

This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Distributed adaptive estimation based on the APA algorithm over diffusion networks with changing topology

PLEASE CITE THE PUBLISHED VERSION

PUBLISHER

© IEEE

VERSION

VoR (Version of Record)

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Li, Leilei, and Jonathon Chambers. 2019. "Distributed Adaptive Estimation Based on the APA Algorithm over Diffusion Networks with Changing Topology". figshare. <https://hdl.handle.net/2134/5858>.

This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.



CC creative commons
COMMONS DEED

Attribution-NonCommercial-NoDerivs 2.5

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:

BY: **Attribution.** You must attribute the work in the manner specified by the author or licensor.

Noncommercial. You may not use this work for commercial purposes.

No Derivative Works. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

DISTRIBUTED ADAPTIVE ESTIMATION BASED ON THE APA ALGORITHM OVER DIFFUSION NETWORKS WITH CHANGING TOPOLOGY

LeiLei Li and Jonathon A. Chambers

Advanced Signal Processing Group, Electronic & Electrical Engineering Department,
Loughborough University, {l.li2, j.a.chambers}@lboro.ac.uk

ABSTRACT

In this paper, we present a novel distributed affine projection algorithm (APA) to solve distributed estimation problem within dynamic diffusion networks. In addition, mean-square stability of the proposed algorithm is also studied through exploitation of the energy conservation approach due to Sayed. Simulations confirm that the novel algorithm achieves a greatly improved performance as compared with a noncooperative scheme.

Index Terms— Adaptive filtering, distributed adaptive estimation, affine projection, diffusion networks, dynamic networks

1. INTRODUCTION

In a network of nodes spreading over a physical area, the objective is to estimate the parameters of interest by exploiting observations of temporal data collected from nodes with different spatial locations, statistical profiles of which are possibly different. However, applications in some environments suffer limited capabilities of communications and complexity due to tight energy and bandwidth constraints, especially in wireless sensor networks. Such constraints lead to the development of distributed adaptive algorithms based on least mean square (LMS) and recursive least squares (RLS) rules for incremental networks as in [1], [2]. It is well known that an incremental topology requires a Hamiltonian cycle to be established to connect all the nodes in the network at every iteration. This mode of cooperation is only likely to be suitable for small size networks and limits the autonomy of the network topology. Therefore, peer-to-peer distributed algorithms based on diffusion protocols have been derived and studied in [3], [4], which take advantage of the cooperation among the individual adaptive nodes.

In the case of a single adaptive filter, the advantage of the APA algorithm is well known: it achieves an improved convergence performance as compared with the LMS algorithm for some coloured input signals but obtains a reasonable convergence performance by using less computational cost than the RLS algorithm for certain tap-lengths. In [5], a new incremental distributed algorithm based on an APA rule was

developed with the purpose to obtain a good balance between computational cost and convergence performance. As a consequence, a new diffusion type learning algorithm relying on the APA rule is derived in this paper to solve the distributed estimation problem, where the cooperation strategy between nodes is a peer-to-peer diffusion protocol and nodes communicate only with their peer neighbours at every iteration.

2. ESTIMATION PROBLEM

Consider the distributed estimation problem as in [6]: the objective is to seek an unknown $M \times 1$ vector w over an N -node network by solving

$$\min_w E \|\mathbf{d} - \mathbf{U}w\|^2 \quad (1)$$

where two global data matrices

$$\mathbf{d} = \text{col}\{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_N\}, \quad (N \times 1) \quad (2)$$

$$\mathbf{U} = \text{col}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N\}, \quad (N \times M) \quad (3)$$

are formulated by the zero-mean random measurement data \mathbf{d}_k and regression data \mathbf{u}_k , the time realizations of which are denoted by $\{d_{k,i}, u_{k,i}\}$, $k = 1, \dots, N$ with time index i . The optimal solution w^o satisfying the normal equations [7] is therefore obtained by solving

$$R_{du} = R_u w^o \quad (4)$$

with $R_u = E\mathbf{U}^*\mathbf{U}$ and $R_{du} = E\mathbf{U}^*\mathbf{d}$.

Our purpose is to design a distributed adaptive scheme relying on the APA rule to achieve a good estimate approximating the solution w^o of (4) at every node within the network. Let $\psi_k^{(i-1)}$ denote the local estimate at node k in time $i-1$. Node k therefore obtains a set of unbiased estimates $\{\psi_l^{(i-1)}\}_{l \in \mathcal{N}_{k,i-1}}$ from its neighborhood $\mathcal{N}_{k,i-1}$, which is defined as the set of all nodes connecting to node k at any given time $i-1$. At node k , a local combining function f_k is used to fuse these local estimates, yielding the aggregate estimate $\phi_k^{(i-1)}$,

$$\phi_k^{(i-1)} = f_k(\psi_l^{i-1}; l \in \mathcal{N}_{k,i-1}) \quad (5)$$

where $\mathcal{N}_{k,i-1}$ is time dependent and includes node k itself. In this work, we use linear combiners and f_k is therefore replaced by some weighted combination, yielding

$$\phi_k^{(i-1)} = \sum_{l \in \mathcal{N}_{k,i-1}} c_{k,l}(i-1) \psi_l^{(i-1)} \quad (6)$$

with weighted coefficients $c_{k,l}(i-1) \geq 0$ and $\sum_{l \in \mathcal{N}_{k,i-1}} c_{k,l}(i-1) = 1$. Let $n_{k,i-1}$ and $n_{l,i-1}$ denotes the degree of the neighbourhoods, i.e. $n_{k,i-1} = |\mathcal{N}_{k,i-1}|$. One choice for the combiner is the Metropolis rule, as in [8], if k links l

$$c_{k,l}(i-1) = \begin{cases} \frac{1}{\max(n_{k,i-1}, n_{l,i-1})}, & k \neq l \\ 1 - \sum_{l \in \mathcal{N}_{k,i-1}, l \neq k} c_{k,l}(i-1), & k = l \end{cases} \quad (7)$$

otherwise $c_{k,l}(i-1) = 0$ if k does not link l . Other combiners are the Laplacian and the nearest neighbour rules.

Recall the update recursion of the diffusion LMS strategy as in [3], which can be rewritten in theory as follows:

$$\psi_k^{(i)} = \phi_k^{(i-1)} - \mu_k [\nabla J_k(\phi_k^{(i-1)})]^* \quad \text{for node } k \quad (8)$$

where μ_k denotes the local step-size. With the purpose of improving the convergence performance, a Newton's search based approach [7] is therefore adopted in (8), yielding,

$$\begin{aligned} \psi_k^{(i)} &= \phi_k^{(i-1)} - \mu_k [\epsilon I + \nabla^2 J_k(\phi_{k-1}^{(i)})]^{-1} [\nabla J_k(\phi_{k-1}^{(i)})]^* \\ &= \phi_k^{(i-1)} + \mu_k (\epsilon I + R_{u,k})^{-1} [R_{du,k} - R_{u,k} \phi_{k-1}^{(i)}] \end{aligned} \quad (9)$$

where $R_{u,k} = E \mathbf{u}_k^* \mathbf{u}_k$, $R_{du,k} = E \mathbf{d}_k \mathbf{u}_k^*$ and ϵ denotes a regularization parameter with small positive value. For an APA-based method, recursion (9) becomes,

$$\psi_k^{(i)} = \phi_k^{(i-1)} + \mu_k U_{k,i}^* (\epsilon I + U_{k,i} U_{k,i}^*)^{-1} [d_{k,i} - U_{k,i} \phi_k^{(i-1)}] \quad (10)$$

where the local $T \times M$ block data matrix and $T \times 1$ data vector are,

$$U_{k,i} = \begin{pmatrix} u_{k,i} \\ u_{k,i-1} \\ \vdots \\ u_{k,i-T+1} \end{pmatrix}, \quad d_{k,i} = \begin{pmatrix} d_k(i) \\ d_k(i-1) \\ \vdots \\ d_k(i-T+1) \end{pmatrix} \quad (11)$$

and ϵ is employed to avoid the inversion of a rank deficient matrix $U_{k,i} U_{k,i}^*$.

3. PERFORMANCE ANALYSIS

Due to space limitation, we just provide the main steps of the performance analysis in this paper. More details can be found in [9]. Firstly, we define the following global representations, in terms of the stochastic quantities,

$$\begin{aligned} \psi_c^i &= \text{col}\{\psi_1^{(i-1)}, \dots, \psi_N^{(i-1)}\}, & \phi_c^i &= \text{col}\{\phi_1^{(i-1)}, \dots, \phi_N^{(i-1)}\} \\ \mathbf{U}_c^i &= \text{diag}\{\mathbf{U}_{1,i}, \dots, \mathbf{U}_{N,i}\}, & \mathbf{d}_c^i &= \text{col}\{\mathbf{d}_{1,i}, \dots, \mathbf{d}_{N,i}\} \end{aligned}$$

where \mathbf{U}_c^i is an $NT \times NM$ block diagonal matrix. An $NM \times NM$ diagonal matrix D is defined by

$$D = \text{diag}\{\mu_1 I_M, \dots, \mu_N I_M\} \quad (12)$$

to collect the local step-sizes. To facilitate analysis, the network topology is assumed to be static (i.e. $c_{k,l}(i) = c_{k,l}$). It should be highlighted that this assumption does not compromise the algorithm derivation or its operation, and is used for analysis only. Using the above definitions, the global recursion is therefore formulated as follows:

$$\psi_c^i = G \psi_c^{i-1} + D \mathbf{U}_c^{i*} (\epsilon I_{TN} + \mathbf{U}_c^i \mathbf{U}_c^{i*})^{-1} (\mathbf{d}_c^i - \mathbf{U}_c^i G \psi_c^{i-1}) \quad (13)$$

where $G = C \otimes I_M$ is the $NM \times NM$ network topology matrix and the symmetric combining matrix C , is formed by $\{c_{k,l}\}$.

For the later reference, the global weight error vector is defined as

$$\tilde{\psi}_c^i = w_c^o - \psi_c^i. \quad (14)$$

where $w_c^o = Q w^o$ with an $NM \times M$ transition matrix $Q = \text{col}\{I_M, \dots, I_M\}$. Let Λ_v denote a $TN \times TN$ diagonal matrix, whose entries are the noise variances $\{\sigma_{v,k}^2\}$ for $k = 1, \dots, N$, and given by

$$\Lambda_v = \text{diag}\{\sigma_{v,1}^2 I_T, \dots, \sigma_{v,N}^2 I_T\}. \quad (15)$$

Using block vectorization ($\text{bvec}\{\cdot\}$) and block Kronecker product (\odot) within space-time energy conservation arguments, we obtain the mean-square behaviour of the diffusion APA algorithm:

$$\begin{aligned} E \|\tilde{\psi}_c^i\|_\sigma^2 &= E \|\tilde{\psi}_c^{i-1}\|_{\text{bvec}^{-1}\{F\sigma\}}^2 + \gamma^T \sigma \\ F &= (G^T \odot G^*) [I_{N^2 M^2} - (Z^T \odot I_{NM})(D^T \odot I_{NM}) \\ &\quad - (I_{NM} \odot Z)(I_{NM} \odot D) + E(Z^T \odot Z) \cdot (D^T \odot D)] \end{aligned} \quad (16)$$

where $Z = EZ$ with $\mathbf{Z} = \mathbf{U}_c^{i*} (\epsilon I_{TN} + \mathbf{U}_c^i \mathbf{U}_c^{i*})^{-1} \mathbf{U}_c^i$ and

$$\gamma = (D^T \odot D) \cdot E[(\mathbf{W}^T \otimes \mathbf{W}^*)] \cdot \gamma_v \quad (18)$$

with $\mathbf{W} = (\epsilon I_{TN} + \mathbf{U}_c^i \mathbf{U}_c^{i*})^{-1} \mathbf{U}_c^i$ and $\gamma_v = \text{vec}\{\Lambda_v\}$. The global transient behaviour of the adaptive network is shown in $\{\tilde{\psi}_c^{i-1}, \tilde{\psi}_c^i\}$ by expressions (16) and (17), which can be used below to study the mean-square behavior of diffusion APA.

Let $\lambda(A)$ denote all eigenvalues of a matrix A . With the purpose to show the convergence in the mean-square sense, F in (16) can be rewritten as

$$F = (G^T \odot G^*) H \quad (19)$$

where H is an Hermitian matrix, given by

$$\begin{aligned} H &= I_{N^2 M^2} - (Z^T \odot I_{NM})(D^T \odot I_{NM}) \\ &\quad - (I_{NM} \odot Z)(I_{NM} \odot D) + \Pi \cdot (D^T \odot D) \end{aligned} \quad (20)$$

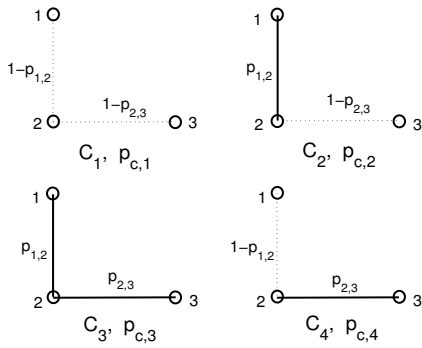


Fig. 1. The subnetworks C_j with corresponding probability $p_{c,j}$

so that F should satisfy

$$\boxed{-1 < \lambda(F) < 1} \quad (21)$$

to guarantee the mean-square stability. In other words, the spectrum of F must be strictly inside the unit disc. Using matrix 2-norms, the following result is obtained

$$|\lambda_{\max}((G^T \odot G^*)H)| \leq |\lambda_{\max}(H)| \quad (22)$$

where $\lambda_{\max}(A)$ denotes the largest eigenvalue of matrix A . Thus, the diffusion cooperation strategy leads to a stabilizing effect on the network. In addition, cooperation reduces the eigenmode of mean-square weight error evolutions as compared with its noncooperative counterpart, which will be verified by simulation results.

4. DYNAMIC NETWORK TOPOLOGY

In a dynamic network topology model (i.e. probabilistic diffusion network as in [10]) it is assumed that the nature of links between nodes is determined randomly due to link failures or time delays. At time i , when the connection between undirected nodes k and l is established with probability $p_{k,l}$, the value of $\mathbf{c}_{k,l}(i)$ is set to be equal to $c_{k,l}$, where $c_{k,l}$ denotes the link weighted coefficient as in (6). Otherwise, $\mathbf{c}_{k,l}(i)$ is zero with probability $1 - p_{k,l}$. Therefore, for undirected nodes k and l , the elements of the corresponding combining matrix C_i are formed by, for $k, l = 1, \dots, N$:

$$\mathbf{c}_{k,l}(i) = \begin{cases} c_{k,l} & \text{with } p_{k,l} \\ 0 & \text{with } 1 - p_{k,l} \end{cases} \quad (23)$$

where $c_{k,l} = c_{l,k}$ and $p_{k,l} = p_{l,k}$. In an N -node network, n_l is defined as the maximum number of links and C_j is a subnetwork matrix for $j = 1, \dots, 2^{n_l}$. A simple example is

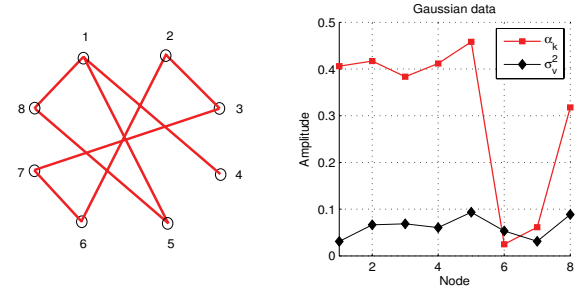


Fig. 2. Network topology (left) and statistical settings.

shown in Figure 1, which describes a 3-node network with $n_l = 2$. The probability $p_{c,j}$ of C_j depends on $\{p_{k,l}\}$, for instance, $p_{c,3} = p_{1,2}p_{2,3}$ is the probability of the subnetwork C_3 . In this manner, the mean topology matrix $G_c = E(G_i^T \odot G_i^*)$ for the dynamic network topology is given by,

$$G_c = \sum_{j=1}^{2^{n_l}} p_{c,j} (G_j^T \odot G_j^*) \quad (24)$$

which is therefore used to replace the corresponding terms in the above equation (22) to obtain the similar analysis of the mean-square stability.

5. SIMULATIONS

In this section, computer simulations are carried out in a system identification scenario, where a correlated input signal at a local node is generated as a Gaussian first-order Markov process, which allows the local covariance matrix $R_{u,k}$ to be a Toeplitz matrix with entries $r_k(m) = \sigma_{u,k}^2 \alpha_k^{|m|}$ for $m = 0, \dots, M-1$, where α_k denotes the correlation index and the variance of the local input signal is set as $\sigma_{u,k}^2 = 1$. In addition, for all APA-based schemes the regularization parameter $\epsilon = 0.001$ is chosen as a small value to reduce its effect on step-sizes. Moreover, all the coefficients of the adaptive filters within the network are initialized to zeros. An NLMS-based strategy can be regarded as a special case of an APA-based scheme with $T = 1$. All the simulated results in this work are obtained by averaging 100 Monte Carlo runs.

Figure 2 presents a network topology, where various algorithms are performed to estimate a 6×1 unknown vector w^o in an 8-node network by using $\mu_k = 0.2$. The corresponding statistical settings of Gaussian input and noise signals are plotted on the right of Figure 2. The mean-square eigenmodes of the network are shown in Figure 3. The expectation terms, involved in (17), are calculated by ensemble averaging. For the probabilistic APA algorithms, the probabilities $p = 0.1, 0.5$ are chosen. It is clear to see that cooperation also decreases the eigenmodes of the mean-square weight error evolution. Figure 4 illustrates the global transient performance of various algorithms in 200 time samples. These results confirm

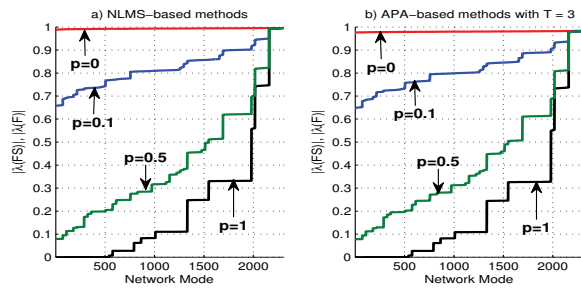


Fig. 3. The $N^2 M^2$ modes of F : a) NLMS-based schemes; b) APA-based schemes with $T = 3$. The value $p_{k,l} = p$ denotes the probability of the link between nodes k and l .

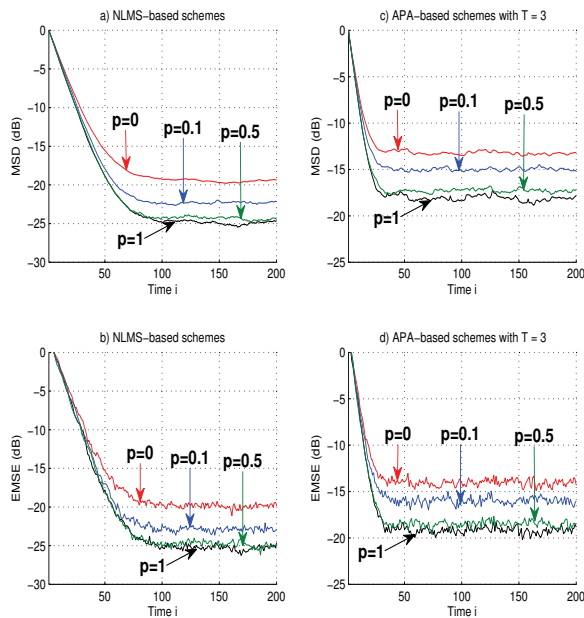


Fig. 4. Global transient performance: a) MSD for NLMS-based schemes; b) EMSE for NLMS-based schemes; c) MSD for APA-based schemes with $T = 3$; d) EMSE for APA-based schemes with $T = 3$.

that the cooperative method, even with the small probability $p = 0.1$, achieves an improved performance over the noncooperative one ($p = 0$).

6. CONCLUSIONS

This paper described a new diffusion adaptive learning algorithm based on APA for dynamic distributed networks and exploited the weighted space-time energy conservation approach of Lopes and Sayed [3] to analyze the performance

of diffusion APA. This approach yields insight into the energy flow between nodes. One main contribution of this paper is to study the mean-square stability of diffusion APA, which is consistent with simulations. Compared with the non-cooperative APA scheme, diffusion APA achieves a great improvement in terms of not only convergence rate but also steady-state performance.

7. REFERENCES

- [1] C. G. Lopes and A. H. Sayed, "Incremental adaptive strategies over distributed networks," *IEEE Trans. on Signal Processing*, vol. 55, no. 8, pp. 4064 – 4077, August 2007.
- [2] A. H. Sayed and C. G. Lopes, "Distributed recursive least-squares strategies over adaptive networks," in *Proc. Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, October-November 2006, pp. 233 – 237.
- [3] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis," *IEEE Trans. on Signal Processing*, vol. 56, no. 7, 2008.
- [4] F. Cattivelli, C. G. Lopes, and A. H. Sayed, "Diffusion recursive least-squares for distributed estimation over adaptive networks," *IEEE Trans. on Signal Processing*, vol. 56, no. 5, pp. 1865 – 1877, May 2008.
- [5] L. Li, J. A. Chambers, C. G. Lopes and A. H. Sayed, "Distributed estimation over an adaptive incremental network based on the affine projection algorithm," *IEEE Trans. on Signal Processing*, 2009, in review.
- [6] L. Li and J. A. Chambers, "A new incremental affine projection based adaptive algorithm for distributed networks," *Fast Communication in Signal Processing*, vol. 88, no. 10, pp. 2599–2603, 2008.
- [7] A. H. Sayed, *Fundamentals of Adaptive Filters*, Wiley, NJ, 2003.
- [8] L. Xiao, S. Boyd and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. 4th Int. Symp. Information Processing in Sensor Networks*, Los Angeles, USA, April 2005, pp. 63 – 70.
- [9] L. Li, *Adaptive Algorithms and Variable Structures for Distributed Estimation*, PhD Thesis, submitted to Loughborough University, UK, 2009.
- [10] C. G. Lopes and A. H. Sayed, "Diffusion adaptive networks with changing topologies," in *Proc. IEEE Int. Cof. Acoustic, Speech, Signal Processing (ICASSP)*, Las Vegas, April 2008, pp. 3285 – 3288.