社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.
CAS2005-2, VLD2005-13, SIP2005-26(2005-6)
June 2005

複数の凸制約条件付き適応信号処理問題の解法と応用

—そのII：収束定理の証明

スラヴァキス　コンスタンティノス†　山田　功†　小倉　信彦††

† 東京工業大学集積システム専攻　〒 152–8552 東京都目黒区大岡山 2–12–1、S3–60
†† 武蔵工業大学環境情報学部　〒 224–0015 横浜市都筑区牛久保西 3–3–1
E-mail: †{slavakis,isao}@comm.ss.titech.ac.jp, ††ogura@yc.musashi-tech.ac.jp

あらまし　先に筆者らは、複数の凸制約条件下で、非負値凸関数列の漸近的最小化問題を解決するために「(複数の凸制約条件付き) 適応射影劣勾配法」を提案し、その大要を報告している [Slavakis & Yamada, 2005 (Technical Report of IEICE-SIP, Jan. 2005)]。小文では、「複数の凸制約条件を同時に満足するベクトルの集合」が「ヒルベルト空間に定義されたある種の非拡大写像の不動点集合」となっていることに注目し、「(非拡大写像の不動点集合上の) 適応射影劣勾配法の収束定理」の厳密な証明を与えている。小文の結果は、適応射影劣勾配法が多様な凸制約条件に柔軟に対応できることを数学的に保証しているばかりでなく、(多様な制約条件を考慮することが必要な) 多くの適応信号処理問題 (「ステレオ音響エコー消去問題」や「アレイアンテナの適応ロバストビーム形成問題」など) を統一的に解決するための基礎を与えている。

キーワード　適応射影劣勾配法、漸近的最小化問題、適応フィルタ、不動点、非拡大写像。

Theory and applications of set theoretic adaptive filtering with multiple a-priori convex constraints

—Part II: Proof of convergence theorem

Konstantinos SLAVAKIS†, Isao YAMADA†, and Nobuhiko OGURA††

† Tokyo Institute of Technology, Dept. of Communications & Integrated Systems.　S3–60, 2–12–1
Ookayama, Meguro-ku, Tokyo, 152–8552 Japan
†† Musashi Institute of Technology, Faculty of Enviromental and Information Studies, 3–3–1,
Ushikubo-Nishi, Tsuzuki-ku, Yokohama 224–0015, Japan
E-mail: †{slavakis,isao}@comm.ss.titech.ac.jp, ††ogura@yc.musashi-tech.ac.jp

Abstract　Recently, the Adaptive Projected Subgradient Method (APSM) over multiple closed convex constraints has been proposed in order to tackle the problem of asymptotically minimizing a sequence of continuous, non-negative, and convex functions over multiple closed convex sets [Slavakis & Yamada, 2005 (Technical Report of IEICE-SIP, Jan. 2005)]. In this paper, by the fact that points satisfying multiple closed convex constraints can be seen as the fixed point set of strongly attracting nonexpansive mappings in a real Hilbert space, we provide with the proofs regarding the convergence theorem of the APSM over the fixed point set of strongly attracting nonexpansive mappings. In this way, these rigorous results firmly support the excellent performance of the APSM to various adaptive signal processing applications with multiple a-priori convex constraints like stereo echo cancelling and robust adaptive beamforming.

Key words　Asymptotic Minimization, Adaptive Filtering, Fixed Point Theory, Nonexpansive Mapping, Subgradient.

1. Introduction

The basic principle behind projection based adaptive filtering methods [1–6] is the successive metric projection onto a series of closed convex subsets of a real Hilbert space $\mathcal{H}$. It is assumed that the series of closed convex sets contain the estimandum (system to be identified) with high probability. The widely used nowadays Affine Projection Algorithm (APA) [2, 3], for example, generates a sequence of estimates by taking, at each time instant $k \in \mathbb{Z}$, the relaxed metric projection onto a linear variety (a special closed convex set defined as the translation of a closed subspace of $\mathcal{H}$) to which the estimandum surely belongs in noiseless situations [6, Appendix B], [7]. Each linear variety is the intersection of a finite number r of hyperplanes. The NLMS [1], for example, is a special case of the APA for $r = 1$. If one, now, recalls that the metric projection mapping provides with a minimizer of the metric distance function to a closed convex set (a nonnegative, continuous, and convex function), then the projection based adaptive filtering methods can be seen as a special case of the general problem of the convexly constrained asymptotic minimization of a sequence of continuous, nonnegative, convex objectives.

An algorithmic solution to this important problem was developed recently by the Adaptive Projected Subgradient Method (APSM) over the fixed point set of strongly attracting nonexpansive mappings in a real Hilbert space [7–9] extending thus the novel work of [10, 11]. The proposed APSM generates a strongly convergent point sequence that asymptotically minimizes a certain sequence of nonnegative continuous convex functions over the fixed point set of strongly attracting nonexpansive mappings. A side effect of the method is the asymptotic minimization of the certain sequence of convex objectives over the nonempty intersection of a finite number of closed convex sets. It provides therefore with new directions and extends the range of projection based adaptive filtering schemes to problems where the estimandum is known to satisfy multiple convex constraints. Many existing adaptive filtering algorithms such as NLMS, APA, Projected NLMS [4], Constrained NLMS [5], Adaptive Parallel Subgradient Algorithm [6], Adaptive Parallel Outer Projection Algorithm [6, 12], Adaptive Parallel Min-Max Projection Algorithm [11] and their embedded constraint versions [11] can be obtained as simple examples of the APSM by an appropriate design of the convex objectives. More than that, the algorithm has recently exhibited remarkably fast and stable convergence properties for highly demanding signal processing applications like adaptive stereophonic echo cancellation [13], and adaptive beamforming [14].

Due to space limitations, this paper gives with a maximal amount of detail the main proofs of the convergence theorem of the method in order to firmly support its recent successful application to real signal processing problems. The full discussion can be found in [9].

2. Preliminaries

Let the set of all integers, nonnegative integers, and real numbers be denoted by $\mathbb{Z}$, $\mathbb{Z}_{\geqq 0}$, and $\mathbb{R}$ respectively.

This paper is based on the assumption of a real Hilbert space $\mathcal{H}$ [15], equipped with an inner product denoted by $\langle x, y\rangle$, $\forall x, y \in \mathcal{H}$. The induced norm will be denoted by $\|x\| := \langle x, x\rangle^{1/2}$, $\forall x \in \mathcal{H}$. Given a nonempty $S \subset \mathcal{H}$ and an $x \in \mathcal{H}$, let $d(x, S) := \inf\{\|x - y\| : y \in S\}$. Moreover, given an $x \in \mathcal{H}$ and an $\varepsilon > 0$, let $B(x, \varepsilon) := \{y \in \mathcal{H} : \|x - y\| < \varepsilon\}$. Also, let $B[x, \varepsilon] := \{y \in \mathcal{H} : \|x - y\| \leqq \varepsilon\}$. Given a nonempty $S \subset \mathcal{H}$, an $x \in S$, and a sufficiently small $\varepsilon > 0$, assume that $B_S(x, \varepsilon) := \{y \in S : \|x - y\| < \varepsilon\} = B(x, \varepsilon) \cap S \neq \emptyset$. The *relative interior* of a nonempty $A \subset \mathcal{H}$ with respect to (w.r.t.) S is defined as $\mathrm{ri}_S(A) := \{x \in A : \exists\varepsilon > 0$ such that (s.t.) $B_S(x, \varepsilon) \subset A\}$. Note here that $\mathrm{int}(A) := \mathrm{ri}_{\mathcal{H}}(A)$ denotes the *interior* of A. Given $a \in \mathcal{H}$ and $\beta \in \mathbb{R}$, a *hyperplane* Π is the set $\Pi := \{x \in \mathcal{H} : \langle a, x\rangle = \beta\}$.

A set $C \subset \mathcal{H}$ is called *convex* if $\forall x, y \in C$, and $\forall\mu \in [0, 1]$, $\mu x + (1 - \mu)y \in C$. A function $\Theta : C \to \mathbb{R} \cup \{\infty\}$ is called *convex* if $\forall x, y \in C$, and $\forall\mu \in [0, 1]$, $\Theta(\mu x + (1 - \mu)y) \leqq \mu\Theta(x) + (1 - \mu)\Theta(y)$.

[Fact 1] [11] Let $C \subset \mathcal{H}$ be a nonempty closed convex set. Assume that there exist $\rho > 0$ and $\tilde{u} \in C$ s.t. $B(\tilde{u}, \rho) \subset C$. Assume $v \in \mathcal{H} \setminus C$ and $t \in (0, 1)$ s.t. $u_t := (1 - t)\tilde{u} + tv \notin C$. Then, $d(v, C) > \rho\frac{1-t}{t} = \rho\frac{\|u_t - v\|}{\|u_t - \tilde{u}\|} > 0$. $\square$

2.1 Nonexpansive mappings

A *fixed point* of a mapping $T : \mathcal{H} \to \mathcal{H}$ is an $x \in \mathcal{H}$ s.t. $T(x) = x$. The *fixed point set* of T is defined as $\mathrm{Fix}(T) := \{x \in \mathcal{H} : T(x) = x\}$. The mapping T is called *nonexpansive* if $\|T(x) - T(y)\| \leqq \|x - y\|$, $\forall x, y \in \mathcal{H}$. If a nonexpansive mapping $T : \mathcal{H} \to \mathcal{H}$ has a fixed point, then $\mathrm{Fix}(T)$ is closed and convex [16, Proposition 1.5.3]. T will be called *attracting nonexpansive* [17] if T is nonexpansive with $\mathrm{Fix}(T) \neq \emptyset$ and $\|T(x) - f\| < \|x - f\|$, $\forall(x, f) \in (\mathcal{H} \setminus \mathrm{Fix}(T)) \times \mathrm{Fix}(T)$, and *strongly attracting* or *η-attracting* nonexpansive [17] if T is nonexpansive with $\mathrm{Fix}(T) \neq \emptyset$ and there exists $\eta > 0$ s.t. $\eta\|x - T(x)\|^2 \leqq \|x - f\|^2 - \|T(x) - f\|^2$, $\forall(x, f) \in \mathcal{H} \times \mathrm{Fix}(T)$. We will consider the identity mapping $I : \mathcal{H} \to \mathcal{H}$ to be η-attracting nonexpansive for an arbitrary $\eta > 0$ and with $\mathrm{Fix}(I) = \mathcal{H}$. For any nonempty closed convex set $C \subset \mathcal{H}$, the *metric projection onto* C is the mapping $P_C : \mathcal{H} \to C$ which maps $x \in \mathcal{H}$ to the uniquely existing $P_C(x) \in C$ s.t. $\|x - P_C(x)\| = d(x, C)$. Note that P_C is an 1-attracting nonexpansive mapping [17].

2.2 Quasi-nonexpansive mappings

A mapping $T : \mathcal{H} \to \mathcal{H}$ is called *quasi-nonexpansive* [11, 18, 19] if $\mathrm{Fix}(T) \neq \emptyset$ and $\|T(x) - f\| \leqq \|x - f\|$, $\forall(x, f) \in \mathcal{H} \times \mathrm{Fix}(T)$. If, in particular, $\|T(x) - f\| < \|x - f\|$, $\forall(x, f) \in (\mathcal{H} \setminus \mathrm{Fix}(T)) \times \mathrm{Fix}(T)$, then T is called *attracting quasi-nonexpansive* (strongly quasi-nonexpansive in [20]). Going even further, if there exists an $\eta > 0$ s.t. $\eta\|x - T(x)\|^2 \leqq \|x - f\|^2 - \|T(x) - f\|^2$, $\forall(x, f) \in \mathcal{H} \times \mathrm{Fix}(T)$, then T is called *η-attracting quasi-nonexpansive*. The mapping T will be called *α-averaged quasi-nonexpansive* if there exists $\alpha \in [0, 1)$ and a quasi-nonexpansive mapping $\mathcal{N} : \mathcal{H} \to \mathcal{H}$ s.t. $T = (1 - \alpha)I + \alpha\mathcal{N}$. In particular, we will call an $\frac{1}{2}$-averaged quasi-nonexpansive mapping *firmly quasi-nonexpansive* (Class $\mathfrak{T}$ in [18]).

[Fact 2] [19] If $T_1 : \mathcal{H} \to \mathcal{H}$ is η_1-attracting quasi-nonexpansive and if $T_2 : \mathcal{H} \to \mathcal{H}$ is η_2-attracting quasi-nonexpansive, then T_2T_1 is $\frac{\eta_1\eta_2}{\eta_1+\eta_2}$-attracting quasi-nonexpansive. $\square$

[Fact 3] [19] Let $T : \mathcal{H} \to \mathcal{H}$ be a firmly quasi-nonexpansive mapping. If $R := (1 - \mu)I + \mu T$, $\mu \in (0, 2)$, then R is $\frac{2-\mu}{\mu}$-attracting quasi-nonexpansive with $\mathrm{Fix}(R) = \mathrm{Fix}(T)$, i.e., $\forall(x, f) \in \mathcal{H} \times \mathrm{Fix}(T)$,

$$\begin{aligned}(2 - \mu)\mu\|x - T(x)\|^2 &= \frac{2 - \mu}{\mu}\|x - R(x)\|^2 \\ &\leqq \|x - f\|^2 - \|R(x) - f\|^2 .\end{aligned}$$

$\square$

2.3 Subgradients

Let $\Theta : \mathcal{H} \to \mathbb{R}$ be a continuous convex function. The *subdif-*

ferential [21, 22] of Θ is the set-valued operator $\partial\Theta : \mathcal{H} \to 2^{\mathcal{H}}$ s.t. $y \mapsto \partial\Theta(y) := \{g \in \mathcal{H} : \langle x - y, g\rangle + \Theta(y) \leqq \Theta(x), \forall x \in \mathcal{H}\}$. For any $y \in \mathcal{H}$, any point in $\partial\Theta(y)$ will be called a *subgradient of* Θ *at* y. Note that for any continuous convex function Θ, $\partial\Theta(y) \neq \emptyset$, $\forall y \in \mathcal{H}$ [21, Proposition I.5.2]. Note also that $0 \in \partial\Theta(y)$ iff $y \in \arg\min_{x\in\mathcal{H}} \Theta(x)$. The function Θ has a unique subgradient at y, if Θ is Gâteaux differentiable at y [21, Proposition I.5.3]. This unique subgradient is nothing but the Gâteaux differential $\Theta'(y)$.

[Fact 4] [22, Example VI.3.3] Given a closed convex set $C \subset \mathcal{H}$, let the *normal cone to* C *at* $x \in C$ be the set $N_C(x) := \{z \in \mathcal{H} : \langle z, y - x\rangle \leqq 0,\ \forall y \in C\}$. Then, $\forall u \in \mathcal{H}$,

$$\partial d(u, C) = \begin{cases} N_C(u) \cap B[0, 1], & u \in C, \\ \frac{u - P_C(u)}{d(u,C)}, & u \notin C. \end{cases}$$

(Although the proof in [22, Example VI.3.3] is performed for a real Euclidean space, one can easily extend it, by following the same steps, to a general infinite dimensional real Hilbert space.) □

Assume now that we have $\mathrm{lev}_{\leqq 0}(\Theta) := \{x \in \mathcal{H} : \Theta(x) \leqq 0\} \neq \emptyset$. Then, define the mapping $T_{\mathrm{sp}(\Theta)} : \mathcal{H} \to \mathcal{H}$ by

$$T_{\mathrm{sp}(\Theta)}(u) := \begin{cases} u - \frac{\Theta(u)}{\|\Theta'(u)\|^2}\Theta'(u), & u \notin \mathrm{lev}_{\leqq 0}(\Theta), \\ u, & u \in \mathrm{lev}_{\leqq 0}(\Theta), \end{cases}$$

where $\Theta'(u)$ is a selection from the set of the subgradients of Θ at u. The mapping $T_{\mathrm{sp}(\Theta)}$ is called the *subgradient projection w.r.t.* Θ. It can be verified that $T_{\mathrm{sp}(\Theta)}$ is firmly quasi-nonexpansive with $\mathrm{Fix}\left(T_{\mathrm{sp}(\Theta)}\right) = \mathrm{lev}_{\leqq 0}(\Theta)$ [18, Proposition 2.3], [20, Lemma 2.8].

The following result generalizes Proposition 2 in [11]. It extends the result from the metric projection P_C onto a closed convex set $C \subset \mathcal{H}$, i.e., an 1-attracting nonexpansive mapping, to the general case of an η-attracting nonexpansive T.

[Proposition 5] Assume that $T : \mathcal{H} \to \mathcal{H}$ is an η-attracting nonexpansive mapping with $\mathrm{Fix}(T) \neq \emptyset$. Let $T_{\mathrm{sp}(\Theta)} : \mathcal{H} \to \mathcal{H}$ be the subgradient projection w.r.t. a continuous convex function $\Theta : \mathcal{H} \to \mathbb{R}$, with $\mathrm{Fix}(T) \cap \mathrm{lev}_{\leqq 0}(\Theta) \neq \emptyset$. For any $\lambda \in (0, 2)$, define the mapping $\widehat{T}_\lambda : \mathcal{H} \to \mathcal{H}$ as

$$\begin{aligned} \widehat{T}_\lambda(u) &:= T\left((1 - \lambda)I + \lambda T_{\mathrm{sp}(\Theta)}\right)(u) \\ &= \begin{cases} T\left(u - \lambda\frac{\Theta(u)}{\|\Theta'(u)\|^2}\Theta'(u)\right), & u \notin \mathrm{lev}_{\leqq 0}(\Theta), \\ T(u), & u \in \mathrm{lev}_{\leqq 0}(\Theta), \end{cases} \end{aligned}$$

where $\Theta'(u) \in \partial\Theta(u)$. Then, the mapping $\widehat{T}_\lambda$ is $\frac{(2-\lambda)\eta}{2-\lambda(1-\eta)}$-attracting quasi-nonexpansive and $\mathrm{Fix}\left(\widehat{T}_\lambda\right) = \mathrm{Fix}(T) \cap \mathrm{lev}_{\leqq 0}(\Theta)$. □

Proof.

Since T is η-attracting nonexpansive, it is also η-attracting quasi-nonexpansive. Moreover, we also know from Fact 3 that $R_{\mathrm{sp}(\Theta)} := (1 - \lambda)I + \lambda T_{\mathrm{sp}(\Theta)}$ is $\frac{2-\lambda}{\lambda}$-attracting quasi-nonexpansive with $\mathrm{Fix}\left(R_{\mathrm{sp}(\Theta)}\right) = \mathrm{Fix}\left(T_{\mathrm{sp}(\Theta)}\right) = \mathrm{lev}_{\leqq 0}(\Theta)$. Thus by Fact 2, the composition $\widehat{T}_\lambda = TR_{\mathrm{sp}(\Theta)}$ is $\frac{(2-\lambda)\eta}{2-\lambda(1-\eta)}$-attracting quasi-nonexpansive with $\mathrm{Fix}(\widehat{T}_\lambda) = \mathrm{Fix}(T) \cap \mathrm{Fix}\left(R_{\mathrm{sp}(\Theta)}\right) = \mathrm{Fix}(T) \cap \mathrm{lev}_{\leqq 0}(\Theta)$. □

2.4 Fejér monotone sequences

Let $C \subset \mathcal{H}$ be a closed convex set. A sequence $(u_n)_{n\geqq 0} \subset \mathcal{H}$ is called *Fejér monotone w.r.t.* C if $\|u_{n+1} - z\| \leqq \|u_n - z\|$, $\forall z \in C, \forall n \geqq 0$.

[Fact 6] [11] Assume that a sequence $(u_n)_{n\geqq 0} \subset \mathcal{H}$ and a closed convex set $C \subset \mathcal{H}$ satisfy the following condition: $\exists \kappa > 0$, s.t. $\forall z \in C$, and $\forall n \geqq 0$), $\kappa\|u_{n+1} - u_n\|^2 \leqq \|u_n - z\|^2 - \|u_{n+1} - z\|^2$. If there exists also a hyperplane $\Pi \subset \mathcal{H}$ s.t. $\mathrm{ri}_\Pi(C) \neq \emptyset$, then $(u_n)_{n\geqq 0}$ converges strongly to a point in $\mathcal{H}$. □

3. The APSM over the Fixed Point Set of Strongly Attracting Nonexpansive Mappings

The following theorem extends Theorem 2 in [11] from the case of the metric projection $P_C : \mathcal{H} \to C$ onto a nonempty closed convex set $C \subset \mathcal{H}$, which is an 1-attracting nonexpansive mapping, to the general case of an η-attracting nonexpansive $T : \mathcal{H} \to \mathcal{H}$ ($\eta > 0$).

[Theorem 7] (*The Adaptive Projected Subgradient Method over the Fixed Point Set of Strongly Attracting Nonexpansive Mappings*) Let $\Theta_n : \mathcal{H} \to [0, \infty)$, $n \in \mathbb{Z}_{\geqq 0}$, be a sequence of continuous convex functions and let $T : \mathcal{H} \to \mathcal{H}$ be an η-attracting nonexpansive mapping with $\mathrm{Fix}(T) \neq \emptyset$. For an arbitrary $u_0 \in \mathcal{H}$, the APSM defines the sequence $(u_n)_{n\in\mathbb{Z}_{\geqq 0}}$ as

$$u_{n+1} := \begin{cases} T\left(u_n - \lambda_n\frac{\Theta_n(u_n)}{\|\Theta'_n(u_n)\|^2}\Theta'_n(u_n)\right), & \Theta'_n(u_n) \neq 0, \\ T(u_n), & \Theta'_n(u_n) = 0, \end{cases} \tag{1}$$

where $\Theta'_n(u_n) \in \partial\Theta_n(u_n)$ and $\lambda_n \in [0, 2]$, $\forall n \in \mathbb{Z}_{\geqq 0}$. Define also

$$\Omega_n := \left\{u \in \mathrm{Fix}(T) : \Theta_n(u) = \Theta_n^* := \inf_{x\in\mathrm{Fix}(T)} \Theta_n(x)\right\}.$$

1. (*Monotone approximation*) Assume the following conditions.

a. $\exists N_0 \in \mathbb{Z}_{\geqq 0}$ s.t. $\Omega := \bigcap_{n\geqq N_0} \Omega_n \neq \emptyset$ and $\Theta_n^* = 0$, $\forall n \geqq N_0$.

b. $\lambda_n \in (0, 2)$, $\forall n \geqq N_0$. Then,

$$\left\|u_{n+1} - u^*_{(n)}\right\| \leqq \left\|u_n - u^*_{(n)}\right\|,\ \forall u^*_{(n)} \in \Omega_n, \forall n \geqq N_0.$$

c. In addition, assume that $u_n \notin \Omega_n$ for $n \geqq N_0$. Then,

$$\left\|u_{n+1} - u^*_{(n)}\right\| < \left\|u_n - u^*_{(n)}\right\|,\ \forall u^*_{(n)} \in \Omega_n.$$

2. (*Strong convergence*) Assume the condition in Theorem 7.1.a. Assume also:

a. $\exists \varepsilon_1, \varepsilon_2 > 0$ s.t. $\lambda_n \in [\varepsilon_1, 2 - \varepsilon_2]$, $\forall n \geqq N_0$.

b. There exists a hyperplane $\Pi \subset \mathcal{H}$ s.t. $\mathrm{ri}_\Pi(\Omega) \neq \emptyset$.

Then, the sequence $(u_n)_{n\in\mathbb{Z}_{\geqq 0}}$ converges strongly to a point $\widehat{u} \in \mathrm{Fix}(T)$.

3. (*Asymptotic optimality*) Assume the conditions in Theorem 7.1.a and in Theorem 7.2.a.

a. Suppose that $(\Theta'_n(u_n))_{n\in\mathbb{Z}_{\geqq 0}}$ is bounded. Then, $\lim_{n\to\infty} \Theta_n(u_n) = 0$.

b. Assume also the condition in Theorem 7.2.b and that $(\Theta'_n(\widehat{u}))_{n\in\mathbb{Z}_{\geqq 0}}$ is bounded. Then, $\lim_{n\to\infty} \Theta_n(\widehat{u}) = 0$.

4. (*Characterization of* $\widehat{u}$) Assume the conditions in Theorem 7.1.a, in Theorem 7.2.a, and the existence of $\widehat{u} = \lim_{n\to\infty} u_n \in \mathrm{Fix}(T)$. Assume also that

a. $\tilde{u} \in \mathrm{int}(\Omega) \neq \emptyset$.

b. $\forall \varepsilon > 0$, $\forall r > 0$, $\exists \delta > 0$ s.t.

$$\inf_{\substack{d(u_n, \mathrm{lev}_{\leqq 0}(\Theta_n)) \geqq \varepsilon,\ \|u_n - \tilde{u}\| \leqq r, \\ n \geqq N_0}} \Theta_n(u_n) \geqq \delta.$$

If also $\lim_{n\to\infty} \Theta_n(u_n) = 0$, then $\widehat{u} \in \overline{\liminf_{n\to\infty} \Omega_n}$, where the overline denotes the closure, in the strong topology of $\mathcal{H}$, of $\liminf_{n\to\infty} \Omega_n := \bigcup_{n=0}^{\infty} \bigcap_{m\geqq n} \Omega_m$. □

Proof of Theorem 7.1.

Notice that $\forall n \geqq N_0$, $\Omega_n = \operatorname{Fix}(T) \cap \operatorname{lev}_{\leqq 0}(\Theta_n)$. Assume, now, $\Theta_n(u_n) > 0$ for $n \geqq N_0$. Then, u_n is not a global minimizer of Θ_n, since we already know by assumption that $\forall u^* \in \Omega \neq \emptyset$, $\Theta_n(u^*) = \Theta_n^* = 0$. Thus, $0 \notin \partial\Theta_n(u_n) \Rightarrow \Theta'_n(u_n) \neq 0$, which in turn implies by (1) that $u_{n+1} = T\left(u_n - \lambda_n \frac{\Theta_n(u_n)}{\|\Theta'_n(u_n)\|^2}\Theta'_n(u_n)\right)$. Assume that $\Theta_n(u_n) = 0$. If $\Theta'_n(u_n) = 0$, then $u_{n+1} = T(u_n)$. Moreover, if $\Theta'_n(u_n) \neq 0$, then we have $u_{n+1} = T\left(u_n - \lambda_n \frac{0}{\|\Theta'_n(u_n)\|^2}\Theta'_n(u_n)\right) = T(u_n)$. Hence, (1) takes the form

$$u_{n+1} := \begin{cases} T\left(u_n - \lambda_n \frac{\Theta_n(u_n)}{\|\Theta'_n(u_n)\|^2}\Theta'_n(u_n)\right), & \Theta_n(u_n) > 0, \\ T(u_n), & \Theta_n(u_n) = 0, \end{cases}$$
$$= T\left((1-\lambda_n)I + \lambda_n T_{\operatorname{sp}(\Theta_n)}\right)(u_n) = \widehat{T}_{\lambda_n}(u_n), \quad \forall n \geqq N_0.$$

By Proposition 5 and for any $n \geqq N_0$, the mapping $\widehat{T}_{\lambda_n}$ is $\frac{(2-\lambda_n)\eta}{2-\lambda_n(1-\eta)}$-attracting quasi-nonexpansive w.r.t.

$$\operatorname{Fix}\left(\widehat{T}_{\lambda_n}\right) = \operatorname{Fix}(T) \cap \operatorname{lev}_{\leqq 0}(\Theta_n) = \Omega_n. \tag{2}$$

Thus, $\forall u^*_{(n)} \in \Omega_n$ and $\forall n \geqq N_0$,

$$\frac{(2-\lambda_n)\eta}{2-\lambda_n(1-\eta)}\|u_n - u_{n+1}\|^2 \leqq \left\|u_n - u^*_{(n)}\right\|^2 - \left\|u_{n+1} - u^*_{(n)}\right\|^2. \tag{3}$$

Notice that $\frac{(2-\lambda_n)\eta}{2-\lambda_n(1-\eta)} > 0$, $\forall \lambda_n \in (0,2)$. Then, by (3), Theorem 7.1.b follows. If we also assume $u_n \notin \Omega_n = \operatorname{Fix}\left(\widehat{T}_{\lambda_n}\right)$, then $u_n \neq \widehat{T}_{\lambda_n}(u_n) = u_{n+1}$. Hence, by using (3) again, Theorem 7.1.c is obtained. □

Proof of Theorem 7.2.

By (2), $\Omega = \bigcap_{n \geqq N_0} \operatorname{Fix}\left(\widehat{T}_{\lambda_n}\right)$. Thus Theorem 7.2.a and (3) imply that

$$\kappa\|u_n - u_{n+1}\|^2 \leqq \|u_n - u^*\|^2 - \|u_{n+1} - u^*\|^2, \forall u^* \in \Omega, \forall n \geqq N_0, \tag{4}$$

where $\kappa := \frac{\varepsilon_2\eta}{\max\{(2-(1-\eta)\varepsilon_1),(\varepsilon_2(1-\eta)+2\eta)\}} > 0$. Now, (4), Theorem 7.2.b, and Fact 6 suggest that $(u_n)_{n\in\mathbb{Z}_{\geqq 0}}$ converges strongly to a point $\widehat{u} \in \mathcal{H}$.

Define $\mathcal{J} := \{n \in \mathbb{Z}_{\geqq 0} : \Theta'_n(u_n) \neq 0\}$. Assume $\operatorname{card}(\mathcal{J}) = \infty$, where $\operatorname{card}(\cdot)$ denotes the cardinal number of a set. Then, $\mathcal{J}$ can be considered as the subsequence $\mathcal{J} = (n_m)_{m\in\mathbb{Z}_{\geqq 0}}$, which for the sake of simplicity in notations is denoted by $(m)_{m\in\mathcal{J}}$. Next we show that $\lim_{m\to\infty} \frac{\Theta_m(u_m)}{\|\Theta'_m(u_m)\|} = 0$.

Assume for a contradiction that $\exists \varepsilon > 0$ and a subsequence $(m_l)_{l\in\mathbb{Z}_{\geqq 0}}$ s.t. $\frac{\Theta_{m_l}(u_{m_l})}{\left\|\Theta'_{m_l}(u_{m_l})\right\|} \geqq \varepsilon$, $\forall l \geqq 0$. Then, by the nonexpansivity of T, $\forall u^* \in \Omega \subset \operatorname{Fix}(T)$, $\forall l \geqq 0$,

$$\begin{aligned} \left\|u_{m_l+1} - u^*\right\|^2 &\leqq \left\|u_{m_l} - \lambda_{m_l}\frac{\Theta_{m_l}(u_{m_l})}{\left\|\Theta'_{m_l}(u_{m_l})\right\|^2}\Theta'_{m_l}(u_{m_l}) - u^*\right\|^2 \\ &= \|u_{m_l} - u^*\|^2 + \lambda_{m_l}^2\frac{\Theta_{m_l}^2(u_{m_l})}{\left\|\Theta'_{m_l}(u_{m_l})\right\|^2} \\ &\quad - 2\lambda_{m_l}\frac{\Theta_{m_l}(u_{m_l})}{\left\|\Theta'_{m_l}(u_{m_l})\right\|^2}\left\langle u_{m_l} - u^*, \Theta'_{m_l}(u_{m_l})\right\rangle. \end{aligned} \tag{5}$$

By the definition of the subgradient we have that $\forall l \geqq 0$, $\left\langle u^* - u_{m_l}, \Theta'_{m_l}(u_{m_l})\right\rangle + \Theta_{m_l}(u_{m_l}) \leqq \Theta_{m_l}(u^*)$. There exists $l_0 \in \mathbb{Z}_{\geqq 0}$ s.t. $l \geqq l_0 \Rightarrow m_l \geqq N_0$. Thus, $\forall l \geqq l_0$, $\left\langle u^* - u_{m_l}, \Theta'_{m_l}(u_{m_l})\right\rangle \leqq -\Theta_{m_l}(u_{m_l})$. Now, by (5) and $\forall l \geqq l_0$,

$$\lambda_{m_l}(2-\lambda_{m_l})\frac{\Theta_{m_l}^2(u_{m_l})}{\left\|\Theta'_{m_l}(u_{m_l})\right\|^2} \leqq \|u_{m_l} - u^*\|^2 - \left\|u_{m_l+1} - u^*\right\|^2,$$

and finally

$$\varepsilon_1\varepsilon_2\varepsilon^2 \leqq \|u_{m_l} - u^*\|^2 - \left\|u_{m_l+1} - u^*\right\|^2. \tag{6}$$

The sequence $\left(\|u_n - u^*\|^2\right)_{n\geqq N_0}$ is nonnegative and nonincreasing by (4); thus convergent. Take the limit in (6) for $l \to \infty$ to obtain the absurd result $0 < \varepsilon_1\varepsilon_2\varepsilon^2 \leqq 0$, which leads to a contradiction. Hence, in our original notation, $\lim_{m\to\infty}\frac{\Theta_{n_m}(u_{n_m})}{\left\|\Theta'_{n_m}(u_{n_m})\right\|} = 0$. It is also easy to verify by

$$\left\|\lambda_{n_m}\frac{\Theta_{n_m}(u_{n_m})}{\left\|\Theta'_{n_m}(u_{n_m})\right\|^2}\Theta'_{n_m}(u_{n_m})\right\| \leqq (2-\varepsilon_2)\frac{\Theta_{n_m}(u_{n_m})}{\left\|\Theta'_{n_m}(u_{n_m})\right\|},$$

that

$$\lim_{m\to\infty}\lambda_{n_m}\frac{\Theta_{n_m}(u_{n_m})}{\left\|\Theta'_{n_m}(u_{n_m})\right\|^2}\Theta'_{n_m}(u_{n_m}) = 0.$$

The APSM suggests that

$$u_{n_m+1} = T\left(u_{n_m} - \lambda_{n_m}\frac{\Theta_{n_m}(u_{n_m})}{\left\|\Theta'_{n_m}(u_{n_m})\right\|^2}\Theta'_{n_m}(u_{n_m})\right).$$

Since the original sequence $(u_n)_{n\in\mathbb{Z}_{\geqq 0}}$ converges strongly to $\widehat{u}$, any subsequence of $(u_n)_{n\in\mathbb{Z}_{\geqq 0}}$ will also converge strongly to $\widehat{u}$. Recall that T is nonexpansive, thus continuous in the strong topology of $\mathcal{H}$, and take the limit above as $m \to \infty$ to obtain that

$$\begin{aligned} \widehat{u} &= \lim_{m\to\infty} T\left(u_{n_m} - \lambda_{n_m}\frac{\Theta_{n_m}(u_{n_m})}{\left\|\Theta'_{n_m}(u_{n_m})\right\|^2}\Theta'_{n_m}(u_{n_m})\right) \\ &= T\left(\lim_{m\to\infty} u_{n_m} - \lim_{m\to\infty}\lambda_{n_m}\frac{\Theta_{n_m}(u_{n_m})}{\left\|\Theta'_{n_m}(u_{n_m})\right\|^2}\Theta'_{n_m}(u_{n_m})\right) \\ &= T(\widehat{u}). \end{aligned}$$

This clearly implies that the limiting point $\widehat{u} \in \operatorname{Fix}(T)$.

Assume, now, that $\operatorname{card}(\mathcal{J}) < \infty$. Then there exists N_1 s.t. $\forall n \geqq N_1$, $\Theta'_n(u_n) = 0$. For $n \geqq N_1$, $u_{n+1} = T(u_n) \Rightarrow \lim_{n\to\infty} u_{n+1} = \lim_{n\to\infty} T(u_n) \Rightarrow \widehat{u} = T(\lim_{n\to\infty} u_n) = T(\widehat{u}) \Rightarrow \widehat{u} \in \operatorname{Fix}(T)$. This completes the proof of Theorem 7.2. □

Proof of Theorem 7.3.

Since the sequence $(\Theta'_n(u_n))_{n\in\mathbb{Z}_{\geqq 0}}$ is bounded, there exists $\gamma > 0$ s.t. $\|\Theta'_n(u_n)\| \leqq \gamma$, $\forall n \in \mathbb{Z}_{\geqq 0}$.

Assume again that $\operatorname{card}(\mathcal{J}) = \infty$. We have already seen above that $\mathcal{J}$ can be denoted by $(n_m)_{m\in\mathbb{Z}_{\geqq 0}}$ and that $\lim_{m\to\infty}\frac{\Theta_{n_m}(u_{n_m})}{\left\|\Theta'_{n_m}(u_{n_m})\right\|} = 0$. This means that given $\varepsilon > 0$, there exists M_0 s.t. $\forall m \geqq M_0$, $\frac{\Theta_{n_m}(u_{n_m})}{\left\|\Theta'_{n_m}(u_{n_m})\right\|} < \frac{\varepsilon}{\gamma}$. Thus, $\forall m \geqq M_0$, $0 \leqq \Theta_{n_m}(u_{n_m}) < \frac{\varepsilon}{\gamma}\left\|\Theta'_{n_m}(u_{n_m})\right\| \leqq \frac{\varepsilon}{\gamma}\gamma = \varepsilon$. This implies that $\lim_{m\to\infty}\Theta_{n_m}(u_{n_m}) = 0$. Clearly, for all $N_0 \leqq n \in \mathbb{Z}_{\geqq 0} \setminus \mathcal{J}$, $\Theta_n(u_n) = 0$, which leads finally to $\lim_{n\to\infty}\Theta_n(u_n) = 0$.

For the case where $\operatorname{card}(\mathcal{J}) < \infty$, there exists N_1 s.t. $\forall n \geqq N_1$, $\Theta'_n(u_n) = 0$. Thus $\forall n \geqq \max\{N_0, N_1\}$, $\Theta_n(u_n) = 0 \Rightarrow \lim_{n\to\infty}\Theta_n(u_n) = 0$.

Assume also that there exists $\gamma' > 0$ s.t. $\|\Theta'_n(\widehat{u})\| \leqq \gamma'$, $\forall n \in \mathbb{Z}_{\geqq 0}$. By the definition of the subgradient, $\langle u_n - \widehat{u}, \Theta'_n(\widehat{u})\rangle + \Theta_n(\widehat{u}) \leqq \Theta_n(u_n)$, $\forall n \in \mathbb{Z}_{\geqq 0}$. Obviously, we have that

$$\begin{aligned} 0 \leqq \Theta_n(\widehat{u}) &\leqq \Theta_n(u_n) - \langle u_n - \widehat{u}, \Theta'_n(\widehat{u})\rangle \\ &\leqq \Theta_n(u_n) + \|\widehat{u} - u_n\|\,\|\Theta'_n(\widehat{u})\| \\ &\leqq \left(\Theta_n(u_n) + \gamma'\|\widehat{u} - u_n\|\right) \xrightarrow[n\to\infty]{} 0. \end{aligned}$$

Thus, $\lim_{n\to\infty}\Theta_n(\widehat{u}) = 0$, which completes the proof of Theorem 7.3. $\square$

Proof of Theorem 7.4.

One can verify by definition that Ω_n is convex $\forall n \geqq 0$; thus $\liminf_{n\to\infty}\Omega_n$ is also convex. Similarly, $\mathrm{lev}_{\leqq 0}(\Theta_n)$ is convex $\forall n \geqq 0$. By Theorem 7.4.a, there exists a $\rho > 0$ s.t. $B(\tilde{u},\rho) \subset \Omega$.

Assume for a contradiction that $\widehat{u} \notin \overline{\liminf_{n\to\infty}\Omega_n}$. Then, $\exists t \in (0,1)$ s.t. $u_t := (1-t)\tilde{u} + t\widehat{u} \notin \overline{\liminf_{n\to\infty}\Omega_n} \supset \liminf_{n\to\infty}\Omega_n$. Since $\tilde{u}, \widehat{u} \in \mathrm{Fix}(T)$, $u_t \in \mathrm{Fix}(T)$. Since, also, $\lim_{n\to\infty} u_n = \widehat{u}$, $\exists L_0 \geqq N_0$ s.t. $\|u_n - \widehat{u}\| \leqq \frac{\rho}{2}\frac{1-t}{t}$, $\forall n \geqq L_0$. By the result $u_t \notin \liminf_{n\to\infty}\Omega_n$, we have that for any $L_1 \geqq L_0$, $\exists n_1 = n_1(L_1) \geqq L_1$, s.t. $u_t \notin \Omega_{n_1} = \mathrm{Fix}(T) \cap \mathrm{lev}_{\leqq 0}(\Theta_{n_1})$. Thus $u_t \notin \mathrm{lev}_{\leqq 0}(\Theta_{n_1})$.

Now, by Fact 1 and $B(\tilde{u},\rho) \subset \Omega \subset \Omega_{n_1} \subset \mathrm{lev}_{\leqq 0}(\Theta_{n_1})$,

$$\begin{aligned} d(u_{n_1}, \mathrm{lev}_{\leqq 0}(\Theta_{n_1})) &\geqq d(\widehat{u}, \mathrm{lev}_{\leqq 0}(\Theta_{n_1})) - \|u_{n_1} - \widehat{u}\| \\ &\geqq \rho\frac{1-t}{t} - \frac{\rho}{2}\frac{1-t}{t} = \frac{\rho}{2}\frac{1-t}{t} =: \varepsilon > 0. \end{aligned}$$

We also have

$$\|\tilde{u} - u_{n_1}\| \leqq \|\tilde{u} - \widehat{u}\| + \frac{\rho}{2}\frac{1-t}{t} =: r > 0.$$

Similarly, fix $L_2 > n_1$. Then, as above, $\exists n_2 = n_2(L_2) \geqq L_2$ s.t. $d(u_{n_2}, \mathrm{lev}_{\leqq 0}(\Theta_{n_2})) \geqq \varepsilon$ and $\|\tilde{u} - u_{n_2}\| \leqq r$. In other words, we can construct a subsequence $(n_l)_{l\geqq 1}$ s.t.

$$d(u_{n_l}, \mathrm{lev}_{\leqq 0}(\Theta_{n_l})) \geqq \varepsilon,\ \|\tilde{u} - u_{n_l}\| \leqq r, \text{ and } n_l \geqq N_0,\ \forall l \geqq 1.$$

Obviously, by Theorem 7.4.b, there exists a $\delta > 0$ s.t. $\Theta_{n_l}(u_{n_l}) \geqq \delta$, $\forall l \geqq 1$. This contradicts $\lim_{n\to\infty}\Theta_n(u_n) = 0$. We therefore obtain $\widehat{u} \in \overline{\liminf_{n\to\infty}\Omega_n}$ which completes the proof of Theorem 7. $\square$

4. An Example of the APSM

Due to lack of space, only one example of the proposed APSM is shown. However, it is general enough to cover many existing projection based adaptive filtering algorithms. For more examples and for the full discussion the interested reader is referred to [9].

Assume $\mathcal{J}_n \subset \mathbb{Z}_{\geqq 0}$ s.t. $\mathrm{card}(\mathcal{J}_n) < \infty$, $\forall n \in \mathbb{Z}_{\geqq 0}$. Let, then, the closed convex set $S_\iota^{(n)} \subset \mathcal{H}$ and its associated metric projection $P_{S_\iota^{(n)}}$, $\forall \iota \in \mathcal{J}_n$, $\forall n \in \mathbb{Z}_{\geqq 0}$. Define $\omega_\iota^{(n)} \in (0,1)$, $\forall \iota \in \mathcal{J}_n$, $\forall n \in \mathbb{Z}_{\geqq 0}$, s.t. $\sum_{\iota\in\mathcal{J}_n}\omega_\iota^{(n)} = 1$, $\forall n \in \mathbb{Z}_{\geqq 0}$. Assume also an η-attracting nonexpansive mapping $T: \mathcal{H} \to \mathcal{H}$ with $\mathrm{Fix}(T) \neq \emptyset$. In the following, we abide by the condition that $\mathrm{Fix}(T) \cap \left(\bigcap_{\iota\in\mathcal{J}_n, n\geqq 0} S_\iota^{(n)}\right) \neq \emptyset$. A successful strategy for the generation of such sequences of closed convex sets $(S_\iota^{(n)})_{n\in\mathbb{Z}_{\geqq 0}} \subset \mathcal{H}$ in adaptive filtering schemes for system identification is demonstrated in [6]. Therein, the design is based on the information of the statistical properties of the noise process that usually contaminates environments where system identification is desired.

[Example 8] Fix $n \geqq 0$. Given $u_n \in \mathcal{H}$, define for any $u \in \mathcal{H}$ the convex function:

$$\Theta_n(u) := \begin{cases} \frac{1}{L_n}\sum_{\iota\in\mathcal{J}_n}\omega_\iota^{(n)} d\left(u_n, S_\iota^{(n)}\right) d\left(u, S_\iota^{(n)}\right), \\ \qquad\qquad u_n \notin \bigcap_{\iota\in\mathcal{J}_n} S_\iota^{(n)}, \\ 0, \qquad\qquad \text{otherwise,} \end{cases} \tag{7}$$

where $L_n := \sum_{\iota\in\mathcal{J}_n}\omega_\iota^{(n)} d\left(u_n, S_\iota^{(n)}\right)$. Define also $\mathcal{I}_n := \left\{\iota \in \mathcal{J}_n : u_n \notin S_\iota^{(n)}\right\}$. Clearly, if $u_n \in \bigcap_{\iota\in\mathcal{J}_n} S_\iota^{(n)}$, then $\mathcal{I}_n = \emptyset$. If $u_n \notin \bigcap_{\iota\in\mathcal{J}_n} S_\iota^{(n)}$, then $L_n = \sum_{\iota\in\mathcal{I}_n}\omega_\iota^{(n)} d\left(u_n, S_\iota^{(n)}\right)$.

If $u_n \notin \bigcap_{\iota\in\mathcal{J}_n} S_\iota^{(n)}$, then $\mathrm{lev}_{\leqq 0}(\Theta_n) = \bigcap_{\iota\in\mathcal{I}_n} S_\iota^{(n)}$, $\Omega_n = \mathrm{Fix}(T) \cap \left(\bigcap_{\iota\in\mathcal{I}_n} S_\iota^{(n)}\right)$, $\Theta_n^* = 0$, and $\Omega = \mathrm{Fix}(T) \cap \left(\bigcap_{\iota\in\mathcal{I}_n, n\geqq 0} S_\iota^{(n)}\right)$. For no ambiguity, if $\mathcal{I}_n = \emptyset$ we let $\bigcap_{\iota\in\mathcal{I}_n} S_\iota^{(n)} = \mathcal{H}$.

Obviously, if $u_n \in \bigcap_{\iota\in\mathcal{J}_n} S_\iota^{(n)}$, we drop to the trivial case $\Theta_n'(u) = 0, \forall u \in \mathcal{H}$. However, in the case where $u_n \notin \bigcap_{\iota\in\mathcal{J}_n} S_\iota^{(n)}$, we obtain by [21, (I.5.21) and Proposition I.5.6] that

$$\partial\Theta_n(u) = \frac{1}{L_n}\sum_{\iota\in\mathcal{I}_n}\omega_\iota^{(n)} d\left(u_n, S_\iota^{(n)}\right)\partial d\left(u, S_\iota^{(n)}\right),\ \forall u \in \mathcal{H}. \tag{8}$$

In particular, if $u \notin \bigcap_{\iota\in\mathcal{J}_n} S_\iota^{(n)}$, we can make the following choice among the subgradients in $\partial\Theta_n(u)$ (see Fact 4),

$$\Theta_n'(u) = \sum_{\left\{\iota\in\mathcal{I}_n : u\notin S_\iota^{(n)}\right\}} \frac{\omega_\iota^{(n)} d\left(u_n, S_\iota^{(n)}\right)}{L_n}\frac{u - P_{S_\iota^{(n)}}(u)}{d\left(u, S_\iota^{(n)}\right)}.$$

Thus, if $u_n \notin \bigcap_{\iota\in\mathcal{J}_n} S_\iota^{(n)}$,

$$\begin{aligned} \Theta_n'(u_n) &= \frac{1}{L_n}\sum_{\iota\in\mathcal{I}_n}\omega_\iota^{(n)}\left(u_n - P_{S_\iota^{(n)}}(u_n)\right) \\ &= \frac{1}{L_n}\sum_{\iota\in\mathcal{J}_n}\omega_\iota^{(n)}\left(u_n - P_{S_\iota^{(n)}}(u_n)\right). \end{aligned}$$

Now, for an arbitrary $u_0 \in \mathcal{H}$, use the APSM to define inductively the series of functions Θ_n, as introduced in (7). The following algorithm is then formed.

$$u_{n+1} = T\left(u_n + \mu_n\left(\sum_{\iota\in\mathcal{J}_n}\omega_\iota^{(n)} P_{S_\iota^{(n)}}(u_n) - u_n\right)\right),\ \forall n \geqq 0,$$

where $\mu_n \in [0, 2\mathcal{M}_n^{(1)}]$, and

$$\mathcal{M}_n^{(1)} := \begin{cases} \frac{\sum_{\iota\in\mathcal{J}_n}\omega_\iota^{(n)}\left\|P_{S_\iota^{(n)}}(u_n) - u_n\right\|^2}{\left\|\sum_{\iota\in\mathcal{J}_n}\omega_\iota^{(n)} P_{S_\iota^{(n)}}(u_n) - u_n\right\|^2}, & u_n \notin \bigcap_{\iota\in\mathcal{J}_n} S_\iota^{(n)}, \\ 1, & \text{otherwise.} \end{cases}$$

$\square$

[Proposition 9] Let the sequence $(u_n)_{n\geqq 0}$ defined by the APSM in the Example 8. Throughout, we assume $\mathrm{Fix}(T) \cap \left(\bigcap_{\iota\in\mathcal{J}_n, n\geqq 0} S_\iota^{(n)}\right) \neq \emptyset$. Assume also that $(\mu_n)_{n\geqq 0} \subset \left[\mathcal{M}_n^{(1)}\varepsilon_1, \mathcal{M}_n^{(1)}(2-\varepsilon_2)\right]$ in Example 8 for some $\varepsilon_1, \varepsilon_2 > 0$.

1. Then,

a. $\lim_{n\to\infty}\Theta_n(u_n) = 0$.

Assume that there exists a hyperplane $\Pi \subset \mathcal{H}$ s.t. $\tilde{u} \in \mathrm{ri}_\Pi(\Omega) \neq \emptyset$. Then the following statements hold.

b. There exists $\widehat{u} \in \mathrm{Fix}(T)$ s.t. $\lim_{n\to\infty} u_n = \widehat{u}$.

c. $\lim_{n\to\infty}\Theta_n(\widehat{u}) = 0$.

2. Assume that there exists $\tilde{u} \in \mathrm{int}(\Omega) \neq \emptyset$, for Ω defined in Example 8. Assume also that $\omega_0 := \inf_{\iota\in\mathcal{J}_n, n\geqq 0}\omega_\iota^{(n)} > 0$. Then, $\widehat{u} = \lim_{n\to\infty} u_n \in \overline{\liminf_{n\to\infty}\Omega_n}$. $\square$

Proof of Proposition 9.1.a.

By definition, and $\forall n \geqq 0$, we have $\mu_n = \lambda_n\mathcal{M}_n^{(1)}$ for Example 8. Thus, the original relaxation coefficient λ_n in APSM lies in the range $\lambda_n \in [\varepsilon_1, 2-\varepsilon_2]$, $\forall n \geqq 0$. Since we have also assumed $\mathrm{Fix}(T) \cap \left(\bigcap_{\iota\in\mathcal{J}_n, n\geqq 0} S_\iota^{(n)}\right) \neq \emptyset$, it can be readily verified that the conditions in Theorem 7.1.a and in Theorem 7.2.a are satisfied.

Notice that whenever $u_n \in \bigcap_{\iota\in\mathcal{J}_n} S_\iota^{(n)}$, then we obtain the trivial case $\Theta_n'(u) = 0, \forall u \in \mathcal{H}$. Assume, now, $u_n \notin \bigcap_{\iota\in\mathcal{J}_n} S_\iota^{(n)}$. If $u \in \bigcap_{\iota\in\mathcal{J}_n} S_\iota^{(n)}$, then we have by (8) and Fact 4 that

$$\partial\Theta_n(u) \subset \frac{1}{L_n}\sum_{\iota\in\mathcal{I}_n}\left(\omega_\iota^{(n)} d\left(u_n, S_\iota^{(n)}\right) B[0,1]\right) = B[0,1],\ \forall u \in \mathcal{H}.$$

Now, assume that $u_n \notin \bigcap_{\iota \in \mathcal{J}_n} S_\iota^{(n)}$ and $u \notin \bigcap_{\iota \in \mathcal{J}_n} S_\iota^{(n)}$. Then,

$$\left\|\Theta_n'(u)\right\| \leqq \sum_{\left\{\iota \in \mathcal{I}_n : u \notin S_\iota^{(n)}\right\}} \frac{\omega_\iota^{(n)} d\left(u_n, S_\iota^{(n)}\right)}{L_n} \frac{\left\|u - P_{S_\iota^{(n)}}(u)\right\|}{d\left(u, S_\iota^{(n)}\right)} \leqq 1.$$

Concluding, for Example 8,

$$\left\|\Theta_n'(u)\right\| \leqq 1,\ \forall u \in \mathcal{H},\ \forall n \geqq 0. \tag{9}$$

In particular, $\|\Theta_n'(u_n)\| \leqq 1$, $\forall n \geqq 0$. Hence, by Theorem 7.3.a, we get $\lim_{n\to\infty} \Theta_n(u_n) = 0$ for Example 8. □

Proof of Proposition 9.1.b and Proposition 9.1.c.

We have already seen above that the conditions in Theorem 7.1.a and in Theorem 7.2.a are satisfied. Moreover, by assumption, the condition in Theorem 7.2.b also holds. Thus, Proposition 9.1.b is proved for Example 8.

By (9), the sequence $(\Theta_n'(\widehat{u}))_{n\in\mathbb{Z}_{\geqq 0}}$ is bounded for Example 8. Hence, by Theorem 7.3.b, Proposition 9.1.c holds. □

Proof of Proposition 9.2.

An inspection of Theorem 7.4.b suggests that our task will be over as soon as we show that $\forall \varepsilon > 0$, $\forall r > 0$, $\exists \delta > 0$ s.t.

$$\inf_{\substack{d(u_n, \mathrm{lev}_{\leqq 0}(\Theta_n)) \geqq \varepsilon,\ \|u_n - \tilde{u}\| \leqq r, \\ n \geqq 0}} \Theta_n(u_n) \geqq \delta.$$

For any $\varepsilon, r > 0$, define $\mathcal{O} := \{n \in \mathbb{Z}_{\geqq 0} : d(u_n, \mathrm{lev}_{\leqq 0}(\Theta_n)) \geqq \varepsilon, \|u_n - \tilde{u}\| \leqq r\}$. Fix $n \in \mathcal{O}$ arbitrarily. Notice that there exists $\rho > 0$ s.t. $B(\tilde{u}, \rho) \subset \Omega = \mathrm{Fix}(T) \cap \left(\bigcap_{k \geqq 0} \mathrm{lev}_{\leqq 0}(\Theta_k)\right) \subset \mathrm{lev}_{\leqq 0}(\Theta_n)$. It is easy also to verify that $\frac{\varepsilon}{2r} < 1$ by

$$2r > r \geqq \|u_n - \tilde{u}\| \geqq d\left(u_n, \mathrm{lev}_{\leqq 0}(\Theta_n)\right) \geqq \varepsilon.$$

Let, then, $t := 1 - \frac{\varepsilon}{2r}$ and notice that the point $u_t := (1-t)\tilde{u} + t u_n$ satisfies $\|u_t - u_n\| \leqq \frac{\varepsilon}{2}$. This implies that $u_t \notin \mathrm{lev}_{\leqq 0}(\Theta_n)$. Notice that for Example 8 $\mathrm{lev}_{\leqq 0}(\Theta_n) = \bigcap_{\iota \in \mathcal{I}_n} S_\iota^{(n)}$. Hence, $u_t \notin \mathrm{lev}_{\leqq 0}(\Theta_n)$ is equivalent to $\exists \iota_0 \in \mathcal{I}_n$ s.t. $u_t \notin S_{\iota_0}^{(n)}$. However, by Fact 1 and $B(\tilde{u}, \rho) \subset \Omega$, we obtain that

$$d\left(u_n, S_{\iota_0}^{(n)}\right) > \rho \frac{1-t}{t} = \rho \frac{\varepsilon}{2r - \varepsilon}.$$

Now,

$$\begin{aligned} \Theta_n(u_n) &= \frac{1}{L_n} \sum_{\iota \in \mathcal{I}_n} \omega_\iota^{(n)} d^2\left(u_n, S_\iota^{(n)}\right) \geqq \frac{1}{L_n} \omega_{\iota_0}^{(n)} d^2\left(u_n, S_{\iota_0}^{(n)}\right) \\ &> \frac{\omega_0 \rho^2}{r} \frac{\varepsilon^2}{(2r - \varepsilon)^2}, \end{aligned} \tag{10}$$

since $d\left(u_n, S_\iota^{(n)}\right) \leqq d\left(u_n, \mathrm{lev}_{\leqq 0}(\Theta_n)\right) \leqq \|u_n - \tilde{u}\| \leqq r$, $\forall \iota \in \mathcal{I}_n$, and thus $L_n < r$. Hence, we can choose a $\delta > 0$, e.g., $0 < \delta \leqq \frac{\omega_0 \rho^2}{r} \frac{\varepsilon^2}{(2r-\varepsilon)^2}$, s.t. $\Theta_n(u_n) \geqq \delta > 0$, $\forall n \in \mathcal{O}$. This completes the proof. □

文献

[1] J. Nagumo and J. Noda. A learning method for system identification. *IEEE Trans. Autom. Control*, 12(3):282–287, 1967.

[2] T. Hinamoto and S. Maekawa. Extended theory of learning identification. *Trans. IEE Japan*, 95-C(10):227–234, 1975. in Japanese.

[3] K. Ozeki and T. Umeda. An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties. *IEICE Trans.*, 67-A(5):126–132, 1984. in Japanese.

[4] S. C. Park and J. F. Doherty. Generalized projection algorithm for blind interference suppression in DS/CDMA communications. *IEEE Trans. Circuits Syst. II*, 44(6):453–460, 1997.

[5] M-L. R. Campos, S. Werner, and J. A. Apolinário. Constrained adaptive algorithm employing Householder transformation. *IEEE Trans. Signal Processing*, 50(9):2187–2195, 2002.

[6] I. Yamada, K. Slavakis, and K. Yamada. An efficient robust adaptive filtering algorithm based on parallel subgradient projection techniques. *IEEE Trans. Signal Processing*, 50(5):1091–1101, 2002.

[7] K. Slavakis and I. Yamada. Theory and applications of set theoretic adaptive filtering with multiple a-priori convex constraints. Technical Report SIP2004-105 (2005-01), IEICE, Jan. 2005.

[8] K. Slavakis, I. Yamada, N. Ogura, and M. Yukawa. Adaptive Projected Subgradient Method and set theoretic adaptive filtering with multiple convex constraints. In *Proceedings of the 38th Asilomar Conference on Signals, Systems and Computers*, pages 960–964, Nov. 2004.

[9] K. Slavakis, I. Yamada, and N. Ogura. The Adaptive Projected Subgradient Method over the fixed point set of strongly attracting nonexpansive mappings. Submitted for publication.

[10] I. Yamada. Adaptive Projected Subgradient Method: A unified view for projection based adaptive algorithms. *The Journal of IEICE*, 86(8):654–658, Aug. 2003. In Japanese.

[11] I. Yamada and N. Ogura. Adaptive Projected Subgradient Method for asymptotic minimization of sequence of nonnegative convex functions. *Numerical Functional Analysis and Optimization*, 25(7&8):593–617, 2004.

[12] N. Ogura and I. Yamada. Adaptive parallel outer projection algorithm based on supporting hyperplane approximation. Technical report, IEICE (DSP/EA), May 2003.

[13] M. Yukawa, K. Slavakis, and I. Yamada. Stereo echo canceler by Adaptive Projected Subgradient Method with multiple room-acoustics information. To be presented at IWAENC, Eindhoven, The Netherlands, Sept. 2005.

[14] K. Slavakis, M. Yukawa, and I. Yamada. Efficient robust adaptive beamforming by the Adaptive Projected Subgradient Method: A set-theoretic time-varying approach over multiple a priori constraints. To be presented at the technical meeting of SIgnal Processing (SIP), IEICE, Ehime University, July 2005.

[15] E. Kreyszig. *Introductory Functional Analysis with Applications.* Wiley Classics Library. John Wiley & Sons, New York, 1989.

[16] K. Goebel and S. Reich. *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*, volume 83 of *Pure and Applied Math.* Marcel Dekker, New York, 1984.

[17] H. H. Bauschke and J. M. Borwein. On projection algorithms for solving convex feasibility problems. *SIAM Review*, 38(3):367–426, Sept. 1996.

[18] H. H. Bauschke and P. L. Combettes. A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces. *Mathematics of Operations Research*, 26(2):248–264, May 2001.

[19] I. Yamada and N. Ogura. Hybrid Steepest Descent Method for variational inequality problem over the fixed point set of certain quasi-nonexpansive mappings. *Numerical Functional Analysis and Optimization*, 25(7&8):619–655, 2004.

[20] V. V. Vasin and A. L. Ageev. *Ill-Posed Problems with A Priori Information.* VSP, 1995.

[21] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems.* SIAM, Philadelphia, 1999.

[22] J-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms*, volume 1. Springer-Verlag, Berlin, 1993.