PAPER *Special Section on Papers Selected from the 19th Symposium on Signal Processing*

Efficient Blind MAI Suppression in DS/CDMA Systems by Embedded Constraint Parallel Projection Techniques

Masahiro YUKAWA[†a)], ***Student Member***, **Renato L. G.** CAVALCANTE[†b)], ***Nonmember***, ***and*** **Isao** YAMADA[†c)], ***Regular Member***

SUMMARY This paper presents two novel blind set-theoretic adaptive filtering algorithms for suppressing "Multiple Access Interference (MAI)", which is one of the central burdens in DS/CDMA systems. We naturally formulate the problem of MAI suppression as an asymptotic minimization of a sequence of cost functions under some linear constraint defined by the desired user's signature. The proposed algorithms embed the constraint into the direction of update, and thus the adaptive filter moves toward the optimal filter without stepping away from the constraint set. In addition, using parallel processors, the proposed algorithms attain excellent performance with linear computational complexity. Geometric interpretation clarifies an advantage of the proposed methods over existing methods. Simulation results demonstrate that the proposed algorithms achieve (i) much higher speed of convergence with rather better bit error rate performance than other blind methods and (ii) much higher speed of convergence than the non-blind NLMS algorithm (indeed, the speed of convergence of the proposed algorithms is comparable to the non-blind RLS algorithm).

key words: *blind MAI suppression, DS/CDMA system, linearly constrained algorithms, adaptive projected subgradient method*

1. Introduction

The goal of this paper is to develop a blind Multiple Access Interference (MAI) suppressing algorithm, being "efficient" in the sense of (i) low computational complexity and (ii) high speed of convergence, for Direct Sequence Code-Division Multiple-Access (DS/CDMA) systems.

One of the noticeable advantages of CDMA systems is that users can share time and frequency by exploiting distinct spreading codes, or, in other words, users can transmit their information symbols at the same time and frequency. CDMA receivers, on the other hand, are usually affected by interference originated from transmitted symbols of other users. This is commonly referred to as MAI and it is known to deteriorate the overall capacity. A great deal of effort has been devoted to MAI suppression [1–11].

To realize high throughput systems, blind methods for MAI suppression, which do not require a training sequence (or pilot signals), have been particularly in great demand [6–11]. In 1995, Honig *et. al.* proposed a blind adaptive multiuser detection method [6], in which the problem is formulated as a constrained optimization with a linear constraint defined by the desired user's signature. In 1997, Park and Doherty proposed a simple set-theoretic blind method called *Space Alternating Generalized Projection (SAGP)* [7], which utilizes generalized projections onto non-convex sets (see Remark 2 and [12]). The SAGP exhibits better performance in the steady state at the expense of slower convergence rate than the method in [6]. In [13], it is reported that fast algorithms are necessary to keep good performance especially in wireless communications.

In 1998, Apolinário Jr. *et. al.* proposed the *Constrained Normalized Least Mean Square (CNLMS) algorithm* [14], which embeds the constraint used in [6] into the direction of update, providing fast convergence. Unfortunately, the CNLMS does not yet achieve sufficient speed of convergence because it takes just one datum into account at each iteration. In 2004, on the other hand, a fast blind MAI suppression method was proposed [15], which we call *Blind Parallel Projection (B2P) algorithm*. The B2P developed the idea of the SAGP by using a certain *parallel structure* and *convexification*, leading to excellent performance. The filter recursion (update) of the B2P is constructed by two steps at each iteration (cf. Remark 2): (i) shift the filter in descent directions of cost functions and (ii) enforce it in the constraint set.

This paper presents two embedded constraint blind algorithms for an adaptive MAI suppression filter. *Embedded constraint* and *parallel structure* are the keys to realize fast convergence with linear order complexity (see Remarks 1 and 2). The proposed algorithms develop the idea of the CNLMS for acceleration of convergence by taking into account more than one datum with several parallel processors at each iteration. Actually, the algorithms are derived from a set-theoretic adaptive filtering scheme named *Adaptive Projected Subgradient Method (APSM)* [16–18], which has been successfully applied to *the stereophonic acoustic echo cancellation problem* [19, 20]. Roughly speaking, the algorithms minimize asymptotically a sequence of cost functions that are defined by the received data at every sampling

Manuscript received December 10, 2004.
Manuscript revised March 4, 2005.
Final manuscript received April 28, 2005.

[†]The authors are with the Department of Communications and Integrated Systems, Tokyo Institute of Technology, Tokyo, 152-8552, Japan.
a) E-mail: masahiro@comm.ss.titech.ac.jp
b) E-mail: renato@comm.ss.titech.ac.jp
c) E-mail: isao@comm.ss.titech.ac.jp

time. Each iteration is constructed by two stages as follows. The first stage of the algorithms estimates the amplitude of the transmitted signal (as in [7]) and the transmitted bits. By using these estimates and the constraint used in [6], closed convex sets called *stochastic property sets* [see (9) in Sec. 3] are newly designed and, based on the distances to these sets, a reasonable cost function is defined. The second stage updates the MAI suppression filter in a descent direction of the cost function. The proposed algorithms have no need to enforce the filter in the constraint set unlike the SAGP or the B2P, since the constraint is embedded into the direction of update; i.e., the filter does not step away from the constraint set. Geometric interpretation clarifies an advantage of the proposed algorithms over the CNLMS, the SAGP and the B2P algorithms (see Remark 2). Simulation results exemplify dramatical improvements expected by the geometric interpretation.

Preliminary versions of this paper are presented in [21, 22].

2. Background

2.1 System Model

A Binary Phase-Shift Keying (BPSK) short-code DS/CDMA system is briefly summarized below. The system model considered in this paper is exactly the same as the one presented in [7, 11, 15]. Without loss of generality, assume that the desired user's signature $\boldsymbol{s}_1$ satisfies $\|\boldsymbol{s}_1\| = 1$ as in [7]. The received data process $(\boldsymbol{r}[i])_{i\in\mathbb{N}} \subset \mathbb{R}^N$ (N: the length of signature) is

$$\boldsymbol{r}[i] = A_1 b_1[i]\boldsymbol{s}_1 + \sum_{l=2}^{L} A_l \bar{b}_l[i]\bar{\boldsymbol{s}}_l + \boldsymbol{n}[i], \ \forall i \in \mathbb{N}, \quad (1)$$

where

$A_1 > 0$: amplitude of the 1-st (desired) user

$b_1[i] \in \{-1, 1\}$: i-th transmitted bit of the desired user

$\boldsymbol{s}_1 \in \left\{-\frac{1}{\sqrt{N}}, \frac{1}{\sqrt{N}}\right\}^N$: signature of the desired user

$\boldsymbol{n}[i] \in \mathbb{R}^N$: i-th noise vector.

Moreover, A_l ($2 \le l \le L$) is the amplitude of the l-th interference, and $\bar{b}_l[i]$ and $\bar{\boldsymbol{s}}_l$ are respectively the i-th interfering symbol bit and the interfering vector generated by l-th interfering user's parameters such as associated data bits and signature. In the presence of K users, the number of interferences $L-1$ can range from $K-1$ to $2(K-1)$, due to relative delays of the $K-1$ interfering users [4].

The problem addressed in this paper is to suppress efficiently the MAI, $\sum_{l=2}^{L} A_l\, \bar{b}_l[i]\bar{\boldsymbol{s}}_l$ in (1), with a linear filter without amplifying the noise $\boldsymbol{n}[i]$ severely.

2.2 Adaptive Projected Subgradient Method [16–18]

Let $C \subset \mathbb{R}^N$ be a nonempty closed convex set; a set $K \subset \mathbb{R}^N$ is said to be *convex* provided that $\forall \boldsymbol{x}, \boldsymbol{y} \in K$, $\forall \nu \in (0, 1)$, $\nu\boldsymbol{x} + (1-\nu)\boldsymbol{y} \in K$. Then, the *projection operator* $P_C : \mathbb{R}^N \to C$ maps a vector $\boldsymbol{x} \in \mathbb{R}^N$ to the unique vector $P_C(\boldsymbol{x}) \in C$ s.t. $\mathrm{d}(\boldsymbol{x}, C) := \|\boldsymbol{x} - P_C(\boldsymbol{x})\| = \min_{\boldsymbol{y}\in C} \|\boldsymbol{x} - \boldsymbol{y}\|$, where $\|\boldsymbol{x}\| := \langle \boldsymbol{x}, \boldsymbol{x}\rangle^{1/2}$, $\forall \boldsymbol{x} \in \mathbb{R}^N$ ($\langle \boldsymbol{x}, \boldsymbol{y}\rangle := \boldsymbol{x}^T\boldsymbol{y}$, $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$, and the superscript T stands for transposition). Also let $\Theta_n : \mathbb{R}^N \to [0, \infty)$ be a continuous *convex* function and $\partial\Theta(\boldsymbol{y})$ the *subdifferential*† of Θ at $\boldsymbol{y}$; a function $\Theta : \mathbb{R}^N \to \mathbb{R}$ is said to be *convex* if $\Theta(\nu\boldsymbol{x} + (1-\nu)\boldsymbol{y}) \le \nu\Theta(\boldsymbol{x}) + (1-\nu)\Theta(\boldsymbol{y})$, $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$ and $\forall \nu \in (0, 1)$. Then, the following scheme asymptotically minimizes $(\Theta_n)_{n\in\mathbb{N}}$ over C.

Scheme 1: (Adaptive Projected Subgradient Method (APSM) [16–18]) Generate a sequence $(\boldsymbol{h}_n)_{n\in\mathbb{N}}$ by

$$\boldsymbol{h}_{n+1} := \begin{cases} P_C\left(\boldsymbol{h}_n - \lambda_n \frac{\Theta_n(\boldsymbol{h}_n)}{\|\Theta_n'(\boldsymbol{h}_n)\|^2}\Theta_n'(\boldsymbol{h}_n)\right), \\ \qquad \text{if } \Theta_n'(\boldsymbol{h}_n) \neq \boldsymbol{0}, \\ \boldsymbol{h}_n, \quad \text{otherwise}, \end{cases} \quad (2)$$

where $\boldsymbol{h}_0 \in \mathbb{R}^N$, $\Theta_n'(\boldsymbol{h}_n) \in \partial\Theta_n(\boldsymbol{h}_n)$ and $\lambda_n \in [0, 2]$ is the *relaxation parameter*.

Basic properties of Scheme 1 are given in Appendix A.

3. Proposed Embedded Constraint Adaptive Algorithms

This section provides two set-theoretic algorithms for adaptation of a blind MAI suppression filter $\boldsymbol{h}_n \in \mathbb{R}^N$, where $n \in \mathbb{N}$ denotes the iteration number. All available data for the adaptation are assumed to be the sequence of received vectors $(\boldsymbol{r}[i])_{i\in\mathbb{N}}$ and the desired user's signature $\boldsymbol{s}_1$ (NOTE: In the absence of Inter-Chip Interference (ICI), the signature coincides with the spreading code and may be readily available [10]).

3.1 Set Design

To avoid the *self-nulling*††(i.e., canceling the desired user's signals), the following constraint is commonly imposed on the filter (e.g., [6]):

$$\boldsymbol{h}_n \in C_s := \{\boldsymbol{h} \in \mathbb{R}^N : \langle \boldsymbol{h}, \boldsymbol{s}_1\rangle = 1\}, \ \forall n \in \mathbb{N}. \quad (3)$$

†The *subdifferential* of Θ at $\boldsymbol{y}$ is the set of all the *subgradients* of Θ at $\boldsymbol{y}$; $\partial\Theta(\boldsymbol{y}) := \{\boldsymbol{a} \in \mathbb{R}^N : \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{a}\rangle + \Theta(\boldsymbol{y}) \le \Theta(\boldsymbol{x}), \forall \boldsymbol{x} \in \mathbb{R}^N\}$.

††In the case when the amplitude of some interference is greater than that of a desired user, the filter may track not the desired user but the interference. In such a case, the desired user's signal is suppressed. The set C_s can avoid such a situation.

Actually, $\langle \boldsymbol{h}_n, \boldsymbol{s}_1 \rangle$ can be any positive constant, however, for simplicity, we let $\langle \boldsymbol{h}_n, \boldsymbol{s}_1 \rangle = 1$. For any $\boldsymbol{h}_n \in C_s$, $\forall i \in \mathbb{N}$,

$$\langle \boldsymbol{h}_n, \boldsymbol{r}[i] \rangle = A_1 b_1[i] + \sum_{l=2}^{L} A_l \bar{b}_l[i] \langle \boldsymbol{h}_n, \bar{\boldsymbol{s}}_l \rangle + \langle \boldsymbol{h}_n, \boldsymbol{n}[i] \rangle . \quad (4)$$

For suppressing the MAI without amplifying noise severely, the second and third terms on the right side of (4) should be reduced as much as possible. Thus, a Minimum Mean-Squared Error (MMSE) filter is defined as follows [9]:

$$\boldsymbol{h}^* \in \operatorname*{argmin}_{\boldsymbol{h} \in C_s} E\left\{ (\langle \boldsymbol{h}, \boldsymbol{r}[i] \rangle - A_1 b_1[i])^2 \right\}, \quad (5)$$

where $E\{\cdot\}$ denotes the expectation; see Appendix B for relationship between the MMSE and the Minimum Output Energy (MOE) optimal filters. Since A_1 and $b_1[i]$ in (5) are not available, we use the following estimates [7]:

$$\widehat{A}_{1,n+1} := \widehat{A}_{1,n} + \gamma \left(|\langle \boldsymbol{h}_n, \boldsymbol{r}[n] \rangle| - \widehat{A}_{1,n} \right), \ \forall n \in \mathbb{N}, \quad (6)$$

$$\widehat{b}_{1,n}[i] := \operatorname{sgn} \langle \boldsymbol{h}_n, \boldsymbol{r}[i] \rangle , \ \forall n \in \mathbb{N}, \quad (7)$$

where $\widehat{A}_{1,n}$ ($\widehat{A}_{1,0} = 0$) and $\widehat{b}_{1,n}[i]$ are respectively estimates of the amplitude A_1 and the i-th transmitted bit $b_1[i]$ at iteration number n, and $\gamma \in (0,1]$ is the *forgetting factor*; see Remark 3. For simplicity, we define the signum function sgn: $\mathbb{R} \to \{-1, 1\}$ as, if $a > 0$, sgn $a = 1$, otherwise, sgn $a = -1$ ($\forall a \in \mathbb{R}$). With the estimates in (6) and (7), the problem is reformulated as finding a point in

$$\operatorname*{argmin}_{\boldsymbol{h} \in C_s} E\left\{ \left(\langle \boldsymbol{h}, \boldsymbol{r}[i] \rangle - \widehat{A}_{1,n+1} \widehat{b}_{1,n}[i] \right)^2 \right\}. \quad (8)$$

Instead of the expectation in (8), we newly introduce the following stochastic property sets [cf. Remark B.1 (d)]:

$$C_\rho^{(n)}[i] := \left\{ \boldsymbol{h} \in \mathbb{R}^N : \left(\langle \boldsymbol{h}, \boldsymbol{r}[i] \rangle - \widehat{A}_{1,n+1} \widehat{b}_{1,n}[i] \right)^2 \leq \rho \right\},$$
$$\forall n \in \mathbb{N}, \ \forall i \in \mathcal{I}_n := \{n, n-1, \cdots, n-q+1\}, \quad (9)$$

where $\mathcal{I}_n$ is the so-called *control sequence* (cf. [23]) with q elements (see Remark 3) and $\rho \geq 0$ is a parameter that determines the reliability of the set to contain the MMSE optimal filter $\boldsymbol{h}^*$ in (5). Intuitively, an increase of ρ inflates the set $C_\rho^{(n)}[i]$, and thus we call ρ *inflation parameter* (ρ should be described as $\rho_{n,i}$ because it can be designed independently for each set: in the following, however, such subscripts are omitted for notational simplicity).

Since C_s is completely reliable to contain $\boldsymbol{h}^*$, our strategy is to use C_s as a hard (absolute) constraint set and $\{C_\rho^{(n)}[i]\}_{i \in \mathcal{I}_n}$ as a collection of sets to which the distances should be reduced.

3.2 Proposed Algorithms

Let us derive the proposed algorithms from Scheme 1 with the sets in (3) and (9). Given $q \in \mathbb{N} \setminus \{0\}$, let $\{\omega_\iota^{(n)}\}_{\iota \in \mathcal{I}_n} \subset (0,1]$ satisfying $\sum_{\iota \in \mathcal{I}_n} \omega_\iota^{(n)} = 1$, $\forall n \in \mathbb{N}$, be the weights. Define the cost function

$$\Theta_n(\boldsymbol{h}) := \begin{cases} \displaystyle\sum_{\iota \in \mathcal{I}_n} \frac{\omega_\iota^{(n)}}{L_n^{(1)}} \mathrm{d}(\boldsymbol{h}_n, C_\rho^{(n)}[\iota] \cap C_s) \\ \mathrm{d}(\boldsymbol{h}, C_\rho^{(n)}[\iota] \cap C_s), \\ \text{if } L_n^{(1)} := \displaystyle\sum_{\iota \in \mathcal{I}_n} \omega_\iota^{(n)} \mathrm{d}(\boldsymbol{h}_n, C_\rho^{(n)}[\iota] \cap C_s) \neq 0, \\ 0, \qquad \text{otherwise}, \end{cases} \quad (10)$$

where $\mathrm{d}(\boldsymbol{h}, C_\rho^{(n)}[\iota] \cap C_s) \left(= \left\| \boldsymbol{h} - P_{C_\rho^{(n)}[\iota] \cap C_s}(\boldsymbol{h}) \right\| \right)$, $\forall \iota \in \mathcal{I}_n$, is the distance function of the variable vector $\boldsymbol{h} \in \mathbb{R}^N$ to the set $C_\rho^{(n)}[\iota] \cap C_s$ (which should be reduced). When $L_n^{(1)} \neq 0$ ($\Leftrightarrow$ $\boldsymbol{h}_n \notin \bigcap_{\iota \in \mathcal{I}_n} C_\rho^{(n)}[\iota] \cap C_s$), the weighting $\frac{\omega_\iota^{(n)}}{L_n^{(1)}} \mathrm{d}(\boldsymbol{h}_n, C_\rho^{(n)}[\iota] \cap C_s)$ is given to each distance function, where $L_n^{(1)}$ is the normalizing factor; the sets far from $\boldsymbol{h}_n$ have large weighting. When $L_n^{(1)} = 0$, we have $\boldsymbol{h}_n \in \bigcap_{\iota \in \mathcal{I}_n} C_\rho^{(n)}[\iota] \cap C_s$, hence nothing is left to do in this case. A subgradient of Θ_n at $\boldsymbol{h}_n$ is given by $\Theta'_n(\boldsymbol{h}_n) = \frac{1}{L_n^{(1)}} \sum_{\iota \in \mathcal{I}_n} \omega_\iota^{(n)} \left(\boldsymbol{h}_n - P_{C_\rho^{(n)}[\iota] \cap C_s}(\boldsymbol{h}_n) \right) \in \partial\Theta_n(\boldsymbol{h}_n)$ if $L_n^{(1)} \neq 0$; for details, see [18, p.607, Example 3].

Application of $C = \mathbb{R}^N$ and $\Theta_n(\boldsymbol{h})$ in (10) to Scheme 1 yields the following algorithm.

Algorithm 1: (Blind Parallel Constrained Projection Algorithm)

Requirements: the control sequence $\mathcal{I}_n$, the weights $\omega_\iota^{(n)} > 0$ s.t. $\sum_{\iota \in \mathcal{I}_n} \omega_\iota^{(n)} = 1$, the signature $\boldsymbol{s}_1$, the projection matrix $\boldsymbol{Q} := \boldsymbol{I} - \boldsymbol{s}_1 \boldsymbol{s}_1^T$ ($\boldsymbol{I}$: the identity matrix, NOTE: $\|\boldsymbol{s}_1\| = 1$), the inflation parameter $\rho \geq 0$, the step size $\lambda_n \in [0,2]$ and the forgetting factor $\gamma \in (0,1]$.

Initialization: $\widehat{A}_{1,0} = 0$, $\boldsymbol{h}_0 = \boldsymbol{s}_1 \in C_s$

Algorithm:

1) Estimation of A_1 and $b_1[\iota]$

$$\widehat{A}_{1,n+1} = \widehat{A}_{1,n} + \gamma \left(|\langle \boldsymbol{h}_n, \boldsymbol{r}[n] \rangle| - \widehat{A}_{1,n} \right)$$
$$\widehat{b}_{1,n}[\iota] = \operatorname{sgn} \langle \boldsymbol{h}_n, \boldsymbol{r}[\iota] \rangle , \quad \iota \in \mathcal{I}_n$$

2) Update of filter

$$\boldsymbol{h}_{n+1} = \boldsymbol{h}_n + \lambda_n \mathcal{M}_n^{(1)} \left(\sum_{\iota \in \mathcal{I}_n} \omega_\iota^{(n)} P_{C_\rho^{(n)}[\iota] \cap C_s}(\boldsymbol{h}_n) - \boldsymbol{h}_n \right), \quad (11)$$

where, for any $\boldsymbol{h} \in C_s$,

Table 1 Adaptive Blind Algorithms. $P_{C_s}(\boldsymbol{x}) = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{s}_1$, $\forall \boldsymbol{x} \in \mathbb{R}^N$.

Algorithm	Adaptation rule
OPM-GP	$\boldsymbol{h}_{n+1} = P_{C_s}\{\boldsymbol{h}_n + \mu(P_{H_n}(\boldsymbol{h}_n) - \boldsymbol{h}_n)\} = \boldsymbol{h}_n - \mu\frac{\langle \boldsymbol{h}_n, \boldsymbol{r}[n]\rangle}{\Vert\boldsymbol{r}[n]\Vert^2}\boldsymbol{Q}\boldsymbol{r}[n]$ where $H_n := \{\boldsymbol{h} : \langle \boldsymbol{h}, \boldsymbol{r}[n]\rangle = 0\}$
SAGP	$\widehat{A}_{1,n+1}$ & $\widehat{b}_{1,n}$ are computed by (6) & (7) $\boldsymbol{h}_{n+1} = P_{C_s}\left\{\boldsymbol{h}_n + \mu\left(P^{(g)}(\boldsymbol{h}_n) - \boldsymbol{h}_n\right)\right\}$, where $P^{(g)}(\boldsymbol{h}_n) := \begin{cases} P_{H_n^{(+)}}(\boldsymbol{h}_n) = \boldsymbol{h}_n - \frac{\langle \boldsymbol{h}_n, \boldsymbol{r}[n]\rangle - \widehat{A}_{1,n+1}}{\Vert\boldsymbol{r}[n]\Vert^2}\boldsymbol{r}[n], \text{ if } \langle \boldsymbol{h}_n, \boldsymbol{r}[n]\rangle > 0, \\ P_{H_n^{(-)}}(\boldsymbol{h}_n) = \boldsymbol{h}_n - \frac{\langle \boldsymbol{h}_n, \boldsymbol{r}[n]\rangle + \widehat{A}_{1,n+1}}{\Vert\boldsymbol{r}[n]\Vert^2}\boldsymbol{r}[n], \text{ otherwise,} \end{cases}$ with $H_n^{(\pm)} := \{\boldsymbol{h} : \langle \boldsymbol{h}, \boldsymbol{r}[n]\rangle = \pm\widehat{A}_{1,n+1}\}$
CNLMS	$\widehat{A}_{1,n+1}$ & $\widehat{b}_{1,n}$ are computed by (6) & (7) $\boldsymbol{h}_{n+1} = \boldsymbol{h}_n + \mu\left(P_{C_0^{(n)}[n]\cap C_s}(\boldsymbol{h}_n) - \boldsymbol{h}_n\right)$
B2P	$\widehat{A}_{1,n+1}$ is computed by (6) $\boldsymbol{h}_{n+1} = P_{C_s}\left\{\boldsymbol{h}_n + \lambda_n\left(\sum_{j=0}^{q-1} w_j^{(n)} P_{C_j^{(n)}}(\boldsymbol{h}_n) - \boldsymbol{h}_n\right)\right\}$ where $C_j^{(n)} := \{\boldsymbol{h} : \vert\langle \boldsymbol{h}, \boldsymbol{r}[n-j]\rangle\vert \le \widehat{A}_{1,n+1}\}$

$$P_{C_\rho^{(n)}[\iota]\cap C_s}(\boldsymbol{h}) = \begin{cases} \boldsymbol{h} - \dfrac{\langle \boldsymbol{h}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] - \sqrt{\rho}}{\boldsymbol{r}[\iota]^T\boldsymbol{Q}\boldsymbol{r}[\iota]}\boldsymbol{Q}\boldsymbol{r}[\iota], \\ \qquad \text{if } \langle \boldsymbol{h}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] > \sqrt{\rho}, \\ \boldsymbol{h} - \dfrac{\langle \boldsymbol{h}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] + \sqrt{\rho}}{\boldsymbol{r}[\iota]^T\boldsymbol{Q}\boldsymbol{r}[\iota]}\boldsymbol{Q}\boldsymbol{r}[\iota], \\ \qquad \text{if } \langle \boldsymbol{h}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] < -\sqrt{\rho}, \\ \boldsymbol{h}, \qquad \text{otherwise,} \end{cases} \tag{12}$$

$$\mathcal{M}_n^{(1)} := \begin{cases} \dfrac{\sum_{\iota\in\mathcal{I}_n} \omega_\iota^{(n)} \left\Vert P_{C_\rho^{(n)}[\iota]\cap C_s}(\boldsymbol{h}_n) - \boldsymbol{h}_n\right\Vert^2}{\left\Vert\sum_{\iota\in\mathcal{I}_n} \omega_\iota^{(n)} P_{C_\rho^{(n)}[\iota]\cap C_s}(\boldsymbol{h}_n) - \boldsymbol{h}_n\right\Vert^2}, \\ \qquad \text{if } \boldsymbol{h}_n \notin \bigcap_{\iota\in\mathcal{I}_n} C_\rho^{(n)}[\iota] \cap C_s, \\ 1, \qquad \text{otherwise.} \end{cases}$$

NOTE: For all $n \in \mathbb{N}$, $\boldsymbol{h}_n \in C_s$ holds, since (i) $\boldsymbol{h}_0 \in C_s$ and (ii) $\boldsymbol{h}_n \in C_s \Rightarrow \boldsymbol{h}_{n+1} \in C_s$ from (3) and (11). The proof of (12) is given in Appendix C. A weighted average of multiple projections as in (11) is referred to as *parallel projection* [24], since it can be computed in parallel by using q concurrent processors.

On the other hand, application of $C = \mathbb{R}^N$ and $\Theta_n(\boldsymbol{h}) := \Phi_n(P_{C_s}(\boldsymbol{h}))$, where

$$\Phi_n(\boldsymbol{h}) := \begin{cases} \displaystyle\sum_{\iota\in\mathcal{I}_n} \frac{\omega_\iota^{(n)}}{L_n^{(2)}} \mathrm{d}(\boldsymbol{h}_n, C_\rho^{(n)}[\iota])\mathrm{d}(\boldsymbol{h}, C_\rho^{(n)}[\iota]), \\ \text{if} \quad L_n^{(2)} := \sum_{\iota\in\mathcal{I}_n} \omega_\iota^{(n)} \mathrm{d}(\boldsymbol{h}_n, C_\rho^{(n)}[\iota]) \neq 0, \\ 0, \qquad \text{otherwise,} \end{cases}$$

to Scheme 1 yields the following algorithm (for details about the derivation of the algorithm, see [18, p.610, Example 5]).

Algorithm 2: (Blind Constrained Parallel Projection Algorithm)

Requirements & Initialization: the same as Algorithm 1

Algorithm:

1) Estimation of A_1 and $b_1[\iota]$: the same as Algorithm 1

2) Update of filter

$$\boldsymbol{h}_{n+1} = \boldsymbol{h}_n + \lambda_n \mathcal{M}_n^{(2)} P_{\widetilde{C}_s}\left(\sum_{\iota\in\mathcal{I}_n} \omega_\iota^{(n)} P_{C_\rho^{(n)}[\iota]}(\boldsymbol{h}_n) - \boldsymbol{h}_n\right), \tag{13}$$

where $\widetilde{C}_s := \{\boldsymbol{h} \in \mathbb{R}^N : \langle \boldsymbol{h}, \boldsymbol{s}_1\rangle = 0\}$ is a translated linear subspace of C_s and

$$P_{\widetilde{C}_s}(\boldsymbol{h}) = \boldsymbol{Q}\boldsymbol{h},$$

$$P_{C_\rho^{(n)}[\iota]}(\boldsymbol{h}) = \begin{cases} \boldsymbol{h} - \dfrac{\langle \boldsymbol{h}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] - \sqrt{\rho}}{\Vert\boldsymbol{r}[\iota]\Vert^2}\boldsymbol{r}[\iota], \\ \qquad \text{if } \langle \boldsymbol{h}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] > \sqrt{\rho}, \\ \boldsymbol{h} - \dfrac{\langle \boldsymbol{h}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] + \sqrt{\rho}}{\Vert\boldsymbol{r}[\iota]\Vert^2}\boldsymbol{r}[\iota], \\ \qquad \text{if } \langle \boldsymbol{h}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] < -\sqrt{\rho}, \\ \boldsymbol{h}, \qquad \text{otherwise,} \end{cases}$$

$$\mathcal{M}_n^{(2)} := \begin{cases} \dfrac{\sum_{\iota\in\mathcal{I}_n} \omega_\iota^{(n)} \left\Vert P_{C_\rho^{(n)}[\iota]}(\boldsymbol{h}_n) - \boldsymbol{h}_n\right\Vert^2}{\left\Vert P_{\widetilde{C}_s}\left(\sum_{\iota\in\mathcal{I}_n} \omega_\iota^{(n)} P_{C_\rho^{(n)}[\iota]}(\boldsymbol{h}_n) - \boldsymbol{h}_n\right)\right\Vert^2}, \\ \text{if } \sum_{\iota\in\mathcal{I}_n} \omega_\iota^{(n)} P_{C_\rho^{(n)}[\iota]}(\boldsymbol{h}_n) - \boldsymbol{h}_n \notin \widetilde{C}_s^{\perp}, \\ 1, \qquad \text{otherwise.} \end{cases}$$

For any linear subspace $M \subset \mathbb{R}^N$, $M^\perp \subset \mathbb{R}^N$ is defined as $M^\perp := \{\boldsymbol{x} \in \mathbb{R}^N : \langle \boldsymbol{x}, \boldsymbol{m}\rangle = 0, \forall \boldsymbol{m} \in M\}$. Algorithm 2 belongs to the family of *Embedded Constraint Adaptive Projected Subgradient Method (EC-APSM)* [16–18]. Moreover, Algorithm 1 can be generalized into a new family of embedded constraint algorithms shown in Appendix D.

Remark 1: (Computational Complexity)
Note that the computation of $\boldsymbol{Q}\boldsymbol{a} = \boldsymbol{a} - \boldsymbol{s}_1(\boldsymbol{s}_1^T\boldsymbol{a})$, $\forall \boldsymbol{a} \in \mathbb{R}^N$, requires $2N$ multiplications. Moreover,

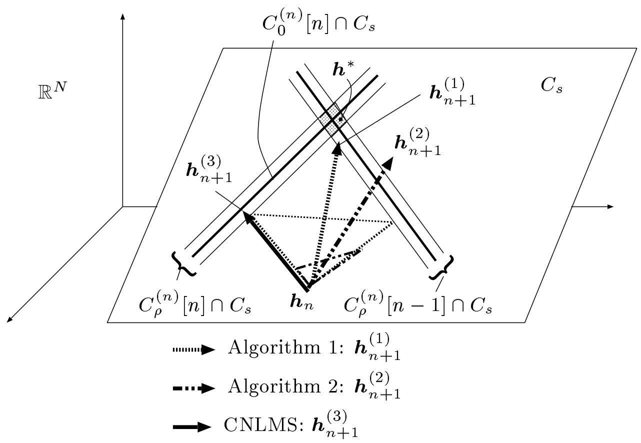

Fig. 1 A geometric interpretation of embedded constraint methods: the proposed algorithms and the CNLMS algorithm. The dotted area shows $C_\rho^{(n)}[n] \cap C_\rho^{(n)}[n-1] \cap C_s$.

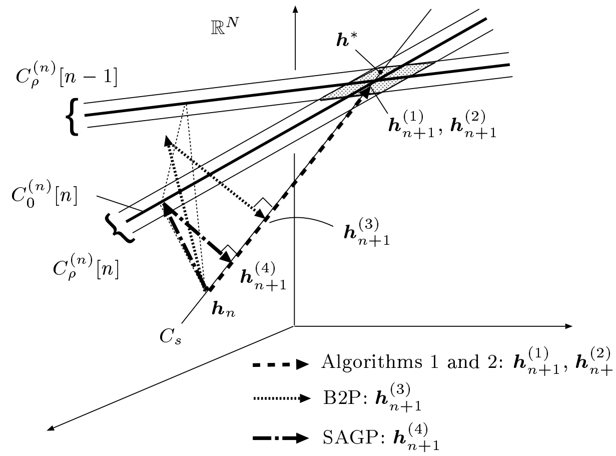

Fig. 2 A geometric interpretation of non-embedded constraint methods (the SAGP and the B2P) and the proposed methods. The dotted area shows $C_\rho^{(n)}[n] \cap C_\rho^{(n)}[n-1]$.

$\forall \iota \in \mathcal{I}_n \setminus \{n\}$, "$\boldsymbol{r}[\iota]^T \boldsymbol{Q} \boldsymbol{r}[\iota]$ and $\boldsymbol{Q}\boldsymbol{r}[\iota]$ in Algorithm 1" and "$\|\boldsymbol{r}[\iota]\|^2$ in Algorithm 2" are computed at the previous iterations. Hence, we see that both Algorithms 1 and 2 require $(4q+5)N$ multiplications at each iteration. Furthermore, note that each term in the summation in (11) [or (13)] can be computed in parallel (independently). Therefore, with q concurrent processors, the number of multiplications imposed on each processor is reduced to $9N$ no matter how many projections are used; i.e., the complexity order imposed on each processor is linear. This implies that the proposed algorithms are suitable for real-time implementation. On the other hand, the RLS-based-MMSE method [8] and the subspace approach [9], which are well-known blind methods, require $O(N^2)$ and $(4L+3)N+O(L)$ multiplications, respectively. Moreover, for good performance, the subspace approach needs to detect the exact number of strong interferences, which increases the overall system complexity.

Table 1 gives a unified view, with projection operators, to the following blind algorithms: the normalized OPM-based gradient projection (OPM-GP) [6, 7], the SAGP [7], the blind CNLMS that is based on the idea of [14] combined with our defining sets in (9), and the B2P [15]. The OPM-GP [7] is a normalized version of the blind MOE algorithm [6]; the algorithms are called respectively *projected NLMS* and *projected LMS* in [25]. It is not hard to see that the CNLMS is a special case of Algorithm 1 with $q=1$ and $\rho=0$. It should be remarked that the steady state performance of the B2P and the SAGP may be different, since the algorithms use different sets as shown in Table 1. The SAGP utilizes the so-called *generalized projection* $P^{(g)}(\boldsymbol{h}_n)$ (see e.g., [12]), which gives a nearest point from $\boldsymbol{h}_n$ in the non-convex set $H_n^{(+)} \cup H_n^{(-)}$. The generalized projection is not a strict projection because it is not always unique (cf. the definition of projection in Sec. 2-B). In fact, if $\langle \boldsymbol{h}_n, \boldsymbol{r}[n] \rangle = 0$, there exist two nearest points from $\boldsymbol{h}_n$ in $H_n^{(+)} \cup H_n^{(-)}$; $P_{H_n^{(+)}}(\boldsymbol{h}_n)$ and $P_{H_n^{(-)}}(\boldsymbol{h}_n)$. Fortunately, a geometric comparison of the SAGP with the proposed algorithms is possible (see Remark 2), since $C_0^{(n)}[n]$ coincides with $H_n^{(+)}$ [or $H_n^{(-)}$] when $\mathrm{sgn}\langle \boldsymbol{h}_n, \boldsymbol{r}[n] \rangle = 1$ (or $\mathrm{sgn}\langle \boldsymbol{h}_n, \boldsymbol{r}[n] \rangle = -1$), by (7), (9) and Table 1. It is easily seen that $C_0^{(n)}$ used in the B2P is a closed convex set bounded by the hyperplanes $H_n^{(+)}$ and $H_n^{(-)}$ used in the SAGP.

Figures 1 and 2 illustrate geometric interpretations of the proposed algorithms compared with a simple embedded constraint method (the CNLMS) and non-embedded constraint methods (the SAGP and the B2P), respectively. A geometric interpretation of the OPM-GP is also possible; the set H_n is nothing but the translated subspace of $H_n^{(+)}$ [or $H_n^{(-)}$]. For visual clarity, however, it is omitted. For the proposed algorithm and the B2P, the uniform weights, $\omega_\iota^{(n)} = 1/2$ ($\forall \iota = 1, 2$), are employed with $q=2$ parallel processors. For the B2P, the step size is set to $\mathcal{M}_n$. For the other methods, the step sizes are set to 1. The MMSE optimal filter $\boldsymbol{h}^*$ is assumed to satisfy $\boldsymbol{h}^* \in C_\rho^{(n)}[n] \cap C_\rho^{(n)}[n-1] \cap C_s$. All algorithms are assumed to have, if necessary, a common amplitude estimation $\widehat{A}_{1,n+1}$ and a correct bit estimation $\widehat{b}_{1,n}[n]$. A remark on geometric comparisons is given below.

Remark 2: (Geometric Comparisons)
Referring to Fig. 1, we see that the proposed algorithms generate closer points to the MMSE optimal filter $\boldsymbol{h}^*$ than the CNLMS due to its parallel structure; i.e., the proposed algorithms utilize multiple data simultaneously. As also seen in the figure, Algorithm 1 takes an averaged direction of exact projections onto $\{C_\rho^{(n)}[\iota] \cap C_s\}_{\iota \in \mathcal{I}_n}$, while Algorithm 2 takes an averaged direction of relaxed projections. The "relaxation" de-

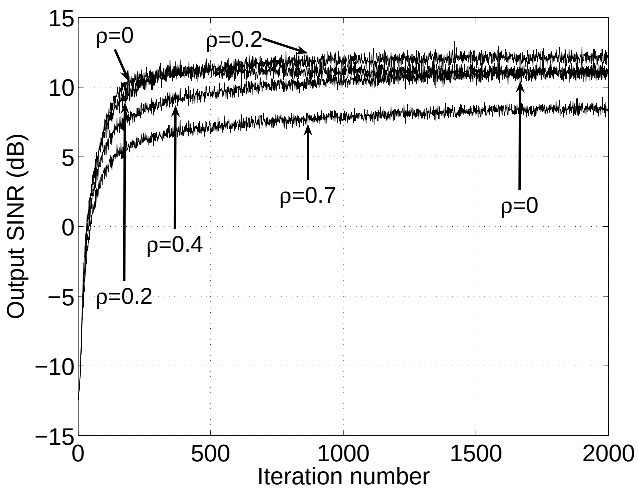

Fig. 3 SINR curves of Algorithm 1 with different values of inflation parameter ρ under SNR=15 dB.

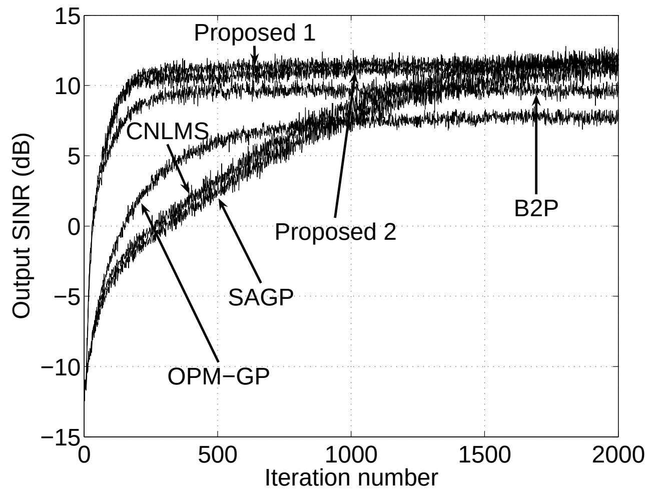

Fig. 4 Proposed algorithms versus other blind methods in SINR under SNR=15 dB.

pends on the angle between $\boldsymbol{s}_1$ (the normal vector of the hyperplane C_s) and $\boldsymbol{r}[\iota]$ (the one of the boundary hyperplanes of $C_\rho^{(n)}[\iota]$).

Referring to Fig. 2, we see that the B2P generates a closer point to $\boldsymbol{h}^*$ than the SAGP due to its parallel structure. The proposed algorithms generate even closer points than the B2P due to its embedded constraint structure in addition to its parallel structure. We also see that the SAGP and the B2P are constructed by two steps; the second step $P_{C_s}(\cdot)$ in Table 1 is to enforce the filter in the constraint set. On the other hand, the CNLMS and the proposed algorithms update the filter along the constraint set, and hence they are constructed by one step.

Finally, from our observation, a simple strategy for the design of γ and q [cf. (6) and (9)] is given below.

Remark 3: (On Design of γ and q)
From Remark B.1 (d) in Appendix B, $\widehat{A}_{1,n+1} \approx A_1$ should be valid for good steady state performance, which can be obtained with small γ, although it may decrease the speed of convergence [7]. From our experience, q leads to good performance when $T_{a.c.}/qT_b > 0.1$, where $T_{a.c.}$ and T_b denote the period when the channels are almost constant and the bit period, respectively.

To exemplify the discussion in Remark 2, we present numerical comparisons in the following section.

4. Simulation Results

This section provides the results of some computer simulations, all of which are performed under the following conditions. The number of interfering users is $(K-1)=5$, and all users have amplitude 10 times greater than the amplitude of the desired signal $A_1=1$. Signals are modulated by 31-length Gold sequences

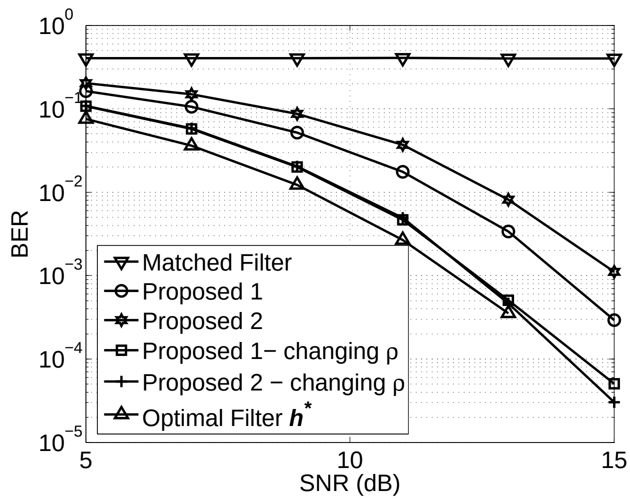

Fig. 5 BER curves of the proposed algorithms with ρ (i) fixed throughout simulations and (ii) switched after convergence.

$(N=31)$, which are chosen randomly. For all algorithms, $\boldsymbol{h}_0 = \boldsymbol{s}_1 \in C_s$ is employed and, if the estimation of the amplitude is needed, the forgetting factor is set to $\gamma = 0.01$, by following the way in [7, 15].

4.1 Effects of Inflation Parameter

First, the effects of the inflation parameter ρ in (9) are examined. Figure 3 compares the output Signal to Interference-plus-Noise Ratio (SINR) performance of Algorithm 1, which at the n-th iteration is obtained by

$$\mathrm{SINR}_n := \frac{\sum_{u=1}^{U} \left\langle \boldsymbol{h}_n^{(u)}, \boldsymbol{s}_1 \right\rangle^2}{\sum_{u=1}^{U} \left[\frac{\left\langle \boldsymbol{h}_n^{(u)}, \boldsymbol{r}^{(u)}[n] - A_1^{(u)} b_1^{(u)}[n] \boldsymbol{s}_1 \right\rangle}{A_1^{(u)}} \right]^2}.$$

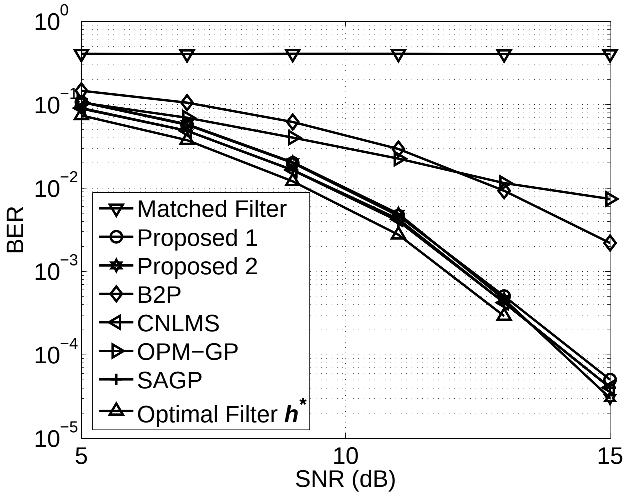

Fig. 6 Proposed algorithms versus other blind methods in BER.

Here $\boldsymbol{h}_n^{(u)}$ and $\boldsymbol{r}^{(u)}[n]$ are the respective vectors at the u-th realization, $A_1^{(u)}$ and $b_1^{(u)}[n]$ are respectively the amplitude and the n-th transmitted bits of the desired user at the u-th realization, and $U = 500$ is the number of realizations. For simplicity, the path delays of users 2 to 6 are integer multiples of the chip rate $aT_c \in [0, T_b]$, $a \in \mathbb{N}$, which were chosen randomly with equal probability among the given multiples at every realizations. The simulations are performed under Signal (or bit energy) to Noise Ratio (SNR) $:= 10\log_{10}\frac{A_1^2}{\sigma_n^2}$ =15 dB, where σ_n^2 is the variance of additive noise. Different fixed values, $\rho = 0$, 0.2, 0.4 and 0.7, are assigned to the inflation parameter. For simplicity, we set $\boldsymbol{r}[i] = \boldsymbol{r}[1]$ for $i \le 1$, and $\omega_\iota^{(n)} = \frac{1}{q}$, $\forall \iota \in \mathcal{I}_n$. The step size $\lambda_n = 0.2$ (see also below) is employed with $q = 16$ parallel projections.

We observe that, although "$\rho = 0$" exhibits the fastest initial convergence in the experiments, "$\rho = 0.2$" achieves better steady state performance ("$\rho = 0.4$" and "$\rho = 0.7$" are also expected to achieve higher SINR than "$\rho = 0$" after more iterations). Considering the performance in the initial and steady states, "$\rho = 0.2 \approx 6\sigma_n^2$" may be an effective fixed value in this simulation. Note, however, that ρ should be designed by taking into account influence of MAI and estimation errors in $\widehat{A}_{1,n+1}$ & $\widehat{b}_{1,n}[i]$ as well as noise. Hence, the design of inflation parameter needs additional discussion, which will be addressed in a future work; a simple fundamental analysis on this designing problem is reported in [26]. With an appropriately designed inflation parameter, the step size λ_n can naturally be set to 1; $\boldsymbol{h}^*$ may not belong to the simple sets we designed herein, and $\lambda_n = 0.2$ realizes robustness against such a situation in our simulations.

A simple review of Fig. 3 brings a natural suggestion that excellent performance in both initial and steady states will be simultaneously realized by assigning "$\rho = 0$" at the beginning and "an appropriate value of ρ" after convergence; this suggestion is consistent with the results in [26]. To verify this suggestion, additional experiments are performed below.

4.2 Proposed Methods with Change of Inflation Parameter & Comparison with Other Blind Methods

We assign 0 at the beginning and 0.7 after iteration number 500 to the inflation parameter ρ, and the other parameters are the same as employed in Fig. 3. Figure 4 compares the SINR performance, under SNR =15 dB, of the proposed algorithms with the ones presented in Table 1 (For comparisons with another major blind method, the Constant Modulus with Amplitude Estimation (CMAE) [11], see [15]). For Algorithm 2 and the B2P, the same parameters as Algorithm 1 are employed (For the B2P, the step size is set to $\lambda_n = 0.2\mathcal{M}_n$). For the OPM-GP, the SAGP and the CNLMS, step sizes are set to 0.2 for a fair comparison. As expected from Remark 2, we observe that the proposed algorithms outperform all other methods in terms of speed of convergence, while attaining excellent SINR in the steady state. Moreover, the additional computational complexity imposed by the proposed algorithms can be somehow alleviated by using processors in parallel (see Remark 1). As suggested in the end of Sec. 4-A, we observe that the steady state performance of Algorithm 1 is improved by around 1 dB, although, judged from Fig. 3, the choice of $\rho = 0.7$ may not be the best.

To highlight the steady state performance, the Bit Error Rate (BER) performance is depicted in Figs. 5 and 6 over SNR ranging from 5 to 15 dB. To capture the steady state performance in a fair manner, 6000 bits are transmitted at each realization and the last 1000 bits for 100 realizations are used to calculate the BER. For a comparison, the line by the optimal filter $\boldsymbol{h}^*$ is depicted, which is computed by (B.1) and $\boldsymbol{R_r} = A_1^2\boldsymbol{s}_1\boldsymbol{s}_1^T + \sum_{l=2}^{L} A_l^2\bar{\boldsymbol{s}}_l\bar{\boldsymbol{s}}_l^T + \sigma_n^2\boldsymbol{I}$, with full information, based on the independence assumption.

Figure 5 compares the BER of the proposed algorithms with "changing the inflation parameter ρ as in Fig. 4" and "fixing ρ to 0". We see that the BER performance is significantly improved due to the change of ρ. In Fig. 6, the BER performance of the proposed algorithms with changing ρ is compared with the blind methods employed in Fig. 4. Referring to Figs. 4 and 6, we observe that the proposed algorithms achieve much faster convergence in SINR than the SAGP and the CNLMS as well as almost the same BER performance as the optimal filter. Also we observe that the proposed algorithms drastically outperform the OPM-GP and the B2P in BER. Reviewing Fig. 3 and considering that the CNLMS is a special case of Algorithm 1 with $q = 1$ (see (11) and Table 1), another suggestion

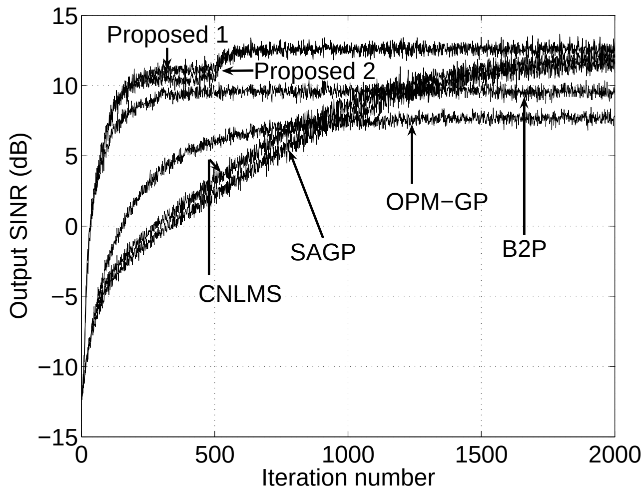

Fig. 7 "Proposed algorithms with q switched to 1 after convergence" versus other blind methods in SINR under SNR=15 dB.

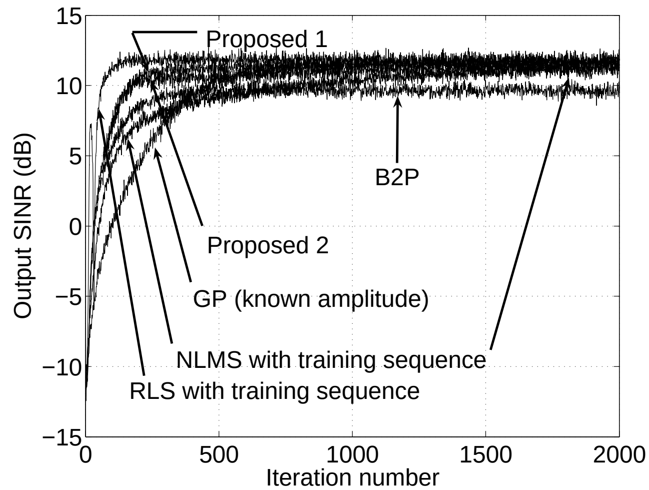

Fig. 8 Proposed algorithms versus non-blind methods in SINR under SNR=15 dB. RLS: $\delta = 1.0 \times 10^{-4}$ and $\lambda = 0.98$. NLMS: $\mu = 0.6$. GP: $\mu = 0.6$.

is brought that the steady state performance of Algorithm 1 will also be improved by switching q to 1 after convergence.

To verify this second suggestion, further experiments for the proposed algorithms are performed under SNR $= 15$ dB, where the number of parallel projections is set to $q = 16$ at the beginning and it is switched to 1 at iteration number 500 and the inflation parameter is fixed to $\rho = 0$ throughout the simulations. The other parameters are the same as in Fig. 4. Figure 7 compares the SINR performance of the proposed algorithms with the blind methods used in Fig. 4. We observe that the performance in the steady state is efficiently improved by decreasing the number of parallel projections after convergence, as expected by the second suggestion. This switching strategy is easily realized in hardware implementation.

4.3 Comparison with Non-Blind Methods

Finally, Fig. 8 compares the proposed algorithms, under SNR=15 dB, with the non-blind (semi-blind) algorithms; Generalized Projection (GP) algorithm [7] with known amplitude of desired user, the Normalized Least Mean Square (NLMS) and the Recursive Least Squares (RLS) algorithms [27] with training sequences. For the non-blind methods, parameters are adjusted to achieve the fastest noticeable rate of convergence. For the proposed algorithms and the B2P, the employed parameters are the same as in Fig. 4. We observe that the proposed algorithms achieve rather faster convergence than the non-blind NLMS, and exhibit comparable speed of convergence to the non-blind RLS. These remarkable improvements are accomplished by the embedded constraint and parallel structures.

5. Concluding Remarks

This paper has presented two blind adaptive filtering algorithms for the MAI suppression in DS/CDMA systems. Since the proposed algorithms are based on the parallel projection with the embedded constraint structure, they achieve closer points to the MMSE optimal filter than the existing methods at each iteration. Simulation results have demonstrated that the proposed algorithms exhibit excellent performance.

The extensive experiments presented in this paper suggest that the A-PCP (see Appendix D) and the EC-APSM may include excellent embedded constraint algorithms. Those two families of embedded constraint algorithms (i.e., A-PCP and EC-APSM) are expected to be useful not only in communications but also in a wide range of applications. In the presented simulations, we focus on the uniform weights ($\omega_\iota^{(n)} = \frac{1}{q}$, $\forall \iota \in \mathcal{I}_n$; see the previous section) for simplicity. For further improvements, an efficient strategic weighting technique such as the one proposed in [28, 29] would be effective. An extension of the proposed algorithms is possible, by following the way in [30], to the more general case when the signature at the receiver is complex-valued as in multipath channels. Furthermore, to employ the proposed algorithms in multipath channels, channel estimation techniques should be used and the algorithms must be robust against errors in the channel estimation. An extension of the proposed algorithms to such cases will be addressed in a future work.

Appendix A: Properties of Scheme 1

Scheme 1 has the following properties [16–18].

(a) (Monotonicity)

$$\left\| \boldsymbol{h}_{n+1} - \boldsymbol{h}^{*(n)} \right\| \leq \left\| \boldsymbol{h}_n - \boldsymbol{h}^{*(n)} \right\|, \ \forall n \in \mathbb{N},$$

$\forall \boldsymbol{h}^{*(n)} \in \Omega_n := \{\boldsymbol{h} \in C : \Theta_n(\boldsymbol{h}) = \inf_{\boldsymbol{x} \in C} \Theta_n(\boldsymbol{x})\}$.

(b) (Asymptotic minimization)
Suppose $(\Theta'_n(\boldsymbol{h}_n))_{n\in\mathbb{N}}$ is bounded and $\exists N_0$ s.t. (i) $\inf_{\boldsymbol{x}\in C} \Theta_n(\boldsymbol{x}) = 0$, $\forall n \geq N_0$ and (ii) $\Omega := \bigcap_{n\geq N_0} \Omega_n \neq \emptyset$. Then, we have

$$\lim_{n\to\infty} \Theta_n(\boldsymbol{h}_n) = 0.$$

Note that Θ'_n used to derive Algorithm 1 (or Algorithm 2) in Sec. 3 is automatically bounded [17].

Appendix B: MMSE and MOE Detectors

Let us show a simple observation.

Observation 1: Suppose (I) the auto-correlation matrix $\boldsymbol{R_r} := E\{\boldsymbol{r}[i]\boldsymbol{r}[i]^T\}$ is full rank ($\Rightarrow \boldsymbol{h}^*$ is unique), and (II) $\boldsymbol{y_{rb}} := E\{\boldsymbol{r}[i]b_1[i]\} = \beta \boldsymbol{s}_1$, $\exists \beta \in \mathbb{R}$. Then, for any given $\alpha \in \mathbb{R}$,

$$\boldsymbol{h}^* = \frac{\boldsymbol{R_r}^{-1}\boldsymbol{s}_1}{\boldsymbol{s}_1^T\boldsymbol{R_r}^{-1}\boldsymbol{s}_1} = \underbrace{\operatorname*{argmin}_{\boldsymbol{h}\in C_s} E\left\{(\langle \boldsymbol{h}, \boldsymbol{r}[i]\rangle - \alpha b_1[i])^2\right\}}_{=:\, \boldsymbol{h}^*_\alpha}. \quad \text{(B.1)}$$

Sketch of proof:
By the condition (I) and "Lagrangian multiplier" methodology (e.g., [31]), we can easily obtain

$$\boldsymbol{h}^*_\alpha = \frac{\boldsymbol{R_r}^{-1}\boldsymbol{s}_1}{\boldsymbol{s}_1^T\boldsymbol{R_r}^{-1}\boldsymbol{s}_1} + \alpha\left[\boldsymbol{R_r}^{-1}\boldsymbol{y_{rb}} - \frac{\boldsymbol{s}_1^T\boldsymbol{R_r}^{-1}\boldsymbol{y_{rb}}}{\boldsymbol{s}_1^T\boldsymbol{R_r}^{-1}\boldsymbol{s}_1}\boldsymbol{R_r}^{-1}\boldsymbol{s}_1\right],$$

and, by imposing the condition (II), we readily verify

$$\boldsymbol{h}^*_\alpha = \boldsymbol{h}^* = \frac{\boldsymbol{R_r}^{-1}\boldsymbol{s}_1}{\boldsymbol{s}_1^T\boldsymbol{R_r}^{-1}\boldsymbol{s}_1}.$$

□

Remark B.1:

(a) Without (I), $\boldsymbol{h}^*_\alpha$ is not necessarily unique.
(b) The condition (II) holds under slow time-varying fading situations with the following assumption: $E\{b_1[i]b_l[i]\} = 0$, $\forall l \in \{2, 3, \cdots, L\}$, and $E\{b_1[i]\boldsymbol{n}[i]\} = \boldsymbol{0}$.
(c) The filters $\boldsymbol{h}^*_0$ and $\boldsymbol{h}^*_{A_1} (= \boldsymbol{h}^*)$ are called the MOE detector and the (constrained) MMSE detector, respectively. Observation 1 shows that the MMSE and the MOE detectors coincide under (I) and (II).
(d) By $\boldsymbol{h}^* = \boldsymbol{h}^*_\alpha$ ($\forall \alpha \in \mathbb{R}$) under (I) and (II), a natural question would be: Does the set

$$\widetilde{C}_\rho^{(n)}[i] := \left\{\boldsymbol{h} \in \mathbb{R}^N : \left(\langle \boldsymbol{h}, \boldsymbol{r}[i]\rangle - \alpha\widehat{b}_{1,n}[i]\right)^2 \leq \rho\right\}$$

with an arbitrarily chosen α contain the optimal filter $\boldsymbol{h}^*$? If "yes", we could get an optimistic conclusion that the amplitude estimation $\widehat{A}_{1,n+1}$ is not necessary. Unfortunately, however, the answer is "no", of which the reason is as follows. By (4), $\langle \boldsymbol{h}, \boldsymbol{r}[i]\rangle - \alpha\widehat{b}_{1,n}[i]$ has the term $A_1 b_1[i] - \alpha\widehat{b}_{1,n}[i]$ in addition to the terms of MAI and noise. Hence, bounding $\left(\langle \boldsymbol{h}, \boldsymbol{r}[i]\rangle - \alpha\widehat{b}_{1,n}[i]\right)^2$ by small ρ does not necessarily suppress MAI sufficiently (without amplifying noise severely) when $|A_1 - \alpha| \gg 0$, which implies, from the context between (4) and (5), that α should be close to A_1 in order to ensure $\boldsymbol{h}^* \in \widetilde{C}_\rho^{(n)}[i]$. Therefore, high accuracy of the estimation of A_1 is essential for good steady state performance.

Appendix C: Proof of Equation (12)

Suppose $\boldsymbol{h} \in C_s$. For notational simplicity, in this section, we represent the stochastic property set $C_\rho^{(n)}[\iota]$ as C [see (9)]. The set C is a closed convex set bounded by two hyperplanes

$$H_+ := \{\boldsymbol{x} \in \mathbb{R}^N : \langle \boldsymbol{x}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] = \sqrt{\rho}\},$$
$$H_- := \{\boldsymbol{x} \in \mathbb{R}^N : \langle \boldsymbol{x}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] = -\sqrt{\rho}\}.$$

(a) Assume $-\sqrt{\rho} \leq \langle \boldsymbol{h}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] \leq \sqrt{\rho}$ ($\Leftrightarrow \boldsymbol{h} \in C$). In this case,

$$P_{C\cap C_s}(\boldsymbol{h}) = \boldsymbol{h}.$$

In the other cases, $P_{C\cap C_s}(\boldsymbol{h}) = P_{H_{\mathrm{sgn}}\cap C_s}(\boldsymbol{h})$, where H_{sgn} (sgn: $+$ or $-$) is the nearest hyperplane, from $\boldsymbol{h}$, of the two H_+ and H_-.

(b) Assume $\langle \boldsymbol{h}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] > \sqrt{\rho}$ ($\Rightarrow \boldsymbol{h} \notin C$). In this case, the nearest hyperplane is obviously H_+, and hence $P_{C\cap C_s}(\boldsymbol{h}) = P_{H_+\cap C_s}(\boldsymbol{h})$. Since

$$H_+ \cap C_s = \{\boldsymbol{x} : \boldsymbol{x}^T[\boldsymbol{s}_1, \boldsymbol{r}[\iota]] = \left[1, \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] + \sqrt{\rho}\right]\},$$

we have (cf. e.g., [32, p.65 Theorem 2])

$$P_{H_+\cap C_s}(\boldsymbol{h}) = \boldsymbol{h} - \boldsymbol{G}(\boldsymbol{G}^T\boldsymbol{G})^{-1}(\boldsymbol{G}^T\boldsymbol{h} - \boldsymbol{v}),$$

where $\boldsymbol{G} := [\boldsymbol{r}[\iota], \boldsymbol{s}_1]$ and $\boldsymbol{v} := \begin{bmatrix} \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] + \sqrt{\rho} \\ 1 \end{bmatrix}$.

Using $\langle \boldsymbol{s}_1, \boldsymbol{h}\rangle = 1$, $\|\boldsymbol{s}_1\| = 1$ and $\boldsymbol{I} - \boldsymbol{s}_1\boldsymbol{s}_1^T = \boldsymbol{Q}$ (see Requirements in Algorithm 1), we obtain

$$P_{C\cap C_s}(\boldsymbol{h}) = \boldsymbol{h} - \frac{\langle \boldsymbol{h}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] - \sqrt{\rho}}{\boldsymbol{r}[\iota]^T\boldsymbol{Q}\boldsymbol{r}[\iota]}\boldsymbol{Q}\boldsymbol{r}[\iota].$$

(c) Assume $\langle \boldsymbol{h}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] < -\sqrt{\rho}$ ($\Rightarrow \boldsymbol{h} \notin C$). In this case, the nearest hyperplane is obviously H_-, and hence $P_{C\cap C_s}(\boldsymbol{h}) = P_{H_-\cap C_s}(\boldsymbol{h})$. In analogy with (b), we can verify

$$P_{C\cap C_s}(\boldsymbol{h}) = \boldsymbol{h} - \frac{\langle \boldsymbol{h}, \boldsymbol{r}[\iota]\rangle - \widehat{A}_{1,n+1}\widehat{b}_{1,n}[\iota] + \sqrt{\rho}}{\boldsymbol{r}[\iota]^T \boldsymbol{Q}\boldsymbol{r}[\iota]}\boldsymbol{Q}\boldsymbol{r}[\iota],$$

which completes the proof. □

Appendix D: New Family of Embedded Constraint Algorithms

Let us consider the following problem.

Problem 1: Suppose q sets $\{S_\iota(n)\}_{\iota=1}^q \subset \mathbb{R}^N$ are defined for each $n \in \mathbb{N}$. Find a sequence $(\boldsymbol{h}_n)_{n\in\mathbb{N}} \subset \mathbb{R}^N$ that asymptotically minimizes the distance to $(\{S_\iota(n)\}_{\iota=1}^q)_{n\in\mathbb{N}}$ over a linear variety† V.

Setting $V = C_s$ and $S_\iota(n) = C_\rho^{(n)}[\iota]$, $\forall n \in \mathbb{N}$, $\forall\iota \in \mathcal{I}_n (:= \{1, 2, \cdots, q\})$, Problem 1 is reduced to the one in Sec. 3. Conversely, using V and $S_\iota(n)$ instead of C_s and $C_\rho^{(n)}[\iota]$, $\forall n \in \mathbb{N}$, $\forall\iota \in \mathcal{I}_n$, in Algorithm 1, respectively, we obtain the following scheme to solve Problem 1.

Scheme 2: (Adaptive Parallel Constrained Projection [A-PCP] Method) Generate a sequence $(\boldsymbol{h}_n)_{n\in\mathbb{N}}$ by

$$\boldsymbol{h}_{n+1} = \boldsymbol{h}_n + \lambda_n \mathcal{M}_n \left(\sum_{\iota=1}^q \omega_\iota^{(n)} P_{S_\iota(n)\cap V}(\boldsymbol{h}_n) - \boldsymbol{h}_n \right),$$

$\forall n \in \mathbb{N}$, where $\boldsymbol{h}_0 \in V$, $\mu_n \in [0, 2]$ and

$$\mathcal{M}_n := \begin{cases} \dfrac{\sum_{\iota=1}^q \omega_\iota^{(n)} \left\| P_{S_\iota(n)\cap V}(\boldsymbol{h}_n) - \boldsymbol{h}_n \right\|^2}{\left\| \sum_{\iota=1}^q \omega_\iota^{(n)} P_{S_\iota(n)\cap V}(\boldsymbol{h}_n) - \boldsymbol{h}_n \right\|^2}, \\ \qquad \text{if } \boldsymbol{h}_n \notin \bigcap_{\iota=1}^q S_\iota(n) \cap V, \\ 1, \qquad \text{otherwise.} \end{cases}$$

If, in Scheme 2, the projection onto $S_\iota(n) \cap V$ is computationally expensive, an outer approximating closed half-space can be used instead of $S_\iota(n)$ as in the *adaptive parallel subgradient projection algorithm* (see [33]). When $S_\iota(n)$ ($\forall n \in \mathbb{N}$) is a hyperplane, the choice of $q = 1$ in Scheme 2 derives the CNLMS algorithm [14].

Acknowledgement

The authors would like to express their deep gratitude to Prof. K. Sakaniwa of Tokyo Institute of Technology for fruitful discussions. This work was supported in part by JSPS grants-in-Aid (178440).

†Given $\boldsymbol{v} \in \mathbb{R}^N$ and a closed subspace $M \subset \mathbb{R}^N$, the translation of M by $\boldsymbol{v}$ defines the *linear variety* $V := \boldsymbol{v} + M := \{\boldsymbol{v} + \boldsymbol{m} : \boldsymbol{m} \in M\}$. Suppose $\dim(M^\perp) = 1$, where $M^\perp = \{\boldsymbol{x} \in \mathbb{R}^N : \langle \boldsymbol{x}, \boldsymbol{m}\rangle = 0, \forall \boldsymbol{m} \in M\}$. Then, V is called *hyperplane*, which can be expressed as $V = \{\boldsymbol{x} \in \mathbb{R}^N : \langle \boldsymbol{a}, \boldsymbol{x}\rangle = c\}$ for some $(\boldsymbol{0} \neq)\boldsymbol{a} \in \mathbb{R}^N$ and $c \in \mathbb{R}$.

References

[1] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 32, no. 1, pp. 85–96, Jan. 1986.
[2] R. Lupas and S. Verdu, "Linear multiuser detectors in synchronous code-decision multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 35, no. 1, pp. 123–136, Jan. 1989.
[3] R. Lupas and S. Verdu, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Commun.*, vol. 38, no. 4, pp. 496–508, Apr. 1990.
[4] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3178–3188, Dec. 1994.
[5] S. L. Miller, "An adaptive direct-sequence code division multiple-access receiver for multiuser interference rejection," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 1746–1754, 1995.
[6] M. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, no. 4, pp. 944–960, July 1995.
[7] S. C. Park and J. F. Doherty, "Generalized projection algorithm for blind interference suppression in DS/CDMA communications," *IEEE Trans. Circuits and Systems–II*, vol. 44, no. 6, pp. 453–460, June 1997.
[8] H. V. Poor and X. Wang, "Code-aided interference suppression for DS/CDMA communications —part II: Parallel blind adaptive implementations," *IEEE Trans. Commun.*, vol. 45, no. 9, pp. 1112–1122, Sep. 1997.
[9] X. Wang and H. V. Poor, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 677–690, Mar. 1998.
[10] M. Honig and M. K. Tsatsanis, "Adaptive techniques for multiuser CDMA receivers," *IEEE Signal Processing Magazine*, vol. 17, pp. 49–61, May 2000.
[11] H. Sakai and T. Sakaguchi, "Performance analysis of a new adaptive algorithm for blind multiuser detection," in *Proc. EUSIPCO*, Sept. 2002, vol. I, pp. 379–382, Toulouse.
[12] H. Stark and Y. Yang, *Vector Space Projections – A Numerical Approach to Signal and Image Processing, Neural Nets and Optics*, New York: Wiley, 1998.
[13] I. Oppermann, P. van Rooyen, and R. Kohno, "Guest editorial spread spectrum for global communication II," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 1, pp. 1–5, Jan. 2000.
[14] J. A. Apolinário Jr. and S. Werner, P. S. R. Diniz, and T. I. Laakso, "Constrained normalized adaptive filters for CDMA mobile communications," in *Proc. EUSIPCO*, 1998, vol. IV, pp. 2053–2056.
[15] R. L. G. Cavalcante, I. Yamada, and K. Sakaniwa, "A fast blind MAI reduction based on adaptive projected subgradient method," *IEICE Trans. Fundamentals*, vol. E87-A, no. 8, pp. 1973–1980, Aug. 2004.
[16] I. Yamada, "Adaptive projected subgradient method – a unified view for projection based adaptive algorithms," *Journal of IEICE*, vol. 86, no. 8, pp. 654–658, 2003, in Japanese.
[17] I. Yamada and N. Ogura, "Adaptive projected subgradient method and its applications to set theoric adaptive filtering," in *Proc. Thirty Seventh Asilomar Conf. Signals, Syst., Comput.*, Nov. 2003, pp. 600–606.
[18] I. Yamada and N. Ogura, "Adaptive projected subgradient method for asymptotic minimization of sequence of nonnegative convex functions," *Numer. Funct. Anal. Optim.*, vol. 25, no. 7&8, pp. 593–617, 2004.

[19] M. Yukawa and I. Yamada, "Efficient adaptive stereo echo canceling schemes based on simultaneous use of multiple state data," *IEICE Trans. Fundamentals*, vol. E87-A, no. 8, pp. 1949–1957, Aug. 2004.

[20] M. Yukawa, N. Murakoshi, and I. Yamada, "An efficient fast stereo echo canceler by pairwise optimal weight realization technique," in *Proc. EUSIPCO*, Sept. 2005, to be presented.

[21] M. Yukawa, R. L. G. Cavalcante, and I. Yamada, "Efficient blind DS/CDMA receivers by embedded constraint adaptive parallel projection techniques," in *Proc. IEICE SIP Symposium*, Nov. 2004, in CD-ROM.

[22] M. Yukawa, R. L. G. Cavalcante, and I. Yamada, "Efficient adaptive blind MAI suppression in DS/CDMA by embedded constraint parallel projection techniques," in *Proc. IEEE ICASSP*, Philadelphia, PA, Mar. 2005, pp. 929–932.

[23] Y. Censor and S. A. Zenios, *Parallel optimization: Theory, algorithm, and optimization*, Oxford University Press, 1997.

[24] P. L. Combettes, "The foundations of set theoretic estimation," *Proc. IEEE*, vol. 81, no. 2, pp. 182–208, 1993.

[25] M. L. R. de Campos, S. Werner, and J. A. Apolinário Jr., "Constrained adaptive algorithms employing householder transformation," *IEEE Trans. Signal Processing*, vol. 50, no. 9, pp. 2187–2195, Sept. 2002.

[26] M. Yukawa and I. Yamada, "A note on inflation parameter for adaptive parallel subgradient projection algorithms — An optimal design in case of single projection," Tech. Rep., IEICE, Jan. 2005, SIP2004-104.

[27] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 3rd edition, 1996.

[28] M. Yukawa and I. Yamada, "Acceleration of adaptive parallel projection algorithms by pairwise optimal weight realization," in *Proc. EUSIPCO*, Vienna, Austria, Sept. 2004, pp. 713–716.

[29] M. Yukawa and I. Yamada, "Acceleration technique for adaptive parallel subgradient projection algorithm and its properties," in *Proc. IEICE SIP Symposium*, Nov. 2004, in CD-ROM.

[30] R. L. G. Cavalcante, M. Yukawa, and I. Yamada, "An efficient receiver for multipath BPSK-DS/CDMA systems by adaptive projected subgradient method," in *Proc. IEICE SIP Symposium*, Nov. 2004, in CD-ROM.

[31] J. B. Hiriart-Urruty and C. Lemaréchal, *Convex analysis and minimization algorithms I —Fundamentals*, Springer-Verlag, Germany, 1993.

[32] D. G. Luenberger, *Optimization by Vector Space Methods*, John Wiley & Sons, New York, 1969.

[33] I. Yamada, K. Slavakis, and K. Yamada, "An efficient robust adaptive filtering algorithm based on parallel subgradient projection techniques," *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 1091–1101, May 2002.

Masahiro Yukawa received the B.E. and M.E. degrees from Tokyo Institute of Technology in 2002 and 2004, respectively. Since 2005, he has been a Research Fellow of the Japan Society for the Promotion of Science (JSPS). He is currently pursuing the Ph.D. degree in Department of Communications and Integrated Systems at Tokyo Institute of Technology. His research interests include Mathematical Adaptive Signal Processing with Applications to Acoustics/Communications/Adaptive Beamforming. He is a student member of IEEE/IEICE.

Renato L. G. Cavalcante was born in Rio de Janeiro, Brazil, in 1979. He received the B.Sc. degree in Electronics Engineering from Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos, Brazil, in 2002. Since 2003, he has been a recipient of the Japanese Government Ministry of Education, Science and Culture (Monbusho) Scholarship. He is currently pursuing the M.E. degree in electrical and electronic engineering at the Tokyo Institute of Technology. His research interests are in Digital Communication and Digital Signal Processing.

Isao Yamada received the B.E. degree in computer science in 1985 from University of Tsukuba, and the M.E. and Ph.D.degrees in Electrical and Electronic Engineering from Tokyo Institue of Technology, in 1987 and 1990, respectively. In 1990, he joined the Department of Electrical and Electronic Engineering at Tokyo Institute of Technology, as a research associate. Currently he is an associate professor in the Department of Communications and Integrated Systems at Tokyo Institute of Technology. He received the Excellent Paper Awards, in 1990 and 1994, and the Young Researcher Award, in 1992, from IEICE and the ICF Research Award, in 2004. His current research interests are in Mathematical/Multidimensional/Statistical/Adaptive/Array Signal Processing, Image Processing, Optimization Theory, Inverse Problem. He is a member of IEEE, AMS, SIAM, JSIAM, SITA and IEICE. He serves as an Associate Editor for The International Journal on Multidimensional Systems and Signal Processing (Springer).