

A Matrix Pseudoinversion Lemma and Its Application to Block-Based Adaptive Blind Deconvolution for MIMO Systems

Kiyotaka Kohno, *Member, IEEE*, Mitsuru Kawamoto, *Member, IEEE*, and Yujiro Inouye, *Member, IEEE*

Abstract—The matrix inversion lemma gives an explicit formula of the inverse of a positive definite matrix \mathbf{A} added to a block of dyads (represented as $\mathbf{B}\mathbf{B}^H$) as follows: $(\mathbf{A} + \mathbf{B}\mathbf{B}^H)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{I} + \mathbf{B}^H\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{A}^{-1}$. It is well known in the literature that this formula is very useful to develop a block-based recursive least squares algorithm for the block-based recursive identification of linear systems or the design of adaptive filters. We extend this result to the case when the matrix \mathbf{A} is singular and present a matrix pseudoinversion lemma along with some illustrative examples. Based on this result, we propose a block-based adaptive multichannel superexponential algorithm. We present simulation results for the performance of the block-based algorithm in order to show the usefulness of the matrix pseudoinversion lemma.

Index Terms—Adaptive superexponential algorithm, blind deconvolution, block-based recursive least squares algorithm, matrix pseudoinversion lemma.

I. INTRODUCTION

THE FAMILIAR matrix inversion lemma states that the inverse of a positive definite $n \times n$ matrix \mathbf{A} added to a block of dyads (represented as $\mathbf{B}\mathbf{B}^H$) can be represented as

$$(\mathbf{A} + \mathbf{B}\mathbf{B}^H)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{I} + \mathbf{B}^H\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{A}^{-1} \quad (1)$$

where \mathbf{B} is an $n \times l$ matrix and the superscript H denotes the complex conjugate transpose (or Hermitian) operation. It is widely known in the literature that this formula is very useful to develop a block-based recursive least squares algorithm for the block-based recursive identification [1], [2] or the design of adaptive filters [3].

In the late 1980s, Ogawa extended the matrix inversion lemma in (1) to the case when \mathbf{A} is positive semidefinite [4]. However, his extension is valid under the condition that the range of \mathbf{A} includes the range of \mathbf{B} , but this condition is very restrictive and is not satisfied for adaptive signal processing in a nonstationary environment.

Manuscript received July 31, 2009; revised February 26, 2010; accepted April 16, 2010. Date of current version July 16, 2010. This work was supported in part by the Japan Society for the Promotion of Science Research Projects under Grant-in-Aid 21500088. This paper was recommended by Associate Editor W. X. Zheng.

K. Kohno is with the Department of Electronic Control Engineering, Yonago National College of Technology, Yonago 683-8502, Japan (e-mail: kohno@yonago-k.ac.jp).

M. Kawamoto is with the Information Technology Research Institute, National Institute of Advanced Industrial Science and Technology, Tokyo 135-0064, Japan (e-mail: m.kawamoto@aist.go.jp).

Y. Inouye is with the Department of Electronic and Control Systems Engineering, Shimane University, Matsue 690-8504, Japan (e-mail: inouye@riko.shimane-u.ac.jp).

Digital Object Identifier 10.1109/TCSI.2010.2050222

TABLE I

COMPARISON OF THE NUMBERS OF FLOATING-POINT OPERATIONS [$\times 10^8$]

The length of the block	The length of the deconvolver	The proposed method	The method using "pinv"
$l = 1$	$L = 7$	2.9063	8.8792
$l = 2$	$L = 3$	0.1860	0.4104
$l = 2$	$L = 7$	1.6341	4.4555
$l = 2$	$L = 14$	11.0880	30.8100
$l = 4$	$L = 7$	1.0019	2.2706

TABLE II

COMPARISON OF THE EXECUTION TIMES (IN SECONDS)

The length of the block	The length of the deconvolver	The proposed method	The method using "pinv"
$l = 1$	$L = 7$	0.3413	0.7007
$l = 2$	$L = 3$	0.1012	0.1042
$l = 2$	$L = 7$	0.2017	0.3555
$l = 2$	$L = 14$	0.7152	1.8108
$l = 4$	$L = 7$	0.1246	0.1987

In this paper, we extend the matrix inversion lemma in (1) to the case when the matrix \mathbf{A} is positive semidefinite without the aforementioned condition for the ranges of the relevant matrices and present a matrix pseudoinversion lemma together with some illustrative examples [5]. Such a singular case may occur in a situation where a problem dealt with is overdetermined in the sense that it has more equations than unknowns [6]. In particular, we encountered this singular situation when we developed a sample-based adaptive version of the superexponential method for the blind deconvolution of multi-input–multioutput (MIMO) systems, where the number of its outputs is greater than the number of its inputs. It should be noted that our previous work on the matrix pseudoinversion lemma is restricted to the case when the added matrix $\mathbf{B}\mathbf{B}^H$ is a single dyad (i.e., \mathbf{B} is a column vector) [7]–[9]. Therefore, we can confirm some differences between the present work and the previous one, e.g., see Remark 2 and Tables I and II.

After the presentation of the matrix pseudoinversion lemma, we apply this lemma to block-based adaptive blind deconvolution of a MIMO system. About recent research works of the blind techniques for MIMO systems, a lot of blind deconvolution techniques for MIMO systems have been proposed until now. For example, the multichannel blind deconvolution using a general state-space approach has been proposed [10]. Also, some blind channel estimation methods have been proposed, e.g., [11]–[13]. In this paper, we propose a block-based adaptive version of the multichannel superexponential algorithm (MSEA) for the blind deconvolution.

We also include simulation results for the performance of the proposed algorithm in order to show its effectiveness, where we compare the performance of the proposed algorithm using the lemma with that of the algorithm using the built-in function in MATLAB version 7.1.0 for calculating pseudoinverses of the relevant matrices instead of using the lemma. We should note that a performance result of a block-based adaptive multi-channel superexponential deflation algorithm is shown in [14].

This paper uses the following notation: Let Z denote the set of all integers, C denote the set of all complex numbers, and $C^{m \times n}$ denote the set of all $m \times n$ matrices with complex components. The superscripts T , $*$, H , and \dagger denote the transpose, the complex conjugate, the complex conjugate transpose (Hermitian), and the (Moore–Penrose) pseudoinverse operations of a matrix, respectively. The symbol \oplus denotes the direct sum of subspaces or the direct sum of matrices, and the superscript \perp denotes the orthogonal complement of a subspace [15]. A matrix $D \in C^{m \times r}$ is called a *dyad* (or *dyadic matrix*) if D can be represented as $D = bc^H$ with $b \in C^n$ and $c \in C^r$. Thus, a Hermitian dyadic matrix D can be described as $D = bb^H$. The range space (or image) and the null space (or kernel) of $A \in C^{n \times r}$ are denoted by $R(A)$ and $N(A)$, respectively [16]. Let $i = \overline{1, n}$ stand for $i = 1, 2, \dots, n$.

II. MATRIX PSEUDOINVERSION LEMMA: A GENERAL CASE WITH A BLOCK OF DYADS

The following lemma gives an explicit formula of the pseudoinverse of a positive semidefinite Hermitian matrix A added to a block of Hermitian dyads (represented as BB^H).

Lemma 1: Let $A \in C^{n \times n}$ be a positive semidefinite Hermitian matrix and $B \in C^{n \times l}$ be a matrix decomposed uniquely as

$$B = B_1 \oplus B_2 \text{ with } R(B_1) \subset R(A) \text{ and } R(B_2) \subset R(A)^\perp. \quad (2)$$

Let Q be defined as

$$\begin{aligned} Q &:= A + BB^H \\ &= A + (B_1 + B_2)(B_1 + B_2)^H \\ &= A + B_1 B_1^H + B_2 B_2^H + B_1 B_2^H + B_2 B_1^H \in C^{n \times n}. \end{aligned} \quad (3)$$

Then, the pseudoinverse Q^\dagger of the matrix Q is explicitly expressed, depending on the values of matrices B_1 and B_2 , as follows:

1) If $B_2 = 0$, then

$$Q^\dagger = A^\dagger - A^\dagger B_1 \left(I + B_1^H A^\dagger B_1 \right)^{-1} B_1^H A^\dagger. \quad (4)$$

2) If $B_2 \neq 0$ and $B_1 = 0$, then

$$Q^\dagger = A^\dagger + \left(B_2^H \right)^\dagger B_2^\dagger. \quad (5)$$

3) If $B_2 \neq 0$ and $B_1 \neq 0$, then

$$Q^\dagger = Q_B^\dagger - Q_B^\dagger [B_1, B_2] Q_D^{-1} [B_1, B_2]^H Q_B^\dagger \quad (6)$$

where Q_B^\dagger and Q_D^{-1} are respectively defined by

$$\begin{aligned} Q_B^\dagger &:= \left(A + B_1 B_1^H + B_2 B_2^H \right)^\dagger \\ &= A^\dagger - A^\dagger B_1 \left(I + B_1^H A^\dagger B_1 \right)^{-1} B_1^H A^\dagger + \left(B_2^H \right)^\dagger B_2^\dagger \end{aligned} \quad (7)$$

$$\begin{aligned} Q_D^{-1} &:= \left(P + [B_1, B_2]^H Q_B^\dagger [B_1, B_2] \right)^{-1} \\ &= \left[\begin{array}{c|c} -\Delta^{-1} B_2^H Q_B^\dagger B_2 & \Delta^{-1} \\ \hline I + B_1^H Q_B^\dagger B_1 \Delta^{-1} B_2^H Q_B^\dagger B_2 & -B_1^H Q_B^\dagger B_1 \Delta^{-1} \end{array} \right] \\ &\in C^{2l \times 2l} \end{aligned} \quad (8)$$

with

$$P := \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \in R^{2l \times 2l} \quad (9)$$

$$\Delta := I - B_2^H Q_B^\dagger B_2 B_1^H Q_B^\dagger B_1. \quad (10)$$

Here, $R^{2l \times 2l}$ denotes the set of all $2l \times 2l$ matrices with real components.

Remark 1: It is possible to rewrite (8) and (10) as follows:

$$Q_D^{-1} = \left[\begin{array}{c|c} -B_2^H Q_B^\dagger B_2 \dot{\Delta}^{-1} & I + B_2^H Q_B^\dagger B_2 \dot{\Delta}^{-1} B_1^H Q_B^\dagger B_1 \\ \hline \dot{\Delta}^{-1} & -\dot{\Delta}^{-1} B_1^H Q_B^\dagger B_1 \end{array} \right] \quad (11)$$

$$\dot{\Delta} := I - B_1^H Q_B^\dagger B_1 B_2^H Q_B^\dagger B_2. \quad (12)$$

The proof of Lemma 1 is very lengthy and so is relegated to Appendix A.

Remark 2: A technical important fact in Lemma 1 is that there really exists the inverse of the matrix Q_D defined as

$$Q_D := P + [B_1, B_2]^H Q_B^\dagger [B_1, B_2] \quad (13)$$

in (8) even if P is not positive definite. The proof of the existence is not easy and requires a notion of orthogonal projectors along with a geometric approach to linear transformations [16]. This is a key point which is different from the case where B is a column vector [7]–[9]. The proof of the nonsingularity of matrix Q_D is shown in Appendix B.

It can be seen that the first and second expressions of the pseudoinverse given in Lemma 1 can be included as special cases in the third expression of the pseudoinverse given in (6). Namely, we have the following theorem.

Theorem 1: Under the same conditions in Lemma 1, it follows that

$$\begin{aligned} Q^\dagger &= (A + BB^H)^\dagger \\ &= Q_B^\dagger - Q_B^\dagger [B_1, B_2] Q_D^{-1} [B_1, B_2]^H Q_B^\dagger \end{aligned} \quad (14)$$

where Q_B^\dagger and Q_D^{-1} are defined by (7) and (8), respectively. The proof of Theorem 1 is shown in Appendix C.

Remark 3: In the late 1980s, Ogawa extended the matrix inversion lemma to the singular case and presented an operator pseudoinversion lemma [4]. Instead of the adding term BB^H in (3), he treated a more general adding term $BC^{-1}B^H$ where B is an operator and C is a positive definite operator, but he gave the operator pseudoinversion lemma under the condition

$$R(B) \subset R(A). \quad (15)$$

Therefore, case 1) of Lemma 1 is included in the case that he considered, but the other two cases, namely, 2) and 3), are not treated by him. We should note that the aforementioned condition does not hold true in a nonstationary environment in the blind deconvolution of MIMO systems.

In order to present illustrations of the matrix pseudoinversion lemma, we demonstrate illustrative examples in Appendix D for the reader's convenience.

III. APPLICATION TO BLOCK-BASED ADAPTIVE BLIND DECONVOLUTION

We consider a MIMO system with n inputs and m outputs as described by

$$\mathbf{y}(t) = \sum_{k=-\infty}^{\infty} \mathbf{H}^{(k)} \mathbf{s}(t-k) + \mathbf{n}(t), \quad t \in Z \quad (16)$$

where $\mathbf{s}(t)$ is an n -column vector of input (or source) signals, $\mathbf{y}(t)$ is an m -column vector of output signals, and $\{\mathbf{H}^{(k)}\}$ is an $m \times n$ matrix sequence of impulse responses.

The transfer function of the channel is defined by

$$\mathbf{H}(z) = \sum_{k=-\infty}^{\infty} \mathbf{H}^{(k)} z^{-k}, \quad z \in C. \quad (17)$$

To recover the source signals, we process the output signals by an $n \times m$ deconvolver (or equalizer) $\mathbf{W}(z)$ described by

$$\begin{aligned} \mathbf{z}(t) &= \sum_{k=-\infty}^{\infty} \mathbf{W}^{(k)} \mathbf{y}(t-k) \\ &= \sum_{k=-\infty}^{\infty} \mathbf{G}^{(k)} \mathbf{s}(t-k), \quad t \in Z \end{aligned} \quad (18)$$

where $\{\mathbf{G}^{(k)}\}$ is the impulse response of the cascade system of the unknown system $\mathbf{H}(z)$ and the deconvolver $\mathbf{W}(z)$. The impulse response $\{\mathbf{G}^{(k)}\}$ of the cascade system is defined by

$$\mathbf{G}^{(k)} := \sum_{\tau=-\infty}^{\infty} \mathbf{W}^{(\tau)} \mathbf{H}^{(k-\tau)}, \quad k \in Z. \quad (19)$$

The objective of multichannel blind deconvolution is to construct a deconvolver $\mathbf{W}(z)$ that recovers the original source signals only from the measurements of the corresponding outputs. For the time being, it is assumed for theoretical analysis that the noise term $\mathbf{n}(t)$ in (16) is absent. However, all the signals and the parameters of the systems are allowed to be complex valued.

We put the following assumptions on the channel, the source signals, and the deconvolver.

A1) The transfer function $\mathbf{H}(z)$ is stable and has full column rank on the unit circle $|z| = 1$ (this implies that the unknown system has less inputs than outputs, i.e., $n \leq m$ and that there exists a left stable inverse of the unknown system).

A2) The input sequence $\{\mathbf{s}(t)\}$ is a complex zero-mean non-Gaussian random vector process with element processes $\{s_i(t)\}$, with $i = \overline{1, n}$ being mutually independent. Moreover, each element process $\{s_i(t)\}$ is an independent and identically distributed process with a nonzero variance σ_i^2 and a nonzero fourth-order cumulant γ_i . The variances σ_i^2 's and the fourth-order cumulants γ_i 's are unknown.

A3) The deconvolver $\mathbf{W}(z)$ is a FIR system of sufficient length L so that the truncation effect can be ignored.

Remark 4: As to A1), if the channel $\mathbf{H}(z)$ is FIR, then a condition of the existence of a FIR deconvolver is $\text{rank } \mathbf{H}(z) = n$ for all nonzero $z \in C$ [17]. Moreover, if $\mathbf{H}(z)$ is irreducible, then there exists an equalizer $\mathbf{W}(z)$ of length $L \leq n(K -$

1), where K is the length of the channel [17]. Moreover, it is shown in [18] and [19] that there exists generically (or except for pathological cases) an equalizer $\mathbf{W}(z)$ of length $L = \lceil n(K-1)/(m-n) \rceil$, where $\lceil x \rceil$ stands for the smallest integer that is greater than or equal to x .

Based on assumption A3), let us consider a FIR deconvolver with the transfer function $\mathbf{W}(z)$ given by

$$\mathbf{W}(z) = \sum_{k=L_1}^{L_2} \mathbf{W}^{(k)} z^{-k} \quad (20)$$

where L_1 and L_2 are the first and last superscripted numbers, respectively, of the tap coefficients, i.e., $\mathbf{W}^{(k)}$'s, of the deconvolver $\mathbf{W}(z)$ and where the length $L := L_2 - L_1 + 1$ is taken to be sufficiently large. Let $\tilde{\mathbf{w}}_i$ be the mL -column vector consisting of the tap coefficients (corresponding to the i th output) of the deconvolver defined by

$$\tilde{\mathbf{w}}_i := [\mathbf{w}_{i,1}^T, \mathbf{w}_{i,2}^T, \dots, \mathbf{w}_{i,m}^T]^T \in \mathbf{C}^{mL} \quad (21)$$

$$\mathbf{w}_{i,j} = [w_{i,j}^{(L_1)}, w_{i,j}^{(L_1+1)}, \dots, w_{i,j}^{(L_2)}]^T \in \mathbf{C}^L \quad (22)$$

where $w_{i,j}^{(k)}$ is the (i, j) th element of matrix $\mathbf{W}^{(k)}$.

Inouye and Tanebe [20] proposed the MSEA for finding the tap coefficient vectors, i.e., $\tilde{\mathbf{w}}_i$'s, of the deconvolver $\mathbf{W}(z)$, each iteration of which consists of the following two steps:

$$\tilde{\mathbf{w}}_i^{[1]} = \tilde{\mathbf{R}}^\dagger \tilde{\mathbf{d}}_i, \quad i = \overline{1, n}, \quad (23)$$

$$\tilde{\mathbf{w}}_i^{[2]} = \frac{\tilde{\mathbf{w}}_i^{[1]}}{\sqrt{\tilde{\mathbf{w}}_i^{[1]H} \tilde{\mathbf{R}} \tilde{\mathbf{w}}_i^{[1]}}}, \quad i = \overline{1, n}, \quad (24)$$

where $(\cdot)^{[1]}$ and $(\cdot)^{[2]}$ stand for the results of the first and the second step, respectively. Let $\tilde{\mathbf{y}}(t)$ be the mL -column vector consisting of the L consecutive inputs of the deconvolver defined by

$$\tilde{\mathbf{y}}(t) := [\bar{\mathbf{y}}_1(t)^T, \bar{\mathbf{y}}_2(t)^T, \dots, \bar{\mathbf{y}}_m(t)^T]^T \in \mathbf{C}^{mL} \quad (25)$$

$$\bar{\mathbf{y}}_i(t) := [y_i(t-L_1), y_i(t-L_1-1), \dots, y_i(t-L_2)]^T \in \mathbf{C}^L \quad (26)$$

where $y_i(t)$ is the i th element of the output vector $\mathbf{y}(t)$ of the channel in (16). Then, the correlation matrix $\tilde{\mathbf{R}}$ (defined in [20, (41) and (42)]) is represented as

$$\tilde{\mathbf{R}} = E \left[\tilde{\mathbf{y}}^*(t) \tilde{\mathbf{y}}^T(t) \right] \in \mathbf{C}^{mL \times mL} \quad (27)$$

and the fourth-order cumulant vector $\tilde{\mathbf{d}}_i$ is defined by

$$\tilde{\mathbf{d}}_i := \left[\mathbf{d}_{i,1}^T, \mathbf{d}_{i,2}^T, \dots, \mathbf{d}_{i,m}^T \right]^T \in \mathbf{C}^{mL} \quad (28)$$

whose j th block element $\mathbf{d}_{i,j}$ is the L -column vector with r th element $[d_{i,j}]_r$ defined by

$$[d_{i,j}]_r = \text{cum} \left(z_i(t), z_i(t), z_i^*(t), y_j^*(t-r) \right), \quad r = 0, \pm 1, \pm 2, \dots \quad (29)$$

and $\tilde{\mathbf{d}}_i$ is represented as

$$\begin{aligned} \tilde{\mathbf{d}}_i &= E \left[|z_i(t)|^2 z_i(t) \tilde{\mathbf{y}}^*(t) \right] \\ &\quad - 2E \left[|z_i(t)|^2 \right] E \left[z_i(t) \tilde{\mathbf{y}}^*(t) \right] \\ &\quad - E \left[z_i^2(t) \right] E \left[z_i^*(t) \tilde{\mathbf{y}}^*(t) \right] \in \mathbf{C}^{mL} \end{aligned} \quad (30)$$

where $E[x]$ denotes the expectation of a random variable x . We note that the last term can be ignored in case of $E[z_i^2(t)] = 0$ for all $i = \overline{1, n}$, in which case $E[z_i^2(t)] = 0$ for all $i = \overline{1, n}$.

In addition, we obtain (from [20, (32) and (39)])

$$\tilde{\mathbf{d}}_i \in R(\tilde{\mathbf{R}}), \quad i = \overline{1, n}. \quad (31)$$

We should note that the stationarity of the input process $\{\mathbf{s}(t)\}$ (or assumption A2) ensures relation (31) and that relation (31) means that the vector $\tilde{\mathbf{w}}_i^{[1]}$ obtained by (23) satisfies

$$\tilde{\mathbf{R}}\tilde{\mathbf{w}}_i^{[1]} = \tilde{\mathbf{d}}_i. \quad (32)$$

Consider the batch algorithm in (23) and (24). Equation (24) constrains a weighted norm of vector $\tilde{\mathbf{w}}_i$ to equal one, and thus, we assume that this constraint is always satisfied using a normalization or an automatic gain control of $\tilde{\mathbf{w}}_i$ at each discrete (or sample) time t . To develop an adaptive version of (23), we must specify the dependence of each time t and rewrite (23) as

$$\tilde{\mathbf{w}}_i(t) = \tilde{\mathbf{R}}^\dagger(t)\tilde{\mathbf{d}}_i(t), \quad i = \overline{1, n}. \quad (33)$$

On the other hand, a block-based adaptive algorithm for designing adaptive filters is one of many efficient adaptive filtering algorithms aimed at increasing convergence speed and reducing the computational complexity just as the block-based least mean square algorithm shown in [3, p. 347]. The basic principle of the block-based algorithm for designing an adaptive filter is that the filter coefficients remain unchanged during the processing of each data block and are updated only once per block [3]. Suppose l is the block length. Then, the original discrete (or sample) time t is related to the k th block of data as

$$t = kl + i, \quad i = \overline{1, l-1}; k \in \mathbf{Z}. \quad (34)$$

The index k is referred to as the *block index*. Following this principle along with the notation in (34), we develop a block-based adaptive MSEA for the blind deconvolution of the system (16).

Let k denote the block index. We can rewrite (33) as

$$\tilde{\mathbf{w}}_i(k) = \tilde{\mathbf{R}}^\dagger(k)\tilde{\mathbf{d}}_i(k), \quad i = \overline{1, n}. \quad (35)$$

Then, we should obtain recursion formulas for the block updating of matrix $\tilde{\mathbf{R}}(k)$ and vector $\tilde{\mathbf{d}}_i(k)$ in (35), respectively

$$\tilde{\mathbf{R}}(k) = (1 - \alpha_k)\tilde{\mathbf{R}}(k-1) + \alpha_k\tilde{\mathbf{B}}^*(k)\tilde{\mathbf{B}}^\top(k) \quad (36)$$

$$\begin{aligned} \tilde{\mathbf{d}}_i(k) &= (1 - \alpha_k)\tilde{\mathbf{d}}_i(k-1) \\ &+ \alpha_k \sum_{j=0}^{l-1} \tilde{\mathbf{y}}^* \{(k-1)l + j\} \tilde{z}_i \{(k-1)l + j\} \end{aligned} \quad (37)$$

where

$$\begin{aligned} \tilde{\mathbf{B}}(k) &= [\tilde{\mathbf{y}} \{(k-1)l\}, \tilde{\mathbf{y}} \{(k-1)l + 1\}, \dots, \\ &\quad \tilde{\mathbf{y}} \{(k-1)l + l - 1\}] \in \mathbf{C}^{mL \times l} \end{aligned} \quad (38)$$

$$\begin{aligned} \tilde{z}_i(j) &:= \left(|z_i(j)|^2 - 2 \langle |z_i(j)|^2 \rangle \right) z_i(j) \\ &- \langle z_i^2(j) \rangle z_i^*(j). \end{aligned} \quad (39)$$

Here, $\langle |z_i(j)|^2 \rangle$ and $\langle z_i^2(j) \rangle$ denote the estimates of $E[|z_i(j)|^2]$ and $E[z_i^2(j)]$ at time t , respectively, and α_k is a positive number close to, but greater than zero, which ac-

counts for some exponential weighting factor or forgetting factor [3]. For example, we may take $\alpha_k = 1/kl$.

Because we consider the case when the number of inputs n is less than the number of outputs m , i.e., $n < m$, the correlation matrix $\tilde{\mathbf{R}}(k)$ is not of full rank and a singular matrix [7]. Therefore, we may apply the matrix pseudoinversion lemma to the recursive equation (36).

By applying Theorem 1 to (36) for obtaining a recursive formula for the block updating of pseudoinverse $\mathbf{P}(k) = \tilde{\mathbf{R}}^\dagger(k)$, we have the following lemma.

Lemma 2: Let \mathbf{Q} , \mathbf{Q}^\dagger , \mathbf{A} , \mathbf{A}^\dagger , \mathbf{B} , \mathbf{B}_1 , and \mathbf{B}_2 in Lemma 1 be respectively defined as

$$\mathbf{Q} = \tilde{\mathbf{R}}(k) \quad (40)$$

$$\mathbf{Q}^\dagger = \mathbf{P}(k) = \tilde{\mathbf{R}}^\dagger(k) \quad (41)$$

$$\mathbf{A} = (1 - \alpha_k)\tilde{\mathbf{R}}(k-1) \quad (42)$$

$$\mathbf{A}^\dagger = \frac{1}{1 - \alpha_k}\tilde{\mathbf{R}}^\dagger(k-1) = \frac{1}{1 - \alpha_k}\mathbf{P}(k-1) \quad (43)$$

$$\mathbf{B} = \mathbf{B}(k) = \sqrt{\alpha_k}\tilde{\mathbf{B}}^*(k) \quad (44)$$

$$\mathbf{B}_1 = \mathbf{B}_1(k) = \tilde{\mathbf{R}}(k-1)\mathbf{P}(k-1)\mathbf{B}(k) \quad (45)$$

$$\mathbf{B}_2 = \mathbf{B}_2(k) = \left\{ \mathbf{I} - \tilde{\mathbf{R}}(k-1)\mathbf{P}(k-1) \right\} \mathbf{B}(k). \quad (46)$$

Then, substituting these definitions into Lemma 1, the recursion for the pseudoinverse $\mathbf{P}(k) = \tilde{\mathbf{R}}^\dagger(k)$ of the correlation matrix $\tilde{\mathbf{R}}(k)$ from $\mathbf{P}(k-1)$ is explicitly expressed as follows:

$$\begin{aligned} \mathbf{P}(k) &= \mathbf{P}_B^\dagger(k) - \mathbf{P}_B^\dagger(k) [\mathbf{B}_1(k), \mathbf{B}_2(k)] \mathbf{P}_D^{-1}(k) \\ &\quad \times [\mathbf{B}_1(k), \mathbf{B}_2(k)]^H \mathbf{P}_B^\dagger(k) \end{aligned} \quad (47)$$

where $\mathbf{P}_B^\dagger(k)$ and $\mathbf{P}_D^{-1}(k)$ are respectively defined by

$$\begin{aligned} \mathbf{P}_B^\dagger(k) &:= \frac{1}{1 - \alpha_k} [\mathbf{P}(k-1) - \mathbf{P}(k-1)\mathbf{B}_1(k) \\ &\quad \times \mathbf{P}_A^{-1}(k)\mathbf{B}_1^H(k)\mathbf{P}(k-1)] \\ &\quad + \left(\mathbf{B}_2^H(k) \right)^\dagger \mathbf{B}_2^\dagger(k) \end{aligned} \quad (48)$$

$$\mathbf{P}_D^{-1}(k) := \left[\begin{array}{c|c} -\Delta^{-1}(k)\mathbf{E}_2(k) & \Delta^{-1}(k) \\ \hline \mathbf{I} + \mathbf{E}_1(k)\Delta^{-1}(k)\mathbf{E}_2(k) & -\mathbf{E}_1(k)\Delta^{-1}(k) \end{array} \right] \quad (49)$$

with

$$\Delta(k) := \mathbf{I} - \mathbf{E}_2(k)\mathbf{E}_1(k) \quad (50)$$

where

$$\mathbf{E}_1(k) = \mathbf{B}_1^H(k)\mathbf{P}_B^\dagger(k)\mathbf{B}_1(k) \quad (51)$$

$$\mathbf{E}_2(k) = \mathbf{B}_2^H(k)\mathbf{P}_B^\dagger(k)\mathbf{B}_2(k). \quad (52)$$

These equations are initialized by their values appropriately selected or calculated by the batch algorithm in (23) and (24) at initial block index k_0 and used for $k = k_0 + 1, k_0 + 2, \dots$

The proof of Lemma 2 is easy and follows from Lemma 1 along with simple calculations.

Based on Lemma 2 along with (35)–(39), we have the following theorem which gives a recursion formula for the block updating of the tap vector $\tilde{\mathbf{w}}_i(k)$ for $i = \overline{1, n}$.

Theorem 2: The recursion for $\tilde{\mathbf{w}}_i(k)$ is

$$\begin{aligned} \tilde{\mathbf{w}}_i(k) = & \mathbf{P}(k) \tilde{\mathbf{R}}(k) \tilde{\mathbf{w}}_i(k-1) + \mathbf{k}(k) \\ & \times \left[\sum_{j=0}^{l-1} \tilde{\mathbf{y}}^* \{(k-1)l+j\} \tilde{z}_i \{(k-1)l+j\} \right. \\ & \left. - \tilde{\mathbf{B}}^*(k) \tilde{\mathbf{B}}^T(k) \tilde{\mathbf{w}}_i(k-1) \right] \end{aligned} \quad (53)$$

where

$$\mathbf{k}(k) := \alpha_k \mathbf{P}(k) \quad (54)$$

$$\tilde{z}_i(j) := \left(|z_i(j)|^2 - 2 \langle |z_i(j)|^2 \rangle \right) z_i(j) - \langle z_i^2(j) \rangle z_i^*(j) \quad (55)$$

$$\langle |z_i(j)|^2 \rangle := (1 - \beta_k) \langle |z_i(j-1)|^2 \rangle + \beta_k |z_i(j)|^2 \quad (56)$$

$$\langle z_i^2(j) \rangle := (1 - \beta_k) \langle z_i^2(j-1) \rangle + \beta_k z_i^2(j). \quad (57)$$

Here, β_k is a positive constant greater than α_k and $\mathbf{P}(k)$ is calculated from (47). The proof of Theorem 2 is shown in Appendix E.

Remark 5: The recursive algorithm proposed by Shalvi and Weinstein [21] can be shown to correspond to the particular case of Theorem 1 where $n = m = 1$, $l = 1$, and the correlation matrices $\tilde{\mathbf{R}}(k)$'s are nonsingular.

Remark 6: The case where the length of the block $l = 1$ of the matrix pseudoinversion lemma is restricted to the case when the added matrix $\mathbf{B}\mathbf{B}^H$ is a single dyad (i.e., \mathbf{B} is a column vector) [7]–[9]. An illustrative example and the simulation results of the matrix pseudoinversion lemma in this case and the application to adaptive blind deconvolution of MIMO systems are shown in [7]–[9].

IV. SIMULATION RESULTS

To demonstrate the usefulness of the matrix pseudoinversion lemma, some computer simulations for obtaining the pseudoinverse $\mathbf{P}(k) = \tilde{\mathbf{R}}^\dagger(k)$ of the correlation matrix $\tilde{\mathbf{R}}(k)$ in (36) by using Lemma 2 were conducted. We note here that we do not use Theorem 2, because we are interested not in finding $\tilde{\mathbf{w}}_i(k)$'s but in calculating $\tilde{\mathbf{R}}^\dagger(k)$. Results of calculating $\tilde{\mathbf{w}}_i(k)$'s and recovering original sources $\mathbf{s}(k)$'s are found in [14].

We considered a MIMO system $\mathbf{H}(z)$ with two inputs ($n = 2$) and five outputs ($m = 5$) and assumed that the system $\mathbf{H}(z)$ is of FIR and the length of channel is three, i.e., $\mathbf{H}^{(k)}$'s in (17) were set to be

$$\begin{aligned} \mathbf{H}(z) &= \sum_{k=0}^2 \mathbf{H}^{(k)} z^{-k} \\ &= \begin{bmatrix} 1.00 + 0.15z^{-1} + 0.10z^{-2} & 0.65 + 0.25z^{-1} + 0.15z^{-2} \\ 0.50 - 0.10z^{-1} + 0.20z^{-2} & 1.00 + 0.25z^{-1} + 0.10z^{-2} \\ 0.60 + 0.10z^{-1} + 0.40z^{-2} & 0.10 + 0.20z^{-1} + 0.10z^{-2} \\ 0.50 + 0.10z^{-1} - 0.20z^{-2} & 0.30 + 0.40z^{-1} + 0.50z^{-2} \\ 0.80 + 0.10z^{-1} + 0.40z^{-2} & 0.10 - 0.20z^{-1} + 0.10z^{-2} \end{bmatrix}. \end{aligned} \quad (58)$$

Two source signals were 4- and 8-PSK signals, respectively. The length of the equalizer is seven ($L = 7$). For obtaining

the pseudoinverse of the correlation matrix, the initial values of $\tilde{\mathbf{R}}$, $\tilde{\mathbf{d}}_i$, and \mathbf{P} were estimated using 30 data samples. The value of α_k was chosen as $\alpha_k = 1/kl$ for each k . As a measure of performance, we use the following sum of the Frobenius norms of the four error matrices:

$$M := \frac{\|\mathbf{F}_1(k)\| + \|\mathbf{F}_2(k)\| + \|\mathbf{F}_3(k)\| + \|\mathbf{F}_4(k)\|}{\|\tilde{\mathbf{R}}(k)\|} \quad (59)$$

where $\|\mathbf{F}_i(k)\|$ denotes the Frobenius norm of matrix $\mathbf{F}_i(k)$ and $\mathbf{F}_1(k)$, $\mathbf{F}_2(k)$, $\mathbf{F}_3(k)$, and $\mathbf{F}_4(k)$, respectively, are the error matrices defined by using the Moore–Penrose conditions as follows [16]:

$$\mathbf{F}_1(k) = \tilde{\mathbf{R}}(k) \mathbf{P}(k) \tilde{\mathbf{R}}(k) - \tilde{\mathbf{R}}(k) \quad (60)$$

$$\mathbf{F}_2(k) = \mathbf{P}(k) \tilde{\mathbf{R}}(k) \mathbf{P}(k) - \mathbf{P}(k) \quad (61)$$

$$\mathbf{F}_3(k) = \left(\tilde{\mathbf{R}}(k) \mathbf{P}(k) \right)^H - \tilde{\mathbf{R}}(k) \mathbf{P}(k) \quad (62)$$

$$\mathbf{F}_4(k) = \left(\mathbf{P}(k) \tilde{\mathbf{R}}(k) \right)^H - \mathbf{P}(k) \tilde{\mathbf{R}}(k). \quad (63)$$

In this application, from A2), the input process $\{\mathbf{s}(t)\}$ is stationary, and this means that the random process $\{\tilde{\mathbf{y}}^*(t)\}$ is also stationary. Thus, it follows from (27) that $\tilde{\mathbf{y}}^*(t)$ belongs almost surely (a.s.) (or with probability 1) to $R(\tilde{\mathbf{R}})$, i.e.,

$$\tilde{\mathbf{B}}^*(t) \in R(\tilde{\mathbf{R}}), \quad \text{a.s.} \quad (64)$$

This means that, from (38), (42), and (44)

$$\mathbf{B} \in R(\mathbf{A}). \quad (65)$$

Therefore, we can assume in this application that the component matrix $\mathbf{B}_2 = \mathbf{B}_2(k)$ always vanishes, i.e., $\mathbf{B}_2(k) = 0$, and we can use the recursion formula (47) for calculating the pseudoinverse $\mathbf{P}(k) = \tilde{\mathbf{R}}^\dagger(k)$ at each iteration (or block index) k .

We compared the performance of the proposed method (i.e., the method using the matrix pseudoinversion lemma) with the performance of the method using the built-in function “pinv” in MATLAB Version 7.1.0 for calculating the pseudoinverse $\mathbf{P}(k) = \tilde{\mathbf{R}}^\dagger(k)$ of correlation matrix $\tilde{\mathbf{R}}(k)$. The pseudoinverses are calculated iteratively (or recursively) for each iteration (or recursion) number k for the two methods.

Fig. 1 shows the performance results of the performance measure M for the proposed method with $l = 1$ (a) and $l = 2$ (b) and for the latter method (c) by using 500 data samples. The almost converged values of the performance measure M at discrete time $t = 500$ also are superimposed on the three figures.

We also compared the performances of the two methods in computational complexity by using the built-in function “flops” in MATLAB Version 5.2 and in execution time by using a personal computer with a 2.66-GHz processor and 3.4-GB main memories used in simulation experiments.

Table I shows the average of the numbers of floating-point operations (flops) over 100 independent Monte Carlo runs using 500 data samples of the outputs for each Monte Carlo run by changing the length of the block l from 1 to 4 and the length of the deconvolver L from 3 to 14.

Table II shows the average of the execution times over 100 independent Monte Carlo runs using 500 data samples of the outputs for each Monte Carlo run by changing the length of the block l from 1 to 4 and the length of the deconvolver L from 3 to 14.

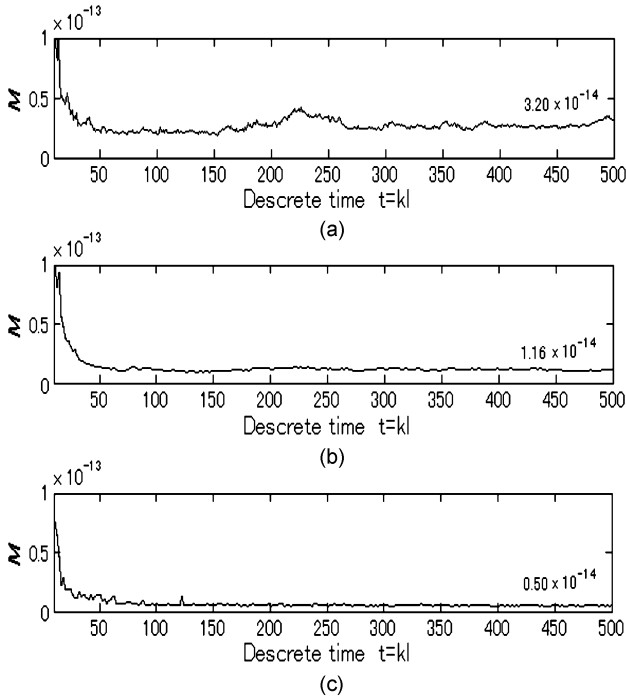


Fig. 1. Performance measure M of (a) the proposed method with $l = 1$, (b) the proposed method with $l = 2$, and (c) the method using “pinv” ($l = 2$) for the time-invariant MIMO system.

It can be seen from Fig. 1(a) and (b) and Tables I and II that the performance measure M , the average of the numbers of floating-point operations, and the average of the execution times of the proposed method are better as the length of the block l increases for the same length of the deconvolver $L = 7$.

In the meantime, it can be seen from Fig. 1(b) and (c) that the accuracy of the matrix pseudoinversion lemma is equivalent to the built-in function “pinv.” However, it can be seen from Tables I and II that the average of the numbers of floating-point operations and the average of the execution times for the proposed method are better than those for the method using built-in function “pinv,” respectively.

The computational complexity of the method using the built-in function “pinv” increases more than the computational complexity of the proposed method when the length of the deconvolver L increases. We consider that one of reasons why the matrix pseudoinversion lemma is superior to the built-in function “pinv” in the numbers of floating-point operations and the execution times is that it is not necessary to calculate the pseudoinverse \mathbf{A}^\dagger in Lemma 1, because the results of the previous iteration $\mathbf{P}(k-1) = \tilde{\mathbf{R}}^\dagger(k-1)$ can be used instead of \mathbf{A}^\dagger in Lemma 2.

Next, we considered a time-variant MIMO system, where the system is the same as that in (58) except that the last matrix $\mathbf{H}^{(2)}$ of the impulse response of the channel was varied by adding 0.05 to all its elements at discrete time $t = 200$. In this case, the random process $\{\tilde{\mathbf{y}}^*(t)\}$ is not stationary, i.e., $\mathbf{B}_2(k) \neq 0$. Fig. 2(a) and (b) shows the performance results of the performance measure M for the proposed method and the method using the built-in function “pinv” with $L = 7$ and $l = 2$ by using 500 data samples of the outputs of the MIMO time-variant system. The almost converged values of the performance measure M at discrete time $t = 500$ for both methods are superimposed on the two figures. Fig. 2(c) shows the values of the rank of $\tilde{\mathbf{R}}$ normalized by the length of the deconvolver L in this case.

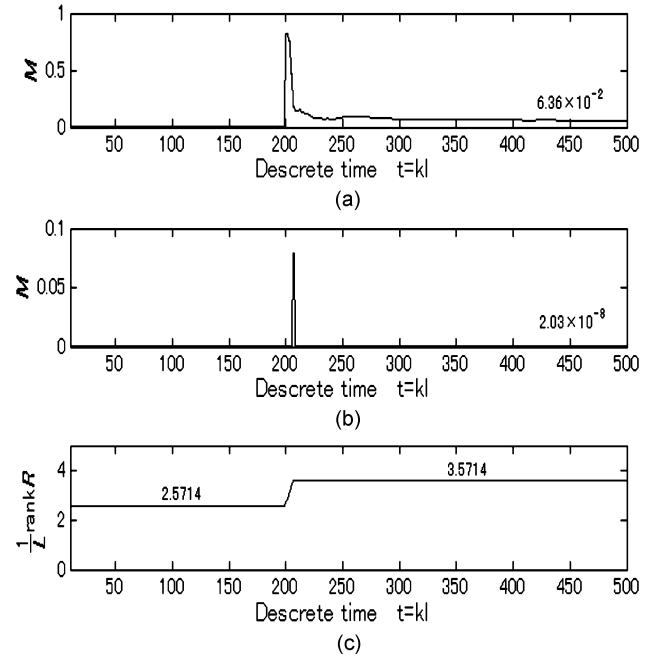


Fig. 2. Performance measure M of (a) the proposed method with $l = 2$ and (b) the method using “pinv” ($l = 2$). (c) Values of $1/L$ rank $\tilde{\mathbf{R}}$ for the time-variant MIMO system.

It can be seen from Fig. 2(a) and (b) that the accuracy of the matrix pseudoinversion lemma is a little worse than the built-in function “pinv,” but the formula (14) of the matrix pseudoinversion lemma can treat the changes of the channel.

From these results, the matrix pseudoinversion lemma is useful to calculate the pseudoinverse of a correlation matrix for block-based adaptive algorithms of blind deconvolution in time-invariant MIMO systems and even time-variant MIMO systems.

V. CONCLUSION

We have extended the matrix inversion lemma to the case when the matrix \mathbf{A} in $\mathbf{A} + \mathbf{B}\mathbf{B}^H$ is singular and presented a matrix pseudoinversion lemma together with some illustrative examples. In order to show the usefulness of this lemma, we applied it to develop a block-based adaptive superexponential algorithm for the blind deconvolution of a MIMO system. It has been shown through computer simulations that the matrix pseudoinversion lemma is useful for block-based adaptive algorithms of blind deconvolution in time-invariant MIMO systems and even time-variant MIMO systems.

APPENDIX A THE PROOF OF LEMMA 1

First, we consider the case when $\mathbf{B}_2 = 0$. Put

$$\mathbf{Q}_A := \mathbf{A} + \mathbf{B}_1 \mathbf{B}_1^H. \quad (66)$$

Since \mathbf{A} is Hermitian, it is unitarily diagonalizable, i.e., there exist a unitary matrix \mathbf{U} and a diagonal matrix $\mathbf{\Lambda}$ such that

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \quad (67)$$

which is called the spectral decomposition (or eigenvalue decomposition) of \mathbf{A} . In (67), the diagonal elements λ_i 's are arranged in the decreasing order as

$$\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \quad (68)$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$. If $\text{rank } \mathbf{A} = r < n$, then $\lambda_i = 0$ for $r < i \leq n$. Thus, we have \mathbf{U} and $\mathbf{\Lambda}$ partitioned as

$$\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2], \quad \mathbf{U}_1 \in \mathcal{C}^{m \times r}; \mathbf{U}_2 \in \mathcal{C}^{m \times (n-r)} \quad (69)$$

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 \end{bmatrix}, \quad \mathbf{\Lambda}_1 \in \mathcal{C}^{r \times r}; \mathbf{\Lambda}_2 \in \mathcal{C}^{(n-r) \times (n-r)} \quad (70)$$

where $\mathbf{\Lambda}_2 = \mathbf{0}$. Therefore, from (67), we have

$$\mathbf{A} = [\mathbf{U}_1, \mathbf{U}_2] \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{bmatrix} = \mathbf{U}_1 \mathbf{\Lambda}_1 \mathbf{U}_1^H. \quad (71)$$

Since $\mathbf{B}_1 \in R(\mathbf{A})$, there exists $\mathbf{\Xi}_1 \in \mathcal{C}^{r \times l}$ such that

$$\mathbf{B}_1 = \mathbf{U}_1 \mathbf{\Xi}_1. \quad (72)$$

From (66), (71), and (72), we have

$$\begin{aligned} \mathbf{Q}_A &= \mathbf{A} + \mathbf{B}_1 \mathbf{B}_1^H \\ &= \mathbf{U}_1 \mathbf{\Lambda}_1 \mathbf{U}_1^H + \mathbf{U}_1 \mathbf{\Xi}_1 (\mathbf{U}_1 \mathbf{\Xi}_1)^H \\ &= \mathbf{U}_1 \left(\mathbf{\Lambda}_1 + \mathbf{\Xi}_1 \mathbf{\Xi}_1^H \right) \mathbf{U}_1^H \\ &= [\mathbf{U}_1, \mathbf{U}_2] \left[\begin{array}{c|c} \mathbf{\Lambda}_1 + \mathbf{\Xi}_1 \mathbf{\Xi}_1^H & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right] \begin{bmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{bmatrix}. \end{aligned} \quad (73)$$

Here, we note the following properties of matrix pseudoinverses [22]:

P1) If \mathbf{U} and \mathbf{V} are unitary, then

$$(\mathbf{UAV})^\dagger = \mathbf{V}^\dagger \mathbf{A}^\dagger \mathbf{U}^\dagger = \mathbf{V}^H \mathbf{A}^\dagger \mathbf{U}^H. \quad (74)$$

P2) If $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2$ where $\mathbf{A}_1^H \mathbf{A}_2 = \mathbf{0}$ and $\mathbf{A}_1 \mathbf{A}_2^H = \mathbf{0}$, then

$$\mathbf{A}^\dagger = \mathbf{A}_1^\dagger + \mathbf{A}_2^\dagger. \quad (75)$$

P3) Let $r = \text{rank } \mathbf{X}$ and let

$$\mathbf{X} = \mathbf{YZ} \in \mathcal{C}^{m \times m}, \quad \mathbf{Y} \in \mathcal{C}^{m \times r}; \mathbf{Z} \in \mathcal{C}^{r \times m} \quad (76)$$

be a rank decomposition of $\mathbf{X} \in \mathcal{C}^{m \times m}$, i.e., $\text{rank } \mathbf{Y} = \text{rank } \mathbf{Z} = r$. Then

$$\mathbf{X}^\dagger = \mathbf{Z}^\dagger \mathbf{Y}^\dagger. \quad (77)$$

Using (73) and the aforementioned properties, we have

$$\begin{aligned} \mathbf{Q}_A^\dagger &= [\mathbf{U}_1, \mathbf{U}_2] \begin{bmatrix} \mathbf{\Lambda}_1 + \mathbf{\Xi}_1 \mathbf{\Xi}_1^H & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{bmatrix} \\ &= [\mathbf{U}_1, \mathbf{U}_2] \begin{bmatrix} \left(\mathbf{\Lambda}_1 + \mathbf{\Xi}_1 \mathbf{\Xi}_1^H \right)^\dagger & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{bmatrix} \\ &= \mathbf{U}_1 \left(\mathbf{\Lambda}_1 + \mathbf{\Xi}_1 \mathbf{\Xi}_1^H \right)^\dagger \mathbf{U}_1^H \\ &= \mathbf{U}_1 \left(\mathbf{\Lambda}_1 + \mathbf{\Xi}_1 \mathbf{\Xi}_1^H \right)^{-1} \mathbf{U}_1^H. \end{aligned} \quad (78)$$

The last equality in (78) comes from the fact that $\mathbf{\Lambda}_1 + \mathbf{\Xi}_1 \mathbf{\Xi}_1^H$ is nonsingular, because $\mathbf{\Lambda}_1$ is nonsingular. Using the matrix inversion lemma [3], [23], we have

$$\left(\mathbf{\Lambda}_1 + \mathbf{\Xi}_1 \mathbf{\Xi}_1^H \right)^{-1} = \mathbf{\Lambda}_1^{-1} - \mathbf{\Lambda}_1^{-1} \mathbf{\Xi}_1 \left(\mathbf{I} + \mathbf{\Xi}_1^H \mathbf{\Lambda}_1^{-1} \mathbf{\Xi}_1 \right)^{-1} \mathbf{\Xi}_1^H \mathbf{\Lambda}_1^{-1} \quad (79)$$

where \mathbf{I} is the $l \times l$ identity matrix.

On the other hand, from (71), (72), and P1), we have

$$\mathbf{A}^\dagger = \mathbf{U}_1 \mathbf{\Lambda}_1^{-1} \mathbf{U}_1^H \quad (80)$$

$$\mathbf{U}_1^H \mathbf{B}_1 = \mathbf{U}_1^H \mathbf{U}_1 \mathbf{\Xi}_1 = \mathbf{\Xi}_1 \quad (81)$$

$$\mathbf{\Xi}_1^H \mathbf{\Lambda}_1^{-1} \mathbf{\Xi}_1 = \mathbf{B}_1^H \mathbf{U}_1 \mathbf{\Lambda}_1^{-1} \mathbf{U}_1^H \mathbf{B}_1 = \mathbf{B}_1^H \mathbf{A}^\dagger \mathbf{B}_1. \quad (82)$$

Using (72) and (78)–(82), we obtain

$$\begin{aligned} \mathbf{Q}_A^\dagger &= \mathbf{U}_1 \left\{ \mathbf{\Lambda}_1^{-1} - \mathbf{\Lambda}_1^{-1} \mathbf{\Xi}_1 \left(\mathbf{I} + \mathbf{\Xi}_1^H \mathbf{\Lambda}_1^{-1} \mathbf{\Xi}_1 \right)^{-1} \mathbf{\Xi}_1^H \mathbf{\Lambda}_1^{-1} \right\} \mathbf{U}_1^H \\ &= \mathbf{U}_1 \mathbf{\Lambda}_1^{-1} \mathbf{U}_1^H - \mathbf{U}_1 \mathbf{\Lambda}_1^{-1} \mathbf{U}_1^H \mathbf{U}_1 \mathbf{\Xi}_1 \left(\mathbf{I} + \mathbf{\Xi}_1^H \mathbf{\Lambda}_1^{-1} \mathbf{\Xi}_1 \right)^{-1} \\ &\quad \times \mathbf{\Xi}_1^H \mathbf{U}_1^H \mathbf{U}_1 \mathbf{\Lambda}_1^{-1} \mathbf{U}_1^H \\ &= \mathbf{A}^\dagger - \mathbf{A}^\dagger \mathbf{B}_1 \left(\mathbf{I} + \mathbf{B}_1^H \mathbf{A}^\dagger \mathbf{B}_1 \right)^{-1} \mathbf{B}_1^H \mathbf{A}^\dagger. \end{aligned} \quad (83)$$

Equation (83) is equivalent to (4), because $\mathbf{Q} = \mathbf{Q}_A$ in this case when $\mathbf{B}_2 = \mathbf{0}$.

Now we consider the proof of the proposition in the general case where $\mathbf{B}_2 \neq \mathbf{0}$. The proof is carried out in two stages. At the first stage, we get the pseudoinverse of the matrix \mathbf{Q}_B defined by

$$\begin{aligned} \mathbf{Q}_B &:= \mathbf{A} + \mathbf{B}_1 \mathbf{B}_1^H + \mathbf{B}_2 \mathbf{B}_2^H \\ &= \left(\mathbf{A} + \mathbf{B}_1 \mathbf{B}_1^H \right) \oplus \mathbf{B}_2 \mathbf{B}_2^H \\ &= \mathbf{Q}_A \oplus \mathbf{B}_2 \mathbf{B}_2^H \end{aligned} \quad (84)$$

where \oplus denotes the direct sum of two matrices, i.e., $\mathbf{A}_1 + \mathbf{A}_2 = \mathbf{A}_1 \oplus \mathbf{A}_2$ if $\mathbf{A}_1^H \mathbf{A}_2 = \mathbf{0}$ and $\mathbf{A}_1 \mathbf{A}_2^H = \mathbf{0}$. In the second stage, we obtain the pseudoinverse of the matrix \mathbf{Q} which is represented as

$$\mathbf{Q} = \mathbf{Q}_B + [\mathbf{B}_1, \mathbf{B}_2] \mathbf{P} [\mathbf{B}_1, \mathbf{B}_2]^H \quad (85)$$

where \mathbf{P} is the $2l \times 2l$ permutation matrix defined by

$$\mathbf{P} := \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \in \mathcal{R}^{2r \times 2r} \quad (86)$$

and \mathbf{I} is the $l \times l$ identity matrix.

The first stage of the proof is carried out as follows. If \mathbf{B}_2 is of full column rank, then it follows from P3) that

$$\left(\mathbf{B}_2 \mathbf{B}_2^H \right)^\dagger = \left(\mathbf{B}_2^H \right)^\dagger \mathbf{B}_2^\dagger. \quad (87)$$

It should be noted that (87) holds true even if \mathbf{B}_2 is not of full column rank (and, thus, even if $\mathbf{B}_2 = \mathbf{0}$). This statement is verified by using the singular-value decomposition (SVD) of the matrix \mathbf{B}_2 as follows. Let

$$\mathbf{B}_2 = \mathbf{U} \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{V}^H \quad (88)$$

be an SVD of \mathbf{B}_2 , where \mathbf{U} and \mathbf{V} are unitary and $\mathbf{\Sigma}$ is a diagonal matrix with diagonal elements $\sigma_i > 0$ for $i = \overline{1, r_2}$, where $r_2 := \text{rank } \mathbf{B}_2$. Then, it follows that

$$\begin{aligned} \mathbf{B}_2 \mathbf{B}_2^H &= \mathbf{U} \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{V}^H \mathbf{V} \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}^H \\ &= \mathbf{U} \begin{bmatrix} \mathbf{\Sigma}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}^H \end{aligned} \quad (89)$$

which yields from P1)

$$\left(\mathbf{B}_2 \mathbf{B}_2^H\right)^\dagger = \mathbf{U} \begin{bmatrix} \mathbf{\Sigma}^{-2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}^H. \quad (90)$$

We can obtain the following from (88) and P1):

$$\mathbf{B}_2^\dagger = \mathbf{V} \begin{bmatrix} \mathbf{\Sigma}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}^H \quad (91)$$

$$\left(\mathbf{B}_2^H\right)^\dagger = \left\{ \mathbf{V} \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}^H \right\}^\dagger = \mathbf{U} \begin{bmatrix} \mathbf{\Sigma}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{V}^H. \quad (92)$$

Since $\mathbf{V}^H \mathbf{V} = \mathbf{I}$, it is clear from (90)–(92) that

$$\left(\mathbf{B}_2 \mathbf{B}_2^H\right)^\dagger = \left(\mathbf{B}_2^H\right)^\dagger \mathbf{B}_2^\dagger. \quad (93)$$

Therefore, from (83), (84), (87), and P2), we have

$$\begin{aligned} \mathbf{Q}_B^\dagger &= \mathbf{Q}_A^\dagger \oplus \left(\mathbf{B}_2 \mathbf{B}_2^H\right)^\dagger \\ &= \mathbf{Q}_A^\dagger + \left(\mathbf{B}_2^H\right)^\dagger \mathbf{B}_2^\dagger \\ &= \mathbf{A}^\dagger - \mathbf{A}^\dagger \mathbf{B}_1 \left(\mathbf{I} + \mathbf{B}_1^H \mathbf{A}^\dagger \mathbf{B}_1\right)^{-1} \mathbf{B}_1^H \mathbf{A}^\dagger + \left(\mathbf{B}_2^H\right)^\dagger \mathbf{B}_2^\dagger. \end{aligned} \quad (94)$$

Equation (94) is equivalent to (5) in the case where $\mathbf{B}_1 = \mathbf{0}$, because $\mathbf{Q} = \mathbf{Q}_B$ in this case [see (3) and (84)].

The second stage of the proof is carried out as follows. Since $R(\mathbf{B}_2) \subset R(\mathbf{A})^\perp = R(\mathbf{U}_1)^\perp = R(\mathbf{U}_2)$, there exists $\mathbf{\Xi}_2 \in \mathbf{C}^{(n-r) \times l}$ such that

$$\mathbf{B}_2 = \mathbf{U}_2 \mathbf{\Xi}_2. \quad (95)$$

From (72) and (95), we have

$$\mathbf{B} = \mathbf{B}_1 \oplus \mathbf{B}_2 = \mathbf{U}_1 \mathbf{\Xi}_1 \oplus \mathbf{U}_2 \mathbf{\Xi}_2 = [\mathbf{U}_1, \mathbf{U}_2] \begin{bmatrix} \mathbf{\Xi}_1 \\ \mathbf{\Xi}_2 \end{bmatrix}. \quad (96)$$

From (3), (71), and (96), we obtain

$$\begin{aligned} \mathbf{Q} &= \mathbf{A} + \mathbf{B} \mathbf{B}^H \\ &= [\mathbf{U}_1, \mathbf{U}_2] \left\{ \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{\Xi}_1 \\ \mathbf{\Xi}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Xi}_1^H & \mathbf{\Xi}_2^H \end{bmatrix} \right\} \begin{bmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{bmatrix} \\ &= [\mathbf{U}_1, \mathbf{U}_2] \tilde{\mathbf{Q}} \begin{bmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{bmatrix} \end{aligned} \quad (97)$$

where

$$\tilde{\mathbf{Q}} := \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{\Xi}_1 \\ \mathbf{\Xi}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Xi}_1^H & \mathbf{\Xi}_2^H \end{bmatrix}. \quad (98)$$

Using P1) and (97), we have

$$\mathbf{Q}^\dagger = [\mathbf{U}_1, \mathbf{U}_2] \tilde{\mathbf{Q}}^\dagger \begin{bmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{bmatrix}. \quad (99)$$

Then, $\tilde{\mathbf{Q}}$ becomes

$$\begin{aligned} \tilde{\mathbf{Q}} &= \left[\begin{array}{c|c} \mathbf{\Lambda}_1 + \mathbf{\Xi}_1 \mathbf{\Xi}_1^H & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{\Xi}_2 \mathbf{\Xi}_2^H \end{array} \right] \\ &\quad + \begin{bmatrix} \mathbf{\Xi}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Xi}_2 \end{bmatrix} \mathbf{P} \begin{bmatrix} \mathbf{\Xi}_1^H & \mathbf{0} \\ \mathbf{0} & \mathbf{\Xi}_2^H \end{bmatrix} \\ &= \tilde{\mathbf{Q}}_B + \begin{bmatrix} \mathbf{\Xi}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Xi}_2 \end{bmatrix} \mathbf{P} \begin{bmatrix} \mathbf{\Xi}_1^H & \mathbf{0} \\ \mathbf{0} & \mathbf{\Xi}_2^H \end{bmatrix} \end{aligned} \quad (100)$$

where

$$\tilde{\mathbf{Q}}_B := \left[\begin{array}{c|c} \mathbf{\Lambda}_1 + \mathbf{\Xi}_1 \mathbf{\Xi}_1^H & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{\Xi}_2 \mathbf{\Xi}_2^H \end{array} \right]. \quad (101)$$

Since

$$\begin{aligned} \text{rank } \tilde{\mathbf{Q}}_B &= \text{rank} \left(\mathbf{\Lambda}_1 + \mathbf{\Xi}_1 \mathbf{\Xi}_1^H \right) + \text{rank} \left(\mathbf{\Xi}_2 \mathbf{\Xi}_2^H \right) \\ &= r + l_2 \end{aligned} \quad (102)$$

with

$$l_2 := \text{rank } \mathbf{B}_2 = \text{rank } \mathbf{\Xi}_2 \quad (103)$$

there exists a unitary matrix \mathbf{V} and a diagonal matrix $\tilde{\mathbf{\Lambda}}$ such that

$$\begin{aligned} \tilde{\mathbf{Q}}_B &= \mathbf{V} \tilde{\mathbf{\Lambda}} \mathbf{V}^H \\ &= [\mathbf{V}_1, \mathbf{V}_2] \begin{bmatrix} \tilde{\mathbf{\Lambda}}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^H \\ \mathbf{V}_2^H \end{bmatrix} \\ &= \mathbf{V}_1 \tilde{\mathbf{\Lambda}}_1 \mathbf{V}_1^H \end{aligned} \quad (104)$$

where $\mathbf{V}_1 \in \mathbf{C}^{n \times (r+l_2)}$, $\mathbf{V}_2 \in \mathbf{C}^{n \times (n-r-l_2)}$, and $\tilde{\mathbf{\Lambda}}_1$ is a non-singular diagonal matrix belonging to $\mathbf{R}^{(r+l_2) \times (r+l_2)}$. Since

$$R \left(\begin{bmatrix} \mathbf{\Xi}_1 \\ \mathbf{0} \end{bmatrix} \right), R \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{\Xi}_2 \end{bmatrix} \right) \subset R(\tilde{\mathbf{Q}}_B) = R(\mathbf{V}_1) \quad (105)$$

there exist $\mathbf{H}_1 \in \mathbf{C}^{(r+l_2) \times l}$ and $\mathbf{H}_2 \in \mathbf{C}^{(r+l_2) \times l}$ such that

$$\begin{bmatrix} \mathbf{\Xi}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Xi}_2 \end{bmatrix} = \mathbf{V}_1 [\mathbf{H}_1, \mathbf{H}_2]. \quad (106)$$

Therefore, from (100), (104), and (106), we have

$$\begin{aligned} \tilde{\mathbf{Q}} &= \tilde{\mathbf{Q}}_B + \begin{bmatrix} \mathbf{\Xi}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Xi}_2 \end{bmatrix} \mathbf{P} \begin{bmatrix} \mathbf{\Xi}_1^H & \mathbf{0} \\ \mathbf{0} & \mathbf{\Xi}_2^H \end{bmatrix} \\ &= \mathbf{V}_1 \tilde{\mathbf{\Lambda}}_1 \mathbf{V}_1^H + \mathbf{V}_1 [\mathbf{H}_1, \mathbf{H}_2] \mathbf{P} [\mathbf{H}_1, \mathbf{H}_2]^H \mathbf{V}_1^H \\ &= \mathbf{V}_1 \left\{ \tilde{\mathbf{\Lambda}}_1 + [\mathbf{H}_1, \mathbf{H}_2] \mathbf{P} [\mathbf{H}_1, \mathbf{H}_2]^H \right\} \mathbf{V}_1^H. \end{aligned} \quad (107)$$

Put

$$\mathbf{Q}_C := \tilde{\mathbf{\Lambda}}_1 + [\mathbf{H}_1, \mathbf{H}_2] \mathbf{P} [\mathbf{H}_1, \mathbf{H}_2]^H \in \mathbf{C}^{(r+l_2) \times (r+l_2)}. \quad (108)$$

Since $\mathbf{B}_2 \neq \mathbf{0}$, we have $R(\mathbf{B}) \not\subset R(\mathbf{A})$. Therefore

$$\text{rank } \mathbf{Q} = \text{rank} (\mathbf{A} + \mathbf{B} \mathbf{B}^H) = \text{rank } \mathbf{A} + l_2 = r + l_2. \quad (109)$$

This means that, from (97) and (98)

$$\text{rank } \tilde{\mathbf{Q}} = \text{rank} (\mathbf{U}^H \mathbf{Q} \mathbf{U}) = \text{rank } \mathbf{Q} = r + l_2 \quad (110)$$

which implies that, from (107) and (108)

$$\text{rank } \mathbf{Q}_C = \text{rank } \tilde{\mathbf{Q}} = r + l_2. \quad (111)$$

Therefore, \mathbf{Q}_C is nonsingular. By a generalized matrix inversion lemma in [3] and [23]

$$\mathbf{Q}_C^{-1} = \tilde{\Lambda}_1^{-1} - \tilde{\Lambda}_1^{-1} [\mathbf{H}_1, \mathbf{H}_2] \mathbf{Q}_D^{-1} [\mathbf{H}_1, \mathbf{H}_2]^H \tilde{\Lambda}_1^{-1} \quad (112)$$

where

$$\mathbf{Q}_D := \mathbf{P}^{-1} + [\mathbf{H}_1, \mathbf{H}_2]^H \tilde{\Lambda}_1^{-1} [\mathbf{H}_1, \mathbf{H}_2] \in \mathbf{C}^{2l \times 2l}. \quad (113)$$

Here, we note that it can be shown that \mathbf{Q}_D is also nonsingular if \mathbf{Q}_C is nonsingular (See Appendix B). Since $\mathbf{P}^{-1} = \mathbf{P}$, \mathbf{Q}_D becomes

$$\begin{aligned} \mathbf{Q}_D &= \mathbf{P} + [\mathbf{H}_1, \mathbf{H}_2]^H \tilde{\Lambda}_1^{-1} [\mathbf{H}_1, \mathbf{H}_2] \\ &= \left[\begin{array}{c|c} \mathbf{H}_1^H \tilde{\Lambda}_1^{-1} \mathbf{H}_1 & \mathbf{I} + \mathbf{H}_1^H \tilde{\Lambda}_1^{-1} \mathbf{H}_2 \\ \hline \mathbf{I} + \mathbf{H}_2^H \tilde{\Lambda}_1^{-1} \mathbf{H}_1 & \mathbf{H}_2^H \tilde{\Lambda}_1^{-1} \mathbf{H}_2 \end{array} \right] \in \mathbf{C}^{2l \times 2l}. \end{aligned} \quad (114)$$

From (104)–(114) along with P3), we have

$$\begin{aligned} \tilde{\mathbf{Q}}^\dagger &= (\mathbf{V}_1 \mathbf{Q}_C \mathbf{V}_1^H)^\dagger \\ &= (\mathbf{V}_1^H)^\dagger (\mathbf{V}_1 \mathbf{Q}_C)^\dagger \\ &= \mathbf{V}_1 \mathbf{Q}_C^{-1} \mathbf{V}_1^H \\ &= \mathbf{V}_1 \tilde{\Lambda}_1^{-1} \mathbf{V}_1^H - \mathbf{V}_1 \tilde{\Lambda}_1^{-1} [\mathbf{H}_1, \mathbf{H}_2] \\ &\quad \times \mathbf{Q}_D^{-1} [\mathbf{H}_1, \mathbf{H}_2]^H \tilde{\Lambda}_1^{-1} \mathbf{V}_1^H \\ &= \tilde{\mathbf{Q}}_B^\dagger - \tilde{\mathbf{Q}}_B^\dagger \begin{bmatrix} \Xi_1 & \mathbf{0} \\ \mathbf{0} & \Xi_2 \end{bmatrix} \mathbf{Q}_D^{-1} \begin{bmatrix} \Xi_1^H & \mathbf{0} \\ \mathbf{0} & \Xi_2^H \end{bmatrix} \tilde{\mathbf{Q}}_B^\dagger \end{aligned} \quad (115)$$

where we have used the following from (104) and (106):

$$\tilde{\mathbf{Q}}_B^\dagger = \mathbf{V}_1 \tilde{\Lambda}_1^{-1} \mathbf{V}_1^H \quad (116)$$

$$\begin{aligned} \mathbf{V}_1 \tilde{\Lambda}_1^{-1} [\mathbf{H}_1, \mathbf{H}_2] &= \mathbf{V}_1 \tilde{\Lambda}_1^{-1} \mathbf{V}_1^H \mathbf{V}_1 [\mathbf{H}_1, \mathbf{H}_2] \\ &= \tilde{\mathbf{Q}}_B^\dagger \begin{bmatrix} \Xi_1 & \mathbf{0} \\ \mathbf{0} & \Xi_2 \end{bmatrix}. \end{aligned} \quad (117)$$

Using (99) and (115) along with P3), we obtain

$$\begin{aligned} \mathbf{Q}^\dagger &= \mathbf{U} \tilde{\mathbf{Q}}^\dagger \mathbf{U}^H \\ &= \mathbf{U} \tilde{\mathbf{Q}}_B^\dagger \mathbf{U}^H - \mathbf{U} \tilde{\mathbf{Q}}_B^\dagger \begin{bmatrix} \Xi_1 & \mathbf{0} \\ \mathbf{0} & \Xi_2 \end{bmatrix} \\ &\quad \times \mathbf{Q}_D^{-1} \begin{bmatrix} \Xi_1^H & \mathbf{0} \\ \mathbf{0} & \Xi_2^H \end{bmatrix} \tilde{\mathbf{Q}}_B^\dagger \mathbf{U}^H \\ &= \mathbf{Q}_B^\dagger - \mathbf{Q}_B^\dagger [\mathbf{B}_1, \mathbf{B}_2] \mathbf{Q}_D^{-1} [\mathbf{B}_1, \mathbf{B}_2]^H \mathbf{Q}_B^\dagger \end{aligned} \quad (118)$$

where we have used

$$\mathbf{U} \tilde{\mathbf{Q}}_B^\dagger \mathbf{U}^H = \mathbf{Q}_B^\dagger \quad (119)$$

$$\mathbf{U} \tilde{\mathbf{Q}}_B^\dagger \begin{bmatrix} \Xi_1 & \mathbf{0} \\ \mathbf{0} & \Xi_2 \end{bmatrix} = \mathbf{Q}_B^\dagger [\mathbf{B}_1, \mathbf{B}_2]. \quad (120)$$

The first equation (119) holds true, because P1) holds true and

$$\begin{aligned} \mathbf{U} \tilde{\mathbf{Q}}_B \mathbf{U}^H &= [\mathbf{U}_1, \mathbf{U}_2] \left[\begin{array}{c|c} \Lambda_1 + \Xi_1 \Xi_1^H & \mathbf{0} \\ \hline \mathbf{0} & \Xi_2 \Xi_2^H \end{array} \right] \begin{bmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{bmatrix} \\ &= \mathbf{U}_1 (\Lambda_1 + \Xi_1 \Xi_1^H) \mathbf{U}_1^H + \mathbf{U}_2 \Xi_2 \Xi_2^H \mathbf{U}_2^H \\ &= \mathbf{A} + \mathbf{B}_1 \mathbf{B}_1^H + \mathbf{B}_2 \mathbf{B}_2^H \\ &= \mathbf{Q}_B. \end{aligned} \quad (121)$$

The second equation (120) also holds true, because

$$\begin{aligned} \mathbf{U} \tilde{\mathbf{Q}}_B^\dagger \begin{bmatrix} \Xi_1 & \mathbf{0} \\ \mathbf{0} & \Xi_2 \end{bmatrix} &= \mathbf{U} \tilde{\mathbf{Q}}_B^\dagger \mathbf{U}^H \mathbf{U} \begin{bmatrix} \Xi_1 & \mathbf{0} \\ \mathbf{0} & \Xi_2 \end{bmatrix} \\ &= \mathbf{Q}_B^\dagger [\mathbf{U}_1, \mathbf{U}_2] \begin{bmatrix} \Xi_1 & \mathbf{0} \\ \mathbf{0} & \Xi_2 \end{bmatrix} \\ &= \mathbf{Q}_B^\dagger [\mathbf{B}_1, \mathbf{B}_2]. \end{aligned} \quad (122)$$

On the other hand, from (106) and (117), we obtain

$$\begin{aligned} \mathbf{H}_1^H \tilde{\Lambda}_1^{-1} \mathbf{H}_2 &= [\Xi_1^H, \mathbf{0}] \mathbf{V}_1 \tilde{\Lambda}_1^{-1} \mathbf{V}_1^H \begin{bmatrix} \mathbf{0} \\ \Xi_2 \end{bmatrix} \\ &= [\Xi_1^H, \mathbf{0}] \tilde{\mathbf{Q}}_B^\dagger \begin{bmatrix} \mathbf{0} \\ \Xi_2 \end{bmatrix} \\ &= [\Xi_1^H \tilde{\mathbf{Q}}_B^\dagger, \mathbf{0}] \begin{bmatrix} \mathbf{0} \\ \Xi_2 \end{bmatrix} \\ &= 0. \end{aligned} \quad (123)$$

Similarly, we obtain

$$\begin{aligned} \mathbf{H}_2^H \tilde{\Lambda}_1^{-1} \mathbf{H}_1 &= [\mathbf{0}, \Xi_2^H] \mathbf{V}_1 \tilde{\Lambda}_1^{-1} \mathbf{V}_1^H \begin{bmatrix} \Xi_1 \\ \mathbf{0} \end{bmatrix} \\ &= [\mathbf{0}, \Xi_2^H] \tilde{\mathbf{Q}}_B^\dagger \begin{bmatrix} \Xi_1 \\ \mathbf{0} \end{bmatrix} \\ &= [\mathbf{0}, \Xi_2^H \tilde{\mathbf{Q}}_B^\dagger] \begin{bmatrix} \Xi_1 \\ \mathbf{0} \end{bmatrix} \\ &= 0. \end{aligned} \quad (124)$$

Furthermore

$$\begin{aligned} \mathbf{H}_1^H \tilde{\Lambda}_1^{-1} \mathbf{H}_1 &= [\Xi_1^H, \mathbf{0}] \mathbf{V}_1 \tilde{\Lambda}_1^{-1} \mathbf{V}_1^H \begin{bmatrix} \Xi_1 \\ \mathbf{0} \end{bmatrix} \\ &= [\Xi_1^H, \mathbf{0}] \tilde{\mathbf{Q}}_B^\dagger \begin{bmatrix} \Xi_1 \\ \mathbf{0} \end{bmatrix} \\ &= \mathbf{B}_1^H \mathbf{U} \tilde{\mathbf{Q}}_B^\dagger \mathbf{U}^H \mathbf{B}_1 \\ &= \mathbf{B}_1^H \mathbf{Q}_B^\dagger \mathbf{B}_1 \end{aligned} \quad (125)$$

where we have used from (96)

$$\begin{bmatrix} \Xi_1 \\ \mathbf{0} \end{bmatrix} = \mathbf{U}^H \mathbf{B}_1 \quad (126)$$

and (119). Similarly, we obtain

$$\begin{aligned} H_2^H \tilde{\Lambda}_1^{-1} H_2 &= \begin{bmatrix} \mathbf{0}, \Xi_2^H \\ \mathbf{0}, \Xi_2 \end{bmatrix} V_1 \tilde{\Lambda}_1^{-1} V_1^H \begin{bmatrix} \mathbf{0} \\ \Xi_2 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{0}, \Xi_2^H \\ \mathbf{0}, \Xi_2 \end{bmatrix} \tilde{Q}_B^\dagger \begin{bmatrix} \mathbf{0} \\ \Xi_2 \end{bmatrix} \\ &= B_2^H U \tilde{Q}_B^\dagger U^H B_2 \\ &= B_2^H Q_B^\dagger B_2. \end{aligned} \quad (127)$$

Using (123)–(125) and (127), we have the following from (114):

$$\begin{aligned} Q_D &= P + [H_1, H_2]^H \tilde{\Lambda}_1^{-1} [H_1, H_2] \\ &= P + [B_1, B_2]^H Q_B^\dagger [B_1, B_2] \\ &= \left[\begin{array}{c|c} B_1^H Q_B^\dagger B_1 & I \\ \hline I & B_2^H Q_B^\dagger B_2 \end{array} \right] \in \mathcal{C}^{2l \times 2l}. \end{aligned} \quad (128)$$

If Q_D is nonsingular, we obtain the inverse of Q_D as follows:

$$Q_D^{-1} = \left[\begin{array}{c|c} -\Delta^{-1} B_2^H Q_B^\dagger B_2 & \Delta^{-1} \\ \hline I + B_1^H Q_B^\dagger B_1 \Delta^{-1} B_2^H Q_B^\dagger B_2 & -B_1^H Q_B^\dagger B_1 \Delta^{-1} \end{array} \right] \quad (129)$$

with

$$\Delta := I - B_2^H Q_B^\dagger B_2 B_1^H Q_B^\dagger B_1. \quad (130)$$

This completes the proof.

APPENDIX B

THE PROOF OF THE NONSINGULARITY OF MATRIX Q_D

We use the following lemma.

Lemma B1: Let

$$E_1 := B_1^H Q_B^\dagger B_1 \in \mathcal{C}^{l \times l} \quad (131)$$

$$E_2 := B_2^H Q_B^\dagger B_2 \in \mathcal{C}^{l \times l} \quad (132)$$

for brevity. Then, the following statements hold true.

- 1) E_1 is a Hermitian matrix such that $I - E_1 > 0$.
- 2) E_2 is the orthogonal projection of \mathcal{C}^l onto $R(B_2^H)$.

Proof: It is clear from the definitions that E_1 and E_2 are both Hermitian. Because

$$\begin{aligned} Q_B^\dagger &= (A + B_1 B_1^H)^\dagger \oplus (B_2 B_2^H)^\dagger \\ &= \left\{ A^\dagger - A^\dagger B_1 (I + B_1^H A^\dagger B_1)^{-1} B_1^H A^\dagger \right\} \oplus (B_2^H)^\dagger B_2^\dagger \end{aligned} \quad (133)$$

we have

$$\begin{aligned} B_1^H Q_B^\dagger B_1 &= B_1^H A^\dagger B_1 - B_1^H A^\dagger B_1 (I + B_1^H A^\dagger B_1)^{-1} \\ &\quad \times B_1^H A^\dagger B_1 \end{aligned} \quad (134)$$

where we used

$$B_1^H (B_2^H)^\dagger B_2^\dagger B_1 = 0. \quad (135)$$

Let

$$F_1 := B_1^H A^\dagger B_1 \quad (136)$$

$$F_2 := B_2^H A^\dagger B_2 \quad (137)$$

for brevity. Then, it follows from (134) along with the definitions of E_1 and F_1 that

$$\begin{aligned} I - E_1 &= I - F_1 + F_1 (I + F_1)^{-1} F_1 \\ &= I - (I + F_1) (I + F_1)^{-1} F_1 + F_1 (I + F_1)^{-1} F_1 \\ &= I - (I + F_1)^{-1} F_1 \\ &= (I + F_1)^{-1} (I + F_1) - (I + F_1)^{-1} F_1 \\ &= (I + F_1)^{-1} > 0. \end{aligned} \quad (138)$$

It follows from (133) and the definition of E_2 that

$$\begin{aligned} E_2 &= B_2^H Q_B^\dagger B_2 \\ &= B_2^H \left\{ (A + B_1 B_1^H)^\dagger \oplus (B_2 B_2^H)^\dagger \right\} B_2 \\ &= B_2^H (B_2^H)^\dagger B_2^\dagger B_2 \\ &= P_{R(B_2^H)} \end{aligned} \quad (139)$$

where we used $P_{R(B_2^H)} = B_2^\dagger B_2$, $P_{R(B_2^H)} = P_{R(B_2^H)}$, and $P_{R(B_2^H)} = P_{R(B_2^H)} P_{R(B_2^H)}$. It is clear from (139) that E_2 is the orthogonal projection of \mathcal{C}^l onto $R(B_2^H)$. This completes the proof.

Now let us prove the regularity of Q_D . We use the following lemma on the inverse of the block matrix.

Lemma B2: Let $A \in \mathcal{C}^{n \times n}$, $B \in \mathcal{C}^{r \times r}$, $C \in \mathcal{C}^{r \times n}$, $D \in \mathcal{C}^{n \times r}$, and $\Delta := B - CA^{-1}D \in \mathcal{C}^{r \times r}$, where we assume A^{-1} exists. If Δ^{-1} exists, then

$$\left[\begin{array}{c|c} A & D \\ \hline C & B \end{array} \right]^{-1} = \left[\begin{array}{c|c} A^{-1} + E\Delta^{-1}F & -E\Delta^{-1} \\ \hline -\Delta^{-1}F & -\Delta^{-1} \end{array} \right] \quad (140)$$

where $E = A^{-1}D$ and $F = CA^{-1}$.

This lemma is found in [23, p. 656]. It follows from (128) and the definitions of E_1 and E_2 that

$$Q_D \begin{bmatrix} \mathbf{0} & I \\ I & \mathbf{0} \end{bmatrix} = \left[\begin{array}{c|c} I & B_1^H Q_B^\dagger B_1 \\ \hline B_2^H Q_B^\dagger B_2 & I \end{array} \right] = \begin{bmatrix} I & E_1 \\ E_2 & I \end{bmatrix}. \quad (141)$$

Let

$$T := \begin{bmatrix} I & E_1 \\ E_2 & I \end{bmatrix} \in \mathcal{C}^{2l \times 2l} \quad (142)$$

for brevity. Instead of showing the nonsingularity of Q_D , we show that T^{-1} exists. Applying Lemma B2 to T in (142), we show that there exists the inverse of Δ defined by

$$\Delta := I - E_2 E_1. \quad (143)$$

Applying Lemma B1 to Δ in (143), Δ becomes

$$\Delta = I - P_{R(B_2^H)} E_1. \quad (144)$$

We shall show that $N(\Delta) = \{0\}$. Therefore, we suppose

$$\Delta \xi = 0 \quad \text{for } \xi \in \mathcal{C}^l. \quad (145)$$

Then, it follows from (144) that

$$\xi = P_{R(B_2^H)} E_1 \xi. \quad (146)$$

Because ξ can be decomposed as

$$\xi = \xi_1 \oplus \xi_2 \quad (147)$$

where $\xi_1 := P_{R(\mathbf{B}_2^H)}^\dagger \xi$ and $\xi_2 := P_{R(\mathbf{B}_2^H)} \xi$, it follows from (146) that

$$\xi_1 = 0 \quad (148)$$

which gives the following, along with (146) and (147)

$$\xi_2 = P_{R(\mathbf{B}_2^H)} \mathbf{E}_1 \xi_2. \quad (149)$$

Since $\xi_2 := P_{R(\mathbf{B}_2^H)} \xi_2$, this gives

$$\xi_2^H \xi_2 = \xi_2^H \mathbf{E}_1 \xi_2 \quad (150)$$

which is equivalent to

$$\xi_2^H (\mathbf{I} - \mathbf{E}_1) \xi_2 = 0. \quad (151)$$

Applying 1) of Lemma B1 to (151), we have

$$\xi_2^H = 0. \quad (152)$$

Therefore, we obtain $\xi = 0$, which shows $N(\Delta) = \{0\}$. This completes the proof of the nonsingularity of \mathbf{Q}_D .

APPENDIX C THE PROOF OF THEOREM 1

First, we consider case 1). If $\mathbf{B}_2 = 0$, then \mathbf{Q}_B^\dagger and \mathbf{Q}_D^{-1} are represented from (7) and (8), respectively, as follows:

$$\begin{aligned} \mathbf{Q}_B^\dagger &= (\mathbf{A} + \mathbf{B}_1 \mathbf{B}_1^H)^\dagger \\ &= \mathbf{A}^\dagger - \mathbf{A}^\dagger \mathbf{B}_1 (\mathbf{I} + \mathbf{B}_1^H \mathbf{A}^\dagger \mathbf{B}_1)^{-1} \mathbf{B}_1^H \mathbf{A}^\dagger \end{aligned} \quad (153)$$

$$\mathbf{Q}_D^{-1} = \left[\begin{array}{c|c} 0 & \mathbf{I} \\ \hline \mathbf{I} & -\mathbf{B}_1^H \mathbf{Q}_B^\dagger \mathbf{B}_1 \end{array} \right]. \quad (154)$$

Then, from (6), (153), and (154), we have \mathbf{Q}^\dagger as follows:

$$\begin{aligned} \mathbf{Q}^\dagger &= \mathbf{Q}_B^\dagger - \mathbf{Q}_B^\dagger [\mathbf{B}_1, 0] \left[\begin{array}{c|c} -0 & \mathbf{I} \\ \hline \mathbf{I} & -\mathbf{B}_1^H \mathbf{Q}_B^\dagger \mathbf{B}_1 \end{array} \right] [\mathbf{B}_1, 0]^H \mathbf{Q}_B^\dagger \\ &= \mathbf{Q}_B^\dagger \\ &= \mathbf{A}^\dagger - \mathbf{A}^\dagger \mathbf{B}_1 (\mathbf{I} + \mathbf{B}_1^H \mathbf{A}^\dagger \mathbf{B}_1)^{-1} \mathbf{B}_1^H \mathbf{A}^\dagger \end{aligned} \quad (155)$$

which is identical to (4).

Second, we consider case 2). If $\mathbf{B}_2 \neq 0$ and $\mathbf{B}_1 = 0$, then \mathbf{Q}_B^\dagger and \mathbf{Q}_D^{-1} are represented from (7) and (8), respectively, as follows:

$$\begin{aligned} \mathbf{Q}_B^\dagger &= (\mathbf{A} + \mathbf{B}_2 \mathbf{B}_2^H)^\dagger \\ &= \mathbf{A}^\dagger + (\mathbf{B}_2^H)^\dagger \mathbf{B}_2^\dagger, \end{aligned} \quad (156)$$

$$\mathbf{Q}_D^{-1} = \left[\begin{array}{c|c} -\mathbf{B}_2^H \mathbf{Q}_B^\dagger \mathbf{B}_2 & \mathbf{I} \\ \hline \mathbf{I} & 0 \end{array} \right]. \quad (157)$$

Then, from (6), (156), and (157), we have \mathbf{Q}^\dagger as follows:

$$\begin{aligned} \mathbf{Q}^\dagger &= \mathbf{Q}_B^\dagger - \mathbf{Q}_B^\dagger [0, \mathbf{B}_2] \left[\begin{array}{c|c} -\mathbf{B}_2^H \mathbf{Q}_B^\dagger \mathbf{B}_2 & \mathbf{I} \\ \hline \mathbf{I} & 0 \end{array} \right] [0, \mathbf{B}_2]^H \mathbf{Q}_B^\dagger \\ &= \mathbf{Q}_B^\dagger \\ &= \mathbf{A}^\dagger + (\mathbf{B}_2^H)^\dagger \mathbf{B}_2^\dagger \end{aligned} \quad (158)$$

which is identical to (5). This completes the proof.

APPENDIX D ILLUSTRATIVE EXAMPLES

We consider a single matrix \mathbf{A} with three different values of matrix \mathbf{B} as follows:

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} [1 \ 1 \ 1 \ 1] + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} [0 \ 0 \ 0 \ 1] \\ &= \mathbf{A}_b + \mathbf{x}\mathbf{x}^H \end{aligned} \quad (159)$$

(Case 1)

$$\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \quad (160)$$

(Case 2)

$$\mathbf{B} = \begin{bmatrix} 1 & -2 \\ 1 & 1 \\ -2 & 1 \\ 0 & 0 \end{bmatrix} \quad (161)$$

(Case 3)

$$\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \quad (162)$$

where \mathbf{A}_b and \mathbf{x} are a matrix and a vector which are defined by

$$\mathbf{A}_b = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (163)$$

$$\mathbf{x} = [0 \ 0 \ 0 \ 1]^H. \quad (164)$$

Then, putting $\mathbf{Q} := \mathbf{A} + \mathbf{B}\mathbf{B}^H$, the matrix \mathbf{Q} is calculated for each vector \mathbf{B} , as follows:

(Case 1)

$$\begin{aligned} \mathbf{Q} &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 4 & 4 & 4 & 7 \end{bmatrix} \end{aligned} \quad (165)$$

(Case 2)

$$\begin{aligned} \mathbf{Q} &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 1 & -2 \\ 1 & 1 \\ -2 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & -2 & 0 \\ -2 & 1 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 6 & 0 & -3 & 1 \\ 0 & 3 & 0 & 1 \\ -3 & 0 & 6 & 2 \\ 1 & 1 & 1 & 2 \end{bmatrix} \end{aligned} \quad (166)$$

(Case 3)

$$\begin{aligned} \mathbf{Q} &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 4 & 5 & 6 \\ 4 & 6 & 8 & 10 \\ 5 & 8 & 11 & 14 \\ 6 & 10 & 14 & 19 \end{bmatrix}. \end{aligned} \quad (167)$$

Since $\mathbf{A}\mathbf{A}^\dagger$ is the orthogonal projection of \mathbf{C}^4 onto $R(\mathbf{A})$ and $(\mathbf{I} - \mathbf{A}\mathbf{A}^\dagger)$ is the orthogonal projection of \mathbf{C}^4 onto $R(\mathbf{A})^\perp$ [15], [16], it follows from (2) that the matrix \mathbf{B} is decomposed into two matrices \mathbf{B}_1 and \mathbf{B}_2 by using the following relations:

$$\mathbf{B}_1 = \mathbf{A}\mathbf{A}^\dagger\mathbf{B} \quad \text{with} \quad R(\mathbf{B}_1) \subset R(\mathbf{A}) \quad (168)$$

$$\mathbf{B}_2 = (\mathbf{I} - \mathbf{A}\mathbf{A}^\dagger)\mathbf{B} \quad \text{with} \quad R(\mathbf{B}_2) \subset R(\mathbf{A})^\perp \quad (169)$$

where

$$\mathbf{A}^\dagger = (\mathbf{A}_b + \mathbf{x}\mathbf{x}^H)^\dagger = \frac{1}{9} \begin{bmatrix} 2 & 2 & 2 & -3 \\ 2 & 2 & 2 & -3 \\ 2 & 2 & 2 & -3 \\ -3 & -3 & -3 & 9 \end{bmatrix} \quad (170)$$

(see the matrix pseudoinversion lemma for $(\mathbf{A} + \mathbf{b}\mathbf{b}^h)^\dagger$ in [9]) and \mathbf{I} is the 4×4 identity matrix.

First, we consider the first case, i.e., case 1. From (168)–(170), we have

$$\mathbf{B}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \quad \mathbf{B}_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (171)$$

where we note that $R(\mathbf{B}) \subset R(\mathbf{A})$. Therefore, the pseudoinverse \mathbf{Q}^\dagger of the matrix $\mathbf{Q} = \mathbf{A} + \mathbf{B}\mathbf{B}^H$ is obtained by applying (4) which is the case when $\mathbf{B}_2 = 0$ in Lemma 1 as follows:

$$\begin{aligned} \mathbf{Q}^\dagger &= \mathbf{A}^\dagger - \mathbf{A}^\dagger\mathbf{B}_1 \left(\mathbf{I} + \mathbf{B}_1^H\mathbf{A}^\dagger\mathbf{B}_1 \right)^{-1} \mathbf{B}_1^H\mathbf{A}^\dagger \\ &= \frac{1}{9} \begin{bmatrix} 2 & 2 & 2 & -3 \\ 2 & 2 & 2 & -3 \\ 2 & 2 & 2 & -3 \\ -3 & -3 & -3 & 9 \end{bmatrix} - \frac{1}{9} \begin{bmatrix} 2 & 2 & 2 & -3 \\ 2 & 2 & 2 & -3 \\ 2 & 2 & 2 & -3 \\ -3 & -3 & -3 & 9 \end{bmatrix} \\ &\quad \times \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix} \frac{1}{9} \right. \\ &\quad \left. \times \begin{bmatrix} 2 & 2 & 2 & -3 \\ 2 & 2 & 2 & -3 \\ 2 & 2 & 2 & -3 \\ -3 & -3 & -3 & 9 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \right)^{-1} \\ &= \frac{1}{45} \begin{bmatrix} 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ -12 & -12 & -12 & 27 \end{bmatrix}. \end{aligned} \quad (172)$$

As an evaluation of the value of the pseudoinverse \mathbf{Q}^\dagger , we consider the following error matrices defined by using the Moore–Penrose conditions [16], [22]:

$$\mathbf{E}_1 := \mathbf{Q}\mathbf{Q}^\dagger\mathbf{Q} - \mathbf{Q} \quad (173)$$

$$\mathbf{E}_2 := \mathbf{Q}^\dagger\mathbf{Q}\mathbf{Q}^\dagger - \mathbf{Q}^\dagger \quad (174)$$

$$\mathbf{E}_3 := (\mathbf{Q}\mathbf{Q}^\dagger)^H - \mathbf{Q}\mathbf{Q}^\dagger \quad (175)$$

$$\mathbf{E}_4 := (\mathbf{Q}^\dagger\mathbf{Q})^H - \mathbf{Q}^\dagger\mathbf{Q}. \quad (176)$$

We note that all the error matrices vanish when the value of \mathbf{Q}^\dagger is correct. Substituting (165) and (172) into (173)–(176), we obtain

$$\begin{aligned} \mathbf{E}_1 &= \begin{bmatrix} 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 4 & 4 & 4 & 7 \end{bmatrix} - \frac{1}{45} \begin{bmatrix} 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ -12 & -12 & -12 & 27 \end{bmatrix} \\ &\quad \times \begin{bmatrix} 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 4 & 4 & 4 & 7 \end{bmatrix} - \begin{bmatrix} 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 4 & 4 & 4 & 7 \end{bmatrix} \\ &= 0 \end{aligned} \quad (177)$$

$$\begin{aligned}
 \mathbf{E}_2 &= \frac{1}{45} \begin{bmatrix} 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ -12 & -12 & -12 & 27 \end{bmatrix} \begin{bmatrix} 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 4 & 4 & 4 & 7 \end{bmatrix} \\
 &\times \frac{1}{45} \begin{bmatrix} 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ -12 & -12 & -12 & 27 \end{bmatrix} \\
 &- \frac{1}{45} \begin{bmatrix} 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ -12 & -12 & -12 & 27 \end{bmatrix} \\
 &= 0
 \end{aligned} \tag{178}$$

$$\begin{aligned}
 \mathbf{E}_3 &= \left(\begin{bmatrix} 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 4 & 4 & 4 & 7 \end{bmatrix} \frac{1}{45} \begin{bmatrix} 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ -12 & -12 & -12 & 27 \end{bmatrix} \right)^H \\
 &- \begin{bmatrix} 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 4 & 4 & 4 & 7 \end{bmatrix} \frac{1}{45} \begin{bmatrix} 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ -12 & -12 & -12 & 27 \end{bmatrix} \\
 &= 0
 \end{aligned} \tag{179}$$

$$\begin{aligned}
 \mathbf{E}_4 &= \left(\frac{1}{45} \begin{bmatrix} 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ -12 & -12 & -12 & 27 \end{bmatrix} \begin{bmatrix} 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 4 & 4 & 4 & 7 \end{bmatrix} \right)^H \\
 &- \frac{1}{45} \begin{bmatrix} 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ 7 & 7 & 7 & -12 \\ -12 & -12 & -12 & 27 \end{bmatrix} \begin{bmatrix} 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 4 & 4 & 4 & 7 \end{bmatrix} \\
 &= 0.
 \end{aligned} \tag{180}$$

From (177)–(180), these results prove that the value of the pseudoinverse \mathbf{Q}^\dagger is correct in Case 1.

Similarly, we obtain the pseudoinverse \mathbf{Q}^\dagger and evaluate its value in the second and third cases, i.e., Cases 2 and 3. From (168)–(170), two matrices \mathbf{B}_1 and \mathbf{B}_2 are obtained in each case as follows:

(Case 2)

$$\mathbf{B}_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \mathbf{B}_2 = \begin{bmatrix} 1 & -2 \\ 1 & 1 \\ -2 & 1 \\ 0 & 0 \end{bmatrix} \tag{181}$$

(Case 3)

$$\mathbf{B}_1 = \begin{bmatrix} 2 & 1 \\ 2 & 1 \\ 2 & 1 \\ 4 & 1 \end{bmatrix} \quad \mathbf{B}_2 = \begin{bmatrix} -1 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}. \tag{182}$$

The pseudoinverse \mathbf{Q}^\dagger of the matrix $\mathbf{Q} = \mathbf{A} + \mathbf{B}\mathbf{B}^H$ is calculated from (5) in Case 2 and from (6)–(10) in Case 3 as follows:

(Case 2)

$$\mathbf{Q}^\dagger = \frac{1}{9} \begin{bmatrix} 3 & 1 & 2 & -3 \\ 1 & 4 & 1 & -3 \\ 2 & 1 & 3 & -3 \\ -3 & -3 & -3 & 9 \end{bmatrix} \tag{183}$$

(Case 3)

$$\mathbf{Q}^\dagger = \frac{1}{12} \begin{bmatrix} 19 & 0 & -19 & 8 \\ 0 & 2 & 4 & -4 \\ -19 & 4 & 27 & -16 \\ 8 & -4 & -16 & 12 \end{bmatrix} \tag{184}$$

where Case 2 corresponds to the case when $\mathbf{B}_2 \neq 0$ and $\mathbf{B}_1 = 0$ in Lemma 1 and Case 3 corresponds to the case when $\mathbf{B}_2 \neq 0$ and $\mathbf{B}_1 \neq 0$ in Lemma 1.

Then, the values of the pseudoinverse \mathbf{Q}^\dagger are evaluated by using the error matrices in (173)–(176). Because all the error matrices become zero, it can be seen that the values of the pseudoinverse \mathbf{Q}^\dagger are correct in both cases.

Therefore, these examples may convince ourselves that Lemma 1 is true.

APPENDIX E THE PROOF OF THEOREM 2

Substituting (37) and $\mathbf{P}(k) = \tilde{\mathbf{R}}^\dagger(k)$ to (35), we have

$$\begin{aligned}
 \tilde{\mathbf{w}}_i(k) &= \mathbf{P}(k) \left[(1 - \alpha_k) \tilde{\mathbf{d}}_i(k-1) \right. \\
 &\quad \left. + \alpha_k \sum_{j=0}^{r-1} \tilde{\mathbf{y}}^* \{ (k-1)r+j \} \tilde{\mathbf{z}}_i \{ (k-1)r+j \} \right].
 \end{aligned} \tag{185}$$

On the other hand, from (32) and (35)–(39), we obtain the following:

$$\begin{aligned}
 \tilde{\mathbf{d}}_i(k-1) &= \tilde{\mathbf{R}}(k-1) \tilde{\mathbf{w}}_i(k-1) \\
 &= \frac{1}{1 - \alpha_k} \left\{ \tilde{\mathbf{R}}(k) - \alpha_k \tilde{\mathbf{B}}^*(k) \tilde{\mathbf{B}}^T(k) \right\} \tilde{\mathbf{w}}_i(k-1).
 \end{aligned} \tag{186}$$

Substituting (186) to (185), the recursion formula for the time updating of the tap vector $\tilde{\mathbf{w}}_i(k)$ is obtained as follows:

$$\begin{aligned}
 \tilde{\mathbf{w}}_i(k) &= \mathbf{P}(k) \left[\left\{ \tilde{\mathbf{R}}(k) - \alpha_k \tilde{\mathbf{B}}^*(k) \tilde{\mathbf{B}}^T(k) \right\} \tilde{\mathbf{w}}_i(k-1) \right. \\
 &\quad \left. + \alpha_k \sum_{j=0}^{r-1} \tilde{\mathbf{y}}^* \{ (k-1)l+j \} \tilde{\mathbf{z}}_i \{ (k-1)l+j \} \right] \\
 &= \mathbf{P}(k) \tilde{\mathbf{R}}(k) \tilde{\mathbf{w}}_i(k-1) + \mathbf{k}(k) \\
 &\quad \times \left[\sum_{j=0}^{r-1} \tilde{\mathbf{y}}^* \{ (k-1)l+j \} \tilde{\mathbf{z}}_i \{ (k-1)l+j \} \right. \\
 &\quad \left. - \tilde{\mathbf{B}}^*(k) \tilde{\mathbf{B}}^T(k) \tilde{\mathbf{w}}_i(k-1) \right]
 \end{aligned} \tag{187}$$

where

$$\mathbf{k}(k) := \alpha_k \mathbf{P}(k). \tag{188}$$

This completes the proof.

REFERENCES

- [1] L. Ljung, *System Identification: Theory for the User*. Upper Saddle River, NJ: Prentice-Hall, 1987.
- [2] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications, and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [3] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.
- [4] H. Ogawa, "An operator pseudo-inversion lemma," *SIAM J. Appl. Math.*, vol. 48, no. 6, pp. 1527–1531, Dec. 1988.

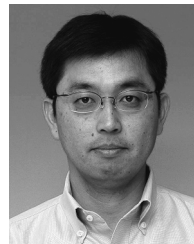
- [5] K. Kohno, Y. Inouye, and M. Kawamoto, "A matrix pseudo-inversion lemma and its application to block-based adaptive blind deconvolution for MIMO systems," in *Proc. ISCAS*, New Orleans, LA, May 27–30, 2007, pp. 3490–3493.
- [6] D.-Z. Feng, H.-Q. Zhang, X.-D. Zhang, and Z. Bao, "An extended recursive least-squares algorithm," *Signal Process.*, vol. 81, no. 5, pp. 1075–1081, May 2001.
- [7] K. Kohno, Y. Inouye, M. Kawamoto, and T. Okamoto, "Adaptive super-exponential algorithms for blind deconvolution of MIMO systems," in *Proc. ISCAS*, Vancouver, BC, Canada, May 23–26, 2004, vol. 5, pp. 680–683.
- [8] K. Kohno, Y. Inouye, and M. Kawamoto, "An adaptive super-exponential deflation algorithm for blind deconvolution of MIMO systems using the matrix pseudo-inversion lemma," in *Proc. ISCAS*, Kobe, Japan, May 23–26, 2005, pp. 5870–5873.
- [9] K. Kohno, Y. Inouye, and M. Kawamoto, "A matrix pseudo-inversion lemma for positive semidefinite Hermitian matrices and its application to adaptive blind deconvolution of MIMO systems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 1, pp. 424–435, Feb. 2008.
- [10] A. C. Tsoi and L. Ma, "A balanced approach to multichannel blind deconvolution," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 2, pp. 599–613, Mar. 2008.
- [11] Y. S. Chen and C.-A. Lin, "Blind-channel identification for MIMO single-carrier zero-padding block-transmission systems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 6, pp. 1571–1579, Jul. 2008.
- [12] Q. Ling and T. Li, "Blind-channel estimation for MIMO systems with structured transmit delay scheme," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 8, pp. 2344–2355, Sep. 2008.
- [13] J.-L. Yu and Y.-C. Lin, "Space-time-coded MIMO ZP-OFDM systems: Semiblind channel estimation and equalization," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 7, pp. 1360–1372, Jul. 2009.
- [14] K. Kohno, M. Kawamoto, and Y. Inouye, "A block-based adaptive super-exponential deflation algorithm for blind deconvolution of MIMO systems using the matrix pseudo-inversion lemma," in *Proc. ISCAS*, Paris, France, May 30–June 2 2010, pp. 801–804.
- [15] P. Lancaster and M. Tismenetsky, *The Theory of Matrix*. San Diego, CA: Academic, 1985.
- [16] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*. New York: Dover, 1979.
- [17] Y. Inouye and R.-W. Liu, "A system-theoretic foundation for blind equalization of an FIR MIMO channel system," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 4, pp. 425–436, Apr. 2002.
- [18] M. Martone, "An adaptive algorithm for antenna array low-rank processing in cellular TDMA base stations," *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 627–643, May 1998.
- [19] G. Hari Kumar and Y. Bresler, "FIR perfect signal reconstruction from multiple convolutions: Minimum deconvolver orders," *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 215–218, Jan. 1998.
- [20] Y. Inouye and K. Tanebe, "Super-exponential algorithms for multichannel blind deconvolution," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 881–888, Mar. 2000.
- [21] O. Shalvi and E. Weinstein, "Super-exponential methods for blind deconvolution," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 504–519, Mar. 1993.
- [22] R. Penrose, "A generalized inverse for matrices," *Math. Proc. Camb. Philos. Soc.*, vol. 51, no. 3, pp. 406–413, Jul. 1955.
- [23] T. Kailath, *Linear Systems*. New York: Prentice-Hall, 1980.



Kiyotaka Kohno (S'04–M'07) received the B.Eng. and M.Eng. degrees in electrical electronics and systems engineering from Nagaoka University of Technology, Nagaoka, Japan, in 1981 and 1983, respectively, and the D.Eng. degree in electronic functions and systems engineering from Shimane University, Matsue, Japan, in 2007.

From April 1983 to March 1990, he was a Research Worker with Matsushita Electric Industrial Company, Ltd. Since April 1990, he has been with Yonago National College of Technology, Hikona, Japan, where he was an Assistant Professor from April 1990 to March 1994 and an Associate Professor from April 1994 to March 2007 and where he is currently a Professor of electronic control engineering. His current research interests are in blind signal processing, blind source separation, communication systems, and image processing.

Dr. Kohno is a member of the European Association for Signal Processing (EURASIP), the Institute of Electronics, Information and Communication Engineers in Japan, and the Information Processing Society of Japan.



Mitsuru Kawamoto (M'97) received the B.Eng., M.Eng., and D.Eng. degrees in control engineering from Kyushu Institute of Technology, Kiakyushu, Japan, in 1992, 1994, and 1997, respectively.

He was a Research Scientist with the Bio-Mimetic Control Research Center, Institute of Physical and Chemical Research (RIKEN), Saitama, Japan, from April 1997 to December 1999. He was an Assistant Professor from January 2000 to March 2002 and an Associate Professor from April 2002 to March 2005 with the Department of Electronic and Control

Systems Engineering, Shimane University, Matsue, Japan. From April 2004 to March 2005, he was with the Department of Electrical Engineering and Electronics, The University of Liverpool, Liverpool, U.K., as a Visiting Researcher. Since March 2005, he has been a Research Scientist with the National Institute of Advanced Industrial Science and Technology. His current research interests include ubiquitous computing, blind signal processing, higher order statistics, sound signal processing, and communication systems.

Dr. Kawamoto is a member of The Society of Instrument and Control Engineers, the Institute of Electronics, Information and Communication Engineers, and the Acoustical Society of Japan.



Yujiro Inouye (M'76–LM'10) received the B.Eng., M.Eng., and D.Eng. degrees in control engineering from Osaka University, Osaka, Japan, in 1966, 1968, and 1971, respectively.

From April 1971 to March 1997, he was with the Faculty of Engineering Science, Osaka University. From April 1997 to March 2008, he was with the Faculty of Science and Engineering, Shimane University, Matsue, Japan, where he is currently an Emeritus Professor of electronic and control systems engineering. His major research interests are in blind signal processing, blind source separation, neural networks, and communications.

Dr. Inouye is a member of the Institute of Electronics, Information and Communication Engineers in Japan and of the Institute of Systems, Control, and Information Engineers in Japan. He is a past Chair of the Technical Committee on Blind Signal Processing of the IEEE Circuits and Systems Society and a past Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I: FUNDAMENTAL THEORY AND APPLICATIONS.