IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 5, 2017, accepted June 28, 2017, date of publication August 14, 2017, date of current version September 27, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2723051

Distributed Blind Estimation Over Sensor Networks

YING LIU[1] AND LILI LI[2]

[1]College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China
[2]School of Naval Architecture and Ocean Engineering, Zhejiang Ocean University, Zhoushan 316000, China

Corresponding author: Ying Liu (yingliu@zju.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61471320, Grant 61571392, and Grant 61631003, in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LY17F010009 and Grant LQ15F030007, and in part by the Fundamental Research Funds for the Central Universities under Grant 2017QNA5009.

ABSTRACT Distributed estimation, where a set of nodes collaboratively estimate some parameters of interest from noisy measurements, has received much attention in both science and engineering. Recent studies mainly focus on distributed non-blind or training-based estimation, that is, a training signal and its desired output are both known to the distributed receivers. However, in some applications, it is physically difficult, if not impossible, to get a training signal in prior. Besides, the use of training signals consumes much channel bandwidth. So, it is more preferable for the receivers to perform distributed estimation without the assistance and expense of training sequences, i.e., *distributed blind estimation*. In this paper, the problem of distributed blind estimation over sensor networks is considered, and a kind of distributed diffusion generalized Sato algorithm is proposed to design a blind equalizer for channel equalization and source signal estimation. The stability of the proposed method in mean and mean-square senses is analyzed theoretically, and its performance is verified numerically by a series of simulations.

INDEX TERMS Blind estimation, distributed network, diffusion, blind equalization, generalized Sato algorithm, sensor networks.

I. INTRODUCTION

Distributed in-network processing is a major tenet for wireless and sensor networks, where a group of interconnected nodes cooperatively perform a predefined task, for example, to estimate a certain parameter vector of interest from noisy measurements. It can exploit the flexible cooperative learning and information processing across a set of spatially distributed sensors with the ability of sensing, learning, and communication. In this manner, distributed processing reduces the amount of data communication over the networks, thus saving bandwidth and energy, extending the network lifetime and reducing latency. Moreover, as each node shares with some computation and information transmission, distributed processing avoids the fragility to the dysfunction of the center node that exists in the classical centralized processing. Owing to these merits, distributed processing has been considered to be an effective approach for in-network data fusion and processing, and been widely used in many applications, such as precision agriculture and environmental monitoring, military surveillance and distributed cooperative sensing in cognitive radio networks [1], [2].

Based on the conceptual structure of distributed in-network processing, the study of distributed adaptive algorithms for parameter estimation, i.e. *distributed estimation*, has aroused much attention. In the past few years, many distributed adaptive schemes have been proposed, such as incremental LMS [3]–[5], incremental RLS [3], incremental affine projection algorithm [6], diffusion LMS [3], [7]–[10], diffusion RLS [11]–[12], distributed information theoretical learning [13], [14], and distributed consensus strategies [15].

Recent studies have mainly focused on the problems of distributed estimation that both the training signal (reference signal) and the desired output are known in advance by the receivers so as to adapt the estimate for the unknown parameters [3], [7], [8], [11], [16], [17]. When used in this way, these algorithms are classified as *non-blind*, or training-based methods. 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 [18]–[20]. On the other hand, even if a reference signal is available, the use of a training signal may scarify valuable channel capacity [18], [21]–[24].

Therefore, developing distributed adaptive algorithms for signal estimation without training sequences, i.e *distributed blind estimation*, is somehow indispensable and thus more preferable in practice.

In the literature, many kinds of blind adaptive algorithms have been proposed for equalizing the channel and furthermore for estimating the transmitted source signal [21], [25]–[31]. Considering that many signals in digital communication, such as Pulse Amplitude Modulation (PAM), Quadrature Phase Shift Keying (QPSK), and Quadrature Amplitude Modulation (QAM), are of some constant modulus properties, a very important kind of technique that has been proposed so far is based on the minimization of a nonlinear error function between the output of the equalizer and the statistics of the transmitted data constellation using a stochastic gradient approach. This kind of method includes the well-known Sato's algorithm [25], the generalized Sato algorithm (GSA) [26], the constant modulus algorithm (CMA) [28] and their modifications [27], [29]–[31]. It is also worth pointing out that most of these blind equalization algorithms are performed based on the measurements of a single receiver but not a network with multiple receivers.

Compared with traditional training-based adaptive strategies, it is theoretically challenging to analyze the performance of blind adaptive algorithms as a nonlinear cost function is minimized in adaption. When it comes to the case of a sensor network, combined with the effect of distributed cooperation among sensors, the performance analysis becomes more difficult. In [32], a distributed constant modulus algorithm (d-CMA) using the *incremental* cooperation protocol has been proposed, in which a Hermitian cyclic path through the network is required. But, such a cyclic trajectory is vulnerable. Once a sensor fails, a new cyclic path must be re-established, which is NP-hard especially when a large-size network is considered. In [33], a recursive consensus-based distributed blind equalization algorithm has been proposed. But, to obtain the equalizer for each sensor, a high-dimensional vector must be computed and transmitted among neighbors, which consumes much computation and data transmission.

In this paper, the problem of distributed blind estimation over sensor networks is considered. To release the restriction on the network topology in incremental cooperation [32], a *diffusion* cooperation rule, where each sensor cooperates with its direct neighbors, is used. Besides, to reduce the number of computation and transmission, only the tap coefficients of the channel equalizers among the neighbors are transmitted. The main contribution of this paper is summarized as follows.

1) Analogous to the training-based distributed LMS (d-LMS), it is to develop some distributed forms of blind adaptive algorithms over sensor networks to estimate the transmitted data symbols from a common source in different transmission environments. To perform the task of estimation, some distributed blind equalizers are designed for equalizing the channel and estimating the source signal using GSA method. But instead of minimizing the mean-squared error in the d-LMS, the distributed GSAs adapt the complex-valued tap coefficients of the channel equalizer by quantifying a nonlinear error function related to the complex equalizer outputs and the statistics of the transmitted data constellation.
2) It is to provide a theoretical framework to analyze the stability and convergence of the proposed distributed blind adaptive algorithms in both mean and mean-square senses, which is believed to be more challenging in blind signal processing.

This paper is organized as follows. In Sec. II, the system model is described and the problem of distributed in-network blind estimation through the design of blind equalizer is formulated. In Sec. III, a kind of distributed GSA using the diffusion cooperation is derived. The performance of the proposed algorithm in mean and mean-square senses is analyzed theoretically in Sec. IV, and verified numerically by simulations in Sec. V. Finally, this paper is concluded in Sec. VI.

Notation: In this paper, we use small boldface letters to denote vectors. Boldface capital letters denote matrices. The superscript "T" and "$*$" denote the transposition of a matrix or vector, and the complex conjugate-transposition, respectively. The notation $||x||$ denotes the Euclidean norm of a vector. The operators $\otimes$, vec$\{\cdot\}$, diag$\{\cdot\}$ and $E[\cdot]$ denote the Kronecker product of two matrices, the standard vectorization operation, the (block) diagonal matrix, and expectation, respectively. I_n denotes an n by n identity matrix. Other notation will be introduced if necessary.

II. 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 a FIR channel with impulse response $\{h(n)\}$, see Fig. 1. With reference to Fig. 1, the output $u_k(n)$ collected at each sensor k, the input data $s(n)$ and the FIR channel $\{h(n)\}$ are related by

$$\begin{aligned} u_k(n) &= \sum_{l=0}^{L-1} h(l)s(n-l) + v_k(n) \\ &= h(n) \circledast s(n) + v_k(n), \end{aligned} \tag{1}$$

where notation '$\circledast$' denotes the convolution operation, $v_k(n)$ 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 $h(n)$. Here, complex-valued $s(n)$ and $h(n)$ are considered, and $s(n)$ is of some constant modulus properties. In this paper, we focus on the QAM constellation, which is a typical class of

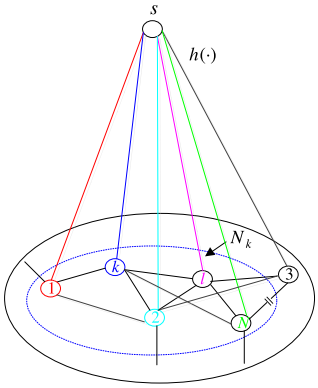


FIGURE 1. Network and system settings. 's' denotes the far-end common source, $h(\cdot)$ denotes the transmission channel, $\mathcal{N}_k$ denotes the set of the neighbors for node k.

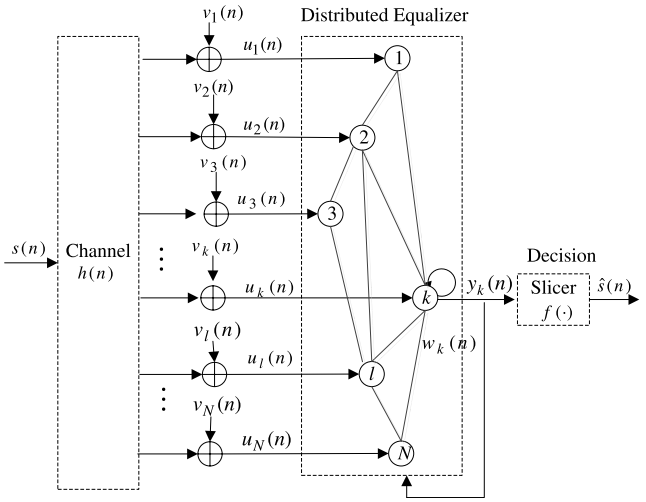


FIGURE 2. System model for distributed estimation.

modulation signal with constant envelope. We consider the case that the common source $s(n)$ lies in a far-end field. That is, the distance between the source and the network of sensors is much larger than those among the sensors. Under this situation, it is reasonable to assume that the filter $h(n)$ applied to the source is approximately the same at each sensor.

For QAM data sequences, it is obvious that we can estimate them based on the information of a single sensor instead of a sensor network [25]–[28]. The use of one receiver is of course low-cost from the aspect of implementation. But, it may require a long time series of measurements to achieve the desired performance. In practice, for efficiency reason, the common message $s(n)$ broadcast from the source (an access point) may transmit only limited information. For example, the source might be an unmanned aerial vehicle (UAV) with a fast speed flying over a wireless sensor network (WSN) deployed on the ground, and the broadcasting of information lasts for only a very short period of time [34], [35]. In this case, to achieve an acceptable estimation performance, a fast convergence rate is required. Moreover, since sensors are usually bared in harsh environments, they can be easily damaged. In some scenarios, it is impossible to request retransmission when errors are detected, and thus the loss of information cannot be recovered if only one receiver is used. Although we can replace the sensors to avoid this to some extent, it is difficult and costly to perform the manual maintenance in some dangerous environments, for instance, the sensors are located on the top of the mountains or on the deep of the oceans.

Alternatively, the adoption of a distributed sensor network is a good candidate, which has been widely used in wireless cooperative communications [34], [35]. Firstly, it is more robust to the failure of a single sensor. Secondly, it can achieve fast convergence with a very short length of measurements by exploiting the spatial diversity existing in distributed sensors. Thirdly, as the sensors are low-cost devices, the use of a sensor network does not increase much cost.

Considering this, in this paper, a distributed sensor network is adopted to collaboratively estimate the source signal $s(n)$ only based on the measurements $\{u_k(n)\}$, without knowing the channel $h(n)$. Referring to the theory of blind signal processing, we would like to recover $s(n)$ by designing a distributed blind equalizer to equalize the channel, reduce the intersymbol interference (ISI) and furthermore make an optimal estimate of $s(n)$. The schematic structure of the proposed algorithm is given in Fig. 2. As shown in Fig. 2, each sensor k designs a distributed channel equalizer $\boldsymbol{w}_k(\cdot)$ by fusing the information from a subset of k's neighbors and make the equalizer output $y_k(n)$ become a suitable estimate of $s(n)$. Then, following by a suitable slicer $f(\cdot)$, it is to give an optimal estimate for the common source $s(n)$, denoted as $\hat{s}(n)$. It is also noted that here, the transmission channels used in inter-sensor communication are assumed to be perfect, i.e, without noise and distortion.

III. DIFFUSION GENERALIZED SATO ALGORITHM

In this section, we would like to develop a distributed blind adaptive algorithm to design a blind equalizer and estimate the source signal based on the collections of a sensor network. The generalized Sato's algorithm, originally proposed in [26], is used to design a blind equalizer, but implemented collaboratively with a *diffusion* cooperation strategy. That is, each sensor k in the network is allowed to access the information of its neighbors $l \in \mathcal{N}_k$ ($\mathcal{N}_k$ denotes the neighbors of node k, i.e. those nodes directly connected to k, including itself) to design a distributed blind equalizer and furthermore to estimate the source signal.

As stated in Sec. II, since the transmission channels are identical at different sensors, all the sensors correspond to the same optimal channel equalizer $\boldsymbol{w}^o$, and thus the optimization for seeking $\boldsymbol{w}^o$ is a *single-task* optimization problem. In the following, based on the principle of d-LMS for non-blind single-task optimization [8], [9], we derive a distributed blind adaptive GSA using the diffusion cooperation rule.

Referring to the d-LMS [8], [9], two parts of cost are considered. In the first part, we try to approximate the global

cost at each sensor k by linearly combining the weighted local cost of node k's neighbors $l \in \mathcal{N}_k$. In the second part, we incorporate a norm constraint on the differences of the equalizers between sensor k and its neighbors l so as to drive the tap weights of the equalizers converge to each other. To summarize, each node k can proceed to minimize the following cost function

$$J_k(\boldsymbol{w}) = \sum_{l\in\mathcal{N}_k} c_{lk} E[|\gamma \mathrm{csgn}(y_l(n)) - y_l(n)|^2] + \sum_{l\in\mathcal{N}_k\backslash k} v_{lk}||\boldsymbol{w} - \boldsymbol{\varphi}_l||^2, \quad (2)$$

where the equalizer output

$$y_l(n) = \boldsymbol{u}_l(n)\boldsymbol{w}, \quad (3)$$

with $\boldsymbol{u}_l(n) = [u_l(n), \ldots, u_l(n-M+1)]$ being the equalizer input vector at sensor l and $\boldsymbol{w} = [w_0, \ldots, w_{M-1}]^T$ being the vector of complex tap coefficients. The function "csgn" denotes the complex sign function for a complex data, which is defined as

$$\mathrm{csgn}(x_r + jx_i) = \mathrm{sgn}(x_r) + j\mathrm{sgn}(x_i), \quad (4)$$

where x_r and x_i are the real and the imaginary parts of a complex data x, respectively. The parameter γ is a positive constant depending on the input data symbol constellation, which is defined as

$$\gamma = \frac{E[s_r^2(n)]}{E[|s_r(n)|]} = \frac{E[s_i^2(n)]}{E[|s_i(n)|]}, \quad (5)$$

where $s_r(n)$ and $s_i(n)$ are the real and imaginary parts of the complex data symbol $s(n)$, respectively.

Note that the notation $l \in \mathcal{N}_k \setminus k$ in (2) denotes the neighbors for node k except itself, $\boldsymbol{\varphi}_l$ denotes the intermediate estimate for $\boldsymbol{w}$ at sensor l, and the non-negative real coefficients c_{lk} and v_{lk} denote the coupling from the neighbors of sensor k, which satisfy

$$V^T\mathbf{1}_N = \mathbf{1}_N, \quad C\mathbf{1}_N = \mathbf{1}_N, \quad v_{lk} = 0, \quad c_{lk} = 0, \quad \text{if } l \notin \mathcal{N}_k, \quad (6)$$

where $\mathbf{1}_N$ denotes an $N \times 1$ vector with unit entries, C and V are $N \times N$ matrices with entities c_{lk} and v_{lk}, respectively.

Using an iterative steepest-descent algorithm to minimize the cost function (2), we get

$$\boldsymbol{w}_k(n) = \boldsymbol{w}_k(n-1) - \mu_k[\nabla_{\boldsymbol{w}} J_k(\boldsymbol{w}_k(n-1))]^*, \quad (7)$$

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

$$[\nabla_{\boldsymbol{w}} J_k(\boldsymbol{w}_k(n-1))]^* = -\sum_{l\in\mathcal{N}_k} c_{lk}\boldsymbol{u}_l^*(n)[\gamma \mathrm{csgn}(y_l(n)) - y_l(n)] - \sum_{l\in\mathcal{N}_k\backslash k} v_{lk}(\boldsymbol{\varphi}_l - \boldsymbol{w}_k(n-1)). \quad (8)$$

Substituting (8) into (7), we can obtain a recursion for the estimate of $\boldsymbol{w}$ at node k and iteration n,

$$\boldsymbol{w}_k(n) = \boldsymbol{w}_k(n-1) + \mu_k \sum_{l\in\mathcal{N}_k} c_{lk}\boldsymbol{u}_l^*(n)[\gamma \mathrm{csgn}(y_l(n)) - y_l(n)] + \mu_k \sum_{l\in\mathcal{N}_k\backslash k} v_{lk}(\boldsymbol{\varphi}_l - \boldsymbol{w}_k(n-1)). \quad (9)$$

Similar to d-LMS [8], we accomplish the update in two steps by generating an intermediate estimate $\boldsymbol{\varphi}_k(n)$ as follows

$$\begin{cases} \boldsymbol{\varphi}_k(n) = \boldsymbol{w}_k(n-1) + \mu_k \sum_{l\in\mathcal{N}_k} c_{lk}\boldsymbol{u}_l^*(n)[\gamma \mathrm{csgn}(y_l(n)) - y_l(n)] \\ \boldsymbol{w}_k(n) = \boldsymbol{\varphi}_k(n) + \mu_k \sum_{l\in\mathcal{N}_k\backslash k} v_{lk}(\boldsymbol{\varphi}_l - \boldsymbol{w}_k(n-1)), \end{cases} \quad (10)$$

where

$$y_l(n) = \boldsymbol{u}_l(n)\boldsymbol{w}_k(n-1). \quad (11)$$

Then, by replacing $\boldsymbol{\varphi}_l$ and $\boldsymbol{w}_k(n-1)$ in (10) by the intermediate estimate $\boldsymbol{\varphi}_l(n)$ and $\boldsymbol{\varphi}_k(n)$, respectively, the second equation of (10) can be further rewritten into

$$\boldsymbol{w}_k(n) = (1 - \mu_k + \mu_k v_{kk})\boldsymbol{\varphi}_k(n) + \mu_k \sum_{l\in\mathcal{N}_k\backslash k} v_{lk}\boldsymbol{\varphi}_l(n). \quad (12)$$

Define $a_{kk} = 1 - \mu_k + \mu_k v_{kk}$, $a_{lk} = \mu_k v_{lk}$ for $l \neq k$. Equation (12) becomes

$$\boldsymbol{w}_k(n) = \sum\nolimits_{l\in\mathcal{N}_k} a_{lk}\boldsymbol{\varphi}_l(n). \quad (13)$$

If the tap weights $\boldsymbol{w}_k(n)$ converges to the optimal channel equalizer $\boldsymbol{w}^o$, the equalizer output $y_k(n)$ converges to the region that corresponds to the true $s(n)$ with an aid of a suitable slicer $f(\cdot)$ such that $s(n)$ can be recovered more reliably. Note that the selection of slicer $f(\cdot)$ depends on the constellation of the transmitted data symbols [25]–[27]. For example, for the 4-QAM, the slicer can be selected as $f = \mathrm{csgn}(\cdot)$.

The implementation of distributed blind estimation algorithm is summarized in Algorithm 1. Note that from (10) and (13), each sensor first adapts its estimate by exchanging the information of its neighbors and combine fuses the estimates of its neighbors linearly as its own local estimate. We name such an adapt-then-combine (ATC) diffusion GSA as ATC-GSA. Reversing the order of adaptation and combination, we can obtain combine-then-adapt (CTA) diffusion GSA, denoted as CTA-GSA. Its implementation is given in Algorithm 2.

Remark 1: As shown in Algorithms 1 and 2, ATC diffusion GSA and CTA diffusion GSA for distributed blind estimation show similar structure as the ATC diffusion LMS and CTA diffusion LMS for distributed non-blind estimation, respectively [8]. However, due to inaccessibility to the channel information, the instantaneous error is now a nonlinear function of the equalizer output and the statistics of the transmitted data constellation instead of a linear function of the desired output and the estimated output in distributed non-blind LMS.

Algorithm 1 Adapt-Then-Combine Diffusion GSA (ATC-GSA)

1) **Initialization:** The channel equalizer $\boldsymbol{w}_k(0), k = 1, \ldots N$ is initialized such that the center tap equals to one and the other taps equal to zero. For each time instant $n \geq 1$ and each node k, repeat the following:
2) **Blind channel equalization:**
   a) **Adaption:**

$$\boldsymbol{\varphi}_k(n) = \boldsymbol{w}_k(n-1) + \mu_k \sum_{l \in \mathcal{N}_k} c_{lk} \boldsymbol{u}_l^*(n) \cdot [\gamma \mathrm{csgn}(y_l(n)) - y_l(n)],$$

   where

$$y_l(n) = \boldsymbol{u}_l(n) \boldsymbol{w}_k(n-1).$$

   b) **Combination:**

$$\boldsymbol{w}_k(n) = \sum_{l \in \mathcal{N}_k} a_{lk} \boldsymbol{\varphi}_l(n).$$

3) **Source estimation:** Based on $\boldsymbol{w}_k(n)$ at any sensor after convergence, an estimate of the transmitted data symbol $s(n)$ can be obtained using a suitable slicer $f(\cdot)$. For example, using the final estimate of sensor k after N_m iteration, $s(n)$ can be estimated as

$$\hat{s}(n) = f(\boldsymbol{u}_k(n) \boldsymbol{w}_k(N_m)).$$

Algorithm 2 Combine-Then-Adapt Diffusion GSA (CTA-GSA)

1) **Initialization:** The channel equalizer $\boldsymbol{w}_k(0), k = 1, \ldots N$ is initialized such that the center tap equals to one and the other taps equal to zero. For each time $n \geq 1$ and each node k, repeat the following:
2) **Blind channel equalization:**
   a) **Combination:**

$$\boldsymbol{\varphi}_k(n-1) = \sum_{l \in \mathcal{N}_k} a_{lk} \boldsymbol{w}_l(n-1).$$

   b) **Adaption:**

$$\boldsymbol{w}_k(n) = \boldsymbol{\varphi}_k(n-1) + \mu_k \sum_{l \in \mathcal{N}_k} c_{lk} \boldsymbol{u}_l^*(n) \cdot [\gamma \mathrm{csgn}(y_l(n)) - y_l(n)],$$

   where

$$y_l(n) = \boldsymbol{u}_l(n) \boldsymbol{\varphi}_k(n-1).$$

3) **Source estimation:** Based on $\boldsymbol{w}_k(n)$ at any sensor after convergence, an estimate of the transmitted data symbol $s(n)$ can be obtained using a suitable slicer $f(\cdot)$. For example, using the final estimate of sensor k after N_m iteration, $s(n)$ can be estimated as

$$\hat{s}(n) = f(\boldsymbol{u}_k(n) \boldsymbol{w}_k(N_m)).$$

IV. PERFORMANCE ANALYSIS

In this section, the performance of the proposed diffusion GSA in terms of mean and mean-square stability, and mean-square deviation (MSD) is analyzed. In the following, instead of analyzing a specified algorithm (ATC-GSA or CTA-GSA), we give a systematic analysis of the proposed diffusion GSA by resorting to a generalized structure, which include these two algorithms. Based on Algorithms 1 and 2, for each sensor k, the design of distributed blind equalizer can be rewritten into the following generalized form

$$\begin{cases} \boldsymbol{\phi}_k(n-1) = \sum_{l \in \mathcal{N}_k} q_{1,lk} \boldsymbol{w}_l(n-1), \\ \boldsymbol{\psi}_k(n) = \boldsymbol{\phi}_k(n-1) + \mu_k \sum_{l \in \mathcal{N}_k} b_{lk} \boldsymbol{u}_l^*(n) \cdot [\gamma \mathrm{csgn}(y_l(n)) - y_l(n)], \\ \boldsymbol{w}_k(n) = \sum_{l \in \mathcal{N}_k} q_{2,lk} \boldsymbol{\psi}_l(n), \end{cases} \tag{14}$$

and the non-negative real coefficients $q_{1,lk}$, $q_{2,lk}$ and b_{lk} correspond to the l and k entities of matrices Q_1, Q_2 and B, respectively, which satisfy

$$\begin{aligned} &Q_1^T \mathbf{1}_N = \mathbf{1}_N, \quad Q_2^T \mathbf{1}_N = \mathbf{1}_N, \quad B\mathbf{1}_N = \mathbf{1}_N, \\ &q_{1,lk} = 0, \quad q_{2,lk} = 0, \quad b_{lk} = 0, \quad \text{if } l \notin \mathcal{N}_k. \end{aligned} \tag{15}$$

Several choices for selecting cooperation weights from graph theory have been suggested in the literature, such as the Metropolis rule [7], [8], the relative degree [8], and the Laplacian matrix [7], [8]. Since the Metropolis rule has shown to be better than the others, it is adopted in this paper. Using different settings of matrices, the general structure of diffusion GSA can be specialized into different algorithms, which is summarized in Table 1.

TABLE 1. Different GSAs with different choices of matrices.

	Q_1	Q_2	B
ATC-GSA	I_N	A	C
CTA-GSA	A	I_N	C
Global GSA	I_N	I_N	$\mathbf{1}\mathbf{1}^T/N$
No-cooperation GSA	I_N	I_N	I_N

To generalize the performance analysis, we define the following global quantities:

$$\begin{aligned} \boldsymbol{w}^{(o)} &= \mathrm{col}\{\boldsymbol{w}^o, \ldots, \boldsymbol{w}^o\}, \quad \boldsymbol{\phi}_n = \mathrm{col}\{\boldsymbol{\phi}_1(n), \ldots, \boldsymbol{\phi}_N(n)\}, \\ \boldsymbol{\psi}_n &= \mathrm{col}\{\boldsymbol{\psi}_1(n), \ldots, \boldsymbol{\psi}_N(n)\}, \\ \boldsymbol{w}_n &= \mathrm{col}\{\boldsymbol{w}_1(n), \ldots, \boldsymbol{w}_N(n)\}, \\ \mathcal{M} &= \mathrm{diag}\{\mu_1 I_M, \ldots, \mu_N I_M\}, \\ U_n &= \mathrm{col}\{\boldsymbol{u}_l(n), \ldots, \boldsymbol{u}_N(n)\}, \\ \mathcal{D}_n &= \mathrm{diag}\{\sum_{l \in \mathcal{N}_1} b_{l1} \boldsymbol{u}_l^*(n) \boldsymbol{u}_l(n), \ldots, \sum_{l \in \mathcal{N}_N} b_{lN} \boldsymbol{u}_l^*(n) \boldsymbol{u}_l(n)\}, \\ \mathcal{G}_n &= \mathrm{col}\{\sum_{l \in \mathcal{N}_1} b_{l1} \boldsymbol{u}_l^*(n) \mathrm{csgn}(y_l(n)), \ldots, \\ &\qquad \sum_{l \in \mathcal{N}_N} b_{lN} \boldsymbol{u}_l^*(n) \mathrm{csgn}(y_l(n))\}, \end{aligned} \tag{16}$$

and the extended weighting matrices

$$\mathcal{Q}_1 = Q_1 \otimes I_M, \quad \mathcal{Q}_2 = Q_2 \otimes I_M, \quad \mathcal{B} = B \otimes I_M. \tag{17}$$

Based on the above global quantities (16), the design of blind equalizer for the whole network using the diffusion GSA can be rewritten into the following global form

$$\begin{cases} \boldsymbol{\phi}_{n-1} = \mathcal{Q}_1^T \boldsymbol{w}_{n-1}, \\ \boldsymbol{\psi}_n = \boldsymbol{\phi}_{n-1} - \mathcal{M}\mathcal{D}_n \boldsymbol{\phi}_{n-1} + \gamma \mathcal{M}\mathcal{G}_n, \\ \boldsymbol{w}_n = \mathcal{Q}_2^T \boldsymbol{\psi}_n. \end{cases} \tag{18}$$

By cancelling the intermediate variables $\boldsymbol{\phi}_{n-1}$ and $\boldsymbol{\psi}_n$ in (18), we have

$$\boldsymbol{w}_n = \mathcal{Q}_2^T \mathcal{Q}_1^T \boldsymbol{w}_{n-1} - \mathcal{Q}_2^T \mathcal{M}\mathcal{D}_n \mathcal{Q}_1^T \boldsymbol{w}_{n-1} + \gamma \mathcal{Q}_2^T \mathcal{M}\mathcal{G}_n. \tag{19}$$

Define the weight error vector

$$\widetilde{\boldsymbol{w}}_n = \boldsymbol{w}^{(o)} - \boldsymbol{w}_n. \tag{20}$$

Let $\boldsymbol{w}^{(o)}$ subtract both sides of (19). We have the mean weight error $\widetilde{\boldsymbol{w}}_n$ evolve as

$$\widetilde{\boldsymbol{w}}_n = \mathcal{Q}_2^T (I_{MN} - \mathcal{M}\mathcal{D}_n)\mathcal{Q}_1^T \widetilde{\boldsymbol{w}}_{n-1} + \mathcal{Q}_2^T \mathcal{M}(\mathcal{D}_n \boldsymbol{w}^{(o)} - \gamma \mathcal{G}_n). \tag{21}$$

To perform the following analysis, some assumptions commonly adopted in adaptive filtering [3], [7]–[10] and blind equalization are assumed at first [36]–[39].

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, $\boldsymbol{u}_k(n)$. Besides, the central tap weight of $\boldsymbol{w}_k(n)$ is normalized to 1 such that all the nodes converge to the same channel equalizer without phase shift.

A-4: The components of the transformed tap weight vector for each node k, $\boldsymbol{w}_k^j(n)$ are uncorrelated, and also those of different nodes are 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. Under this condition, it is further assumed that the equalizer is initialized such that $s(n) = \boldsymbol{u}_k(n)\boldsymbol{w}^o$ rather than a time-shifted version.

A. MEAN STABILITY ANALYSIS

Based on *A-1*, *A-3* and *A-4*, the expectation of (21) yields

$$E[\widetilde{\boldsymbol{w}}_n] = \mathcal{Q}_2^T (I_{MN} - \mathcal{M}\mathcal{D})\mathcal{Q}_1^T E[\widetilde{\boldsymbol{w}}_{n-1}] + \mathcal{Q}_2^T \mathcal{M}(\mathcal{D}\boldsymbol{w}^{(o)} - \gamma \mathcal{G}), \tag{22}$$

where

$$\begin{aligned} \mathcal{G} &= E[\mathcal{G}_n] \\ &= \mathrm{col}\{\sum_{l\in\mathcal{N}_1} b_{l1} E[\boldsymbol{u}_l^*(n)\mathrm{csgn}(y_l(n))], \ldots, \\ &\qquad \sum_{l\in\mathcal{N}_N} b_{lN} E[\boldsymbol{u}_l^*(n)\mathrm{csgn}(y_l(n))]\}, \end{aligned} \tag{23}$$

and

$$\mathcal{D} = E[\mathcal{D}_n] = \mathrm{diag}\{\sum_{l\in\mathcal{N}_1} b_{l1} R_{u,l}, \ldots, \sum_{l\in\mathcal{N}_N} b_{lN} R_{u,l}\}, \tag{24}$$

with the covariance matrix $R_{u,l} = E[\boldsymbol{u}_l^*(n)\boldsymbol{u}_l(n)]$.

Theorem 1: Assume data model (1) and the above assumptions *A-1*-*A-4* hold. Then, the mean of the weight error of the distributed blind equalizer using the diffusion GSA (14) converges with limitation

$$E[\boldsymbol{w}_\infty] = \boldsymbol{w}^{(o)} - (I_{MN} - \mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D})\mathcal{Q}_1^T)^{-1} \times \mathcal{Q}_2^T \mathcal{M}(\mathcal{D}\boldsymbol{w}^{(o)} - \gamma \mathcal{G}), \tag{25}$$

for any initial condition and any choice of matrices Q_1 and Q_2 satisfying (15) if

$$0 < \mu_k < \frac{2}{\lambda_{\max}(\sum_{l\in\mathcal{N}_k} b_{lk} R_{u,l})}, \quad k = 1, \ldots, N. \tag{26}$$

Proof: Based on *A-2*, $\boldsymbol{u}_k(n)$ is bounded, and thus the covariance matrix $R_{u,l}$ and $\mathcal{G}$ are both bounded (Note that $|\mathrm{csgn}(\cdot)|$ is bounded by 1). For sufficiently small enough step-size μ_k, the term $\mathcal{Q}_2^T \mathcal{M}(\mathcal{D}\boldsymbol{w}^{(o)} - \gamma\mathcal{G})$ in (22) is bounded within a small value and thus the stability of $E[\tilde{\boldsymbol{w}}_n]$ is determined by the matrix $\mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D})\mathcal{Q}_1^T$.

Let $G = I_{MN} - \mathcal{M}\mathcal{D}$. Based on the above analysis, we have $E[\tilde{\boldsymbol{w}}_n]$ converge if the maximal eigenvalue of matrix

$$|\lambda_{\max}(\mathcal{Q}_2^T G \mathcal{Q}_1^T)| < 1. \tag{27}$$

Using the matrix-2 norm,[1] we have [7], [8]

$$||\mathcal{Q}_2^T G \mathcal{Q}_1^T||_2 \le ||\mathcal{Q}_1||_2 \cdot ||G||_2 \cdot ||\mathcal{Q}_2||_2. \tag{28}$$

Since the matrix $\mathcal{D}$ is Hermitian, G is also Hermitian. Recalling that $\mathcal{Q}_1 = Q_1 \otimes I_M$ and $\mathcal{Q}_2 = Q_2 \otimes I_M$, we have $\lambda(\mathcal{Q}_1) = \lambda(Q_1)$, $\lambda(\mathcal{Q}_2) = \lambda(Q_2)$. Moreover, from (15), we have $||Q_1||_2 \le 1$ and $||Q_2||_2 \le 1$. Since matrix G is symmetric, (28) reduces to the following

$$|\lambda_{\max}(\mathcal{Q}_2^T G \mathcal{Q}_1^T)| \le |\lambda_{\max}(G)|. \tag{29}$$

That is to say,

$$|\lambda_{\max}(G)| < 1, \tag{30}$$

is a sufficient condition to ensure the convergence of diffusion GSA. This requires that the step-size μ_k should satisfy (26). Under this condition, when $n \to \infty$, it is derived that diffusion GSA converges to a biased estimate of $\boldsymbol{w}^{(o)}$ with limitation given in (25). ■

[1] The 2-norm of a matrix A is defined as the largest singular value of A.

Remark 2: From (29), it is noted that ATC-GSA with $Q_1 = B = I_N$ and $Q_2 = A$ or CTA-GSA with $Q_1 = A$ and $Q_2 = B = I_N$ has smaller spectral radius than the no-cooperation GSA (Nc-GSA) with $Q_1 = Q_2 = B = I_N$, which indicates that the diffusion GSA converges much faster than the Nc-GSA without cooperation.

As shown in Theorem 1, $\boldsymbol{w}_n$ converges to a biased estimate of $\boldsymbol{w}^{(o)}$, and the bias depends on the envelope of the transmitted QAM data constellation, which can be classified into two types. In the first type, named as Type-I, the modulus of the real and imaginary parts of each complex QAM data are equal and thus we have $\gamma = |s_i(n)| = |s_r(n)|$. For instance, the 4-QAM belongs to this type. In the second type, termed as Type-II, the modulus of the real and imaginary parts of the complex QAM data can be different, while their statistics γ as defined in (5) equal to the same constant. Such kind of QAM includes the 16-QAM and 64-QAM. In the following, the mean stability of these two kinds of signals is analyzed individually.

To carry out the analysis, the following assumption is assumed. As proved in Theorem 1, by choosing suitable μ_l satisfying (26), each equalizer $\boldsymbol{w}_l(n)$ converges to the optimal equalizer $\boldsymbol{w}^o$ in the mean sense with a small bias. Under this condition, it is reasonable to assume that after a certain number of iterations for achieving convergence, $\boldsymbol{w}_l(n)$ is close enough to the optimal $\boldsymbol{w}^o$ such that the equalizer output $y_l(n)$ converges to the half space that has the correct sign as $s(n)$ in majority [26], [40]. That is,

$$\Pr[\text{csgn}(y_l(n)) \neq \text{csgn}(s(n))] < \epsilon, \tag{31}$$

where ϵ is a small positive constant.

1) TYPE-I QAM DATA SYMBOLS

For Type-I, we have $\gamma = |s_r(n)| = |s_i(n)|$ at each iteration n. By choosing suitable μ_k satisfying (26), we can ensure that the assumption (31) holds after a certain number of iterations for achieving convergence. Then, we have the following approximation

$$\gamma\,\text{csgn}(y_l(n)) = \gamma\,\text{csgn}(s(n)) + z_l(n) = s(n) + z_l(n), \tag{32}$$

where $z_l(n)$ is of zero-mean and variance $\sigma^2_{z,l}$, as $s(n)$ is of zero-mean as given in *A-1*. It is also assumed that $z_l(n)$ is independent on $\boldsymbol{u}_l(n)$.

Then, based on *A-5* and (32), we have the element of $\mathcal{D}_n\boldsymbol{w}^{(o)} - \gamma\mathcal{G}_n$ at each node k and iteration n as follows:

$$\begin{aligned}\mathcal{D}_{k,n}\boldsymbol{w}^o - \gamma\mathcal{G}_{k,n} &= \sum_{l\in\mathcal{N}_k} b_{lk}\boldsymbol{u}_l^*(n)(s(n) - s(n) - z_l(n)] \\ &= -\sum_{l\in\mathcal{N}_k} b_{lk}\boldsymbol{u}_l^*(n)z_l(n).\end{aligned} \tag{33}$$

Let

$$\begin{aligned}\mathcal{H}_{uz}(n) = \text{col}\{&\sum_{l\in\mathcal{N}_1} b_{l1}\boldsymbol{u}_l^*(n)z_l(n), \ldots, \\ &\sum_{l\in\mathcal{N}_N} b_{lN}\boldsymbol{u}_l^*(n)z_l(n)\}.\end{aligned} \tag{34}$$

According to *A-2* and *A-3*, the expectation (22) becomes

$$\begin{aligned}E[\widetilde{\boldsymbol{w}}_n] &= \mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D})\mathcal{Q}_1^T E[\widetilde{\boldsymbol{w}}_{n-1}] - \mathcal{Q}_2^T\mathcal{M}E[\mathcal{H}_{uz}(n)] \\ &= \mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D})\mathcal{Q}_1^T E[\widetilde{\boldsymbol{w}}_{n-1}],\end{aligned} \tag{35}$$

where the second equality holds as $z_l(n)$ is of zero-mean and also independent of $\boldsymbol{u}_l(n)$.

Then, by selecting suitable step-size μ_k satisfying (26), $E[\widetilde{\boldsymbol{w}}_n]$ converges to zero. That is to say, the estimator $\boldsymbol{w}_n$ converges to an asymptotically unbiased estimate of the optimal equalizer $\boldsymbol{w}^{(o)}$ for the Type-I QAM data.

2) TYPE-II QAM DATA SYMBOLS

For Type-II, the statistics γ is a constant. Then, based on the assumption (31), we have

$$\gamma\,\text{csgn}(y_l(n)) = \gamma\,\text{csgn}(s(n)) + z_l(n). \tag{36}$$

Different from Type-I, the second equality in (32) does not hold, since $\gamma \neq |s_r(n)| \neq |s_i(n)|$ at each iteration n.

Letting

$$\begin{aligned}\mathcal{H}_{us}(n) = [s(n) - \gamma\,\text{csgn}(s(n))]\text{col}\{&\sum_{l\in\mathcal{N}_1} b_{l1}\boldsymbol{u}_l^*(n), \ldots, \\ &\sum_{l\in\mathcal{N}_N} b_{lN}\boldsymbol{u}_l^*(n)\},\end{aligned} \tag{37}$$

the term $\mathcal{D}_n\boldsymbol{w}^{(o)} - \gamma\mathcal{G}_n$ becomes

$$\mathcal{D}_n\boldsymbol{w}^{(o)} - \gamma\mathcal{G}_n = \mathcal{H}_{us}(n) - \mathcal{H}_{uz}(n), \tag{38}$$

where $\mathcal{H}_{uz}(n)$ is given in (34).

Based on *A-2*, *A-3* and (38), the expectation (22) becomes the following for Type-II

$$\begin{aligned}E[\widetilde{\boldsymbol{w}}_n] &= \mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D})\mathcal{Q}_1^T E[\widetilde{\boldsymbol{w}}_{n-1}] \\ &\quad + \mathcal{Q}_2^T\mathcal{M}E[H_{us}(n)] - \mathcal{Q}_2^T\mathcal{M}E[\mathcal{H}_{uz}(n)] \\ &= \mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D})\mathcal{Q}_1^T E[\widetilde{\boldsymbol{w}}_{n-1}] + \mathcal{Q}_2^T\mathcal{M}E[\mathcal{H}_{us}(n)].\end{aligned} \tag{39}$$

Comparing the mean weight error for Type-II with that for Type-I given in (35), we find that an additional bias term related to $\mathcal{H}_{us}(n)$ is introduced in (39), which reflects the difference between $s(n)$ and the quantity $\gamma\,\text{csgn}(s(n))$. Since the mean of term $s(n) - \gamma\,\text{csgn}(s(n))$ does not equal to zero, the second term on the right hand side (RHS) of (39) always exists. Therefore, the estimate $\boldsymbol{w}_n$ yields to a biased estimate of the optimal equalizer $\boldsymbol{w}^{(o)}$ for Type-II QAM data. From this point of view, it is guessed that the performance of diffusion GSA for estimating Type-I source signal outperforms that for Type-II. This is also verified by simulations in Sec. V.

B. MEAN-SQUARE STABILITY ANALYSIS

By resorting to the energy conservation analysis, we have the MSD of the weight error vector

$$\begin{aligned}E||\widetilde{\boldsymbol{w}}_n||^2_\Sigma &= E||\widetilde{\boldsymbol{w}}_{n-1}||^2_{\Sigma'} + E[(\mathcal{D}_n\boldsymbol{w}^{(o)} - \gamma\mathcal{G}_n)^T \\ &\quad \times \mathcal{M}\mathcal{Q}_2\Sigma\mathcal{Q}_2^T\mathcal{M}(\mathcal{D}_n\boldsymbol{w}^{(o)} - \gamma\mathcal{G}_n)] \\ &\quad + 2E[(\mathcal{D}_n\boldsymbol{w}^{(o)} - \gamma\mathcal{G}_n)^T\mathcal{M}\mathcal{Q}_2\Sigma \\ &\quad \cdot \mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D}_n)\mathcal{Q}_1^T\widetilde{\boldsymbol{w}}_{n-1}],\end{aligned} \tag{40}$$

where

$$\begin{aligned}\Sigma' = {} & \mathcal{Q}_1\mathcal{Q}_2\Sigma\mathcal{Q}_2^T\mathcal{Q}_1^T - \mathcal{Q}_1\mathcal{D}\mathcal{M}\mathcal{Q}_2\Sigma\mathcal{Q}_2^T\mathcal{Q}_1^T \\ & - \mathcal{Q}_1\mathcal{Q}_2\Sigma\mathcal{Q}_2^T\mathcal{M}\mathcal{D}\mathcal{Q}_1^T \\ & + \mathcal{Q}_1 E[\mathcal{D}_n\mathcal{M}\mathcal{Q}_2\Sigma\mathcal{Q}_2^T\mathcal{M}\mathcal{D}_n]\mathcal{Q}_1^T, \end{aligned} \tag{41}$$

and notation $||x||_\Sigma^2 = x^*\Sigma x$ represents a weighted vector norm for any Hermitian positive-definite matrix Σ.

Let $\sigma = \text{vec}\{\Sigma\}$ denote vectorization of matrix Σ, which is obtained by stacking the columns of its matrix argument on top of each other. Using the vectorization operator and the following Kronecker product property [41], [42]

$$\text{vec}(A\Sigma B) = (B^T \otimes A)\sigma, \tag{42}$$

we can rewrite those terms on RHS of (40) as follows

$$\begin{aligned}&\text{vec}\{\mathcal{Q}_1\mathcal{Q}_2\Sigma\mathcal{Q}_2^T\mathcal{Q}_1^T\} \\ &\quad = (\mathcal{Q}_1\otimes\mathcal{Q}_1)(\mathcal{Q}_2\otimes\mathcal{Q}_2)\sigma, \end{aligned} \tag{43a}$$

$$\begin{aligned}&\text{vec}\{\mathcal{Q}_1\mathcal{D}\mathcal{M}\mathcal{Q}_2\Sigma\mathcal{Q}_2^T\mathcal{Q}_1^T\} \\ &\quad = (\mathcal{Q}_1\otimes\mathcal{Q}_1)(\mathcal{D}^T\mathcal{M}\otimes I_{MN})(\mathcal{Q}_2\otimes\mathcal{Q}_2)\sigma, \end{aligned} \tag{43b}$$

$$\begin{aligned}&\text{vec}\{\mathcal{Q}_1\mathcal{Q}_2\Sigma\mathcal{M}\mathcal{Q}_2^T\mathcal{D}\mathcal{Q}_1^T\} \\ &\quad = (\mathcal{Q}_1\otimes\mathcal{Q}_1)(I_{MN}\otimes\mathcal{D}\mathcal{M})(\mathcal{Q}_2\otimes\mathcal{Q}_2)\sigma, \end{aligned} \tag{43c}$$

and

$$\begin{aligned}&\text{vec}\{E[\mathcal{Q}_1\mathcal{D}_n\mathcal{M}\mathcal{Q}_2\Sigma\mathcal{Q}_2^T\mathcal{M}\mathcal{D}_n\mathcal{Q}_1^T]\} \\ &\quad = (\mathcal{Q}_1\otimes\mathcal{Q}_1)E[(\mathcal{D}_n^T\mathcal{M})\otimes(\mathcal{D}_n\mathcal{M})](\mathcal{Q}_2\otimes\mathcal{Q}_2)\sigma. \end{aligned} \tag{43d}$$

The second term in (40) gives

$$E[(\mathcal{D}_n\boldsymbol{w}^{(o)} - \gamma\mathcal{G}_n)^T\mathcal{M}\mathcal{Q}_2\Sigma\mathcal{Q}_2^T\mathcal{M}(\mathcal{D}_n\boldsymbol{w}^{(o)} - \gamma\mathcal{G}_n)] = \alpha^T\sigma \tag{44}$$

where $\alpha = \text{vec}\{\mathcal{Q}_2^T\mathcal{M}(\mathcal{D}\boldsymbol{w}^{(o)} - \gamma\mathcal{G})(\mathcal{D}\boldsymbol{w}^{(o)} - \gamma\mathcal{G})^T\mathcal{M}\mathcal{Q}_2\}$, and the third term in (40) gives

$$\begin{aligned}&2E[(\mathcal{D}_n\boldsymbol{w}^{(o)} - \gamma\mathcal{G}_n)^T\mathcal{M}\mathcal{Q}_2\Sigma\mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D}_n)\mathcal{Q}_1^T\widetilde{\boldsymbol{w}}_{n-1}] \\ &\qquad = \beta^T\sigma \end{aligned} \tag{45}$$

where $\beta = 2\mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D})\mathcal{Q}_1^T E[\widetilde{\boldsymbol{w}}_{n-1}] \otimes \{\mathcal{Q}_2^T\mathcal{M}(\mathcal{D}\boldsymbol{w}^{(o)} - \gamma\mathcal{G})\}$.

In summary, based on the above results, the equation (40) can be described by the following recursion

$$E||\widetilde{\boldsymbol{w}}_n||_\sigma^2 = E||\widetilde{\boldsymbol{w}}_{n-1}||_{F\sigma}^2 + \alpha^T\sigma + \beta^T\sigma, \tag{46}$$

where

$$\begin{aligned}F = {} & (\mathcal{Q}_1\otimes\mathcal{Q}_1)\{I_{M^2N^2} - I_{MN}\otimes(\mathcal{D}\mathcal{M}) - (\mathcal{D}^T\mathcal{M})\otimes I_{MN} \\ & + E[(\mathcal{D}_n^T\mathcal{M})\otimes(\mathcal{D}_n\mathcal{M})](\mathcal{Q}_2\otimes\mathcal{Q}_2). \end{aligned} \tag{47}$$

Let

$$\begin{aligned}S = {} & I_{M^2N^2} - I_{MN}\otimes(\mathcal{D}\mathcal{M}) - (\mathcal{D}^T\mathcal{M})\otimes I_{MN} \\ & + E[(\mathcal{D}_n^T\mathcal{M})\otimes(\mathcal{D}_n\mathcal{M})]. \end{aligned} \tag{48}$$

To ensure the stability in the mean-square sense, we should select suitable step-size μ_k and combination matrices $\mathcal{Q}_1$ and $\mathcal{Q}_2$ such that all the eigenvalues of F satisfy

$$|\lambda_{\max}(F)| < 1. \tag{49}$$

Therefore, to ensure the stability in the mean and mean-square senses, it is required that μ_k should satisfy both (30) and (49). Similar to the analysis in the mean stability and using the matrix 2-norm, we have

$$\begin{aligned}||F||_2 &= ||(\mathcal{Q}_1\otimes\mathcal{Q}_1)\cdot S\cdot(\mathcal{Q}_2\otimes\mathcal{Q}_2)||_2 \\ &\le ||\mathcal{Q}_1\otimes\mathcal{Q}_1||_2\cdot||S||_2\cdot||(\mathcal{Q}_2\otimes\mathcal{Q}_2)||_2 \\ &\le ||\mathcal{Q}_1||_2^2\cdot||\mathcal{Q}_2||_2^2\cdot||S||_2. \end{aligned} \tag{50}$$

Recalling that $\mathcal{Q}_1 = Q_1\otimes I_N$ and $\mathcal{Q}_2 = Q_2\otimes I_N$, we have $||\mathcal{Q}_1||_2 = ||Q_1||_2$, $||\mathcal{Q}_2||_2 = ||Q_2||_2$. Since matrix S is symmetric, we obtain that

$$|\lambda_{\max}(F)| \le ||Q_1||_2^2\cdot||Q_2||_2^2\cdot|\lambda_{\max}(S)|. \tag{51}$$

Equation (51) implies that the stability of the overall system is governed by the property of the matrix S and the designed cooperation protocol (represented by Q_1 and Q_2). By choosing suitable combination protocol satisfying (15), such as the Metropolis and Laplacian rules, we have $||Q_1||, ||Q_2|| \le 1$, and thus

$$|\lambda_{\max}(F)| \le |\lambda_{\max}(S)|. \tag{52}$$

Furthermore, for sufficiently small step-size μ_k satisfying (26), we can approximate S by the following

$$S \approx (I_{MN} - \mathcal{D}^T\mathcal{M})\otimes(I_{MN} - \mathcal{D}\mathcal{M}), \tag{53}$$

which is stable if, and only if, $I_{MN} - \mathcal{D}\mathcal{M}$ is stable. This condition is consistent with the condition (26) for ensuring the mean stability.

To summarize, by choosing a suitable cooperation protocol that ensures $||Q_1||, ||Q_2|| \le 1$ and the step-size μ_k satisfying (26), the mean and mean-square stability of the in-network distributed GSA can be ensured.

Remark 3: If we choose the combination matrix $B = I_N$, S in (53) corresponds to the case without cooperation. According to (52), by selecting suitable combination matrices Q_1 and Q_2, the spectral radius of F in the diffusion GSA is generally smaller than that in the Nc-GSA without cooperation. Therefore, diffusion GSA converges much faster than the Nc-GSA in the mean-square convergence. This suggests that the cooperation based on the diffusion protocol (14) and (15) has a stabilizing effect on the mean and mean-square stability in the network, which is consistent with the results in distributed non-blind signal processing [8].

C. MEAN-SQUARE DEVIATION ANALYSIS FOR DIFFERENT SOURCES

The mean-square deviations for Type-I and Type-II mainly differ in the last term of RHS in (40). Next, we give the specific results of the considered two types of QAM signals with different envelope properties.

1) TYPE-I QAM DATA SYMBOLS

Based on (32) and (33), the vector α in (44) equals to

$$\alpha = \text{vec}\{\mathcal{Q}_2^T\mathcal{M}\mathcal{H}_u^T\mathcal{M}\mathcal{Q}_2\}, \tag{54}$$

where

$$\begin{aligned}\mathcal{H}_u &= E[\mathcal{H}_{uz}(n)\mathcal{H}_{uz}^*(n)] \\ &= \mathcal{B}^T \text{diag}\{\sigma_{z,1}^2 R_{u,1}, \ldots, \sigma_{z,N}^2 R_{u,N}\}\mathcal{B}, \end{aligned} \quad (55)$$

and for sufficiently small step-size μ_k satisfying (26), we have

$$\begin{aligned}\beta &= 2\mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D})\mathcal{Q}_1^T E[\widetilde{\boldsymbol{w}}_{n-1}] \otimes \{\mathcal{Q}_2^T \mathcal{M}(\mathcal{D}\boldsymbol{w}^{(o)} - \gamma\mathcal{G})\} \\ &= 2\mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D})\mathcal{Q}_1^T E[\widetilde{\boldsymbol{w}}_{n-1}] \otimes \{\mathcal{Q}_2^T \mathcal{M} E[\mathcal{H}_{uz}(n)]\} \\ &= 0. \end{aligned} \quad (56)$$

2) TYPE-II QAM DATA SYMBOLS

For Type-II, based on (38), the term α in (44) becomes

$$\begin{aligned}\alpha &= \text{vec}\{\mathcal{Q}_2^T \mathcal{M}\mathcal{H}_u^T \mathcal{M}\mathcal{Q}_2\} \\ &\quad + \text{vec}\{\mathcal{Q}_2^T \mathcal{M} E[\mathcal{H}_{us}(n)\mathcal{H}_{us}^*(n)]^T \mathcal{M}\mathcal{Q}_2\}, \end{aligned} \quad (57)$$

and β can be approximated by the following with a sufficient small step-size μ_k

$$\begin{aligned}\beta &= 2\mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D})\mathcal{Q}_1^T E[\widetilde{\boldsymbol{w}}_{n-1}] \otimes \{\mathcal{Q}_2^T \mathcal{M}(\mathcal{D}\boldsymbol{w}^{(o)} - \gamma\mathcal{G})\} \\ &= 2\mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D})\mathcal{Q}_1^T E[\widetilde{\boldsymbol{w}}_{n-1}] \\ &\quad \otimes \{\mathcal{Q}_2^T \mathcal{M}(E[\mathcal{H}_{uz}(n)] + E[\mathcal{H}_{us}(n)])\} \\ &= 2\mathcal{Q}_2^T(I_{MN} - \mathcal{M}\mathcal{D})\mathcal{Q}_1^T E[\widetilde{\boldsymbol{w}}_{n-1}] \otimes \{\mathcal{Q}_2^T \mathcal{M} E[\mathcal{H}_{us}(n)]\}. \end{aligned} \quad (58)$$

By selecting suitable step-size μ_k satisfies (26), we can ensure that matrix $(I_{M^2N^2} - F)$ is invertible, and thus the steady-state MSD of the diffusion GSA equals to

$$E||\widetilde{\boldsymbol{w}}_\infty|| = (I_{M^2N^2} - F)^{-1}(\alpha + \beta)^T \sigma. \quad (59)$$

Based on the above analysis, it is noticed that an additional term $\beta > 0$ exists for Type-II but not in Type-I, which leads to a larger MSD in Type-II. This conclusion is also verified by simulations in the following Sec. V.

V. NUMERICAL SIMULATIONS

In this section, two simulative examples are performed to show the effectiveness of the proposed diffusion GSAs for blind signal estimation through the design of distributed blind equalizer.

In the simulations, a sensor network consisting of 16 sensors uniformly randomly distributed over 50m $\times$ 50m square area, is adopted. Considering that the power of a sensor is restricted, the network topology is constructed by restricting the maximal transmission range $R = 10$m. That is, two sensors are allowed to communicate with each other if their Euclidean distance is within the maximum transmission range R. Otherwise, there is no information exchange (no connection) between them. Finally, we obtain the network topology as given in Fig. 3(a).

The additive measurement noise $v_k(n)$ is selected such that the signal-to-noise ratio (SNR) for each sensor randomly distributes within (6, 10)dB, see Fig. 3(b). The step-size for each sensor k is set as $\mu_k = 0.001$. The complex impulse

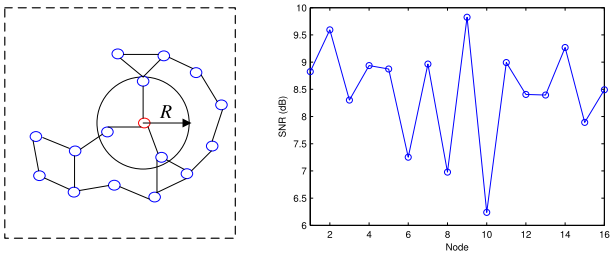


FIGURE 3. Network settings. (a) Network topology (b) Initial SNR vs node.

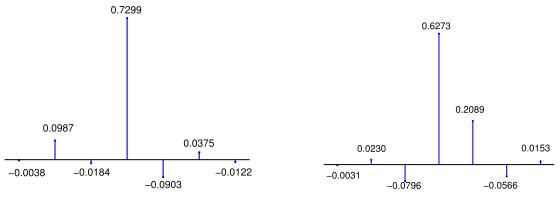
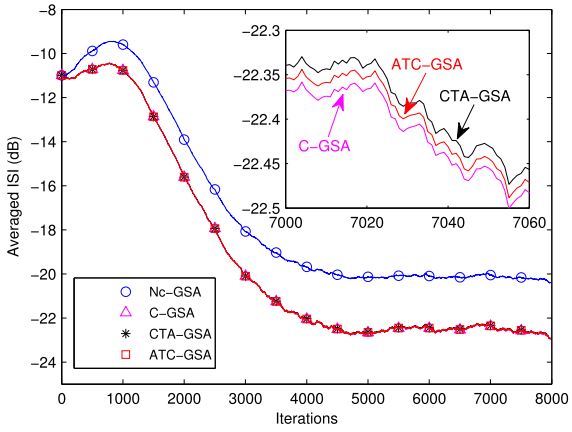


FIGURE 4. Channel impulse response. (a) Real part. (b) Imaginary part.

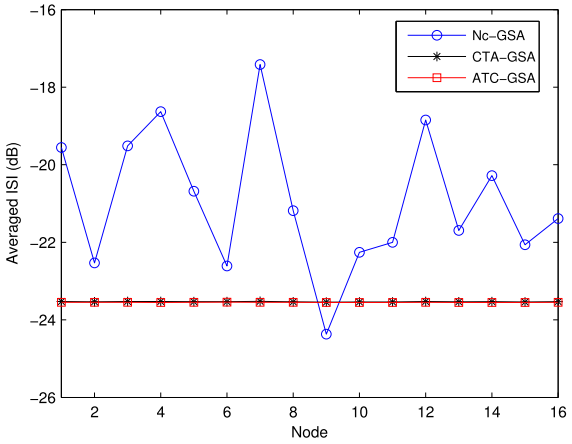
response of the transmission channel $\{h(n)\}$ used in the simulation is depicted in Fig. 4. In our simulation, a 20-tap complex equalizer of a transversal filter structure is used and initialized such that the center tap is set to one and the other taps are set to zeros. The coefficients $q_{1,lk}$ and $q_{2,lk}$ of matrices Q_1 and Q_2 are chosen according to the Metropolis rule, while the matrix B is simply set to be a unit matrix, i.e. $B = I_N$. The simulation results using the no-cooperation GSA (Nc-GSA), the centralized GSA (where the measurements $\{u_k(n)\}$ of all the N sensors in the network are conveyed to a fusion center for processing), denoted as C-GSA, and the diffusion GSA, including ATC-GSA and CTA-GSA are presented and compared. In the following, two types of source signals, 4-QAM and 16 QAM, are used to testify the performance of different algorithms, respectively.

A. EXAMPLE 1: 4-QAM DATA SYMBOLS

In the first example, Type-I source signal $s(n)$ generated from 4-QAM data constellation is considered. We use the measure of residual intersymbol interference (ISI), typically adopted in the context of blind equalization [27], [30], [43] to evaluate the convergence behaviors of different blind adaptive algorithms. Fig. 5(a) shows the averaged transient ISI of the whole network over 50 simulations, while Fig. 5(b) depicts the averaged steady-state ISI at each sensor, which is computed by the ISI after convergence (5000 iterations). From the simulation results, we find that without cooperation, the ISI is relatively larger. By cooperating with the neighbors in the network using the diffusion cooperation, the ISI has been significantly reduced, which is about 2dB better than the Nc-GSA. As shown in Fig. 5, CTA-GSA and ATC-GSA show similar performance, and they both converge to the global optimization algorithm, C-GSA. Moreover, from Fig. 5(b),



(a)



(b)

FIGURE 5. Simulation results for 4-QAM constellation. (a) Averaged transient ISI vs iteration. Note that the simulation results for Nc-GSA, ATC-GSA, and CTA-GSA from 7000 to 7060 iterations are enlarged in the subfigure for illustration. (b) Averaged steady-state ISI vs node.

it is observed that the steady-state ISIs at different sensors for the Nc-GSA differ significantly from each other, which are dependent on the initial SNRs and the network topological structure as given in Fig. 3. Although the initial SNRs differ significantly among sensors, less difference is noticed after equalization using the diffusion cooperation.

Based on the obtained equalizers, we can recover the source signal $s(n)$ using the final estimate of equalizer at any sensor, denoted as $\hat{s}(n)$. To compare the performance of different algorithms, the signal error rate (SER) is computed, which is defined as

$$\text{SER} = \frac{\text{the number of the error symbols in } \hat{s}(n)}{\text{the total number of the source signal } s(n)} \times 100\%. \tag{60}$$

Note that for the diffusion GSAs and C-GSA, to further improve the accuracy of estimation, we can determine the transmitted data symbols at each sensor k by selecting the recovered symbol with the largest probability based on the estimates of k's neighbors. The averaged SERs computed

TABLE 2. Averaged SER over sensor networks using different algorithms for 4-QAM.

	Nc-GSA	C-GSA	CTA-GSA	ATC-GSA
SER(%)	2.55%	0.06%	0.06%	0.06%

from the N sensors in the network using different algorithms are summarized in Table 2. The results of SER are consistent with that of the channel equalization as reflected by ISI given in Fig. 5. That is, the better the performance of channel equalization (the smaller the ISI), the more accurate the estimation (the smaller the SER). As given in Fig. 5 and Table 2, using a distributed diffusion cooperation, the proposed diffusion GSA algorithms show better performance in reducing the ISI and furthermore estimating the source, which can be comparable to the C-GSA.

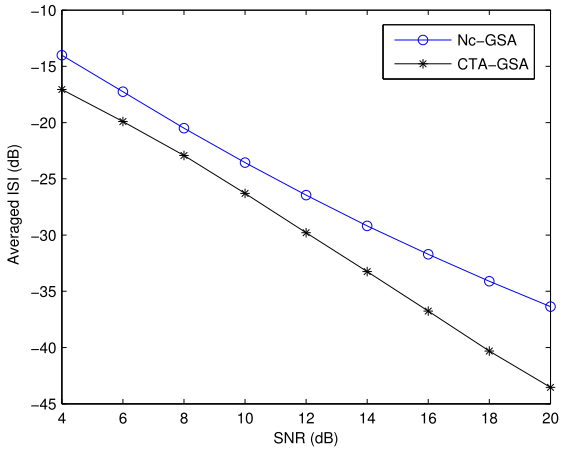


FIGURE 6. Averaged steady-state ISI vs SNR using different algorithms.

We also verify the effect of SNR to the performance of the proposed blind equalizers. As the results evaluated by the measure of ISI are consistent with that evaluated by the SER, we only use the measure of ISI for performance evaluation. In the simulation, the same initial SNR is set for each node, and changes from 4dB to 20dB. Different methods are performed and the averaged steady-states ISI over 50 simulations are given in Fig. 6. It is remarked that, as the CTA-GSA, ATC-GSA and C-GSA perform approximately the same, we only give the result of CTA-GSA for illustration. From Fig. 6, it is found that, comparing to the Nc-GSA, CTA-GSA has improved the performance of equalization irrespective of the initial SNRs. Moreover, with the increasing of the initial SNR, the performance enhancement also increases a little.

The robustness of diffusion GSA in blind equalization is also testified in Fig. 7. In this study, 4 out of 16 sensors are assumed to be abnormal. Note that if kth-sensor is abnormal, it loses the measurements of the source but only collects the noise output, i.e. $\boldsymbol{u}_k(n) = v_k(n)$. It is also noted that the abnormal nodes still have the ability to compute and communicate with their neighbors. Therefore, for an abnormal sensor k, based on the information transmitted from its

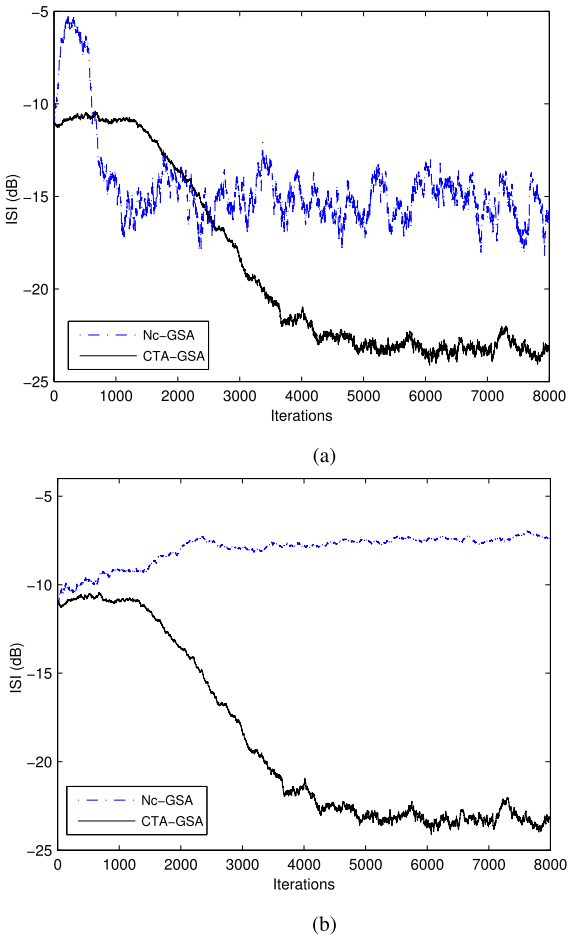


FIGURE 7. Robustness analysis. (a) Transient ISI vs iteration for a normal sensor. (b) Transient ISI vs iteration for an abnormal sensor.

neighbors $\boldsymbol{u}_l(n)$, $l \in \mathcal{N}_k$, it is still possible to give an estimate of the channel equalizer and furthermore recover the source signal using the proposed diffusion GSAs.

Fig. 7(a) and (b) respectively compare the ISI of a normal sensor and an abnormal sensor in the network using the Nc-GSA and CTA-GSA. The combination matrices B and Q_2 for CTA-GSA are selected by the Metropolis rule, while $Q_1 = I_N$. Note that since the results of the ATC-GSA and C-GSA are quite similar to the CTA-GSA, their corresponding results are omitted here for clarity. As shown in Fig. 7(a), if a sensor works normally, even without cooperating with other sensors, it can reduce the ISI to some extent, see the result of the Nc-GSA. But, it is inferior to the CTA-GSA. On the other hand, if a sensor works abnormally, the Nc-GSA performs poorly, and the ISI does not converge as the sensor only observes noise. However, even if a sensor fails to collect the information of the measurements, it still works well by cooperating with the other sensors using the CTA-GSA, see Fig. 7(b). The simulation results suggest that the CTA-GSA is more robust, and thus more reliable in practical applications.

Besides, to testify the applicability of the proposed diffusion GSA in a real far-end network environment, in which the

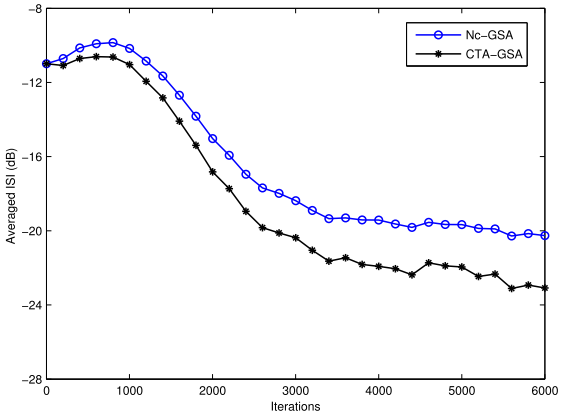


FIGURE 8. Averaged transient ISI vs iteration using different algorithms for 4-QAM constellation with some channel perturbations.

channels share some similarity but not exactly the same, we also perform the simulation by adding different perturbations on the channels $h(n)$ at different nodes. In the simulation, the perturbation is generated from a zero-mean Gaussian with a small variance $\sigma_{h,k} = 0.01$. Similar to the above simulation, the results of Nc-GSA and CTA-GSA are given in Fig. 8 for comparison. From the simulation results, we find that CTA-GSA also works in this case, indicating the effectiveness of distributed cooperation in real far-end sensor networks, which allows for some difference among nodes.

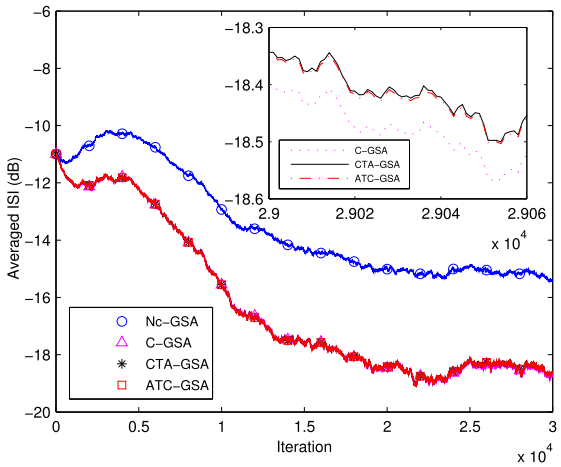


FIGURE 9. Averaged transient ISI vs iteration using different algorithms for 16-QAM constellation. Note that the simulation results for C-GSA, ATC-GSA, and CTA-GSA from 2.900×10^4 to 2.906×10^4 iterations are enlarged in the subfigure for illustration.

B. EXAMPLE 2: 16-QAM DATA SYMBOLS

In the second example, we apply the diffusion GSAs to estimate the Type-II source signal, which is generated from 16-QAM constellation. The same parametric settings as Example 1 are used. The averaged transient ISIs using different algorithms are depicted in Fig. 9, while the corresponding SERs are summarized in Table 3.

TABLE 3. Averaged SER over sensor networks using different algorithms for 16-QAM.

	Nc-GSA	C-GSA	CTA-GSA	ATC-GSA
SER(%)	9.86%	5.96%	6.16%	6.15%

Similar to the results for the 4-QAM, the diffusion generalized Sato algorithms, ATC-GSA and CTA-GSA both outperform the Nc-GSA and converge to the C-GSA. However, comparing to the results in Fig. 5 for Example 1, the performance of the diffusion GSAs for the 16-QAM constellation is inferior to that for the 4-QAM, which is caused by the difference between the statistics γ and the true modulus of the real and imaginary parts of the transmitted data symbols $s(n)$ as explained in Sec. IV-A. On the meanwhile, from Table 3, it is noticed that the SERs of the estimated data series are also larger than those in Example 1. This is because the decision of the data symbol $s(n)$ is dependent on the magnitude of the equalizer output. Since the gain of total transmission system, i.e. the channel model convolved by the designed channel equalizer, is not exactly equal to 1, the magnitude of the equalizer output may oscillate around the space that corresponds to the right symbol $s(n)$, which causes a larger SER. In contrast, the decision of the 4-QAM data symbols is only dependent on the complex sign of the equalizer output $y_k(n)$, but insensitive to its magnitude. Therefore, a smaller SER can be achieved for the 4-QAM.

VI. CONCLUSION

In this paper, a kind of diffusion blind adaptive algorithm, based on the generalized Sato algorithm, is proposed to design distributed blind equalizer to estimate the transmitted QAM data sequences, where the channel models are assumed to be unknown. The performance of the proposed diffusion GSA in mean and mean-square senses is analyzed theoretically. Then, a series of numerical simulations are presented to show the effectiveness of the proposed algorithms in blind equalization and source estimation. Results show that by cooperating with the nodes in the neighborhood, diffusion GSA can efficiently reduce the ISI as well as improve the accuracy of estimation. Moreover, it is more robust to sensor failure and thus more applicable in practical applications.

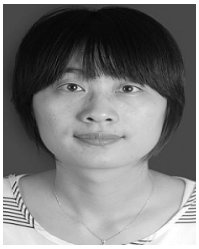
REFERENCES

[1] A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione, "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, X. Zhao, and Z. J. Towfic, "Diffusion strategies for adaptation and learning over networks," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 155–171, May 2013.
[3] A. H. Sayed and C. G. Lopes, "Adaptive processing over distributed networks," *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. 90, no. 8, pp. 1504–1510, Aug. 2007.
[4] 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.
[5] Y. Liu and W. K. S. Tang, "Enhanced incremental LMS with norm constraints for distributed in-network estimation," *Signal Process.*, vol. 94, pp. 373–385, Jan. 2014.
[6] 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. Signal Process.*, vol. 58, no. 1, pp. 151–164, Jan. 2010.
[7] 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.
[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] A. Rastegarnia, M. A. Tinati, and A. Khalihi, "A diffusion least-mean-square algorithm for distributed estimation over senor networks," *Int. J. Elect., Comput., Syst. Eng.*, vol. 2, no. 1, pp. 15–19, 2008.
[10] Y. Liu, C. Li, and Z. Zhang, "Diffusion sparse least-mean squares over networks," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4480–4485, Aug. 2012.
[11] F. S. Cattivelli, C. G. Lopes, and A. H. Sayed, "Diffusion recursive least-squares for distributed estimation over adaptive networks," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1865–1877, May 2008.
[12] 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, Mar. 2014.
[13] C. Li, P. Shen, Y. Liu, and Z. Zhang, "Diffusion information theoretic learning for distributed estimation over network," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4011–4024, Aug. 2013.
[14] P. Shen and C. Li, "Distributed information theoretic clustering," *IEEE Trans. Signal Process.*, vol. 62, no. 13, pp. 3442–3453, Jul. 2014.
[15] I. D. Schizas, G. Mateos, and G. B. Giannakis, "Distributed LMS for consensus-based in-network adaptive processing," *IEEE Trans. Signal Process.*, vol. 8, no. 6, pp. 2365–2381, Jun. 2009.
[16] 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.
[17] 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.
[18] K. Abed-Meraim, W. Qiu, and Y. Hua, "Blind system identification," *Proc. IEEE*, vol. 85, no. 8, pp. 1310–1322, Aug. 1997.
[19] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.*, vol. 13, no. 3, pp. 43–64, May 1996.
[20] L. C. Wood and S. Treitel, "Seismic signal processing," *Proc. IEEE*, vol. 63, no. 4, pp. 649–661, Apr. 1975.
[21] R. Johnson, Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: A review," *Proc. IEEE*, vol. 86, no. 10, pp. 1927–1950, Oct. 1998.
[22] Y. Xiao, G. Z. Liu, and Z. X. Li, "Blind equalization for underwater acoustic communication by genetic algorithm optimizing neural network," *Appl. Acoust.*, vol. 25, no. 6, pp. 340–345, Jun. 2006.
[23] Y. B. Zhang, J. W. Zhao, Y. C. Guo, and J. M. Li, "Blind adaptive MMSE equalization of underwater acoustic channels based on the linear prediction method," *J. Marine Sci. Appl.*, vol. 10, no. 1, pp. 113–120, Jan. 2011.
[24] Y. Xiao and M. Li, "Recursive least squares fractionally-spaced blind equalization algorithm for underwater acoustic communication," *J. Inf. Comput. Sci.*, vol. 10, no. 18, pp. 6077–6084, Dec. 2013.
[25] Y. Sato, "A method of self-recovering equalization for multilevel amplitude-modulation systems," *IEEE Trans. Commun.*, vol. COM-23, no. 6, pp. 679–682, Jun. 1975.
[26] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Trans. Commun.*, vol. COM-32, no. 8, pp. 871–883, Aug. 1984.
[27] Z. Ding and Y. Li, *Blind Equalizaiton and Indentification*. New York, NY, USA: Marcel Dekker, 2001.
[28] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. COM-28, no. 11, pp. 1867–1875, Nov. 1980.
[29] J. R. Treichler and M. G. Larimore, "New processing techniques based on the constant modulus adaptive algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 420–431, Apr. 1985.
[30] S. Abrar and A. K. Nandi, "An adaptive constant modulus blind equalization algorithm and its stochastic stability analysis," *IEEE Signal Process. Lett.*, vol. 17, no. 1, pp. 55–58, Jan. 2010.
[31] W. Rao, W. Q. Tan, Y. M. Li, and H. J. Gao, "New modified constant modulus algorithm for underwater acoustic communications," in *Proc. Int. Conf. Internet Comput. Inf. Services (ICICIS)*, 2011, pp. 563–566.

[32] R. Abdolee and B. Champagne, "Distributed blind adaptive algorithms based on constant modulus for wireless sensor networks," in *Proc. 6th Int. Conf. Wireless Mobile Commun.*, Sep. 2010, pp. 303–308.
[33] 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.
[34] S. C. Draper, B. J. Frey, and F. R. Kschischang, "Interactive decoding of a broadcast message," in *Proc. 41st Allerton Conf. Commun. Control Comput.*, Oct. 2003, pp. 1–11.
[35] R. Dabora and S. D. Servetto, "Broadcast channels with cooperating decoders," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5438–5454, Dec. 2006.
[36] 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.
[37] R. Cusani and A. Laurenti, "Convergence analysis of the CMA blind equalizer," *IEEE Trans. Commun.*, vol. 43, nos. 2–4, pp. 1304–1307, Feb. 1995.
[38] O. Dabeer and E. Masry, "Convergence analysis of the constant modulus algorithm," *IEEE Trans. Inf. Theory*, vol. 49, no. 6, pp. 1447–1464, Jun. 2003.
[39] V. H. Nascimento and M. T. M. Silva, "Stochastic stability analysis for the constant-modulus algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4984–4989, Oct. 2008.
[40] M. G. Larimore and J. R. Treichler, "Convergence behavior of the constant modulus algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 8. Boston, MA, USA, Apr. 1983, pp. 13–16.
[41] D. S. Tracy and R. P. Singh, "A new matrix product and its applications in partitioned matrix differentiation," *Statist. Neerlandica*, vol. 26, no. 4, pp. 143–157, 1972.
[42] R. H. Koning, H. Neudecker, and T. Wansbeek, "Block Kronecker products and the vecb operator," Dept. Econ., Inst. Econ. Res., Univ. Groningen, Groningen, The Netherlands, Tech. Rep. 351, 1990.
[43] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Inf. Theory*, vol. IT-36, no. 2, pp. 312–321, Mar. 1990.

YING LIU received the B.S. degree in communication engineering from the College of Information Engineering, Dalian Maritime University, China, in 2003, and the M.S. degree in signal and information processing from the College of Electrical Engineering, Shantou University, China, in 2006, and the Ph.D. degree from Department of Electronic Engineering, City University of Hong Kong, in 2010.

She is currently an Associate Professor with the College of Information Science and Electronic Engineering, Zhejiang University. Her research interest is statistical signal processing and distributed information processing.



LILI LI received the B.S. degree in communication engineering from the College of Information Engineering, Dalian Maritime University, China, in 2003, and the M.S. degree in optical engineering from the College of Information Engineering, Zhejiang University, China, in 2006.

She is a currently a Senior Experimentalist with the School of Naval Architecture and Mechanical–Electrical Engineering, Zhejiang Ocean University. Her research interest is chaotic signal processing and statistic information processing.