

Performance Evaluation of Adaptive Filters for Sparse Wireless Channel Estimation

Markus V. S. Lima[†], Tadeu N. Ferreira^{*}, Wallace A. Martins[†], Marcele O. K. Mendonça[†], Paulo S. R. Diniz[†]

[†]Poli & COPPE

Federal University of Rio de Janeiro (UFRJ)
P.O. Box 68504
Rio de Janeiro, RJ, Brazil
Zip 21941-972

^{*}Engineering School

Fluminense Federal University (UFF)
Rua Passo da Pátria 56 block E room 406
Niterói, RJ, Brazil
Zip 24210-240

Abstract—The progressive increase of data rates in wireless communication systems has induced channel models with sampled impulse responses which are mostly sparse. This paper presents a unified derivation of adaptive filters exploiting sparsity in the complex domain, and compares the performance of classic and state-of-the-art adaptive algorithms for estimating sparse wireless channels as well as their tracking ability in this inherently time-varying environment. Simulation results confirm the efficiency of the sparsity-aware algorithms.

Index Terms—adaptive filtering, sparsity, set-membership, wireless channel estimation, NLMS

I. INTRODUCTION

The identification of systems with sparse impulse responses has been largely explored in several adaptive solutions [1]. The problem of echo cancellation is the original environment where those *sparsity-aware algorithms* have been applied [2] [3] [4].

Sparsity in wireless communication channels has become an important issue to be addressed considering the demands for ever-increasing transmission data rates. In [5], different models for sparse wireless channels are presented. In [6], a comparative analysis is conducted for sparsity-aware algorithms focusing on real-valued channels with slow variations and considering white and correlated inputs.

The main contribution of this paper is comparing the performance of adaptive algorithms for sparse wireless channel estimation in terms of misalignment. Since the pilot signals used for channel estimation are not strongly correlated, normalized least-mean-square (NLMS)-inspired algorithms are used. As compared to [6], this work focuses on long-term evolution (LTE)-like environments, with time-varying channels, enabling the assessment of both the algorithms' tracking ability and accuracy performance for channel estimation. An additional contribution of this work is to derive the complex-valued versions of some sparsity-aware adaptive algorithms, while also assessing their performance in baseband setups.

The paper is organized as follows: in Section II, wireless channel models are presented, highlighting their sparsity. Section III derives in a unified framework the adaptive algorithms used in the forthcoming analyses, considering complex-valued signals and parameters. In Section IV, the main results obtained in the comparison are presented. The concluding remarks of this work are drawn in Section V.

II. SPARSITY IN WIRELESS CHANNELS

As data rates increase in current wireless communication systems, there is also a corresponding increase in the number of near-zero samples of the channel model perceived at the digital receiver front-end. A case in point is [5] where the authors show that wireless multiple-input multiple-output (MIMO) channels commonly used in the so-called 4th generation (4G) of mobile cellular systems can be regarded as sparse. In that context, the channel is modeled in a four-dimensional domain comprising variables for transmitting angle, receiving angle, delay, and Doppler. One of the simplifications of the modeling presented in [5] and [7] addresses the single-antenna case, which is the approach used in this paper.

The equivalent baseband received signal for single-input single-output (SISO) channels is assumed to be

$$x(t) = \sum_{m=0}^{M-1} \sum_{d=-D}^D H(m, d) e^{j2\pi dt/T} s(t - m/W) + v(t), \quad (1)$$

where $s(t)$ is the transmitted signal, m denotes discrete delays (sampled with period T), d denotes a discrete frequency-domain variable accounting for Doppler effects, $H(m, d)$ is the channel response with bandwidth W , and $v(t)$ is an additive white Gaussian noise (AWGN) corrupting the received signal.

LTE has become the most popular 4G system, achieving data rates over 50 Mbps (megabit-per-second). Those rates give rise to a wide range of values of (m, d) for which $H(m, d)$ presents small amplitudes, thus inducing a sparse/compressible model.

In this paper, channel estimation is performed in time-domain, generating $h(t, m)$, defined as:

$$h(t, m) = \sum_{d=-D}^D H(m, d) e^{j2\pi dt/T}. \quad (2)$$

In some of the scenarios tested here, $h(t, m)$ varies slowly, thus yielding an approximately time-invariant impulse response during a sufficiently small time-interval Δt . In this context, one can write $h(t, m) \approx h_0(m)$ for $t_0 \leq t < t_0 + \Delta t$.

A. The Standardized Pedestrian LTE Channel

The frequency bands used in the deployment of LTE usually range from 700 MHz to 2.5 GHz [8]. Movements of either transmitter or receiver in an LTE-based communication link induce channel models that are strongly time-varying and that suffer from Doppler effects.

Doppler effects are included in the model by inducing slight time variations in the channel impulse response (CIR). For that reason, the standardized extended pedestrian A (EPA) channel is employed since it presents relatively weaker Doppler effects, as compared to other LTE-based standard channels [9].

A simplified version of the EPA channel is considered in which only the paths whose power spectral densities are greater than half the power of the line-of-sight path are used. Hence, only 3 paths are considered. A quite common central frequency is used at 800 MHz. Therefore, a 57-tap channel is modeled at 800 MHz with non-zero taps at positions (1, 25, 57) with relative powers (0, -1, -2) dB.

Although LTE transceivers usually employ multiple antennas, this paper considers a SISO communication model, for the sake of simplicity. Adaptive algorithms can be successfully used to track time variations of the channel. Here, a delay-domain estimation of the channel is used, since the adaptive algorithms chosen work in the time-domain. This may be tailored to LTE, in which channel estimation is usually performed in the frequency-domain, by using the dual of Eq. (2) in the Doppler-domain.

Next section contains the derivation of complex-valued versions of state-of-the-art adaptive algorithms that account for the underlying model sparsity.

III. COMPLEX-VALUED ALGORITHMS FOR SPARSE CHANNEL ESTIMATION

As presented in Section II, the digital baseband channel model of a RF passband transmission has complex coefficients. Therefore, in order to properly estimate the CIR, a complex-valued adaptive filter should be used. In case the CIR is sparse, there exist several algorithms that take such sparsity into account in order to improve convergence speed and/or estimation accuracy. They are widely known as *sparsity-aware algorithms*. Many sparsity-aware algorithms were originally derived considering real variables. In this section we extend several recently proposed algorithms to the case where both filter coefficients and input signals are complex. *We do this in a unified manner in such a way that penalty-based and proportionate algorithms arise naturally.* We focus on NLMS-type algorithms, since pilot signals used for channel estimation are usually white or weakly correlated so that the data-reuse paradigm provided by affine projection (AP) algorithms would not result in a significant improvement, but would certainly increase the computational burden. Thus, the AP data-reuse parameter L was set to zero, yielding NLMS-based algorithms [10].

A. Derivation of Complex-Valued Sparsity-Aware Algorithms

Let us consider an adaptive filter whose complex coefficients $\mathbf{w}(k) \in \mathbb{C}^N$, whenever updated, solve the following optimization problem

$$\begin{aligned} \min_{\mathbf{w}(k+1)} \quad & \frac{1}{2} \|\mathbf{w}(k+1) - \mathbf{w}(k)\|_{\mathbf{G}^{-1}(k)}^2 + \frac{1}{2} \alpha p(\mathbf{w}(k+1)), \\ \text{s.t.} \quad & d(k) - \mathbf{x}^T(k) \mathbf{w}^*(k+1) = \gamma(k), \end{aligned} \quad (3)$$

where $d(k) \in \mathbb{C}$ is the desired/reference signal, $\mathbf{x}(k) \in \mathbb{C}^N$ is the input vector of the adaptive filter, and $\gamma(k) \in \mathbb{C}$ is a set-membership (SM) parameter that should satisfy $|\gamma(k)| \leq \bar{\gamma}$, with $\bar{\gamma} \in \mathbb{R}_+$ representing the SM error threshold, i.e., the amount of error that we consider acceptable given the problem uncertainties known beforehand or estimated online. In addition, the norm induced by matrix $\mathbf{G}^{-1}(k)$ is defined as $\|\mathbf{z}\|_{\mathbf{G}^{-1}(k)}^2 \triangleq \mathbf{z}^H \mathbf{G}^{-1}(k) \mathbf{z}$ with $\mathbf{G}^{-1}(k) \in \mathbb{C}^{N \times N}$ being a positive-definite Hermitian matrix, $\alpha \in \mathbb{R}_+$ is the weight given to the penalty term, and $p(\cdot)$ is any well-behaved sparsity-promoting penalty function.¹ *Observe that the formulation in (3) is very powerful and general, as it is comprised of penalty, proportionate, and set-membership terms.*

By using the method of Lagrange multipliers, the problem in (3) can be rewritten as an unconstrained optimization problem. In this way, our goal is to find $\mathbf{w}(k+1)$ that minimizes the Lagrangian \mathcal{L} , which is defined as

$$\begin{aligned} \mathcal{L} \triangleq \quad & \frac{1}{2} \|\mathbf{w}(k+1) - \mathbf{w}(k)\|_{\mathbf{G}^{-1}(k)}^2 + \frac{1}{2} \alpha p(\mathbf{w}(k+1)) + \\ & + \frac{1}{2} \lambda^*(k) [d^*(k) - \mathbf{x}^H(k) \mathbf{w}(k+1) - \gamma^*(k)] + \\ & + \frac{1}{2} \lambda(k) [d(k) - \mathbf{x}^T(k) \mathbf{w}^*(k+1) - \gamma(k)], \end{aligned} \quad (4)$$

where $\lambda(k) \in \mathbb{C}$ is the Lagrange multiplier. In (4) we took the real part of the term corresponding to the constraints because the involved variables are complex [10].

The optimal $\mathbf{w}(k+1)$ is found by differentiating \mathcal{L} wrt. $\mathbf{w}^*(k+1)$ [10] and equating the result to zero. That is, by computing $\frac{\partial \mathcal{L}}{\partial \mathbf{w}^*(k+1)} = 0$ we obtain

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \lambda(k) \mathbf{G}(k) \mathbf{x}(k) - \alpha \mathbf{G}(k) \nabla p(\mathbf{w}(k+1)), \quad (5)$$

where $\nabla p(\mathbf{w}(k+1)) \triangleq \frac{\partial p(\mathbf{w}(k+1))}{\partial \mathbf{w}^*(k+1)}$ is the gradient of p . Similarly, $\frac{\partial \mathcal{L}}{\partial \lambda^*(k)} = 0$ is computed:

$$\frac{1}{2} [d^*(k) - \mathbf{x}^H(k) \mathbf{w}(k+1) - \gamma^*(k)] = 0, \quad (6)$$

which can be rewritten in a more convenient form as follows,

$$\mathbf{x}^H(k) \mathbf{w}(k+1) = d^*(k) - \gamma^*(k). \quad (7)$$

Multiplying both sides of (5) by $\mathbf{x}^H(k)$ on the left, and then using (7) yields

$$\begin{aligned} e^*(k) - \gamma^*(k) &= \lambda(k) \mathbf{x}^H(k) \mathbf{G}(k) \mathbf{x}(k) \\ &\quad - \alpha \mathbf{x}^H(k) \mathbf{G}(k) \nabla p(\mathbf{w}(k+1)), \end{aligned} \quad (8)$$

¹By well-behaved penalty we mean an almost everywhere differentiable penalty function, such as the l_1 norm.

where $e^*(k) \triangleq d^*(k) - \mathbf{x}^H(k)\mathbf{w}(k)$ is the conjugated error.

By assuming $\mathbf{x}^H(k)\mathbf{G}(k)\mathbf{x}(k) \neq 0$ and defining $S(k) \triangleq [\mathbf{x}^H(k)\mathbf{G}(k)\mathbf{x}(k)]^{-1} \in \mathbb{R}_+$, we obtain

$$\lambda(k) = S(k)[e^*(k) - \gamma^*(k)] + \alpha S(k)\mathbf{x}^H(k)\mathbf{G}(k)\nabla p(\mathbf{w}(k+1)). \quad (9)$$

Substituting (9) into (5) yields

$$\begin{aligned} \mathbf{w}(k+1) &= \mathbf{w}(k) + \mathbf{G}(k)\mathbf{x}(k)S(k)[e^*(k) - \gamma^*(k)] \\ &\quad + \alpha\mathbf{G}(k)\mathbf{x}(k)S(k)\mathbf{x}^H(k)\mathbf{G}(k)\nabla p(\mathbf{w}(k+1)) \\ &\quad - \alpha\mathbf{G}(k)\nabla p(\mathbf{w}(k+1)). \end{aligned} \quad (10)$$

It is worth mentioning that, in practice, $S(k)$ is replaced by $S'(k) \triangleq [\mathbf{x}^H(k)\mathbf{G}(k)\mathbf{x}(k) + \delta]^{-1}$, where $\delta \in \mathbb{R}_+$ is a regularization used to avoid numerical problems that occur when $\mathbf{x}^H(k)\mathbf{G}(k)\mathbf{x}(k)$ approaches zero.

From Eq. (10), complex versions of most sparsity-aware algorithms are obtained, as shown in the following subsections.

B. NLMS Algorithm

By setting $\mathbf{G}(k) = \mathbf{I}$ (no proportionate term), $\alpha = 0$ (no penalty term), and $\bar{\gamma} = 0 \Rightarrow \gamma(k) = 0$ (no SM term), the following recursion is obtained

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mathbf{x}(k)S'(k)e^*(k). \quad (11)$$

In addition, a relaxation parameter μ can be added to control the step-size of the algorithm leading to

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{\mu}{\mathbf{x}^H(k)\mathbf{x}(k) + \delta} e^*(k)\mathbf{x}(k), \quad (12)$$

which corresponds to the complex NLMS algorithm [10] that will be used as benchmark in Section IV.

C. Proportionate NLMS Algorithms

If $\alpha = 0$ and $\mathbf{G}(k)$ is a diagonal matrix whose diagonal entries $g_{nn}(k)$ are proportional to their related coefficients $w_n(k)$ (n^{th} entry of $\mathbf{w}(k)$), then we get the proportionate family of algorithms whose general recursion is

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mathbf{G}(k)\mathbf{x}(k)S'(k)[e^*(k) - \gamma^*(k)], \quad (13)$$

if $|e(k)| > \bar{\gamma}$, and $\mathbf{w}(k+1) = \mathbf{w}(k)$ whenever $|e(k)| \leq \bar{\gamma}$. According to the set-membership principles, we should use $\gamma(k) = \bar{\gamma}\text{sign}(e(k))$, which allows us to rewrite (13) as

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{\mu(k)}{\mathbf{x}^H(k)\mathbf{G}(k)\mathbf{x}(k) + \delta} \mathbf{G}(k)\mathbf{x}(k)e^*(k), \quad (14)$$

where

$$\mu(k) = \begin{cases} 1 - \frac{\bar{\gamma}}{|e^*(k)|} & \text{if } |e^*(k)| > \bar{\gamma}, \\ 0 & \text{otherwise.} \end{cases} \quad (15)$$

The recursion in (14) corresponds to the SM-PNLMS algorithm for complex variables [11], [12].

Regarding matrix $\mathbf{G}(k)$, there are many different ways of defining its diagonal entries—see [13], [14], [15], [16] for some examples. A widely used form is given by [12]

$$g_{nn}(k) = \frac{1 - \kappa\alpha(k)}{N} + \frac{\kappa\alpha(k)|w_n(k)|}{\|\mathbf{w}(k)\|_1}, \quad (16)$$

where κ is a predefined parameter. Observe that if $\mathbf{G}(k) = \mathbf{I}$, then (14) reduces to the complex SM-NLMS [17], [10].

D. Penalty-based Algorithms

In this family of algorithms, we consider $\mathbf{G}(k) = \mathbf{I}$, $\alpha \neq 0$, and $\nabla p(\mathbf{w}(k+1)) \approx \nabla p(\mathbf{w}(k))$ so that (10) becomes

$$\begin{aligned} \mathbf{w}(k+1) &= \mathbf{w}(k) + \mathbf{x}(k)S'(k)[e^*(k) - \gamma^*(k)] \\ &\quad - \alpha [\mathbf{I} - \mathbf{x}(k)S'(k)\mathbf{x}^H(k)] \nabla p(\mathbf{w}(k)), \end{aligned} \quad (17)$$

if $|e(k)| > \bar{\gamma}$ (i.e., an update is required), or $\mathbf{w}(k+1) = \mathbf{w}(k)$, otherwise.

For the sake of clarity, let us start considering $\bar{\gamma} = 0$, implying $\gamma(k) = 0$ (non-SM algorithms). The zero attractor (ZA) algorithms are obtained by choosing p as the l_1 norm leading to $\nabla p(\mathbf{w}(k)) = \text{sign}(\mathbf{w}(k))$.² An interesting alternative is to choose $p = F_\beta$ [18], where F_β is an approximation to the l_0 “norm” and the parameter β controls the tradeoff between smoothness and accuracy of the approximation. There are many suitable functions F_β —see [18], [19]—, but for the simulations presented in the following section we consider just the Geman-McClure function (GMF), which leads to

$$(\nabla p(\mathbf{w}(k)))_n = (\nabla F_\beta(\mathbf{w}(k)))_n = \frac{\beta \text{sign}(w_n(k))}{(\beta|w_n(k)| + 1)^2}, \quad (18)$$

where $(\nabla p(\mathbf{w}(k)))_n$ represents the n^{th} entry of the gradient vector. In this case, one obtains the NLMS sparse system identification (NLMS-SSI). In [19], it is shown that, in some situations, the term $\alpha\mathbf{x}(k)S'(k)\mathbf{x}^H(k)$ may be discarded with no harm to the convergence, which gives rise to the quasi NLMS-SSI (QNLMS-SSI). Another alternative, which is related to the approximation to the l_0 “norm” [18], is to choose $(\nabla p(\mathbf{w}(k)))_n = (\text{sign}(w_n(k)))/(1 + \epsilon|w_n(k)|)$ with small positive ϵ , which is known as reweighted ZA-NLMS (RZA-NLMS) algorithm [20].³

The SM versions can be derived by considering $\bar{\gamma} \neq 0$ and $\gamma(k) = \bar{\gamma}\text{sign}(e(k))$. In this way, the sparse SM-NLMS (SSM-NLMS) is obtained when $p = F_\beta$, as in (18) for example. Moreover, when the term $\alpha\mathbf{x}(k)S'(k)\mathbf{x}^H(k)$ is discarded, the SSM quasi NLMS (SSM-QNLMS) is unveiled [18].⁴

E. Additional Topics

We put some emphasis on the aforementioned algorithms because they are representative of classic and state-of-the-art adaptive algorithms for sparse system identification. Formulation in (10) allows us to derive complex versions of others sparsity-aware algorithms, for instance, by combining both proportionate and penalty terms, as proposed in [21], [22].

Advantages of the l_0 “norm” penalty over the l_1 norm are addressed in [18], [23], [19]. The complexity of some of these

²Observe that the entries of $\text{sign}(\mathbf{w}(k))$ are given by $\text{sign}(w_n(k)) = w_n(k)/|w_n(k)|$ when $w_n(k) \neq 0$, or $\text{sign}(w_n(k)) = 0$ when $w_n(k) = 0$.

³The examples of $(\nabla p(\mathbf{w}(k)))_n$ are the standard expressions for real-valued variables. For complex variables, a formal derivation would show that the complex $(\nabla p(\mathbf{w}(k)))_n$ is half of the real $(\nabla p(\mathbf{w}(k)))_n$. However, this is not a big issue, as the factor 1/2 can be incorporated into α .

⁴Observe that algorithms of the AP-type are generalizations of their NLMS versions that allow the reuse of previous input vectors.

algorithms can be found in [18], including the SM algorithms, which reduce computational cost by not updating the filter coefficients at every iteration.

IV. RESULTS

In this section, simulations are conducted in order to compare the adaptive algorithms' performances for sparse wireless channel estimation. The main figure of merit chosen for the comparison is the Misalignment [18], computed as

$$\text{Misalignment}(k) = \frac{1}{R} \sum_{r=1}^R \frac{\|\mathbf{h}^{(r)}(k) - \mathbf{w}^{(r)}(k)\|_2^2}{\|\mathbf{h}^{(r)}(k)\|_2^2}, \quad (19)$$

where R denotes the number of Monte-Carlo runs, $\mathbf{w}^{(r)}(k)$ and $\mathbf{h}^{(r)}(k)$ are, respectively, the coefficient vector and the channel realization at k^{th} iteration for the r^{th} Monte-Carlo run.

The first simulation scenario comprises a very sparse channel, similar to the ones used in a previous comparison [18], with only one non-zero tap out of 40. The position of the non-zero tap was varied twice, one at iteration 300 and another at iteration 550, while keeping the magnitude of these taps fixed. The main difference is the complex model for the channel taps, since our channel is the baseband equivalent to RF transmission channels. The input signal is binary phase shift-keying (BPSK) symbols. In this simulation, 2500 symbols were transmitted. Misalignment was calculated according to (19), with $R = 2000$, with results presented in Fig. 1(a).

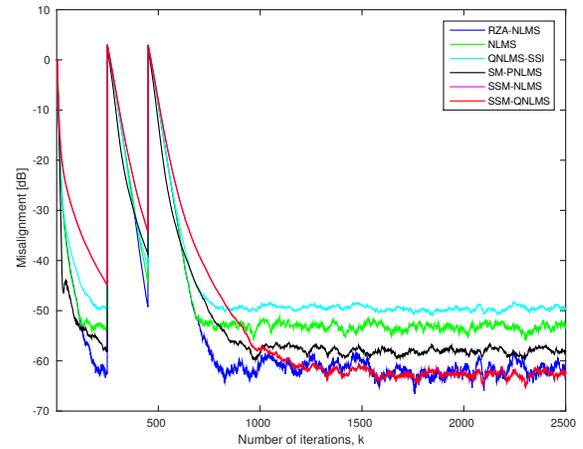
The algorithms used in these simulations present several different parameters. They were all set up in order to present similar mean-squared error (MSE) convergence rates in the beginning of transient states, as depicted in Fig. 1(b). This criterion allows one to draw a fair comparison among the different algorithms, since they all are conceived based on MSE-related cost-functions.⁵ The SM threshold was set as $\bar{\gamma} = \sqrt{0.8\sigma_v^2}$ and the noise power $\sigma_v^2 = 0.01$.

Based on Fig. 1(a), one can verify that SM-based algorithms presented worse tracking ability. On the other hand, SM-based algorithms, together with RZA-NLMS, presented lower final misalignment. The percentage of iterations with updates for SSM-NLMS and SSM-QNLMS were both 56.1%, whereas for SM-PNLMS it was 60.2%.

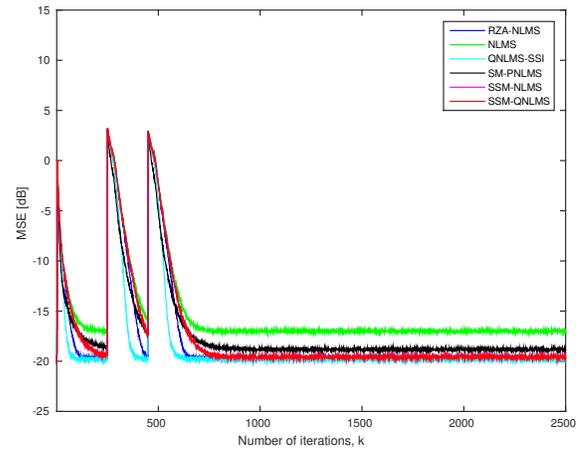
Simulations were also performed using quadrature amplitude modulation (QAM) constellation. The misalignment results are shown in Fig. 2. Once more, 2500 symbols were transmitted and $R = 2000$ Monte-Carlo runs were used. The results were similar to the simulation with BPSK transmission, except for a better performance of SM-PNLMS. As for the percentage of updates, the values for SSM-NLMS and SSM-QNLMS were 51.3%. For SM-PNLMS, there were updates in 54.3% of the iterations.

In order to test the algorithms behavior when the channel is time-variant, the EPA channel described in Subsection II-A, is considered in the next simulation. The EPA channel was

⁵However, in this particular setup, we verified through other simulations that SM-PNLMS could achieve an even lower misalignment if other criteria were used.



(a) Misalignment.



(b) MSE.

Fig. 1. Results for BPSK input signal.

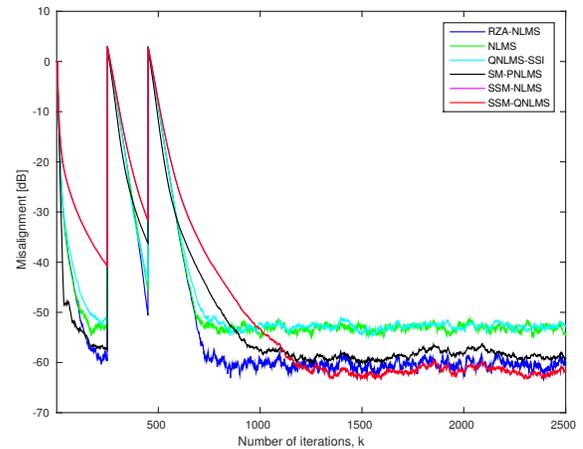


Fig. 2. Misalignment curves measured for QAM input.

modeled to present a coherence time around 75 ns. The CIR remains constant by 60 iterations and then it suffers a small variation so that the non-zero taps may be modeled as a first-order autoregressive process $h_n(k) = \lambda_1 h_n(k-59) + \lambda_2 v_{w,i}(k)$, where $\lambda_1 = 0.9995$, $\lambda_2 = 0.005$ and $v_{w,i}(k)$

is a complex white noise with unit power. Moreover, strong variability on the channel was added on iterations 1200 and 2400, by re-randomizing the magnitude values of the non-zero taps and shifting their position by 1 sample.

In this simulation, 4000 BPSK symbols were transmitted and the result was averaged along $R = 6000$ Monte-Carlo runs. The misalignment curves are presented in Fig. 3, showing that all algorithms presented very good tracking ability throughout the iterations. SSM-NLMS and SSM-QNLMS achieved lower convergence rates. On the other hand, SSM-NLMS and SSM-QNLMS updated in only 62.9% of the iterations, while SM-PNLMS updated in 63.5% of the iterations.

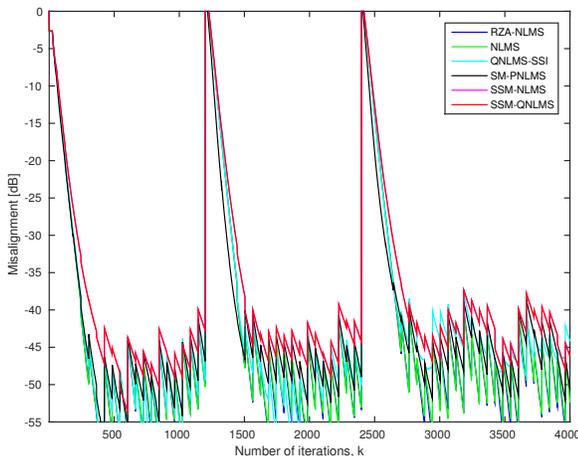


Fig. 3. Misalignment curves measured for BPSK input and EPA channel.

In the simulations with time-invariant channel, both RZA-NLMS and set-membership algorithms achieved lower levels of MSE and misalignment. In addition, considering only MSE measurements, QNLMS-SSI algorithm also presented competitive performance. However, when the channel becomes time-variant, QNLMS-SSI, RZA-NLMS, and NLMS algorithms presented best misalignment performance.

V. CONCLUDING REMARKS

This paper presented complex-valued versions of some state-of-the-art sparsity-aware adaptive algorithms in a unified framework and their comparative analyses when applied to wireless channel estimation with high degree of sparsity. NLMS-based algorithms have been tested in simple scenarios, and also in a time-varying channel. The algorithms that exploit sparsity presented the best results, both in terms of final MSE as well as misalignment. Set-membership algorithms yielded a trade-off between convergence speed and lower computational complexity. The standard NLMS algorithm presented competitive performance in terms of tracking.

ACKNOWLEDGMENT

The authors thank agencies CNPq, Capes, and Faperj.

REFERENCES

[1] D. Etter, "Identification of sparse impulse response system using adaptive delay filter," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Tampa, USA, March 1985, pp. 1169–1172.

[2] S. L. Gay, "An efficient, fast converging adaptive filter for network echo cancellation," in *Thirty-Second Asilomar Conference on Signals, Systems and Computers (ACSSC 1998)*, vol. 1, 1998, pp. 394–398.

[3] P. Naylor, J. Cui, and M. Brookes, "Adaptive algorithms for sparse echo cancellation," *Signal Processing*, vol. 86, no. 6, pp. 1182–1192, June 2006.

[4] C. Paleologu, S. Ciochina, and J. Benesty, "An efficient proportionate affine projection algorithm for echo cancellation," *IEEE Signal Processing Letters*, vol. 17, no. 2, pp. 165–168, 2010.

[5] W. U. Bajwa, A. Sayeed, and R. Nowak, "Sparse multipath channels: Modeling and estimation," in *Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop, 2009. DSP/SPE 2009. IEEE 13th*, Marco Island, FL, USA, January 2009, pp. 320–325.

[6] A. Al-Shabli and B. T. et al., "Sparse nlms adaptive algorithms for multipath wireless channel estimation," in *IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Abu Dhabi, UAE, October 2015, pp. 839–844.

[7] A. Sayeed, "A virtual representation for time- and frequency-selective correlated mimo channels," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2003)*, Hong Kong, China, April 2003, pp. IV648–IV651.

[8] E. Dahlman, S. Parkvall, and J. Skold, *4G: LTE/LTE-Advanced for Mobile Broadband*. Oxford, UK: Academic Press/Elsevier, 2013.

[9] 3GPP/ETSI, "Etsi ts 136 104 v.8.3.0," *LTE: E-UTRA - Base Station (BS) Radio Transmission and Reception*, pp. 1–68, 2008–2011.

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

[11] S. Werner, J. A. Apolinario Jr, P. S. R. Diniz, and T. Laakso, "A set-membership approach to normalized proportionate adaptation algorithms," in *European Signal Processing Conference (EUSIPCO 2005)*, Antalya, Turkey, Sept. 2005, pp. 1–4.

[12] 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.

[13] D. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 5, pp. 508–518, Sept. 2000.

[14] J. Benesty and S. Gay, "An improved PNLM algorithm," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2002)*, vol. 2, Dallas, USA, May 2002, pp. 1881–1884.

[15] L. Liu, M. Fukumoto, and S. Saiki, "An improved mu-law proportionate NLMS algorithm," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)*, 2008, pp. 3797–3800.

[16] A. Khong and P. Naylor, "Efficient use of sparse adaptive filters," in *Fortieth Asilomar Conference on Signals, Systems and Computers (ACSSC 2006)*, 2006, pp. 1375–1379.

[17] 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.

[18] 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.

[19] M. V. S. Lima, W. A. Martins, and P. S. R. Diniz, "Affine projection algorithms for sparse system identification," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2013)*, Vancouver, Canada, May 2013, pp. 5666–5670.

[20] R. Meng, R. de Lamare, and V. Nascimento, "Sparsity-aware affine projection adaptive algorithms for system identification," in *Sensor Signal Processing for Defence*, London, U.K., Sept. 2011, pp. 1–5.

[21] K. Pelekanakis and M. Chitre, "New sparse adaptive algorithms based on the natural gradient and the l_0 -norm," *IEEE Journal of Oceanic Engineering*, vol. 38, no. 2, pp. 323–332, April 2013.

[22] T. N. Ferreira, M. V. S. Lima, P. S. R. Diniz, and W. A. Martins, "Low-complexity proportionate algorithms with sparsity-promoting penalties," in *2016 IEEE International Symposium on Circuits and Systems (IS-CAS)*, May 2016, pp. 253–256.

[23] M. V. S. Lima, I. Sobron, W. A. Martins, and P. S. R. Diniz, "Stability and mse analyses of affine projection algorithms for sparse system identification," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 6399–6403.