A DATA REUSAGE ALGORITHM BASED ON INCREMENTAL COMBINATION OF LMS FILTERS

Polytechnic School
Dept. of Electronics Systems
São Paulo — Brazil

Luiz F. O. Chamon chamon@usp.br

Humberto F. Ferro hferro@usp.br

Cássio G. Lopes cassio@lps.usp.br

ABSTRACT

This work introduces a new data reusage algorithm based on the incremental combination of LMS filters which is able to outperform the Affine Projection Algorithm (APA), another well-known data reusage AF. First, the true gradient data reusage LMS (TRUE-LMS) is shown to be equivalent to an over-regularized APA. In the sequel, an incremental counterpart of its recursion is motivated by distributed optimization and adaptive networks. Simulations in different scenarios show the efficiency of the proposed data reusage algorithm, which is able to either match or outperform the APA in the mean-square sense at lower computational complexity.

INTRODUCTION

Data reusage:

- **Definition:** using either a single data pair ($\{u, d\}$) more than once or a set of past pairs.
- **Goal:** improve convergence by further exploiting the available data.
- **The Affine Projection Algorithm:**
  - One of the most celebrated data reusage AFs due to its superior performance (e.g., for *speech echo cancellation*, the APA has a performance comparable to Fast RLS filter with 3 times less operations).
  - Other data reusage algorithms are rarely compared to the APA.

Combination of AFs:

- **Definition:** set of AFs combined by a mixing parameter.
- Used when the accurate design of a single filter is difficult or the resulting algorithm's complexity is too high

PROBLEM FORMULATION

Adaptive filters

$$\boxed{w_{n,i} = w_{n,i-1} + \mu_n p_n}$$

$w_{n,i} \rightarrow M \times 1$ coefficients estimate of the n^{th} filter at iteration i

$\mu_n \rightarrow n^{th}$ filter's step size

$p_n = -B_n \nabla^* J(w_{n,i-1}) \rightarrow$ update direction

$J(w_{n,i-1}) \rightarrow$ cost function—usually $\mathrm{E}\,|e_n(i)|^2$

$e_n(i) = d_n(i) - u_{n,i} w_{n,i-1} \rightarrow$ estimation error

$u_i \rightarrow 1 \times M$ input regressor at iteration i

$d(i) = u_i w^o + v(i) \rightarrow$ desired signal at iteration i

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

$v(i) \rightarrow$ i.i.d. measurement noise

Data reusage

Problem: finding a coefficients estimate w_i that minimizes the mean-square error based on a previous estimate w_{i-1} and the data set $\{U_i, d_i\}$.

$$U_i = \begin{bmatrix} u_i \\ \vdots \\ u_{i-K+1} \end{bmatrix} \updownarrow K \qquad d_i = \begin{bmatrix} d(i) \\ \dots \\ d(i-K+1) \end{bmatrix} \updownarrow K$$

$\longleftarrow M \longrightarrow$

Algorithms

$$w_i = w_{i-1} + \mu u_i^* e(i) \qquad \text{(LMS)}$$

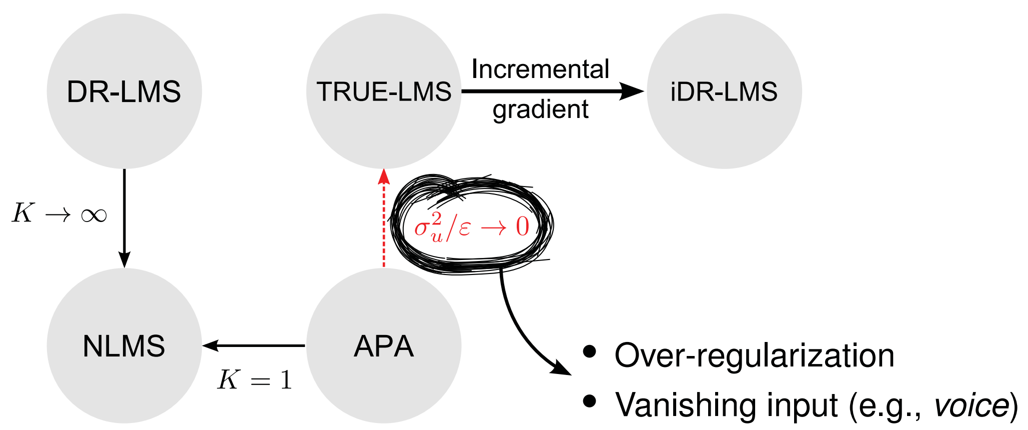
$$w_i = w_{i-1} + \mu \frac{u_i^*}{\|u_i\|^2 + \epsilon} e(i) \qquad \text{(NLMS)}$$

$$\begin{aligned} w_{0,i} &= w_{i-1} \\ w_{n,i} &= w_{n-1,i} + \mu u_i^* [d(i) - u_i w_{n-1,i}] \qquad \text{(DR-LMS)} \\ w_i &= w_{K,i} \end{aligned}$$

$$w_i = w_{i-1} + \mu U_i^* e_i \qquad \text{(TRUE-LMS)}$$

$$w_i = w_{i-1} + \mu U_i^* (\epsilon I + U_i U_i^*)^{-1} e_i \qquad \text{(APA)}$$

APA AND TRUE-LMS



APA $\rightarrow$ TRUE-LMS

Focusing on the APA update and choosing ($\mu = \mu' \epsilon$)

$$\lim_{\sigma_u^2/\epsilon \to 0} \mu' \epsilon U_i^* (\epsilon I + U_i U_i^*)^{-1} e_i =$$

$$\mu' U_i^* \left[\lim_{\sigma_u^2/\epsilon \to 0} \epsilon (\epsilon I + U_i U_i^*)^{-1} \right] e_i =$$

$$\mu' U_i^* \left[I + \lim_{\sigma_u^2/\epsilon \to 0} \epsilon^{-1} U_i U_i^* \right]^{-1} e_i$$

Definition 1. $f(x)$ is said to be *superlinear* if $\lim_{x \to \pm\infty} \frac{|x|}{f(x)} = 0$.

Theorem 1. *Take* $\mathbf{u}$ *a zero-mean w.s.s. RV with* $\sigma_u^2 = \mathrm{E}\,|\mathbf{u}|^2$. *Over any path where* σ_u^2/ϵ *is a superlinear function of* $1/M$,

$$\lim_{\sigma_x^2/\epsilon \to 0} \epsilon^{-1} U_i U_i^* = 0 \text{ (a.s.).} \qquad (1)$$

Proof. By the strong law of large numbers,

$$\lim_{M \to +\infty} \frac{UU^*}{M} = \begin{bmatrix} \mathrm{E}\, u_i u_i^* & \cdots & \mathrm{E}\, u_i u_{i-M+1}^* \\ \vdots & \ddots & \vdots \\ \mathrm{E}\, u_{i-M+1} u_i^* & \cdots & \mathrm{E}\, u_{i-M+1} u_{i-M+1}^* \end{bmatrix} \text{ (a.s.)}$$

Every element of this autocorrelation matrix is upper bounded by the diagonal elements (σ_u^2). It is, therefore, sufficient to show that (1) would converge for the special case of σ_u^2. Indeed,

$$\lim_{\sigma_u^2/\epsilon \to 0} \frac{u_i u_i^*}{\epsilon} = \lim_{(M, \sigma_u^2/\epsilon) \to (+\infty, 0)} M \frac{u_i u_i^*}{M \epsilon} = \lim_{(M, \sigma_u^2/\epsilon) \to (+\infty, 0)} M \frac{\sigma_u^2}{\epsilon}$$

Taking $\sigma_u^2/\epsilon = f(1/M)$ yields

$$\lim_{(M, \sigma_u^2/\epsilon) \to (+\infty, 0)} M \frac{\sigma_u^2}{\epsilon} = \lim_{M \to +\infty} \frac{f(1/M)}{1/M} = 0$$

for any superlinear function f. □

$$\boxed{\therefore \text{ APA } \xrightarrow{\sigma_u^2/\epsilon \to 0} w_i = w_{i-1} + \mu' U_i^* e_i}$$

THE INCREMENTAL DR-LMS

Distributed estimation $\begin{cases} \nabla J_n(w) & \text{(true gradient)} \\ \nabla J_n(w_{n-1}) & \text{(incremental gradient)} \end{cases}$

$J(w) = \sum_n J_n(w)$

iDR-LMS: the incremental counterpart of TRUE-LMS

$$\boxed{\begin{aligned} w_{0,i} &= w_{i-1} \\ w_{n,i} &= w_{n-1,i} + \mu_n u_k^* [d(k) - u_k w_{n-1,i}] \\ w_i &= w_{N,i} \end{aligned}}$$

where $k = i - \text{mod}\,(n-1, K)$ and $n = 1, \dots, N$

– For $N = K$ the DR-LMS from Schnaufer & Jenkings is recovered.

✓ Matches or outperforms the APA for $N \ll K^2$, i.e., at lower computational cost

✓ LMS-based: robust and stable (even in reduced precision environments)

✓ Can be seen as a particular case of the adaptive network algorithm INC-LMS where all nodes statistics are the same

× Cannot be parallelized

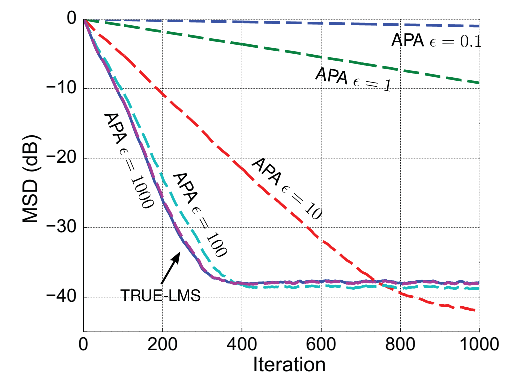
Complexity comparison

- iDR-LMS: $O(NM)$
- APA: $O(K^2 M)$
- Fast APA: $O(KM) + O(3K^2)$

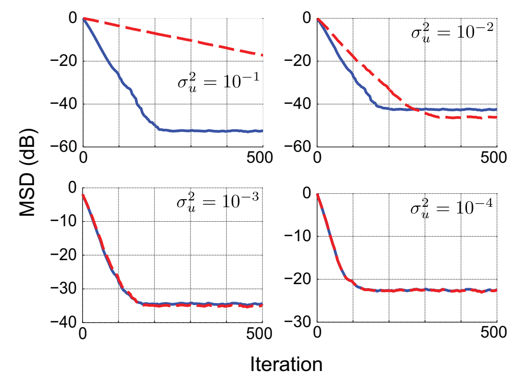
SIMULATIONS

APA $\rightarrow$ TRUE-LMS

Over-regularization
$M = 16$, $K = 4$,
$\sigma_u^2 = 1$, $\sigma_v^2 = 10^{-3}$,
$\mu' = 0.05$, $\mu_{APA} = \mu' \epsilon$



Vanishing input
$M = 16$, $K = 4$,
$\sigma_v^2 = 10^{-6}$, $\epsilon = 0.1$,
$\mu_{APA} = \mu_{TRUE} = 0.3/\sigma_u^2$.



iDR-LMS $\times$ APA

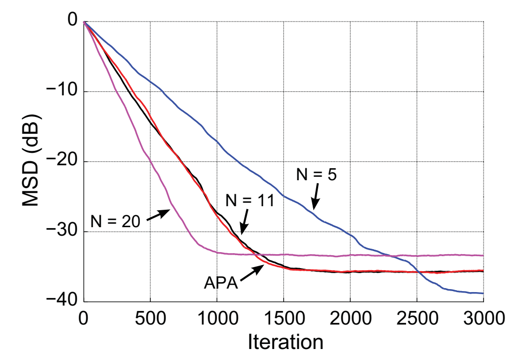
$M = 100$, $w^o = \text{col}\{1\}/\sqrt{M}$, $\mu_0 = 0.05$, $\mu_{APA} = \mu_0$,
$\mu_n = \mu_0 / M \sigma_u^2$, $\{x, v\} \sim$ Gaussian, $\sigma_x^2 = 1$, $\sigma_v^2 = 10^{-3}$.

White input: $u(i) = x(i)$

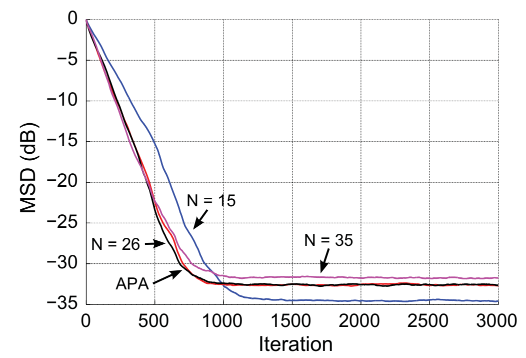
Correlated input: $u(i) = \alpha\, u(i-1) + \sqrt{1 - \alpha^2}\, x(i)$, $\alpha = 0.95$.

Non-stationary: $w^o(i) = w^o(i-1) + q$, $\sigma_q^2 = 10^{-6}$, $\alpha = 0.95$, $\mu_{APA} = 0.1$, $\mu_n = 0.003$.

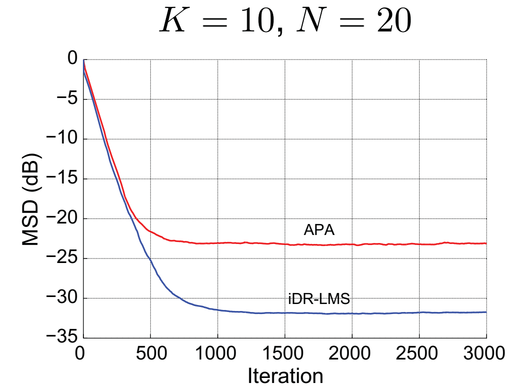
White input – $K = 10$



White input – $K = 20$



Correlated input
$K = 10$, $N = 20$



Non-stationary
$M = 60$, $K = 10$, $N = 20$