, has actually $(L+1)^2$ nonzero product unknowns. By removing the zero entries in $vec(\tilde{\mathbf{G}})$, and the corresponding indexed columns of the matrix $\mathbf{Q}$, equation (3.14) can be simplified to a set of N^2 scalar equations in $(L+1)^2$ unknowns. Indeed, with $\mathbf{g}_i$ defined in (3.22), we have

$$\mathbf{g}_i\mathbf{g}_{\bar{i}}^H = \begin{bmatrix} \mathbf{h}_i\mathbf{h}_{\bar{i}}^H & \mathbf{0}_{(L+1)\times(N-L-1)} \\ \mathbf{0}_{(N-L-1)\times(L+1)} & \mathbf{0}_{N-L-1} \end{bmatrix}, \quad 1 \le i, \bar{i} \le 2 \tag{4.4}$$

and hence

$$\tilde{\mathbf{G}} = \mathbf{g}_2\mathbf{g}_1^T - \mathbf{g}_1\mathbf{g}_2^T = \begin{bmatrix} \mathbf{H} & \mathbf{0}_{(L+1)\times(N-L-1)} \\ \mathbf{0}_{(N-L-1)\times(L+1)} & \mathbf{0}_{N-L-1} \end{bmatrix}, \tag{4.5}$$

where $\mathbf{H}$ is defined in (3.15). Based on (4.5) and by definition of the $vec(\cdot)$ operation, equation (3.14) can be shown (after some direct manipulations) to be reduced into

$$\underbrace{\mathbf{Q}\mathbf{J}_1\left(\mathbf{I}_{L+1} \otimes \mathbf{J}_2\right)}_{:=\tilde{\mathbf{Q}}} vec(\mathbf{H}) = vec(\mathbf{R}_{\mathbf{y}}(1)) \tag{4.6}$$

in which

$$\mathbf{J}_1 = \begin{bmatrix} \mathbf{I}_{N(L+1)} \\ \mathbf{0}_{N(N-L-1)\times N(L+1)} \end{bmatrix} \in \mathbb{R}^{N^2 \times N(L+1)} \quad \text{and} \quad \mathbf{J}_2 := \begin{bmatrix} \mathbf{I}_{L+1} \\ \mathbf{0}_{(N-L-1)\times(L+1)} \end{bmatrix} \in \mathbb{R}^{N\times(L+1)}.$$

(4.7)

The matrix $\tilde{\mathbf{Q}} \in \mathbb{R}^{N^2 \times (L+1)^2}$ in (4.7) is obtained by deleting $N^2 - (L+1)^2$ columns from $\mathbf{Q}$. It is thus of full column rank whenever $\mathbf{Q}$ is nonsingular and, if so, the product channel coefficients can be computed via

$$vec(\mathbf{H}) = \left(\tilde{\mathbf{Q}}^T \tilde{\mathbf{Q}}\right)^{-1} \tilde{\mathbf{Q}}^T vec(\mathbf{R}_{\mathbf{y}}(1)) . \tag{4.8}$$

Compared with the direct matrix inversion method (4.3), the solution (4.8) can yield better estimation accuracy at the expense of computational complexity (see Appendix B for complexity evaluation). Based on (4.3) and (4.8), the selection of precoder $p(n)$ for optimal numerical robustness is discussed next.

Chapter 4

Optimal Precoder Design

If the conjugate cross-correlation matrix $\mathbf{R}_{\mathbf{y}}(1)$ is perfectly obtained, both solutions (4.3) and (4.8) are exact. In practice, however, only a finite number of data samples can be used for estimating $\mathbf{R}_{\mathbf{y}}(1)$; equations (3.14) and (4.6) should be accordingly modified as

$$vec(\hat{\mathbf{R}}_{\mathbf{y}}(1)) = \mathbf{Q}vec(\tilde{\mathbf{G}}) + \mathbf{w}\,, \tag{5.1}$$

and

$$vec(\hat{\mathbf{R}}_{\mathbf{y}}(1)) = \tilde{\mathbf{Q}}vec(\mathbf{H}) + \mathbf{w}\,, \tag{5.2}$$

where $\hat{\mathbf{R}}_{\mathbf{y}}(1)$ is an estimate of $\mathbf{R}_{\mathbf{y}}(1)$ and $\mathbf{w}$ accounts for the data mismatch due to finite-sample estimation. Given the error-corrupted $\hat{\mathbf{R}}_{\mathbf{y}}(1)$, it is impossible to recover the actual product channel coefficients. Instead, with (5.1) and (5.2), the estimated solutions are respectively

$$vec(\hat{\tilde{\mathbf{G}}}) \coloneqq \mathbf{Q}^{-1}vec(\hat{\mathbf{R}}_{\mathbf{y}}(1)) = vec(\tilde{\mathbf{G}}) + \mathbf{Q}^{-1}\mathbf{w} \tag{5.3}$$

and

$$vec\left(\hat{\mathbf{H}}\right) \coloneqq \left(\tilde{\mathbf{Q}}^T\tilde{\mathbf{Q}}\right)^{-1}\tilde{\mathbf{Q}}^T vec(\hat{\mathbf{R}}_{\mathbf{y}}(1)) = vec\left(\mathbf{H}\right) + \left(\tilde{\mathbf{Q}}^T\tilde{\mathbf{Q}}\right)^{-1}\tilde{\mathbf{Q}}^T\mathbf{w}\,. \tag{5.4}$$

In what follows, we consider two different modeling schemes of $\mathbf{w}$, and propose the associated optimal criteria for designing $p(n)$ against imperfect data estimation.

A. Minimal Worst-Case Sensitivity to Error Perturbation

We will first treat $\mathbf{w}$ as an unknown "deterministic" perturbation since the statistical property of the data estimation error is in general difficult to characterize. From this standpoint, typical solution robustness measures for (5.3) and (5.4) are the *condition numbers* of the matrices $\mathbf{Q}$ and $\tilde{\mathbf{Q}}$, respectively (see, e.g., [18] and [23]). Small $\kappa(\mathbf{Q})$ and $\kappa(\tilde{\mathbf{Q}})$, in particular, are known to ensure small worst-case sensitivity of the error-perturbed solution to data mismatch [18, p-338]. Since both $\mathbf{Q}$ and $\tilde{\mathbf{Q}}$ depend entirely on $p(n)$, a natural approach to improving the channel estimation accuracy is to choose $p(n)$ so that both $\kappa(\mathbf{Q})$ and $\kappa(\tilde{\mathbf{Q}})$ are kept as small as possible. This type of optimization problem would seem formidable to tackle since the condition number of a matrix is in general a highly nonlinear function in the entries. Toward a tractable design formulation, we note the crucial fact: since $\tilde{\mathbf{Q}}$ contains a subset of columns of $\mathbf{Q}$ (see (4.6)), it follows [23, p-27]

$$\kappa(\tilde{\mathbf{Q}}) \le \kappa(\mathbf{Q}). \tag{5.5}$$

Inequality (5.5) suggests that, to jointly improve the accuracy of solutions (5.3) and (5.4), it is plausible to just minimize $\kappa(\mathbf{Q})$ because a small $\kappa(\mathbf{Q})$ will also guarantee $\kappa(\tilde{\mathbf{Q}})$ to be small. Such a design strategy, on the one hand, can bypass direct minimization of $\kappa(\tilde{\mathbf{Q}})$ which would appear rather intractable. More importantly, it will allow us to exploit the eigenvalue characteristics of the BCCB matrix $\mathbf{Q}$ (in Theorem 4.1) to analytically derive a solution, as is shown below. Hence, we specifically propose to minimize $\kappa(\mathbf{Q})$, subject to the following two

constraints

$$\sum_{n=0}^{N-1} p(n)^2 = N\,, \tag{5.6a}$$

and

$$\min\ p(n)^2 \geq \delta \quad \text{for some} \quad 0 < \delta < 1\,. \tag{5.6b}$$

The constraint (5.6a) normalizes the average transmit power within one block to unity, and the constraint (5.6b) imposes a minimal threshold on the floor power. In the context of single channel blind identification based on modulation-induced-cyclostatonarity, the two constraints have been used in [10], [24], and [37] for precoder design against the channel noise effect.

To derive the optimal solution, we shall first specify $\kappa(\mathbf{Q})$ in terms of the eigenvales of the matrix $\mathbf{Q}$. Since $\mathbf{Q}$ is BCCB, it can be factorized as $\mathbf{Q} = (\mathbf{F} \otimes \mathbf{F})\mathbf{\Lambda}\left(\mathbf{F}^H \otimes \mathbf{F}^H\right)$ for some diagonal $\mathbf{\Lambda}$ [12, p-181]. This then implies that $\mathbf{Q}$ is a normal matrix [18, p-100], as can be seen by

$$\begin{aligned} \mathbf{Q}^H\mathbf{Q} &= (\mathbf{F} \otimes \mathbf{F})\mathbf{\Lambda}^H \underbrace{\left(\mathbf{F}^H \otimes \mathbf{F}^H\right)(\mathbf{F} \otimes \mathbf{F})}_{=\mathbf{I}}\mathbf{\Lambda}\left(\mathbf{F}^H \otimes \mathbf{F}^H\right) \\ &= (\mathbf{F} \otimes \mathbf{F})\mathbf{\Lambda}^H\mathbf{\Lambda}\left(\mathbf{F}^H \otimes \mathbf{F}^H\right) = (\mathbf{F} \otimes \mathbf{F})\mathbf{\Lambda}\mathbf{\Lambda}^H\left(\mathbf{F}^H \otimes \mathbf{F}^H\right) = \mathbf{Q}\mathbf{Q}^H; \end{aligned} \tag{5.7}$$

in deriving (5.7), we have used the identity $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$ [19, p-244]. As $\mathbf{Q}$ is normal, it is known that [18, p-340]

$$\kappa(\mathbf{Q}) = \rho(\mathbf{Q})\,\rho\left(\mathbf{Q}^{-1}\right), \tag{5.8}$$

in which $\rho(\mathbf{M}) \coloneqq \max\{|\,\lambda\,|\colon \lambda\text{'s are eigenvalues of the matrix } \mathbf{M}\}$. Equation (5.8) links the condition number of $\mathbf{Q}$ with the extreme magnitudes of the associated

eigenvalues which, according to Theorem 4.1, are exactly the maximum and minimum among the N elements $\left\{|\,\mathbf{p}(1)\,|\,,|\mathbf{p}(\omega)\,|\,,\cdots,|\mathbf{p}(\omega^{N-1})\,|\right\}$, where $\mathbf{p}(z)$ is the polynomial defined in (4.2). More precisely, we have

$$\kappa(\mathbf{Q})=\frac{\max\left|\mathbf{p}(\omega^k)\right|}{\min\left|\mathbf{p}(\omega^k)\right|},\ \text{for}\ \ 0\le k\le N-1\,. \tag{5.9}$$

To find the minimal $\kappa(\mathbf{Q})$ based on (5.9), we shall further characterize $\mathbf{p}(\omega^k)$'s under the two constraints in (5.6). With (5.6a), it is easy to see from (4.2) that, for $k=0$,

$$\mathbf{p}(\omega^0)=\mathbf{p}(1)=\sum_{n=0}^{N-1}p(n)^2=N\,. \tag{5.10}$$

The following lemma provides an upper bound on $\left|\mathbf{p}(\omega^k)\right|$ for $1\le k\le N-1$; the result is crucial for deriving the minimal $\kappa(\mathbf{Q})$ (the proof of lemma is shown in Appendix C).

Lemma 5.1: For any $p(n)$ satisfying (5.6a) and (5.6b), we have

$$\left|\mathbf{p}(\omega^k)\right|\le N(1-\delta)\ \ \text{for all}\ \ 1\le k\le N-1\,. \tag{5.11}$$

□

With (5.9), (5.10), and (5.11), the minimal achievable $\kappa(\mathbf{Q})$, and the corresponding optimal $p(n)$, are shown in the following theorem.

Theorem 5.2: Under the constraints (5.6a) and (5.6b), the minimal condition number associated with the matrix $\mathbf{Q}$ is given by

$$\kappa_{\min}(\mathbf{Q}) = \frac{1}{1-\delta}, \tag{5.12}$$

which is attained by the following two-level solution: for a fixed but arbitrary $0 \le m \le N-1$,

$$p(m)^2 = N - (N-1)\delta \text{, and } \; p(n)^2 = \delta \;\; \text{for} \;\; n \neq m \,. \tag{5.13}$$

[Proof]: We claim that *i)* $\kappa(\mathbf{Q}) \ge 1/(1-\delta)$ for any $p(n)$ satisfying (5.6a) and (5.6b), and *ii)* equality is attained by the solution (5.13); the theorem thus follows. To show claim *i)*, it is noted from (5.9) and (5.10) that

$$\kappa(\mathbf{Q}) = \frac{\max\left|\mathbf{p}(\omega^k)\right|}{\min\left|\mathbf{p}(\omega^k)\right|} \ge \frac{\left|\mathbf{p}(\omega^0)\right|}{\min\left|\mathbf{p}(\omega^k)\right|} = \frac{N}{\min\left|\mathbf{p}(\omega^k)\right|}. \tag{5.14}$$

Also, (5.10) and (5.11) imply

$$\min\left|\mathbf{p}(\omega^k)\right| \le N(1-\delta) \text{, i.e., } \; \frac{1}{\min\left|\mathbf{p}(\omega^k)\right|} \ge \frac{1}{N(1-\delta)}. \tag{5.15}$$

Claim *i)* then follows immediately from (5.14) and (5.15). To prove claim *ii)*, it is noted that solution (5.13) yields, for any $k \neq 0$,

$$\begin{aligned} \mathbf{p}(\omega^k) = \sum_{n=0}^{N-1} p(n)^2 \omega^{-kn} &= \left\{N - (N-1)\delta\right\}\omega^{-km} + \delta \sum_{n \neq m} \omega^{-kn} \\ &= \left\{N(1-\delta)\right\}\omega^{-km} + \delta \sum_{n=0}^{N-1} \omega^{-kn} = \left\{N(1-\delta)\right\}\omega^{-km}, \end{aligned} \tag{5.16}$$

where the last equality follows since $\sum_{n=0}^{N-1} \omega^{-kn} = 0$ for any $k \neq 0$. Equations (5.10) and (5.16) show that, with solution (5.13), we have $\max\left|\mathbf{p}(\omega^k)\right| = \left|\mathbf{p}(\omega^0)\right| = N$ and $\min\left|\mathbf{p}(\omega^k)\right| = N(1-\delta)$, and hence

$$\kappa(\mathbf{Q}) = \frac{\max\left|\mathbf{p}(\omega^k)\right|}{\min\left|\mathbf{p}(\omega^k)\right|} = \frac{N}{N(1-\delta)} = \frac{1}{1-\delta}. \tag{5.17}$$

The proof is thus completed. □

Theorem 5.2 shows that $\kappa_{\min}(\mathbf{Q})$ depends entirely on the minimal power threshold δ, irrespective of the dimension of $\mathbf{Q}$ (and hence the symbol block length N). A small δ, in particular, is seen to yield small $\kappa_{\min}(\mathbf{Q})$ and thus improves the channel estimation accuracy.

B. Minimization of Mean Square Error

In this subsection, we alternatively formulate $\mathbf{w}$ as a zero-mean white noise vector with covariance matrix $\sigma_w^2\mathbf{I}$, and resort to the well-known minimum mean square error principle, see, e.g., [6], for constructing a solution. Although a theoretical justification of such statistical data error assumption is difficult to establish, our simulation study does confirm this tendency.

Since $\mathbf{w}$ is white, the mean square errors incurred by solutions (5.3) and (5.4) are, respectively,

$$E\left\|vec(\hat{\tilde{\mathbf{G}}}) - vec(\tilde{\mathbf{G}})\right\|^2 = \sigma_w^2 Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right] \tag{5.18}$$

and

$$E\left\|vec(\hat{\mathbf{H}}) - vec(\mathbf{H})\right\|^2 = \sigma_w^2 Tr\left[\left(\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}\right)^{-1}\right] \tag{5.19}$$

Toward an utmost reduction of the white noise effect, the precoder $p(n)$ should thus

be chosen to jointly minimize

$$Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right] \text{ and } Tr\left[\left(\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}\right)^{-1}\right], \tag{5.20}$$

subject to the constraints (5.6a) and (5.6b). Minimization of this type of cost functions has been considered in least-squares based channel estimation, e.g., [6] and [22, chap. 9], among others. The reported solution approach therein is via the following inequality: since both $\mathbf{Q}^H\mathbf{Q}$ and $\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}$ are positive definite, it follows

$$Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right] \geq \sum_i \left[\mathbf{Q}^H\mathbf{Q}\right]_{i,i}^{-1} \text{ and } Tr\left[\left(\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}\right)^{-1}\right] \geq \sum_i \left[\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}\right]_{i,i}^{-1}, \tag{5.21}$$

and equalities in (5.21) hold whenever $\mathbf{Q}^H\mathbf{Q}$ and $\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}$, respectively, are diagonal [28, p-1041]. If the power normalization equation (5.6a) is the only design concern, it is easy to check that the impulse sequence

$$p(m)^2 = N \text{ , and } p(n)^2 = 0 \text{ for } n \neq m, \tag{5.22}$$

where $0 \leq m \leq N-1$ is fixed but arbitrary, simultaneously diagonalizes $\mathbf{Q}^H\mathbf{Q}$ and $\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}$, and is thus the jointly minimizer. However, given the additional threshold power requirement (5.6b), one cannot rely on this principle for finding a solution since, subject to the BCCB structure of $\mathbf{Q}$ and $p(n)^2 > 0$, it is impossible to choose $p(n)$ to render both $\mathbf{Q}^H\mathbf{Q}$ and $\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}$ diagonal. In what follows we propose an alternative strategy to address the considered optimization problem. Our approach is grounded on a key fact shown in the next lemma, which directly establishes an inequality relation analogue to (5.5) regarding the two cost functions in (5.20) (the proof is given in Appendix D).

Lemma 5.3: Let $\mathbf{M}$ be a square nonsingular matrix, and $\bar{\mathbf{M}}$ be constructed from $\mathbf{M}$ by deleting an arbitrary subset of its columns. Then

$$Tr\left[\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\right] \le Tr\left[\left(\mathbf{M}^H\mathbf{M}\right)^{-1}\right]. \tag{5.23}$$

□

Lemma 5.3 asserts $Tr\left[\left(\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}\right)^{-1}\right]$ is upper bounded by $Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right]$. This thus suggests a suboptimal, but would be more simple and efficient, way of precoder design: we can simply choose $p(n)$ to minimize $Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right]$, since $Tr\left[\left(\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}\right)^{-1}\right]$ would in turn be kept small. The main advantage of the proposed design formulation, as expected, is that we can directly take profit of the BCCB property of $\mathbf{Q}$ to derive a closed-form solution. Indeed, since $Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right]$ is the sum of the eigenvalues associated with $\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}$ which, according to Theorem 4.1, are exactly the N replicas of the N-tuple $\left\{|\,\mathbf{p}(\omega^n)\,|^{-2}\right\}_{0\le n\le N-1}$, we have

$$Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right] = \sum_{k=0}^{N-1}\frac{N}{|\,\mathbf{p}(\omega^k)\,|^2}. \tag{5.24}$$

Equation (5.24) rewrites $Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right]$ in terms of $\mathbf{p}(\omega^n)$'s; we can then further exploit equation (5.10) and Lemma 5.1 to construct an optimal solution, as is shown in the next theorem.

Theorem 5.4: The optimal $p(n)$ minimizing $Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right]$, subject to constraints (5.6a) and (5.6b), is the two-level solution (5.13). The resultant minimal mean square error is

$$MSE_{\min} = \frac{\sigma_w^2}{N} + \frac{\sigma_w^2(N-1)}{N(1-\delta)^2}. \tag{5.25}$$

[Proof]: From (5.10), we have

$$Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right]=\sum_{k=0}^{N-1}\frac{N}{|\,\mathbf{p}(\omega^k)\,|^2}=\frac{N}{|\,\mathbf{p}(1)\,|^2}+\sum_{k=1}^{N-1}\frac{N}{|\,\mathbf{p}(\omega^k)\,|^2}=\frac{1}{N}+\sum_{k=1}^{N-1}\frac{N}{|\,\mathbf{p}(\omega^k)\,|^2}. \tag{5.26}$$

From Lemma 5.1, it follows

$$\frac{1}{\left|\mathbf{p}(\omega^k)\right|^2}\geq\frac{1}{N^2(1-\delta)^2},\ \ 1\leq k\leq N-1, \tag{5.27}$$

with equality attained by the two-level sequence (5.13) (this is easily seen from (5.16)). From (5.26) and (5.27), the minimal $Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right]$ is thus

$$Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right]_{\min}=\frac{1}{N}+\sum_{k=1}^{N-1}\frac{1}{N(1-\delta)^2}=\frac{1}{N}+\frac{N-1}{N(1-\delta)^2}. \tag{5.28}$$

Equation (5.25) follows directly from (5.28), and this thus proves the theorem. □

Recall that the impulse sequence (5.22) is optimal with regard to the power normalization constraint (5.6a). When an additional power threshold is imposed, it turns out that the best choice is the "impulse-like" two-level solution (5.13). With (5.25), the resultant $MSE_{\min}$ is seen to decrease whenever δ is decreased. Hence, a small δ not only limits solution sensitivity to deterministic error perturbation (as we have shown in the previous subsection), but also improves the estimation accuracy against white data estimation error. From the equalization point of view, it is however undesirable to keep δ unlimitedly small; this will be further discussed in the next section.

Remarks:

(a) From Theorems 5.2 and 5.4, it is somewhat surprising to see that, although the

objective functions $\kappa(\mathbf{Q})$ and $Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right]$ are quite different in nature, the respective minimizing solutions, under constraints (5.6a) and (5.6b), are the same the two-level form choice (5.13); this is due to the BCCB property of the matrix $\mathbf{Q}$.

(b) The two-level solution (5.13) minimizes both $\kappa(\mathbf{Q})$ and $Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right]$, but its optimality with respect to $\kappa(\tilde{\mathbf{Q}})$ and $Tr\left[\left(\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}\right)^{-1}\right]$ appears intractable to verify. Our simulation results seem to indicate that it is indeed the minimizing solution.

(c) Since $\kappa(\tilde{\mathbf{Q}}) \le \kappa(\mathbf{Q})$ and $Tr\left[\left(\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}\right)^{-1}\right] \le Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right]$, solution (5.4) can yield better estimation accuracy than (5.3); numerical simulations (see Simulation 2) also evidence this tendency.

(d) The optimal solution (5.13) does not depend on the index m at which the peak power occurs: any $0 \le m \le N-1$ allows for an utmost mitigation for data estimation error. However, since the trailing components in each symbol block will be duplicated as CP, the peak power in (5.13) should not be located within the corresponding index region so as to conserve the power resource.

(e) In the study of single channel blind identification via modulation-induced-cyclostationarity, the two-level sequence (5.13) is shown to be optimal for mitigating the channel noise effect for the serial transmission case [10], [24], and also for the FDE based block transmission [37]. □

Chapter 6

Equalization Aspect

Toward symbol recovery in FDE-STBC systems, one commonly used approach is via frequency-domain per-tone equalization [2], [15] based on (2.8), commonly in conjunction with linear ZF or MMSE criterion. In this section we resort to ZF-PEP analysis [35] for investigating the equalization performance regarding the optimal solution (5.13).

To proceed, based on (2.2), we shall first expand the linearly combined frequency-domain signal model (2.8) into

$$\underbrace{\mathbf{D}^H\begin{bmatrix}\mathbf{Y}_k\\ \mathbf{Y}_{k+1}^*\end{bmatrix}}_{:=\mathbf{Z}_{k,k+1}} = \begin{bmatrix}\tilde{\mathbf{D}} & \mathbf{0}_N\\ \mathbf{0}_N & \tilde{\mathbf{D}}\end{bmatrix}\begin{bmatrix}\mathbf{X}_k\\ \mathbf{X}_{k+1}\end{bmatrix} + \mathbf{D}^H\begin{bmatrix}\mathbf{V}_k\\ \mathbf{V}_{k+1}^*\end{bmatrix} = \underbrace{\begin{bmatrix}\tilde{\mathbf{D}} & \mathbf{0}_N\\ \mathbf{0}_N & \tilde{\mathbf{D}}\end{bmatrix}\begin{bmatrix}\mathbf{FP} & \mathbf{0}_N\\ \mathbf{0}_N & \mathbf{FP}\end{bmatrix}}_{:=\mathbf{\Phi}}\underbrace{\begin{bmatrix}\mathbf{s}_k\\ \mathbf{s}_{k+1}\end{bmatrix}}_{:=\mathbf{s}_{k,k+1}} + \underbrace{\mathbf{D}^H\begin{bmatrix}\mathbf{V}_k\\ \mathbf{V}_{k+1}^*\end{bmatrix}}_{:=\bar{\mathbf{V}}_{k,k+1}} \tag{6.1}$$

or by dropping the block index k and $k+1$ for notational simplicity,

$$\mathbf{Z} = \mathbf{\Phi}\mathbf{s} + \bar{\mathbf{V}}\,. \tag{6.2}$$

The PEP measures the probability that a symbol block $\mathbf{s}$ is transmitted but another $\tilde{\mathbf{s}} \neq \mathbf{s}$ is detected. Given the channel realizations $\mathbf{h}_1$ and $\mathbf{h}_2$, the conditional PEP is by definition given by

$$\Pr\left[\mathbf{s} \to \tilde{\mathbf{s}} \mid \mathbf{h}_1, \mathbf{h}_2\right] = \Pr\left[\|\tilde{\mathbf{s}} - \hat{\mathbf{s}}\| < \|\mathbf{s} - \hat{\mathbf{s}}\| \mid \mathbf{h}_1, \mathbf{h}_2\right], \tag{6.3}$$

where $\hat{\mathbf{s}}$ is the estimate of $\mathbf{s}$ under the ZF metric and, from (6.2), is given by

$$\hat{\mathbf{s}} := \mathbf{\Phi}^{-1}\mathbf{Z} = \mathbf{s} + \mathbf{\Phi}^{-1}\bar{\mathbf{V}}\,. \tag{6.4}$$

By following the procedures as in [35] and define $\bar{d} := \|\mathbf{s} - \tilde{\mathbf{s}}\|$, the conditional PEP in (6.3) can be upper bounded by

$$\Pr\left[\mathbf{s} \to \tilde{\mathbf{s}} \mid \mathbf{h}_1, \mathbf{h}_2\right] \le Q\left(\overline{d}\left(\sqrt{4\sigma_v^2 \left\|\mathbf{P}^{-1}\right\|_F^2 \left\|\tilde{\mathbf{D}}^{-1}\right\|_F^2}\right)^{-1}\right) = Q\left(\overline{d}\left(\sqrt{4\sigma_v^2 \left(\sum_{n=0}^{N-1} p(n)^{-2}\right)\left\|\tilde{\mathbf{D}}^{-1}\right\|_F^2}\right)^{-1}\right), \quad (6.5)$$

where $Q(\cdot)$ denotes the Gaussian tail, and the equality in (6.5) follows directly from (2.1). For a given channel pair, and hence $\tilde{\mathbf{D}}$, the upper bound in (6.5) is minimized if the quantity $\sum_{n=0}^{N-1} p(n)^{-2}$ attains the minimum. Since, by the Cauchy's inequality,

$$\left(p(0)^{-2} + \cdots + p(N-1)^{-2}\right)\left(p(0)^2 + \cdots + p(N-1)^2\right) \ge (1 + \cdots + 1)^2 = N^2, \quad (6.6)$$

and $\sum_{n=0}^{N-1} p(n)^2 = N$ (cf. (5.6a)), we have $\left(p(0)^{-2} + \cdots + p(N-1)^{-2}\right) \ge N$, with equality holds if and only if $p(n) = 1$ for $0 \le n \le N-1$. This shows the equal power scheme is optimal from an equalization point of view. Any form of precoding induced power variation, therefore, will incur a loss in the decision performance. The precoder (5.13), however, turns out to be the worst-case choice, as can be seen from the following theorem (see Appendix E for a proof).

Theorem 6.1: For all $p(n)$ satisfying (5.6a) and (5.6b), the solution (5.13) maximizes the quantity $\sum_{n=0}^{N-1} p(n)^{-2}$, leading to

$$\max \sum_{n=0}^{N-1} p(n)^{-2} = \frac{N-1}{\delta} + \frac{1}{N-(N-1)\delta}. \quad (6.7)$$

□

Based on (6.7), simple manipulation shows the maximum value, when viewed as a function of δ, will increase as δ is decreased. As a result, a small δ, although improving channel estimation accuracy, will enlarge the PEP upper bound in (6.5), and hence bring potentially poor equalization performance. This thus imposes a tradeoff in selecting δ; our simulation study (see Simulation 5) indicates that $\delta = 0.7 \sim 0.8$ are the compromising choices.

Chapter 7

Simulation Results

This section uses several numerical simulations to illustrate the performance of the proposed method. The symbol block length and the channel order, respectively, are set to be $N = 32$ and $L = 8$; the inserted CP spans 8 symbol periods and the source constellation is QPSK. Unless otherwise stated, we will consider a block fading environment in which the channel taps, modeled as i.i.d. zero-mean unit-variance complex Gaussian random variables, remain constant over a burst of K symbol blocks and can vary independently between different bursts. The identification performance is measured by the normalized mean square error (NMSE), namely, $\text{NMSE} \coloneqq \frac{1}{2I}\sum_{l=1}^{2}\sum_{i=1}^{I}\left\|\hat{\mathbf{h}}_l^{(i)} - \mathbf{h}_l^{(i)}\right\|^2 \cdot \left\|\mathbf{h}_l^{(i)}\right\|^{-2}$, where $\mathbf{h}_l^{(i)}$ is the realization of the lth channel in the ith data packet, $\hat{\mathbf{h}}_l^{(i)}$ is the corresponding estimate, and I is the total number of trials. Throughout the simulations, the peak power index of the optimal precoder (5.13) is chosen to be $m = 0$; the signal-to-noise ratio (SNR) is defined as $\text{SNR} \coloneqq (E\left\|\mathbf{h}_1\right\|^2 + E\left\|\mathbf{h}_2\right\|^2)/2\sigma_v^2$. Simulations I~V investigate the intrinsic aspects pertaining to the proposed method, and we simply use the least-squares fit technique for matrix ambiguity removal, as is done in [4], [13], [21]; in Simulations I~IV, we set $I = 200$.

Simulation 1-Effectiveness of the Optimal Precoder (5.13): This simulation illustrates the effectiveness of the proposed optimal precoder (5.13). For SNR=10 dB and

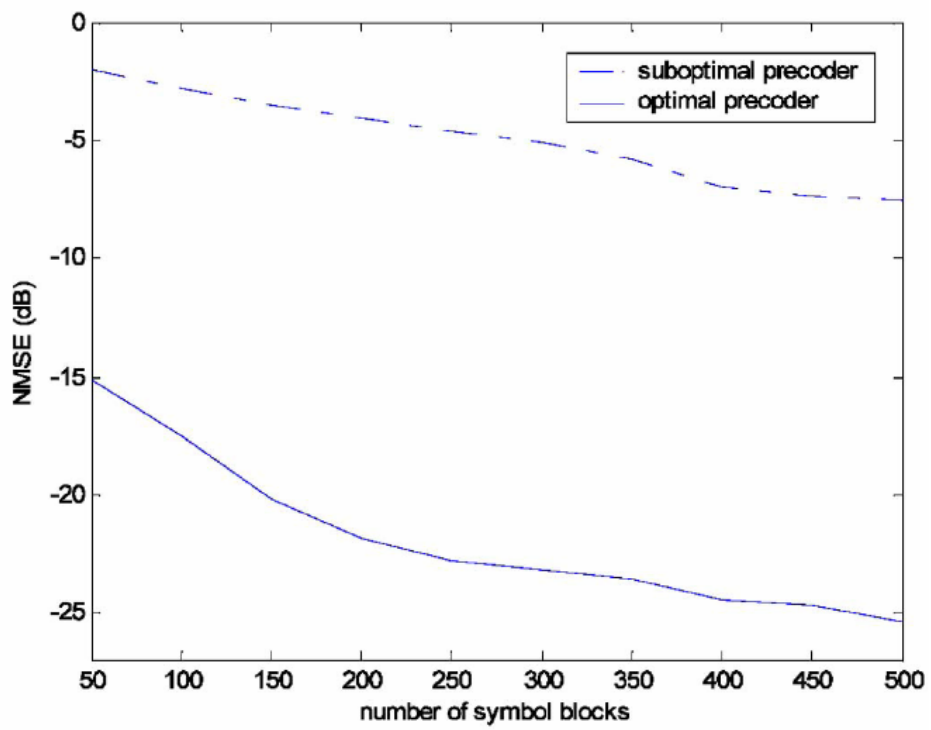

Fig. 2. Channel NMSE (optimal and suboptimal precoders).

$\delta = 0.6$, we consider the optimal sequence (5.13) and another sub-optimal choice given as $p(n)^2 = 0.6$ for $0 \le n \le 15$ and $p(n)^2 = 1.4$ for $16 \le n \le 31$. Figure 2 shows the computed NMSE with various numbers of symbol blocks K (the product channel coefficients are computed via (5.4)). It can be seen that the optimal solution (5.13) significantly improves the performance.

Simulation 2-Performance Comparison of Solutions (5.3) and (5.4): This simulation compares the estimation performance of solutions (5.3) and (5.4). Figure 3 shows the respective NMSE, versus number of symbol blocks, for three power thresholds δ: 0.3, 0.6, and 0.92 (SNR is fixed at 10 dB). The result shows that the performances of the two methods are very close for $\delta = 0.3$ and 0.6; however, solution (5.4) seems to yield smaller NMSE when $\delta = 0.92$. This is because, for $\delta = 0.3$ and 0.6, the associated condition number pair $\left(\kappa(\mathbf{Q}), \kappa(\tilde{\mathbf{Q}})\right)$ are $(1.4286, 1.1370)$ and

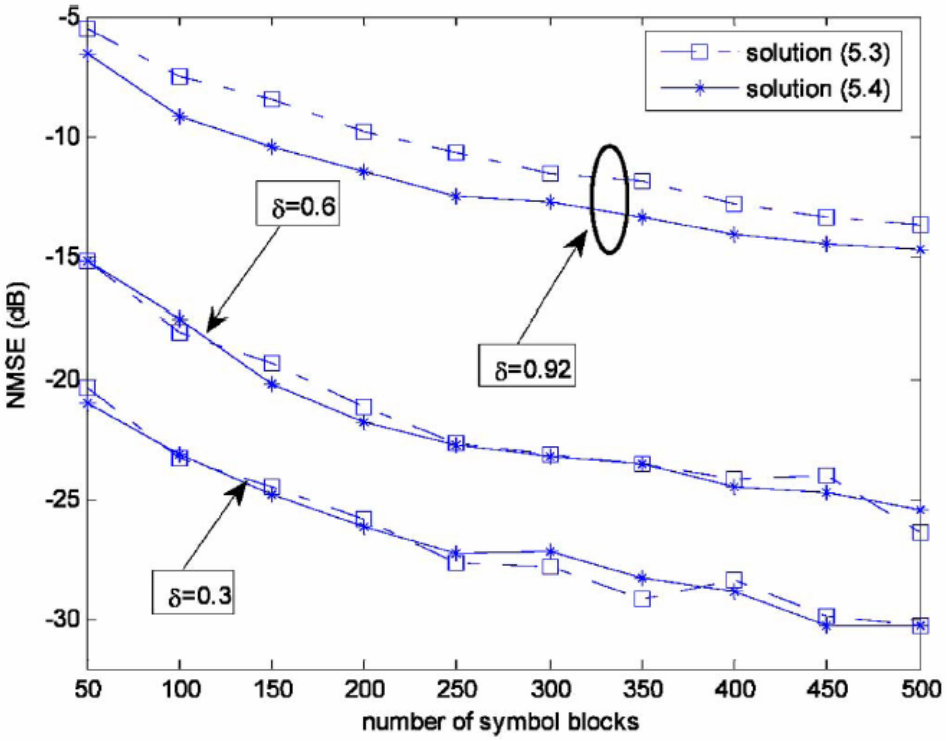

Fig. 3. Channels NMSE by two solutions.

$(2.5, 1.5737)$: both the two matrices $\mathbf{Q}$ and $\tilde{\mathbf{Q}}$ remain well conditioned, and can largely limit the error effect. However, for $\delta = 0.92$, we have $\left(\kappa(\mathbf{Q}), \kappa(\tilde{\mathbf{Q}})\right) = (12.5, 6.683)$: the matrix $\mathbf{Q}$ tends to be ill-conditioned, and solution (5.3) becomes more susceptible to data errors (solution (5.4) will be adopted in subsequent simulations).

Simulation 3-Robustness Against Channel Order Overestimation: This simulation tests the proposed method when channel order is overestimated. We consider two different levels of SNR: 0 dB and 15 dB. For the overestimated channel order $8 \le \hat{L} \le 15$, Figure 4 shows the respective computed NMSE ($K = 500$ and $\delta = 0.8$). It can be seen that the proposed method is quite robust with respect to channel order overestimation: the NMSE increment is only about 3 dB as $\hat{L}$ increases from 8 to 15.

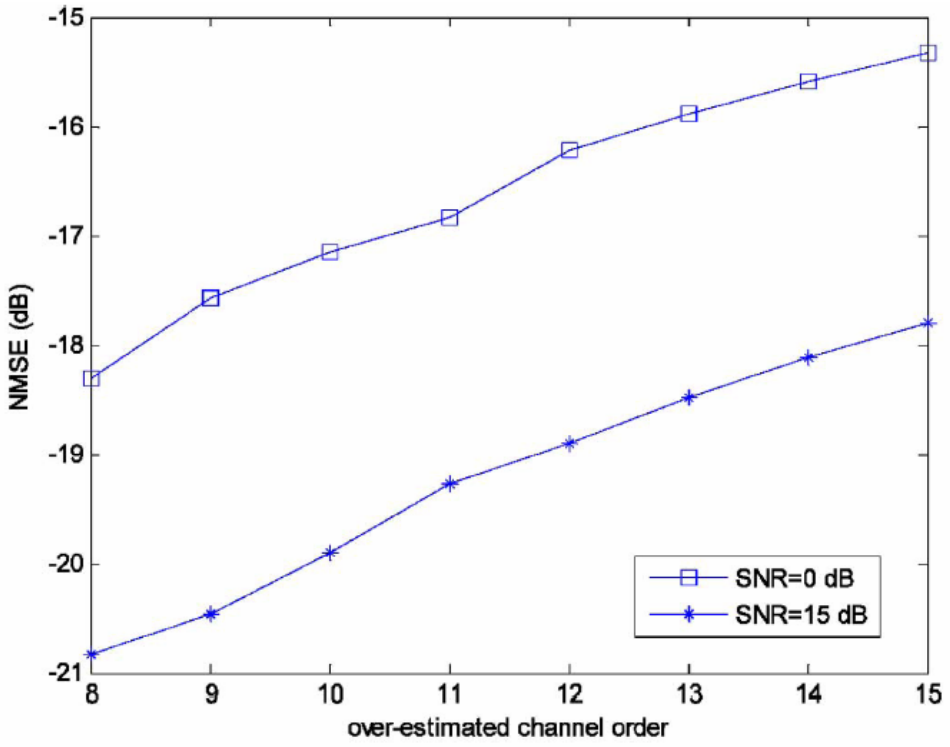

Fig. 4. NMSE in the presence of channel order overestimation.

Simulation 4-Estimation Performance Against Blind Subspace Method with Transmit Redundancy [40]: This simulation compares channel estimation performances of the proposed scheme with the identical-precoder subspace method [40, p-1218], in which FFT precoding matrix is adopted to convert the multi-carrier scheme into single-carrier FDE-STBC systems considered in this paper. To implement the algorithm in [40], the last 8 entries in each symbol block are set to be zero; this introduces the minimal amount of transmit redundancy for fulfilling the associated channel identifiability condition (cf. [40, p-1218]). For fixed SNR= 10 dB, Figure 5 shows the computed NMSE versus number of symbol blocks; the proposed method, depicted with solid lines, is implemented with various choices of δ. We can first see from the figure that the performance of the proposed method is improved as δ decreases: this is because small δ results in small $\kappa(\mathbf{Q})$, and also reduces the mean square error incurred by

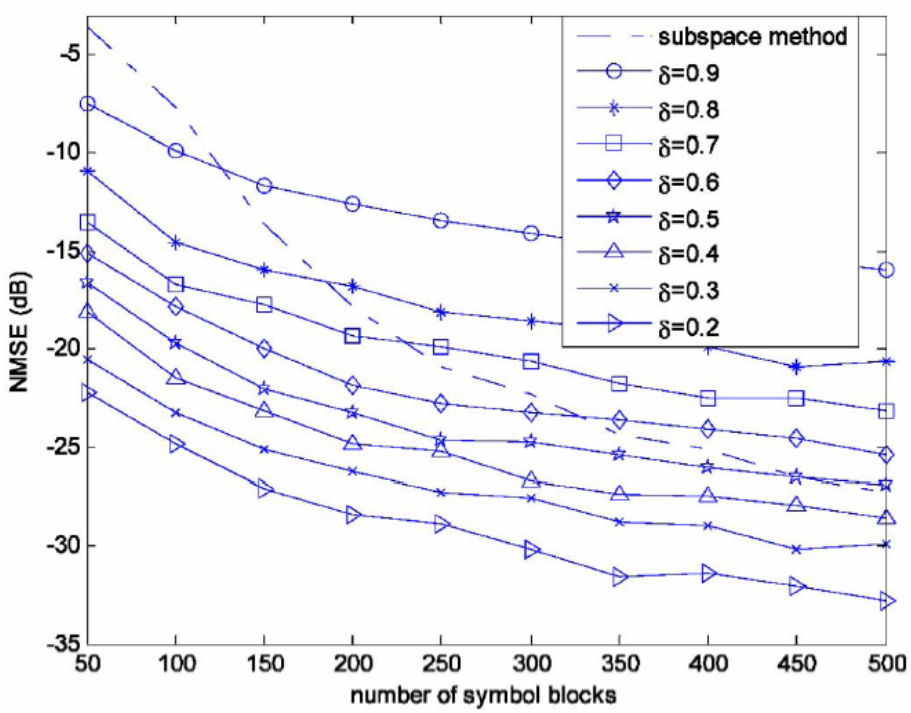

Fig. 5. NMSE of two methods at different numbers of symbol blocks.

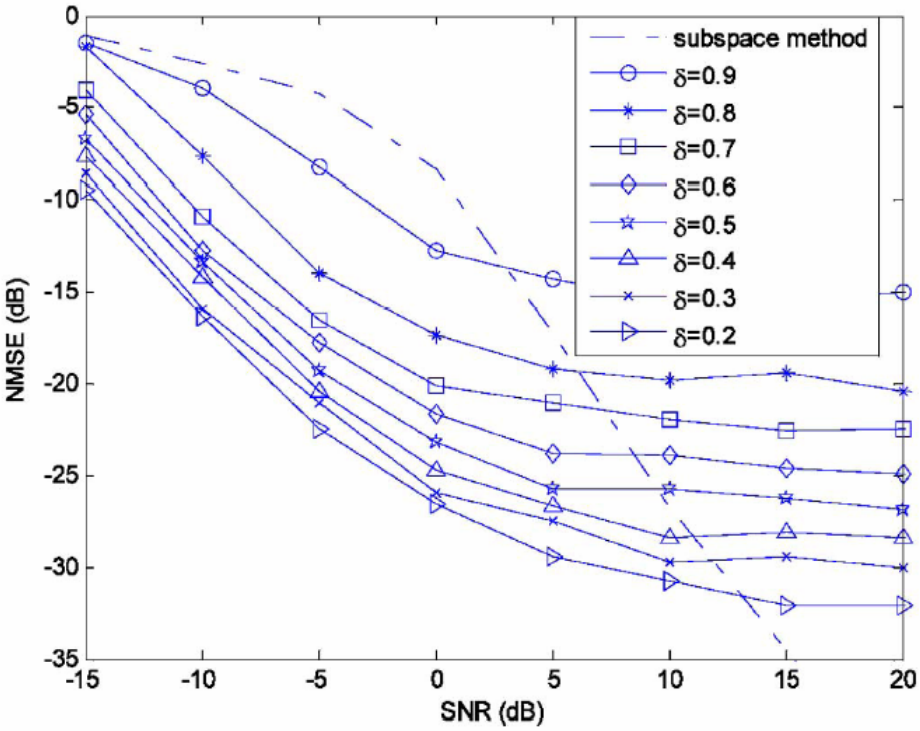

Fig. 6. NMSE of two methods at different SNR levels.

white noise perturbation. Compared with the subspace method [40], the proposed approach can better track the channel with a small number of received data blocks.

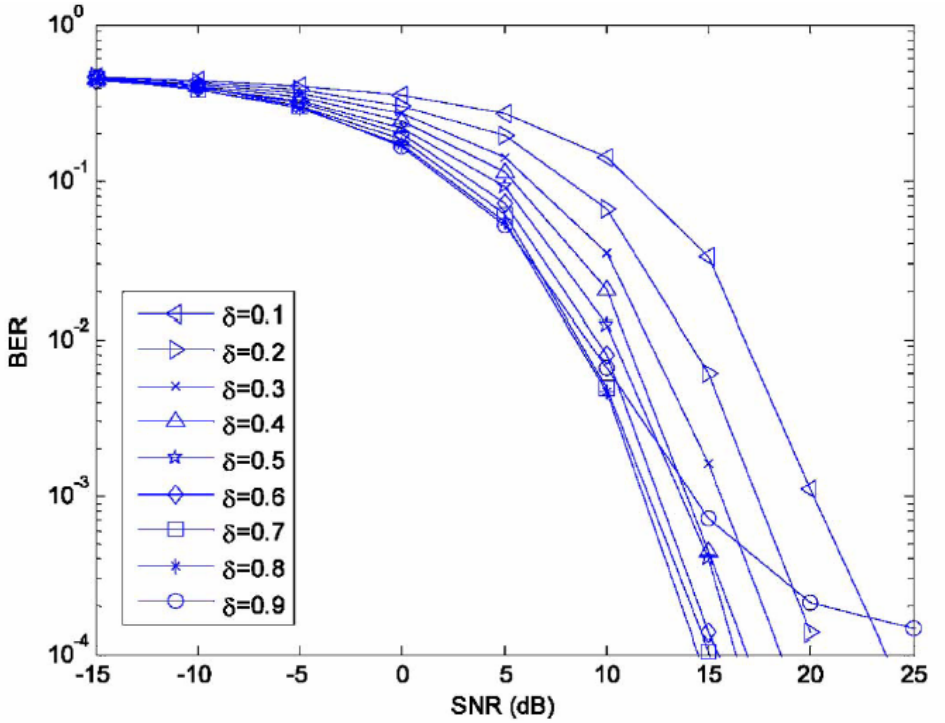

Fig. 7. BER performance of the proposed method at different minimal power threshold δ.

Figure 6 shows the NMSE of the two methods at different SNR levels ($K = 500$). The result shows that, in the medium-to-low SNR region, our method performs better even with the moderate choice $\delta = 0.7$. When SNR increases, the output NMSE of [40] exhibits a fast decay. This is not unexpected since the method [40] is "deterministic" in nature: it benefits from the finite-sample-convergence property, and can usually yield impressive estimation accuracy when SNR is high [31]. A similar tendency is also observed in [31, p-1942] when non-redundant diagonal precoding based identification is compared with the (deterministic) multi-channel subspace methods [29] and [38].

Simulation 5-On Selection of Power Threshold δ*:* This simulation considers the optimal precoder (5.13) and illustrates the impact of δ on equalization performance. Figure 7 shows the bit-error-rate (BER) curves for $0.1 \le \delta \le 0.9$; we set $K = 500$, $I = 1000$, and use frequency-domain ZF equalizer [15] for symbol recovery. It can

be seen that, although a large δ results in a less accurate channel estimate, the BER is however improved as δ increases from 0.1 to 0.8. This would reflect the ZF-PEP analysis given in Section VI: a large δ can on the other hand limit the power penalty and improve the equalization performance. However, if δ is too large $(\delta > 0.8)$, the channels will be poorly estimated, resulting in large decision error floor. Hence $\delta = 0.7 \sim 0.8$ seem to be the compromising choices, as far as equalization performance is concerned.

Simulation 6-Equalization Performance Comparison: In this simulation we compare the proposed method (implemented using the optimal precoder (5.13) and $\delta = 0.8$) with the blind identical-precoder subspace algorithm in [40] and the training based scheme [11] in terms of BER. To implement the method [11], pilot symbols (64 in total) are placed in the initial two blocks in each data burst and are optimally designed according to [11, p-730]. To fairly compare the three methods under a fixed spectral efficiency, we will similarly use 64 training symbols, placed also in the initial block pair per data burst, for ambiguity removal in the two blind approaches; for simplicity we just choose the optimal sequence reported in [11]. We note that, in the transmit-redundancy based blind scheme [40], 8 entries in each symbol block are zero-padded for facilitating channel identification: only 24 entries per symbol block can thus be used for carrying source data. To compensate for possible spectral efficiency loss, 20 elements out of which are loaded with 8-PSK symbols, whereas the remaining four are BPSK modulated: this maintains an overall data rate of 64 bits/block, as in the training scheme [11] and the proposed method (both with QPSK

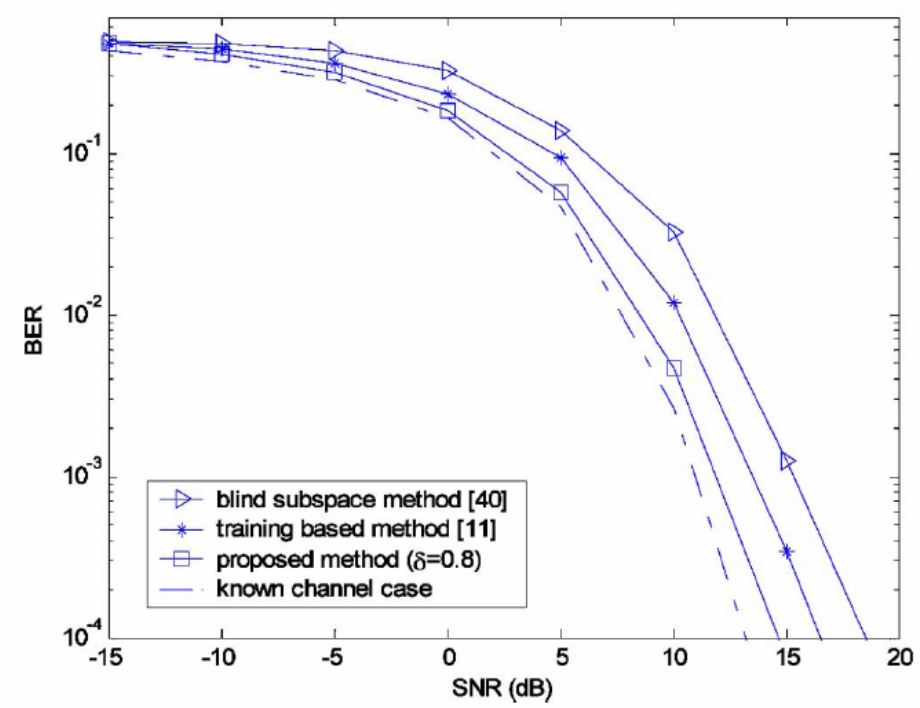

Fig. 8. BER performances of three methods (i.i.d. Gaussian channel model).

constellation). Figure 8 shows the BER results of the three methods at different SNR levels ($K = 500$ and $I = 1000$); the known channel case is also included as the benchmark performance. The proposed method, as we can see, leads to the lowest BER. The performance advantage over the training method [11] would come from the reduction of number of unknowns from $2(L+1)$ to three: this would further smoothen the noise effect and thus improve performance. In contrast with [40], our method yields about a 3~4 dB SNR gain; this benefits from the non-redundant transmission of the proposed scheme: for a target data rate one can otherwise use lower order constellations to buy more BER floor margin. Finally we observe that, compared with the known channel case, our method seems to incur no more than 1 dB penalty. We repeat the above experiment with the exponential delay power profile channel model [30]. Figure 9 shows the resultant BER curves, which are seen to exhibit a similar tendency as in the

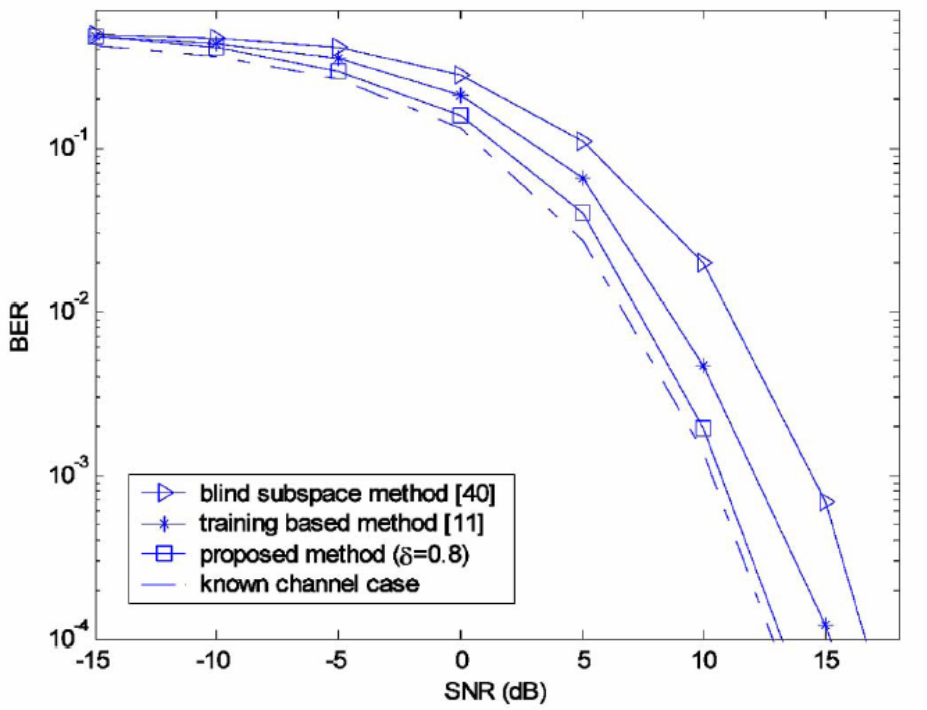

Fig. 9. BER performances of three methods (exponential delay-power profile channel model).

i.i.d. Gaussian channel case.

Simulation 7-Equalization Performance in Slowly Time-Varying Channels: This simulation compares the proposed method (with optimal precoder (5.13) and $\delta = 0.8$) against [11] and [40] in a slowly time-varying channel environment. We assume each channel tap varies according to the Jake's model with a maximal Doppler frequency of 52 Hz; this corresponds to a moving speed 3 m/sec and a carrier frequency of 5.2 GHz (the same simulation environment is considered in [40]). The number of symbol blocks in each data burst is set to be $K = 300$. To track the channel variation, in both blind methods the receive data statistics are adaptively updated based on the rectangular sliding windowing scheme suggested in [40, p-1218], with the window size set to be 150. Channel estimation is performed each time a new sub-burst of 50 symbol blocks are available. As in the previous simulation, the

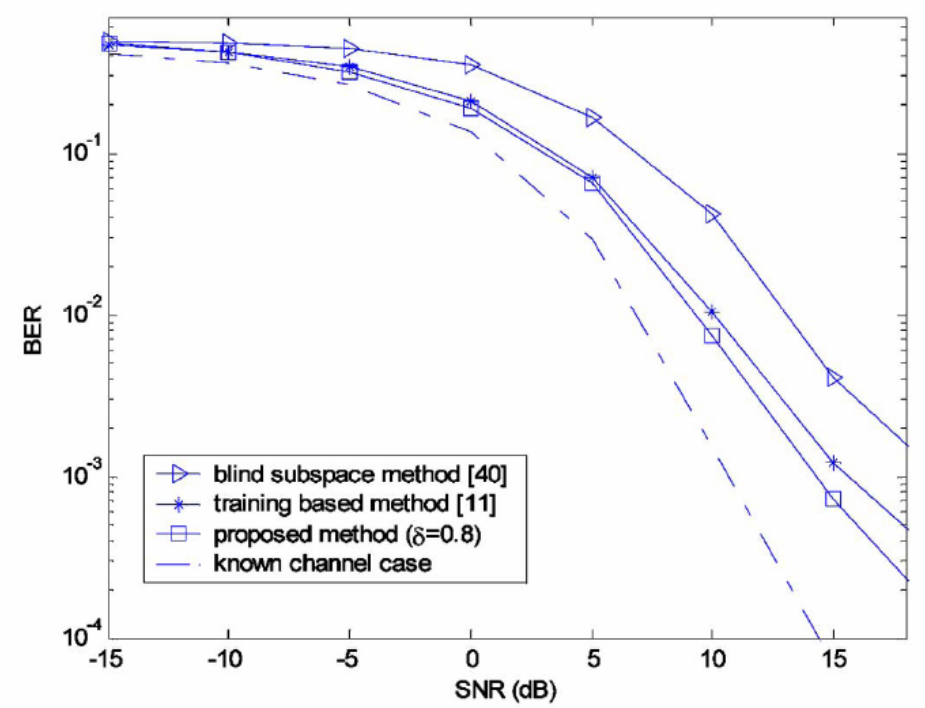

Fig. 10. BER performances of three methods (slowly time-varying channel with Jake's model).

optimal pilot sequence designed in [11] is placed in the initial symbol block pair per sub-burst, for training based estimation as well as for ambiguity removal in the two blind schemes; for the transmit-redundancy based solution [40], symbol constellations are likewise loaded for comparison under a fixed data rate. Figure 10 shows the respective BER curves (averaged over $I = 2000$ trials). Compared with the quasi-static case (Figures 8 and 9), the performances of the three methods degrade due to the time-varying channel characteristic; the proposed method, still, leads to the lowest error rate.

Simulation 8-On PAPR Performance: This simulation investigates the PAPR results when the optimal precoder (5.13) is used. For the considered system parameters ($N = 32$, $L = 8$) and with Nyquist pulse shaping filter, the values of PAPR for various choices of δ with respect to different symbol constellations are tabulated in [37, p-1124]. The results show that the two-level precoder (5.13) does increase PAPR

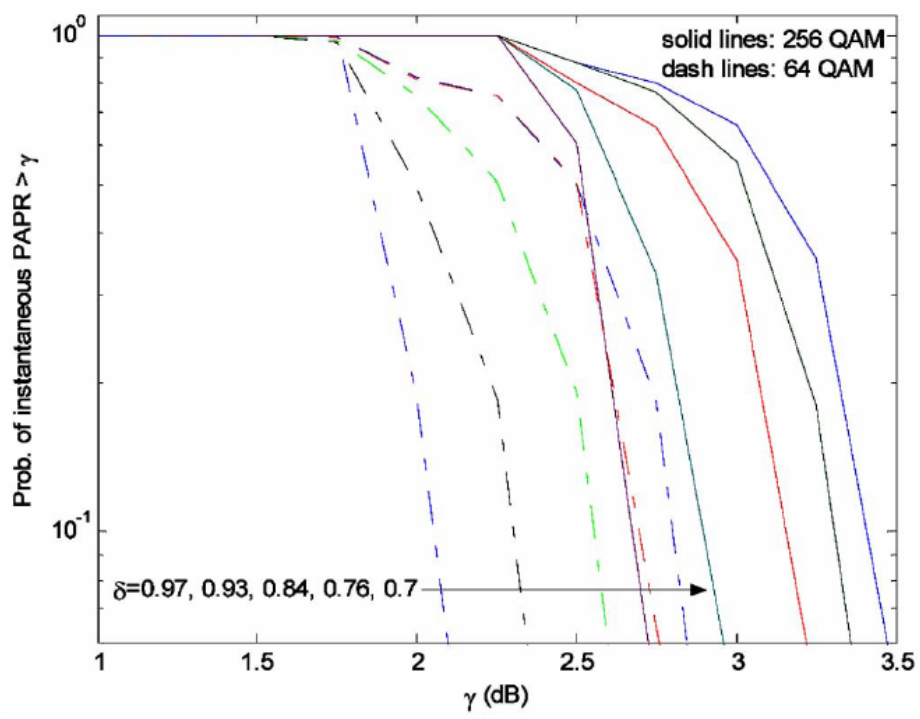

Fig. 11. Tendency of instantaneous PAPR for two QAM constellations.

over the equal-power case; in particular, the choice $\delta = 0.3$ raises the PAPR to a level comparable to that of multi-carrier scenario. It is noted that the PAPR does not faithfully reflect the actual signal amplitude variation in all cases; this is because the probability that such a peak occurs could depend on the block length as well [36]. A more realistic performance metric for describing the actual power amplification portrait is the *instantaneous* PAPR [36, p-383]. For 64-QAM and 256-QAM constellations, Figure 11 shows the probability that the instantaneous PAPR exceeds a prescribed value γ with respect to five different choices of power thresholds: $\delta = 0.7, 0.76, 0.84, 0.93,$ and 0.97 (the precoding induced power spike, i.e., $N-(N-1)\delta$, are respectively 10.3, 8.44, 6, 3, 1.8). As we can see from the figure, it is likely that no more than 1 dB power back-off is required as δ decreases from 0.97 to 0.7 (or equivalently, the spike value increases from 1.8 to 10.3); a slight impact on

the incurred instantaneous PAPR is also observed when different symbol constellations are used. As a result, with moderate choices of δ, the proposed optimal precoder (5.13) does not seem to induce a large power back-off in practice. It is noted that, although the instantaneous PAPR will increase when δ falls below 0.7, small δ should be precluded for maintaining the BER performance (see Figure 7).

Chapter 8

Conclusions

Blind channel estimation for STBC transmission in a MISO environment is a challenging research problem. This paper presents a solution for FDE-STBC systems. The proposed method relies on non-redundant diagonal precoding and i.i.d. source assumption, and exploits the resultant linear signal structure in the conjugate cross-correlation matrix of the received data. The circulant channel matrix property, which is unique to FDE based block transmission, leads to an identification equation set with a BCCB nature. Such a distinctive system matrix structure simplifies the characterization of channel identifiability condition (in terms of precoder coefficients), and can also alleviate the underlying algorithm complexity. The proposed channel estimate has an appealing property: it is exact when perfect data statistic is available and channel noise is circularly Gaussian. In the presence of finite-sample estimation errors, our channel estimation framework easily incorporates the data mismatch effect, and allows for natural precoder design formulation and criteria for improving estimation accuracy. Through analysis the optimization problems can be formulated to exploit the BCCB matrix property and are then analytically solved. The proposed solution tends to optimize the channel estimation robustness against deterministic error perturbation, and also minimize the mean square error when data mismatch is modeled as a white noise. The PEP analysis shows a trade-off regarding the proposed optimal error-resistant precoder: it incurs the worst-case power penalty for symbol decision. Through numerical simulations compromising choices for precoder parameters are determined. Simulation results show that the proposed approach

compares favorably with existing training and blind methods fitted for FDE-STBC systems.

Appendix

A: Proof of Theorem 4.

We will denote by $BCCB_{N,N}$ the set of all $N^2 \times N^2$ block circulant matrices with circulant blocks [12, p-184], each characterized by N circulant matrices of dimension $N \times N$. The proof of Theorem 4.1 is based on the following lemma.

Lemma A.1 [12, p-185]: If $\mathbf{X} \in BCCB_{N,N}$, then $\mathbf{X}$ can be diagonalized by $\mathbf{F} \otimes \mathbf{F}$. More precisely, let $\left\{\mathbf{C}_0, \cdots, \mathbf{C}_{N-1}\right\}$ be the set of $N \times N$ circulant matrices on the top row block of $\mathbf{X}$, and let $\mathbf{\Lambda}_n$ be the diagonal matrix containing the eigenvalues of $\mathbf{C}_n$. Then we have

$$\mathbf{X} = (\mathbf{F} \otimes \mathbf{F})\left(\sum_{n=0}^{N-1} \mathbf{\Omega}_N^n \otimes \mathbf{\Lambda}_n\right)\left(\mathbf{F}^{-1} \otimes \mathbf{F}^{-1}\right), \tag{A.1}$$

where $\mathbf{\Omega}_N \coloneqq diag\left\{\begin{bmatrix}1 & \omega & \omega^2 & \cdots & \omega^{N-1}\end{bmatrix}^T\right\}$. Conversely, any matrix of the form $(\mathbf{F} \otimes \mathbf{F})\bar{\mathbf{\Lambda}}\left(\mathbf{F}^{-1} \otimes \mathbf{F}^{-1}\right)$ for some diagonal $\bar{\mathbf{\Lambda}}$ belongs to $BCCB_{N,N}$.

□

[Sketch Proof of Theorem 4.1]: It can be seen from (3.14) that the matrix $\mathbf{Q}$ is characterized by the N circulant matrices $\left\{p(0)^2\mathbf{I}_N, p(1)^2\mathbf{J}, \cdots, p(N-1)^2\mathbf{J}^{N-1}\right\}$ on its top row block. If we stack the N eigenvalues of the circulant matrix $p(n)^2\mathbf{J}^n$ into a vector, say $\mathbf{u}_n$, by definition of $\mathbf{J}$ (cf. (3.7)) it can be verified that

$$\mathbf{u}_0 = \sqrt{N} \cdot p(0)^2\mathbf{f}_0 \text{ and } \mathbf{u}_n = \sqrt{N} \cdot p(n)^2\mathbf{f}_{N-n}, \text{ for } 1 \le n \le N-1, \tag{A.2}$$

where $\mathbf{f}_n \coloneqq 1/\sqrt{N} \cdot \begin{bmatrix}1 & \omega^n & \omega^{2n} & \cdots & \omega^{(N-2)n} & \omega^{(N-1)n}\end{bmatrix}^T \in \mathbb{C}^N$ is the $(n+1)$th column of $\mathbf{F}^{-1}$, $0 \le n \le N-1$. Based on (A.2), Lemma A.1 and by going through essentially the same steps as in [37, Appendix A], it can be shown that the eigenvalues of $\mathbf{Q}$ are the N^2 entries of the vector

$$\mathbf{q}_{\mathbf{Q}} := \left[\left(\sqrt{N}\mathbf{F}\mathbf{p}\right)^T \quad \left(\left(\mathbf{J}^T\right)\sqrt{N}\mathbf{F}\mathbf{p}\right)^T \quad \cdots \quad \left(\left(\mathbf{J}^T\right)^{N-1}\sqrt{N}\mathbf{F}\mathbf{p}\right)^T \right]^T, \tag{A.3}$$

and the assertion thus follows. □

B: On Computation of Product Unknowns

In the following theorem we will show that $\mathbf{Q}^{-1}$ also belongs to the BCCB category. Moreover, it possesses an identical structure as that of $\mathbf{Q}$, and is completely specified by N scalar parameters which are very easy to compute.

Theorem B.1: Let $\mathbf{p}(z)$ be the polynomial defined as in (4.2). Assume $p(n)$ is chosen so that $\mathbf{Q}$ is nonsingular. Then $\mathbf{Q}^{-1}$ is also a BCCB matrix with $\left[\alpha_0\mathbf{I}_N, \alpha_1\mathbf{J}, \cdots, \alpha_{N-1}\mathbf{J}^{N-1}\right] \in \mathbb{R}^{N\times N^2}$ as the top row block, where

$$\underbrace{[\alpha_0 \quad \cdots \quad \alpha_{N-1}]^T}_{:=\boldsymbol{\alpha}} = (\sqrt{N})^{-1}\mathbf{F}^{-1}[\mathbf{p}(1)^{-1} \quad \mathbf{p}(\omega)^{-1} \quad \cdots \quad \mathbf{p}(\omega^{N-1})^{-1}]^T. \tag{B.1}$$

□

[Proof]: From Lemma A.1, it is easy to see that $\mathbf{Q} \in BCCB_{N,N}$ implies $\mathbf{Q}^{-1} \in BCCB_{N,N}$, and hence we can write $\mathbf{Q}^{-1} = (\mathbf{F}\otimes\mathbf{F})\bar{\boldsymbol{\Lambda}}\left(\mathbf{F}^{-1}\otimes\mathbf{F}^{-1}\right)$ for some diagonal $\bar{\boldsymbol{\Lambda}}$. It suffices to check that, for the BCCB matrix with $\left\{\alpha_0\mathbf{I}_N, \alpha_1\mathbf{J}, \cdots, \alpha_{N-1}\mathbf{J}^{N-1}\right\}$ on the top row block, the resultant $\bar{\boldsymbol{\Lambda}}$ satisfies $\bar{\boldsymbol{\Lambda}}\boldsymbol{\Lambda}_{\mathbf{Q}} = \mathbf{I}_{N^2}$, where $\boldsymbol{\Lambda}_{\mathbf{Q}} = diag\left\{\mathbf{q}_{\mathbf{Q}}\right\}$ and $\mathbf{q}_{\mathbf{Q}}$ is in (A.3). By following the procedures as in Appendix A, it can be shown $\bar{\boldsymbol{\Lambda}} = diag\{\mathbf{q}\}$, where

$$\mathbf{q} = \left[\left(\sqrt{N}\mathbf{F}\boldsymbol{\alpha}\right)^T \quad \left(\left(\mathbf{J}^T\right)\sqrt{N}\mathbf{F}\boldsymbol{\alpha}\right)^T \quad \cdots \quad \left(\left(\mathbf{J}^T\right)^{N-1}\sqrt{N}\mathbf{F}\boldsymbol{\alpha}\right)^T \right]^T. \tag{B.2}$$

By definition of $\boldsymbol{\alpha}$ in (B.1), it is easy to see that the lth entry of $\mathbf{q}$ is simply the

reciprocal of the lth entry of $\mathbf{q}_{\mathbf{Q}}$ in (A.3). The proof is thus completed.

□

Theorem B.1 asserts that, to invert the $N^2 \times N^2$ matrix $\mathbf{Q}$, it merely requires to compute N scalars α_i's via (B.1). This calls for two N-point FFT operations, one for computing $\mathbf{p}(\omega^n)$'s form $p(n)$ and the other for α_n's based on $\mathbf{p}(\omega^n)^{-1}$'s; additional N multiplications are also needed for evaluating $\mathbf{p}(\omega^n)^{-1}$'s from $\mathbf{p}(\omega^n)$'s. When the symbol block length N is chosen to be a power of two, the number of flop counts is $4N\log_2 N - N$. The main computational cost for (4.8), on the other hand, lies in inverting an $(L+1)^2 \times (L+1)^2$ Hermitian Toeplitz matrix $\tilde{\mathbf{Q}}^T\tilde{\mathbf{Q}}$; the complexity can be limited to $4(L+1)^4$ by using the Levinson algorithm [17, p-196]. □

C: Proof of Lemma 5.1

The assertion relies on the following key observation: *any given* $p(n)$ satisfying (5.6) can be constructed by "squeezing" the peak power of the two-level solution (5.13) so that the ground powers at other time instants are "raised" to the prescribed levels. More precisely, let $p(n)$ be an admissible sequence such that $\delta < p(n)^2 < N-(N-1)\delta$ for $n \in \mathcal{I}$, where the index set $\mathcal{I}$ is a subset of $\{0,\cdots,N-1\} \setminus \{m\}$. Then $p(n)$ can be expressed as

$$p(m)^2 = N-(N-1)\delta - \sum_{n\in\mathcal{I}} \Delta_n\,, \tag{C.1}$$

and

$$p(n)^2 = \delta + \Delta_n \quad \text{for} \quad n \in \mathcal{I}\,, \text{and} \quad p(n)^2 = \delta \quad \text{for} \quad n \notin \mathcal{I}\,, \tag{C.2}$$

where $\Delta_n > 0$ models the excessive power over the ground level δ for $n \in \mathcal{I}$. The sequence of the form (C.1) and (C.2) satisfies the constraints (5.6a) and (5.6b); in particular, since $p(m)^2 \geq \delta$ is required, we can infer from (C.1) that

$$\sum_{n\in\mathcal{I}} \Delta_n \leq N(1-\delta)\,. \tag{C.3}$$

We assume for the moment that $m = 0$; as one will see, the result for the

$1 \le m \le N-1$ case easily follows. Associated with $p(n)$ in (C.1) and (C.2), we have, for $1 \le k \le N-1$,

$$\mathbf{p}(\omega^k) = \sum_{n=0}^{N-1} p(n)^2 \omega^{-kn} = \left[N - (N-1)\delta - \sum_{n \in \mathcal{I}} \Delta_n \right] + \sum_{n \in \mathcal{I}} (\delta + \Delta_n) \omega^{-kn} + \sum_{n \notin \mathcal{I}} \delta \omega^{-kn}$$

,

$$= \left[N(1-\delta) + \sum_{n \in \mathcal{I}} (\omega^{-kn} - 1)\Delta_n \right] + \underbrace{\left[\delta + \sum_{n \in \mathcal{I}} \delta \omega^{-kn} + \sum_{n \notin \mathcal{I}} \delta \omega^{-kn} \right]}_{= \delta \sum_{n=0}^{N-1} \omega^{-kn} = 0} = \left[N(1-\delta) + \sum_{n \in \mathcal{I}} (\omega^{-kn} - 1)\Delta_n \right]$$

. (C.4)

Let us define the nonnegative number

$$d \coloneqq N(1-\delta) - \sum_{n \in \mathcal{I}} \Delta_n , \tag{C.5}$$

Since $\omega^{-kn} = \cos n\theta_k - j \sin n\theta_k$, where $\theta_k \coloneqq 2\pi k / N$, and with (C.5), it follows from (C.4) that

$$\left| \mathbf{p}(\omega^k) \right|^2 = \left[d + \sum_{n \in \mathcal{I}} \Delta_n \cos n\theta_k \right]^2 + \left[\sum_{n \in \mathcal{I}} \Delta_n \sin n\theta_k \right]^2$$

$$= d^2 + 2d \left(\sum_{n \in \mathcal{I}} \Delta_n \cos n\theta_k \right) + \left(\sum_{n \in \mathcal{I}} \Delta_n \cos n\theta_k \right)^2 + \left(\sum_{n \in \mathcal{I}} \Delta_n \sin n\theta_k \right)^2 . \tag{C.6}$$

Observe that

$$\left(\sum_{n \in \mathcal{I}} \Delta_n \cos n\theta_k \right)^2 + \left(\sum_{n \in \mathcal{I}} \Delta_n \sin n\theta_k \right)^2 = \sum_{n \in \mathcal{I}} \Delta_n^2 + 2 \sum_{n_l, n_m \in \mathcal{I}} \Delta_{n_l} \Delta_{n_m} (\cos n_l\theta_k \cos n_m\theta_k + \sin n_l\theta_k \sin n_m\theta_k)$$

$$= \sum_{n \in \mathcal{I}} \Delta_n^2 + 2 \sum_{n_l, n_m \in \mathcal{I}} \Delta_{n_l} \Delta_{n_m} \cos(n_l - n_m)\theta_k \;\le\; \sum_{n \in \mathcal{I}} \Delta_n^2 + 2 \sum_{n_l, n_m \in \mathcal{I}} \Delta_{n_l} \Delta_{n_m} = \left(\sum_{n \in \mathcal{I}} \Delta_n \right)^2 , \tag{C.7}$$

and that

$$2d \left(\sum_{n \in \mathcal{I}} \Delta_n \cos n\theta_k \right) \le 2d \left(\sum_{n \in \mathcal{I}} \Delta_n \right). \tag{C.8}$$

From (C.7) and (C.8), $\left| \mathbf{p}(\omega^k) \right|^2$ in (C.6) can be upper bounded as

$$\left|\mathbf{p}(\omega^k)\right|^2 \le \ d^2 + 2d\left(\sum_{n\in\mathcal{I}}\Delta_n\right) + \left(\sum_{n\in\mathcal{I}}\Delta_n\right)^2 = \ \left(d + \sum_{n\in\mathcal{I}}\Delta_n\right)^2 = N^2(1-\delta)^2, \qquad \text{(C.9)}$$

in which the last equality follows from the definition of d in (C.5). This thus proves the lemma, under the assumption $m = 0$ in (C.1). For $1 \le m \le N-1$, equation (C.4) is then accordingly modified as

$$\mathbf{p}(\omega^k) = \left[N(1-\delta) + \sum_{n\in\mathcal{I}}(\omega^{-(k-m)n} - 1)\Delta_n\right]\omega^{-mn} \qquad \text{(C.10)}$$

By going through the same procedures as in (C.5)~(C.8) the conclusion (C.9) will follow. □

D: Proof of Lemma 5.3

Without loss of generality we assume $\mathbf{M}$ is split as $\mathbf{M} = \left[\bar{\mathbf{M}} \quad \mathbf{M}_d\right]$, in which $\mathbf{M}_d$ contains the columns to be deleted; otherwise we can multiply $\mathbf{M}$ from the right by a permutation matrix to put it in this partition. It thus follows

$$\mathbf{M}^H\mathbf{M} = \begin{bmatrix}\bar{\mathbf{M}}^H \\ \mathbf{M}_d^H\end{bmatrix}\left[\bar{\mathbf{M}} \quad \mathbf{M}_d\right] = \begin{bmatrix}\bar{\mathbf{M}}^H\bar{\mathbf{M}} & \bar{\mathbf{M}}^H\mathbf{M}_d \\ \mathbf{M}_d^H\bar{\mathbf{M}} & \mathbf{M}_d^H\mathbf{M}_d\end{bmatrix}. \qquad \text{(D.1)}$$

Since $\mathbf{M}$ is nonsingular, $\mathbf{M}^H\mathbf{M}$ is positive definite. By the inversion lemma for block matrix [20, p-572], we have

$$\left(\mathbf{M}^H\mathbf{M}\right)^{-1} = \begin{bmatrix}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}} - \bar{\mathbf{M}}^H\mathbf{M}_d\left(\mathbf{M}_d^H\mathbf{M}_d\right)^{-1}\mathbf{M}_d^H\bar{\mathbf{M}}\right)^{-1} & \times \\ \times & \left(\mathbf{M}_d^H\mathbf{M}_d - \mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\right)^{-1}\end{bmatrix}, \qquad \text{(D.2)}$$

in which the notation "$\times$" stands for the block off-diagonal submatrices irrelevant to the proof procedures. From (D.2), we have

$$Tr\left[\left(\mathbf{M}^H\mathbf{M}\right)^{-1}\right] = Tr\left[\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}} - \bar{\mathbf{M}}^H\mathbf{M}_d\left(\mathbf{M}_d^H\mathbf{M}_d\right)^{-1}\mathbf{M}_d^H\bar{\mathbf{M}}\right)^{-1}\right] + Tr\left[\left(\mathbf{M}_d^H\mathbf{M}_d - \mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\right)^{-1}\right] \qquad \text{(D.3)}$$

Since $\left(\mathbf{M}^H\mathbf{M}\right)^{-1}$ is positive definite, so are its principle submatrices and (D.3) implies

$$Tr\left[\left(\mathbf{M}^H\mathbf{M}\right)^{-1}\right] \geq Tr\left[\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}} - \bar{\mathbf{M}}^H\mathbf{M}_d\left(\mathbf{M}_d^H\mathbf{M}_d\right)^{-1}\mathbf{M}_d^H\bar{\mathbf{M}}\right)^{-1}\right]. \tag{D.4}$$

Using the matrix inversion lemma [20, p-571], inequality (D.4) can be further expanded into

$$Tr\left[\left(\mathbf{M}^H\mathbf{M}\right)^{-1}\right] \geq Tr\left[\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}} - \bar{\mathbf{M}}^H\mathbf{M}_d\left(\mathbf{M}_d^H\mathbf{M}_d\right)^{-1}\mathbf{M}_d^H\bar{\mathbf{M}}\right)^{-1}\right]$$

$$= Tr\left[\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1} + \left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\left(\mathbf{M}_d^H\mathbf{M}_d - \mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\right)^{-1}\mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\right]$$

$$= Tr\left[\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\right] + Tr\left[\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\left(\mathbf{M}_d^H\mathbf{M}_d - \mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\right)^{-1}\mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\right] \tag{D.5}$$

Since $\left(\mathbf{M}_d^H\mathbf{M}_d - \mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\right)^{-1}$ is a principle submatrix of $\left(\mathbf{M}^H\mathbf{M}\right)^{-1}$ (cf. (D.1)), it is positive definite and so is $\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\left(\mathbf{M}_d^H\mathbf{M}_d - \mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\right)^{-1}\mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}$. The result then follows from (D.5). □

E: Proof of Theorem 6.1

We will prove the theorem by induction. We will first show that (6.7) holds for an arbitrary admissible three-level sequence which, according to (C.1) and (C.2), can be parameterized as

$$p(m)^2 = N - (N-1)\delta - \Delta_{n_0}, \tag{E.1}$$

and

$$p(n_0)^2 = \delta + \Delta_{n_0}, \text{ and } \; p(n)^2 = \delta \;\; \text{for} \;\; n \notin \left\{m, n_0\right\}, \tag{E.2}$$

where $0 \leq n_0 \neq m \leq N-1$. From (C.3), we must have $\Delta_{n_0} \leq N(1-\delta)$. With (E.1) and (E.2), it follows that

$$\sum_{n=0}^{N-1} p(n)^{-2} = \frac{N-2}{\delta} + \frac{1}{\delta + \Delta_{n_0}} + \frac{1}{N-(N-1)\delta - \Delta_{n_0}} = \frac{N-2}{\delta} + \frac{N(1-\delta)}{-\Delta_{n_0}^2 + N(1-\delta)\Delta_{n_0} + [N-(N-1)\delta]\delta}$$

(E.3)

It thus suffices to show $J(\Delta_{n_0}) \coloneqq -\Delta_{n_0}^2 + N(1-\delta)\Delta_{n_0} + [N-(N-1)\delta]\delta$ is minimized by either $\Delta_{n_0} = 0$ or $\Delta_{n_0} = N(1-\delta)$: this confirms that the maximizing $p(n)$ reduces to the two-level form (5.13). Indeed, simple manipulations show that the minimal value of $J(\Delta_{n_0})$ within the interval $\Delta_{n_0} \in [0, N(1-\delta)]$ is $[N-(N-1)\delta]\delta$, which is attained by the two boundary points. Now we assume that (6.7) holds for an arbitrary K-level sequence $(3 < K \le N)$, which is described as

$$p(m)^2 = N-(N-1)\delta - \sum_{k=1}^{K-2} \Delta_{n_k}, \tag{E.4}$$

and

$$p(n)^2 = \delta + \Delta_{n_k} \text{ for } k \in \{1,\cdots,K-2\}, \text{ and } p(n)^2 = \delta \text{ for } n \notin n_k\text{'s}. \tag{E.5}$$

Hence we have

$$\sum_{n=0}^{N-1} p(n)^{-2} = \frac{N-(K-1)}{\delta} + \sum_{k=1}^{K-2} \frac{1}{\delta+\Delta_{n_k}} + \left(N-(N-1)\delta - \sum_{k=1}^{K-2}\Delta_{n_k}\right)^{-1} \le \frac{N-1}{\delta} + \frac{1}{N-(N-1)\delta}. \tag{E.6}$$

For any $(K+1)$-level sequence given by

$$p(m)^2 = N-(N-1)\delta - \sum_{k=1}^{K-1} \Delta_{n_k}, \tag{E.7}$$

$$p(n)^2 = \delta + \Delta_{n_k} \text{ for } k \in \{1,\cdots,K-1\}, \text{ and } p(n)^2 = \delta \text{ for } n \notin n_k\text{'s}. \tag{E.8}$$

we have

$$\begin{aligned}\sum_{n=0}^{N-1} p(n)^{-2} &= \frac{N-K}{\delta} + \sum_{k=1}^{K-1} \frac{1}{\delta+\Delta_{n_k}} + \left(N-(N-1)\delta - \sum_{k=1}^{K-1}\Delta_{n_k}\right)^{-1} \\ &= \underbrace{\left\{\frac{N-(K-1)}{\delta} + \sum_{k=1}^{K-2} \frac{1}{\delta+\Delta_{n_k}} + \left(N-(N-1)\delta - \sum_{k=1}^{K-2}\Delta_{n_k}\right)^{-1}\right\}}_{\le \frac{N-1}{\delta} + \frac{1}{N-(N-1)\delta}} \\ &\quad + \underbrace{\left\{\frac{1}{\delta+\Delta_{n_{K-1}}} - \frac{1}{\delta} + \left(N-(N-1)\delta - \sum_{k=1}^{K-1}\Delta_{n_k}\right)^{-1} - \left(N-(N-1)\delta - \sum_{k=1}^{K-2}\Delta_{n_k}\right)^{-1}\right\}}_{:=\Theta}.\end{aligned} \tag{E.9}$$

With (E.9), it thus remains to check $\Theta \le 0$. It is easy to verify that

$$\Theta = \frac{-\Delta_{n_{K-1}}}{(\delta + \Delta_{n_{K-1}})\delta} + \frac{\Delta_{n_{K-1}}}{\left(N - (N-1)\delta - \sum_{k=1}^{K-1} \Delta_{n_k}\right)\left(N - (N-1)\delta - \sum_{k=1}^{K-2} \Delta_{n_k}\right)}. \quad \text{(E.10)}$$

Subject to the power threshold constraint (5.6b), we must have,

$$N - (N-1)\delta - \sum_{k=1}^{K-1} \Delta_{n_k} \geq \delta \text{, and hence } \; N - (N-1)\delta - \sum_{k=1}^{K-2} \Delta_{n_k} \geq \delta + \Delta_{n_{K-1}}. \quad \text{(E.11)}$$

The two inequalities in (E.11) imply

$$\left(N - (N-1)\delta - \sum_{k=1}^{K-1} \Delta_{n_k}\right)^{-1} \left(N - (N-1)\delta - \sum_{k=1}^{K-2} \Delta_{n_k}\right)^{-1} \leq \frac{1}{(\delta + \Delta_{n_{K-1}})\delta}. \quad \text{(E.12)}$$

The assertion follows immediately from (E.10) and (E.12). □

References

[1] S. Alamouti, “A simple transmit diversity technique for wireless communications,” *IEEE J. Selected Areas in Communications*, vol. 16, no. 10, pp. 1451-1458, Oct. 1998.

[2] N. Al-Dhahir, “Single-carrier frequency-domain equalization for space-time block coded transmission over frequency-selective fading channels,” *IEEE Communications Letters*, vol. 5, no. 7, pp. 304-306, July 2001.

[3] N. Al-Dhahir, “Overview and comparison of equalization schemes for space-time coded signals with applications to EDCG,” *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2477-2488, Oct. 2002.

[4] N. Ammar and Z. Ding, “Channel estimation under space-time block coded transmission,” *Proc. Sensor Array and Multi-channel Signal Processing Workshop 2002*, pp. 422-426.

[5] N. Ammar and Z. Ding, “Frequency selective channel estimation in time-reversed space-time coding,” *Proc. WCNC 2004*, pp. 1838-1843.

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

[7] E. Beres and R. Adve, “Blind channel estimation for orthogonal STBC in MISO systems,” *Proc. Globecom 2004*, pp. 2323-2328.

[8] H. Bolcskei, R. W. Heath, and A. J. Paulraj, “Blind channel estimation in spatial multiplexing systems using nonredundant antenna precoding,” in *Proc. Asilomar Conf. Singlas, Systems, and Computers*, vol. 2, Pacific Grove, CA, Oct. 1999, pp. 1127-1132.

[9] H. Bolcskei, R. W. Heath, and A. J. Paulraj, “Blind channel identification and equalization in OFDM-based multiantenna systems,” *IEEE Trans. Signal Processing*, vol. 50, no. 1, pp. 96-109, Jan. 2002.

[10] A. Chevreuil, E. Serpedin, P. Loubaton, and G. B. Giannakis, “Blind channel identification and equalization using periodic modulation precoders: performance analysis,” *IEEE Trans. Signal Processing*, vol. 48, no. 6, pp. 1570-1586, June 2000.

[11] J. Coon, M. Beach, and J. McGeehan, “Optimal training sequences for channel estimation in cyclic-prefix based single-carrier systems with transmit diversity,” *IEEE Signal Processing Letters*, vol. 11, no. 9, pp. 729-732, Sept. 2004.

[12] P. J. Davis, *Circulant Matrices*, John Wieley & Sons, Inc, 1979.

[13] Z. Ding, “Matrix outer-product decomposition method for blind multiple channel identification,” *IEEE Trans. Signal Processing*, vol. 45, no. 12, pp. 3053-3061, Dec. 1997.

[14] Z. Ding and Y. Li, *Blind Equalization and Identification*, Marcel Dekker, Inc., 2001.

[15] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, “Frequency domain equalization for single-carrier broadband wireless systems,” *IEEE Communications Magazine*, pp. 58-66, April 2002.

[16] G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, *Signal Processing Advances in Wireless and Mobil Communication Volume I: Trends in Channel Identification and Equalization*, Prentice Hall PTR, 2001.

[17] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3-rd edition, The Johns Hopkins University Press, 1996.

[18] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press,

1985.

[19] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.

[20] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall Inc., 1993.

[21] T. P. Krauss and M. D. Zoltowski, "Bilinear approach to multiuser second-order statistics-based blind channel estimation," *IEEE Trans. Signal Processing*, vol. 48, no. 9, pp. 2473-2486, Sept. 2000.

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

[23] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*, Prentice Hall, 1974.

[24] C. A. Lin and J. Y. Wu, "Blind identification with periodic modulation: A time-domain approach," *IEEE Trans. Signal Processing*, vol. 50, no. 11, pp. 2875-2888, Nov. 2002.

[25] Y. P. Lin and S. M. Phoong, "BER minimized OFDM systems with channel independent precoders," *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2369-2380, Sept. 2003.

[26] E. Lindskog and A. Paulraj, "A transmit diversity scheme for delay spread channels," *Proc. ICC'00*, pp. 307-311, New Orleans, LA.

[27] Z. Liu, G. B. Giannakis, S. Barbarossa, and A. Scaglione, "Transmit-antenna space-time block coding for generalized OFDM in the presence of unknown multipath," *IEEE J. Selected Areas in Communications*, vol. 19, no. 7, pp. 1352-1364, July, 2001.

[28] Z. Q. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong, "Transceiver

optimization for block-based multiple access through ISI channels," *IEEE Trans. on Signal Processing*, vol. 52, no. 4, pp. 1037-1052, April 2004.

[29] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, no. 2, pp. 516-525, Feb. 1995.

[30] B. O'Hara and A. Petrick, *The IEEE 802.11 Handbook, A Designer's Companion*, IEEE Press, 1999.

[31] E. Serpedin and G. B. Giannakis, "Blind channel identification and equalization with modulation-induced cyclostationarity," *IEEE Trans. on Signal Processing*, vol. 46, no. 7, pp. 1930-1944, July 1998.

[32] S. Shahbazpanahi, A. B. Gershman, and J. H. Manton, "Closed-form blind decoding of orthogonal space-time block codes," *Proc. ICASSP 2004*, vol. IV, pp. 473-476.

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

[34] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Information Theory*, vol. 45, no. 7, pp. 1456-1467, July, 1999.

[35] C. Tepedelenlioglu, "Maximum multipath diversity with linear equalization in precoded OFDM systems," *IEEE Trans. Information Theory*, vol. 50, no. 1, pp. 232-234, Jan. 2004.

[36] Z. Wang, X. Ma, and G. B. Giannakis, "OFDM or single-carrier block transmission ?" *IEEE Trans. Communications*, vol. 52, no. 3, pp. 380-394, March 2004.

[37] J. Y. Wu and T. S. Lee, “Periodic-modulation based blind channel identification for single-carrier block transmission with frequency-domain equalization,” *IEEE Trans. Signal Processing*, vol. 54, no. 3, pp. 1114-1130, March 2006.

[38] G. Xu, H. Liu, L. Tong, and T. Kailath, “A least-squares approach to blind channel identification,” *IEEE Trans. Signal Processing*, vol. 43, no. 12, pp. 2982-2993, Dec. 1995.

[39] S. Zhou and G. B. Giannakis, “Single-carrier space-time block-coded transmissions over frequency-selective fading channels,” *IEEE Trans. Information Theory*, vol. 49, no. 1, pp. 164-179, Jan. 2003.

[40] S. Zhou, B. Muquet, and G. B. Giannakis, “Subspace-based (semi-) blind channel estimation for block precoded space-time OFDM,” *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 1215-1228, May 2002.

計畫成果自評

一、研究內容與原計畫相符程度

大致相符。

二、達成預期目標情況

1. 本研究結果是專門設計給 FDE-STBC 所使用。
2. 參與學生可獲得未來信號處理及通訊領域必要的訓練和技術。
3. 所提出的方法可作為實現下世代無線通訊標準高速且高可靠度的參考方案。

三、研究成果之學術或應用價值

此通道估測的公式化建立在重組共軛交叉相關矩陣的線性方程式集合以及通道脈衝響應，使之成為一個具有區塊循環循環區塊（block-circulant with circulant-block (BCCB)）的特殊結構。這樣允許了一個簡單的僅視先期編碼參數而定的可辨識條件，也提供了一個自然而有效的最佳先期編碼器的設計架構來改善當不完全資料估測發生時的解答正確性。

四、是否適合在學術期刊發表或申請專利

適合在學術期刊發表。

五、主要發現或其他有關價值

共軛交叉相關介於兩個時間區塊信號的先期編碼可以產生線性信號結構和循環行列式的通道矩陣特性，並且在通道雜訊為循環高斯而接收機資料統計可以完整的得到時可以產生精確解。

可供推廣之研發成果資料表

□ 可申請專利　■ 可技術移轉　日期：97 年 10 月 29 日

國科會補助計畫	計畫名稱：具頻域等化機制之單載波空時區塊碼之盲蔽式多通道判別：以非多餘傳送前置編碼為基礎的研究 計畫主持人：李大嵩教授、吳卓諭助理教授 計畫編號：NSC 95-2221-E-009-047-MY2 學門領域：
技術/創作名稱	具頻域等化機制之單載波空時區塊碼之盲蔽式多通道判別
發明人/創作人	吳卓諭、李大嵩
技術說明	中文： 此通道估測的公式化建立在重組共軛交叉相關矩陣的線性方程式集合以及通道脈衝響應，使之成為一個具有區塊循環循環區塊（block-circulant with circulant-block (BCCB)）的特殊結構。這樣允許了一個簡單的僅視先期編碼參數而定的可辨識條件，也提供了一個自然而有效的最佳先期編碼器的設計架構來改善當不完全資料估測發生時的解答正確性。
	英文： The channel estimation formulation builds on rearranging the set of linear equations relating the entries of conjugate cross-correlation matrix and products of channel impulse responses into one with a distinctive block-circulant with circulant-block (BCCB) structure. This allows a simple identifiability condition depending on precoder parameters alone, and also provides a natural yet effective optimal precoder design framework for improving solution accuracy when imperfect data estimation occurs.
可利用之產業及可開發之產品	(1) 寬頻無線通訊產業、行動無線通訊產業。 (2) 行動式 OFDM 通訊系統、WiMAX 通訊系統。
技術特點	基於非冗長對角先期編碼和獨立同分佈（i.i.d.）訊源的假設，提出以單載波頻域等化器為基礎的空時區塊碼系統的盲蔽式通道估測法。此方法開發了共軛交叉相關介於兩個時間區塊信號的先期編碼所產生的線性信號結構和循環行列式的通道矩陣特性，並且在通道雜訊為循環高斯而接收機資料統計可以完整的得到時可以產生精確解。
推廣及運用的價值	可攜式無線通訊產品都將面臨通道的時變性，因此簡單又可靠的通道估測將是降低成本及提昇通訊品質的一大關鍵。

※ 1. 每項研發成果請填寫一式二份，一份隨成果報告送繳本會，一份送　貴單位研發成果推廣單位（如技術移轉中心）。

※ **2. 本項研發成果若尚未申請專利，請勿揭露可申請專利之主要內容。**

※ 3. 本表若不敷使用，請自行影印使用。

出席國際學術會議心得報告

計畫編號	NSC 95-2221-E-009-047-MY2
計畫名稱	具頻域等化機制之單載波空時區塊碼之盲蔽式多通道判別：以非多餘傳送前置編碼為基礎的研究
出國人員姓名 服務機關及職稱	吳卓諭 國立交通大學電信系助理教授
會議時間地點	自 97 年 3 月 30 日至 97 年 4 月 4 日 at Las Vegas, US
會議名稱	IEEE ICASSP 2008
發表論文題目	Energy-constrained MMSE decentralized estimation via partial sensor noise variance knowledge

一、參加會議經過

(i) Paper presentation,

(ii) Exchange ideas with senior professors (e.g., P. K. Varshney) about current research trends in sensor networks.

二、與會心得

We should dedicate more efforts to fundamental research

Energy-Constrained MMSE Decentralized Estimation via Partial Sensor Noise Variance Knowledge†

Jwo-Yuh Wu, Qian-Zhi Huang, and Ta-Sung Lee

Department of Communications Engineering, National Chiao Tung University

1001, Ta-Hsueh Road, Hsinchu, Taiwan

Emails: jywu@cc.nctu.edu.tw; qianzhi.yux@gmail.com; tslee@mail.nctu.edu.tw.

***Abstract*-This paper studies the energy-constrained MMSE decentralized estimation problem with the best-linear-unbiased-estimator fusion rule, under the assumptions that i) each sensor can only send a quantized version of its raw measurement to the fusion center (FC), and ii) exact knowledge of the sensor noise variance is unknown at the FC but only an associated statistical description is available. The problem setup relies on maximizing the reciprocal of the MSE averaged with respect to the prescribed noise variance distribution. While the considered design metric is shown to be highly nonlinear in the local sensor transmit energy (or bit loads), we leverage several analytic approximation relations to derive a associated tractable lower bound; through maximizing this bound a closed-form solution is then obtained. Our analytical results reveal that sensors with bad link quality are shut off to conserve energy, whereas the energy allocated to those active nodes is proportional to the individual channel gain. Simulation results are used to illustrate the performance of the proposed scheme.**

Index Terms: **Decentralized estimation; Sensor networks; Energy efficiency; Quantization; Convex optimization.**

I. INTRODUCTION

Low energy/power cost is a critical concern for various application-specific designs of sensor networks [15], [16]. In the decentralized estimation scenario, wherein each sensor can transmit only a compressed version of its raw measurement to the fusion center (FC) owing to bandwidth and power limitations, several energy-efficient estimation schemes have been reported in the literature [1], [7], [10], [11], [13], [14]. Since the transmission energy is proportional to the message length [2], [13], all these works are formulated within a quantization bit assignment setup, with the optimal bit load determined via the knowledge of *instantaneous* local sensor noise characteristics, e.g., the noise variance if the fusion rule follows the best-linear-unbiased- estimator (BLUE) principle [5, chap. 6]. To maintain the estimation performance against the variation of sensing conditions, repeated update of the noise profile is needed: this inevitably incurs more training overhead and hence extra energy consumption. The design of distributed estimation algorithms independent of the instantaneous noise parameters remains an open problem [13, p-419]. Relying on partial noise variance knowledge in the form of the statistical distribution, the problem of minimizing total transmission energy under an allowable average distortion level (measured in terms of a mean-square-error (MSE) based criterion averaged with respect to some prescribed statistical distribution) is recently considered in [11].

This paper complements the study of [11] by addressing the counterpart problem: how to find the optimal bit load which minimizes the average distortion under a fixed total energy budget. The main contribution of the current work can be summarized as follows: (i) while the design metric, in the form of the reciprocal of the MSE averaged with respect to the distribution, is shown in [11] to be highly nonlinear in the sensor bit load, we leverage several analytic approximation relations to derive an associated tractable low bound, (ii) by maximizing this lower bound the problem can be further formulated in the form of convex optimization which yields a closed-form solution. Our analytic results reveal that, toward utmost estimation accuracy under a limited energy budget, sensors with bad link quality should be shut off, and energy allocated to those active nodes should be proportional to the individual channel gain; a similar energy conservation policy is also found in the previous works [7], [11], [13]. Numerical simulation evidences the effectiveness of the proposed scheme: it outperforms the uniform allocation strategy in an energy-limited environment.

II. SYSTEM SCENARIO

Consider a wireless sensor network, in which N spatially deployed sensors cooperate with a FC for estimating an unknown deterministic parameter θ. The local observation at the ith node is

$$x_i = \theta + n_i, \quad 1 \le i \le N, \tag{2.1}$$

where n_i is a zero-mean measurement noise with variance σ_i^2. Due to bandwidth and power limitations each sensor quantizes its observation into a b_i-bit message, and then transmits this locally processed data to the FC to generate a final estimate of θ. In this paper the uniform quantization scheme with nearest- rounding [9], is adopted; the quantized message at the ith sensor can thus be modeled as

$$m_i = x_i + q_i, \quad 1 \le i \le N, \tag{2.2}$$

where q_i is the quantization error uniformly distributed with zero mean and variance $\sigma_{q_i}^2 = R^2/(12 \cdot 4^{b_i})$ [9], where $[-R/2, R/2]$ is the available signal amplitude range common to all sensors. The adopted quantizer model (2.2) and the uniform quantization error assumption, though being valid only

† This work is sponsored jointly by the National Science Council under grants NSC-96-2752-E-002-009, NSC-96-2628-E-009-003-MY3, by the Ministry of Education of Taiwan under the MoE ATU Program, and by MediaTek research center at National Chiao Tung University, Taiwan.

when the number of quantization bits is sufficiently large [9], are widely used in the literature due to analytical tractability. Assuming that the channel link between the *i*th sensor and the FC is corrupted by a zero-mean additive noise v_i with variance σ_v^2 The received data from all sensor outputs can thus be expressed in a vector form as[a]

$$[y_1 \cdots y_N]^T = [1 \cdots 1]^T \theta + \underbrace{[n_1 \cdots n_N]^T}_{:=\mathbf{n}} + \underbrace{[q_1 \cdots q_N]^T}_{:=\mathbf{q}} + \underbrace{[v_1 \cdots v_N]^T}_{:=\mathbf{v}}, \quad (2.3)$$

where $(\cdot)^T$ denotes the transpose. This paper focuses on linear fusion rules for parameter recovery. More specifically, by assuming that the noise components $\{\mathbf{n}, \mathbf{q}, \mathbf{v}\}$ in (2.3) are mutually independent and the respective samples n_i's, q_i's, and v_i's are also independent across sensors, the parameter θ is retrieved via the BLUE [5, p-138] scheme via

$$\hat{\theta} = \left(\sum_{i=1}^{N} \frac{y_i}{\sigma_i^2 + \sigma_v^2 + \beta 4^{-b_i}} \right) \left(\sum_{i=1}^{N} \frac{1}{\sigma_i^2 + \sigma_v^2 + \beta 4^{-b_i}} \right)^{-1}; \quad (2.4)$$

the incurred MSE is thus [5, p-138]

$$E\left|\hat{\theta} - \theta\right|^2 = \left(\sum_{i=1}^{N} \frac{1}{\sigma_i^2 + \sigma_v^2 + \beta 4^{-b_i}} \right)^{-1}, \quad \beta := R^2/12. \quad (2.5)$$

A commonly used statistical description for sensing noise variance is [7], [13]:

$$\sigma_i^2 = \delta + \alpha z_i, \quad 1 \le i \le N, \quad (2.6)$$

where δ models the network-wide noise variance threshold, α controls the underlying variation from the nominal minimum, and $z_i \sim \chi_1^2$ are i.i.d. central Chi-Square distributed random variables each with degrees-of-freedom equal to one [6, p-24]. The proposed energy-constrained MMSE decentralized estimation scheme is based on the noise variance model (2.6) and is discussed next.

III. MAIN RESULTS

A. Problem Setup

We assume that the *i*th sensor sends the b_i-bit message m_i by using QAM with a constellation size 2^{b_i}. The consumed energy is thus [2], [13],

$$E_i = w_i\left(2^{b_i} - 1\right) \text{ for some } w_i, \quad 1 \le i \le N; \quad (3.1)$$

the energy density w_i is defined as [2]

$$w_i := \rho d_i^{\kappa_i} \cdot \ln\left(2/P_b\right), \quad (3.2)$$

in which ρ is a constant depending on the noise profile, d_i is the distance between the *i*th node and the FC, κ_i is the *i*th path loss exponent, and P_b is the target bit error rate assumed common to all sensor-to-FC links. With (2.5) and (3.1), the energy allocated to the *i*th sensor is thus determined by the number of quantization bits b_i. For a fixed set of sensing noise variances σ_i^2's, the problem of MMSE decentralized estimation, under an allowable total energy budget E_T, can be formulated as

$$\text{Minimize } \left(\sum_{i=1}^{N} \frac{1}{\sigma_i^2 + \sigma_v^2 + \beta 4^{-b_i}} \right)^{-1}, \text{ s.t. } \sum_{i=1}^{N} w_i\left(2^{b_i} - 1\right) \le E_T,$$
$$\text{and } b_i \in \mathbb{Z}_0^+, \ 1 \le i \le N, \quad (3.3)$$

or equivalently,

$$\text{Maximize } \sum_{i=1}^{N} \frac{1}{\sigma_i^2 + \sigma_v^2 + \beta 4^{-b_i}}, \text{ s.t. } \sum_{i=1}^{N} w_i\left(2^{b_i} - 1\right) \le E_T$$
$$\text{and } b_i \in \mathbb{Z}_0^+, \ 1 \le i \le N, \quad (3.4)$$

where $\mathbb{Z}_0^+$ denotes the set of all nonnegative integers. To obtain a universal solution irrespective of instantaneous noise conditions, we will consider the following optimization problem, in which the equivalent distortion cost function in (3.4) is instead averaged with respect to the noise variance statistic characterized in (2.6):

$$\text{Maximize } \int_{\mathbf{z}} \sum_{i=1}^{N} \frac{1}{\tilde{\delta} + \alpha z_i + \beta 4^{-b_i}} p(\mathbf{z}) \, d\mathbf{z},$$
$$\text{s.t. } \sum_{i=1}^{N} w_i\left(2^{b_i} - 1\right) \le E_T, \ b_i \in \mathbb{Z}_0^+, \ 1 \le i \le N, \quad (3.5)$$

where $\tilde{\delta} := \delta + \sigma_v^2$ and $\mathbf{z} := [z_1 \cdots z_N]^T$ with $p(\mathbf{z})$ denoting the associated distribution. To solve (3.5), the first step is to find an analytic expression of the equivalent mean MSE metric. Since $z_i \sim \chi_1^2$ is i.i.d. and $p_{\chi_1^2}(z) = \frac{1}{\sqrt{2\pi z}} \exp(-z/2) u(z)$ [6, p-24], where $u(z)$ denotes the unit step function, it can be shown that (see [12] for a proof)

$$\int_{\mathbf{z}} \sum_{i=1}^{N} \frac{1}{\tilde{\delta} + \alpha z_i + \beta 4^{-b_i}} p(\mathbf{z}) \, d\mathbf{z} = \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{N} \int_0^{\infty} \frac{e^{-z_i/2}}{\left(\alpha z_i + \tilde{\delta} + \beta 4^{-b_i}\right)\sqrt{z_i}} \, dz_i$$
$$= \frac{2\pi \cdot e^{\left(\tilde{\delta} + \beta 4^{-b_i}\right)/2\alpha} \cdot Q\left(\sqrt{\left(\tilde{\delta} + \beta 4^{-b_i}\right)/\alpha}\right)}{\sqrt{\alpha\left(\tilde{\delta} + \beta 4^{-b_i}\right)}}, \quad (3.6)$$

where $Q(x) := \int_x^{\infty} \frac{e^{-t^2/2}}{\sqrt{2\pi}} \, dt$ is the Gaussian tail function. Based on (3.6), problem (3.5) can be equivalently rewritten as

$$\text{Maximize } \sqrt{2\pi} \cdot \sum_{i=1}^{N} \frac{e^{\left(\tilde{\delta} + \beta 4^{-b_i}\right)/2\alpha} \cdot Q\left(\sqrt{\left(\tilde{\delta} + \beta 4^{-b_i}\right)/\alpha}\right)}{\sqrt{\alpha\left(\tilde{\delta} + \beta 4^{-b_i}\right)}},$$
$$\text{under } \sum_{i=1}^{N} w_i\left(2^{b_i} - 1\right) \le E_T, \text{ and } b_i \in \mathbb{Z}_0^+, \ \forall \ i. \quad (3.7)$$

The optimization problem (3.7) appears rather formidable to tackle because the cost function is highly nonlinear in b_i. In what follows we will propose an alternative formulation which is more tractable and can yield an analytic solution.

B. Alternative Formulation

The proposed approach is grounded on the following approximation to $Q(\cdot)$ function [8, p-115]:

$$Q(x) \approx \frac{1}{\sqrt{2\pi}} \left[\frac{e^{-x^2/2}}{(1 - \pi^{-1})x + \pi^{-1}\sqrt{x^2 + 2\pi}} \right]; \quad (3.8)$$

the approximation (3.8) is quite accurate since the peak relative error is less than 1.2% for $x \ge 0$, and is almost identical to zero whenever $x \ge 5$. Based on (3.8) together with some straightforward manipulations, the cost function in (3.7) can be well approximated by

a. As in [1], [7], and [13] we assume orthogonal channel access among all the sensor-to-fusion links, which can be realized via, e.g., TDMA or CDMA with orthogonal spreading.

$$\sqrt{2\pi}\cdot\sum_{i=1}^{N}\frac{e^{(\tilde{\delta}+\beta 4^{-b_i})/2\alpha}\cdot Q\left(\sqrt{\left(\tilde{\delta}+\beta 4^{-b_i}\right)/\alpha}\right)}{\sqrt{\alpha\left(\tilde{\delta}+\beta 4^{-b_i}\right)}}$$
$$\approx\sum_{i=1}^{N}\frac{1}{(1-\pi^{-1})\left(\tilde{\delta}+\beta 4^{-b_i}\right)+\pi^{-1}\sqrt{\left(\tilde{\delta}+\beta 4^{-b_i}\right)^2+2\pi\alpha\left(\tilde{\delta}+\beta 4^{-b_i}\right)}}. \quad (3.9)$$

The main advantage of the approximation (3.9) is that it can lead to an associated lower bound in a more tractable form; through otherwise maximizing this lower bound we can eventually obtain a closed-form solution. More specifically, it can be shown that (see [12])

$$\sum_{i=1}^{N}\frac{1}{(1-\pi^{-1})\left(\tilde{\delta}+\beta 4^{-b_i}\right)+\pi^{-1}\sqrt{\left(\tilde{\delta}+\beta 4^{-b_i}\right)^2+2\pi\alpha\left(\tilde{\delta}+\beta 4^{-b_i}\right)}}$$
$$\geq\sum_{i=1}^{N}\frac{1}{(1-\pi^{-1})\left(\tilde{\delta}+\beta 4^{-b_i}\right)+\pi^{-1}\left[\left(\tilde{\delta}+\beta 4^{-b_i}\right)+\pi\alpha\right]}$$
$$=\sum_{i=1}^{N}\frac{1}{\left(\tilde{\delta}+\beta 4^{-b_i}\right)+\alpha}=\sum_{i=1}^{N}\frac{4^{b_i}}{\beta+(\alpha+\tilde{\delta})4^{b_i}}. \quad (3.10)$$

Based on (3.10) we will instead focus on maximizing the lower bound, and thus reformulate the optimization problem as

$$\text{Maximize}\ \sum_{i=1}^{N}\frac{4^{b_i}}{\beta+(\alpha+\tilde{\delta})4^{b_i}},\ \text{s.t.}\ \sum_{i=1}^{N}w_i\left(2^{b_i}-1\right)\leq E_T,\ \text{and}$$
$$b_i\in\mathbb{Z}_0^+,\ 1\leq i\leq N. \quad (3.11)$$

To facilitate analysis we first observe that, since $b_i\in\mathbb{Z}_0^+$, it follows $\sum_{i=1}^{N}w_i\left(2^{b_i}-1\right)\leq\sum_{i=1}^{N}w_i\left(4^{b_i}-1\right)$: this implies we can replace the total energy constraint in (3.11) by the following one without violating the overall energy budget requirement:

$$\sum_{i=1}^{N}w_i\left(4^{b_i}-1\right)\leq E_T. \quad (3.12)$$

With the aid of (3.12) and by performing a change of variable with $B_i:=4^{b_i}-1$, the optimization problem becomes

$$\text{Maximize}\ \sum_{i=1}^{N}\frac{B_i+1}{(\alpha+\beta+\tilde{\delta})+(\alpha+\tilde{\delta})B_i},$$
$$\text{subject to}\ \sum_{i=1}^{N}w_iB_i\leq E_T,\ \text{and}\ B_i\geq 0,\ 1\leq i\leq N. \quad (3.13)$$

In (3.13), the intermediate variable B_i is relaxed to be a nonnegative real number so as to render the problem tractable; once the optimal real-valued B_i (and hence b_i) is computed, the associated bit loads can be obtained through upper integer rounding, as in [7], [11], [13]. The major advantage of the alternative problem formulation (3.13) is that it admits the form of convex optimization and can moreover lead to a closed-form solution, as is shown next.

C. Optimal Solution

Based on the standard Lagrainge techniques, the optimal solution to (3.13) can be obtained as follows (see [12] for detailed proof). Let us assume $w_1\geq w_2\geq\cdots\geq w_N$ without loss of generality, and define the function

$$f(K):=\frac{E_T\left(1+\frac{\beta}{\alpha+\tilde{\delta}}\right)^{-1}+\sum_{j=K}^{N}w_j}{\sqrt{w_K}\sum_{j=K}^{N}\sqrt{w_j}},\ 1\leq K\leq N. \quad (3.14)$$

Let $1\leq K_1\leq N$ be the unique integer such that $f(K_1-1)<1$ and $f(K_1)\geq 1$; if $f(K)\geq 1$ for all $1\leq K\leq N$, then simply set $K_1=1$ (the existence and uniqueness of such K_1 when otherwise is shown in [12]). Then the optimal solution pair $\left(\lambda^{opt},B_i^{opt}\right)$ is given by

$$\sqrt{\lambda^{opt}}=\frac{\sqrt{\beta}}{\alpha+\tilde{\delta}}\left(\sum_{j=K_1}^{N}\sqrt{w_j}\right)\left(E_T+\left(1+\frac{\beta}{\alpha+\tilde{\delta}}\right)\sum_{j=K_1}^{N}w_j\right)^{-1}, \quad (3.15)$$

and

$$B_i^{opt}=\begin{cases}0, & 1\leq i\leq K_1-1,\\ \frac{1}{\alpha+\tilde{\delta}}\sqrt{\frac{\beta}{\lambda^{opt}w_i}}-\left(1+\frac{\beta}{\alpha+\tilde{\delta}}\right), & K_1\leq i\leq N.\end{cases} \quad (3.16)$$

With $B_i=4^{b_i}-1$ and $\tilde{\delta}=\delta+\sigma_v^2$, the optimal bit load b_i^{opt} can be directly obtained from (3.16); the resultant average distortion level is thus (cf. (3.7))

$$\overline{MSE}=\left(\sqrt{2\pi}\cdot\sum_{i=K_1}^{N}\frac{e^{(\delta+\sigma_v^2+\beta 4^{-b_i^{opt}})/2\alpha}\cdot Q\left(\sqrt{\left(\delta+\sigma_v^2+\beta 4^{-b_i^{opt}}\right)/\alpha}\right)}{\sqrt{\alpha\left(\delta+\sigma_v^2+\beta 4^{-b_i^{opt}}\right)}}\right)^{-1} \quad (3.17)$$

IV. DISCUSSIONS AND SIMULATION

1. We note that the minimal achievable average MSE is attained whenever all the raw sensor measurements with infinite-precision are available to the FC (i.e., the case when $b_i=\infty$, $1\leq i\leq N$). Hence, by setting $b_i=\infty$ in the mean MSE formula specified in (3.7), we have the following performance bound
$$MSE_{\min}=\left[Ne^{(\delta+\sigma_v^2)/2\alpha}Q\left(\sqrt{(\delta+\sigma_v^2)/\alpha}\right)\sqrt{\frac{2\pi}{\alpha(\delta+\sigma_v^2)}}\right]^{-1}. \quad (4.1)$$
Formula (4.1) reveals the impacts of the noise model parameters α and δ on the estimation performance. Specifically, it is easy to see from (4.1) that the minimal MSE increases with α: this implies the estimation accuracy degrades as the sensing environment becomes more and more inhomogeneous. Furthermore, it can be checked that $MSE_{\min}$ also increases with the minimal noise power threshold δ. This is reasonable since a large δ implies poor measurement quality of *all* sensor data, and hence a less accurate parameter estimate. We note that, although these facts are inferred based on the idealized distortion measure (4.1), a similar tendency is also observed for $\overline{MSE}$ in (3.17) attained with sensor data quantization (see the numerical results below).
2. Recall from (3.2) that the energy density factor w_i is proportional to the path loss gain d_i^{κ} (assuming $\kappa_i=\kappa$ throughout all links). Large values of w_i, therefore, correspond to sensors deployed far away from the FC (with large d_i), usually with poor background channel gains. In light of this point, the proposed optimal solution (3.16) is intuitively attractive: sensors associated with the (K_1-1) th largest w_i's are turned off to conserve energy. We note that a similar energy conservation strategy via shutting off sensors alone poor channel links is also found

in [7], [11], [13]. Also, we further note from (3.16) that, for those active nodes, the assigned message length is inversely proportional to $\sqrt{w_i}$: this is intuitively reasonable since sensors with better link conditions should be allocated with more bits (energy) to improve the estimation accuracy.

3. We compare the simulated performance of the proposed solution (3.16) against the uniform energy allocation scheme with bit load determined through

$$w_i\left(2^{b_i}-1\right)=E_T/N\,,\quad 1\le i\le N\,. \tag{4.2}$$

In each run we simply choose $w_i=d_i^{\kappa}$ with $\kappa=2$, and d_i's are uniformly drawn from the interval $[1,10]$ as in [13]. In the following experiments we set the number of sensors to be $N=200$, link noise $\sigma_v^2=0.05$, and consider three different levels of total energy: $E_T=\gamma\sum_{i=1}^{N}w_i$ with $\gamma=0.25,\ 1,\ 3$, which respectively correspond to the low, medium, and high energy regimes. With fixed $\delta=2$, Figure 1-(a) shows the computed mean MSE as α varies from 0 to 8, whereas Figure 1-(b) depicts the MSE for fixed $\alpha=2$ and $0.5\le\delta\le 8$. The results show that, as expected, the estimation accuracy improves as E_T increases. Also, the proposed solution (3.16) outperforms (4.2), especially when E_T is small; it is thus more effective in an energy-limited environment. We finally note that the simulated MSE increases with both α and δ: this coincides with the asserted facts in the previous discussions. Figure 2 further depicts the histogram of the computed bit loads; it appears that a large fraction of the active nodes are assigned with one or two quantization bits (hence with BPSK or 4-QAM modulations adopted).

REFERENCES

[1] V. Aravinthan, S. K. Jayaweera, and K. Al Tarazi, "Distributed estimation in a power constrained sensor networks," *Proc. VTC 2006-Spring*, vol. 3, pp. 1048-1052.

[2] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Communications,* vol. 4, no. 5, pp. 2349-2360, Sep. 2005.

[3] S. Cui, J. Xiao, A. Goldsmith, Z. Q. Luo, and V. Poor, "Energy-efficient joint estimation in sensor networks: Analog vs. digital," *Proc. ICASSP 2005*, vol. IV, pp. 745-748.

[4] S. Cui, J. Xiao, A. Goldsmith, Z. Q. Luo, and V. Poor, "Estimation diversity with multiple heterogeneous sensors," *Proc. ICC 2006*.

[5] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall PTR, 1993.

[6] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Prentice-Hall PTR, 1998.

[7] A. Krasnopeev, J. J. Xiao, and Z. Q. Luo, "Minimum energy decentralized estimation in sensor network with correlated sensor noise," *Proc. ICASSP05'*, vol. 3, pp. 673-676.

[8] A. Leon-Garcia, *Probability and Random Process for Electrical Engineering*, 2nd edition, Addison-Wesley, 1994.

[9] S. J. Orfanidis, *Introduction to Signal Processing*, Prentice-Hall, Inc., 1996.

[10] P. Venkitasubramaniam, G. Mergen, L. Tong, and A. Swami, "Quantization for distributed estimation in large scale sensor networks," *Proc. ICISIP 2005*, pp. 121-127.

[11] J. Y. Wu, Q. Z. Huang, and T. S. Lee, "Minimal energy decentralized estimation based on sensor noise variance statistics," *IEEE Trans. Signal Processing*, to appear, 2008; see also *Proc. ICASSP 2007*, vol. 2, pp. 1001-1004.

[12] J. Y. Wu, Q. Z. Huang, and T. S. Lee, "Energy constrained decentralized best-linear-unbiased estimation via partial sensor noise variance knowledge," *IEEE Signal Processing Letters*, to appear, 2008.

[13] J. J. Xiao, S. Cui, Z. Q. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Trans. Signal Processing*, vol. 54, no. 2, pp. 413-422, Feb. 2006.

[14] J. J. Xiao and Z. Q. Luo, "Universal decentralized estimation in an inhomogeneous sensing environment," *IEEE Trans. Information Theory*, vol. 51, no. 10, pp. 3564-3575, Dec. 2005.

[15] J. J. Xiao, A. Ribeiro, Z. Q. Luo, and G. B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 27-41, July 2006.

[16] Q. Zhao, A. Swami, and L. Tong, "The interplay between signal processing and networking in sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 84-93, July 2006.

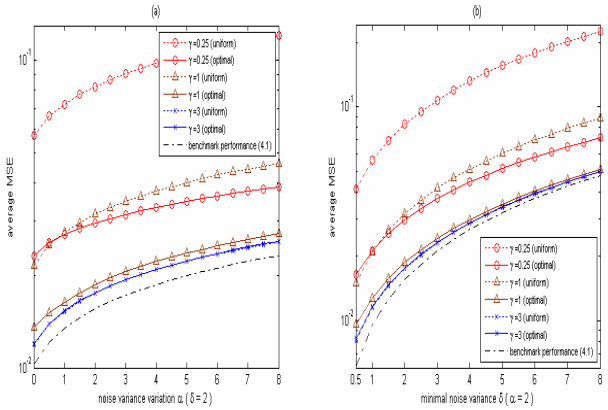

Figure 1. Performance comparison of the proposed solution (3.16) with the uniform allocation scheme (4.2).

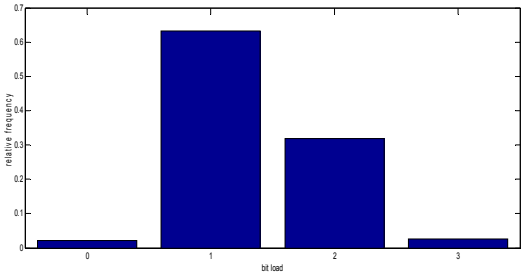

Figure 2. Histogram of the quantization bits.