DISTRIBUTED BLIND EQUALIZATION IN NETWORKED SYSTEMS

Ying Liu [*1] and Yunlong Cai [*2]

[*] Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China
[1] yingliu@zju.edu.cn, [2] ylcai@zju.edu.cn

ABSTRACT

In this paper, we study the problem of distributed blind equalization in single-input multi-output (SIMO) systems, wherein the channels of networked systems share some similarities. This corresponds to a multi-task optimization problem. To tackle this problem, an adaptive distributed generalized Sato algorithm (d-GSA) using the diffusion cooperation rule is proposed. In the proposed d-GSA, only the scalar of equalizer output is combined and transmitted among neighbors, which significantly reduces the cost of computation and communication. The performance of d-GSA is analyzed theoretically and verified by numerical simulations. Results show that the d-GSA outperforms the corresponding non-cooperative GSA.

***Index Terms*—** Blind equalization, networked systems, diffusion, generalized Sato algorithm, multi-task optimization, node-specific.

1. INTRODUCTION

Recently, with the advent of large-scale sensor networks, distributed estimation, where the sensors collaboratively estimate a certain parameter vector of interest from noisy measurements, has received much attention. It can exploit the flexible cooperative learning and information processing across a set of spatially distributed sensors with the ability of sensing, learning, adaption and communication. In such a manner, distributed estimation can properly reduce the amount of data communication over the networks, save bandwidth and energy and extend the network lifetime [1, 2].

Yet, it is noted that most of existing distributed estimation algorithms are nonblind or training-based. That is, to adapt the estimate for the unknown parameters, both the input signal (reference signal) and the desired output are known in advance by the receivers [3]-[7]. However, the use of a reference signal has some drawbacks. On the one hand, it may be physically infeasible to obtain the training signal in prior in some practical applications [8]-[10]. On the other hand, even if a reference signal is available, the use of a training signal may scarify valuable channel capacity [8, 11].

Recently, the theory of blind signal processing has been incorporated into the context of distributed in-network processing, and several distributed blind adaptive algorithms have been proposed [12]-[15]. As we are interested in the single-input multi-output (SIMO) systems in this paper, only the studies on the SIMO systems are briefly reviewed. In [12], Abdolee and Champagne have proposed a kind of distributed constant modulus algorithm (d-CMA) using the *incremental* cooperation protocol. Although this kind of d-CMA requires a relatively low communication overhead, a Hermitian cyclic path through the network is demanded. Such a cyclic trajectory is vulnerable to link and node failures. Once a sensor fails, a new cyclic path must be re-established. But such a path finding problem is NP-hard and time-consuming, especially for a large-size network. Besides, in this paper, all of the channels are assumed to be the same expect a certain phase shift, which is impractical for real cases. To tackle these problems, in [15], a recursive consensus-based distributed blind equalization algorithm has been proposed. But, to obtain the equalizer for each sensor, a high-dimensional quantity must be computed and transmitted among neighbors, which consumes much computation and data transmission.

To reduce the number of computation and data transmission, in this paper, an efficient distributed generalized Sato algorithm (d-GSA) is proposed for the networked SIMO, wherein the source signal is of constant modulars and neighboring channels share some similarities. Considering that the source signal is the same, only the equalizer output (an estimate of the source signal) instead of the tap weights of the channel equalizers, is exchanged and combined among neighbors in the proposed d-GSA, thus significantly reducing the number of data transmission in the d-CMA [12].

The rest of this paper is organized as follows. In Sec. 2, the problem of distributed blind channel equalization for SIMO systems is formulated. In Sec. 3, a new d-GSA is proposed, and its performance is analyzed in Sec. 4. In Sec. 5, some simulations are presented to show the effectiveness of the proposed algorithm. Finally, some conclusion is drawn in Sec. 6.

Notation: In this paper, we use small and capital boldface letters to denote vectors and matrices, respectively. The superscript "T" and "*" denote the transposition and the complex conjugate-transposition of a matrix or vector, respectively. The operators vec$\{\cdot\}$, diag$\{\cdot\}$, col$(\cdot)$, $E[\cdot]$ and $\otimes$ denote the standard vectorization operation, the (block) diagonal matrix, the column vector, the expectation and the Kronecker product of two matrices, respectively. The capital I denotes an identity matrix with suitable dimension.

2. PROBLEM FORMULATION

Consider a network consisting of N sensors spatially distributed over a region with a certain topological structure. Note that here the network topology is described by an undirected graph, and the edge is defined if two sensors exchange information between each other. All the sensors are interested in the common message sender $s(n)$ through its own specified FIR channel with impulse response $h_k(\cdot)$, which gives rise to an output $u_k(n)$, i.e.

$$u_k(n) = \sum_{l=0}^{L-1} h_k(l)s(n-l) + v_k(n) \tag{1}$$

where $v_k(n)$ denotes an additive measurement noise, which is i.i.d. and follows a complex circular Gaussian distribution, i.e. $v_k(n) \sim \mathcal{C}(0, \sigma_{v,k}^2)$. Note that the measurements $u_k(n)$ can be real or complex depending on the input $s(n)$ and the channel $h_k(n)$. The complex $s(n)$ is of constant modulus. In this paper, we take the 4-quadrature amplitude modulation (4-QAM) as an example.

Based on the complex measurements $\{u_k(n)\}$ of the network, our target is to design a blind equalizer $\boldsymbol{w}_k(\cdot)$ for each channel k by fusing the information from a subset of k's neighbors so as to reduce the intersymbol interference (ISI).

Besides, it is also assumed that there exists some similarities (denoted by notation '$\sim$') between the optimal channel equalizers $\boldsymbol{w}_k^o$ and its neighbor $\boldsymbol{w}_l^o$, i.e.

$$\boldsymbol{w}_k^o \sim \boldsymbol{w}_l^o, \quad l \in \mathcal{N}_k. \tag{2}$$

Note that this assumption is reasonable for the case that the common source $s(n)$ lies in a far-end field. In this case, the neighboring channels can be assumed to be similar due to the similar channel fading, which further implies that the equalizers of the neighbors share some similarities. Under this assumption, it is expected that the cooperation among sensors help improving the performance of equalization.

In addition, to perform the performance analysis, some assumptions commonly adopted in adaptive filtering [3]-[5] and blind equalization [12, 16] are assumed at first.

A-1: The channel is time-invariant and $\{s(n)\}$, $\{v_k(n)\}$, and $\{\boldsymbol{u}_k(n)\}$ are stationary and have zero mean. The input sequence $\{s(n)\}$ and the additive noise $\{v_k(n)\}$ are temporally and spatially independent identically distributed (i.i.d.) with zero mean. We also assume that $\{s(n)\}$ and $\{v_k(n)\}$ are independent of each other.

A-2: The equalizer input vector $\boldsymbol{u}_k(n)$ conditioned on the source signal $\{s(n)\}$ is a complex Gaussian random vector (RV), and it is also upper bounded by a constant.

A-3: For each node, the tap weight vector $\boldsymbol{w}_k(n)$ is independent of the equalizer input, $\boldsymbol{u}_k(n)$.

A-4: The components of the transformed tap weight vector for each node k are uncorrelated, and those of different nodes are also uncorrelated.

A-5: Assuming the noise power $\sigma_{v,k}^2$ is small enough such that the zero-forcing solution $\boldsymbol{w}^o \in \mathbb{C}^M$ is the global minimizers of the cost function.

3. DISTRIBUTED GENERALIZED SATO ALGORITHM

In this section, a distributed version of generalized Sato algorithm using the diffusion cooperation rule is proposed.

In the study of distributed in-network processing, it is essential to find a certain common quantity for data fusion among neighboring sensors. In the d-CMA proposed in [12], as all the sensors are interested in the same channel equalizer, the weights of the channel equalizers $\boldsymbol{w}_k(n)$ are combined. But, in our considered case, as the channels are different, the channel equalizers are also node-specific. So, the direct combination of equalizers $\boldsymbol{w}_k(n)$ is not applicable.

Yet, considering that we are interested in the common source signal $s(n)$, it is required that the equalizer outputs at different sensors, which are estimates for $s(n)$, should be the same or at least similar, i.e.

$$y_k(n) \sim y_l(n), \quad l \in \mathcal{N}_k,$$

such that the recovered data symbols with an aid of the slicer at different sensors can achieve consensus. So, the equalizer output $y_k(n)$ can be selected as the quantity to be combined.

Based on the above analysis and the principle of the generalized Sato algorithm [17, 18], we can seek $\boldsymbol{w}_k^o$ at each sensor k by minimizing the following cost function between the aggregated equalizer output (the estimate of the source) and the statistics of the transmitted data constellation, which is expressed as

$$J_k(\boldsymbol{w}) = E[|\gamma \text{csgn}(\tilde{y}_k(n)) - \tilde{y}_k(n)|^2], \tag{3}$$

where "csgn" denotes the complex sign function for the aggregated equalizer output, which is expressed as

$$\tilde{y}_k(n) = \sum_{l\in\mathcal{N}_k} a_{lk} y_l(n) = \sum_{l\in\mathcal{N}_k} a_{lk}\boldsymbol{u}_l(n)\boldsymbol{w}_l(n), \tag{4}$$

and the non-negative coefficients a_{lk} satisfy

$$a_{lk} = 0 \,\text{if}\, l \notin \mathcal{N}_k, \quad A\mathbf{1} = \mathbf{1}, \quad \mathbf{1}^T A = \mathbf{1}^T. \tag{5}$$

In the implementation, using the steepest-descent algorithm, we can obtain a recursion for the estimate of $\boldsymbol{w}_k$ at each iteration n

$$\boldsymbol{w}_k(n+1) = \boldsymbol{w}_k(n) - \mu_k[\nabla_{\boldsymbol{w}} J_k(\boldsymbol{w}_k(n))]^*, \tag{6}$$

where $0 < \mu_k < 1$ is a step-size parameter, and $\nabla_{\boldsymbol{w}} J_k(\boldsymbol{w}_k)$ denotes the gradient of $J_k(\boldsymbol{w}_k)$ with respect to $\boldsymbol{w}_k(n)$, which is given by

$$[\nabla_{\boldsymbol{w}} J_k(\boldsymbol{w}_k(n))]^* = -\boldsymbol{u}_k^*(n)[\gamma \text{csgn}(\tilde{y}_k(n)) - \tilde{y}_k(n)]. \tag{7}$$

Substituting (7) into (6), we have the channel equalizer adapt according to

$$\boldsymbol{w}_k(n+1) = \boldsymbol{w}_k(n) + \mu_k \boldsymbol{u}_k^*(n)[\gamma \text{csgn}(\tilde{y}_k(n)) - \tilde{y}_k(n)]. \tag{8}$$

This algorithm is named as *distributed generalized Sato algorithm*, denoted as d-GSA for short. To summarize, its implementation procedure is given in Algorithm 1.

Algorithm 1 Distributed generalized Sato algorithm (d-GSA)

Initialization: For each node k, the equalizer $\boldsymbol{w}_k(0)$ is initialized such that the center tap being one and the other taps being zero.
For each time $n \geq 1$ and each node k, repeat the following:

1. **Compute equalizer output:**
   $y_k(n) = \boldsymbol{u}_k(n)\boldsymbol{w}_k(n-1)$, $k = 1, \ldots, N$.
2. **Combination:** $\tilde{y}_k(n) = \sum_{l\in\mathcal{N}_k} a_{lk} y_l(n)$.
3. **Adaption:**
   $\boldsymbol{w}_k(n+1) = \boldsymbol{w}_k(n) + \mu_k \boldsymbol{u}_k^*(n)[\gamma \text{csgn}(\tilde{y}_k(n)) - \tilde{y}_k(n)]$.

4. PERFORMANCE ANALYSIS

In this section, the performance of d-GSA in both mean and mean-square senses is analyzed.

For the 4-QAM message sender $s(n)$, we have $|s_r(n)| = |s_i(n)| = \gamma$. Then, we can make the following approximation

$$\begin{aligned}\gamma \text{csgn}(y_l(n)) =& \gamma \text{csgn}(s(n)) + z_l(n) \\ =& s(n) + z_l(n),\end{aligned} \tag{9}$$

where $z_l(n)$ is a Gaussian random variable with zero-mean and variance $\sigma_{z,l}^2$, since $s(n)$ is uniformly symmetrically distributed with zero mean as stated in *A-1*. Besides, $z_l(n)$ is also independent of $\boldsymbol{u}_l(n)$, according to *A-2* and *A-3*.

Based on assumption *A-5*, we assume that by choosing suitable step-size μ_k, the estimate of the channel equalizer $\boldsymbol{w}_k(n)$ converges close enough to the optimum $\boldsymbol{w}_k^o$ such that the complex sign of the equalizer output before and after combination keep invariant [18, 19], i.e.

$$\text{csgn}(\tilde{y}_k(n)) = \text{csgn}(y_l(n)), \quad l \in \mathcal{N}_k. \tag{10}$$

Let us define

$$\begin{aligned}e_k(n) =& \gamma \text{csgn}(\tilde{y}_k(n)) - \tilde{y}_k(n) \\ =& \gamma \text{csgn}(\tilde{y}_k(n)) - \sum_{l\in\mathcal{N}_k} a_{lk} y_l(n).\end{aligned} \tag{11}$$

Then, by selecting a_{lk} satisfying (5) and based on (10), (11) reduces to

$$\begin{aligned} e_k(n) =&\gamma \text{csgn}(y_l(n)) - \sum_{l\in\mathcal{N}_k} a_{lk} y_l(n) \\ =&\sum_{l\in\mathcal{N}_k} a_{lk}(\gamma \text{csgn}(y_l(n)) - y_l(n)). \end{aligned} \tag{12}$$

Using the approximation (9) and based on *A-5*, $e_k(n)$ becomes

$$\begin{aligned} e_k(n) =& \sum_{l\in\mathcal{N}_k} a_{lk}(\gamma \text{csgn}(s(n)) - y_l(n)) \\ =& \sum_{l\in\mathcal{N}_k} a_{lk} \boldsymbol{u}_l(n)(\boldsymbol{w}_l^o - \boldsymbol{w}_l(n)) + \sum_{l\in\mathcal{N}_k} a_{lk} z_l(n). \end{aligned} \tag{13}$$

Let us define the instant weight error vector for each node k as

$$\widetilde{\boldsymbol{w}}_k(n) = \boldsymbol{w}_k^o - \boldsymbol{w}_k(n). \tag{14}$$

Based on (14), we can establish the following relationship

$$\begin{aligned} \widetilde{\boldsymbol{w}}_l(n) =& \boldsymbol{w}_k^o - \boldsymbol{w}_k^o + \boldsymbol{w}_l^o - \boldsymbol{w}_k(n) + \boldsymbol{w}_k(n) - \boldsymbol{w}_l(n) \\ =& \boldsymbol{w}_k^o - \widetilde{\boldsymbol{w}}_{kl}^o - \boldsymbol{w}_k(n) + \widetilde{\boldsymbol{w}}_{kl}(n), \end{aligned}$$

where the instantaneous error between $\boldsymbol{w}_k(n)$ and $\boldsymbol{w}_l(n)$ gives

$$\widetilde{\boldsymbol{w}}_{kl}(n) = \boldsymbol{w}_k(n) - \boldsymbol{w}_l(n), \tag{15}$$

and $\boldsymbol{w}_{kl}^o = \boldsymbol{w}_k^o - \boldsymbol{w}_l^o$ denotes the difference of the optimal channel equalizers between sensor k and l.

Subtracting $\boldsymbol{w}_k^o$ from (8), we have

$$\begin{aligned} \boldsymbol{w}_k^o - \boldsymbol{w}_k(n+1) =& \boldsymbol{w}_k^o - \boldsymbol{w}_k(n) - \mu_k \boldsymbol{u}_k^*(n) e_k(n) \\ =& [1 - \mu_k \boldsymbol{u}_k^*(n) \sum_{l\in\mathcal{N}_k} a_{lk}\boldsymbol{u}_l(n)](\boldsymbol{w}_k^o - \boldsymbol{w}_k(n)) \\ & - \mu_k \boldsymbol{u}_k^*(n) \sum_{l\in\mathcal{N}_k} a_{lk}\boldsymbol{u}_l(n)\widetilde{\boldsymbol{w}}_{kl}(n) \\ & + \mu_k \boldsymbol{u}_k^*(n) \sum_{l\in\mathcal{N}_k} a_{lk}\boldsymbol{u}_l(n)\widetilde{\boldsymbol{w}}_{kl}^o \\ & - \mu_k \boldsymbol{u}_k^*(n) \sum_{l\in\mathcal{N}_k} a_{lk} z_l(n). \end{aligned} \tag{16}$$

Besides, we also define the following global vectors:

$$\begin{aligned} \boldsymbol{w}_n =& \text{col}\{\boldsymbol{w}_1(n), \ldots, \boldsymbol{w}_N(n)\},\ \widetilde{\boldsymbol{w}}_n = \text{col}\{\widetilde{\boldsymbol{w}}_1(n), \ldots, \widetilde{\boldsymbol{w}}_N(n) \\ \boldsymbol{w}^o =& \text{col}\{\boldsymbol{w}_1^o, \ldots, \boldsymbol{w}_N^o\},\ \mathcal{M} = \text{diag}\{\mu_1 I_M, \ldots, \mu_N I_M\} \\ \overline{\mathcal{D}}_n =& \text{diag}\{\sum_{l\in\mathcal{N}_1} a_{l1}\boldsymbol{u}_1^*(n)\boldsymbol{u}_l(n), \ldots, \sum_{l\in\mathcal{N}_N} a_{lN}\boldsymbol{u}_N^*(n)\boldsymbol{u}_l(n)\}, \\ \overline{\mathcal{H}}_{uz}(n) =& \text{col}\{\sum_{l\in\mathcal{N}_1} a_{l1}\boldsymbol{u}_1^*(n) z_l(n), \ldots, \sum_{l\in\mathcal{N}_N} a_{lN}\boldsymbol{u}_N^* z_l(n)\}, \\ \overline{\mathcal{H}}_{\widetilde{\boldsymbol{w}}}(n) =& \text{col}\{\sum_{l\in\mathcal{N}_1} a_{l1}\boldsymbol{u}_1^*(n)\boldsymbol{u}_l(n)\widetilde{\boldsymbol{w}}_{1l}^o, \ldots, \sum_{l\in\mathcal{N}_N} a_{lN}\boldsymbol{u}_N^*(n)\boldsymbol{u}_l(n)\widetilde{\boldsymbol{w}}_{Nl}^o\}, \\ \overline{\mathcal{H}}_{\widetilde{\boldsymbol{w}}'}(n) =& \text{col}\{\sum_{l\in\mathcal{N}_1} a_{l1}\boldsymbol{u}_1^*(n)\boldsymbol{u}_l(n)\widetilde{\boldsymbol{w}}_{1l}(n), \ldots \\ & \sum_{l\in\mathcal{N}_N} a_{lN}\boldsymbol{u}_N^*(n)\boldsymbol{u}_l(n)\widetilde{\boldsymbol{w}}_{Nl}(n)\}. \end{aligned} \tag{17}$$

Based on (16) and (17), we have the global weight error

$$\widetilde{\boldsymbol{w}}_{n+1} = (I_{MN} - \mathcal{M}\overline{\mathcal{D}}_n)\widetilde{\boldsymbol{w}}_n - \overline{\mathcal{H}}_{uz}(n) + \overline{\mathcal{H}}_{\widetilde{\boldsymbol{w}}}(n) - \overline{\mathcal{H}}_{\widetilde{\boldsymbol{w}}'}(n), \tag{18}$$

where I_{MN} denotes a MN-dimensional identity matrix.

Taking the expectations of those terms in (18) and based on the assumptions *A-1-A-4*, we have

$$\begin{aligned} \overline{\mathcal{D}} =& \text{diag}\{\sum_{l\in\mathcal{N}_1} a_{l1} E[\boldsymbol{u}_1^*(n)\boldsymbol{u}_l(n)], \ldots, \sum_{l\in\mathcal{N}_N} a_{lN} E[\boldsymbol{u}_N^*(n)\boldsymbol{u}_l(n)]\}, \\ \overline{\mathcal{H}}_{uz} =& E[\overline{\mathcal{H}}_{uz}(n)] = 0, \\ \overline{\mathcal{H}}_{\widetilde{\boldsymbol{w}}} =& \text{col}\{\sum_{l\in\mathcal{N}_1} a_{l1} E[\boldsymbol{u}_1^*(n)\boldsymbol{u}_l(n)]\widetilde{\boldsymbol{w}}_{1l}^o, \ldots, \\ & \sum_{l\in\mathcal{N}_N} a_{lN} E[\boldsymbol{u}_N^*(n)\boldsymbol{u}_l(n)]\widetilde{\boldsymbol{w}}_{Nl}^o\}, \\ \overline{\mathcal{H}}_{\widetilde{\boldsymbol{w}}'} =& \text{col}\{\sum_{l\in\mathcal{N}_1} a_{l1} E[\boldsymbol{u}_1^*(n)\boldsymbol{u}_l(n)\widetilde{\boldsymbol{w}}_{1l}(n)], \ldots, \\ & \sum_{l\in\mathcal{N}_N} a_{lN} E[\boldsymbol{u}_N^*(n)\boldsymbol{u}_l(n)\widetilde{\boldsymbol{w}}_{Nl}(n)]\} \\ =& \overline{\mathcal{H}}_{\widetilde{\boldsymbol{w}}}. \end{aligned} \tag{19}$$

Based on (19), we have the expectation of global error vector

$$E[\widetilde{\boldsymbol{w}}_{n+1}] = (I_{MN} - \mathcal{M}\overline{\mathcal{D}})E[\widetilde{\boldsymbol{w}}_n]. \tag{20}$$

From (20), it is noticed that the mean stability of the estimate depends on the stability of the matrix $I_{MN} - \mathcal{M}\overline{\mathcal{D}}$. If the positive-definiteness of $\overline{\mathcal{D}}$ can be ensured, then by choosing suitable step-size μ_k, matrix $I_{MN} - \mathcal{M}\overline{\mathcal{D}}$ is stable, and thus the weight error vector $\widetilde{\boldsymbol{w}}_n$ converges to zero as $n \to \infty$. In this case, $\boldsymbol{w}_n$ can achieve an asymptotically unbiased estimate of the optimum equalizer $\boldsymbol{w}^o$.

Based on (14), we define the transient mean-square deviation (MSD) for each node k as

$$\text{MSD}_k \triangleq E||\tilde{\boldsymbol{w}}_k(n)||^2. \tag{21}$$

Following a similar analysis as that performed in the distributed LMS algorithms [3, 4], the global MSD of the network is given by

$$E||\widetilde{\boldsymbol{w}}_{n+1}||_\sigma^2 = E||\widetilde{\boldsymbol{w}}_n||_{F\sigma}^2 + \alpha^T\sigma, \tag{22}$$

and when $n \to \infty$, we have

$$E||\widetilde{\boldsymbol{w}}_\infty||^2_{(I_{M^2N^2} - F)\sigma} = \alpha^T\sigma, \tag{23}$$

where $\alpha = \text{vec}\{\mathcal{M}\overline{\mathcal{H}}_u^T\mathcal{M}\}, \overline{\mathcal{H}}_u = \mathcal{A}^T\text{diag}\{\sigma_{z,1}^2 R_{u,1}, \ldots, \sigma_{z,N}^2 R_{u,N}\}\mathcal{A}$ with $\mathcal{A} = A \otimes I_M$, and $F = \{I_{M^2N^2} - I_{MN} \otimes (\overline{\mathcal{D}}\mathcal{M}) - (\overline{\mathcal{D}}^T\mathcal{M}) \otimes I_{MN} + E[(\overline{\mathcal{D}}_n^T\mathcal{M}) \otimes (\overline{\mathcal{D}}_n\mathcal{M})]$.

Remark 1: Comparing to the d-CMA proposed in [12], it is noticed that in the d-GSA, only a scalar equalizer output $y_l(n), l \in \mathcal{N}_k$ should be transmitted at each iteration, which significantly reduces the number of data transmission. Therefore, d-GSA is more efficient than the d-CMA from the viewpoint of power-efficiency.

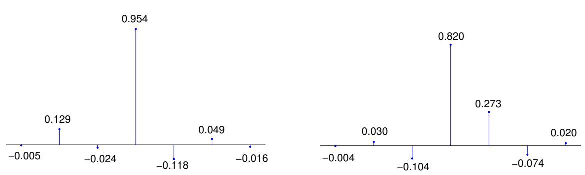
Remark 2: Referring to (20), the stability of the estimate for each node k depends on the second-order moment of the measurements $\boldsymbol{u}_k(n)$ and $\boldsymbol{u}_l(n), l \in \mathcal{N}_k$. Owing to the differences in channel models $h_k(n)$, the second-order moment $\sum_{l\in\mathcal{N}_k} a_{lk}E[\boldsymbol{u}_k^*(n)\boldsymbol{u}_l(n)]$ may not be positive-definite, and thus the positive-definiteness of the matrix $\overline{\mathcal{D}}$ cannot be ensured. To deal with this problem, the original Metropolis rule is modified as

$$\begin{cases} a_{l,k} = 0.5/\max(n_k, n_l) & \text{if } l\ (l \neq k) \text{ connects to } k, \\ a_{l,k} = 1 - \sum_{l \in \mathcal{N}_k \backslash k} a_{l,k} & \text{if } l = k, \\ a_{l,k} = 0 & \text{otherwise,} \end{cases} \tag{24}$$

where n_k and n_l are the degrees for nodes k and l, respectively. From (24), we can see that the combination matrix A is now dominated by the diagonal elements of its own estimates and thus focus more on the covariance matrix of $R_{u,k}$ to ensure the positive-definiteness of the matrix to some extent.

5. NUMERICAL SIMULATIONS

In the simulation, a sensor network consisting of 16 sensors and total 32 links is adopted. The signal-to-noise ratios (SNRs) for the sensors are randomly distributed within $(8\,, 12]$dB. The source signal $s(n)$ is generated from 4-QAM constellation. The signal $s(n)$ is then transmitted to each sensor through a node-specific channel modeled by a transversal filter of length 7. The impulse responses of the transmission channels $\boldsymbol{h}_k = [h_k(0), \ldots, h_k(L-1)]$ over the network are centered at a fixed complex vector $\boldsymbol{h}_o$, that is, $\boldsymbol{h}_k = \boldsymbol{h}_o + \widetilde{\boldsymbol{h}}_k$, where $\boldsymbol{h}_o$ is given in Fig. 1, and the perturbation $\widetilde{\boldsymbol{h}}_k$ follows a complex circular Gaussian distribution, i.e. $\widetilde{\boldsymbol{h}}_k \sim \mathcal{C}(0, \sigma^2_{h,k})$.



(a) (b)

Fig. 1. Averaged channel impulse response $\boldsymbol{h}_o$. (a) Real part. (b) Imaginary part.

In the following simulations, three cases with different values of $\sigma_{h,k}$ are considered. In Case 1, we set $\sigma_{h,k} = 0$ such that the transmission channels $\boldsymbol{h}_k$ are identical. In Case 2, we consider the case that transmission channels are similar, by setting a smaller variance $\sigma_{h,k} = 0.02$. In Case 3, a larger variance $\sigma_{h,k} = 0.16$ is used such that the transmission channels and thus the corresponding channel equalizers differ greatly from each other.

In our simulation, a 20-tap complex equalizer is adopted and initialized so that the center tap is set to one and the other taps are zero. The step-size is set as $\mu_k = 1 \times 10^{-3}$ for each node in the above three cases. Note that for the purpose of performance comparison, the result of the non-cooperative GSA, i.e. nc-GSA, is also depicted.

Since the exact value of $\boldsymbol{w}^o$ is unavailable in practice, we use the measure of residual ISI instead of the MSD to evaluate the performance of the algorithms in blind equalization. Fig. 2 depicts the averaged ISI of the network over 20 independent simulations. It is obvious that the d-GSA shows better performance than the nc-GSA for all the cases, which indicates that the cooperation among neighboring sensors can improve the performance of equalization. Besides, it is noticed that the performance enhancement of the d-GSA reduces, as the differences between channels increase, see the results from Case 1 to Case 3. This is reasonable as the extent of similarities affects the performance of distributed equalization. To be specific, in Case 1, as all the sensors are interested in the same channel equalizer (single task), the performance enhancement by collaborating all the sensors is significant. On the contrary, for Case 3, as there exist relatively large differences in channels between sensors, that is, the tasks differ a lot from each other, the effect of data combination reduces. The Case 2 lies somehow between Case 1 and Case 3.

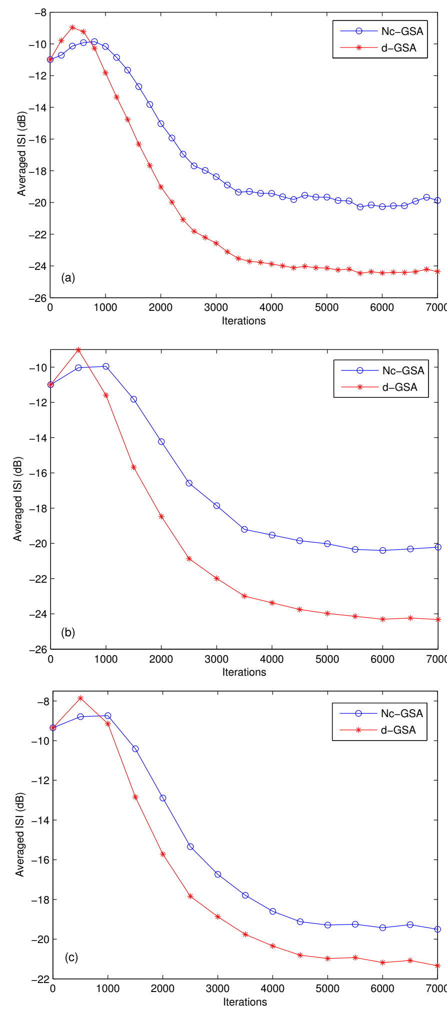


Fig. 2. Simulation results. (a) Averaged transient ISI vs iteration for Case 1. (b) Averaged transient ISI vs iteration for Case 2. (c) Averaged transient ISI vs iteration for Case 3.

6. CONCLUSION

In this paper, we have developed a diffusion generalized Sato algorithm (d-GSA), which seeks the optimal channel equalizer by minimizing the cost function between an aggregated equalizer output and the statistics of the transmitted data constellation. The convergence of the d-GSA has been analyzed and its performance has been verified by numerical simulations. Simulation results have shown that the proposed algorithm shows good equalization performance for SIMO systems, comparing to the non-cooperative GSA.

Acknowledgements
This work was supported by the National Natural Science Foundation of China (Grant Nos. 61471320 and 61571392).

7. REFERENCES

[1] A.G. Dimakis, S. Kar, J.M.F Moura, et al., "Gossip algorithms for distributed signal processing," *Proc. IEEE*, vol. 98, no. 11, pp. 1847-1864, Nov. 2010.

[2] A.H. Sayed, S.Y. Tu, J. Chen, et al., "Diffusion strategies for adaptation and learning over networks," *IEEE Signal Proc. Mag.*, vol. 30, no. 3, pp. 155-171, Mar. 2013.

[3] C.G. Lopes and A.H. Sayed, "Diffusion least-mean squares over adaptive networks: formulation and performance analysis," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3122-3136, Jul. 2008.

[4] 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.

[5] Z. Liu, Y. Liu, and C. Li, "Distributed sparse recursive least-squares over networks," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1386-1395, Feb. 2014.

[6] N. Takahashi, I. Yamada, and A.H. Sayed, "Diffusion least-mean squares with adaptive combiners: formulation and performance analysis," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4795-4810, Sep. 2010.

[7] J. Chen and A.H. Sayed, "Diffusion adaptation strategies for distributed optimization and learning over networks," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4289-4305, Aug. 2012.

[8] K. Abed-Meraim, W. Qui, and Y. Hua, "Blind system identification," *Proc. IEEE*, vol. 85, no. 8, pp. 1310-1322, Aug. 1997.

[9] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.*, vol. 13, pp. 43-64, May 1996.

[10] L.C. Wood and S. Treitel, "Seismic signal processing," *Proc. IEEE*, vol. 63, no. 4, pp. 649-661, Apr. 1975.

[11] C.R. Johnson, P. Schniter, T.J. Endres, et al. "Blind equalization using the constant modulus criterion: a review," *Proc. IEEE*, vol. 86, no. 10, pp. 1927-1950, Oct. 1998.

[12] R. Abdolee and B. Champagne, "Distributed blind adaptive algorithms based on constant modulus for wireless sensor networks," *2010 Sixth International Conference on Wireless and Mobile Communications* 978-0-7695-4182-2/10, 2010.

[13] Z. Ding and Y. Li, *Blind Equalization and Identification*, Marcel Dekker, Inc. 2001.

[14] C. Yu, L. Xie and Y. Soh, "Blind channel and source estimation in networked systems," *IEEE Trans. Signal Process.*, vol. 62, no. 17, pp. 4611-4626, Sep. 2014.

[15] C. Yu and L. Xie, "On recursive blind equalization in sensor networks," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 662-672, Feb. 2015.

[16] V. Weerackody, S.A. Kassam, and K.R. Laker, "Convergence analysis of an algorithm for blind equlization", *IEEE Trans. Commun.*, vol. 39, no. 6, pp. 856-864, Jun. 1991.

[17] Y. Sato, "A method of self-recovering equalization for mulitilevel amplitude-modulation systems," *IEEE Trans. Commun.*, vol. COM-23, pp. 679-682, Jun. 1975.

[18] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Trans. Commun.*, vol. COM-32, pp. 871-883, Aug. 1984.

[19] M.G. Larimore and J.R. Treichler, "Convergence behavior of the constant modulus algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Boston, Apr. 1983, vol. 8, pp. 13-16.