ON PARALLEL-INCREMENTAL COMBINATIONS OF LMS FILTERS THAT OUTPERFORM THE AFFINE PROJECTION ALGORITHM

Signal Processing Lab
University of São Paulo

Luiz F. O. Chamon
chamon@usp.br

Cássio G. Lopes
cassio@lps.usp.br

http://www.lps.usp.br/~chamon

ABSTRACT

The APA is one of the most celebrated data reuse AF due to its performance and computational requirements. Recently, incremental combinations of LMS filters have been used to outperform the APA with lower complexity. This combination, however, brings forth a transient/steady-state trade-off depending on the number of components. This work proposes a parallel-incremental topology with coefficients feedback that decouples these phases performances. Moreover, to address highly correlated signals, a data reusing combination is employed. Numerical experiments show that the novel combination can outperform the convergence rate and misadjustment of the APA and combinations of APAs in different scenarios.

INTRODUCTION

Data reuse:

- Increased performance, low complexity
- Trade-off performance and complexity

Combination of AFs:

- **Definition:** set of AFs aggregated by a supervisor.
- Parallel-independent topology ⇒ convergence stagnation
  - Solutions: Coefficients feedback
- Incremental topology ⇒ improves convergence, worsens misadjustment

ADAPTIVE FILTERING

$$\boldsymbol{w}_i = \boldsymbol{w}_{i-1} + \mu \boldsymbol{u}_i^* e(i) \quad \text{(LMS)}$$

$\boldsymbol{w}_i \rightarrow M \times 1$ coefficient vector at iteration i

$\mu \rightarrow$ step size

$e(i) = d(i) - \boldsymbol{u}_i \boldsymbol{w}_{i-1} \rightarrow$ output estimation error

$\boldsymbol{u}_i \rightarrow 1 \times M$ input regressor—$\mathrm{E}\,|\boldsymbol{u}(i)|^2 = \sigma_u^2$

$d(i) = \boldsymbol{u}_i \boldsymbol{w}^o + v(i) \rightarrow$ desired signal

$\boldsymbol{w}^o \rightarrow M \times 1$ vector that models the unknown system

$v(i) \rightarrow$ i.i.d. measurement noise—$\mathrm{E}\,|v(i)|^2 = \sigma_v^2$

Data reuse

$$\boldsymbol{U}_i = \begin{bmatrix} \boldsymbol{u}_i \\ \vdots \\ \boldsymbol{u}_{i-K+1} \end{bmatrix} \updownarrow K \qquad \boldsymbol{d}_i = \begin{bmatrix} d(i) \\ \cdots \\ d(i-K+1) \end{bmatrix} \updownarrow K$$

$\longleftarrow M \longrightarrow$

APA

$$\boldsymbol{w}_i = \boldsymbol{w}_{i-1} + \mu \boldsymbol{U}_i^* (\epsilon \boldsymbol{I} + \boldsymbol{U}_i \boldsymbol{U}_i^*)^{-1} \boldsymbol{e}_i$$

$$\boldsymbol{e}_i = \boldsymbol{d}_i - \boldsymbol{U}_i \boldsymbol{w}_{i-1}$$

COMBINATIONS OF ADAPTIVE FILTERS

$$\boldsymbol{w}_{n,i} = \boldsymbol{w}_{n,i-1} + \mu_n \boldsymbol{u}_{n,i}^* e_n(i)$$

$n \rightarrow$ components ; $k \rightarrow$ data

$\boldsymbol{w}_{n,i}$, μ_n, $e_n(i) \rightarrow n$-th component

$\boldsymbol{u}_{n,i}$, $d_k(i)$, $v_k(i) \rightarrow k$-th buffered data

$\{\eta_n(i)\} \rightarrow$ supervisor parameters

DR techniques

$\{\boldsymbol{u}_{n,i}, d_n(i)\} = \{\boldsymbol{u}_{i-n+1}, d(i-n+1)\}$ (*data buffering*)

$\{\boldsymbol{u}_{n,i}, d_n(i)\} = \{\boldsymbol{u}_i, d(i)\}$ (*data sharing*)

Topologies

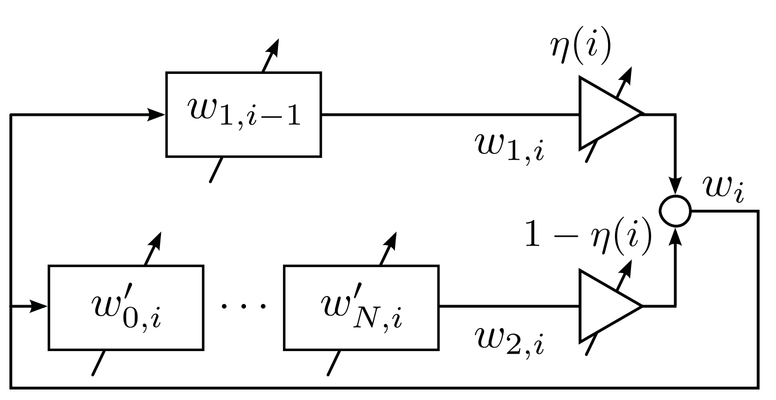
Definition 1 (*Parallel*). LMS + LMS + … *or* $N\{\text{LMS}\}$

$$\begin{aligned} \boldsymbol{w}_{n,i-1} &= \delta(i-rL)\boldsymbol{w}_{i-1} + [1-\delta(i-rL)]\boldsymbol{w}_{n,i-1} \\ \boldsymbol{w}_{n,i} &= \boldsymbol{w}_{n,i-1} + \mu_n \boldsymbol{u}_{n,i}^*[d_n(i) - \boldsymbol{u}_{n,i}\boldsymbol{w}_{n,i-1}] \\ \boldsymbol{w}_i &= \sum_{n=1}^{N} \eta_n(i) \boldsymbol{w}_{n,i} \end{aligned}$$

Definition 2 (*Incremental*). LMS · LMS · … *or* $\{\text{LMS}\}^N$

$$\begin{aligned} \boldsymbol{w}_{0,i} &= \boldsymbol{w}_{i-1} \\ \boldsymbol{w}_{n,i} &= \boldsymbol{w}_{n-1,i} + \eta_n(i)\mu_n \boldsymbol{u}_{n,i}^*[d_n(i) - \boldsymbol{u}_{n,i}\boldsymbol{w}_{n-1,i}] \\ \boldsymbol{w}_i &= \boldsymbol{w}_{N,i}, \end{aligned}$$

PARALLEL-INCREMENTAL COMBINATION



i. *Cyclic coefficients feedback*

$$\boldsymbol{w}_{n,i-1} = \delta(i-rL)\,\boldsymbol{w}_{i-1} + [1-\delta(i-rL)]\,\boldsymbol{w}_{n,i-1}$$

ii. *LMS branch*

$$\boldsymbol{w}_{1,i} = \boldsymbol{w}_{1,i-1} + \mu_1 \boldsymbol{u}_i^*\,[d(i) - \boldsymbol{u}_i \boldsymbol{w}_{1,i-1}]$$

iii. *Incremental LMS with DR branch* (DR-$\{\text{LMS}\}^N$)

$$\begin{aligned} \boldsymbol{w}'_{0,i} &= \boldsymbol{w}_{2,i-1} \\ \boldsymbol{w}'_{n,i} &= \boldsymbol{w}'_{n-1,i} + \eta_n(i)\mu_n \boldsymbol{u}_{i-k}^*[d(i-k) - \boldsymbol{u}_{i-k}\boldsymbol{w}'_{n-1,i}] \\ \boldsymbol{w}_{2,i} &= \boldsymbol{w}'_{N,i} \end{aligned}$$

iv. *Supervisor adaptation*

$$\begin{aligned} p(i) &= \beta p(i-1) + (1-\beta)[y_1(i) - y_2(i)]^2 \\ a(i) &= a(i-1) \\ &\quad + \frac{\mu_a}{p(i)+\epsilon} e(i)[y_1(i) - y_2(i)]\eta(i)[1-\eta(i)] \\ \eta(i) &= \frac{1}{1+e^{-a(i)}} \end{aligned}$$

v. *Parallel combination*

$$\boldsymbol{w}_i = \eta(i)\boldsymbol{w}_{1,i} + [1-\eta(i)]\boldsymbol{w}_{2,i}$$

BRIEF ON ANALYSIS

Mean convergence of the $\{\text{LMS}\}^N$

A.1 (*Data independence*) $\{\boldsymbol{u}_i\}$ is an i.i.d. sequence independent of $v(j)$, $\forall i, j$.

A.2 (*Supervisor separation principle*) For $e_a(i) = \boldsymbol{u}_i \tilde{\boldsymbol{w}}_{i-1}$, $\mathrm{E}[\eta_n(i)\boldsymbol{u}_i] = \mathrm{E}\,\eta_n(i)\,\mathrm{E}\,\boldsymbol{u}_i$ and $\mathrm{E}[\eta_n(i)e_a(i)] = \mathrm{E}\,\eta_n(i)\,\mathrm{E}\,e_a(i)$.

$$\mathrm{E}\,\tilde{\boldsymbol{w}}_i = [\boldsymbol{I} - \mathrm{E}\,\bar{\mu}\boldsymbol{R}_{\boldsymbol{u}} + \boldsymbol{\mathcal{M}}]\,\mathrm{E}\,\tilde{\boldsymbol{w}}_{i-1}$$

$$\boldsymbol{\mathcal{M}} = \sum_{k=2}^{N} \mathrm{E}(-\boldsymbol{u}_i^*\boldsymbol{u}_i)^k\,\mathrm{E}\left[\sum \binom{\{\eta_n\mu_n\}}{k}\right]$$

A.3 (*Small μ and Gaussian data*) $\eta_n\mu_n \ll 1$ and u_i is a Gaussian vector

$$\boldsymbol{R}_{\boldsymbol{u}} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^* \rightarrow \bar{\boldsymbol{w}}_i = \boldsymbol{Q}\tilde{\boldsymbol{w}}_i$$

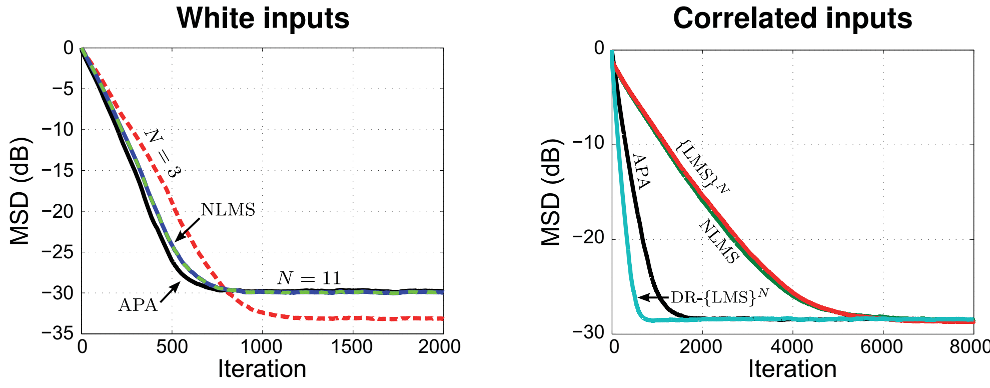
$$\mathrm{E}\,\bar{\boldsymbol{w}}_i = \boldsymbol{A} \cdot \mathrm{E}\,\bar{\boldsymbol{w}}_{i-1}$$

$$\boldsymbol{A} = \boldsymbol{I} - \mathrm{E}\,\bar{\mu}\boldsymbol{\Lambda} + \sum_{m \neq n} \mu_m\mu_n\,\mathrm{E}\,\eta_m\eta_n\left[\boldsymbol{\Lambda}\,\mathrm{Tr}(\boldsymbol{\Lambda}) + \gamma\boldsymbol{\Lambda}^2\right]$$

SIMULATIONS

APA, NLMS, $\{\text{LMS}\}^N$, and DR-$\{\text{LMS}\}^N$: $M = 100$, $K = 10$, $\eta_n = 1$, $\mu_n = \mu$

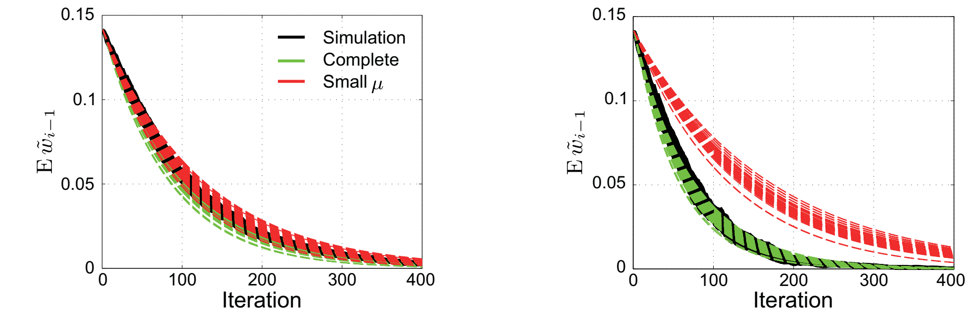
White inputs | Correlated inputs



APA: $\mu = 0.3$; **NLMS:** $\mu = 1$; and $\{\text{LMS}\}^N$: $\mu = 3 \cdot 10^{-3}$, $N = 20$

APA: $\mu = 0.015$; **NLMS:** $\mu = 0.9$; $\{\text{LMS}\}^N$: $\mu = 1.5 \cdot 10^{-3}$, $N = 30$; and **DR-$\{\text{LMS}\}^N$**: $\mu = 2 \cdot 10^{-3}$, $N = 15$

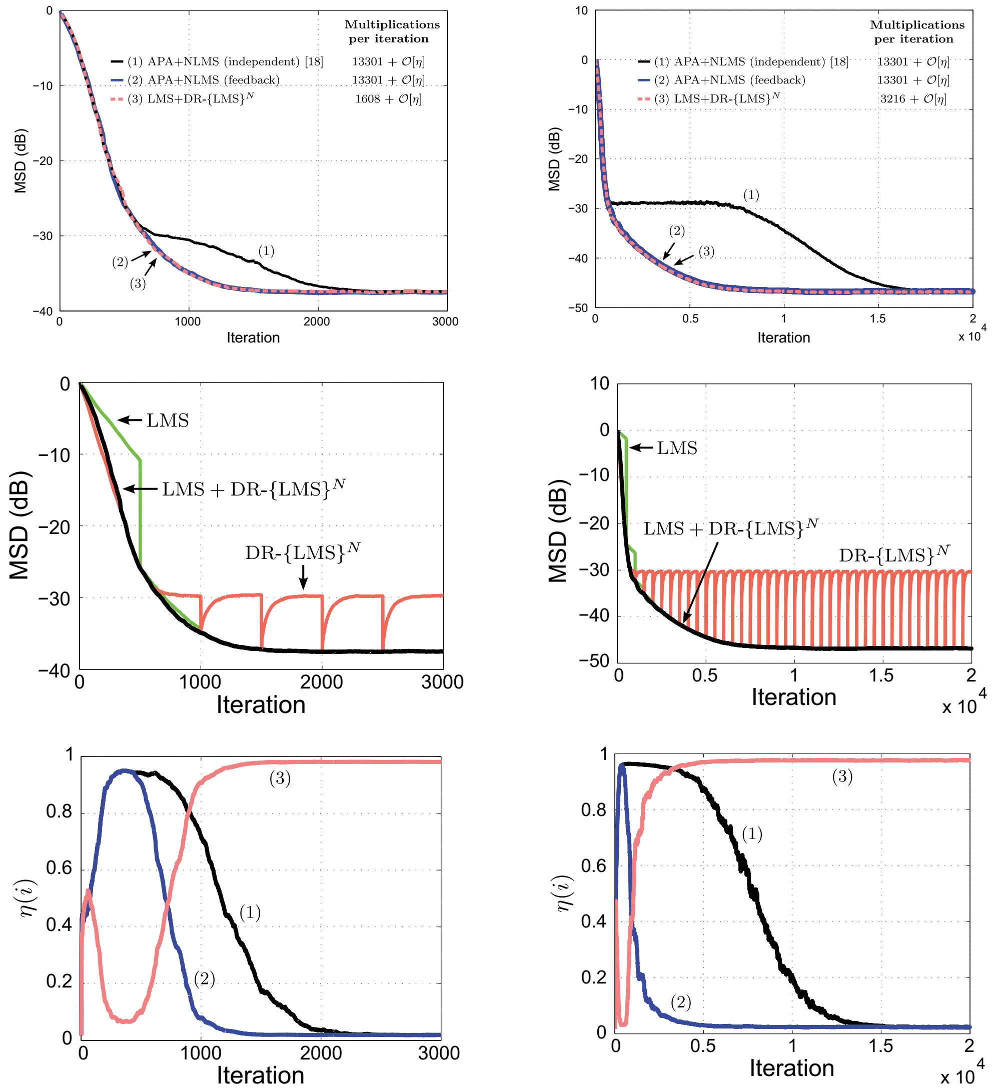
Mean convergence models: $M = 50$, $\eta_n = 1$, white inputs



$\mu_n = 0.01$, $N = 75$ | $\mu_n = 0.02$, $N = 75$

Parallel-incremental combination: $M = 100$, $K = 10$, $\eta_n\mu_n = \mu$

White inputs | Correlated inputs



APA+NLMS: $\mu_{\text{APA}} = 0.3$, $\mu_{\text{NLMS}} = 0.3$; **LMS+DR-$\{\text{LMS}\}^N$**: $\mu_{\text{LMS}} = 3 \cdot 10^{-3}$, $\mu_{\text{DR-}\{\text{LMS}\}^N} = 6 \cdot 10^{-3}$, $N = 7$

APA+NLMS: $\mu_{\text{APA}} = 1$, $\mu_{\text{NLMS}} = 0.04$; **LMS+DR-$\{\text{LMS}\}^N$**: $\mu_{\text{LMS}} = 4 \cdot 10^{-4}$, $\mu_{\text{DR-}\{\text{LMS}\}^N} = 2 \cdot 10^{-3}$, $N = 15$