STEREO ECHO CANCELER BY ADAPTIVE PROJECTED SUBGRADIENT METHOD WITH MULTIPLE ROOM-ACOUSTICS INFORMATION

Masahiro Yukawa *Konstantinos Slavakis* *Isao Yamada*

Department of Communications and Integrated Systems, Tokyo Institute of Technology
Ookayama, Meguro-ku, Tokyo 152-8552, JAPAN.
E-mails: {masahiro, slavakis, isao}@comm.ss.titech.ac.jp

ABSTRACT

In this paper, we propose a novel set-theoretic stereophonic echo canceling algorithm that realizes fast convergence by utilizing multiple constraints based on a priori room-acoustics information. It is reported that the squared impulse response in a typical room decays exponentially on average. Based on this a priori information, we present two simple examples of constraint sets bounding the adaptive filter coefficients. In addition to the constraint sets, we newly introduce different sets, based on the same a priori information, to raise speed of initial convergence. The numerical examples demonstrate that the proposed algorithm reduces the time, to drop by 20dB in system mismatch, by 19[sec.] (11[sec.]) compared with the NLMS (the APA). Also in ERLE, the proposed algorithm dramatically outperforms the conventional algorithms.

1. INTRODUCTION

Constrained adaptive filtering with a priori information has been proven to be effective in signal processing applications such as adaptive beamforming [1, Chapter6]. On the other hand, in the study of *room acoustics*, it is shown that squared impulse response in a typical room decays exponentially on average under some conditions [2], which has been widely used in acoustics; e.g., sound impulse response extension for simulating auditory impression of a virtual environment [3], reverberation time estimation for speech recognition [4], etc. This a priori information is naturally expected to be utilized for stereophonic acoustic echo cancellation (SAEC) problem. This paper presents a set-theoretic SAEC algorithm that efficiently incorporates this information, for raising the convergence speed of adaptive filter, into multiple constraints.

For realistic and high quality acoustic telecommunication, a stereophonic (generally multi-channel) hands-free and full-duplex system is a key technology; see e.g., [1, Chapter4]. One of the most important and challenging issues is to suppress (or cancel if possible) "acoustic echo" from 2 loudspeakers to 2 microphones, and thus stereo echo cancelers must be installed. Motivated by the finding of an intrinsic problem so-called *non-uniqueness problem* [5], a great deal of effort has been devoted to resolve this problem; e.g., [5, 6]. Simply saying, a system of linear equations, to be solved for minimizing the residual echo within a time interval in which the transmission paths are almost constant, has infinitely many solutions depending on the transmission paths. To prevent echo relapses when the transmission paths change, it is strongly desired to estimate the true echo impulse response as "fast" and "accurately" as possible. The major interest in SAEC is to develop an efficient adaptive algorithm to realize the "fast" and "accurate" estimation with $O(N)$ complexity [7] (Note that demand for low complexity is also severe since the filter length N should typically be a few thousand for sufficient echo cancellation).

An efficient SAEC algorithm of $O(N)$ complexity is proposed [8], which is based on simultaneous use of multiple state input-output data by utilizing the adaptive Parallel Subgradient Projection (adaptive PSP) techniques [9]. The algorithm is derived from the Adaptive Projected Subgradient Method (APSM) [10], which generates a strongly convergent point sequence that asymptotically minimizes a certain sequence of nonnegative continuous convex functions over a convex constraint set; in SAEC, e.g., the sequence of functions is defined as the distance to the set of filters that temporarily minimize the error at each time. Recently, in [11], the original APSM [10] was extended from a single convex projector to an η-attracting nonexpansive mapping, to which a concatenation of convex projectors belongs (see Sec. 2-B). Thus, this extension provides great benefit, i.e., the use of "multiple constraints".

In this paper, we propose a powerful set-theoretic SAEC algorithm based on the APSM with multiple constraints. Based on the aforementioned a priori information, we present two examples of constraint sets that bound the adaptive filter coefficients (Sec. 3-A). In addition to the constraint sets, we newly introduce, also based on the same a priori information, different sets along which the adaptive filter quickly approaches the true impulse response especially in the initial phase (see Sec. 3-B). Thanks to the simple structure of those sets, the proposed algorithm causes no serious increase in computational complexity compared with the method in [8] (see Remark 1 in Sec. 3-B). The simulation results demonstrate that the proposed algorithm dramatically outperforms the method in [8] as well as the Normalized Least Mean Square (NLMS) algorithm and the Affine Projection Algorithm (APA) [12] both in system mismatch and in Echo Return Loss Enhancement (ERLE).

2. PRELIMINARIES

Following the problem formulation for SAEC, the APSM with multiple convex constraints [11] is briefly introduced.

A. Stereo echo canceling problem

Without loss of generality, we concentrate on the microphone B1 in the receiving room (Room B); see Fig. 1. The signals are modeled as follows ($k \in \mathbb{N}$: time index, superscript T: transposition):

- talker's voice signal: $\boldsymbol{s}_k \in \mathbb{R}^L$ ($L \in \mathbb{N}^* := \mathbb{N} \setminus \{0\}$)
- i-th transmission path: $\boldsymbol{\theta}_{(i)} \in \mathbb{R}^L$ ($i = 1, 2$)
- signal at mic. (microphone) Ai: $u_k^{(i)} := \boldsymbol{s}_k^T \boldsymbol{\theta}_{(i)} \in \mathbb{R}$ ($i = 1, 2$)
- vector of $u_k^{(i)}$: $\boldsymbol{u}_k^{(i)} := [u_k^{(i)}, \cdots, u_{k-N+1}^{(i)}]^T \in \mathbb{R}^N$ ($N \in \mathbb{N}^*$)
- signal after Unit 1: $\widetilde{\boldsymbol{u}}_k^{(1)} \in \mathbb{R}^N$
- input vector to Room B: $\boldsymbol{u}_k := \begin{bmatrix} \widetilde{\boldsymbol{u}}_k^{(1)} \\ \boldsymbol{u}_k^{(2)} \end{bmatrix} \in \mathcal{H} := \mathbb{R}^{2N}$
- input matrix: $\boldsymbol{U}_k := [\boldsymbol{u}_k, \cdots, \boldsymbol{u}_{k-r+1}] \in \mathbb{R}^{2N \times r}$ ($r \in \mathbb{N}^*$)
- i-th echo path: $\boldsymbol{h}_{(i)}^* \in \mathbb{R}^N$ ($i = 1, 2$)
- estimandum (system to be estimated): $\boldsymbol{h}^* := [\boldsymbol{h}_{(1)}^{*\,T}, \boldsymbol{h}_{(2)}^{*\,T}]^T \in \mathcal{H}$
- adaptive filter (echo canceler): $\boldsymbol{h}_k := [\boldsymbol{h}_k^{(1)T}, \boldsymbol{h}_k^{(2)T}]^T \in \mathcal{H}$
- additive noise at mic. B1: $\boldsymbol{n}_k := [n_k, \cdots, n_{k-r+1}]^T \in \mathbb{R}^r$
- output (observed signal at mic. B1): $\boldsymbol{d}_k := \boldsymbol{U}_k^T \boldsymbol{h}^* + \boldsymbol{n}_k \in \mathbb{R}^r$
- residual error function: $\boldsymbol{e}_k(\boldsymbol{h}) := \boldsymbol{U}_k^T \boldsymbol{h} - \boldsymbol{d}_k \in \mathbb{R}^r$

Here, $\mathcal{H}(:= \mathbb{R}^{2N})$ is a real Hilbert space equipped with the inner product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^T \boldsymbol{y}$, $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{H}$, and its induced norm

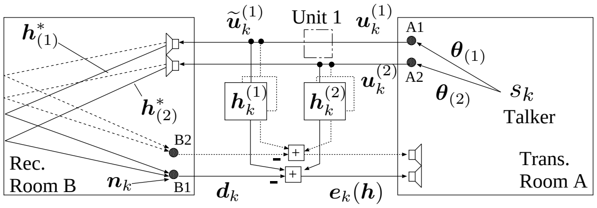


Fig. 1. Stereophonic acoustic echo cancelers. Unit 1 is a preprocessing module that gives a delay to $\boldsymbol{u}_k^{(1)}$ with the cycle period $Q \in \mathbb{N}^*$.

$\|\boldsymbol{x}\| := \left(\boldsymbol{x}^T\boldsymbol{x}\right)^{1/2}$, $\forall \boldsymbol{x} \in \mathcal{H}$. The common notation "**0**" is used to denote both "zero vector" and "zero matrix".

The goal of the SAEC problem is to constantly cancel the echo; i.e., $\langle \boldsymbol{u}_k, \boldsymbol{h}^* \rangle \approx \langle \boldsymbol{u}_k, \boldsymbol{h}_k \rangle, \forall k \in \mathbb{N}$. Since only $(\boldsymbol{u}_k, \boldsymbol{d}_k)_{k\in\mathbb{N}}$ are observable, a common alternative goal is to suppress the residual error; i.e., $\boldsymbol{e}_k(\boldsymbol{h}_k) \approx \boldsymbol{0}$, $\forall k \in \mathbb{N}$. Due to high correlation between two signals $u_k^{(1)}$ and $u_k^{(2)}$, this problem has infinitely many solutions depending on $\boldsymbol{\theta}_{(1)}$ and $\boldsymbol{\theta}_{(2)}$, which is the so-called non-uniqueness problem [5–7]. Without well-approximating $\boldsymbol{h}^*$, echo relapses by change of the transmission paths $\boldsymbol{\theta}_{(1)}$ and $\boldsymbol{\theta}_{(2)}$. Hence, it is strongly desired to keep $\boldsymbol{h}_k$ close to $\boldsymbol{h}^*$.

B. Adaptive projected subgradient method with multiple convex constraints

For any nonempty closed convex[1] set $C \subset \mathcal{H}$, the *projection operator* $P_C : \mathcal{H} \to C$ maps a vector $\boldsymbol{x} \in \mathcal{H}$ to the unique vector $P_C(\boldsymbol{x}) \in C$ s.t. $\mathrm{d}(\boldsymbol{x}, C) := \|\boldsymbol{x} - P_C(\boldsymbol{x})\| = \min_{\boldsymbol{y}\in C} \|\boldsymbol{x} - \boldsymbol{y}\|$. Let $K_1, K_2, \cdots, K_m \subset \mathcal{H}$ be m closed convex sets s.t. $K := \bigcap_{j=1}^m K_j \neq \emptyset$. Also let $\Theta_k : \mathcal{H} \to [0, \infty)$, $k \in \mathbb{N}$, be a continuous *convex*[2] function and $\partial\Theta(\boldsymbol{y})$ the *subdifferential*[3] of Θ at $\boldsymbol{y}$. The following scheme, an extension of the scheme in [10, Theorem 2], provides a vector sequence that minimizes asymptotically the sequence of functions $(\Theta_k)_{k\in\mathbb{N}}$ over K (see the appendix).

Scheme 1 *(Adaptive Projected Subgradient Method with Multiple Convex Constraints* [11]*) For an arbitrary $\boldsymbol{h}_0 \in \mathcal{H}$, generate a sequence $(\boldsymbol{h}_k)_{k\in\mathbb{N}} \subset \mathcal{H}$ by*

$$\boldsymbol{h}_{k+1} := \begin{cases} P_{K_m} \cdots P_{K_2} P_{K_1} \left(\boldsymbol{h}_k - \lambda_k \frac{\Theta_k(\boldsymbol{h}_k)}{\|\Theta_k'(\boldsymbol{h}_k)\|^2} \Theta_k'(\boldsymbol{h}_k) \right), \\ \qquad \text{if } \Theta_k'(\boldsymbol{h}_k) \neq \boldsymbol{0}, \\ P_{K_m} \cdots P_{K_2} P_{K_1} (\boldsymbol{h}_k), \quad \text{otherwise,} \end{cases}$$

where $\Theta_k'(\boldsymbol{h}_k) \in \partial\Theta_k(\boldsymbol{h}_k)$ and $\lambda_k \in [0, 2]$, $\forall k \in \mathbb{N}$.

In [11, Theorem 1], Scheme 1 is presented for a more general mapping called "η-attracting nonexpansive" instead of the concatenation of projectors $P_{K_m} \cdots P_{K_2} P_{K_1}$ (The mapping $P_{K_m} \cdots P_{K_2} P_{K_1}$ is $\frac{1}{m}$-attracting nonexpansive [11, Lemma 2]).

3. STEREO ECHO CANCELER BASED ON APSM WITH MULTIPLE ENERGY CONSTRAINTS

In this section, we present a new set-theoretic stereo echo canceling algorithm based on the APSM with multiple convex constraints (Scheme 1). Note that the convex constraint sets, $K_1, \cdots, K_m$, must be designed simply enough to compute the associated metric projections in real time, although the sets can be designed based on large variety of a priori information about echo paths. Simple examples of constraint sets are given in Sec. 3.A, and the proposed algorithm is then presented in Sec. 3.B.

A. Design of Constraint Sets and Stochastic Property Sets

Let us consider first the *monaural* case for simplicity. It is known that, under the diffuse sound field assumption, the ensemble average E_n of the squared room impulse responses decays exponentially (see [4]); i.e.,

$$E_n := E\{h_{(n)}^2\} = E_1 e^{-\delta n}, \forall n \in \{1, 2, \cdots, N\}, \qquad (1)$$

where $E\{\cdot\}$ denotes expectation, $\{h_{(n)}\}_{n=1}^N$ is a causal room impulse response, and $\delta := \frac{\log 10^6}{T_{60} F_s}$; F_s: the sampling frequency, T_{60}: the time interval in which the reverberant sound energy drops down by 60 dB [2]. For the estimation of T_{60}, a blind algorithm, which does not require the identification of room impulse response unlike the classical techniques, is proposed [4]. (1) implies that $\sum_{n=1}^{N/2} h_{(n)}^2$ is expected to be much greater than $\sum_{n=N/2+1}^{N} h_{(n)}^2$.

Now, return to the *stereophonic* case. Recall that the estimandum is $\boldsymbol{h}^* := [\boldsymbol{h}_{(1)}^{*\,T}, \boldsymbol{h}_{(2)}^{*\,T}]^T \in \mathcal{H}(:= \mathbb{R}^{2N})$. From the above discussion, it is highly expected $\left\|\boldsymbol{h}_{(i),e}^*\right\|^2 \gg \left\|\boldsymbol{h}_{(i),l}^*\right\|^2$ ($\forall i = 1, 2$), where $\boldsymbol{h}_{(i),e}^*, \boldsymbol{h}_{(i),l}^* \in \mathbb{R}^{N/2}$ satisfy $\boldsymbol{h}_{(i)}^* = [\boldsymbol{h}_{(i),e}^{*\,T}, \boldsymbol{h}_{(i),l}^{*\,T}]^T$. Based on this observation, we propose the following constraint set.

Example 1 *(Energy Constraint Sets)*

$$B_{(i)} := \left\{ \boldsymbol{h} \in \mathcal{H} : \boldsymbol{h}^T \boldsymbol{D}_i \boldsymbol{h} \leq \varepsilon_i \right\}, \ \forall i = 1, 2, 3, 4, \qquad (2)$$

where, $\forall i = 1, 2, 3, 4$, $\varepsilon_i \geq 0$ is the energy bound and

$$\boldsymbol{D}_i := \begin{bmatrix} \boldsymbol{0}_{\frac{N(i-1)}{2}} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{\frac{N}{2}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}_{\frac{N(4-i)}{2}} \end{bmatrix} \in \mathbb{R}^{2N \times 2N}.$$

Here, $\boldsymbol{0}_a$ and $\boldsymbol{I}_a$ ($a \in \mathbb{N}$) denote $a \times a$ zero and identity matrices, respectively ($\boldsymbol{0}$ denotes a zero matrix of an appropriate size).

To ensure the membership $\boldsymbol{h}^* \in B_{(i)}$, ε_1 and ε_3 should be much greater than ε_2 and ε_4. The developments in *room acoustics* provide a priori knowledge about statistics of the impulse response $\boldsymbol{h}^*$ to determine ε_i [2]. The definition of $\boldsymbol{D}_i$ is just for simplicity, and it can be devised; e.g., change the size of $\boldsymbol{I}$, or make the diagonal elements of $\boldsymbol{I}$ decay exponentially etc.

Example 2 $B := \{\boldsymbol{h} \in \mathcal{H} : |h_n| \leq \varpi_n, \forall n = 1, 2, \cdots, 2N\}$ *with the upper bound $(\varpi_n)_{n=1}^{2N}$. In this case, the original APSM* [10] *can be used instead of Scheme 1.*

Next is the definition of *stochastic property sets*. Because of corrupted noise, the linear variety $V_k := \{\boldsymbol{h} \in \mathcal{H} : \|\boldsymbol{e}_k(\boldsymbol{h})\|^2 = 0\}$, which is usually employed in the APA (in the NLMS when $r = 1$), has no guarantee to contain the estimandum $\boldsymbol{h}^*$ (For details, see [9]). Hence, we use the following stochastic property sets:

$$C_k(\rho) := \left\{ \boldsymbol{h} \in \mathcal{H} : g_k(\boldsymbol{h}) := \|\boldsymbol{e}_k(\boldsymbol{h})\|^2 - \rho \leq 0 \right\}, \qquad (3)$$

where $\rho \geq 0$ is the *inflation parameter*, which is designed based on statistics of noise process (For details, see [9]. ρ can be varied with time.). Since the projection onto $C_k(\rho)$ requires, in general, huge computational complexity, we employ an approximating projection onto the closed half-space $H_k^-(\boldsymbol{h}) := \{\boldsymbol{x} \in \mathcal{H} :$

[1] A set $C \subset \mathcal{H}$ is said to be *convex* provided that $\forall \boldsymbol{x}, \boldsymbol{y} \in C$, $\forall \nu \in (0, 1)$, $\nu\boldsymbol{x} + (1-\nu)\boldsymbol{y} \in C$.

[2] A function $\Theta : \mathcal{H} \to \mathbb{R}$ is said to be *convex* if $\Theta(\nu\boldsymbol{x} + (1-\nu)\boldsymbol{y}) \leq \nu\Theta(\boldsymbol{x}) + (1-\nu)\Theta(\boldsymbol{y})$, $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{H}$ and $\forall \nu \in (0, 1)$.

[3] The *subdifferential* of Θ at $\boldsymbol{y}$ is the set of all the *subgradients* of Θ at $\boldsymbol{y}$; $\partial\Theta(\boldsymbol{y}) := \{\boldsymbol{a} \in \mathcal{H} : \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{a} \rangle + \Theta(\boldsymbol{y}) \leq \Theta(\boldsymbol{x}), \forall \boldsymbol{x} \in \mathcal{H}\}$.

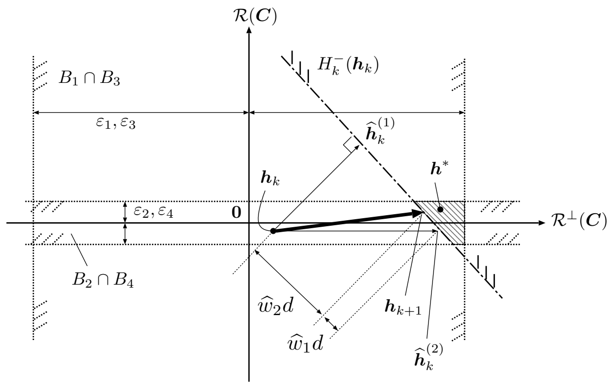


Fig. 2. A geometric interpretation of the proposed algorithm. Here, $d := \left\| \widehat{\boldsymbol{h}}_k^{(1)} - \widehat{\boldsymbol{h}}_k^{(2)} \right\|$.

$\langle \boldsymbol{x} - \boldsymbol{h}, \nabla g_k(\boldsymbol{h}) \rangle + g_k(\boldsymbol{h}) \leq 0\} \supset C_k(\rho)$, which has the following simple closed-form expression:

$$P_{H_k^-(\boldsymbol{h})}(\boldsymbol{h}) = \begin{cases} \boldsymbol{h} - \frac{g_k(\boldsymbol{h})}{\|\nabla g_k(\boldsymbol{h})\|^2} \nabla g_k(\boldsymbol{h}), & \text{if } \boldsymbol{h} \notin H_k^-(\boldsymbol{h}), \\ \boldsymbol{h}, & \text{otherwise}. \end{cases} \quad (4)$$

Here, $\nabla g_k(\boldsymbol{h}) = 2\boldsymbol{U}_k \boldsymbol{e}_k(\boldsymbol{h})$ and $P_{H_k^-(\boldsymbol{h})}(\boldsymbol{h}) \cong P_{C_k(\rho)}(\boldsymbol{h})$; see [9]. Note that $P_{H_k^-(\boldsymbol{h})}(\boldsymbol{h})$ requires only $O(N)$ complexity.

B. Proposed Stereo Echo canceling algorithm

In the following, we focus on Example 1 for simplicity. Let $Q \in \mathbb{N}^*$ denote the cycle period of preprocessing (see [6]). Given $q \in \mathbb{N}^*$, define the control sequences $\mathcal{I}_k^{(\mathrm{c})}, \mathcal{I}_k^{(\mathrm{p})} \subset \mathbb{N}$ (each of which corresponds to the current/previous state of inputs) as $\mathcal{I}_k^{(\mathrm{c})} := \{k, k-1, \ldots, k-q+1\}$ and $\mathcal{I}_k^{(\mathrm{p})} := \emptyset$, if $0 \leq k \leq Q/2$, $\mathcal{I}_k^{(\mathrm{p})} := \mathcal{I}_{k-Q/2}^{(\mathrm{c})}$, if $k > Q/2$. Define the weights $\left\{ w_\iota^{(k)} \right\}_{\iota \in \mathcal{I}_k^{(\mathrm{c})} \cup \mathcal{I}_k^{(\mathrm{p})}} \subset (0,1]$, $\forall k \in \mathbb{N}$, to satisfy $\sum_{\iota \in \mathcal{I}_k^{(\mathrm{c})} \cup \mathcal{I}_k^{(\mathrm{p})}} w_\iota^{(k)} = 1$, $\forall k \in \mathbb{N}$. Let $\mathcal{R}(\boldsymbol{A})$ and $\mathcal{R}^\perp(\boldsymbol{A})$ denote the column space of a matrix $\boldsymbol{A}$ and its orthogonal complement, respectively (i.e., $\mathcal{R}(\boldsymbol{A}) \oplus \mathcal{R}^\perp(\boldsymbol{A}) = \mathcal{H}$). Define $\mathcal{K}_\iota^{(k)} := H_\iota^-(\boldsymbol{h}_k) \cap \mathcal{V}_k$, where $\mathcal{V}_k := \boldsymbol{h}_k + \mathcal{R}^\perp(\boldsymbol{D_1} + \boldsymbol{D_3}) = \boldsymbol{h}_k + \mathcal{R}^\perp(\boldsymbol{C}) = \left\{ \boldsymbol{x} \in \mathcal{H} : \boldsymbol{C}^T \boldsymbol{x} = \boldsymbol{\alpha} \right\}$ with $\boldsymbol{\alpha} := \boldsymbol{C}^T \boldsymbol{h}_k$ and $\boldsymbol{C} := \begin{bmatrix} \mathbf{0}_{\frac{N}{2}} & \boldsymbol{I}_{\frac{N}{2}} & \mathbf{0}_{\frac{N}{2}} & \mathbf{0}_{\frac{N}{2}} \\ \mathbf{0}_{\frac{N}{2}} & \mathbf{0}_{\frac{N}{2}} & \mathbf{0}_{\frac{N}{2}} & \boldsymbol{I}_{\frac{N}{2}} \end{bmatrix}^T \in \mathbb{R}^{2N \times N}$. In the following, despite many choices, we initialize the adaptive filter as $\boldsymbol{h}_0 = \mathbf{0}$ for simplicity. In this case, the observation before (2) suggests that the adaptive filter should be updated much more in the direction along the subspace $\mathcal{R}(\boldsymbol{D}_1 + \boldsymbol{D}_3)$ $(= \mathcal{R}^\perp(\boldsymbol{C}))$ than in the direction along $\mathcal{R}(\boldsymbol{D}_2 + \boldsymbol{D}_4)$ $(= \mathcal{R}(\boldsymbol{C}))$ especially in the initial stage of adaptation; see Fig. 2. From this point of view, the proposed algorithm is given as follows.

Algorithm 1 *Given sequences of input-output data* $(\boldsymbol{u}_k)_{k \in \mathbb{N}}$ *and* $(\boldsymbol{d}_k)_{k \in \mathbb{N}}$*, a sequence* $(\boldsymbol{h}_k)_{k \in \mathbb{N}} \subset \mathcal{H}$ *is iteratively generated as*

$$\boldsymbol{h}_{k+1} := P_{B_4} P_{B_3} P_{B_2} P_{B_1} \left\{ \boldsymbol{h}_k + \lambda_k \left(\widehat{w}_1 \widehat{\boldsymbol{h}}_k^{(1)} + \widehat{w}_2 \widehat{\boldsymbol{h}}_k^{(2)} - \boldsymbol{h}_k \right) \right\}, \quad (5)$$

where $\lambda_k \in [0, 2]$, $\widehat{w}_1$ *and* $\widehat{w}_2$ *are the weights satisfying* $\widehat{w}_2 \gg \widehat{w}_1 \geq 0$ *and* $\widehat{w}_1 + \widehat{w}_2 = 1$*, and*

$$\widehat{\boldsymbol{h}}_k^{(i)} := \boldsymbol{h}_k + \mathcal{M}_k^{(i)} \left(\sum_{\iota \in \mathcal{I}_k^{(c)} \cup \mathcal{I}_k^{(p)}} w_\iota^{(k)} P_{S_{\iota,i}^{(k)}}(\boldsymbol{h}_k) - \boldsymbol{h}_k \right), \quad (6)$$

$\forall i = 1, 2$, $\forall k \in \mathbb{N}$. *Here,* $S_{\iota,1}^{(k)} := H_\iota^-(\boldsymbol{h}_k)$,

$$S_{\iota,2}^{(k)} := \begin{cases} \mathcal{K}_\iota^{(k)} & \text{if } \mathcal{K}_\iota^{(k)} \neq \emptyset \\ H_\iota^-(\boldsymbol{h}_k) & \text{otherwise}, \end{cases}$$

$$\mathcal{M}_k^{(i)} := \begin{cases} \dfrac{\sum_{\iota \in \mathcal{I}_k^{(c)} \cup \mathcal{I}_k^{(p)}} w_\iota^{(k)} \left\| P_{S_{\iota,i}^{(k)}}(\boldsymbol{h}_k) - \boldsymbol{h}_k \right\|^2}{\left\| \sum_{\iota \in \mathcal{I}_k^{(c)} \cup \mathcal{I}_k^{(p)}} w_\iota^{(k)} P_{S_{\iota,i}^{(k)}}(\boldsymbol{h}_k) - \boldsymbol{h}_k \right\|^2}, \\ \qquad \text{if } \boldsymbol{h}_k \notin \bigcap_{\iota \in \mathcal{I}_k^{(c)} \cup \mathcal{I}_k^{(p)}} S_{\iota,i}^{(k)}, \\ 1, \qquad \text{otherwise}. \end{cases}$$

Here, $P_{H_\iota^-(\boldsymbol{h}_k)}(\boldsymbol{h}_k)$ *is computed by* (4) *and, if* $\mathcal{K}_\iota^{(k)} \neq \emptyset$ *(*$\Rightarrow \boldsymbol{D}_{1,3} \nabla g_\iota(\boldsymbol{h}_k) \neq \mathbf{0}$ *if* $\boldsymbol{h}_k \notin H_\iota^-(\boldsymbol{h}_k)$*, where* $\boldsymbol{D}_{1,3} := \boldsymbol{D_1} + \boldsymbol{D_3}$*),* $P_{\mathcal{K}_\iota^{(k)}}(\boldsymbol{h}_k)$ *is computed as*

$$P_{\mathcal{K}_\iota^{(k)}}(\boldsymbol{h}_k) = \begin{cases} \boldsymbol{h}_k - \frac{g_\iota(\boldsymbol{h}_k)}{\|\boldsymbol{D}_{1,3} \nabla g_\iota(\boldsymbol{h}_k)\|^2} \boldsymbol{D}_{1,3} \nabla g_\iota(\boldsymbol{h}_k), \\ \qquad \text{if } \boldsymbol{h}_k \notin H_\iota^-(\boldsymbol{h}_k) \ (\Leftrightarrow g_\iota(\boldsymbol{h}_k) > 0), \\ \boldsymbol{h}_k, \quad \text{otherwise}. \end{cases} \quad (7)$$

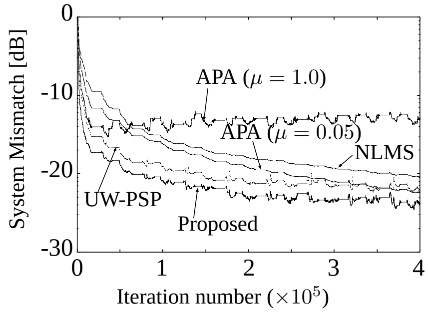
Note that (a) the weights $w_\iota^{(k)}$ can be chosen independently for $\widehat{\boldsymbol{h}}_k^{(1)}$ and $\widehat{\boldsymbol{h}}_k^{(2)}$ and (b) the update rule to generate the sequence $(\widehat{\boldsymbol{h}}_k^{(2)})_{k \in \mathbb{N}}$ belongs to the family of embedded constraint algorithms proposed in [13, Appendix D]. The proof of (7) is omitted because it is verified by simple algebra. Derivation of Algorithm 1 from Scheme 1 is also omitted due to luck of space. Instead, for better understanding, we show a geometric interpretation of the proposed algorithm in Fig.2. We set $q = 1$ and $\lambda_k = 1$ for simplicity. The figure demonstrates that $\widehat{\boldsymbol{h}}_k^{(2)}$ contributes toward letting $\boldsymbol{h}_k$ to move quickly in the direction of $\mathcal{R}^\perp(\boldsymbol{C})$ while $\widehat{\boldsymbol{h}}_k^{(1)}$ contributes toward letting $\boldsymbol{h}_k$ to move slowly in the direction of $\mathcal{R}(\boldsymbol{C})$. Based on the discussion before Algorithm 1, we select the weights for $\widehat{\boldsymbol{h}}_k^{(1)}$ and $\widehat{\boldsymbol{h}}_k^{(2)}$ so that $\widehat{w}_2 \gg \widehat{w}_1$, which is the key to accelerate the speed of initial convergence. Moreover, even if the filter moves out from $\bigcap_{i=1}^4 B_i$ after taking the average of $\widehat{\boldsymbol{h}}_k^{(1)}$ and $\widehat{\boldsymbol{h}}_k^{(2)}$, it is enforced in $\bigcap_{i=1}^4 B_i$ by $P_{B_4} P_{B_3} P_{B_2} P_{B_1}$.

A remark on complexity of Algorithm 1 is given below.

Remark 1 *Thanks to the simple structure of* $\boldsymbol{D}_{1,3}$, $\widehat{\boldsymbol{h}}_k^{(2)}$ *is obtained, like a "by-product", from the results produced in the process to compute* $\widehat{\boldsymbol{h}}_k^{(1)}$. *More precisely, comparing* (4) *with* (7)*, we see that* $P_{\mathcal{K}_\iota^{(k)}}(\boldsymbol{h}_k)$ *is computed by partly using the computation process for* $P_{H_\iota^-(\boldsymbol{h}_k)}(\boldsymbol{h}_k)$. *In addition,* $P_{B_4} P_{B_3} P_{B_2} P_{B_1}(\cdot)$ *is also easy to compute;* $O(N)$ *complexity. Therefore, from a simple inspection of* (5) *and* (6)*, we see that, by using* $2q$ *parallel processors, the computational complexity imposed on each processor is almost identical to the one to compute* $P_{H_\iota^-(\boldsymbol{h}_k)}(\boldsymbol{h}_k)$*, and thus the time to generate* $\boldsymbol{h}_{k+1}$ *from* $\boldsymbol{h}_k$ *is proportional to* $O(N)$ *computational complexity. This implies that the proposed algorithm is suitable for real time implementation, in which the time to update the filter is strictly limited.*

4. NUMERICAL EXAMPLES

In this section, the proposed algorithm is compared with the UW-PSP (Uniformly Weighted Parallel Subgradient Projection) [8], the APA and the NLMS algorithms with a common preprocessing technique proposed in [6]. Input signals in channel 1 are modified with the cycle $Q = 800$; see Sec. 3-B. The tests are performed, for estimating $\boldsymbol{h}^* \in \mathcal{H} := \mathbb{R}^{512} (N = L = 256)$, under the



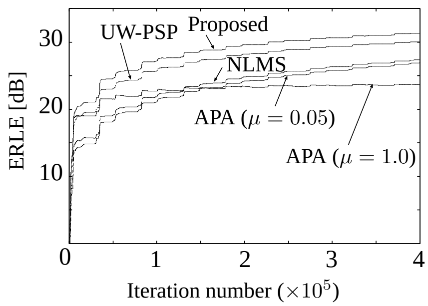


Fig. 3. The proposed algorithm versus the UW-PSP, the APA and the NLMS algorithms under SNR 25dB with common preprocessing. For the proposed and the UW-PSP algorithms, $\lambda_k = 0.4$, $r = 1$, $\rho = 0$ and $q = 10$. For the NLMS, $\mu = 0.2$.

condition of Signal to Noise Ratio (SNR) $:= 10\log_{10}(E\{z_k^2\}/E\{n_k^2\}) = 25$dB, where $z_k := \langle \boldsymbol{u}_k, \boldsymbol{h}^* \rangle$ denotes *pure* echo (echo without noise), respectively. We utilize a male's speech signal, for $(s_k)_{k\in\mathbb{N}}$, recorded at sampling rate 16kHz. To measure the achievement level for echo path identification as well as that of echo cancellation, we evaluate the following two criteria: System Mismatch $(k) := 10\log_{10} \|\boldsymbol{h}^* - \boldsymbol{h}_k\|^2 / \|\boldsymbol{h}^*\|^2$, $\forall k \in \mathbb{N}$, ERLE $(k) := 10\log_{10} \sum_{i=1}^{k} z_i^2 / \sum_{i=1}^{k} (z_i - \langle \boldsymbol{u}_i, \boldsymbol{h}_i \rangle)^2$, $\forall k \in \mathbb{N}$. For the proposed and the UW-PSP algorithms, we set $\lambda_k = 0.4$, $q = 10$, $\forall k \in \mathbb{N}$, and $w_\iota^{(k)} = 1/2q$, $\forall \iota \in \mathcal{I}_k^{(c)} \cup \mathcal{I}_k^{(p)}$, $\forall k \in \mathbb{N}$. The stochastic property sets are designed by $r = 1$ and $\rho = \max\{(r-2)\sigma^2, 0\}$ $(= 0)$, where σ^2 is the variance of noise; for details, see [9]. For the proposed algorithm, we simply set $\varepsilon_1 = \varepsilon_3 = 1.1$, $\varepsilon_2 = \varepsilon_4 = 0.002$, and $\widehat{w}_1 = 0.1$. For the NLMS, the step size is set to $\mu = 0.2$ by following a recommendation given in [6]. For the APA, the step size is set to $\mu = 1$, 0.05 for a comparison. For numerical stability against observable poor excitation of the speech input signals, certain regularization and threshold are utilized, which is the reason for the observable flat intervals.

The results are shown in Fig. 3. We observe that the proposed algorithm reduces the time, to drop by 20dB in system mismatch, by 19[sec.], 11[sec.] and 4[sec.] compared with the NLMS, the APA and the UW-PSP, respectively. Moreover, in the ERLE, the proposed algorithm exhibits much faster convergence in the initial phase than the other methods. Note that the proposed algorithm keeps good steady state performance, while the APA with $\mu = 1$ suffers from serious instability because of the noise (see [9]).

APPENDIX: Asymptotic Optimality of APSM

Recall that $K := \bigcap_{j=1}^{m} K_j (\neq \emptyset)$. The following lemma partly presents the properties of Scheme 1.

Lemma 1 [11, Theorem 1] (cf. [11, Lemma 2])

(I) (Monotone approximation)

$$\left\|\boldsymbol{h}_{k+1} - \boldsymbol{h}_{(k)}^*\right\| \le \left\|\boldsymbol{h}_k - \boldsymbol{h}_{(k)}^*\right\|, \ \forall k \in \mathbb{N},$$
$$\forall \boldsymbol{h}_{(k)}^* \in \Omega_k := \{\boldsymbol{h} \in K : \Theta_k(\boldsymbol{h}) = \inf_{\boldsymbol{x}\in K} \Theta_k(\boldsymbol{x})\}.$$

(II) (Asymptotic optimality)
Suppose (a) $(\Theta_k'(\boldsymbol{h}_k))_{k\in\mathbb{N}}$ is bounded, (b) $\exists \epsilon_1, \epsilon_2 > 0$ s.t. $\lambda_k \in [\epsilon_1, 2-\epsilon_2]$, $\forall k \in \mathbb{N}$, and (c) $\exists N_0 \in \mathbb{N}$ s.t. (i) $\Omega := \bigcap_{k\ge N_0} \Omega_k \neq \emptyset$ and (ii) $\inf_{\boldsymbol{x}\in K} \Theta_k(\boldsymbol{x}) = 0$, $\forall k \ge N_0$. Then, we have

$$\lim_{k\to\infty} \Theta_k(\boldsymbol{h}_k) = 0.$$

Under certain conditions, moreover, it is guaranteed that the sequence $(\boldsymbol{h}_k)_{k\in\mathbb{N}}$ converges strongly to a point $\widehat{\boldsymbol{h}} \in K$.

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).

5. REFERENCES

[1] J. Benesty and Y. Huang, Eds., *Adaptive Signal Processing*, Springer, 2003.

[2] H. Kuttruff, *Room acoustics*, Elsevier, 4th edition, 2000.

[3] Z. Meng, K. Sakagami, M. Morimoto, G. Bi, and K. C. Alex, "Extending the sound impulse response of room using extrapolation," *IEEE Trans. Speech Audio Processing*, vol. 10, no. 3, pp. 167–172, Mar. 2002.

[4] L. Couvreur and C. Couvreur, "Blind model selection for automatic speech recogniton in reverberant environments," *Journal of VLSI Signal Processing*, vol. 36, no. 2–3, pp. 189–203, Mar. 2004.

[5] J. Benesty, D. R. Morgan, and M. M. Sondhi, "A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation," *IEEE Trans. Speech Audio Processing*, vol. 6, no. 3, pp. 156–165, 1998.

[6] A. Sugiyama, Y. Joncour, and A. Hirano, "A stereo echo canceler with correct echo-path identification based on an input-sliding technique," *IEEE Trans. Signal Processing*, vol. 49, no. 11, pp. 2577–2587, 2001.

[7] T. Gänsler and J. Benesty, "Multichannel acoustic echo cancellation: what's new?," in *Proc. IEEE International Workshop on Acoustic Echo and Noise Control*, 2001.

[8] 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.

[9] 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, 2002.

[10] 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.

[11] K. Slavakis, I. Yamada, N. Ogura, and M. Yukawa, "Adaptive projected subgradient method and set theoretic adaptive filtering with multiple convex constraints," in *Proc. Thirty Eighth Asilomar Conf. Signals, Syst., Comput.*, Nov. 2004, pp. 960–964.

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

[13] M. Yukawa, R. L. G. Cavalcante, and I. Yamada, "Efficient blind MAI suppression in DS/CDMA systems by embedded constraint parallel projection techniques," *IEICE Trans. Fundamentals*, to appear.