HOKKAIDO UNIVERSITY

Title	Efficient Hybrid DFE Algorithms in Spatial Multiplexing Systems
Author(s)	Jiang, Wenjie; Asai, Yusuke; Aikawa, Satoru et al.
Citation	IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, E92-A(2), 535-546 https://doi.org/10.1587/transfun.E92.A.535
Issue Date	2009-02-01
Doc URL	https://hdl.handle.net/2115/38895
Rights	Copyright©2009 The Institute of Electronics, Information and Communication Engineers
Type	journal article
File Information	E92-A_p535-546.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

PAPER

Efficient Hybrid DFE Algorithms in Spatial Multiplexing Systems

Wenjie JIANG[†a)], **Yusuke ASAI**[†], **Satoru AIKAWA**[††], ***Members***, ***and*** **Yasutaka OGAWA**[†††], ***Fellow***

SUMMARY The wireless systems that establish multiple input multiple output (MIMO) channels through multiple antennas at both ends of the communication link, have been proved to have tremendous potential to linearly lift the capacity of conventional scalar channel. In this paper, we present two efficient decision feedback equalization algorithms that achieve optimal and suboptimal detection order in MIMO spatial multiplexing systems. The new algorithms combine the recursive matrix inversion and ordered QR decomposition approaches, which are developed for nulling cancellation interaface Bell Labs layered space time (BLAST) and back substitution interface BLAST. As a result, new algorithms achieve total reduced complexities in frame based transmission with various payload lengths compared with the earlier methods. In addition, they enable shorter detection delay by carrying out a fast hybrid preprocessing. Moreover, the operation precision insensitivity of order optimization greatly relaxes the word length of matrix inversion, which is the most computational intensive part within the MIMO detection task.

key words: *complexity reduction, MIMO, spatial multiplexing, signal detection*

1. Introduction

It has been shown that the flat fading multiple input multiple output (MIMO) channels, which are constructed by multiple transmit and receive antennas, have enormous capacity in rich scattering environments [1], [2]. Theoretically, the MIMO capacity scales linearly with $\min\{N_{\mathrm{T}}, N_{\mathrm{R}}\}$ (N_{T} and N_{R} are the numbers of transmit and receive antennas.) relative to the conventional scalar channel. The Bell Labs layered space time (BLAST) transceiver architecture [3]–[5] is designed to approach such capacity without additional power and bandwidth consumption. The original diagonal BLAST (D-BLAST) [3] achieves near-MIMO capacity by using a cyclic association of substreams (called layers), thereby averaging the channel condition of each layer. However, D-BLAST is low efficiency for short frame transmissions, and its implementation difficulty has driven the development of a simplified version: vertical BLAST (V-BLAST) [4], [5], where individual layers are assigned to a fixed transmit antenna and are then emitted in parallel. Unfortunately, the detection algorithm for the V-BLAST architecture is still computationally intensive, especially for systems with large numbers of transmit and/or receive antennas.

In V-BLAST architectures or more general spatial multiplexing (SM) systems, the transmitter divides an information stream into multiple substreams and simultaneously emits them through the corresponding antennas. At the receiver side, the superimposed signals are recovered by a certain detection algorithm. The standard BLAST (we will not distinguish between V-BLAST and BLAST hereafter for simplicity) detection algorithm is a spatial ordered decision feedback equalization (DFE) scheme that recovers symbols in the vector signal one by one using N_{T} iterations. At each iteration, the layer with the smallest mean square error is chosen to be detected and this detection order is proved to be optimal in terms of minimizing the total error probability [4], [5]. Furthermore, the detection interfaces possible to implement spatial DFE can be divided into two categories: the nulling cancellation (NC) based approach [4]–[9] and the back substitution (BS) based approach [10]–[16]. Unfortunately, the naive implementation of either NC-BLAST or BS-BLAST requires multiple matrix inversion or QR decompositions (Instead of QR decomposition, a vector recursion or scalar recursion is used in [14] to carry out the back substitution type detection. Unlike the new algorithms in this paper, recursion algorithms in [14] are not interested in calculating the matrix inversion), and imposes on it a complexity burden of $O(N_{\mathrm{T}}^4)$ or $O(N_{\mathrm{T}}^5)$. To reduce the complexity of NC-BLAST, algorithms proposed in [8], [9] use the recursive matrix inversion (RMI) technique, which calculates matrix inversion under a recursive updating fashion, and reduces the complexity to $O(N_{\mathrm{T}}^3)$. The main advantages of the NC-BLAST-RMI are: 1) lower complexity in preprocessing for order and filter optimization, and 2) shorter detection latency due to fast preprocessing. For BS-BLAST, the straightforward performing of $N_{\mathrm{T}}!$ times QR decomposition (QRD) can be avoided by applying a ordered QRD (OQRD) [12], [13] and again a reduced complexity of $O(N_{\mathrm{T}}^3)$ can be achieved. The use of BS-BLAST-OQRD leads to: 1) lower complexity in data detection of frame based transmission, and 2) availability of wide range of algorithms and array architectures for matrix decomposition [17]–[20].

In this paper, we present two hybrid DFE algorithms for signal detection in MIMO spatial multiplexing systems by combining the RMI and OQRD approaches aiming at obtaining the merit from both. The main contributions are: 1) uniform description of NC and BS interface BLAST detection for MIMO spatial multiplexing systems under a spatial

Manuscript received May 16, 2008.
Manuscript revised October 4, 2008.

[†]The authors are with NTT Network Innovation Laboratories, NTT Corporation, Yokosuka-shi, 239-0847 Japan.
[††]The author is with the Graduate School of Engineering, University of Hyogo, Himeji-shi, 671-2280 Japan.
[†††]The author is with the Graduate School of Information Science and Technology, Hokkaido University, Sapporo-shi, 060-0814 Japan.
a) E-mail: jiang.wenjie@lab.ntt.co.jp

DOI: 10.1587/transfun.E92.A.535

Copyright © 2009 The Institute of Electronics, Information and Communication Engineers

DFE framework, 2) propose two hybrid DFE algorithms by using RMI based double recursion and simple sorting to accelerate the order search, and an ordered Cholesky based QRD to generate DFE filters used in BS interface data detection, 3) detailed performance evaluation of the new algorithm in terms of the error performance, complexity and impact on the operation precision. Finally, the advantages of new hybrid DFE algorithms are pointed out.

The remainder of this paper is organized as follows. Section 2 introduces the complex baseband model of SM systems. In Sect. 3, we review the BLAST algorithm with respect to NC and BS detection interface. Section 4 describes the new hybrid algorithms achieving optimal and suboptimal ordering. We analyze the complexity and performance in Sects. 5 and 6. The paper is concluded in Sect. 7.

Bold uppercase letters denote the matrices. $\mathbf{a}_{(i)}$ and $\mathbf{a}_j$ denote the ith row and jth column vector in matrix $\mathbf{A}$. The (i, j)th entry of $\mathbf{A}$ is expressed as $a_{i,j}$. or $[\mathbf{A}]_{i,j}$. $\mathbf{A}^T$ and $\mathbf{A}^H$ represent the transpose and Hermitian transpose of $\mathbf{A}$, while $\mathbf{A}^{-1}$ denotes the inverse of $\mathbf{A}$. $\mathbf{I}_N$ denotes an identity matrix with dimension N.

2. Spatial Multiplexing Systems

We consider a complex baseband model of SM systems with N_T transmit and $N_\mathrm{R} \geq N_\mathrm{T}$ receive antennas, as shown in Fig. 1. An information bit stream is divided into N_T substreams (called layers) at the transmitter and each substream is mapped into M-QAM/PSK symbols. After being partitioned in a frame of length N_D, all N_T substreams are simultaneously emitted through the corresponding antennas in the same frequency band. On the receiver side, the faded and spatially superimposed substreams are captured by N_R antennas and recovered by the SM detector. Consequently, the original information is regenerated by mixing the demapped substreams. The transmit signal block is represented by an $N_\mathrm{T} \times N_\mathrm{D}$ matrix $\mathbf{S} = [\mathbf{s}(1)\ \mathbf{s}(2) \cdots \mathbf{s}(N_\mathrm{D})]$. The row and column index of $\mathbf{S}$ indicate space (i.e., layer) and time, respectively. The corresponding receive signal block and additive noise are expressed as $N_\mathrm{R} \times N_\mathrm{D}$ matrices $\mathbf{X} = [\mathbf{x}(1)\ \mathbf{x}(2) \cdots \mathbf{x}(N_\mathrm{D})]$ and $\mathbf{N} = [\mathbf{n}(1)\ \mathbf{n}(2) \cdots \mathbf{n}(N_\mathrm{D})]$. The entries of $\mathbf{N}$ are independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and variance σ_n^2. Thus, the noise that disturbs the receive signal is spatially and temporally independent. The relationship between the $N_\mathrm{T} \times 1$ transmit signal $\mathbf{s}(t)$ and the $N_\mathrm{R} \times 1$ receive signal $\mathbf{x}(t)$ is given by:

$$\mathbf{x}(t) = \mathbf{H}\mathbf{s}(t) + \mathbf{n}(t) \quad (1 \leq t \leq N_\mathrm{D}), \tag{1}$$

where $\mathbf{n}(t)$ denotes the $N_\mathrm{R} \times 1$ noise vector. The transmit signal $\mathbf{s}(t)$ follows $E\{\mathbf{s}^H(t)\mathbf{s}(t)\} = N_\mathrm{T}\sigma_s^2 =$ constant, where σ_s^2 denotes the signal power. $\mathbf{H}$ stands for an $N_\mathrm{R} \times N_\mathrm{T}$ channel matrix that contains i.i.d. complex Gaussian fading gains with zero mean and unit variance. In addition, we assume that $\mathbf{H}$ remains constant over one signal frame and changes randomly from frame to frame. For simplicity, we also suppose that the receiver perfectly knows the channel state information and noise variance.

3. Standard BLAST Algorithm

Here, we review the standard BLAST algorithm with respect to the NC and BS detection interface. First, we briefly describe the NC-BLAST algorithm. Then, an equivalent alternative detection interface BS-BLAST is described. For brevity of notation, we drop the time index t from here on.

3.1 Nulling and Cancellation (NC) Interface BLAST (NC-BLAST)

The most straightforward approach to combat the multistream interference in the MIMO SM scenario is the utilizing of the linear equalization (LE) detector. The weight matrix of LE is given by

$$\mathbf{G}_\mathrm{LE} \equiv \mathbf{C}^{-1}\mathbf{H}^H, \tag{2}$$

$$\equiv \mathbf{M}\,\mathbf{H}^H, \tag{3}$$

with $\mathbf{C}$ and $\mathbf{M}$ defined as

$$\mathbf{C} \equiv \mathbf{M}^{-1} \equiv \mathbf{H}^H\mathbf{H} + \gamma\mathbf{I}_{N_\mathrm{T}}, \tag{4}$$

where γ is a criterion switch factor, which denotes ZF for $\gamma = 0$ and MMSE for $\gamma = \sigma_n^2/\sigma_s^2 = \alpha^2$. In addition to interference suppression, the MMSE criterion also balances the noise amplification thereby achieving better error performance than ZF. The error covariance matrix of the linear detector can be written as:

$$\mathbf{\Phi} = E\{(\mathbf{s} - \mathbf{G}\mathbf{x})(\mathbf{s} - \mathbf{G}\mathbf{x})^H\} = \sigma_n^2\,\mathbf{C}^{-1} = \sigma_n^2\mathbf{M}. \tag{5}$$

Clearly, the minimum square error (MSE) of each layer, i.e., layer-wise MSE, corresponds to the diagonal entries of $\mathbf{M}$.

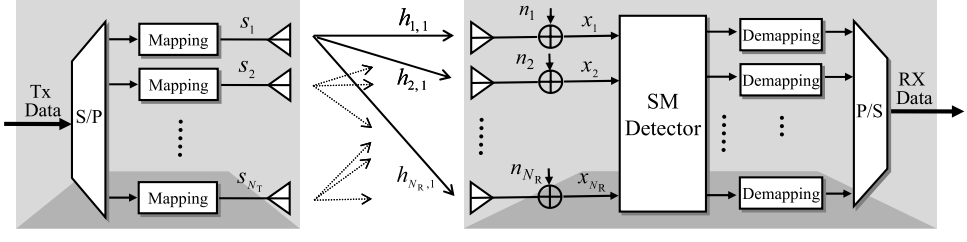


Fig. 1 Baseband model of spatial multiplexing systems.

Table 1 NC-BLAST algorithm.

INPUT: $\mathbf{x}$, $\mathbf{H}$, γ		
OUTPUT: $\hat{\mathbf{s}}$		
Initialization	(1)	$\mathbf{H}^{[N_T]} = \mathbf{H}$, $\mathbf{X}^{[N_T]} = \mathbf{X}$, $\mathbf{G}_{\mathrm{DFE}} = \mathbf{0}_{N_T,N_R}$, $O_1 = (1,\ldots,N_T)$, $O = \varnothing$
Preprocessing	(2)	**for** $\kappa = 1,\ldots,N_T \quad (k = N_T,\ldots,1)$
	(3)	$\mathbf{M}^{[k]} = (\mathbf{H}^{[k],H}\mathbf{H}^{[k]} + \gamma\mathbf{I}_k)^{-1}$
	(4)	$i_\kappa = \arg\min \{[\mathbf{M}^{[k]}]_{p,p} \mid p = 1,\ldots,k\}$
	(5)	$o_\kappa = O_\kappa(i_\kappa)$, $\quad O_{\kappa+1} = O_\kappa \setminus \{o_\kappa\}$
	(6)	$\mathbf{g}_{\mathrm{DFE},(o_\kappa)} = \mathbf{m}^{[k]}_{(i_\kappa)}\mathbf{H}^{[k],H}$
	(7)	$\mathbf{H}^{[k-1]} = \mathbf{H}^{[k]}_{\overline{i_\kappa}}$
	(8)	**end**
Data Detection	(9)	**for** $\kappa = 1,\ldots,N_T \quad (k = N_T,\ldots,1)$
	(10)	$\tilde{\mathbf{s}}_{(o_\kappa)} = \mathbf{g}_{\mathrm{DFE},(o_\kappa)} \cdot \mathbf{X}^{[k]}$
	(11)	$\hat{\mathbf{s}}_{(o_\kappa)} = Q[\tilde{\mathbf{s}}_{(o_\kappa)}]$
	(12)	$\mathbf{X}^{[k-1]} = \mathbf{X}^{[k]} - \mathbf{h}_{o_\kappa}\hat{\mathbf{s}}_{(o_\kappa)}$
	(13)	**end**

The ZF and MMSE BLAST algorithm [4]–[7] is an N_T-step nonlinear approach that successively utilizes the nulling vectors (i.e., rows in LE detector $\mathbf{G}_{\mathrm{LE}}$) to recover the signals of $\mathbf{s}$ in a sequential manner. In addition, the detection order is optimized by always first detecting the layer with minimum MSE, and this ordering is proved to be optimal in terms of minimization of the total error probability. We denote NC interface BLAST as NC-BLAST and define the order sequence $O = (o_1, o_2, \ldots, o_{N_T})$ as the permutation of the integers $(1, 2, \ldots, N_T)$. For instance, $o_2 = 3$ means the 3rd layer is detected second. Let $O_1 = \{1, 2, \ldots N_T\}$. After step $\kappa(\kappa = 1, \ldots, N_T)$, the set contains indices of the undetected layers becomes $O_{\kappa+1} = O_\kappa \backslash \{o_\kappa\}$. The corresponding reduced channel matrix $\mathbf{H}^{[k]}(k = N_T - \kappa + 1)$ is created by removing columns $(\mathbf{h}_{o_1}, \mathbf{h}_{o_2}, \ldots, \mathbf{h}_{o_\kappa})$ from $\mathbf{H}$. Accordingly, we have the following input output relationship associated with the system with reduced transmission dimension.

$$\mathbf{x}^{[k]} = \mathbf{H}^{[k]}\mathbf{s}^{[k]} + \mathbf{n}, \tag{6}$$

where $\mathbf{s}^{[k]}$ and $\mathbf{x}^{[k]}$ are the resultant reduced transmit and receive signals. The corresponding error covariance matrix reads

$$\mathbf{\Phi}^{[k]} = (\mathbf{H}^{[k],H}\mathbf{H}^{[k]} + \gamma\mathbf{I}_k)^{-1} \equiv \sigma_n^2\mathbf{C}^{[k],-1} \equiv \sigma_n^2\mathbf{M}^{[k]}. \tag{7}$$

NC-BLAST detects the signals in $\mathbf{s}$ from s_{o_1} to $s_{o_{N_T}}$. At step κ, we first compute $\mathbf{M}^{[k]}$ by inverting $\mathbf{C}^{[k]}$:

$$\mathbf{M}^{[k]} = (\mathbf{H}^{[k],H}\mathbf{H}^{[k]} + \gamma\mathbf{I}_k)^{-1}. \tag{8}$$

Accordingly, we determine the current best order as follows

$$i_\kappa = \arg\min_o\{[\mathbf{M}^{[k]}]_{o,o} | 1 \le o \le k\}, \tag{9}$$

$$o_\kappa = O_\kappa(i_\kappa), \tag{10}$$

where o_κ denotes the i_κth entry in O_κ. Namely, column i_κ in $\mathbf{H}^{[k]}$ corresponds to column o_κ in $\mathbf{H}$ and is related to the current best estimate. Then, we calculate the nulling vector according to the current best order as

$$\mathbf{g}_{\mathrm{DFE},(o_\kappa)} \equiv \mathbf{m}^{[k]}_{(i_\kappa)} \cdot \mathbf{H}^{[k],H}, \tag{11}$$

where $\mathbf{G}_{\mathrm{DFE}}$ is the NC interface DFE filter matrix. Using, $\mathbf{g}_{\mathrm{DFE},(o_\kappa)}$, we null out the interference caused by the remaining undetected symbols as follows.

$$\tilde{s}_{o_\kappa} = \mathbf{g}_{\mathrm{DFE},(o_\kappa)} \cdot \mathbf{x}^{[k]}. \tag{12}$$

The current estimate is now obtained by decision device $Q[\cdot]$ as in (13), and the interference originating from this estimate is removed from the received signal in (14).

$$\hat{s}_{o_\kappa} = Q[\tilde{s}_{o_\kappa}], \tag{13}$$

$$\mathbf{x}^{[k-1]} = \mathbf{x}^{[k]} - \mathbf{h}_{o_\kappa}\hat{s}_{o_\kappa}, \tag{14}$$

Before going to the next step $\kappa + 1$, we reduce the channel matrix $\mathbf{H}^{[k]}$:

$$\mathbf{H}^{[k-1]} \equiv \mathbf{H}^{[k]}_{\overline{i_\kappa}}, \tag{15}$$

where $\mathbf{H}^{[k]}_{\overline{i_\kappa}}$ denotes removing column i_κ from $\mathbf{H}^{[k]}$ to obtain $\mathbf{H}^{[k-1]}$. After N_T iterations, all the layers are detected according to the optimal order with layer-wise MSE

$$\mathrm{MSE}_{o_\kappa} = \sigma_n^2[\mathbf{M}^{[k]}]_{i_\kappa,i_\kappa} \propto [\mathbf{M}^{[k]}]_{i_\kappa,i_\kappa}. \tag{16}$$

Table 1 summarizes the entire algorithm with respect to the frame based transmission. The adaptation of detection order and DFE filter (called preprocessing) need to be carried out once per signal frame (lines (2)–(8)), while the signal detection should be performed from the head to the tail of the data part (lines (9)–(13)). We emphasize that the most computational intensive part in NC-BLAST is the multiple matrix inversion, i.e., the $\mathbf{M}^{[k]}$ calculation in the preprocessing. To reduce the complexity of NC-BLAST, the recursive matrix inversion (RMI) algorithms computing matrix inverse under

Table 2 BS-BLAST algorithm.

INPUT: $\mathbf{H}$		
OUTPUT: $\hat{\mathbf{s}}$		
Initialization	(1)	$\mathbf{F} = \mathbf{0}_{N_\mathrm{T},N_\mathrm{R}}$, $\mathbf{B} = \mathbf{0}_{N_\mathrm{T},N_\mathrm{T}}$, $O = \varnothing$
Preprocessing	(2)	find $\mathbf{HP}(O) = \mathbf{QR}$ by performing $\frac{1}{2}N_\mathrm{T}^2 + \frac{1}{2}N_\mathrm{T} - 1$ QRD
	(3)	$\mathbf{F} = \mathbf{D}_\mathbf{R}^{-1}\mathbf{Q}^H$, $\mathbf{B} = \mathbf{D}_\mathbf{R}^{-1}\mathbf{R} - \mathbf{I}_{N_\mathrm{T}}$
Data Detection	(4)	$\mathbf{Y} = \mathbf{FX}$
	(5)	**for** $\kappa = 1, \ldots, N_\mathrm{T}$ $(k = N_\mathrm{T}, \ldots, 1)$
	(6)	$\hat{\mathbf{d}}_{(k)} = \sum_{j=k+1}^{N_\mathrm{T}} b_{k,j} \cdot \hat{\mathbf{s}}_{(o_{\kappa-1})}$
	(7)	$\tilde{\mathbf{s}}_{(o_\kappa)} = \mathbf{y}_{(k)} - \hat{\mathbf{d}}_{(k)}$
	(8)	$\hat{\mathbf{s}}_{(o_\kappa)} = \mathrm{Q}[\tilde{\mathbf{s}}_{(o_\kappa)}]$
	(9)	**end**
	(10)	$\hat{\mathbf{S}} = \mathbf{P}[\hat{\mathbf{s}}_{(o_{N_\mathrm{T}})}^T, \hat{\mathbf{s}}_{(o_{N_\mathrm{T}-1})}^T, \ldots, \hat{\mathbf{s}}_{(o_1)}^T]^T$

a recursive updating fashion was proposed in [8], [9], which reduce the complexity from $O(N_\mathrm{T}^4)$ to $O(N_\mathrm{T}^3)$. The main advantages of NC-BLAST-RMI are: 1) lower complexity in preprocessing for order and filter adaptation, and 2) shorter detection latency due to fast preprocessing.

3.2 Back Substitution (BS) Interface BLAST (BS-BLAST)

An alternative but equivalent detection interface of NC-BLAST is the back substitution (BS) [21] interface BLAST (BS-BLAST) [11] (also called QRD interface BLAST), which contains a feedforward filter $\mathbf{F}$, a feedback filter $\mathbf{B}$, and a permutation filter $\mathbf{P}$ as the BS interface DFE filters. The permutation matrix $\mathbf{P}$ facilitates the column exchange in $\mathbf{H}$ according to the optimized order sequence O.

$$\mathbf{P} = [\mathbf{e}_{o_{N_\mathrm{T}}}\ \mathbf{e}_{o_{N_\mathrm{T}-1}} \cdots \mathbf{e}_{o_1}], \tag{17}$$

where $\mathbf{e}_k$ denotes the kth column in $\mathbf{I}_{N_\mathrm{T}}$. We factorize the permutated channel matrix using QR decomposition (QRD) [21]

$$\mathbf{HP} = \mathbf{QR}, \tag{18}$$

where $\mathbf{Q}$ is an $N_\mathrm{R} \times N_\mathrm{T}$ semi-unitary matrix, and $\mathbf{R}$ is an $N_\mathrm{T} \times N_\mathrm{T}$ upper triangular matrix. As a result of permutation, the signals in $\mathbf{s}$ can be detected in the order $(o_1, o_2, \ldots, o_{N_\mathrm{T}})$. The corresponding ZF feedforward filter $\mathbf{F}$ and feedback filter $\mathbf{B}$ are thus expressed as:

$$\mathbf{F} = \mathbf{D}_\mathbf{R}^{-1}\mathbf{Q}^H, \tag{19}$$

$$\mathbf{B} = \mathbf{D}_\mathbf{R}^{-1}\mathbf{R} - \mathbf{I}_{N_\mathrm{T}}, \tag{20}$$

where $\mathbf{D_R}$ is a diagonal matrix with entries $[\mathbf{D_R}]_{k,k} = r_{k,k} (1 \le k \le N_\mathrm{T})$. We obtain the modified received signal $\mathbf{x}$ by linear filtering as follows

$$\mathbf{y} = \mathbf{Fx} = \mathbf{P}^T\mathbf{s} + \mathbf{BP}^T\mathbf{s} + \tilde{\mathbf{n}}, \tag{21}$$

where $\tilde{\mathbf{n}} = \mathbf{D}_\mathbf{R}^{-1}\mathbf{Q}^H\mathbf{n}$. Since the columns in $\mathbf{Q}$ are orthonormal, the components in $\tilde{\mathbf{n}}$ are still spatially uncorrelated, thereby guaranteeing the equivalence between (21) and (1). The component-wise form of (21) can be written as

$$y_k = s_{o_\kappa} + d_k + \tilde{n}_k \quad (1 \le k \le N_\mathrm{T}), \tag{22}$$

Here, the first term stands for the desired signal at the layer o_κ, $d_k = \sum_{j=k+1}^{N_\mathrm{T}} b_{\mathrm{ZF},k,j} \cdot s_{o_{\kappa-1}}$ denotes the interference caused by the already detected signals, and $\tilde{n}_k$ expresses noise. The feedforward filter $\mathbf{F}$ transforms the system into an upper triangular form, such that at each layer the desired signal is free from the interference originating from the upper undetected signals. By exploiting this causality property, we can successively recover the signals through a back substitution based decision feedback. At step $\kappa (\kappa = N_\mathrm{T} - k + 1 = 1, \ldots, N_\mathrm{T})$, we have

$$\tilde{s}_{o_\kappa} = y_k - \hat{d}_k = y_k - \sum_{j=k+1}^{N_\mathrm{T}} b_{\mathrm{ZF},k,j} \cdot \hat{s}_{o_{\kappa-1}}, \tag{23}$$

$$\hat{s}_{o_\kappa} = \mathrm{Q}[\tilde{s}_{o_\kappa}], \tag{24}$$

where $\hat{d}_k$ stands for the estimate of d_k. After passing the decision feedback loop, the N_T estimates are rearranged into their original spatial positions to finish the detection task.

$$\hat{\mathbf{s}} = \mathbf{P}[\hat{s}_{o_{N_\mathrm{T}}}, \hat{s}_{o_{N_\mathrm{T}-1}}, \ldots, \hat{s}_{o_1}]^T. \tag{25}$$

From (22), it is easy to find that the post detection signal to noise ratio (SNR) of layer o_κ is given by:

$$\mathrm{SNR}_{o_\kappa} = \frac{1}{\mathrm{MSE}_{o_\kappa}} = \frac{|r_{k,k}|^2\sigma_s^2}{\sigma_n^2} \propto |r_{k,k}|^2, \tag{26}$$

which implies that maximization of $|r_{k,k}|$ for $k = N_\mathrm{T}$ down to 1 in BS-BLAST, is equivalent to the successive minimization of the diagonal entries of $[\mathbf{M}^{[k]}]$ in NC-BLAST (see (16)).

The BS-BLAST algorithm with respect to the ZF criterion is listed in Table 2. In the preprocessing, The straightforward approach to achieve the ordering of BS-BLAST is to perform $\frac{1}{2}N_\mathrm{T}^2 + \frac{1}{2}N_\mathrm{T} - 1$ QR decompositions [10] (line (2)). Then, we establish the filters $\mathbf{F}$ and $\mathbf{B}$ on line (3). The data is detected by first constructing the spatial causality through

linear filtering (line (4)). Thus, through N_{T} steps decision feedback, the estimates of N_{T} layers are obtained (lines (5)–(9)), and finally the permutation is performed on line (10). The reduced complexity algorithms for BS-BLAST can be found in [12], [13]. The multiple QRDs is avoided by performing only one ordered QRD (OQRD) in conjunction with a sequence of unitary transformations, thereby leading to a lower complexity of $O(N_{\mathrm{T}}^3)$. By using BS-BLAST-OQRD, we obtain: 1) lower complexity in data detection of frame based transmission [see Section 5], and 2) availability of wide range of algorithms and array architectures for matrix decomposition.

In order to adapt to the MMSE criterion, we need to introduce an extended channel matrix [12] as follows.

$$\underline{\mathbf{H}} = \begin{bmatrix} \mathbf{H} \\ \alpha \mathbf{P}^T \end{bmatrix}. \tag{27}$$

To reflect the detection sequence, the column permutated $\underline{\mathbf{H}}$ is decomposed,

$$\underline{\mathbf{H}}\mathbf{P} = \begin{bmatrix} \mathbf{H}\mathbf{P} \\ \alpha \mathbf{I}_{N_{\mathrm{T}}} \end{bmatrix} = \underline{\mathbf{Q}}\underline{\mathbf{R}} = \begin{bmatrix} \mathbf{Q}_1 \\ \mathbf{Q}_2 \end{bmatrix} \underline{\mathbf{R}}, \tag{28}$$

where the $(N_{\mathrm{R}}+N_{\mathrm{T}})\times N_{\mathrm{T}}$ matrix $\underline{\mathbf{Q}}$ has orthonormal columns, and $\underline{\mathbf{R}}$ is an $N_{\mathrm{T}} \times N_{\mathrm{T}}$ upper triangular matrix. $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are submatrices of $\underline{\mathbf{Q}}$ with N_{R} and N_{T} rows. The corresponding MMSE QRD-DFE filters are given by

$$\underline{\mathbf{F}} = \mathbf{D}_{\underline{\mathbf{R}}}^{-1}\mathbf{Q}_1^H, \tag{29}$$

$$\underline{\mathbf{B}} = \mathbf{D}_{\underline{\mathbf{R}}}^{-1}\underline{\mathbf{R}} - \mathbf{I}_{N_{\mathrm{T}}}, \tag{30}$$

where $\mathbf{D}_{\underline{\mathbf{R}}} = \mathrm{diag}(\underline{r}_{1,1}, \underline{r}_{2,2}, \ldots, \underline{r}_{N_{\mathrm{T}},N_{\mathrm{T}}})$. The remaining processing is the same as in the ZF counterpart. Compared with ZF, additional memory and computation is required to meet the MMSE criterion because of the extension of the channel matrix.

4. New Hybrid DFE Algorithms with Optimal and Suboptimal Ordering

Here, we describe two new hybrid DFE algorithms, which combine the conventional RMI and OQRD approaches and obtain the merit from both. Since once the detection order and DFE filter is determined the data detection is performed in a regular way as described in Sect. 3, we focus on the preprocessing, i.e., order and filter optimization, from here on. Hybrid algorithm I (Hybrid-I), which achieves the optimal ordering, uses a double recursion matrix inversion to accelerate the order search and apply a pre-ordered Cholesky based QRD to determine BS interface filters for subsequent data detection. In contrast, hybrid algorithm II (Hybrid-II) replaces the recursive order search by a simple sorting function. As a result, only single recursion is performed and a further complexity reduction is achieved. Although Hybrid-II uses a suboptimal ordering criterion, the performance degradation compared with Hybrid-I is negligible.

4.1 Hybrid Algorithm I: Optimal Ordering

The task of preprocessing is to adapt the order and DFE filters that are used in the data detection. A straightforward approach is the joint optimization of these two as performed in conventional RMI and OQRD. However, here a separate optimization is adopted so that we can combine the order search of NC-BLAST-RMI and filter adaptation of BS-BLAST-OQRD to get the following advantages. First, in data detection BS-BLAST-OQRD is more efficient than NC-BLAST-RMI by using the BS interface. This feature is especially obvious in systems with long data payload, where the order and DFE filters needs to be adapted only once per signal frame, while the data detection should be performed from the head to the tail within the data part. Hence, we use OQRD to form the DFE filters, which enable BS interface data detection and then lead to a total reduced complexity. Second, on the other hand, data detection can only start after the preprocessing and preprocessing is much more computationally intensive than data detection. Therefore, the detection delay mainly depends on the complexity of the preprocessing (see Fig.2). In other words, a fast preprocessing leads to short detection latency. The disadvantage of OQRD is its higher complexity in preprocessing compared with RMI. Since order search is the dominant part in preprocessing, we apply RMI based order search to shorten the detection latency. Third, since the layer ordering improves the total error probability by always first detecting the layer with "good" condition, for channel realizations where the layer-wise conditions (e.g. MSE) are very close to each other, ordering provides very minor or even no impact on error performance. In that case, wrong ordering caused by the limited precision of a digital signal processor will not seriously degrade the error performance. Namely, unlike filter adaptation, the order optimization does not need high precision. Thus, the word length of ordering operations (i.e., RMI), which strongly impacts the design area, hardware complexity and power consumption, can be greatly relaxed.

To achieve the above goals, we first determine the detection order by using the RMI based double recursion algorithm [9]. Redefine $\mathbf{C}$ in a recursive form as follows.

$$\begin{aligned} \mathbf{C}^{[k]} &\equiv \mathbf{H}^{[k],H}\mathbf{H}^{[k]} + \gamma \mathbf{I}_k, \\ &\equiv \begin{bmatrix} \mathbf{H}^{[k-1],H}\mathbf{H}^{[k-1]} + \gamma \mathbf{I}_{k-1} & \mathbf{H}^{[k-1],H}\mathbf{h}_k \\ \mathbf{h}_k^H \mathbf{H}^{[k-1]} & \|\mathbf{h}_k\|^2 + \gamma \end{bmatrix}, \quad (31) \\ &\equiv \begin{bmatrix} \mathbf{C}^{[k-1]} & \mathbf{v}^{[k]} \\ \mathbf{v}^{[k],H} & \beta_k \end{bmatrix} \quad (k = 1, \ldots, N_{\mathrm{T}}), \quad (32) \end{aligned}$$

with

$$\beta_k \equiv \|\mathbf{h}_k\|^2 + \gamma, \tag{33}$$

$$\mathbf{v}^{[k]} \equiv \mathbf{H}^{[k-1],H}\mathbf{h}_k. \tag{34}$$

It must be noted that no ordering is performed here, so that we have $\mathbf{H}^{[k]} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_k] = [\mathbf{H}^{[k-1]}\mathbf{h}_k]$, i.e., $i_\kappa = N_{\mathrm{T}} - \kappa + 1$ in (15). Hence, $\mathbf{C} \equiv \mathbf{C}^{[N_{\mathrm{T}}]}$ is represented

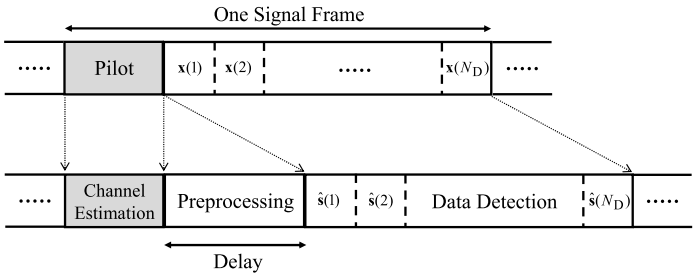


Fig. 2 Signal processing during one signal frame of a frame based MIMO SM systems.

by successively adding the effect of new columns in the order of $(\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{N_\mathrm{T}})$. Next, we write $\mathbf{M}^{[k]} \equiv \mathbf{C}^{[k],-1}$ as

$$\mathbf{M}^{[k]} = \begin{bmatrix} \mathbf{T}^{[k-1],-1} & \mathbf{w}^{[k]} \\ \mathbf{w}^{[k],H} & \xi_k \end{bmatrix}, \tag{35}$$

with

$$\mathbf{T}^{[k-1]} \equiv \mathbf{C}^{[k-1]} - \frac{1}{\beta_k}\mathbf{v}^{[k]}\mathbf{v}^{[k],H}, \tag{36}$$

$$\mathbf{w}^{[k]} \equiv -\frac{1}{\beta_k}\mathbf{T}^{[k-1],-1}\mathbf{v}^{[k]} \tag{37}$$

$$\xi_k \equiv \frac{1}{\beta_k}(1 - \mathbf{v}^{[k],H}\mathbf{w}^{[k]}), \tag{38}$$

where $\mathbf{T}^{[k-1]}$ denotes the Schur complement of β_k in $\mathbf{C}^{[k]}$. Using the matrix inverse lemma on (36), we have

$$\mathbf{T}^{[k-1],-1} = \mathbf{M}^{[k-1]} + \frac{\mathbf{M}^{[k-1]}\mathbf{v}^{[k]}\mathbf{v}^{[k],H}\mathbf{M}^{[k-1]}}{\beta_k - \mathbf{v}^{[k],H}\mathbf{M}^{[k-1]}\mathbf{v}^{[k]}}, \tag{39}$$

After further mathematical development, we get

$$\mathbf{w}^{[k]} = -\xi_k \mathbf{w}'^{[k]}, \tag{40}$$

$$\xi_k = \frac{1}{\beta_k - \mathbf{v}^{[k],H}\mathbf{w}'^{[k]}}, \tag{41}$$

where $\mathbf{w}'^{[k]} \equiv \mathbf{M}^{[k-1]}\mathbf{v}^{[k]}$. Hence, we can rewrite (39) as

$$\mathbf{T}^{[k-1],-1} = \mathbf{M}^{[k-1]} - \mathbf{w}^{[k]}\mathbf{w}'^{[k],H}. \tag{42}$$

Consequently, we attain an N_T-step inflated recursion that calculates $\mathbf{M}^{[k]}$ for $k = 1$ to N_T. At each step k, corresponding to $\mathbf{C}^{[k]}$, we add the effect of the new column $\mathbf{h}_k$ by using (40)–(42): thus, ξ_k, $\mathbf{w}^{[k]}$ and $\mathbf{T}^{[k-1],-1}$ of $\mathbf{M}^{[k]}$ (see (35)) are determined in turn.

Once $\mathbf{M} = \mathbf{M}^{[N_\mathrm{T}]}$ is obtained, we need to deflate it to determine the order sequence. From (42), we know that the task of deflation is to remove the effect of unnecessary columns by modifying (42) as

$$\mathbf{M}^{[k-1]} = \mathbf{T}^{[k-1],-1} + \mathbf{w}^{[k]}\mathbf{w}'^{[k],H}. \tag{43}$$

However, instead of reversing the order for inflation (i.e., $(\mathbf{h}_{N_\mathrm{T}}, \mathbf{h}_{N_\mathrm{T}-1}, \ldots, \mathbf{h}_1)$), the deflation must be performed according to the optimized detection order as $(\mathbf{h}_{o_1}, \mathbf{h}_{o_2}, \ldots, \mathbf{h}_{o_{N_\mathrm{T}}})$. Hence, with the assignment of $\breve{\mathbf{M}}^{[N_\mathrm{T}]} = \mathbf{M}^{[N_\mathrm{T}]}$, at each deflation step, we exchange the i_κth column and row (which are associated with the current best column i_κ in $\mathbf{H}^{[k]}$ and recall that column i_κ in $\mathbf{H}^{[k]}$ corresponds to column o_κ in $\mathbf{H}$.) with the last column and row in $\breve{\mathbf{M}}^{[N_\mathrm{T}]}$ to achieve the ordering.

$$\widehat{\mathbf{M}}^{[k]} \equiv \mathbf{P}_{i_\kappa}^{[k]}\breve{\mathbf{M}}^{[k]}\mathbf{P}_{i_\kappa}^{[k]} \equiv \begin{bmatrix} \widehat{\mathbf{T}}^{[k-1],-1} & \widehat{\mathbf{w}}^{[k]} \\ \widehat{\mathbf{w}}^{[k],H} & \widehat{\xi}_k \end{bmatrix} \quad (k = N_\mathrm{T}, \ldots, 1), \tag{44}$$

where $\mathbf{P}_{i_\kappa}^{[k]}$ is a $k \times k$ permutation matrix given by exchanging column i_κ and the last column in $\mathbf{I}_k$. Then, applying (44) to (40) and (42), we obtain

$$\widehat{\mathbf{w}'}^{[k]} \equiv -\frac{1}{\widehat{\xi}_k}\widehat{\mathbf{w}}^{[k]}, \tag{45}$$

$$\breve{\mathbf{M}}^{[k-1]} \equiv \widehat{\mathbf{T}}^{[k-1],-1} + \widehat{\mathbf{w}}^{[k]}\widehat{\mathbf{w}'}^{[k],H}, \tag{46}$$

which means that we can deflate $\breve{\mathbf{M}}^{[k]}$ for $k = N_\mathrm{T}$ down to 1 by using the already determined quantities in conjunction with the recursive permutation (see (44)).

On the other hand, since the size reduction and permutation, which strongly depend on each channel realization, are irregular and are intractable from an implementation viewpoint, we further refine (46) as follows:

$$\underline{\breve{\mathbf{M}}}^{[k-1]} \equiv \widehat{\mathbf{M}}^{[k]} + \widehat{\mathbf{m}}_k^{[k]}\left(\frac{-1}{\widehat{m}_{k,k}^{[k]}}\widehat{\mathbf{m}}_k^{[k],H}\right), \tag{47}$$

$$= \begin{bmatrix} \breve{\mathbf{M}}^{[k-1]} & \mathbf{0}_{k-1,1} \\ \mathbf{0}_{1,k-1} & 0 \end{bmatrix}. \tag{48}$$

Obviously, (48) is equivalent to (46), but it avoids the size reduction and the corresponding memory movement. Next, we define an alternative form by permuting $\underline{\breve{\mathbf{M}}}^{[k-1]}$ as

$$\mathbf{P}_{i_\kappa}^{[k]}\underline{\breve{\mathbf{M}}}^{[k-1]}\mathbf{P}_{i_\kappa}^{[k]} = \breve{\mathbf{M}}^{[k]} + \breve{\mathbf{m}}_{i_\kappa}^{[k]}\left(\frac{-1}{\breve{m}_{i_\kappa,i_\kappa}^{[k]}}\breve{\mathbf{m}}_{i_\kappa}^{[k],H}\right), \tag{49}$$

$$= \begin{bmatrix} \times & \mathbf{0}_{i_\kappa-1,1} & \times \\ \mathbf{0}_{1,i_\kappa-1} & 0 & \mathbf{0}_{1,k-i_\kappa} \\ \times & \mathbf{0}_{k-i_\kappa,1} & \times \end{bmatrix}, \tag{50}$$

where $\times$ denotes the non-zero entries. It is noteworthy that (50) and (48) only differ in the positions of the kth and i_κth columns and rows. This fact guarantees the equivalence of

Table 3 Preprocessing of Hybrid-I.

INPUT: $\mathbf{H}$, γ		
OUTPUT: $\mathbf{Q}$, $\mathbf{R}$ = upper triangular portion of $\mathbf{R}$, O		
Initialization	(1)	$\mathbf{Q}=\mathbf{H}$, $\mathbf{R}=\mathbf{H}^H\mathbf{H}+\gamma\mathbf{I}$, $\mathbf{M}=\mathbf{0}_{N_T,N_T}$, $O=\varnothing$
Inflation	(2)	**for** $i=1,\ldots,N_T$
	(3)	$\mathbf{a}=\mathbf{M}(1:i-1,1:i-1)\mathbf{R}(1:i-1,i)$
	(4)	$\mathbf{M}(i,i)=1/(\mathbf{R}(i,i)-\mathbf{R}^H(1:i-1,i)\mathbf{a})$
	(5)	$\mathbf{M}(1:i-1,i)=-\mathbf{M}(i,i)\mathbf{a}$, $\mathbf{M}(i,1:i-1)=\mathbf{M}^H(1:i-1,i)$
	(6)	$\mathbf{M}(1:i-1,1:i-1):=\mathbf{M}(1:i-1,1:i-1)-\mathbf{M}(1:i-1,i)\mathbf{a}^H$
	(7)	**end**
Deflation	(8)	**for** $i=1,\ldots,N_T$
	(9)	$o_i=\arg\min\{\mathbf{M}(o,o)\mid o\in O_1\wedge o\neq o_1,\ldots,o_{i-1}\}$
	(10)	$\mathrm{MSE}=\mathbf{M}(o_i,o_i)$, $\mathbf{b}=[\mathbf{M}^T(1:o_i-1,o_i)\ \ 0\ \ \mathbf{M}^T(o_i+1:N_T,o_i)]^T$
	(11)	$\mathbf{M}(o_i,:)=\mathbf{0}_{1,N_T}$, $\mathbf{M}(:,o_i)=\mathbf{0}_{N_T,1}$
	(12)	$\mathbf{M}(1:N_T,1:N_T):=\mathbf{M}(1:N_T,1:N_T)-\mathbf{b}((1/\mathrm{MSE})\nabla)^H$
	(13)	**end**
Permutation	(14)	permute cols. in $\mathbf{Q}$ and $\mathbf{R}$ and rows in $\mathbf{R}$ according to O
CQRD	(15)	**for** $i=1,\ldots,N_T$
	(16)	$\mathbf{R}(i,i:N_T):=(1/\sqrt{\mathbf{R}(i,i)})\mathbf{R}(i,i:N_T)$
	(17)	$\mathbf{Q}(:,i):=(1/\mathbf{R}(i,i))\mathbf{Q}(:,i)$
	(18)	**for** $j=i+1,\ldots,N_T$
	(19)	$\mathbf{R}(j,j:N_T):=\mathbf{R}(j,j:N_T)-\mathbf{R}^*(i,j)\mathbf{R}(i,j:N_T)$
	(20)	$\mathbf{Q}(:,j):=\mathbf{Q}(:,j)-\mathbf{R}(i,j)\mathbf{Q}(:,i)$
	(21)	**end**
	(22)	**end**

* original CQRD proposed in [15] is modified here (lines (15)–(22)) and becomes more compact.

the order search according to either (50) or (48). Hence, at each deflation step, instead of permutation, we can remove the effect of current best column by directly zeroing the corresponding row and column in $\breve{\mathbf{M}}^{[k]}$.

Now with the detection order in hand, we proceed to determine the BS interface DFE filters. The disadvantage of the separate optimization of the order and DFE fitlers is a complexity penalty. However, this penalty can be partially compensated by using an ordered Cholesky based QRD (CQRD) algorithm [15]. In addition, compared with the standard QRD algorithm, the new decomposition involves the ZF and MMSE criteria without introducing the extended channel matrix, thereby avoiding the additional computation and memory required to meet the MMSE criterion. Recall that $\mathbf{C}$ is symmetric positive definite. Thus we have the following relationship:

$$\mathbf{P}^T\mathbf{C}\,\mathbf{P}=\mathbf{P}^T(\mathbf{H}^H\mathbf{H}+\gamma\mathbf{I}_{N_T})\mathbf{P}=\mathbf{R}^H\mathbf{R}, \tag{51}$$

which implies that $\mathbf{R}$ can be alternatively derived by Cholesky decomposition (CD) [21] of the already calculated $\mathbf{C}$. Furthermore, CD is more efficient than the operations associated with the $\mathbf{R}$ calculation in QRD. Once each row of $\mathbf{R}$ has been computed, we can use it to derive $\mathbf{Q}$ by orthogonalizing $\mathbf{H}$ column-wise as is done in the Gram-Schmidt procedure [21]. In particular, for $i=1,2,\ldots,N_T$:

$$\mathbf{q}_i=\widehat{\mathbf{h}}_i/r_{i,i}, \tag{52}$$

with $\widehat{\mathbf{h}}_i$ defined as:

$$\widehat{\mathbf{h}}_i=\mathbf{h}_i-\sum_{j=1}^{i-1}r_{j,i}\cdot\mathbf{q}_j. \tag{53}$$

As a result, a Cholesky based QR decomposition (CQRD) algorithm for BS interface filter generation is established. To reflect the detection order, an ordered CQRD should be performed by permuting the columns of $\mathbf{H}$ according to the order sequence.

We summarize the preprocessing of Hybrid-I in Table 3 using the MATLAB Notation. Following the initialization, the inverse of $\mathbf{M}$ is calculated using the N_T-step inflation on lines (2)–(7). At each step, the effect of a new column is added on lines (3)–(6). Then, the order sequence is determined through the deflated recursion on lines (8)–(13). Once the current best column is found, its effect is removed by zeroing on line (11). Compared with naive matrix inversion, double recursion is not only efficient in terms of the reduced arithmetic operations, but also tractable from an implementation perspective by avoiding the recursive size

reduction and permutation. After determining the order sequence, the parallel permutation is performed on line (14). Finally, we perform CQRD on lines (15)–(22), where **R** is computed row-wise from top to bottom and **Q** is derived column-wise from left to right. Note that both **M** and **R** are Hermitian, so a memory efficient implementation, which only accesses the upper (or lower) triangular portion, is possible.

4.2 Hybrid Algorithm II: Suboptimal Ordering

It is well known that the detection order is crucial for the DFE algorithm in terms of the suppression of error propagation. In particular, the first detected layer is most important in DFE because it influences all the remaining layers. In contrast, the layers detected later affect only a few layers so that the failure of detection therein will not strongly impact the total error performance. Therefore, we can form a suboptimal ordering criterion that replaces the recursive order search by simply sorting the layer-wise MSEs of the full channel matrix **H** in descending order

$$o_\kappa = \arg\min_{\substack{\kappa \\ o}} \{[\mathbf{M}]_{o,o} | o \in O_1\} \quad (\kappa = 1, 2, \ldots, N_\mathrm{T}), \quad (54)$$

where $[\arg\min_a F(\cdot)]$ stands for the argument for which the value of the function $F(\cdot)$ attains its ath minimum value. It is noteworthy that this MSE sorting rule guarantees the optimal BLAST order at the first detected layer which is most vital in terms of reducing the risk of error propagation. Hence, the suboptimal ordering criterion first computes **M** by using the inflated recursion described previously, then sorts the diagonal entries of **M** (i.e., the layer-wise MSEs) to determine the order sequence. Immediately, the deflated recursion and repeated minimum value searches, i.e., lines (8)–(13) in Table 3, can be replaced by a simple sorting function, thereby further lowering complexity.

5. Complexity Comparison

In this section, we show the computational cost that we should pay for various DFE detectors. Each complex multiplication and addition is counted as one floating/fixed point operation (flop), and then a complexity function associated with each algorithm in terms of N_R, N_T, and N_D is derived. The complexities of the data detection in NC and BS interface are $(4N_\mathrm{T}N_\mathrm{R} - N_\mathrm{T} - 2N_\mathrm{R})N_\mathrm{D}$ and $(N_\mathrm{T}^2 + 2N_\mathrm{T}N_\mathrm{R} - 2N_\mathrm{T})N_\mathrm{D}$, respectively. To avoid repetition, we only display the complexity of the preprocessing in the following.

In standard NC-BLAST, the preprocessing contains the computation of $\mathbf{M}^{[k]}$ and $\mathbf{G}_\mathrm{DFE}$. $\mathbf{M}^{[k]}$ is calculated via singular value decomposition (SVD) [21] for $k = N_\mathrm{T}, \ldots, 1$. The determination of $\mathbf{G}_\mathrm{DFE}$ is straightforward. Hence, the total complexity becomes

$$f_\mathrm{NC-BLAST} = \frac{9}{2}N_\mathrm{T}^4 + \frac{8}{3}N_\mathrm{T}^3 N_\mathrm{R} + 9N_\mathrm{T}^3 + 5N_\mathrm{T}^2 N_\mathrm{R} + \frac{9}{2}N_\mathrm{T}^2 + \frac{4}{3}N_\mathrm{T}N_\mathrm{R}, \quad (55)$$

The fast NC-BLAST-RMI algorithms reduce complexity by applying the Sherman-Morrison formula [8] and double recursion [9] to calculate the multiple matrix inversion. The subsequent flops are

$$f_\mathrm{NC-BLAST-RMI[8]} = \frac{7}{6}N_\mathrm{T}^3 + \frac{11}{2}N_\mathrm{T}^2 N_\mathrm{R} - N_\mathrm{T}^2 + \frac{7}{2}N_\mathrm{T}N_\mathrm{R} - \frac{13}{6}N_\mathrm{T} - N_\mathrm{R} + 1, \quad (56)$$

and

$$f_\mathrm{NC-BLAST-RMI[9]} = \frac{4}{3}N_\mathrm{T}^3 + 2N_\mathrm{T}^2 N_\mathrm{R} + N_\mathrm{T}N_\mathrm{R} - \frac{10}{3}N_\mathrm{T} + 1. \quad (57)$$

For BS-BLAST, which straightforward calculates the $\frac{1}{2}N_\mathrm{T}^2 + \frac{1}{2}N_\mathrm{T} - 1$ Gram-Schmidt based QRDs with respect to the reduced channel matrices to determine the order sequence, suffers from a computational load of

$$f_\mathrm{BS-BLAST} = \frac{2}{5}N_\mathrm{T}^5 + \frac{1}{2}N_\mathrm{T}^4 N_\mathrm{R} + \frac{9}{8}N_\mathrm{T}^4 + \frac{3}{2}N_\mathrm{T}^3 N_\mathrm{R} + \frac{3}{4}N_\mathrm{T}^3 + \frac{5}{4}N_\mathrm{T}^2 N_\mathrm{R} - \frac{1}{8}N_\mathrm{T}^2 + \frac{1}{4}N_\mathrm{T}N_\mathrm{R} - \frac{3}{20}N_\mathrm{T} - \frac{7}{2}N_\mathrm{R} - 2. \quad (58)$$

BS-BLAST-OQRD discussed here is a modified version of the algorithm in [12]. Instead of using square root propagation, we utilize the QRD so that the resulting algorithm is more efficient. We derive $\underline{\mathbf{Q}}$ and $\underline{\mathbf{R}}$ by QR decomposing $\underline{\mathbf{H}}$. The inversion of $\underline{\mathbf{R}}$ can be avoided by using the relation $\alpha\mathbf{I}_{N_\mathrm{T}} = \mathbf{Q}_2\underline{\mathbf{R}}$. After performing a sequence of unitary transformations and triangular matrix inversion, we finally obtain the BS interface filters. The resulting complexity is

$$f_\mathrm{BS-BLAST-OQRD} = \frac{11}{3}N_\mathrm{T}^3 + 4N_\mathrm{T}^2 N_\mathrm{R} + 3N_\mathrm{T}^2 + \frac{7}{2}N_\mathrm{T}N_\mathrm{R} + \frac{1}{3}N_\mathrm{T}. \quad (59)$$

In the new Hybrid-I, the detection order is optimized by utilizing the double recursion to calculate the multiple matrix inversion. The BS interface filters are derived using the ordered CQRD, so that it covers both the ZF and MMSE criteria while avoiding additional computations and memory assignment. In Hybrid-II, the deflated recursion is replaced by a simple sorting function, thereby leading to a further reduced complexity. The complexities of two algorithms become

$$f_\mathrm{Hybrid-I} = \frac{5}{3}N_\mathrm{T}^3 + 2N_\mathrm{T}^2 N_\mathrm{R} + N_\mathrm{T}N_\mathrm{R} - \frac{11}{3}N_\mathrm{T} + 1, \quad (60)$$

and

$$f_\mathrm{Hybrid-II} = \frac{4}{3}N_\mathrm{T}^3 + 2N_\mathrm{T}^2 N_\mathrm{R} - \frac{1}{2}N_\mathrm{T}^2 + N_\mathrm{T}N_\mathrm{R} - \frac{17}{6}N_\mathrm{T} + 1. \quad (61)$$

Assuming $N_\mathrm{T} = N_\mathrm{R}$, the speedups of Hybrid-I and Hybrid-II over the NC-BLAST (BS-BLAST) algorithm are about

$1.95N_T + 3.82$ and $2.15N_T + 4.20$ ($0.82N_T^2 + 0.72N_T + 0.55$ and $0.90N_T^2 + 0.79N_T + 0.60$). Compared with NC-BLAST-RMI[8] (BS-BLAST-OQRD), our proposals reduce the computational effort by a factor of about 1.82 and 2.00 (2.09 and 2.30). Furthermore, we emphasize that besides having fewer arithmetic operations, the implementation is further simplified by avoiding the recursive size reduction and/or permutation necessary in NC- BLAST and NC-BLAST-RMI [8]. Moreover, MMSE is involved without additional computations and memory assignment.

Figure 3 shows the results of a numerical experiment counting complex flops with $N_T = N_R$ and N_D=1. In this case, the channel variation is relatively fast so that the DFE filters must be adapted during each data sample. Hence, the complexity of the preprocessing becomes dominant. The computational advantage of both Hybrid-I and Hybrid-II, which is consistent with the theoretical analysis, can be obviously confirmed. Compared with the NC-BLAST-RMI algorithm in [9], only Hybrid-I receives a very minor complexity penalty (caused by the order and filter separate optimization), which means a shorter latency in NC-BLAST-RMI [9] is maintained.

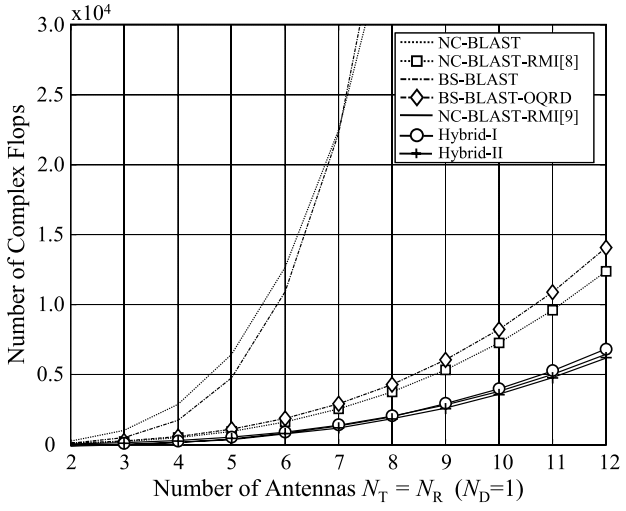


Fig. 3 Number of complex flops for various DFE algorithms in system with $N_T = N_R$ and $N_D = 1$.

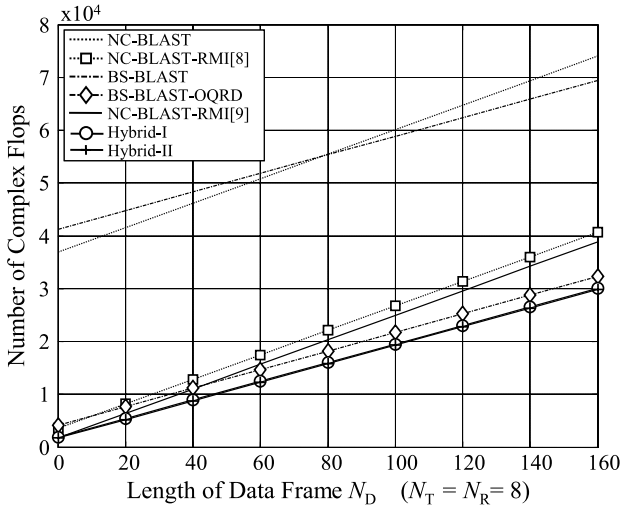


Fig. 4 Number of complex flops for various DFE algorithms in systems with $N_T = N_R = 8$ and $N_D = 0$ to 160.

Figure 4 displays the situations covering systems with varying frame lengths. We assume $N_T = N_R = 8$ and the length of signal frame to vary from $N_D = 0$ to 160. The dominant computational load now shifts from the preprocessing to the data detection with increasing N_D. Again, the new algorithms remarkably outperform the earlier methods. The advantage over NC interface detectors becomes greater with increasing frame length, since the BS interface is less complex than NC interface in data detection. Although the complexity of BS-BLAST-OQRD is close to that of the hybrid detectors, we emphasize that the new algorithms enable a shorter detection delay.

6. Performance Analysis

We verified the performance of the new algorithms through computer simulations. The data were counted under the assumption of $E_b/N_0 = \sigma_s^2 N_R/(\sigma_n^2 \log_2^M)$, where E_b denotes the average energy per bit per receive antenna and 4-QAM symbols are used. Duo to the mathematical equivalency between the NC and BS interface, we only display the BER of NC-BLAST for comparison. Furthermore, since we assume all signal vectors $\mathbf{s}(t)(1 \le t \le N_D)$ in one transmission frame experience the same fading and each element of $\mathbf{s}(t)$ is taken from the same constellation, the BER performance is independent of the frame length N_D. Therefore, we only display the error performance when $N_D = 1$ for brevity.

Figures 5 and 6 show the bit error rate (BER) of various ZF and MMSE detectors in a system with $N_T = N_R = 4$. As expected, all the DFE detectors outperform their linear counterparts. The impact of BLAST ordering is clear; it outperforms unordered DFE by 4.5 dB and 9.5 dB for a BER of 10^{-3} with the ZF and MMSE criterion. On the other hand, for both ZF and MMSE cases, the new Hybrid-I algorithm

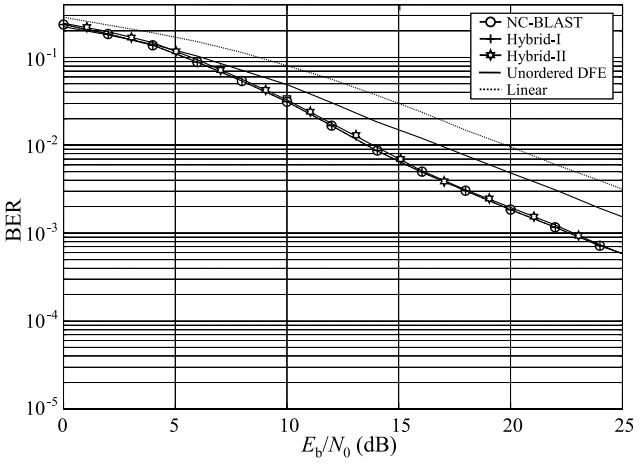


Fig. 5 BER performance of ZF detectors in system with $N_T = N_R = 4$.

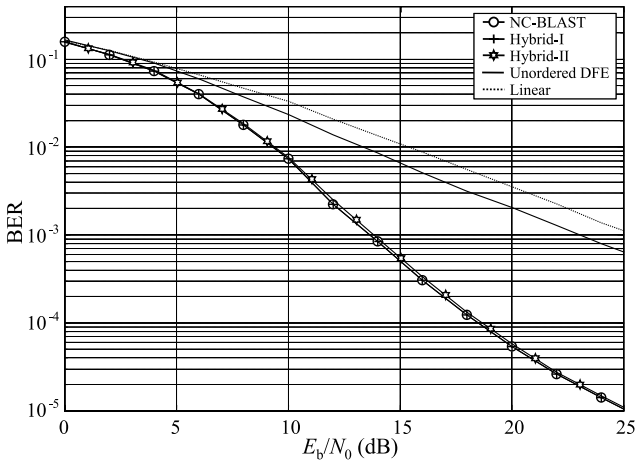


Fig. 6 BER performance of MMSE detectors in system with $N_T = N_R = 4$.

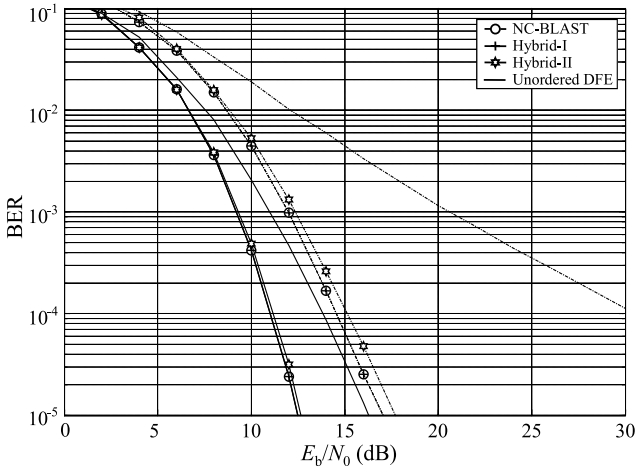


Fig. 7 BERs of MMSE detectors in system with $N_T = N_R = 8$ (dash-dotted), $N_T = 8$, $N_R = 12$ (solid).

achieves the exact BLAST performance and Hybrid-II keeps the performance loss negligible.

Figure 7 displays the performances of MMSE detectors in two systems with $N_T = 8$, $N_R = 8, 12$. The performance advantage of BLAST relative to the unordered DFE is still clear. Again, Hybrid-I achieves the optimal BLAST performance, and even in systems with large numbers of antennas, Hybrid-II keeps the performance penalty reasonably small. Moreover, this loss vanishes with increasing receive antennas.

Next, we verify the impact of finite word length on error performance by using the following quantization scheme. Let $Q(w, f)$ denote a quantization scheme in which totally w bits are used, of which f bits are used for the fractional part of the value. Namely, a value has $(w - f)$ bits of dynamic range and f bits of precision. Since the focus of this paper is not the architectural design of hardware, we only quantize the matrix $\mathbf{C}$ which is then used in NC-BLAST for joint order and DFE filter adaptation and is also used in new hybrid algorithms for order optimization

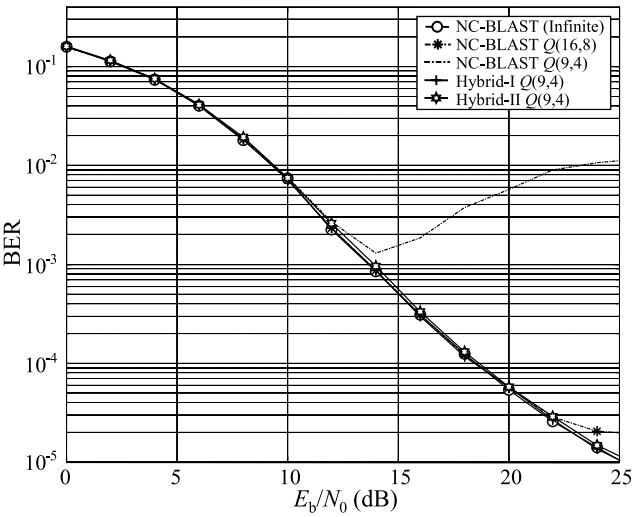


Fig. 8 Impact of finite word length precision on BER performance in system with $N_T = N_R = 4$.

only. For the new algorithms, thanks to the separate optimization of order and filters, we can use the unquantized $\mathbf{C}$ in CQRD. Figure 8 depicts the corresponding BER performance of MMSE detectors in a system with $N_T = N_R = 4$. Even when using quantization scheme $Q(9, 4)$, the new algorithms have only a very minor performance loss within the E_b/N_0 range of interest. For NC-BLAST, quantization scheme $Q(16, 8)$ should be adopted to achieve the same performance level. Otherwise, an unacceptable BER floor appears when using $Q(9, 4)$. It is well known that the reduction of word length leads to an almost linear reduction of the design area, hardware complexity, and power consumption. Furthermore, double recursion based matrix inversion (lines (2)–(13) for Hybrid-I or lines (2)–(7) for Hybrid-II in Table 3) is the most computational intensive part within the MIMO detection task. Therefore, new algorithms not only greatly reduce the arithmetic operations, but also enable one to remarkably compress the hardware and power requirement for each operation therein.

The advantages of the new algorithms are summarized as follows: 1) optimal and near optimal DFE performance with a total reduced number of arithmetic operations in frame based transmission, 2) shorter detection latency by performing fast hybrid preprocessing, 3) reduced hardware and power requirement due to relaxed word length for matrix inversion, 4) simplified implementation by avoiding recursive size reduction and permutation, 5) coverage of ZF and MMSE criteria without additional computations and memory assignment.

7. Conclusions

We proposed two hybrid DFE algorithms for signal detection in MIMO spatial multiplexing systems that achieve optimal and near optimal DFE error performance. The new algorithms combine the RMI and OQRD approaches, and obtain the advantages from both. The main advantages are:

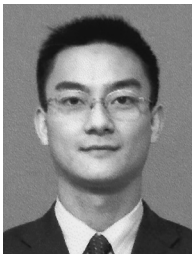
remarkably reduced total complexities in frame based transmission with various data lengths, shorter detection latency by carrying out a fast hybrid preprocessing, and largely reduced hardware and power requirement due to relaxed word length for matrix inversion. With the significant improvement on above important issues, we have made the spatial ordered DFE scheme much closer to the efficient real implementation.

Acknowledgment

The authors would like to thank Mr. Hiroshi Kazama of NTT Access Network Service Systems Laboratories and Mr. Kazuyasu Okada of NTT Network Innovation Laboratories for their continuous encouragement. They also express their deepest appreciation to their colleagues for their support during this work.

References

[1] I.E. Telatar, Capacity of multi-antenna Gaussian channels, Bell Labs., Murray Hill, NJ, 1995.
[2] G.J. Foschini and M.J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," Wirel. Pers. Commun., vol.6, no.3, pp.311–335, March 1998.
[3] G.J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," Bell Labs Tech. J., vol.1, no.2, pp.41–59, 1996.
[4] P.W. Wolniansky, G.J. Foschini, G.D. Golden, and R.A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," Proc. IEEE ISSSE, pp.295–300, Sept. 1998.
[5] G.D. Golden, G.J. Foschini, R.A. Valenzuela, and P.W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," Electron. Lett., vol.35, no.1, pp.14–16, Jan. 1999.
[6] C.Z.W. Hassell, J.S. Thompson, B. Mulgrew, and P.M. Grant, "A comparison of detection algorithms including BLAST for wireless communication using multiple antennas," Proc. IEEE PIMRC, vol.1, pp.698–703, Sept. 2000.
[7] A. Benjebbour, H. Murata, and S. Yoshida, "Comparison of ordered successive receivers for space-time transmission," Proc. IEEE VTC Fall, vol.4, pp.2053–2057, Oct. 2001.
[8] J. Benesty, Y. Huang, and J. Chen, "A fast recursive algorithm for optimum sequential signal detection in a BLAST system," IEEE Trans. Signal Process., vol.51, no.7, pp.1722–1730, July 2003.
[9] W. Jiang, Y. Asai, and S. Aikawa, "Very fast recursion based algorithm for BLAST signal detection in spatial multiplexing systems," IEICE Trans. Commun., vol.E90-B, no.7, pp.1773–1779, July 2007.
[10] G.J. Foschini, G.D. Golden, A. Valenzela, and P.W. Wolniansky, "Simplified processing for high spectral efficiency wireless communications emplying multi-element arrays," IEEE J. Sel. Areas Commun., vol.17, pp.1841–1852, Nov. 1999.
[11] G. Ginis and J.M. Cioffi, "On the relation between V-BLAST and the GDFE," IEEE Commun. Lett., vol.5, pp.364–366, Sept. 2001.
[12] B. Hassibi, "An efficient square-root algorithm for BLAST," Proc. IEEE ICASSP, vol.2, pp.11737–11740, June 2000.
[13] D. Wübben, R. Bohnke, V. Kuhn, and K.D. Kammeyer, "MMSE extension of V-BLAST based on sorted QR decomposition," Proc. IEEE VTC Fall, vol.1, pp.508–512, Oct. 2003.
[14] T. Koike, K. Kiyoshima, and S. Yoshida, "Low-complexity recursive algorithm for MIMO-BLAST receiver," IEICE Technical Report, RCS2006-60, July 2006.
[15] W. Jiang, Y. Asai, T. Onizawa, and S. Aikawa, "Fast algorithm for decision feedback equalization in multiple input multiple output channel," Proc. IEEE VTC Spring, vol.5, pp.2423–2428, May 2006.
[16] W. Jiang, Y. Asai, T. Onizawa, and S. Aikawa, "Efficient detection algorithms for spatial multiplexing systems," Proc. IEEE PIMRC, Sept. 2006.
[17] G. Lightbody, R. Woods, and R. Walke, "Design of a parameterizable silicon intellectual property core for QR-based RLS filtering," IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol.11, no.4, pp.659–678, Aug. 2003.
[18] Z. Guo and P. Nilsson, "A VLSI implementation of MIMO detection for future wireless communictions," Proc. IEEE PIMRC, vol.3, pp.2853–2856, Sept. 2003.
[19] L.M. Davis, "Scaled and decoupled Cholesky and QR decompositions with application to spherical MIMO detection," Proc. IEEE WCNC, vol.1, pp.326–331, March 2003.
[20] S.J. Bellis, W.P. Marnane, and P.J. Fish, "Alternative systolic array for non-square-root Cholesky decomposition," Proc. IEE Comput. Digit. Tech., vol.144, no.2, pp.57–64, March 1997.
[21] G.H. Golub and C.F. Van Loan, Matrix Computations, third ed., Johns Hopkins Univ. Press, 1996.
[22] S. Haykin, Adaptive Filter Theory, fourth ed., Prentice Hall, 2002.



Wenjie Jiang received B.E. and M.E. degrees in electrical engineering from Osaka Prefecture University, Osaka, Japan, in 2001 and 2003, respectively. Since joining NTT Laboratories in 2003, He has been engaged in research and development of high data rate wireless access systems. His current research interests are in signal processing for high data rate wireless systems, and MIMO wireless communications. He received the Young Researchers' Award from the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan in 2008. He is a member of IEEE.



Yusuke Asai received the B.E. and M.E. degrees from Nagoya University, Japan in 1997 and 1999, respectively. Since joining NTT Corporation in 1999, he has been engaged in research and development of coded OFDM systems and broadband MIMO-OFDM techniques for high-speed wireless LAN systems. He is currently a research engineer of NTT Access Network Service Systems Laboratories. He received the Young Researcher's Award from the IEICE of Japan in 2004. He is a member of IEEE.



Satoru Aikawa received the B.E. degree in electrical engineering from Yokohama National University, Japan, in 1984, and the Dr. ENG degree from the University of Tokyo, Japan, in 1995, respectively. Since 1984 he joined NTT Electrical Communications Laboratories. From 1984 to 1993, he has researched and developed FEC and Modem, especially, Trellis Coded 256QAM, for Microwave Radio Relay Systems. From 1993, he has been working on research and developing Wireless LAN Systems, especially, PHY and MAC technologies and their applications. Since October 2006, he is a professor of the Graduate School of Engineering, University of Hyogo. He is a member of IEEE.



Yasutaka Ogawa received the B.E., M.E. and Ph.D. degrees from Hokkaido University, Sapporo, Japan, in 1973, 1975, and 1978, respectively. Since 1979, he has been with Hokkaido University, where he is currently a Professor of the Graduate School of Information Science and Technology. During 1992–1993, he was with ElectroScience Laboratory, the Ohio State University, U.S.A., as a Visiting Scholar, on leave from Hokkaido University. His interests are in adaptive antennas, mobile communications, super-resolution techniques, and MIMO systems. Dr. Ogawa received the Yasujiro Niwa outstanding paper award in 1978, the Young Researchers' Award of IEICE Japan in 1982, and the Best Paper Award from IEICE Japan in 2007. Dr. Ogawa is a senior member of the IEEE.