Optimal Least Squares Deterministic Parameter Estimation from a Class of Block-Circulant-with-Circulant-Block Linear Model

Jwo-Yuh Wu and Ta-Sung Lee
Department of Communication Engineering, National Chiao Tung University,
1001, Ta-Hsueh Road, Hsinchu, Taiwan
Tel: 886-35-712121, ext. 54549; Fax: 886-35-710116
Emails: jywu@cc.nctu.edu.tw; tslee@mail.nctu.edu.tw

Abstract- **This paper investigates the least-squares (LS) estimation of unknown deterministic parameters from a standard linear model characterized by a class of block-circulant-with-circulant-block (BCCB) matrix. We propose a method for designing the BCCB system matrix coefficients to minimize the mean square error incurred by the LS estimate, under certain equality and inequality constraints. By exploiting the eigenvalue characteristic of BCCB matrices, precise analysis is undertaken to derive a closed-form solution. The considered optimization problem arises in the study of blind channel estimation for single-carrier block transmission with cyclic prefix; the presented analysis reveals several key features associated with the BCCB family, and shows an original investigation of the BCCB matrix structure for facilitating linear optimal parameter estimation.**

Index Terms: **Least squares; parameter estimation; circulant matrix; block circulant matrix with circulant blocks; blind channel estimation.**

I. INTRODUCTION

The estimation of linearly mixed parameters subject to additive white Gaussian measurement noise has been addressed in diverse fields in science and engineering. Mathematically, this problem is formulated through the linear model

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}\,, \qquad (1.1)$$

where $\mathbf{y}$ is the observed data vector, $\mathbf{x}$ is the unknown signal of interest, $\mathbf{A}$ is a known matrix of full column rank, and $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \sigma_v^2 \mathbf{I}_k)$, where $\mathbf{I}_k$ denotes the $k \times k$ identity matrix for some k. There have been many criteria for reliably estimating $\mathbf{x}$ based on (1), depending on the priori knowledge known about $\mathbf{x}$ [8], [14]. When $\mathbf{x}$ is treated as deterministic, one popular solution scheme is the least-squares (LS) estimate, namely,

$$\hat{\mathbf{x}}_{LS} := \left(\mathbf{A}^H \mathbf{A}\right)^{-1} \mathbf{A}^H \mathbf{y}\,. \qquad (1.2)$$

Despite its simplicity, the LS solution is attractive for it produces the optimal linear unbiased estimate under white noise assumption [8], [12]. To assess the performance of $\hat{\mathbf{x}}_{LS}$, one commonly used metric is the mean square error (MSE), namely, $E\left\{\left\|\hat{\mathbf{x}}_{LS} - \mathbf{x}\right\|_2^2\right\}$, where $E\{\cdot\}$ is the expectation operator. By equivalently rewriting $\hat{\mathbf{x}}_{LS}$ in (1.2) as

$$\hat{\mathbf{x}}_{LS} = \mathbf{x} + \left(\mathbf{A}^H \mathbf{A}\right)^{-1} \mathbf{A}^H \mathbf{v} \qquad (1.3)$$

and since $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \sigma_v^2 \mathbf{I})$, it is straightforward to show

$$E\left\{\left\|\hat{\mathbf{x}}_{LS} - \mathbf{x}\right\|_2^2\right\} = \sigma_v^2 Tr\left[\left(\mathbf{A}^H \mathbf{A}\right)^{-1}\right], \qquad (1.4)$$

where $Tr[\cdot]$ denotes the trace. The performance of the LS solution thus depends crucially on the matrix $\mathbf{A}$. In many situations it is plausible to judiciously choose $\mathbf{A}$ to improve the utmost solution reliability, e.g., in training based channel estimation the matrix $\mathbf{A}$ contains the pilot symbols and are designed to minimize the MSE in (1.4) under a total transmit power constraint [9], [11].

In this paper, we focus on a special class of linear model, wherein the matrix $\mathbf{A}$ is described by

$$\mathbf{A} = \mathbf{Q}\mathbf{T}\,, \qquad (1.5)$$

in which $\mathbf{T}$ contains an arbitrary column subset of the $N^2 \times N^2$ identity matrix, $\mathbf{Q}$ is a block-circulant-with-circulant-blocks (BCCB) matrix [3, p-184] of dimension $N^2 \times N^2$:

$$\mathbf{Q} := \begin{bmatrix} p(0)^2 \mathbf{I}_N & p(1)^2 \mathbf{J} & \cdots & p(N-2)^2 \mathbf{J}^{N-2} & p(N-1)^2 \mathbf{J}^{N-1} \\ p(N-1)^2 \mathbf{J}^{N-1} & p(0)^2 \mathbf{I}_N & \cdots & p(N-3)^2 \mathbf{J}^{N-3} & p(N-2)^2 \mathbf{J}^{N-2} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ p(2)^2 \mathbf{J}^2 & p(3)^2 \mathbf{J}^3 & \cdots & p(0)^2 \mathbf{I}_N & p(1)^2 \mathbf{J} \\ p(1)^2 \mathbf{J} & p(2)^2 \mathbf{J}^2 & \cdots & p(N-1)^2 \mathbf{J}^{N-1} & p(0)^2 \mathbf{I}_N \end{bmatrix} \qquad (1.6)$$

where $p(n)$, $0 \le n \le N-1$, are some positive real numbers, and $\mathbf{J} \in \mathbb{R}^{N \times N}$ is the circulant permutation matrix defined by

$$\mathbf{J} := \begin{bmatrix} \mathbf{0}_{1\times(N-1)} & 1 \\ \mathbf{I}_{N-1} & \mathbf{0}_{(N-1)\times 1} \end{bmatrix}. \qquad (1.7)$$

We will seek for the optimal $p(n)$ which minimizes

$$MSE := \sigma_v^2 Tr\left[\left(\mathbf{T}^H \mathbf{Q}^H \mathbf{Q} \mathbf{T}\right)^{-1}\right], \qquad (1.8)$$

subject to the following two design constraints

$$\sum_{n=0}^{N-1} p(n)^2 = N\,, \qquad (1.9)$$

$$\min\; p(n)^2 \ge \delta \;\text{ for some }\; 0 < \delta < 1\,. \qquad (1.10)$$

Such a problem arises in the study of blind channel estimation for single-carrier transmission with cyclic prefix [5], [16] and also the related multi-antenna system with space-time block coding [2], [17]. By exploiting the eigen-structure of the BCCB matrix $\mathbf{Q}$, in this paper we propose a method for constructing a closed-form optimal $p(n)$. Block circulant matrices (not necessarily with circulant blocks) have found important applications in computational reduction [4], [13], [15], and in the study of spectral distribution for Toeplitz matrices [1]. The presented study in this paper brings out the nice features of the block circulant family in optimal linear parameter estimation.

1-4244-0504-1/06/$20.00 ©2006 IEEE

II. PROPOSED OPTIMAL SOLUTION

A. Design Approach

Minimization of the cost function of the form (1.8) has been considered in [9], [11]. The reported solution approach therein is via the following inequality: since $\mathbf{T}^H\mathbf{Q}^H\mathbf{Q}\mathbf{T}$ are positive definite, it follows

$$Tr\left[\left(\mathbf{T}^H\mathbf{Q}^H\mathbf{Q}\mathbf{T}\right)^{-1}\right] \geq \sum_i \left[\mathbf{T}^H\mathbf{Q}^H\mathbf{Q}\mathbf{T}\right]_{i,i}^{-1}, \qquad (2.1)$$

and equality holds whenever $\mathbf{T}^H\mathbf{Q}^H\mathbf{Q}\mathbf{T}$ is diagonal [10, p-1041]. If equality (1.9) is the only design concern, it is easy to check that the impulse sequence

$$p(m)^2 = N, \text{ and } p(n)^2 = 0 \text{ for } n \neq m, \qquad (2.2)$$

where $0 \leq m \leq N-1$ is fixed but arbitrary, diagonalizes $\mathbf{T}^H\mathbf{Q}^H\mathbf{Q}\mathbf{T}$ and is thus the minimizing solution. However, given the additional threshold requirement (1.10), 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 $\mathbf{T}^H\mathbf{Q}^H\mathbf{Q}\mathbf{T}$ diagonal. Another possible solution scheme would be via numerical search techniques. However, as the cost function in (1.8) are non-convex in $p(n)$, there do not seem to have efficient methods for finding the global optimum. 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 as we show below will facilitate the exploitation of the BCCB property of $\mathbf{Q}$ to derive a closed-form solution.

Lemma 2.1: 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

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

[Proof]: Without loss of generality we assume $\mathbf{M}$ is split as $\mathbf{M} = \begin{bmatrix}\bar{\mathbf{M}} & \mathbf{M}_d\end{bmatrix}$, 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

$$\mathbf{M}^H\mathbf{M} = \begin{bmatrix}\bar{\mathbf{M}}^H \\ \mathbf{M}_d^H\end{bmatrix}\begin{bmatrix}\bar{\mathbf{M}} & \mathbf{M}_d\end{bmatrix} = \begin{bmatrix}\bar{\mathbf{M}}^H\bar{\mathbf{M}} & \bar{\mathbf{M}}^H\mathbf{M}_d \\ \mathbf{M}_d^H\bar{\mathbf{M}} & \mathbf{M}_d^H\mathbf{M}_d\end{bmatrix}. \qquad (2.3)$$

Since $\mathbf{M}$ is nonsingular, $\mathbf{M}^H\mathbf{M}$ is positive definite. By the inversion lemma for block matrix [8, p-572], we have

$$\left(\mathbf{M}^H\mathbf{M}\right)^{-1} = \begin{bmatrix}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}} - \bar{\mathbf{M}}^H\mathbf{M}_d\left(\mathbf{M}_d^H\mathbf{M}_d\right)^{-1}\mathbf{M}_d^H\bar{\mathbf{M}}\right)^{-1} & \times \\ \times & \left(\mathbf{M}_d^H\mathbf{M}_d - \mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\right)^{-1}\end{bmatrix} \qquad (2.4)$$

in which the notation "$\times$" stands for the block off-diagonal submatrices irrelevant to the proof procedures. From (2.4), we have

$$Tr\left[\left(\mathbf{M}^H\mathbf{M}\right)^{-1}\right] = Tr\left[\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}} - \bar{\mathbf{M}}^H\mathbf{M}_d\left(\mathbf{M}_d^H\mathbf{M}_d\right)^{-1}\mathbf{M}_d^H\bar{\mathbf{M}}\right)^{-1}\right] + Tr\left[\left(\mathbf{M}_d^H\mathbf{M}_d - \mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\right)^{-1}\right] \qquad (2.5)$$

Since $\left(\mathbf{M}^H\mathbf{M}\right)^{-1}$ is positive definite, so are its principle submatrices and (2.5) implies

$$Tr\left[\left(\mathbf{M}^H\mathbf{M}\right)^{-1}\right] \geq Tr\left[\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}} - \bar{\mathbf{M}}^H\mathbf{M}_d\left(\mathbf{M}_d^H\mathbf{M}_d\right)^{-1}\mathbf{M}_d^H\bar{\mathbf{M}}\right)^{-1}\right]. \qquad (2.6)$$

Using the matrix inversion lemma [8, p-571], inequality (2.6) can be further expanded into

$$\begin{aligned} Tr\left[\left(\mathbf{M}^H\mathbf{M}\right)^{-1}\right] &\geq Tr\left[\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}} - \bar{\mathbf{M}}^H\mathbf{M}_d\left(\mathbf{M}_d^H\mathbf{M}_d\right)^{-1}\mathbf{M}_d^H\bar{\mathbf{M}}\right)^{-1}\right] \\ &= Tr\left[\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1} + \left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\left(\mathbf{M}_d^H\mathbf{M}_d - \mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\right)^{-1}\mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\right] \\ &= Tr\left[\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\right] + Tr\left[\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\left(\mathbf{M}_d^H\mathbf{M}_d - \mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\bar{\mathbf{M}}^H\mathbf{M}_d\right)^{-1}\mathbf{M}_d^H\bar{\mathbf{M}}\left(\bar{\mathbf{M}}^H\bar{\mathbf{M}}\right)^{-1}\right] \end{aligned} \qquad (2.7)$$

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. (2.4)), 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 (2.7). □

As $\mathbf{Q}\mathbf{T}$ contains a column subset of $\mathbf{Q}$, Lemma 2.1 asserts that $Tr\left[\left(\mathbf{T}^H\mathbf{Q}^H\mathbf{Q}\mathbf{T}\right)^{-1}\right]$ 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 designing $p(n)$: we can simply choose $p(n)$ to minimize

$$J := Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right], \qquad (2.8)$$

since $Tr\left[\left(\mathbf{T}^H\mathbf{Q}^H\mathbf{Q}\mathbf{T}\right)^{-1}\right]$ would in turn be kept small. The main advantage of the proposed design formulation is that we can directly take advantage of the BCCB structure of $\mathbf{Q}$ to derive a closed-form solution. Toward this end, we shall first express the cost function J in (2.8) in a more tractable form. Since $Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right]$ is the sum of the N^2 eigenvalues associated with $\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}$ [6, p-42], we propose to rewrite J in terms of such spectral characteristics; this is specified next and will lay the foundation for subsequent analytic design.

B. Eigen-Structure of BCCB Matrix

The eigen-property of the BCCB matrix family has been studied in [3]. A distinctive feature of the BCCB matrices is that they are diagonalizable by FFT based operations. This is pinned down via the following lemma; the result will be used for explicitly computing the eigenvalues of the matrix $\mathbf{Q}$.

We will hereafter denote by $BCCB_{N,N}$ the set of all $N^2 \times N^2$ block circulant matrices with circulant blocks, each characterized by N circulant matrices of dimension $N \times N$, $\mathbf{F}$ the $N \times N$ FFT matrix, and $\otimes$ the Kronecker product [7, p-243].

Lemma 2.2 [3, p-185]: If $\mathbf{X} \in BCCB_{N,N}$, then $\mathbf{X}$ can be diagonalized by $\mathbf{F} \otimes \mathbf{F}$. More precisely, let $\{\mathbf{C}_0, \cdots, \mathbf{C}_{N-1}\}$ 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

$$\mathbf{X} = (\mathbf{F} \otimes \mathbf{F}) \left(\sum_{n=0}^{N-1} \mathbf{\Omega}_N^n \otimes \mathbf{\Lambda}_n \right) \left(\mathbf{F}^{-1} \otimes \mathbf{F}^{-1} \right), \quad (2.9)$$

with $\mathbf{\Omega}_N := diag\left\{ \begin{bmatrix} 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \end{bmatrix}^T \right\}$, $\omega := \exp(j2\pi / N)$.

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 $BCCB_{N,N}$. □

Based on Lemma 2.2, we can determine the eigenvalues of the matrix $\mathbf{Q}$ in (1.6). Observe that, despite its BCCB structure, the circulant block submatrices of $\mathbf{Q}$ are further characterized by $\mathbf{J}^n$: this will lead to very elegant eigenvalue characteristics. Roughly speaking, if we define the vector

$$\mathbf{p} := [p(0)^2 \quad p(1)^2 \quad \cdots \quad p(N-1)^2]^T \in \mathbb{R}^N, \quad (2.10)$$

then the N^2 eigenvalues of $\mathbf{Q}$ is completely determined by the N eigenvalues associated with the $N \times N$ circulant matrix with $\mathbf{p}^T$ as the first row. More precisely, we have the following theorem (see Appendix A for a proof).

Theorem 2.3: Let $\mathbf{F}$ be the $N \times N$ FFT matrix; also, associated with the vector $\mathbf{p}$ in (2.10) we define the polynomial

$$\mathbf{p}(z) := p(0)^2 + p(1)^2 z^{-1} + \cdots + p(N-1)^2 z^{-(N-1)}. \quad (2.11)$$

Then the N^2 eigenvalues of the matrix $\mathbf{Q}$ defined in (1.6) are exactly given by the N replicas of the N-tuple $\left\{ \mathbf{p}(1), \mathbf{p}(\omega), \cdots, \mathbf{p}(\omega^{N-1}) \right\}$. □

Based on Theorem 2.3, it follows immediately that the eigenvalues of $\mathbf{Q}^H\mathbf{Q}$ are N replicas of the N-tuple $\left\{ |\mathbf{p}(1)|^2, |\mathbf{p}(\omega)|^2, \cdots, |\mathbf{p}(\omega^{N-1})|^2 \right\}$; the objective function J in (2.8) is thus

$$J = Tr\left[\left(\mathbf{Q}^H \mathbf{Q} \right)^{-1} \right] = \sum_{k=0}^{N-1} \frac{N}{|\mathbf{p}(\omega^k)|^2}. \quad (2.12)$$

Equation (2.12) rewrites $Tr\left[\left(\mathbf{Q}^H \mathbf{Q} \right)^{-1} \right]$ in terms of the "frequency responses" $\mathbf{p}(\omega^n)$'s in a rather simple way: it is just a scaled sum of $|\mathbf{p}(\omega^k)|^{-2}$ over $0 \le k \le N-1$. The derivation of the optimal $p(n)$ is based on equation (2.12) and is shown below.

C. Optimal Solution

The first step toward a solution is to transform the two constraints (1.9) and (1.10) in terms of $\mathbf{p}(\omega^n)$. With (1.9) and (2.11), it is easy to check that, for $k = 0$,

$$\mathbf{p}(\omega^0) = \mathbf{p}(1) = \sum_{n=0}^{N-1} p(n)^2 = N. \quad (2.13)$$

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 optimal solution (see Appendix B for a proof).

Lemma 2.4: For any $p(n)$ satisfying (1.9) and (1.10), we have

$$\left| \mathbf{p}(\omega^k) \right| \le N(1-\delta) \text{ for all } 1 \le k \le N-1. \quad (2.14)$$

□

With (2.13) and (2.14), the optimal $p(n)$ is shown in the next theorem.

Theorem 2.5: The optimal $p(n)$ minimizing $Tr\left[\left(\mathbf{Q}^H \mathbf{Q} \right)^{-1} \right]$, subject to constraints (1.9) and (1.10), is given by the following the two-level form solution: for a fixed but arbitrary $0 \le m \le N-1$,

$$p(m)^2 = N - (N-1)\delta, \text{ and } p(n)^2 = \delta \text{ for } n \ne m, \quad (2.15)$$

leading to

$$J_{\min} = \frac{1}{N} + \frac{(N-1)}{N(1-\delta)^2}. \quad (2.16)$$

[Proof]: We claim that *i)* $Tr\left[\left(\mathbf{Q}^H \mathbf{Q} \right)^{-1} \right] \ge J_{\min}$ for any $p(n)$ satisfying (1.9) and (1.10), and *ii)* equality is attained by the two-level scheme (2.15); the theorem thus follows. To show claim *i)*, we observe form (2.12) and (2.13) that

$$Tr\left[\left(\mathbf{Q}^H \mathbf{Q} \right)^{-1} \right] = \sum_{k=0}^{N-1} \frac{N}{|\mathbf{p}(\omega^k)|^2} = \frac{N}{|\mathbf{p}(1)|^2} + \sum_{k=1}^{N-1} \frac{N}{|\mathbf{p}(\omega^k)|^2} = \frac{1}{N} + \sum_{k=1}^{N-1} \frac{N}{|\mathbf{p}(\omega^k)|^2} \quad (2.17)$$

From (2.14), it follows

$$\left| \mathbf{p}(\omega^k) \right|^{-2} \ge \frac{1}{N^2(1-\delta)^2}, \ 1 \le k \le N-1, \quad (2.18)$$

With (2.17) and (2.18), we have

$$Tr\left[\left(\mathbf{Q}^H \mathbf{Q} \right)^{-1} \right] = \frac{1}{N} + \sum_{k=1}^{N-1} \frac{N}{|\mathbf{p}(\omega^k)|^2} \ge \frac{1}{N} + \frac{N-1}{N(1-\delta)^2}, \quad (2.19)$$

which proves claim *i)*. To show claim *ii)*, it is noted that the two-level solution (2.15) yields, for any $k \ne 0$,

$$\mathbf{p}(\omega^k) = \sum_{n=0}^{N-1} p(n)^2 \omega^{-kn} = \{N - (N-1)\delta\} \omega^{-km} + \delta \sum_{n \ne m} \omega^{-kn} = \{N(1-\delta)\} \omega^{-km} + \delta \sum_{n=0}^{N-1} \omega^{-kn} = \{N(1-\delta)\} \omega^{-km}, \quad (2.20)$$

where the last equality in (2.20) follows since $\sum_{n=0}^{N-1} \omega^{-kn} = 0$ for any $k \ne 0$. Equations (2.13) and (2.20) show $\left| \mathbf{p}(1) \right| = N$ and $\left| \mathbf{p}(\omega^k) \right| = N(1-\delta)$ for $1 \le k \le N-1$; hence the two-level scheme (2.15) attains $J_{\min}$ in (2.16). □

Recall that the impulse sequence (2.2) is optimal whenever (1.9) is the only design concern. When an additional threshold on the magnitude of $p(n)$ is imposed as in (1.10), it turns out that the best choice is the "impulse-like" two-level solution (2.15). With (2.16), the minimal $J_{\min}$ is seen to decrease whenever δ is decreased. It is noted that two-level solution (2.15) minimizes $Tr\left[\left(\mathbf{Q}^H\mathbf{Q}\right)^{-1}\right]$, but its optimality with respect to $Tr\left[\left(\mathbf{T}^H\mathbf{Q}^H\mathbf{Q}\mathbf{T}\right)^{-1}\right]$ appears intractable to verify. Our simulation results indicate that it indeed seems to be the minimizing solution, irrespective of the choices of $\mathbf{T}$.

III. CONCLUSION

In this paper we investigate the optimal LS estimation from a class of BCCB linear model, which is encountered in our recent study in blind channel estimation problems. We show a method for designing the system matrix coefficients to minimize MSE under certain equality and inequality constraints. The proposed approach minimizes an upper bound on MSE, and exploits the BCCB system matrix structure as well as the associated spectral characteristics. The frequency-domain-based formulation in terms of eigenvalues nicely tackles the inequality constraint, allows precise analysis procedures, and eventually leads to an appealing simple closed-form solution. We will try to generalize the results to less restricted families like block circulant matrices or matrices with circulant blocks; this would find potential applications in channel estimation for MIMO cyclic prefix based single- (multi-) carrier block transmission.

APPENDIX A: PROOF OF THEOREM 2.3

The matrix $\mathbf{Q}$ in (1.6) 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. Let $\mathbf{u}_n \in \mathbb{C}^N$ be the vector containing the eigenvalues of the matrix $p(n)^2\mathbf{J}^n$, $0 \le n \le N-1$. We then have [3, p-73], for $0 \le n \le N-1$,

$$\mathbf{u}_n = \sqrt{N}\cdot\mathbf{F}^{-1}\mathbf{r}_n^T, \tag{A.1}$$

where $\mathbf{r}_n$ denotes the first row of $p(n)^2\mathbf{J}^n$. By definition of $\mathbf{J}$ in (1.7), it can be deduced that, for $1 \le n \le N-1$,

$$\mathbf{r}_0 = p(0)^2\mathbf{e}_1^T \text{ and } \mathbf{r}_n = p(n)^2\mathbf{e}_{N+1-n}^T, \tag{A.2}$$

where $\mathbf{e}_l$ denotes the l th unit standard vector in $\mathbb{R}^N$. From (A.1) and (A.2), it follows that, for $1 \le n \le N-1$,

$$\mathbf{u}_0 = \sqrt{N}\cdot p(0)^2\mathbf{f}_0 \text{ and } \mathbf{u}_n = \sqrt{N}\cdot p(n)^2\mathbf{f}_{N-n}, \tag{A.3}$$

where

$$\mathbf{f}_n := \sqrt{N^{-1}}\cdot[1 \quad \omega^n \quad \omega^{2n} \quad \cdots \quad \omega^{(N-2)n} \quad \omega^{(N-1)n}]^T \tag{A.4}$$

is the $(n+1)$ th column of $\mathbf{F}^{-1}$, $0 \le n \le N-1$. By Lemma 2.1, the eigenvalues of $\mathbf{Q}$ are given by the diagonal entries of the matrix $\sum_{n=0}^{N-1}\mathbf{\Omega}_N^n \otimes \mathbf{\Lambda}_n$, where $\mathbf{\Lambda}_n = diag\{\mathbf{u}_n\}$. Since $\mathbf{\Omega}_N^n = \sqrt{N}\cdot diag\{\mathbf{f}_n\}$ (this follows by definition of $\mathbf{\Omega}_N$ and from (A.4)), the eigenvalues of $\mathbf{Q}$ can simply be computed as entries of the vector $\sqrt{N}\cdot\sum_{n=0}^{N-1}\mathbf{f}_n\otimes\mathbf{u}_n$. From (A.3) and (A.4), it can be seen that

$$\mathbf{f}_0\otimes\mathbf{u}_0 = p(0)^2[\mathbf{f}_0^T \quad \mathbf{f}_0^T \quad \cdots \quad \mathbf{f}_0^T]^T, \tag{A.5}$$

and, for $1 \le n \le N-1$,

$$\mathbf{f}_n\otimes\mathbf{u}_n = p(n)^2[\mathbf{f}_{N-n}^T \quad \omega^n\mathbf{f}_{N-n}^T \quad \cdots \quad \omega^{(N-1)n}\mathbf{f}_{N-n}^T]^T. \tag{A.6}$$

From (A.5) and (A.6), the vector $\sqrt{N}\cdot\sum_{n=0}^{N-1}\mathbf{f}_n\otimes\mathbf{u}_n$ can be computed as

$$\sqrt{N}\begin{bmatrix} p(0)^2\mathbf{f}_0 + \sum_{n=1}^{N-1} p(n)^2\mathbf{f}_{N-n} \\ p(0)^2\mathbf{f}_0 + \sum_{n=1}^{N-1}\omega^n p(n)^2\mathbf{f}_{N-n} \\ \vdots \\ p(0)^2\mathbf{f}_0 + \sum_{n=1}^{N-1}\omega^{(N-1)n}p(n)^2\mathbf{f}_{N-n}\end{bmatrix} = \sqrt{N}\begin{bmatrix}\mathbf{F}^{-1}\mathbf{\Gamma}\mathbf{p} \\ \mathbf{F}^{-1}\mathbf{\Omega}_N\mathbf{\Gamma}\mathbf{p} \\ \vdots \\ \mathbf{F}^{-1}\mathbf{\Omega}_N^{N-1}\mathbf{\Gamma}\mathbf{p}\end{bmatrix} \tag{A.7}$$

where $\mathbf{p}$ is given in (2.10) and $\mathbf{\Gamma}\in\mathbb{R}^{N\times N}$ is the Hankel matrix with $[1\ 0\cdots 0]^T$ as the first column and $[0\ 1\ 0\cdots 0]$ as the last row. It can be checked by definition that $\mathbf{F}^{-1}\mathbf{\Omega}_N^n = (\mathbf{J}^T)^n\mathbf{F}^{-1}$ and hence

$$\sqrt{N}\cdot\mathbf{F}^{-1}\mathbf{\Omega}_N^n\mathbf{\Gamma}\mathbf{p} = (\mathbf{J}^T)^n\cdot\sqrt{N}\cdot\mathbf{F}^{-1}\mathbf{\Gamma}\mathbf{p},\ 0 \le n \le N-1. \tag{A.8}$$

From (A.7) and (A.8) we can see that, for $2 \le i \le N$, the entries of the i th N-dimensional block of $\sqrt{N}\cdot\sum_{n=0}^{N-1}\mathbf{f}_n\otimes\mathbf{u}_n$ are simply a permuted version of those in the first one, namely, $\sqrt{N}\cdot\mathbf{F}^{-1}\mathbf{\Gamma}\mathbf{p}$. As a result, the N^2 eigenvalues of $\mathbf{Q}$ thus assume N distinct values only. Since $\mathbf{\Gamma} = \mathbf{F}^2$ [3, p-33], we have $\sqrt{N}\cdot\mathbf{F}^{-1}\mathbf{\Gamma}\mathbf{p} = \sqrt{N}\cdot\mathbf{F}\mathbf{p}$. The assertion thus follows by definition of $\mathbf{F}$. □

APPENDIX B: PROOF OF LEMMA 2.4

The assertion relies on the following key observation: any given $p(n)$ satisfying (1.9) and (1.10) can be constructed by "squeezing" the peak value at $n = m$ of the two-level solution (2.15) so that the ground values at other n's 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

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

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

where $\Delta_n > 0$ models the excessive power over the ground level δ for $n \in \mathcal{I}$. The sequence of the form (B.1) and (B.2) satisfies the constraints (1.9) and (1.10); in particular, since $p(m)^2 \geq \delta$ is required, we can infer from (B.1) that

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

We assume for the moment that $m = 0$; as one will see, the result for the $1 \leq m \leq N-1$ case easily follows. Associated with $p(n)$ in (B.1) and (B.2), we have, for $1 \leq k \leq N-1$,

$$\begin{aligned}
\mathbf{p}(\omega^k) &= \sum_{n=0}^{N-1} p(n)^2 \omega^{-kn} \\
&= \left[N-(N-1)\delta - \sum_{n\in\mathcal{I}} \Delta_n\right] + \sum_{n\in\mathcal{I}} (\delta + \Delta_n)\omega^{-kn} + \sum_{n\notin\mathcal{I}} \delta\omega^{-kn} \\
&= \left[N(1-\delta) + \sum_{n\in\mathcal{I}} (\omega^{-kn}-1)\Delta_n\right] + \underbrace{\left[\delta + \sum_{n\in\mathcal{I}} \delta\omega^{-kn} + \sum_{n\notin\mathcal{I}} \delta\omega^{-kn}\right]}_{=\delta\sum_{n=0}^{N-1}\omega^{-kn}=0} \\
&= \left[N(1-\delta) + \sum_{n\in\mathcal{I}} (\omega^{-kn}-1)\Delta_n\right].
\end{aligned} \tag{B.4}$$

Define the nonnegative number

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

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

$$\begin{aligned}
\left|\mathbf{p}(\omega^k)\right|^2 &= \left[d + \sum_{n\in\mathcal{I}} \Delta_n \cos n\theta_k\right]^2 + \left[\sum_{n\in\mathcal{I}} \Delta_n \sin n\theta_k\right]^2 \\
&= d^2 + 2d\left(\sum_{n\in\mathcal{I}} \Delta_n \cos n\theta_k\right) + \left(\sum_{n\in\mathcal{I}} \Delta_n \cos n\theta_k\right)^2 \\
&\quad + \left(\sum_{n\in\mathcal{I}} \Delta_n \sin n\theta_k\right)^2.
\end{aligned} \tag{B.6}$$

Observe that

$$\begin{aligned}
&\left(\sum_{n\in\mathcal{I}} \Delta_n \cos n\theta_k\right)^2 + \left(\sum_{n\in\mathcal{I}} \Delta_n \sin n\theta_k\right)^2 \\
&= \sum_{n\in\mathcal{I}} \Delta_n^2 + 2\sum_{n_l,n_m\in\mathcal{I}} \Delta_{n_l}\Delta_{n_m}(\cos n_l\theta_k \cos n_m\theta_k + \sin n_l\theta_k \sin n_m\theta_k) \\
&= \sum_{n\in\mathcal{I}} \Delta_n^2 + 2\sum_{n_l,n_m\in\mathcal{I}} \Delta_{n_l}\Delta_{n_m} \cos(n_l - n_m)\theta_k \\
&\leq \sum_{n\in\mathcal{I}} \Delta_n^2 + 2\sum_{n_l,n_m\in\mathcal{I}} \Delta_{n_l}\Delta_{n_m} = \left(\sum_{n\in\mathcal{I}} \Delta_n\right)^2,
\end{aligned} \tag{B.7}$$

and that

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

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

$$\begin{aligned}
\left|\mathbf{p}(\omega^k)\right|^2 &\leq d^2 + 2d\left(\sum_{n\in\mathcal{I}} \Delta_n\right) + \left(\sum_{n\in\mathcal{I}} \Delta_n\right)^2 \\
&= \left(d + \sum_{n\in\mathcal{I}} \Delta_n\right)^2 = N^2(1-\delta)^2,
\end{aligned} \tag{B.9}$$

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

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

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

ACKNOWLEDGMENT

This work is sponsored by the National Science Council of Taiwan under joint grants NSC 95-2752-E-002-009, NSC-95-2623-7-009-008-D, and the MoE ATU Program.

REFERENCES

[1] N. K. Bose and H. J. Boo, "Asymptotic eigenvalue distribution of block Toeplitz matrix," *IEEE Tran. Information Theory*, vol. 44, no. 2, pp. 858-861, March 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] P. J. Davis, *Circulant Matrices*, John Wieley & Sons, Inc, 1979.

[4] T. De Mazancourt and D. Gerlic, "The inverse of a block-circulant matrix," *IEEE Trans. Antennas and Propagation*, vol. AP-31, no. 5, pp. 808-810, Sept. 1983.

[5] 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.

[6] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

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

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

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

[10] 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.

[11] J. H. Manton, "Optimal training sequences and pilot tones for OFDM systems," *IEEE Communications Letters*, vol. 5, no. 4, pp. 151-153, April 2001.

[12] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, 2000.

[13] Y. Tang, B. Vecutic, and Y. Li, "An FFT-based multiuser detection for asynchronous block-spreading CDMA ultra wideband communication systems," *Proc. IEEE International Conference Communications*, 2005, vol. 5, pp. 2872-2876.

[14] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, *Part I*, Wiley Interscience, New York, 2001.

[15] R. Vescovo, "Inversion of block-circulant matrices and circular array approach," *IEEE Trans. Antennas and Propagation*, vol. 45, no. 10, pp. 1565-1567, Oct. 1997.

[16] 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.

[17] J. Y. Wu and T. S. Lee, "Non-redundant precoding assisted blind channel estimation for single-carrier space-time block coded transmission with frequency-domain equalization," *IEEE Trans. Signal Processing*, to appear, 2007.