# CONVEX COMBINATION OF CONSTRAINT VECTORS FOR SET-MEMBERSHIP AFFINE PROJECTION ALGORITHMS

*Tadeu N. Ferreira\*, Wallace A. Martins†‡, Markus V. S. Lima‡, Paulo S. R. Diniz‡*

\* Universidade Federal Fluminense (UFF), Niteroi, Brazil
† Université du Luxembourg (SnT/LU), Luxembourg, Luxembourg
‡ Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil

## ABSTRACT

Set-membership affine projection (SM-AP) adaptive filters have been increasingly employed in the context of online data-selective learning. A key aspect for their good performance in terms of both convergence speed and steady-state mean-squared error is the choice of the so-called constraint vector. Optimal constraint vectors were recently proposed relying on convex optimization tools, which might sometimes lead to prohibitive computational burden. This paper proposes a convex combination of simpler constraint vectors whose performance approaches the optimal solution closely, utilizing much fewer computations. Some illustrative examples confirm that the sub-optimal solution follows the accomplishments of the optimal one.

***Index Terms***— adaptive filtering, set-membership, affine projection, combination of adaptive filters, optimization

## 1. INTRODUCTION

Adaptive filters incorporating data-selection are attracting growing interest thanks to their ability to reduce the computational load of learning processes [1–12]. *Set-membership affine projection* (SM-AP) algorithms [4] are among successful examples within the data-selective class of adaptive filters. The incoming data is evaluated before implementing a new update of the parameter vector associated with the adaptive filter, and for that, each SM-AP algorithm requires a *constraint vector* (CV) to compute the parameter vector when updating [13].

The original SM-AP algorithm [4] uses the so-called *simple choice* CV (SC-CV), which is usually the preferred choice in many variations of the SM-AP algorithm—see, for instance, [5]. In addition to the computational simplicity, such a preference is justified by the good steady-state mean-squared error (MSE) performance induced by the SC-CV [14]. Other heuristically chosen CVs include *fixed modulus error-based* CV (FMEB-CV) [15–17] and *exponential decay* CV (ED-CV) [17, 18]. Besides the good convergence speed and low computational burden induced by those closed-form CVs, the

heuristics employed to define them find reasonable geometrical justification—see, for instance, [17].

SM-AP algorithms employing heuristic CVs (such as SC-CV, FMEB-CV, and ED-CV) generally do not solve the original SM-AP optimization problem. Indeed, the SM-AP updating process should, in principle, lead to parameter vectors that correspond to the *closest points* to a polytope defined by pairs of input and desired signals. Instead, most heuristic CVs lead to parameter vectors lying on the *closest boundary* of the aforementioned polytope, thus weakening the original optimization problem by replacing "the closest point" with "a point belonging to the closest boundary". The work [19] was the first to solve the original SM-AP optimization problem by proposing an *optimal* CV, which is the solution to a box-constrained strictly convex quadratic problem. The work [20] further extended the results from [19] to the set-membership *proportionate* affine projection algorithm (SM-PAPA) [21].

Although the optimal CVs proposed in [19, 20] can be applied in a variety of applications, their corresponding computational burden is much higher than the heuristic CVs'. This can be a big issue, particularly in those online applications where the rate of data acquisition is really high, severely constraining the time to perform the computations between successive updates. The main goal of this work is to join the heuristic and the optimal approaches in order to yield sub-optimal solutions that are easy to compute with few numerical operations while also preserving the good properties of the optimal solution. This goal is accomplished via a convex combination of two heuristic CVs (e.g., SC-CV and ED-CV) with an optimized weight following the methodology employed by the optimal CV in [19, 20].

This work is organized as follows. Section 2 describes the SM-AP algorithm, including its variant SM-PAPA; the closed-form expressions for the heuristic CVs and the convex optimization problem underlying the calculation of the optimal CVs are also provided. Section 3 contains the main proposal of this paper: the closed-form expression for the optimal convex combination of CVs; it is also shown the equivalence of this approach to the well-known convex combination of adaptive filters, and some remarks are made related to the computational complexity of the proposal. Section 4

describes a comprehensive set of simulation results assessing the performance of the sub-optimal proposal and comparing it against heuristic and optimal approaches in a variety of scenarios and considering different figures of merit, such as MSE, misalignment, convergence speed, and percentage of updates. Some concluding remarks are drawn in Section 5.

## 2. SET-MEMBERSHIP AFFINE PROJECTION

The affine projection (AP) algorithm can be thought as a generalization of the normalized least-mean squares (NLMS) algorithm when low computational complexity and fast convergence are both required [13]. The generalization is obtained by reusing input data vectors from previous iterations. Further reduction of the computational burden is achieved when AP employs a data-selective scheme based on set-membership (SM), while maintaining or even increasing the convergence speed [4].

Let the parameter vector of the adaptive filter at iteration $k \in \mathbb{N}$ be $\mathbf{w}(k) \in \mathbb{R}^{N+1}$, where $N \in \mathbb{N}$ is the order of the filter. Let also $\mathbf{x}(k) \in \mathbb{R}^{N+1}$, $\mathbf{d}(k)$, and $\mathbf{e}(k) \in \mathbb{R}^{L+1}$ denote the input, desired, and error vectors, where $L \in \mathbb{N}$ is the data-reuse factor. One can then build the input matrix $\mathbf{X}(k) \in \mathbb{R}^{(N+1) \times (L+1)}$, whose $l^{\text{th}}$ column is $\mathbf{x}(k-l)$, with $l \in \mathcal{L} \triangleq \{0, 1, \cdots, L\}$; these variables are related as follows: $\mathbf{e}(k) = \mathbf{d}(k) - \mathbf{X}^{\text{T}}(k)\mathbf{w}(k)$. The first entry of the error vector $\mathbf{e}(k)$ is called *a priori error* and denoted as $e_0(k)$. The SM strategy is implemented by comparing the magnitude of the *a priori* error with a predefined threshold $\bar{\gamma} \in \mathbb{R}_+$. By denoting the CV as $\boldsymbol{\gamma}(k) \in \mathbb{R}^{L+1}$, which must satisfy $\|\boldsymbol{\gamma}(k)\|_\infty \leq \bar{\gamma}$, one can write the SM-PAPA recursion as [21]

$$\mathbf{w}(k+1) = \begin{cases} \mathbf{w}(k) + \mathbf{G}(k)\mathbf{X}(k)\mathbf{S}(k)[\mathbf{e}(k) - \boldsymbol{\gamma}(k)], & |e_0(k)| > \bar{\gamma}, \\ \mathbf{w}(k), & \text{otherwise,} \end{cases}$$

(1)

where $\mathbf{G}(k) \in \mathbb{R}^{(N+1) \times (N+1)}$ is a diagonal proportionate matrix chosen as in [21], $\mathbf{S}(k) \triangleq \left[ \mathbf{X}^{\text{T}}(k)\mathbf{G}(k)\mathbf{X}(k) + \delta\mathbf{I} \right]^{-1} \in \mathbb{R}^{(L+1) \times (L+1)}$, and $\delta \in \mathbb{R}_+$ is a regularization factor used to avoid numerical issues due to matrix inversion. Note that $\mathbf{S}(k)$ is a positive-definite matrix. The standard SM-AP [4,13] algorithm results from setting $\mathbf{G}(k)$ as the identity matrix.

The definition of the CV $\boldsymbol{\gamma}(k)$ determines the convergence behavior of the algorithm. By using geometrical reasoning, it is possible to conceive different CV definitions [17]. The following ones describe the main heuristic choices, namely: SC-CV, ED-CV, and FMEB-CV. In these definitions, it is assumed that the $l^{\text{th}}$ entry of vectors $\boldsymbol{\gamma}(k)$ and $\mathbf{e}(k)$ are respectively denoted as $\gamma_l(k)$ and $e_l(k)$, for all $l \in \mathcal{L}$.

**Definition 1** (SC-CV). *The SC-CV proposed in [4] and analyzed in [14] is defined as*

$$\gamma_l(k) \triangleq \begin{cases} \bar{\gamma} \operatorname{sign}[e_l(k)] & \textit{if } l = 0, \\ e_l(k) & \textit{for } l \in \mathcal{L} \setminus \{0\}. \end{cases}$$

(2)

**Definition 2** (ED-CV). *The ED-CV proposed in [17] requires that $\bar{\gamma} \leq 1$ and is given by*

$$\gamma_l(k) \triangleq \bar{\gamma}^{l+1} \operatorname{sign}[e_l(k)], \textit{ for } l \in \mathcal{L}.$$

(3)

**Definition 3** (FMEB-CV). *The FMEB-CV proposed in [15] and analyzed in [16] is defined as*

$$\gamma_l(k) \triangleq \bar{\gamma} \operatorname{sign}[e_l(k)], \textit{ for } l \in \mathcal{L}.$$

(4)

The work in [19] (respect. [20]) solved the original SM-AP (respect. SM-PAPA) problem considering that both the parameter vector and the CV are optimization variables. In this context, the optimal CV is defined as follows.

**Definition 4** (Optimal CV). *The optimal CV proposed in [19, 20] is obtained by solving the following strictly convex box-constrained quadratic problem:*

$$\begin{aligned} \underset{\boldsymbol{\gamma}(k)}{\textit{minimize}} \quad & [\boldsymbol{\gamma}(k) - \mathbf{e}(k)]^{\text{T}}\mathbf{S}(k)[\boldsymbol{\gamma}(k) - \mathbf{e}(k)], \\ \textit{subject to} \quad & \|\boldsymbol{\gamma}(k)\|_\infty \leq \bar{\gamma}. \end{aligned}$$

(5)

## 3. PROPOSED CONVEX COMBINATION OF CVS

Although the optimal CV can be numerically computed by solving the convex problem in Definition 4 within an $(L+1)$-dimensional space, using it induces a higher computational burden to online systems. Therefore, solving such an optimization problem could be unrealistic in applications working at high data rates. A compromise solution can be achieved by mixing the heuristic closed-form strategies in Definitions 1, 2, or 3 with the convex optimization strategy in Definition 4 to come up with sub-optimal CVs by performing a search on a one-dimensional subspace, entailing a lower computational complexity. This can be achieved by considering convex combinations of closed-form CVs, as in the following definition.

**Definition 5** (Convex combination — CC). *Given two CVs $\boldsymbol{\gamma}_{\text{a}}(k), \boldsymbol{\gamma}_{\text{b}}(k) \in \mathbb{R}^{L+1}$, their* convex combination *is*

$$\boldsymbol{\gamma}_{\text{c}}(k) = \alpha_k \boldsymbol{\gamma}_{\text{a}}(k) + (1 - \alpha_k)\boldsymbol{\gamma}_{\text{b}}(k), \quad \alpha_k \in [0, 1].$$

(6)

Notice that $\boldsymbol{\gamma}_{\text{c}}(k)$ is itself a CV. Indeed, as both $\boldsymbol{\gamma}_{\text{a}}(k)$, $\boldsymbol{\gamma}_{\text{b}}(k)$ belong to the ball $\{\boldsymbol{\gamma} \in \mathbb{R}^{L+1} \mid \|\boldsymbol{\gamma}\|_\infty \leq \bar{\gamma}\}$, which is a convex set, then any convex combination of them also belongs to the infinity-norm ball, making $\boldsymbol{\gamma}_{\text{c}}(k)$ a valid CV associated with the predefined error threshold $\bar{\gamma}$.

The proposed sub-optimal CV is the optimal CC of two predefined CVs, as detailed in the following definition.

**Definition 6** (CC-CV). *Given two CVs $\boldsymbol{\gamma}_{\text{a}}(k) \neq \boldsymbol{\gamma}_{\text{b}}(k) \in \mathbb{R}^{L+1}$, the* convex combination constraint vector (CC-CV) *is*

$$\boldsymbol{\gamma}^{\text{CC-CV}}(k) = \alpha_k^* \boldsymbol{\gamma}_{\text{a}}(k) + (1 - \alpha_k^*)\boldsymbol{\gamma}_{\text{b}}(k),$$

(7)

*in which $\alpha_k^*$ is the solution to the optimization problem*

$$\begin{aligned} \underset{0 \leq \alpha_k \leq 1}{\textit{minimize}} \quad & [\boldsymbol{\gamma}_{\text{c}}(k) - \mathbf{e}(k)]^{\text{T}}\mathbf{S}(k)[\boldsymbol{\gamma}_{\text{c}}(k) - \mathbf{e}(k)], \\ \textit{subject to} \quad & \boldsymbol{\gamma}_{\text{c}}(k) = \alpha_k \boldsymbol{\gamma}_{\text{a}}(k) + (1 - \alpha_k)\boldsymbol{\gamma}_{\text{b}}(k). \end{aligned}$$

(8)

The optimization problem in (8) admits a closed-form solution as follows.

**Proposition 1** (Optimal weight). *Given the number*

$$\beta_k = \frac{[\mathbf{e}(k) - \boldsymbol{\gamma}_{\mathrm{b}}(k)]^{\mathrm{T}}\mathbf{S}(k)[\boldsymbol{\gamma}_{\mathrm{a}}(k) - \boldsymbol{\gamma}_{\mathrm{b}}(k)]}{[\boldsymbol{\gamma}_{\mathrm{a}}(k) - \boldsymbol{\gamma}_{\mathrm{b}}(k)]^{\mathrm{T}}\mathbf{S}(k)[\boldsymbol{\gamma}_{\mathrm{a}}(k) - \boldsymbol{\gamma}_{\mathrm{b}}(k)]}, \quad (9)$$

*the solution to the optimization problem in* (8) *is*

$$\alpha_k^* = \beta_k\left[u(\beta_k) - u(\beta_k - 1)\right] + u(\beta_k - 1), \quad (10)$$

*where $u(\cdot)$ is the standard unit-step function.*

*Proof.* One can replace the equality constraint in (8) into the cost-function to get the convex quadratic function

$$\begin{aligned}
C(\alpha_k) =& \alpha_k^2[\boldsymbol{\gamma}_{\mathrm{a}}(k) - \boldsymbol{\gamma}_{\mathrm{b}}(k)]^{\mathrm{T}}\mathbf{S}(k)[\boldsymbol{\gamma}_{\mathrm{a}}(k) - \boldsymbol{\gamma}_{\mathrm{b}}(k)] \\
& + 2\alpha_k[\boldsymbol{\gamma}_{\mathrm{b}}(k) - \mathbf{e}(k)]^{\mathrm{T}}\mathbf{S}(k)[\boldsymbol{\gamma}_{\mathrm{a}}(k) - \boldsymbol{\gamma}_{\mathrm{b}}(k)] \\
& + [\boldsymbol{\gamma}_{\mathrm{b}}(k) - \mathbf{e}(k)]^{\mathrm{T}}\mathbf{S}(k)[\boldsymbol{\gamma}_{\mathrm{b}}(k) - \mathbf{e}(k)],
\end{aligned}$$
$$(11)$$

which must be minimized for $\alpha_k \in [0, 1]$. Notice that $C(\beta_k) \leq C(\alpha_k)$ for all $\alpha_k \in \mathbb{R}$. Thus, when $\beta_k \in [0, 1]$, then $\alpha_k^* = \beta_k$ is the solution to the problem in (8). If $\beta_k < 0$, then $C(\alpha_k)$ is a monotonic increasing function within the interval $[0, 1]$, since it lies entirely on the right branch of the parabola. Therefore, the solution to (8) is $\alpha_k^* = 0$. Similarly, for $\beta_k > 1$, the solution to (8) is $\alpha_k^* = 1$. □

The next result shows that using an SM-AP filter along with the proposed CC-CV is equivalent to convexically combining two SM-AP filters as follows.

**Proposition 2.** *An SM-AP filter using the proposed CC-CV in* (7) *is equivalent to the convex combination of two SM-AP filters, one of them with CV $\boldsymbol{\gamma}_{\mathrm{a}}(k)$, and the other with $\boldsymbol{\gamma}_{\mathrm{b}}(k)$.*

*Proof.* As presented in [22], the coefficient vector $\mathbf{w}(k)$ of the convexically combined adaptive filter is given by

$$\mathbf{w}(k) = \lambda_k\mathbf{w}_{\mathrm{a}}(k) + (1 - \lambda_k)\mathbf{w}_{\mathrm{b}}(k), \quad (12)$$

where $\lambda_k \in [0, 1]$. Assume that both $\mathbf{w}_{\mathrm{a}}(k)$ and $\mathbf{w}_{\mathrm{b}}(k)$ are updated by an SM-AP algorithm, with different CVs $\boldsymbol{\gamma}_{\mathrm{a}}(k)$ and $\boldsymbol{\gamma}_{\mathrm{b}}(k)$. Then, $\mathbf{w}_z(k+1) = \mathbf{w}_z(k) + \mathbf{X}(k)\mathbf{S}(k)(\mathbf{e}_z(k) - \boldsymbol{\gamma}_z(k)), z \in \{a, b\}$. By using (12) and considering the equivalent error as the convex combination of the errors as in [22], i.e. $\mathbf{e}(k) = \lambda_k\mathbf{e}_{\mathrm{a}}(k) + (1 - \lambda_k)\mathbf{e}_{\mathrm{b}}(k)$, then

$$\mathbf{w}(k + 1) = \mathbf{w}(k) + \mathbf{X}(k)\mathbf{S}(k)[\mathbf{e}(k) - \boldsymbol{\gamma}_{\mathrm{c}}(k)], \quad (13)$$

for $\lambda_k = \alpha_k^*$. □

The result in Proposition 2 does not hold for the SM-PAPA proposed in [21]. Indeed, different parameter vectors induce different proportionate matrices, breaking the symmetry of some terms in the expression for the convex combination in (12). Nonetheless, one could employ a unified strategy

for forming the proportionate matrices appearing in each filter so that they are equal. If this modification is considered, then Proposition 2 would also hold for the convex combination of SM-PAPA filters.

The optimal CV for the SM-AP proposed in [19] is usually implemented via interior-point methods. As discussed in [23], the computational complexity of common interior-point methods are in-between $\mathcal{O}\left\{(L + 1)^{3.5}\right\}$ and $\mathcal{O}\left\{(L + 1)^4\right\}$. In [20], the gradient projection (GP) method is used, which presents a lower computational complexity of $\mathcal{O}\left\{(L + 1)^3\right\}$, as shown in [24]. The proposed CC-CV requires an even lower computational complexity for its computation in the order of $\mathcal{O}\left\{(L + 1)^2\right\}$, which is the complexity for computing $\beta_k$ in (9). The complexity of computing any of the heuristic CVs presented in Definitions 1-3 is negligible, as it involves only trivial multiplications (no multiplication is needed and the sign of a precomputed parameter is adjusted).

## 4. SIMULATION RESULTS

The proposed CC-CV in Definition 6 is tested in a *system identification scenario* considering that $\boldsymbol{\gamma}_{\mathrm{a}}(k)$ is the SC-CV of Definition 1 and $\boldsymbol{\gamma}_{\mathrm{b}}(k)$ is the ED-CV of Definition 2. This combination is tested and compared with the optimal CV of Definition 4 as well as to both the pure SC-CV and pure ED-CV versions of the SM-AP and SM-PAPA.

In all simulated scenarios, a measurement noise with variance $\sigma_{\mathrm{n}}^2 = 10^{-3}$ is added to the clean desired signal. In addition, 4000 input samples are filtered through a 16-tap finite-impulse response system to be identified in $R = 1000$ independent Monte-Carlo runs. Three different experiments are considered:

- **Exp. 1**: a $4^{\mathrm{th}}$-order autoregressive unit-power input signal, denominated here as AR(4) input, is generated by the response of a recursive filter to a white input, i.e.: $x(k) = 0.95x(k - 1) + 0.19x(k - 2) + 0.09x(k - 3) - 0.5x(k - 4) + w(k)$, where $w(k)$ is the realization of a white random sequence. The unknown system response is $\mathbf{h}(k) = [h_0(k) \cdots h_{15}(k)]^{\mathrm{T}}$, which models a *dispersive* setup given as

  - $k < 800$: $h_m(k) = 1$, for $m \in \{0, 1, \ldots, 15\}$;
  - $k \geq 800$: $h_m(k) = 2$, for $m \in \{0, 1, \ldots, 7\}$, whereas $h_m(k) = 1$, for $m \in \{8, 9, \ldots, 15\}$.

- **Exp. 2**: the AR(4) signal of Exp. 1 is the input of a *sparse* system $\mathbf{h}(k)$ defined as

  - $k < 800$: $h_m(k) = 10^{-4}$, for $m \in \{0, 1, \ldots, 15\}\backslash\{3\}$, and $h_3(k) = 1$;
  - $k \geq 800$: $h_m(k) = 10^{-4}$, for $m \in \{0, 1, \ldots, 15\}\backslash\{3, 10\}$, and $h_3(k) = h_{10}(k) = 1$.

- **Exp. 3**: a unit-power white random sequence is the input of the same sparse system of Exp. 2.
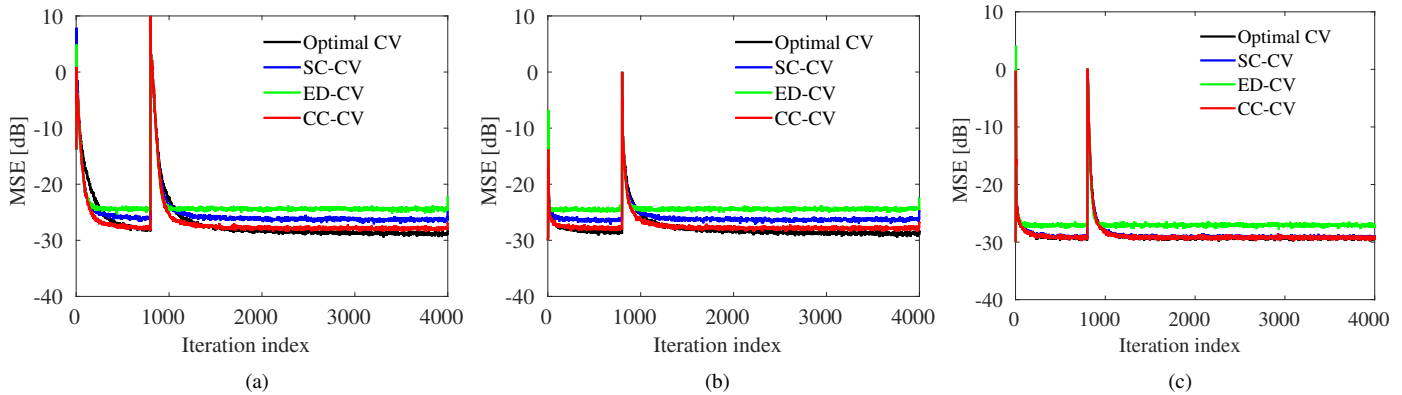
**Fig. 1**. MSE for: (a) **Exp. 1**; (b) **Exp. 2**; (c) **Exp. 3**.

The setup of a white input for a dispersive system is not considered, since it is known that the SC-CV is asymptotically optimal for this case [19].

The SM-AP algorithm (employed in **Exp. 1**) and the SM-PAPA (employed in **Exps. 2** and **3**) use $L = 3$ as the reuse factor for all CVs. In addition, the error threshold is given as $\bar{\gamma} = \sqrt{4\sigma_n^2}$, the regularization factor is $\delta = 5 \cdot 10^{-7}$, and the coefficients are initialized as $\mathbf{w}(0) = \mathbf{0} \in \mathbb{R}^{16}$.

Three figures of merit are considered: (i) MSE, defined as: $\text{MSE}(k) = \frac{1}{R} \sum_{r=1}^{R} e_{0,r}^2(k)$, where $e_{0,r}(k)$ is the *a priori* error of the $r^{\text{th}}$ Monte-Carlo run; (ii) misalignment [17, 19], defined as $\text{Mis}(k) = \frac{1}{R} \sum_{r=1}^{R} \frac{\|\mathbf{w}_r(k)-\mathbf{h}(k)\|_2^2}{\|\mathbf{w}_r(k)\|_2^2}$, where $\mathbf{w}_r(k)$ is the parameter vector for the $r^{\text{th}}$ Monte-Carlo run; and (iii) overall percentage of updates.

Fig. 1(a) shows the MSE results for **Exp. 1**. It is clear that the SM-AP algorithm working with the proposed CC-CV is able to achieve fast convergence, similarly to SC-CV and ED-CV, and faster than the optimal CV. In addition, the steady-state MSE is lower than both SC-CV and ED-CV, being rather close to the steady-state MSE of the optimal CV. Fig. 2 depicts the misalignment results for **Exp. 1**. One can see that the proposal finds a good compromise between convergence speed for the misalignment and the steady-state misalignment, being much faster than the optimal.
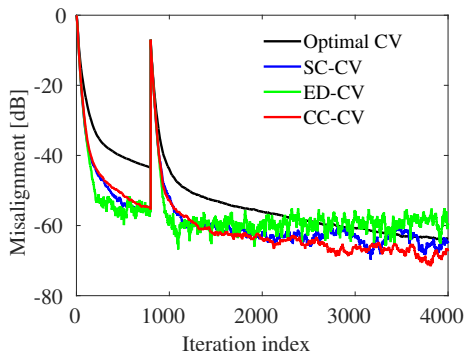
Fig. 1(b) shows the MSE results for **Exp. 2**. These results let clear that the SM-PAPA working with the CC-CV achieves near-optimal performance, outperforming both SC-CV and ED-CV. Fig. 1(c) shows the MSE results for a unit-variance white input. Note that the MSE curves of SC-CV, optimal CV, and the proposed CC-CV are virtually identical.

Table 1 shows the percentage of updates for all three experiments. Note that both the optimal CV and the proposed CC-CV update much less than the SC-CV and ED-CV.

**Table 1**. Overall percentage of updates

| CV | Exp. 1 | Exp. 2 | Exp. 3 |
|---------|--------|--------|--------|
| Optimal | 16.7% | 11.9% | 08.5% |
| SC | 23.2% | 19.9% | 09.0% |
| ED | 30.9% | 28.8% | 16.2% |
| CC | 16.5% | 13.6% | 08.6% |

## 5. CONCLUSIONS

In this paper we proposed an efficient way to determine the constraint vector in set-membership affine projection (SM-AP) adaptive algorithms whose performances match closely the optimal solution. The proposal consists of an optimal convex combination of two predefined constraint vectors. The resulting sub-optimal solution can be computed in closed-form, requiring low computational complexity. It is shown that using an SM-AP algorithm along with the proposed constraint vector is equivalent to using a combination of two SM-AP filters, each one employing the predefined constraint vectors that were initially used for the convex combination. The simulation results in dispersive and sparse scenarios demonstrate the near-optimality of the proposed solution.

### Acknowledgements

**Fig. 2**. Misalignment for **Exp. 1**.

# 6. REFERENCES

[1] E. Fogel and Y.-F. Huang, "On the value of information in system identification–bounded noise case," *Automatica*, vol. 18, no. 2, pp. 229–238, March 1982.

[2] J.R. Deller, "Set-membership identification in digital signal processing," *IEEE ASSP Magazine*, vol. 6, no. 4, pp. 4–20, Oct. 1989.

[3] S. Gollamudi, S. Nagaraj, S. Kapoor, and Y.-F. Huang, "Set-membership filtering and a set-membership normalized LMS algorithm with an adaptive step size," *IEEE Signal Processing Letters*, vol. 5, no. 5, pp. 111–114, May 1998.

[4] S. Werner and P.S.R. Diniz, "Set-membership affine projection algorithm," *IEEE Signal Processing Letters*, vol. 8, no. 8, pp. 231–235, Aug. 2001.

[5] P.S.R. Diniz and S. Werner, "Set-membership binormalized data reusing LMS algorithms," *IEEE Transactions on Signal Processing*, vol. 51, no. 1, pp. 124–134, Jan. 2003.

[6] L. Guo and Y.-F. Huang, "Frequency-domain set-membership filtering and its applications," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1326–1338, April 2007.

[7] M.Z.A. Bhotto and A. Antoniou, "A set-membership affine projection algorithm with adaptive error bound," in *Canadian Conference on Electrical and Computer Engineering (CCECE-2009).*, May 2009, pp. 894–897.

[8] M.Z.A. Bhotto and A. Antoniou, "Robust set-membership affine-projection adaptive-filtering algorithm," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 73–81, Jan 2012.

[9] R. Arablouei and K. Dogançay, "Set-membership recursive least-squares adaptive filtering algorithm," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2012)*, March 2012, pp. 3765–3768.

[10] R. Arablouei and K. Dogançay, "Low-complexity implementation of quasi-OBE algorithm," *Electronics Letters*, vol. 48, no. 11, pp. 621–623, May 2012.

[11] S. Werner, Y.-F. Huang, M.L.R. de Campos, and V. Koivunen, "Distributed parameter estimation with selective cooperation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009)*, April 2009, pp. 2849–2852.

[12] J. R. Deller and Y.-F. Huang, "Set-membership identification and filtering for signal processing applications," *Circuits, Systems and Signal Processing*, vol. 21, no. 1, pp. 69–82, 2002.

[13] P.S.R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, Springer, New York, USA, 4th edition, 2013.

[14] P.S.R. Diniz, "Convergence performance of the simplified set-membership affine projection algorithm," *Circuits, Systems, and Signal Processing*, vol. 30, no. 2, pp. 439–462, April 2011.

[15] M.V.S. Lima and P.S.R. Diniz, "Steady-state analysis of the set-membership affine projection algorithm," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2010)*, Dallas, USA, March 2010, pp. 3802–3805.

[16] M.V.S. Lima and P.S.R. Diniz, "Steady-state MSE performance of the set-membership affine projection algorithm," *Circuits, Systems, and Signal Processing*, vol. 32, no. 4, pp. 1811–1837, Jan. 2013, DOI: 10.1007/s00034-012-9545-4.

[17] M.V.S. Lima, T.N. Ferreira, W.A. Martins, and P.S.R. Diniz, "Sparsity-aware data-selective adaptive filters," *IEEE Transactions on Signal Processing*, vol. 62, no. 17, pp. 4557–4572, Sept. 2014.

[18] M.V.S. Lima and P.S.R. Diniz, "Fast learning set theoretic estimation," in *21st European Signal Processing Conference (EUSIPCO 2013)*, Marrakech, Morocco, Sept. 2013, pp. 1–5.

[19] W.A. Martins, M.V.S. Lima, P.S.R. Diniz, and T.N. Ferreira, "Optimal constraint vectors for set-membership affine projection algorithms," *Signal Processing*, vol. 134, pp. 285–294, May 2017.

[20] M.J.M. Spelta and W.A. Martins, "Optimal constraint vectors for set-membership proportionate affine projection algorithms," in *IEEE Statistical Signal Processing Workshop (SSP-2018)*, Freiburg, Germany, June 2018, pp. 1–5.

[21] S. Werner, J.A. Apolinario Jr., and P.S.R. Diniz, "Set-membership proportionate affine projection algorithms," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2007, no. 1, pp. 1–10, Jan. 2007.

[22] J. Arenas-Garcia, A. Figueiras-Vidal, and A. Sayed, "Mean-square performance of a convex combination of two adaptive filters," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1078–1090, Mar 2006.

[23] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[24] P.H. Calamai and J.J. More, "Projected gradient methods for linearly constrained problems," *Mathematical Programming*, vol. 39, no. 1, pp. 93–116, 1987.