



HAL
open science

On multiplicative update with forgetting factor adaptive step size for least mean-square algorithms

Robin Gerzaguët, Laurent Ros, Fabrice Belvèze, Jean-Marc Brossier

► To cite this version:

Robin Gerzaguët, Laurent Ros, Fabrice Belvèze, Jean-Marc Brossier. On multiplicative update with forgetting factor adaptive step size for least mean-square algorithms. ICT 2018 - 25th International Conference on Telecommunications, Jun 2018, Saint-Malo, France. hal-01820720

HAL Id: hal-01820720

<https://hal.science/hal-01820720v1>

Submitted on 22 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On multiplicative update with forgetting factor adaptive step size for least mean-square algorithms

Robin GERZAGUET*, Laurent ROS†, Fabrice BELVEZE‡, Jean-Marc BROSSIER†

* Univ Rennes, CNRS, IRISA, France, robin.gerzagu@irisa.fr

†Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, France

‡ST-Microelectronics, Grenoble, France

Abstract—This paper deals with a new adaptive step-size overlay dedicated to Least Mean-Square (LMS) algorithms for noise cancellation with reference. LMS algorithms are often used in real-time estimations due to their robustness and simplicity, and are guided by their constant step size. Their transient and asymptotic performances are both linked to the step size, which leads to the well-known trade-off between speed and accuracy. In this paper, we propose a LMS algorithm with a new self-adaptive overlay based on a Multiplicative Update with Forgetting Factor (LMS-MUFF). The adaptive overlay improves the speed and the reactivity of the algorithm. In order to tune the algorithm parameters, we express the scalar case semi-analytical relationship between the LMS-MUFF parameters and a false-alarm probability, as defined from unwanted variations in the step-size during the asymptotic mode. We compare our method to reference methods in the literature and show that it offers better speed and reactivity with the same asymptotic performance.

Index Terms—least mean-square algorithms; adaptive noise cancellation algorithms; self-adaptive step size; steepest descent algorithms.

I. INTRODUCTION

The least mean-square (LMS) algorithm [1], [2] is often used in adaptive signal processing due to its robustness and its low complexity. Thus, in various fields, noise-reduction algorithms that exploit knowledge of the polluters with a natural or synthesized reference [3], [4], [5] and equalization algorithms [6], [7], [8] are often based on LMS. Recently, LMS algorithms have been extensively used for cancellation of self-interference in radio frequency transceivers, where low complexity and exploitation of polluter knowledge are mandatory [9], [10], [11]. Basically, the algorithm consists in estimating the coefficients of the channel between a reference (linked to the polluter) and the observation in order to subtract to the observation the reference filtered by the estimated channel. The LMS algorithm is parameterized by its step size, and many studies have investigated the relationships between step size and performance and the trade-off between the speed of the algorithm and its accuracy, especially when the channel to estimate is time varying [12], [13]. In such a case, the performance can be expressed through the sum of a static term, relative to the influence of the additive noise, and a dynamic term, linked to the time variation of the channel. This leads to an optimal step size, which is related to the minimal mean squared error (MSE) that offers the best asymptotic performance [14], [15], [4], [5].

However, most of the state-of-the-art studies have focused on the asymptotic performance (and the design of the optimal step size), where the algorithm has to track the time-varying channel, and not on the transient performance that is related to the convergence mode. The transient performance has a fundamental role for the original convergence mode (*i.e.*, at the

very beginning of the process) and when the algorithm faces sudden reconfiguration (*i.e.*, when the time-varying channel to be estimated is affected by a sudden change in its model).

To improve the efficiency of the transient behavior of the LMS algorithm, many solutions have exploited self-adaptive step size, where the constant step size of the steepest descent is replaced by a variable step size that is tuned according to some criterion that provides a measure of the adaptation process state. The objective is to have a large step size value during the convergence mode, and a small but well-suited value after convergence (*i.e.* closed to the optimal step-size) [16]. The step-size update can be additive or sometimes multiplicative, and the criteria can be based on counting passages through zero (*i.e.*, on sign changes) [17], [18], on square instantaneous error measures [19], [20], or on other statistics that are based on the filtering of the output of the algorithm [21], [7], [22], [23], [24]. Most of these algorithms also limit the step-size evolution, with a maximal value to ensure convergence, and a minimal value (which is strictly positive in the case of multiplicative updates) to conserve reactivity [25], [26]. These methods often offer very good initial convergence speed, but can suffer from a trade-off between a desired reactivity (in the case of reconfigurations) and obtained MSE after convergence. This trade-off can be embodied by introducing the False Alarm Probability (FAP), defined as the probability for the step to quit its minimal value in asymptotic mode. Monitoring the FAP is matter of importance as false alarms lead to excess MSE, and thus loss of performance. This is particularly the case with additive updates such that the return time (*i.e.*, the time for the step to re-converge after a reconfiguration to its minimal value, which is regarded as an equilibrium) is longer than with multiplicative updates [27].

In this paper, we propose a new adaptive-step process for the LMS algorithm that can cope with reconfigurations and leads to a predictable asymptotic performance, as the step rapidly converges to its minimal limits. The Multiplicative Update with a Forgetting Factor (MUFF) leads to a small number of false alarms and a short return time. We show the benefits of the proposed methods by comparing the MUFF approach with respect to state of the art algorithms in terms of performance and complexity. The model and the LMS-MUFF algorithm are presented in section II. The discussion about the choice of the charge function and the relationship between the FAP and the parameters in the scalar case are discussed in section III. Simulations and comparisons with the classical methods in the literature are presented in section IV. Section V eventually draws some conclusions.

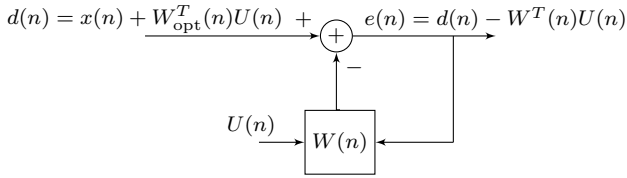


Fig. 1. Noise cancellation structure based on the LMS algorithm.

II. THE LMS-MUFF ALGORITHM

We consider a classical noise reduction context, where the observed complex signal $d(n)$ is expressed as:

$$d(n) = W_{\text{opt}}^T(n)U(n) + x(n), \quad (1)$$

where $U(n)$ is the noise reference input sliding vector with size $L \geq 1$, and $x(n)$ is the desired signal modeled as a white additive complex circular noise of variance σ_x^2 . $W_{\text{opt}}^T(n)$ is assumed to be a Wide Sense Stationary (WSS) time-varying optimal weight vector (of size L) that models the finite impulse response of the channel between reference and observation between two reconfigurations. In noise cancellation problems, which are recalled in Figure 1, the input vector $U(n)$ is known. The (optimal) weight vector $W_{\text{opt}}^T(n)$ is unknown but is adaptively estimated by the L -size weight vector $W(n)$ by minimizing the instantaneous MSE $|e(n)|^2$, which leads to the well-known complex LMS algorithm [2]:

$$e(n) = d(n) - W^T(n)U(n) \quad (2)$$

$$W(n+1) = W(n) + \mu_n U^*(n)e(n), \quad (3)$$

where the real positive coefficient μ_n is the adaptive step size.

The proposed update for the adaptive step size is a MUFF, and it can be expressed as follows:

$$\mu_{n+1} = [\mu_n (\gamma + \varepsilon \|G(n)\|_1)]_{\mu_{\min}}^{\mu_{\max}}, \quad (4)$$

where $0 < \gamma < 1$ is the forgetting factor, $\varepsilon > 0$ is a scale factor, $\|G(n)\|_1 = \sum_{i=1}^L |G_i(n)|$ is the L^1 norm, and $[\bullet]_{\mu_{\min}}^{\mu_{\max}}$ means that the step value is bounded between a maximal value μ_{\max} and a minimal value μ_{\min} , for the reasons explained in I. $G(n)$ is a charge vector that is defined as:

$$G(n) = \frac{\partial W(n)}{\partial \mu}. \quad (5)$$

and can be recursively expressed as:

$$G(n+1) = [1 - \mu_n U^H(n)U(n)]G(n) + U^*(n)e(n) \quad (6)$$

The chosen charge vector, $G(n)$, offers good properties: in asymptotic mode its norm is low, and when the system faces reconfigurations, its norm increases. It can be noted that the LMS-MUFF algorithm is similar to the variable step size (VSS) algorithm [19] and the modified VSS (MVSS) algorithm, as described by [22], whereby the differences lie in the step update and the expression of the charge vector used to update the step size.

If the optimal weight vector is WSS time varying, the asymptotic performance is composed of a static term (relative to the variance of $x(n)$) and a dynamic term (relative to the time variations of the optimal weight). The trade-off between these terms is characterized by an optimal step size of μ_{opt}

that provides minimal asymptotic error variance [5], [11]. In practice, the minimal step size μ_{\min} must be set to have $\mu_{\min} \approx \mu_{\text{opt}}$.

III. ALGORITHM PARAMETERS AND FALSE-ALARM PROBABILITY

A. The principle of MUFF

From Equation (4), the step update can be viewed as the product between the current step μ_n and $K(n)$, which is expressed as:

$$K(n) = \gamma + \varepsilon \|G(n)\|_1 \quad (7)$$

The step update is parameterized by γ and ε . γ controls the return time to equilibrium, as the step-size update is a multiplicative one. And ε has to be chosen for the desired false alarm, as it normalizes the impact of the charge vector norm on the step-size update. The principle of the MUFF is:

- During transient mode, we expect $K(n) > 1$, due to the derivative in (5) implying that $G(n)$ has a large value. The step size increases, which leads to more rapid convergence. The stability of the LMS-MUFF is ensured by the upper limit μ_{\max} [25].
- After convergence, in asymptotic mode, $G(n)$ is expected to reach very weak values, which leads to $\varepsilon \cdot \|G(n)\|_1 < 1 - \gamma$. As a consequence, $K(n)$ becomes < 1 and the step size decreases. As the step is bounded by μ_{\min} , it will converge to this minimal value.

The main difference between our method and the classical methods in the literature is shown by the MUFF. In convergence mode, as the step size decreases and is bounded by a minimal value, it converges to μ_{\min} and remains at μ_{\min} , while $\varepsilon \cdot \|G(n)\|_1$ is lower than $1 - \gamma$. This intrinsic threshold implies that the asymptotic performance of the MUFF algorithm is the same as the asymptotic performance of a constant-step-size LMS parameterized by μ_{\min} . It is however mandatory to use a ε that limits the number of false alarms in asymptotic mode, that occur when $K(n) > 1$ (that leads to $\mu_n > \mu_{\min}$). We denote p_{FA} as the FAP, which is defined as:

$$p_{\text{FA}} = P_r(K > 1) \quad (8)$$

$$p_{\text{FA}} = P_r\left(|G| > \frac{1-\gamma}{\varepsilon}\right). \quad (9)$$

B. Relationship between ε and false-alarm probability in the scalar case

We assume $L = 1$ with an uncorrelated zero mean input, $u(n)$. The charge function, $G(n)$, can be approximated as a complex Gaussian process as Equation (6) can be approximately described as an auto-regressive model with zero mean state noise input, $e(n)u^*(n)$. As a consequence, $K(n)$ is a function of $|G(n)|$, which has a Rayleigh distribution of parameter $\sigma_G/\sqrt{2}$, with σ_G^2 as the variance of the complex Gaussian process $G(n)$.

its cumulative density function (CDF) is expressed as:

$$F(\rho, \sigma_G/\sqrt{2}) = 1 - e^{-\frac{\rho^2}{\sigma_G^2}}. \quad (10)$$

For a given γ and a desired FAP, the maximal acceptable ε , denoted ε_{\max} , is expressed as:

$$p_{\text{FA}} = 1 - F\left(\frac{1-\gamma}{\varepsilon_{\max}}, \sigma_G/\sqrt{2}\right) \\ \iff \varepsilon_{\max} = \frac{1-\gamma}{\sqrt{-\sigma_G^2 \ln(p_{\text{FA}})}}. \quad (11)$$

IV. SIMULATIONS

The relationship between FAP and ε is further analyzed through simulations, and the performance of the proposed algorithm is compared to the classical methods in the literature.

A. Charge function and false alarm in scalar case

1) *Cumulative density function of $G(n)$* : we first consider a scalar case, where the optimal weight W_{opt} is a constant complex coefficient. The input $u(n)$ is assumed to be an uncorrelated unitary variance process, and the observation signal follows Equation (1) with an additive white Gaussian signal $x(n)$ of unitary variance ($\sigma_x^2 = 1$). Figure 2 shows the theoretical and simulated CDF of the real part and the modulus of $G(n)$ defined in Equation (1). The first is compared to a Gaussian CDF with variance $\sigma_G^2/2$, and the second to a Rayleigh distribution with parameter $\sigma_G/\sqrt{2}$, where σ_G^2 is obtained by numerical identification through Monte-Carlo simulations. In both cases, the theory is supported by the simulation.

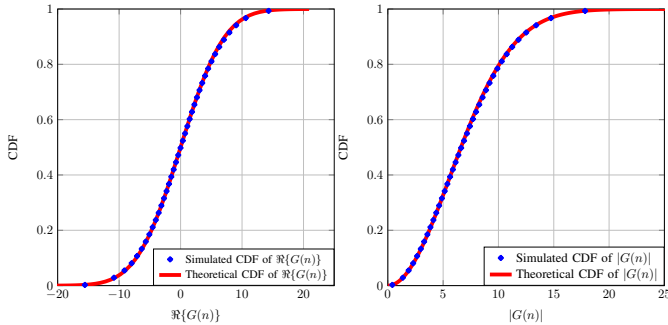


Fig. 2. Left: comparison between simulated CDF of $\Re\{G(n)\}$ and a theoretical Gaussian CDF of variance $\sigma_G^2/2$. Right: comparison between simulated CDF of $|G(n)|$ and a Rayleigh CDF of parameter $\sigma_G/\sqrt{2}$.

2) *Link between empirical charge function variance, LMS step size and signal variance*: we show the relationship between the charge function variance σ_G^2 , the LMS minimal step size μ_{\min} and the signal variance σ_x^2 on Figure 3. For this, we use different values for the step size and the signal variance and evaluate the simulated charge function variance σ_G^2 . Besides, in order to stress out the dependence between these 3 parameters, we also plot the curve $H(\sigma_x^2, \mu_{\min}) = \sigma_x^2/(4 \times \mu_{\min})$. It is shown that in the scalar case, we can approximate σ_G^2 by $H(\sigma_x^2, \mu_{\min})$. It can help in practical cases in order to find an appropriate value of ε for the MUFF overlay.

3) *False-alarm probability in scalar case*: For different values of γ , Figure 4 shows the simulated FAP that is obtained for different desired false-alarm values. For each desired FAP, ε is obtained with Equation (11). The charge function variance is approximated by $\sigma_x^2/(4\mu_{\min})$. We have set $\sigma_x^2 = 1$ and $\mu_{\min} = 10^{-3}$. We see that theory and simulations are in agreement. This also shows that a certain level of false alarms can be maintained with an adequate ε through Equation (11).

4) *Methodology for algorithm parameters specification*: The LMS-MUFF algorithm depends on different parameters linked to the MUFF update or the desired behaviour: $(\gamma, \varepsilon, \mu_{\min}, p_{\text{FA}})$. However, some of these parameters are related one to each other and we propose in this section a simple

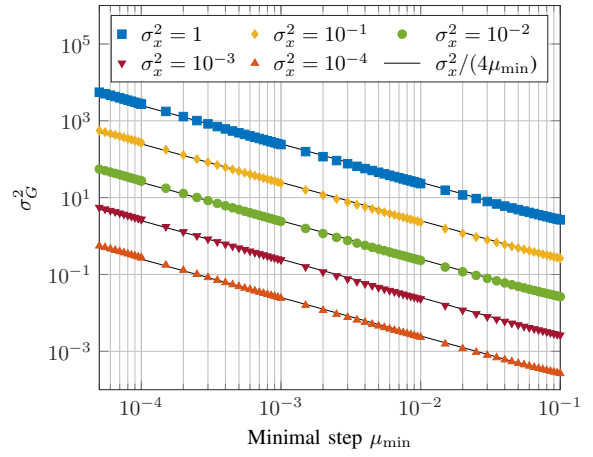


Fig. 3. Simulated relationship between σ_G^2 and μ_{\min} . Approximated relationship $H(\sigma_x^2, \mu_{\min}) = \sigma_x^2/(4 \times \mu_{\min})$ is also plotted for comparison purpose.

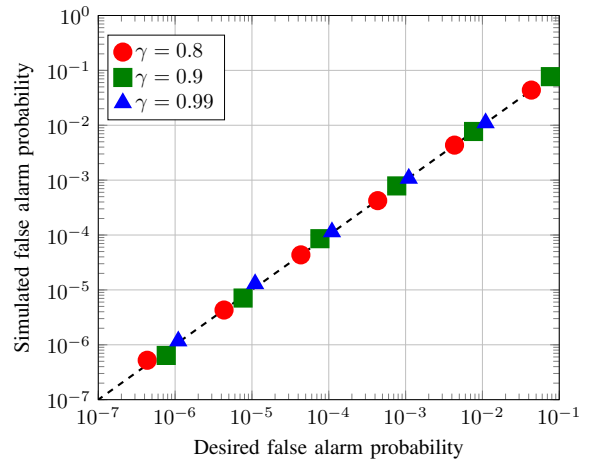


Fig. 4. Desired and obtained FAP for different values of γ and with ε parameterized with Equation (11)

methodology to obtain a complete parametrization in the scalar case:

- The forgetting factor γ should be set to a value close to 1 and pilots the return time to equilibrium. A classical choice in the literature for forgetting factors is 0.99 [5].
- The minimal step size μ_{\min} is chosen to ensure a targeted asymptotic performance (or targeted MSE) for a constant step LMS. The link between the step size and the asymptotic performance is model dependent (see [11] for an example of MSE that is expressed with a time-varying model for LMS-based noise-cancellation problems). If the model is static (or slowly time varying); the MSE can be approximated as $\mu_{\min}\sigma_x^2\sigma_u^2/2$ (with σ_u^2 the variance of the input vector) [28].
- For a given p_{FA} ; ε is then deduced from the aforementioned choices (using Equation (11)) as the charge function variance can be approximated in the scalar case as $\sigma_G^2 \approx \sigma_x^2/(4\mu_{\min})$.

From the key aforementioned parameters $(\gamma, \mu_{\min}, p_{\text{FA}})$, the first one γ is a constant easy to tune and the second parameter μ_{\min} is chosen in the same way as with a constant step approach. Hence; the proposed overlay only relies on one key design parameter; the false alarm probability p_{FA} .

	Additions	Multiplications
LMS	$2L + 10$	$8L$
VSS	$2L + 16$	$8L + 4$
MVSS	$2L + 18$	$8L + 11$
MMSD - NLMS	$5L + 3$	$14L + 2$
LMS-MUFF	$9L + 18$	$13L + 5$

TABLE I
COMPLEXITY COMPARISON (REAL ADDITIONS AND REAL
MULTIPLICATIONS) BETWEEN METHODS

B. Comparison with the methods from the literature (multi-tap case)

We now compare our method with three methods from the literature: the VSS presented in [19], the MVSS presented in [22] and the MMSD-NLMS presented in [20]. For this, we use the same simulation case and the same algorithm parameters as in [22]: W_{opt}^T is a time-invariant moving-average model with four taps, $U(n)$ is a zero-mean white Gaussian reference with unitary variance, and $x(n)$ is the desired signal of unitary variance ($\sigma_x^2 = 1$). The estimation is performed with a LMS algorithm with $L = 4$. Also, to test the reactivity of the different methods, a reconfiguration is applied at the middle of the simulation (*i.e.*, a sudden variation of the optimal weight vector).

To have an asymptotic MSE of approximatively -31 dB, as in [22], we chose $\mu = 3.5 \cdot 10^{-4}$, for constant-step LMS. For the MUFF algorithm, we chose $\mu_{\min} = 3.5 \cdot 10^{-4}$ (same as the constant-step LMS), $\gamma = 0.99$, and $\varepsilon = 10^{-5}$. As specified in [22], for the VSS algorithm, we chose $\alpha = 0.97$, $\beta = 0.99$, and $\gamma = 10^{-3}$. We also computed the MVSS algorithm [22] with $\gamma = 10^{-2}$, to have a more rapid transient performance. For the MMSD-NLMS algorithm [20], we use $\alpha = 0.99$, $\mu_{\max} = 0.1$, and $C = 10$ to have the same asymptotic performance as the constant step LMS.

Figure 5 shows the evolution of the MSE (expressed in dB). It is shown that, as expected, our method offers the same asymptotic performance as a classical constant step-size LMS, but with a faster convergence mode. The asymptotic performance of the proposed method is thus equal to the performance of a constant-step LMS, with constant step to minimal value μ_{\min} . In configuration $\gamma = 10^{-2}$ for MVSS, the literature method is as fast as the LMS-MUFF, but it does not reach the same asymptotic level. However, when algorithms from the literature are tuned to provide the same asymptotic performance as constant-step LMS, the LMS-MUFF offers better transient performance. The proposed method thus offers a better trade-off between speed and accuracy both for initial convergence and after a reconfiguration.

C. Complexity

In Table I, we assess the computational complexity for the aforementioned methods with respect to the number of LMS taps L . Complexity is assessed in terms of real additions and real multiplications. We assume that a complex multiplication can be carried out using three real multiplications and four real additions[29].

The lower complexity increases (w.r.t LMS with constant step) is obtained with VSS algorithm. This algorithm also offers the lower reactivity in case of a reconfiguration (see Figure 5). One can see a difference in terms of complexity increase between VSS and MVSS in one hand and MMSD-NLMS and LMS-MUFF at the other hand. For the former, the step-size updates only use the output of the LMS algorithm ($\{e(n)\}$ for VSS algorithm and $\{e(n), e(n-1)\}$ for MVSS algorithm). For the latter, the step-size updates are computed with the use of the entire reference vector. This is done through the use of a charge function (denoted as \hat{p} in [20] and as $G(n)$ in our paper). This allows a better behaviour as demonstrated in Figure 5. Among the compared methods, LMS-MUFF offers an interesting trade-off between performance (best reactivity and good asymptotic performance) and complexity management (comparable overhead as MMSD-NLMS).

V. CONCLUSIONS

This study focuses on a new adaptive step-size LMS algorithm, which is based on a multiplicative update with forgetting factor (the LMS-MUFF). LMS algorithms are guided by their step size, which leads to a trade-off between the speed and asymptotic final excess mean squared error. We have proposed a new variable step-size overlay, which is bounded by a minimal step, and which offers the same asymptotic performance as a constant-step LMS (with a constant step to the minimal value) but with a more rapid convergence mode, and a strong ability to cope with sudden reconfigurations. In the scalar case, we have derived the semi-analytical relationship between the LMS-MUFF parameters and the false alarm probability, and we have compared our method with classical methods from the literature. It shows that, when tuned to have comparable asymptotic performance, our method offers better speed and reactivity.

REFERENCES

- [1] B. Widrow, P. Mantey, L. Griffiths, and B. Goode, "Adaptive antenna systems," *Proceedings of the IEEE*, vol. 55, no. 12, pp. 2143–2159, Dec 1967.
- [2] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 719 – 720, april 1975.
- [3] S. Boll and D. Pulsipher, "Suppression of acoustic noise in speech using two microphone adaptive noise cancellation," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 28, no. 6, pp. 752–753, Dec 1980.
- [4] A. Mader, H. Puder, and G. U. Schmidt, "Step-size control for acoustic echo cancellation filters : an overview," *Signal Processing*, vol. 80, no. 9, pp. 1697 – 1719, 2000.
- [5] R. Gerzaguet, L. Ros, F. Belveze, and J.-M. Brossier, "On the performance of digital adaptive spur cancellation for multi-standard radio frequency transceivers," *Digital Signal Processing*, vol. 33, pp. Pages 83–97, 2014.
- [6] S. Qureshi, "Adaptive equalization," *Proceedings of the IEEE*, vol. 73, no. 9, pp. 1349–1387, Sept 1985.
- [7] B. Geller, V. Capellano, J. Brossier, A. Essebbat, and G. Jourdain, "Equalizer for video rate transmission in multipath underwater communications," *IEEE Journal of Oceanic Engineering*, vol. 21, no. 2, pp. 150–155, Apr 1996.
- [8] M. Morelli, L. Sanguinetti, and U. Mengali, "Channel estimation for adaptive frequency-domain equalization," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2508–2518, Sept 2005.
- [9] M. Valkama, M. Renfors, and V. Koivunen, "On the performance of interference canceller based I/Q imbalance compensation," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 5, 2000, pp. 2885–2888 vol.5.
- [10] A. Frotzschner and G. Fettweis, "A stochastic gradient LMS algorithm for digital compensation of Tx leakage in Zero-IF-Receiver," in *Proc. IEEE Vehicular Technology Conference (VTC)*, 2008, pp. 1067–1071.

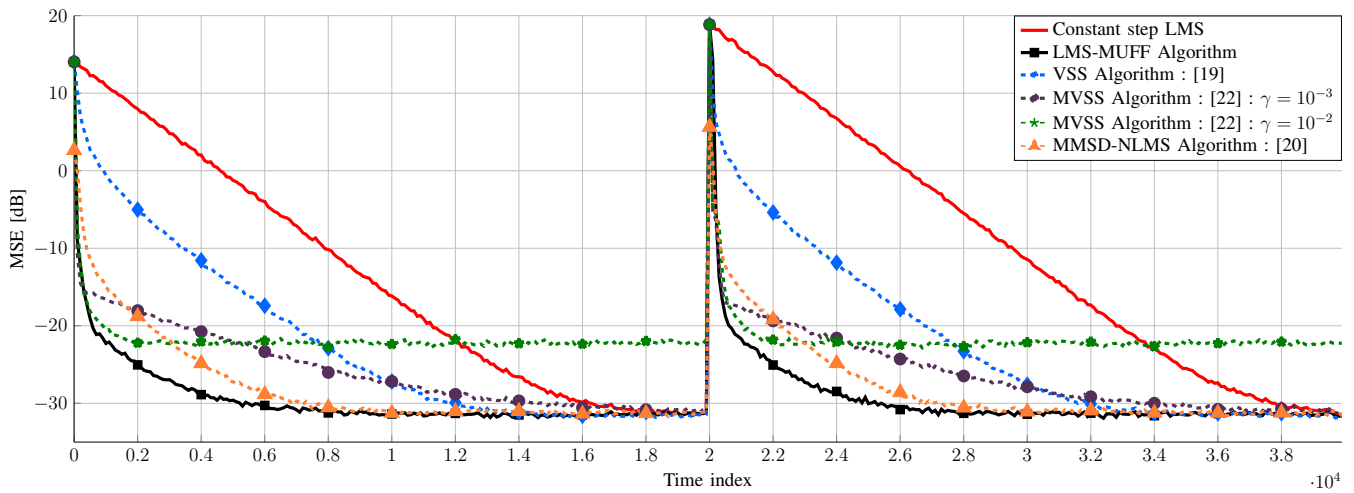


Fig. 5. Comparison of MSE for the MUFF algorithm and the different methods in the literature, with reconfiguration at the middle of the simulation.

- [11] R. Gerzaguet, L. Ros, F. Belveze, and J.-M. Brossier, "Performance of a digital transmitter leakage LMS-based cancellation algorithm for multi-standard radio-frequency transceivers," *Digital Signal Processing*, vol. 51, pp. 35 – 46, 2016.
- [12] M. Hajivandi and W. Gardner, "Measures of tracking performance for the LMS algorithm," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 11, pp. 1953–1958, 1990.
- [13] P. Wei, J. Han, J. Zeidler, and W. Ku, "Comparative tracking performance of the LMS and RLS algorithms for chirped narrowband signal recovery," *IEEE Transactions on Signal Processing*, vol. 50, no. 7, pp. 1602–1609, 2002.
- [14] N. Bershad and O. Macchi, "Adaptive recovery of a chirped sinusoid in noise. II. performance of the LMS algorithm," *IEEE Transactions on Signal Processing*, vol. 39, no. 3, pp. 595–602, 1991.
- [15] O. Macchi, *Adaptive Processing: The Least Mean Squares Approach with Applications in Transmission*, Wiley, Ed., 1995.
- [16] A. Benveniste and P. Metivier, *Adaptive algorithms and stochastic approximations*. Springer, 1991.
- [17] R. Harris, D. Chabries, and F. Bishop, "A variable step (vs) adaptive filter algorithm," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, no. 2, pp. 309–316, 1986.
- [18] B. Farhang-Boroujeny, "Variable-step-size LMS algorithm: new developments and experiments," *IEE Proceedings on Vision, Image and Signal Processing*, vol. 141, no. 5, pp. 311–317, 1994.
- [19] R. Kwong and E. Johnston, "A variable step size LMS algorithm," *IEEE Transactions on Signal Processing*, vol. 40, no. 7, pp. 1633 –1642, jul 1992.
- [20] S. Zhao, D. Jones, S. Khoo, and Z. Man, "New variable step-sizes minimizing mean-square deviation for the LMS-type algorithms," *Circuits, Systems, and Signal Processing*, vol. 33, no. 7, pp. 2251–2265, 2014.
- [21] V. Mathews and Z. Xie, "Stochastic gradient adaptive filters with gradient adaptive step sizes," in *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 1990, pp. 1385–1388 vol.3.
- [22] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm: analysis and simulations," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 631–639, 1997.
- [23] Z. Ramadan and A. Poularikas, "Performance analysis of a new variable step-size LMS algorithm with error nonlinearities," in *Proceedings of the Thirty-Sixth Southeastern Symposium on System Theory*, 2004, pp. 384–388.
- [24] T. Yang and B. Shahrrava, "Performance of variable step-size LMS algorithms for linear adaptive inverse control systems," in *Proc. Canadian Conference on Electrical and Computer Engineering*, 2005, pp. 755–758.
- [25] S. Gelfand, Y. Wei, and J. Krogmeier, "The stability of variable step-size LMS algorithms," *IEEE Transactions on Signal Processing*, vol. 47, no. 12, pp. 3277–3288, 1999.
- [26] S.-L. Goh and D.-P. Mandic, "Stochastic Gradient-Adaptive Complex-Valued nonlinear neural adaptive filters with a Gradient-Adaptive step size," *IEEE Transactions on Neural Networks*, vol. 18, no. 5, pp. 1511–1516, Sep. 2007.
- [27] A. Wee-Peng and B. Farhang-Boroujeny, "A new class of gradient adaptive step-size LMS algorithms," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, pp. 805–810, Apr. 2001.
- [28] S. Haykin, *Adaptive filter theory*. Prentice Hall, 1996.
- [29] S. G. Krantz, *Handbook of Complex Variables*. Birkhäuser Basel, 1999.