

Distributed Adaptive LMF Algorithm for Sparse Parameter Estimation in Gaussian Mixture Noise

Mojtaba Hajibabadi and Hossein Zamiri-Jafarian

Electrical Engineering Department
Ferdowsi University of Mashhad
Mashhad, Iran
mhajibabadi@gmail.com

Abstract— A distributed adaptive algorithm for estimation of sparse unknown parameters in the presence of nonGaussian noise is proposed in this paper based on normalized least mean fourth (NLMF) criterion. At the first step, local adaptive NLMF algorithm is modified by zero norm in order to speed up the convergence rate and also to reduce the steady state error power in sparse conditions. Then, the proposed algorithm is extended for distributed scenario in which more improvement in estimation performance is achieved due to cooperation of local adaptive filters. Simulation results show the superiority of the proposed algorithm in comparison with conventional NLMF algorithms.

Keywords— Cooperative estimation; Gaussian mixture noise; NLMF algorithm; sparse parameters; zero norm

I. INTRODUCTION

Adaptive Filters are using in large applications to endow a system with learning and tracking abilities, especially when the signal statistics are unknown and are expected to vary with time. Over the last several years, a wide range of adaptive algorithms has been developed for diverse demands such as channel equalization, spectral estimation, target localization, and interference cancellation. One group of the basic adaptive algorithms is gradient-based algorithms such as the least mean square (LMS) algorithm. The well-known LMS algorithm is perhaps one of the most familiar and widely used algorithms because of its good performance in many circumstances and its simplicity of implementation. The books [1], [2] should provide an excellent sense of the main concepts in this area.

However, the LMS algorithm is a popular method for adaptive parameter estimation, in many scenarios parameters of unknown systems can be assumed to be sparse, containing only a few large coefficients spreaded among many small ones. Using such prior information about the sparsity of unknown parameters can be helpful to improve estimation performance, but standard LMS filters do not exploit such sparsity information. In the past years, several algorithms have been proposed for sparse adaptive filtering using LMS, which was motivated by recent progress in compressive sensing [3]. The basic idea of these techniques is to introduce a penalty into the cost function of the standard LMS to notice sparsity. This achieves better performance than that of the standard LMS for sparse models [4].

Many approaches for signal processing problems have been studied when the additive noise process is modeled with Gaussian distribution. However, for many real-life situations, the additive noise of the system is found to be dominantly nonGaussian. Some examples of nonGaussian environments are the ocean acoustic noise and the urban radio-frequency (RF) noise. Also, in processing of radar and sonar signals we need to deal with nonGaussian noise [5]. When the additive noise process is nonGaussian, LMS algorithm has a poor performance. In [6], it was shown that for some environments with nonGaussian noise, Least Mean Fourth (LMF) algorithm outperforms LMS Algorithm. One of the main drawbacks of the LMF algorithm is, its stability problem. In [7] Eweda proposed NLMF algorithm that overcomes the stability problem.

It is clear that using a distributed estimation approach, by cooperation between local adaptive filters, provides spatial diversity that results in better performance in comparison with local adaptive filter [8],[9]. In this paper, after modifying NLMF adaptive filter for sparse parameter estimation, it is extended for distributed scenario when additive noise of system is Gaussian mixture noise. Computer simulation results show that our proposed adaptive algorithm achieves better performance compared to the conventional adaptive NLMF algorithms.

This paper is organized as follows. After the introduction, sparse NLMF algorithm for estimation of sparse problems in the presence of nonGaussian noise is developed in section II. A novel distributed sparse NLMF algorithm for a finite mixture of Gaussian noise is proposed in section III. Finally, Simulation and comparison results are given in section IV, followed by conclusions in section V.

II. ADAPTIVE SPARSE NLMF ALGORITHM

Assum that signal $X(n)$ is the input of the system in Fig. 1, with sparse unknown parameters column vector

$$W^o = [w_1, w_2, \dots, w_N]^T \quad (1)$$

that $(\cdot)^T$ represents the transpose operator. An observation of output signal is

$$d(n) = X(n)^T W^o + z(n) \quad (2)$$

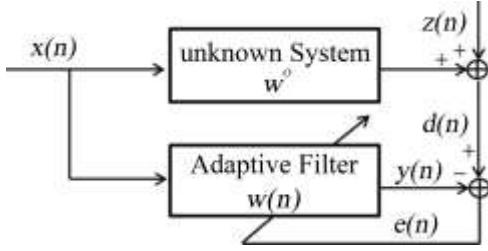


Fig. 1. The block diagram of FIR adaptive filter.

where $X(n) = [x(n), x(n-1), \dots, x(n-N+1)]^T$ denotes the vector of input signal and $z(n)$ is the observation noise assumed to be independent of the input signal $x(n)$. The objective of the algorithm is to estimate the sparse unknown vector, W^o , using the input signal $X(n)$ and desired output signal $d(n)$. Let $W(n) = [w_1(n), w_2(n), \dots, w_N(n)]^T$ be the estimated vector of the adaptive filter at iteration n . In the standard LMF, the cost function $J(n)$ is defined as

$$J(n) = \frac{1}{4} e^4(n) \quad (3)$$

where $e(n)$ is the instantaneous error determined as

$$e(n) = d(n) - y(n) \quad (4)$$

in which $y(n)$ is the output of adaptive filter and it is equal to $y(n) = X(n)^T W(n)$. The filter coefficients vector is then updated by

$$\begin{aligned} W(n+1) &= W(n) - \mu_f \frac{\partial J(n)}{\partial W(n)} \\ &= W(n) + \mu_f e^3(n) X(n) \end{aligned} \quad (5)$$

The stability of (5) is dependent on the statistics of the input and noise signals. In [7] Eweda proposed Normalised version of LMF algorithm, which solves the stability problem of LMF algorithm, as follows:

$$W(n+1) = W(n) + \frac{\mu_f e^3(n) X(n)}{\|X(n)\|_2^2 (\|X(n)\|_2^2 + e^2(n))} \quad (6)$$

in which, the normalizing term is fourth-order in the regressor and second-order in the estimation error, make the algorithm to be stable against the increase of input power and noise power, respectively. The step-size μ_f , controls the transient and steady-state behavior of the algorithm. It is well-known that, $\|\cdot\|_2$ denotes Euclidean norm operator. Equation (6) can be rewritten as NLMS with variable step size $\mu_f(n)$,

$$\begin{aligned} W(n+1) &= W(n) + \frac{\mu_f e^2(n)}{\|X(n)\|_2^2 + e^2(n)} \cdot \frac{e(n) X(n)}{\|X(n)\|_2^2} \\ &= W(n) + \mu_f(n) \cdot \frac{e(n) X(n)}{\|X(n)\|_2^2} \end{aligned} \quad (7)$$

To exploit the unknown parameters sparsity, a new cost function, $J_{sparse}(n)$, is defined by combining the instantaneous fourth-order error with the ℓ_0 -norm penalty,

$$J_{sparse}(n) = \frac{1}{4} e^4(n) + \lambda_f \|W(n)\|_0 \quad (8)$$

where, λ_f , is a regularization parameter, which represents a trade off between estimation error and sparsity of the parameters. Operator $\|\cdot\|_0$ denotes zero-norm operator, which counts the number of nonzero coefficients of vector $W(n)$. Since solving zero-norm algorithm is a hard problem, the zero norm is generally approximated by a continuous function [4]. A popular approximation is:

$$\|W(n)\|_0 \approx \sum_{i=1}^N (1 - e^{-\beta|w_i(n)|}) \quad (9)$$

which leads to the following gradient column vector:

$$\frac{\partial \|W(n)\|_0}{\partial W(n)} = \beta \cdot \text{diag}\{e^{-\beta|w_1(n)|}, \dots, e^{-\beta|w_N(n)|}\} \cdot \text{sgn}(W(n)) \quad (10)$$

in which $\text{diag}\{\cdot\}$ represents a diagonal matrix and $\text{sgn}(\cdot)$ is a sign function. Using the gradient descent updating, the sparse NLMF filter update is given as

$$\begin{aligned} W(n+1) &= W(n) \\ &\quad + \frac{\mu_f e^3(n) X(n)}{\|X(n)\|_2^2 (\|X(n)\|_2^2 + e^2(n))} + \text{Sparse Penalty} \end{aligned} \quad (11)$$

By exerting sparse penalty to the standard NLMF cost function, the solution will be sparse and the gradient descent recursion will accelerate the convergence rate of near-zero coefficients in the sparse system. Equation (11) can be rewritten as follows

$$\begin{aligned} W(n+1) &= W(n) + \frac{\mu_f e^3(n) X(n)}{\|X(n)\|_2^2 (\|X(n)\|_2^2 + e^2(n))} \\ &\quad - \frac{\mu_f}{\|X(n)\|_2^2 (\|X(n)\|_2^2 + e^2(n))} \lambda_f \frac{\partial \|W(n)\|_0}{\partial W(n)} \end{aligned} \quad (12)$$

In the next section, the proposed distributed algorithm is presented based on (12) as an update equation for each local adaptive filter.

III. ADAPTIVE DISTRIBUTED SPARSE NLMF ALGORITHM

It is well-known that using a distributed estimation approach provides spatial diversity that results in better performance in comparison with local adaptive filter [8]. In this section, we propose a novel distributed sparse adaptive algorithm to further improve the filtering performance in the presence of a Gaussian mixture noise. Obviously, the effectiveness of any distributed scenario will depend on cooperation that are allowed among the nodes [9]. Here, we consider a network with K adaptive filters that the goal of each one is to estimate

the unknown sparse vector, W^o . The k th adaptive filter has access to local data $\{d_k(n), X_k(n)\}$ and the sparse unknown vector, W^o , relates to the local data as

$$d_k(n) = X_k(n)^T W^o + z_k(n) \quad (13)$$

where $z_k(n)$ is Gaussian mixture noise that is white in both space and time. In diffusion strategy [8], as shown in Fig. 2, each node k , cooperates with its neighborhood nodes. N_k , is defined as a set of nodes linked to node k , including k itself. In this way, node k combines its *local estimate*, $W_k(n)$ with its neighbors' estimates, $\{W_\ell(n), \ell \in N_k\}$, as follows,

$$\phi_k(n) = \sum_{\ell \in N_k} c_{\ell k} W_\ell(n) \quad (14)$$

The coefficient $c_{\ell k}$ denotes the weight of cooperation, which is chosen as,

$$c_{\ell k} = \begin{cases} \frac{1}{|N_k|} & , \ell \in N_k \\ 0 & , \ell \notin N_k \end{cases} \quad (15)$$

where $|N_k|$ is defined as the size of set N_k . If we define local error signal as:

$$e_k(n) = d_k(n) - \phi_k(n)^T X_k(n) \quad (16)$$

Based on (12), the distributed algorithm recursions can be given as:

$$\begin{aligned} W_k(n+1) &= \phi_k(n) \\ &+ \frac{\mu_f}{\|X_k(n)\|_2^2 (\|X_k(n)\|_2^2 + e_k^2(n))} e_k^3(n) X_k(n) \\ &- \frac{\mu_f}{\|X_k(n)\|_2^2 (\|X_k(n)\|_2^2 + e_k^2(n))} \\ &\times \lambda_f \beta \cdot \text{diag}\{e^{-\beta|\phi_{k_1}(n)|}, \dots, e^{-\beta|\phi_{k_N}(n)|}\} \cdot \text{sign}(\phi_k(n)) \end{aligned} \quad (17)$$

Local Mean Square Deviation (MSD) is defined as a performance criterion,

$$V_k(n) = E \|W^o - W_k(n)\|^2 \quad (18)$$

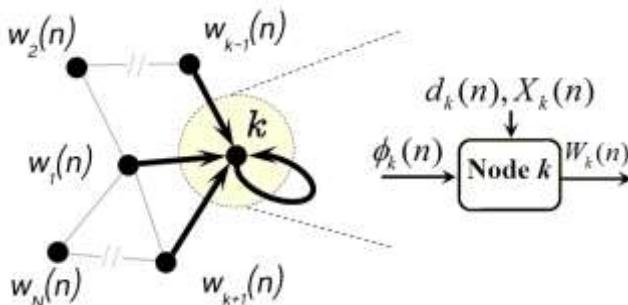


Fig. 2. Distributed network consist of local adaptive filters.

For comparing the performance of the proposed distributed adaptive algorithm (17) with local adaptive filters described in the previous section, we average the local MSD, $V_k(n)$, over the nodes for each time iteration n , and so total MSD is defined as,

$$V(n) = \frac{1}{N} \sum_{k=1}^N V_k(n) \quad (19)$$

IV. COMPUTER SIMULATIONS

In this section, we present our simulations of adaptive algorithms described in the previous sections. For all of the simulations, the input signal $x(n)$ is assumed to be white Gaussian process with unit variance. Sparse parameters vector is chosen as :

$$W^o = [0, 0.9, 0.03, 0.7, 0.01, 0, 0.09, 0, 0, 0.01, 0, 0, 0.01, 0, 0.015, 0]^T, \quad (20)$$

that was shown in Fig. 3. The sparsity ratio of W^o is equal to 2/16 which means vector, W^o contain only 2 large coefficients. The observation noise, assumed to be independent in both space and time, is modeled as a Gaussian mixture defined by [10],

$$f_Z(z) = \sum_{i=1}^3 a_i N(z | \mu_i, \sigma_i^2) \quad (21)$$

that was shown in Fig. 3, where the weights are all positive and sum to one.

$$\sum_{i=1}^3 a_i = 1, \quad a_i \geq 0 \quad (22)$$

The parameters of Gaussian mixture noise are given in the Table 1. For simulation of (17) we consider a network consist of $K = 10$ nodes and $|N_k| = 4$, as shown in Fig. 4.

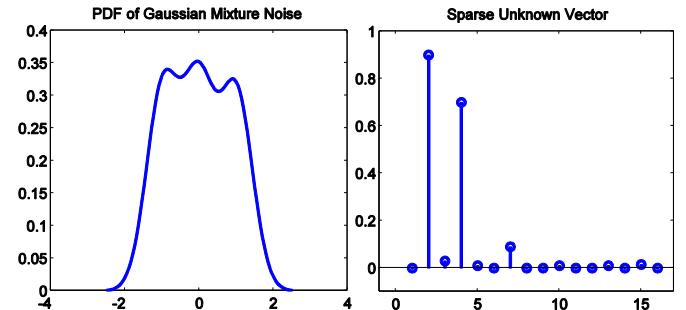


Fig. 3. Probability density function of Gaussian mixture noise (left), and Sparse parameters vector (20) with sparsity ratio 2/16 (Right).

TABLE I. GAUSSIAN MIXTURE MODEL

| Gaussian Number | Gaussian Mixture Components | |
|-----------------|-----------------------------|-----------------------|
| | Mean μ_i | Variance σ_i^2 |
| $i = 1$ | -1 | 0.01 |
| $i = 2$ | 0 | 0.01 |
| $i = 3$ | 1 | 0.01 |

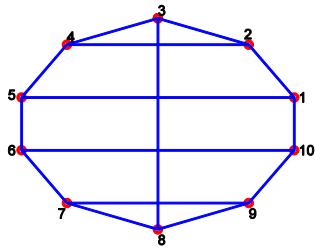


Fig. 4. Distributed network with $K=10$ nodes and $|N_k|=4$.

The performance of the proposed adaptive sparse NLMF algorithms evaluated by computer simulations are shown in Fig. 5, and Fig. 6, for local and distributed scenario, respectively.

In Fig. 5, the local adaptive NLMF algorithm has been simulated based on (12) for Gaussian mixture noise. As seen, our proposed NLMF has superior performance in comparison with NLMS algorithm both in convergence rate and steady state estimation error power.

Fig. 6, shows the performance of the proposed distributed adaptive NLMF algorithm simulated based on (17). As seen, the distributed NLMF algorithm outperforms other local adaptive algorithms regarding convergence rate and MSD criteria.

V. CONCLUSIONS

A distributed adaptive Normalized Least Mean Fourth (NLMF) algorithm for sparse estimation in Gaussian mixture noise has been proposed in this paper. At first step, local adaptive NLMF algorithm has been modified for sparse estimation by zero norm criterion in order to speed up the convergence rate and also to reduce the steady state estimation error power. Then, the proposed algorithm has been extended to distributed scenario in which more improvement in performance has been achieved due to spatial diversity gain. It should be note that although the proposed distributed algorithm has a better performance in both transient and steady states, the computational complexity of distributed algorithms is higher than local ones.

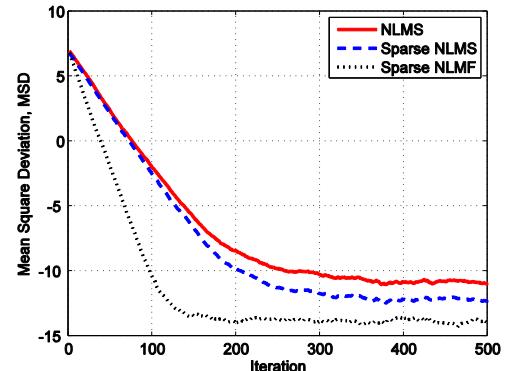


Fig. 5. MSD performances of NLMS, Sparse NLMS and Sparse NLMF algorithms.

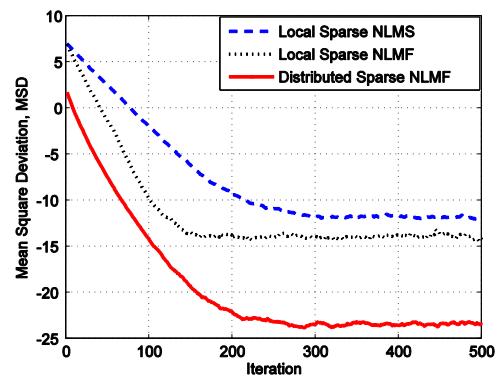


Fig. 6. MSD performances of Local Sparse NLMS, Local Sparse NLMF and Distributed Sparse NLMF algorithms.

REFERENCES

- [1] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*. New York: John Wiley, 1998.
 - [2] A. H. Sayed, *Fundamentals of Adaptive Filtering*. New York: Wiley, 2003.
 - [3] Y. Chen, Y. Gu, and A. O. Hero, "Sparse LMS for system identification," in *Proc. IEEE ICASSP*, pp. 3125–3128, Taipei, Taiwan, Apr. 2009.
 - [4] Y. Gu, J. Jin, and S. Mei, "L₀-norm constraint LMS algorithm for sparse system identification," *IEEE Signal Process. Lett.*, vol. 16, no. 9, pp. 774–777, Sep. 2009.
 - [5] Y. Yardimci, A. E. Cetin, and J. A. Cadzow, "Robust direction-of-arrival estimation in nonGaussian noise," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1443–1451, May 1998.
 - [6] E. Walach and B. Widrow, "The least mean fourth (LMF) adaptive algorithm and its family," *IEEE Trans. Inf. Theory.*, vol. 30, no. 2, pp. 275–283, Mar. 1984.
 - [7] E. Eweda, "Global stabilization of the least mean fourth algorithm," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1473–1477, Mar. 2012.
 - [8] F. S. Cattivelli and A. H. Sayed, "Diffusion LMS strategies for distributed estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1035–1048, Mar. 2010.
 - [9] C. G. Lopes and A. H. Sayed, "Incremental adaptive strategies over distributed networks," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4064–4077, Aug. 2007.
 - [10] D. A. Reynolds and R. C. Rose, "Robust text-independent speaker identification using gaussian mixture speaker models," *IEEE Trans. Speech and Audio Process.*, vol. 3, no. 1, pp. 72–83, Jan. 1995.