Received June 4, 2017, accepted July 24, 2017, date of publication July 31, 2017, date of current version October 12, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2733766

Diffusion Signed LMS Algorithms and Their Performance Analyses for Cyclostationary White Gaussian Inputs

WENYUAN WANG AND HAIQUAN ZHAO, (Senior Member, IEEE)
[1]Key Laboratory of Magnetic Suspension Technology and Maglev Vehicle, Ministry of Education, Chengdu 610031, China
[2]School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China

Corresponding author: Haiquan Zhao (hqzhao_swjtu@126.com)

This work was supported by the National Science Foundation of China under Grant 61571374, Grant 61271340, and Grant 61433011.

ABSTRACT As one of the signed variants of the diffusion least mean square (DLMS) algorithm over networks, the diffusion sign error algorithm has been presented in previous reference. In this paper, we propose two novel signed variants of the DLMS algorithm, i.e., the diffusion signed regressor algorithm and the diffusion sign-sign algorithm. Moreover, this paper analyzes the performance of these three signed variants of the DLMS algorithm for cyclostationary white Gaussian inputs which have periodically time-varying variances. It is assumed that the distributed algorithms are in non-stationary environments. Specifically, the unknown parameter to be identified is time-varying according to the standard random walk model. The analysis models in terms of mean weight behavior and mean square performance are provided, in which, we can find some interesting results. Finally, simulations are carried out to verify the correctness of the proposed analysis model.

INDEX TERMS Distributed network, adaptive filter, sign algorithm, stochastic model, and cyclostationary signals.

I. INTRODUCTION

Distributed adaptive signal processing over the network is an attractive and challenging technique in which the nearby agents implement network-wide computation and control objectives through local information exchange [1]–[3]. It has been applied to a variety of occasions, such as factory automation, robotics, telecommunications and so on. There are several distributed strategies for parameters estimation including incremental strategy, diffusion strategy, and consensus strategy. Using these strategies, researchers have proposed many distributed algorithms including the incremental least mean square (ILMS) algorithm [4], the incremental affine projection algorithm (IAPA) [5], diffusion LMS (DLMS) [6], [7], diffusion recursive least square (DRLS) [8], consensus LMS [9], [10]. In the incremental algorithm, information flows from one node to the next in a sequential fashion [4], thus this algorithm would suffer from the performance degradation when the link failures occur [4], [5]. Two time-scales are required in consensus algorithm [9]. One is used to gather the measurement values of the network nodes, and another is used to iterate fully over the gathered values to obtain agreement till the next repeated process begins [9]–[12]. However, the implementations of two time-scales may result in the unstable behavior [12]. In this work, we mainly study diffusion strategy because it has proven to be more robust and still has stable performance, even when there are some unstable underlying nodes [6]–[8], [12]–[26].

Unfortunately, for today's high data rates, the aforementioned distributed algorithms including DLMS algorithm may be not practical due to their required multiplications which need long bit duration. To address this issue, in single-agent model conditions, the signed algorithms including signed regressor algorithm (SRA), signed error algorithm (SEA) and sign-sign algorithm (SSA), have been studied. The SRA was derived by introducing the signum function for regressor data [27]. The SEA was obtained by introducing the signum function for estimate error [28]–[30]. The SSA was derived by introducing the signum function for both estimate error and regressor data [32], [33]. The study of the transient behavior and the steady state performance for these three algorithms has been provided for the stationary input in the literature [27], [28], [32].

In many practical applications, the man-made signals are typically cyclostationary [34]. Thereby, the cyclostationary signals are widely used in radar, communication and power systems [34]–[40]. In [35], the cyclostationary signals were selected as the quadrature amplitude modulated (QAM) signals. Nassar *et al.* [36] studied the narrowband (NB) power line communication (PLC) in which the additive noises were cyclostationary. Additionally, in the beamforming algorithms proposed in [37] and [38], the input signals were also cyclostationary. The stochastic analyses of the adaptive algorithms for the cyclostationary inputs were also studied in the literature [41]–[49]. The work [41] studied the convergence of the mean weight for the LMS algorithm when the input signal is cyclostationary. In [42]–[45], the mean square performance of the LMS was analyzed for the cyclostationary input when the system was time-variant. In [46], the behaviors of the least mean fourth (LMF) algorithm were studied for the cyclostationary input. The work [49] presented the novel algorithm called time-average LMS (TA-LMS) to enhance the performance of LMS when the input signals were cyclostationary. Also, the detail performance analysis of the TA-LMS algorithm was also provided. Eweda and Bershad [48] presented the analysis of the signed LMS algorithms in the nonstationary conditions for cyclostationary Gaussian inputs. However, the analysis of the distributed algorithms for cyclostationary Gaussian input is absent in the previous literature.

In the multi-agent condition, one of the signed variants of diffusion LMS algorithm has been proposed in [20], which is called diffusion sign error algorithm (DSEA). In addition, the performance analysis of the DSEA for stationary input was also given in [20]. In this paper, we propose two other variants of signed LMS algorithm. The first is diffusion signed regressor algorithm (DSRA). The second is diffusion sign-sign algorithm (DSSA). The DSRA, DSEA and DSSA constitute a family of the diffusion signed LMS algorithms. Although the diffusion signed LMS algorithms may be inferior to DLMS algorithm in terms of performance, they are not required multiplications which need long bit duration. Thereby, the diffusion signed LMS algorithms are suitable for the practical applications. Then, we provide the statistical analyses of the three variants of the diffusion signed LMS algorithm for cyclostationary white Gaussian inputs under nonstationary environments. Compared with mean square error (MSE) and excess MSE (EMSE), the mean weight behavior and mean square weight behavior is easy to analyze. In addition, the MSE and EMSE can be easily derived from the mean weight behavior and mean square weight behavior [54], [55]. Thereby, we choose mean weight behavior and mean square weight behavior to analyze the proposed algorithm. From the results of the analysis, since the effects of the cyclostationary input at different nodes could be offset, we can get that the proposed algorithms can suppress the impact of the cyclostationary input. Simulations are also conducted to illustrate the consistency between the theoretically predicted behaviors and the simulated results.

The rest of the paper is organized as follows. In section II, we give the problem formulation. In section III, we propose two other variants of the diffusion signed LMS algorithm. In section IV, the analysis of DSEA for cyclostationary inputs is provided. In section V, we analyze the performance of the DSRA for cyclostationary inputs. In section VI, we provide the analysis of the DSSA for cyclostationary inputs. In section VII, we carry out the simulations to verify the correctness of the analyses. In section VIII the conclusion of this paper is drawn.

Notation: I_a (a is a positive integer) denotes identity matrix of size $a \times a$. $\otimes$ denotes Kronecker product. $\mathcal{I}_k$ is a matrix of size $M \times MN$. Averagely partitioning $\mathcal{I}_k$ into N square matrices of size $M \times M$, the kth matrix is identity matrix and others are zero matrices. $\circ$ denotes the Hadamard product. $Tr[\bullet]$ denotes the trace operator. $[\bullet]^T$ denotes the transposition for real vectors and matrices. $E[\bullet]$ denotes the expectation operator. $diag\{\bullet\}$ formulates a (block) diagonal matrix with its arguments. $\mathbf{1}_a$ (a is a positive integer) is a row vector with a length of a whose entries are 1. $\mathbf{1}'_{N\times N}$ is a square matrix of size $N \times N$ whose entries are 1. $\mathbb{I}$ is a square matrix of size $NM \times NM$. We partition $\mathbb{I}$ into $N \times N$ blocks, and each block is a square matrix of size $M \times M$. The diagonal blocks of $\mathbb{I}$ are $\mathbf{1}'_{M\times M}$. The others are zero matrices. sgn$[\bullet]$ means the signum function.

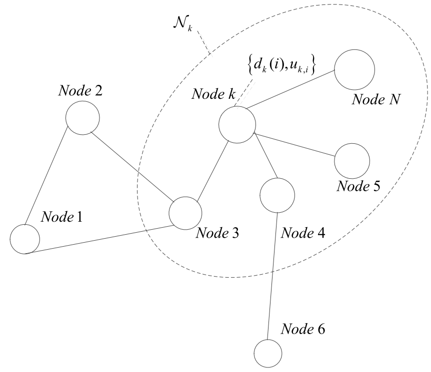


FIGURE 1. Diffusion network with *N* nodes.

II. PROBLEM FORMULATION

A. PROBLEM FORMULATION

We consider the network with N nodes, which is shown in Fig.1. The nodes are distributed over a spatial domain to collectively estimate some interesting parameters which vary over time. At time instant i, node $k$in the network attains the measurement sample $d_k(i)$ and the regression data $\boldsymbol{u}_{k,i}$. The relationship between the $d_k(i)$ and $\boldsymbol{u}_{k,i}$ is given by

$$d_k(i) = \boldsymbol{u}_{k,i}\boldsymbol{w}_i^o + v_k(i) \tag{1}$$

where $v_k(i)$ denotes the measurement noise of the node k and $\boldsymbol{u}_{k,i} = [u_{k,i}, u_{k,i-1}, \ldots, u_{k,i-M+1}]$ has the length of M. The optimal solution, denoted by $\boldsymbol{w}_i^o$, is time-variant according to the standard random walk model [52], which is usually used in analyzing the adaptive filter under non-stationary conditions [42], [44], [48], [53], [55]:

$$\boldsymbol{w}_i^o = \boldsymbol{w}_{i-1}^o + \boldsymbol{q}_i \tag{2}$$

where $\boldsymbol{q}_i$ is the parameter incremental vector.

B. CYCLOSTATIONARY PROCESSES

The input signal $u_{k,i}$ is the cyclostationary random process which satisfies

$$E\left[u_{k,i_1+T}\right] = E\left[u_{k,i_1}\right] \tag{3}$$

$$E\left[u_{k,i_1+T}u_{k,i_2+T}\right] = E\left[u_{k,i_1}u_{k,i_2}\right] \tag{4}$$

for all i_1 and i_2, where T is the period.

To make the analysis tractable, we assume that $\boldsymbol{u}_{k,i}$ is a cyclostationary white Gaussian vector with zero mean and covariance matrix

$R_{u,k}(i)$ which is expressed as

$$\begin{aligned} R_{u,k}(i) &= E\left[\boldsymbol{u}_{k,i}^T \boldsymbol{u}_{k,i}\right] \\ &= diag\left[\sigma_{k,u}^2(i),\ \sigma_{k,u}^2(i-1), \ldots,\ \sigma_{k,u}^2(i-M+1)\right] \end{aligned} \tag{5}$$

where $\sigma_{k,u}^2(i)$ is a periodic function whose period is T. It is assumed that $\sigma_{k,u}^2(i)$ has the following two typical forms in this paper. The first is a sinusoidal power time variation

$$\sigma_{k,u}^2(i) = \beta_k\left[1 + \sin\left(\frac{2\pi i}{T} + \varpi_k\right)\right] \tag{6}$$

where β_k is a positive constant which controls the maximum value of the $\sigma_{k,u}^2(i)$ and ϖ_k is the phase position for the node k. The second is a pulsed power time variation

$$\sigma_{k,u}^2(i) = \begin{cases} \mathcal{H}_k, & for\ qT + \omega_k < i < qT + \alpha_k T + \omega_k \\ \mathcal{L}_k, & for\ qT + \alpha_k T + \omega_k < i < (q+1)T + \omega_k \end{cases}$$
$$0 < \alpha_k < 1, \quad q = 1, 2, \ldots \tag{7}$$

where $\mathcal{H}_k$ and $\mathcal{L}_k$ are the lowest and highest values for $\sigma_{k,u}^2(i)$, respectively, q is a positive integer, and ω_k is the delay time parameter. For tractability, the periods of the cyclostationary signals are classified into large periods, moderate periods and small periods. The periods of the cyclostationary signals can be seen as small periods if $T \ll M$. Conversely, the periods of the cyclostationary signals can be seen as large periods if $T \gg M$. The periods of the cyclostationary signals can be seen as moderate periods if $T \approx M$. Evidently, the large periods are corresponding to the slow variations of input power. On the contrary, the small periods are corresponding to the fast variations of input power. The moderate periods are matched with the moderate variations of input power.

III. A FAMILY OF THE DIFFUSION SIGNED LMS ALGORITHMS

Assuming that the speed for variation of $\boldsymbol{w}_i^o$ is very slow, the centralized LMS algorithm can be used to estimate $\boldsymbol{w}_i^o$ as follows

$$\boldsymbol{w}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu\sum_{k=1}^{N}\boldsymbol{u}_{k,i}^T(d_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{w}_{k,i-1}) \tag{8}$$

where μ is the step size, $\boldsymbol{w}_{k,i}$ denotes the estimate of $\boldsymbol{w}_i^o$ at node k. To lower the computation complexity of this algorithm, the centralized signed regressor algorithm (SRA) can be obtained by introducing the signum function for regressor data, which is multiplication-free,

$$\boldsymbol{w}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu\sum_{k=1}^{N}\text{sgn}\left[\boldsymbol{u}_{k,i}^T\right](d_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{w}_{k,i-1}). \tag{9}$$

Inspired by the merits of distributed algorithm, the DLMS algorithm has been proposed to estimate the $\boldsymbol{w}_i^o$ [7]. Similar to the DLMS algorithm, the proposed adapt-then-combine (ATC) diffusion SRA (DSRA) without information exchange is given by

$$\begin{cases} \boldsymbol{\psi}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu_k\,\text{sgn}\left[\boldsymbol{u}_{k,i}^T\right](d_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{w}_{k,i-1}) \\ \boldsymbol{w}_{k,i} = \sum_{l\in\mathcal{N}_k} a_{l,k}\boldsymbol{\psi}_{l,i}. \end{cases} \tag{10}$$

where μ_k is the step size for node k. $\boldsymbol{\psi}_{k,i}$ is the intermediate estimate of $\boldsymbol{w}_i^o$, $\mathcal{N}_k$ denotes the set of neighbor nodes of node k and $a_{l,k}$ is the combination weight between node l and node k, which satisfies

$$a_{l,k} > 0, \quad \sum_{l=1}^{N} a_{l,k} = 1. \tag{11}$$

Using the same method of deriving DSRA, we can obtain the diffusion sign-sign algorithm (DSSA) by introducing the signum function for estimate errors and regressor data

$$\begin{cases} \boldsymbol{\psi}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu_k\,\text{sgn}\left[\boldsymbol{u}_{k,i}^T\right]\text{sgn}(d_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{w}_{k,i-1}) \\ \boldsymbol{w}_{k,i} = \sum_{l\in\mathcal{N}_k} a_{l,k}\boldsymbol{\psi}_{l,i}. \end{cases} \tag{12}$$

Moreover, the diffusion signed error algorithm (DSEA) which has been proposed in [20] is given by

$$\begin{cases} \boldsymbol{\psi}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu_k\boldsymbol{u}_{k,i}^T\,\text{sgn}(d_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{w}_{k,i-1}) \\ \boldsymbol{w}_{k,i} = \sum_{l\in\mathcal{N}_k} a_{l,k}\boldsymbol{\psi}_{l,i}. \end{cases} \tag{13}$$

Then, the DSRA, DSEA and DSSA constitute a family of the diffusion signed LMS algorithms. In the rest of the paper, the mathematical models of three variants of the diffusion signed LMS algorithm are analyzed. Subsequently, the mean square deviation (MSD) of these three algorithms is derived.

IV. ANALYSIS OF THE DIFFUSION SIGNED ERROR ALGORITHM

A. MEAN WEIGHT BEHAVIOR

Firstly, the intermediate weight error vector at node k is defined as

$$\tilde{\boldsymbol{\psi}}_{k,i} = \boldsymbol{\psi}_{k,i} - \boldsymbol{w}_i^o, \tag{14}$$

and the weight error vector at node k is defined as

$$\tilde{\boldsymbol{w}}_{k,i} = \boldsymbol{w}_{k,i} - \boldsymbol{w}_i^o. \tag{15}$$

Then, the global vectors are expressed as

$$\tilde{\boldsymbol{\psi}}_i = \begin{bmatrix} \tilde{\boldsymbol{\psi}}_{1,i} \\ \tilde{\boldsymbol{\psi}}_{2,i} \\ \vdots \\ \tilde{\boldsymbol{\psi}}_{N,i} \end{bmatrix}, \quad \tilde{\boldsymbol{w}}_i = \begin{bmatrix} \tilde{\boldsymbol{w}}_{1,i} \\ \tilde{\boldsymbol{w}}_{2,i} \\ \vdots \\ \tilde{\boldsymbol{w}}_{N,i} \end{bmatrix}, \quad \boldsymbol{U}_i = \begin{bmatrix} \boldsymbol{u}_{1,i}^T \\ \boldsymbol{u}_{2,i}^T \\ \vdots \\ \boldsymbol{u}_{N,i}^T \end{bmatrix}, \quad \boldsymbol{Q}_i = \begin{bmatrix} \boldsymbol{q}_i \\ \boldsymbol{q}_i \\ \vdots \\ \boldsymbol{q}_i \end{bmatrix}$$

The diagonal matrix $\mathcal{M}$ is also introduced

$$\mathcal{M} = diag\left[\mu_1 I_M, \mu_2 I_M, \ldots, \mu_N I_M\right]. \tag{16}$$

Finally, we introduce the extended weighting matrices

$$\mathcal{P} = P \otimes I_M \tag{17}$$

where P is a matrix whose $\{l,k\}\, th$ entry is $\{a_{l,k}\}$.

In order to make the performance analysis tractable, the following assumptions are stated:

A1: All the input signals $\boldsymbol{u}_{k,i}$ are spatially and temporally independent. The input signals are white Gaussian cyclostationary processes.

A2: All the measurement noises $v_k(i)$ are spatially and temporally independent. In addition, they are Gaussian noise with zero means and variances $\sigma_{k,v}^2$.

A3: The input signals $\boldsymbol{u}_{k,i}$ are independent with all the measurement noises $v_k(i)$.

A4: The parameter incremental vectors $\boldsymbol{q}_i$ are zero mean processes with the covariance matrix $\mathbb{O} = E\left[\boldsymbol{q}_i\boldsymbol{q}_i^T\right] = \sigma_q^2 I_M$.

A5: For node k, weight error quantities $\tilde{\psi}_{k,i}$ and $w_{k,i}$ are statistically independent of $\boldsymbol{u}_{k,i}^T\boldsymbol{u}_{k,i}$. In addition, $\boldsymbol{w}_{k,i}\boldsymbol{w}_{k,i}^T$ is also statistically independent of $\boldsymbol{u}_{k,i}^T\boldsymbol{u}_{k,i}$.

A6: The parameter incremental vectors $\boldsymbol{q}_i$ are statistically independent of $\boldsymbol{u}_{k,i}$.

These assumptions have been widely used in analyzing the distributed algorithms.

In addition, we state the following lemma which will be used in analyzing the performance of DSSA.

Lemma: Assuming that a and b are jointly Gaussian zero mean random variables, we have

$$E\left[\operatorname{sgn}(a)\operatorname{sgn}(b)\right] = \frac{2}{\pi}\sin^{-1}\frac{E\left[ab\right]}{\sqrt{E\left[a^2\right]E\left[b^2\right]}}. \tag{18}$$

The *Lemma* can be easily obtained by the price theorem[1] [50].

[1]Assuming that random variables a and b are jointly Gaussian processes with zero mean, we have

$$E\left[a\operatorname{sgn}(b)\right] = \sqrt{\frac{2}{\pi}}\frac{E\left[ab\right]}{\sqrt{E\left[b^2\right]}}$$

Subtracting $\boldsymbol{w}_i^o$ from both sides of (13) and using equations (14) and (15) result in the following equations

$$\begin{cases} \tilde{\boldsymbol{\psi}}_{k,i} = \tilde{\boldsymbol{w}}_{k,i-1} - \mu_k\boldsymbol{u}_{k,i}^T\operatorname{sgn} \\ \qquad \times\left(\boldsymbol{u}_{k,i}\tilde{\boldsymbol{w}}_{k,i-1} - v_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{q}_i\right) - \boldsymbol{q}_i \\ \tilde{\boldsymbol{w}}_{k,i} = \sum\limits_{l\in\mathcal{N}_k} a_{l,k}\tilde{\boldsymbol{\psi}}_{l,i}. \end{cases} \tag{19}$$

Using the global vectors and matrices which have been mentioned in the foregoing paragraphs yields

$$\begin{cases} \tilde{\boldsymbol{\psi}}_i = \tilde{\boldsymbol{w}}_{i-1} - \mathcal{M}\boldsymbol{\mathcal{A}}_i\boldsymbol{U}_i - \boldsymbol{Q}_i \\ \tilde{\boldsymbol{w}}_i = \mathcal{P}^T\tilde{\boldsymbol{\psi}}_i \end{cases} \tag{20}$$

where

$$\begin{aligned} \boldsymbol{\mathcal{A}}_i = diag\{&\operatorname{sgn}\left(\boldsymbol{u}_{1,i}\tilde{\boldsymbol{w}}_{1,i-1} - v_1(i) - \boldsymbol{u}_{1,i}\boldsymbol{q}_i\right), \\ &\operatorname{sgn}\left(\boldsymbol{u}_{2,i}\tilde{\boldsymbol{w}}_{2,i-1} - v_2(i) - \boldsymbol{u}_{2,i}\boldsymbol{q}_i\right), \ldots, \\ &\operatorname{sgn}\left(\boldsymbol{u}_{N,i}\tilde{\boldsymbol{w}}_{N,i-1} - v_N(i) - \boldsymbol{u}_{N,i}\boldsymbol{q}_i\right)\} \otimes I_M. \end{aligned} \tag{21}$$

Combining two equations in (20) results in

$$\tilde{\boldsymbol{w}}_i = \mathcal{P}^T\tilde{\boldsymbol{w}}_{i-1} - \mathcal{P}^T\mathcal{M}\boldsymbol{\mathcal{A}}_i\boldsymbol{U}_i - \mathcal{P}^T\boldsymbol{Q}_i. \tag{22}$$

Applying the stochastic expectation to (22) yields

$$E\left[\tilde{\boldsymbol{w}}_i\right] = \mathcal{P}^T E\left[\tilde{\boldsymbol{w}}_{i-1}\right] - \mathcal{P}^T\mathcal{M}E\left[\boldsymbol{\mathcal{A}}_i\boldsymbol{U}_i\right] - \mathcal{P}^T E\left[\boldsymbol{Q}_i\right]. \tag{23}$$

Using Assumption A4, we easily obtain

$$E\left[\boldsymbol{Q}_i\right] = 0. \tag{24}$$

Then, let

$$\mathcal{Q}_i = E\left[\boldsymbol{\mathcal{A}}_i\boldsymbol{U}_i\right]. \tag{25}$$

To make analysis clear, we averagely partition $\mathcal{Q}_i$ into N blocks. According to (21), the kth block of $\mathcal{Q}_i$ is given by

$$\left[\mathcal{Q}_i\right]_k = E\left[\operatorname{sgn}\left(\boldsymbol{u}_{k,i}\tilde{\boldsymbol{w}}_{k,i-1} - v_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{q}_i\right)\boldsymbol{u}_{k,i}^T\right]. \tag{26}$$

Then, the jth element of $[\mathcal{Q}_i]_k$ is expressed as

$$\{[\mathcal{Q}_i]_k\}_j = E\left[\operatorname{sgn}\left(\boldsymbol{u}_{k,i}\tilde{\boldsymbol{w}}_{k,i-1} - v_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{q}_i\right)u_{k,i-j+1}\right]. \tag{27}$$

$\{[\mathcal{Q}_i]_k\}_j$ can be calculated by conditioning the expectation on $\tilde{\boldsymbol{w}}_{k,i-1}$ and $\boldsymbol{q}_i$, then averaging over $\tilde{\boldsymbol{w}}_{k,i-1}$ for a given $\boldsymbol{q}_i$ and finally averaging over $\boldsymbol{q}_i$ as follows

$$\begin{aligned} \{[\mathcal{Q}_i]_k\}_j = E\{E\{E[&\operatorname{sgn}\left(\boldsymbol{u}_{k,i}\tilde{\boldsymbol{w}}_{k,i-1} - v_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{q}_i\right) \\ &\times u_{k,i-j+1}|\tilde{\boldsymbol{w}}_{k,i-1},\boldsymbol{q}_i]\,|\boldsymbol{q}_i\}\}. \end{aligned} \tag{28}$$

Using the price theorem, Assumptions A3 and A6 leads to (29), as shown at the top of the next page, where $\left[\tilde{\boldsymbol{w}}_{k,i-1}\right]_j$ and $\left[\boldsymbol{q}_i\right]_j$ denote the jth element of $\tilde{\boldsymbol{w}}_{k,i-1}$ and $\boldsymbol{q}_i$, respectively. $\Gamma_k(i)$ in (29) is expressed as

$$\Gamma_k(i) = \boldsymbol{q}_i^T R_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1} + \tilde{\boldsymbol{w}}_{k,i-1}^T R_{u,k}(i)\boldsymbol{q}_i. \tag{30}$$

$$E\left[\operatorname{sgn}\left(\boldsymbol{u}_{k,i}\tilde{\boldsymbol{w}}_{k,i-1}-v_k(i)-\boldsymbol{u}_{k,i}\boldsymbol{q}_i\right)u_{k,i-j+1}\right]=\sqrt{\frac{2}{\pi}}E\left\{E\left\{\frac{\left[\tilde{\boldsymbol{w}}_{k,i-1}\right]_j\sigma_{k,u}^2(i-j+1)+\left[\boldsymbol{q}_i\right]_j\sigma_{k,u}^2(i-j+1)}{\sqrt{\sigma_{k,v}^2+\left[\tilde{\boldsymbol{w}}_{k,i-1}^TR_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1}\right]+\boldsymbol{q}_i^TR_{u,k}(i)\boldsymbol{q}_i-\Gamma_k(i)}}\middle|\boldsymbol{q}_i\right\}\right\} \tag{29}$$

$$E\left[\operatorname{sgn}\left(\boldsymbol{u}_{k,i}\tilde{\boldsymbol{w}}_{k,i-1}-v_k(i)-\boldsymbol{u}_{k,i}\boldsymbol{q}_i\right)u_{k,i-j+1}\right]=\sqrt{\frac{2}{\pi}}E\left\{\frac{E\left[\tilde{\boldsymbol{w}}_{k,i-1}\right]_j\sigma_{k,u}^2(i-j+1)+\left[\boldsymbol{q}_i\right]_j\sigma_{k,u}^2(i-j+1)}{\sqrt{\sigma_{k,v}^2+E\left[\tilde{\boldsymbol{w}}_{k,i-1}^TR_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1}\right]+\boldsymbol{q}_i^TR_{u,k}(i)\boldsymbol{q}_i-\Gamma'_k(i)}}\right\}. \tag{31}$$

Assuming that the approximation $\tilde{\boldsymbol{w}}_{k,i-1}^TR_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1}\approx E\left[\tilde{\boldsymbol{w}}_{k,i-1}^TR_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1}\right]$ holds, we obtain (31), as shown at the top of this page, where

$$\Gamma'_k(i)=\boldsymbol{q}_i^TR_{u,k}(i)E\left[\tilde{\boldsymbol{w}}_{k,i-1}\right]+E\left[\tilde{\boldsymbol{w}}_{k,i-1}^T\right]R_{u,k}(i)\boldsymbol{q}_i. \tag{32}$$

Assuming that the approximation $\boldsymbol{q}_i^TR_{u,k}(i)\boldsymbol{q}_i\approx E\left[\boldsymbol{q}_i^TR_{u,k}(i)\boldsymbol{q}_i\right]$ holds and using Assumption A4, we have

$$E\left[\operatorname{sgn}\left(\boldsymbol{u}_{k,i}\tilde{\boldsymbol{w}}_{k,i-1}-\boldsymbol{u}_{k,i}\boldsymbol{q}_i-v_k(i)\right)u_{k,i-j+1}\right]=\sqrt{\frac{2}{\pi}}\frac{E\left[\tilde{\boldsymbol{w}}_{k,i-1}\right]_j\sigma_{k,u}^2(i-j+1)}{\sqrt{\sigma_{k,v}^2+E\left[\tilde{\boldsymbol{w}}_{k,i-1}^TR_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1}\right]+E\left[\boldsymbol{q}_i^TR_{u,k}(i)\boldsymbol{q}_i\right]}}. \tag{33}$$

After some algebraic manipulations, we obtain

$$[\mathcal{Q}_i]_k=\sqrt{\frac{2}{\pi}}\frac{R_{u,k}(i)E\left[\tilde{\boldsymbol{w}}_{k,i-1}\right]}{\sqrt{\sigma_{k,v}^2+E\left[\tilde{\boldsymbol{w}}_{k,i-1}^TR_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1}\right]+E\left[\boldsymbol{q}_i^TR_{u,k}(i)\boldsymbol{q}_i\right]}}. \tag{34}$$

Using the Assumptions A1, A5 and A6 yields

$$E\left[\tilde{\boldsymbol{w}}_{k,i-1}^TR_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1}\right]\approx Tr\left[\mathcal{I}_kK(i-1)\mathcal{I}_k^TR_{u,k}(i)\right], \tag{35}$$

$$E\left[\boldsymbol{q}_i^TR_{u,k}(i)\boldsymbol{q}_i\right]\approx Tr\left[\mathbb{O}R_{u,k}(i)\right] \tag{36}$$

where $\boldsymbol{K}(i-1)=E\left[\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^T\right]$.

Combining the results of N nodes and using some algebraic operations, yield

$$\mathcal{Q}_i=\sqrt{\frac{2}{\pi}}\Theta(i)diag\{R_{u,1}(i),R_{u,2}(i),\ldots,R_{u,N}(i)\}E[\tilde{\boldsymbol{w}}_{i-1}] \tag{37}$$

where

$$\Theta(i)=diag\left\{\Xi_1(i),\Xi_2(i),\ldots,\Xi_N(i)\right\}\otimes I_M \tag{38}$$

with

$$\Xi_k(i)=\frac{1}{\sqrt{\sigma_{k,v}^2+Tr\left[\mathcal{I}_k\boldsymbol{K}(i-1)\mathcal{I}_k^TR_{u,k}(i)\right]+Tr\left[\mathbb{O}R_{u,k}(i)\right]}}. \tag{39}$$

Substituting (24) and (37) into (23) results in the recursive expression for $E\left[\tilde{\boldsymbol{w}}_i\right]$ of DSEA

$$E\left[\tilde{\boldsymbol{w}}_i\right]=\mathcal{P}^T\left\{I_{MN}-\mathcal{M}\sqrt{\frac{2}{\pi}}\Theta(i)\times diag\left\{R_{u,1}(i),R_{u,2}(i),\ldots,R_{u,N}(i)\right\}\right\}E\left[\tilde{\boldsymbol{w}}_{i-1}\right] \tag{40}$$

B. MEAN SQUARE WEIGHT BEHAVIOR

Post-multiplying (22) by its transpose, averaging the result and using Assumption A4 lead to the following expression for $\boldsymbol{K}(i)$

$$\boldsymbol{K}(i)=\mathcal{P}^T\boldsymbol{K}(i-1)\mathcal{P}-\mathcal{P}^T\mathcal{M}\mathcal{T}_1\mathcal{P}-\mathcal{P}^T\mathcal{T}_2\mathcal{M}\mathcal{P}+\mathcal{P}^T\mathcal{M}\mathcal{T}_3\mathcal{M}\mathcal{P}+\mathcal{P}^T\mathcal{T}_4\mathcal{P} \tag{41}$$

where

$$\mathcal{T}_1=E\left[\boldsymbol{\mathcal{A}}_i\boldsymbol{U}_i\tilde{\boldsymbol{w}}_{i-1}^T\right], \tag{42}$$

$$\mathcal{T}_2=E\left[\tilde{\boldsymbol{w}}_{i-1}\boldsymbol{U}_i^T\boldsymbol{\mathcal{A}}_i^T\right], \tag{43}$$

$$\mathcal{T}_3=E\left[\boldsymbol{\mathcal{A}}_i\boldsymbol{U}_i\boldsymbol{U}_i^T\boldsymbol{\mathcal{A}}_i^T\right], \tag{44}$$

$$\mathcal{T}_4=E\left[\boldsymbol{Q}_i\boldsymbol{Q}_i^T\right]. \tag{45}$$

The expected values are computed for cyclostationary input as follows.

1) EXPECTATION VALUES $\mathcal{T}_1$ AND $\mathcal{T}_2$

Firstly, we averagely partition $\mathcal{T}_1$ into $N\times N$ blocks where each block is a matrix of size $M\times M$. The $\{s,t\}$ *th* block of $\mathcal{T}_1$ is given by

$$(\mathcal{T}_1)_{s,t}=E\left[\operatorname{sgn}\left(\boldsymbol{u}_{s,i}\tilde{\boldsymbol{w}}_{s,i-1}-v_s(i)-\boldsymbol{u}_{s,i}\boldsymbol{q}_i\right)\boldsymbol{u}_{s,i}^T\tilde{\boldsymbol{w}}_{t,i-1}^T\right]. \tag{46}$$

Obviously, the $\{j,k\}$ *th* element of $(\mathcal{T}_1)_{s,t}$ is expressed as

$$\left[(\mathcal{T}_1)_{s,t}\right]_{j,k}=E\left\{\operatorname{sgn}\left(\boldsymbol{u}_{s,i}\tilde{\boldsymbol{w}}_{s,i-1}-v_s(i)-\boldsymbol{u}_{s,i}\boldsymbol{q}_i\right)u_{s,i-j+1}\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k\right\}. \tag{47}$$

$$\left[(\mathcal{T}_1)_{s,t}\right]_{j,k} = E\left\{E\left\{E\left\{\text{sgn}\left(\boldsymbol{u}_{s,i}\tilde{\boldsymbol{w}}_{s,i-1} - v_s(i) - \boldsymbol{u}_{s,i}\boldsymbol{q}_i\right) u_{s,i-j+1}\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k \middle| \tilde{\boldsymbol{w}}_{s,i-1}, \boldsymbol{q}_i\right\} \middle| \boldsymbol{q}_i\right\}\right\}. \tag{48}$$

$$\left[(\mathcal{T}_1)_{s,t}\right]_{j,k} = \sqrt{\frac{2}{\pi}} E\left\{E\left\{\frac{\left\{\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_j + \left[\boldsymbol{q}_i\right]_j\right\}\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k \sigma_{s,u}^2(i-j+1)}{\sqrt{\sigma_{s,v}^2 + \left[\tilde{\boldsymbol{w}}_{s,i-1}^T R_{u,s}(i)\tilde{\boldsymbol{w}}_{s,i-1}\right] + \boldsymbol{q}_i^T R_{u,s}(i)\boldsymbol{q}_i - \Gamma_s(i)}} \middle| \boldsymbol{q}_i\right\}\right\}. \tag{49}$$

$$\left[(\mathcal{T}_1)_{s,t}\right]_{j,k} = \sqrt{\frac{2}{\pi}} E\left\{\frac{\left\{\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_j + \left[\boldsymbol{q}_i\right]_j\right\}\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k \sigma_{s,u}^2(i-j+1)}{\sqrt{\sigma_{s,v}^2 + E\left[\tilde{\boldsymbol{w}}_{s,i-1}^T R_{u,s}(i)\tilde{\boldsymbol{w}}_{s,i-1}\right] + \boldsymbol{q}_i^T R_{u,s}(i)\boldsymbol{q}_i - \Gamma'_s(i)}}\right\}. \tag{50}$$

Conditioning the expectation on $\tilde{\boldsymbol{w}}_{s,i-1}$ and $\boldsymbol{q}_i$, averaging over $\tilde{\boldsymbol{w}}_{s,i-1}$ for a given $\boldsymbol{q}_i$, and finally averaging over $\boldsymbol{q}_i$, we have (48), as shown at the top of the next page.

Using price theorem and Assumption A1 yields (49), as shown at the top of this page. Assuming that the approximation $\tilde{\boldsymbol{w}}_{s,i-1}^T R_{u,s}(i)\tilde{\boldsymbol{w}}_{s,i-1} \approx E\left[\tilde{\boldsymbol{w}}_{s,i-1}^T R_{u,s}(i)\tilde{\boldsymbol{w}}_{s,i-1}\right]$, (49) implies (50), as shown at the top of this page. Assuming that the approximation $\boldsymbol{q}_i^T R_{u,s}(i)\boldsymbol{q}_i \approx E\left[\boldsymbol{q}_i^T R_{u,s}(i)\boldsymbol{q}_i\right]$ holds and using Assumption **A4**, we have

$$\left[(\mathcal{T}_1)_{s,t}\right]_{j,k} = \sqrt{\frac{2}{\pi}} \frac{E\left\{\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_j\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k\right\}\sigma_{s,u}^2(i-j+1)}{\sqrt{\sigma_{s,v}^2 + E\left[\tilde{\boldsymbol{w}}_{s,i-1}^T R_{u,s}(i)\tilde{\boldsymbol{w}}_{s,i-1}\right] + E\left[\boldsymbol{q}_i^T R_{u,s}(i)\boldsymbol{q}_i\right]}}. \tag{51}$$

Using (39) leads to

$$\left[(\mathcal{T}_1)_{s,t}\right]_{j,k} = \sqrt{\frac{2}{\pi}}\,\Xi_s(i)\sigma_{s,u}^2(i-j+1)E\left[\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_j\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k\right]. \tag{52}$$

Using matrix theory, we have

$$\mathcal{T}_1 = \sqrt{\frac{2}{\pi}} A_i \circ \boldsymbol{K}(i-1) \tag{53}$$

where

$$A_i = \{\Theta(i) col\{B_1, B_2, \ldots, B_N\}\} \otimes \mathbf{1}_{MN} \tag{54}$$

with

$$B_k(i) = col\left\{\sigma_{k,u}^2(i), \sigma_{k,u}^2(i-1), \ldots, \sigma_{k,u}^2(i-M+1)\right\}. \tag{55}$$

According to (42) and (43), we easily obtain

$$\mathcal{T}_2 = \mathcal{T}_1^T \tag{56}$$

2) EXPECTATION VALUE $\mathcal{T}_3$

We averagely partition $\mathcal{T}_3$ into $N \times N$ blocks where each block is a matrix of size $M \times M$. The $\{s, t\}$ *th* block of $\mathcal{T}_3$ is given by

$$(\mathcal{T}_3)_{s,t} = E\Big[\text{sgn}\left(\boldsymbol{u}_{s,i}\tilde{\boldsymbol{w}}_{s,i-1} - v_s(i) - \boldsymbol{u}_{s,i}\boldsymbol{q}_i\right)\boldsymbol{u}_{s,i}^T \bullet \boldsymbol{u}_{t,i}\,\text{sgn}\left(\boldsymbol{u}_{t,i}\tilde{\boldsymbol{w}}_{t,i-1} - v_t(i) - \boldsymbol{u}_{t,i}\boldsymbol{q}_i\right)\Big] \tag{57}$$

If $s = t$, we have

$$(\mathcal{T}_3)_{s,t} = E\left[\boldsymbol{u}_{s,i}^T\boldsymbol{u}_{s,i}\right]. \tag{58}$$

Then, if $s \neq t$, using Assumption **A1** yields

$$(\mathcal{T}_3)_{s,t} = 0. \tag{59}$$

Combining (58) and (59), we have

$$\mathcal{T}_3 = diag\left\{R_{u,1}(i),\ R_{u,2}(i), \ldots,\ R_{u,N}(i)\right\} \tag{60}$$

Using Assumption **A1** results in

$$\mathcal{T}_3 = E\left[U_i U_i^T\right]. \tag{61}$$

3) EXPECTATION VALUE $\mathcal{T}_4$

Using the definition of global vector Q_i, the property of the Kronecker product and the Assumption A4 yields

$$\begin{aligned}\mathcal{T}_4 &= E\left\{\left[\mathbf{1}_N^T \otimes \boldsymbol{q}_i\right]\left[\mathbf{1}_N^T \otimes \boldsymbol{q}_i\right]^T\right\} \\ &= \left[\mathbf{1}'_{N\times N}\right] \otimes E\left[\boldsymbol{q}_i\boldsymbol{q}_i^T\right] \\ &= \sigma_q^2\left(\mathbf{1}'_{N\times N} \otimes I_M\right)\end{aligned} \tag{62}$$

Substituting (53), (56), (61) and (62) into (41), we can easily get the recursive equation for $\boldsymbol{K}(i)$ of the DSEA.

According to [7], the instantaneous *MSD* of the coefficients at node k is defined as

$$MSD_k(i) = E\left[\tilde{\boldsymbol{w}}_{k,i}^T\tilde{\boldsymbol{w}}_{k,i}\right]. \tag{63}$$

Using the algebra theorem yields

$$\begin{aligned}MSD_k(i) &= Tr\left\{E\left[\tilde{\boldsymbol{w}}_{k,i}\tilde{\boldsymbol{w}}_{k,i}^T\right]\right\} \\ &= Tr\left\{\mathcal{I}_k\boldsymbol{K}(i)\mathcal{I}_k^T\right\}.\end{aligned} \tag{64}$$

Similarly, the instantaneous network *MSD* of the coefficients is defined as

$$MSD^{network}(i) = \frac{1}{N}\sum_{k=1}^{N} MSD_k(i) \tag{65}$$

Using the relationship between $MSD_k(i)$ and $MSD^{network}(i)$, we have

$$MSD^{network}(i) = \frac{1}{N} Tr\left\{\boldsymbol{K}(i)\right\} \tag{66}$$

C. BEHAVIOR FOR SLOW VARYING INPUT POWER

According to [45] and [47], the adaptive filter algorithms are usually influenced by cyclostationary input signal with slow variant power. Thereby, it is necessary to investigate this special case in distributed algorithm. For this case, the period of input power is so large that the length of the filter can be omitted. Then, we can get

$$\sigma_{k,u}^2(i) = \sigma_{k,u}^2(i-1) = \ldots = \sigma_{k,u}^2(i-M+1) \quad (67)$$

Substituting (67) into (40) yields

$$E\left[\tilde{\boldsymbol{w}}_i\right] = \mathcal{P}^T\left\{I_{MN} - \sqrt{\frac{2}{\pi}}\mathcal{M}\Theta_{sl}(i)\left(diag\left\{\sigma_{1,u}^2(i),\ \sigma_{2,u}^2(i), \ldots, \sigma_{N,u}^2(i)\right\} \otimes I_M\right)\right\}E\left[\tilde{\boldsymbol{w}}_{i-1}\right] \quad (68)$$

where

$$\Theta_{sl}(i) = diag\left\{\Xi_{sl,1}(i), \Xi_{sl,2}(i), \ldots, \Xi_{sl,N}(i)\right\} \otimes I_M \quad (69)$$

with

$$\Xi_{sl,k}(i) = \frac{1}{\sqrt{\sigma_{k,v}^2 + \sigma_{k,u}^2(i)Tr\left[\mathcal{I}_k \boldsymbol{K}(i-1)\mathcal{I}_k^T\right] + \sigma_{k,u}^2(i)Tr\left[\mathbb{O}\right]}}. \quad (70)$$

Then, plugging (67) into (41) results in

$$\boldsymbol{K}(i) = \mathcal{P}^T\boldsymbol{K}(i-1)\mathcal{P} - \mathcal{P}^T\mathcal{M}\mathcal{T}_{sl,1}\mathcal{P} - \mathcal{P}^T\mathcal{T}_{sl,1}^T\mathcal{M}\mathcal{P} + \mathcal{P}^T\mathcal{M}\mathcal{T}_{sl,3}\mathcal{M}\mathcal{P} + \mathcal{P}^T\mathcal{T}_4\mathcal{P} \quad (71)$$

where

$$\mathcal{T}_{sl,1} = \sqrt{\frac{2}{\pi}} \times \left\{\left[\Theta_{sl}(i)\left\{col\left\{\sigma_{1,u}^2(i), \sigma_{2,u}^2(i), \ldots, \sigma_{N,u}^2(i)\right\} \otimes \mathbf{1}_M^T\right\}\right] \otimes \mathbf{1}_{NM}\right\} \circ \boldsymbol{K}(i-1) \quad (72)$$

$$\mathcal{T}_{sl,3} = diag\left\{\sigma_{1,u}^2(i),\ \sigma_{2,u}^2(i), \ldots,\ \sigma_{N,u}^2(i)\right\} \otimes I_M. \quad (73)$$

Remark: When the variations of $\sigma_{k,u}^2(i)$ for all nodes are synchronous, i.e. $\varpi_1 = \varpi_2 = \ldots = \varpi_N$ and $\omega_1 = \omega_2 = \ldots = \omega_N$ for sinusoidal power time variation and pulsed power time variation, respectively, change trends of the MSD curves for all nodes are same. Moreover, from (71), (72) and (73), we get that the MSD curves for all nodes have ripples. Thereby, the network MSD has ripples. On the contrary, when the variations of power for all nodes are asynchronous, i.e. $\varpi_1 \neq \varpi_2 \neq \ldots \neq \varpi_N$ and $\omega_1 \neq \omega_2 \neq \ldots \neq \omega_N$ for sinusoidal power time variation and pulsed power time variation, the variations of the entries of the $\mathcal{T}_{sl,1}$ and $\mathcal{T}_{sl,3}$ are asynchronous. In the other word, $\sigma_{k_1,u}^2(i)$ and $\sigma_{k_2,u}^2(i)$ for $k_1 \neq k_2$ would not reach the maximum or minimum at the same time. Thus, the ripples of the instantaneous *MSD* of the coefficients at node k_1 and $k_2$would be offset. Then, the curve of network *MSD* would have no ripples or little ripples. In summary, the diffusion algorithms can eliminate the influence of the cyclostationary signal by appropriately adjusting $\{\varpi_k\}$ or $\{\omega_k\}$.

V. ANALYSIS OF THE DIFFUSION SIGNED REGRESSOR ALGORITHM

A. MEAN WEIGHT BEHAVIOR

Plugging (14) and (15) into (10) and using the global vectors in previous section yields

$$\begin{cases} \tilde{\boldsymbol{\psi}}_i = \tilde{\boldsymbol{w}}_{i-1} - \mathcal{M}\left[\mathcal{D}_i\tilde{\boldsymbol{w}}_{i-1} - \mathcal{G}_i - \mathcal{D}_i\boldsymbol{Q}_i\right] - \boldsymbol{Q}_i \\ \tilde{\boldsymbol{w}}_i = \mathcal{P}^T\psi_i \end{cases} \quad (74)$$

where

$$\mathcal{D}_i = diag\left\{\text{sgn}\left[\boldsymbol{u}_{1,i}^T\right]\boldsymbol{u}_{1,i}, \text{sgn}\left[\boldsymbol{u}_{2,i}^T\right]\boldsymbol{u}_{2,i}, \ldots, \text{sgn}\left[\boldsymbol{u}_{N,i}^T\right]\boldsymbol{u}_{N,i}\right\} \quad (75)$$

$$\mathcal{G}_i = col\left\{\text{sgn}\left[\boldsymbol{u}_{1,i}^T\right]v_{1,i}, \text{sgn}\left[\boldsymbol{u}_{2,i}^T\right]v_{2,i}, \ldots, \text{sgn}\left[\boldsymbol{u}_{N,i}^T\right]v_{N,i}\right\} \quad (76)$$

Combining two equations of (74), we have

$$\tilde{w}_i = \mathcal{P}^T\tilde{\boldsymbol{w}}_{i-1} - \mathcal{P}^T\mathcal{M}\left[\mathcal{D}_i\tilde{\boldsymbol{w}}_{i-1} - \mathcal{G}_i - \mathcal{D}_iQ_i\right] - \mathcal{P}^TQ_i \quad (77)$$

Averaging both sides of (77) and using Assumption A4 result in

$$E\left[\tilde{\boldsymbol{w}}_i\right] = \mathcal{P}^TE\left[\tilde{\boldsymbol{w}}_{i-1}\right] - \mathcal{P}^T\mathcal{M}E\left[\mathcal{D}_i\right]E\left[\tilde{\boldsymbol{w}}_{i-1}\right]. \quad (78)$$

Let

$$\mathbb{D}_i = E\left[\mathcal{D}_i\right]. \quad (79)$$

Then, we averagely partition $\mathbb{D}_i$ into $N \times N$ blocks of size $M \times M$. Obviously, the $\{j, j\}$ *th* block of $\mathbb{D}_i$ can be expressed as

$$[\mathbb{D}_i]_{j,j} = E\left[\text{sgn}\left[\boldsymbol{u}_{j,i}^T\right]\boldsymbol{u}_{j,i}\right]. \quad (80)$$

Using Assumption **A1**, we have

$$E\left[\text{sgn}\left(\boldsymbol{u}_{j,i}^T\right)\boldsymbol{u}_{j,i}\right] = diag\left\{E\left[\text{sgn}\left(u_{j,i}\right)u_{j,i}\right], E\left[\text{sgn}\left(u_{j,i-1}\right)u_{j,i-1}\right], \ldots, E\left[\text{sgn}\left(u_{j,i-M+1}\right)u_{j,i-M+1}\right]\right\}. \quad (81)$$

Using the algebraic theory yields

$$E\left[\text{sgn}\left(u_{j,i}\right)u_{j,i}\right] = E\left[\left|u_{j,i}\right|\right] = \sqrt{\frac{2}{\pi}}\sigma_{j,u}(i). \quad (82)$$

Plugging back into (81), we obtain

$$E\left[\text{sgn}\left[u_{j,i}^T\right]u_{j,i}\right] = \sqrt{\frac{2}{\pi}}diag\left\{\sigma_{j,u}(i), \sigma_{j,u}(i-1), \ldots, \sigma_{j,u}(i-M+1)\right\}. \quad (83)$$

Using Assumption A1, $[\mathbb{D}_i]_{j,m} = 0$ for $j \neq m$. Thereby, substituting (83) into (79) leads to the expression for $\mathbb{D}_i$

$$\mathbb{D}_i = diag\left\{\mathcal{Z}_1(i), \mathcal{Z}_2(i), \ldots, \mathcal{Z}_N(i)\right\} \quad (84)$$

where

$$\mathcal{Z}_j(i) = \sqrt{\frac{2}{\pi}} diag\left\{\sigma_{j,u}(i), \sigma_{j,u}(i-1), \ldots, \sigma_{j,u}(i-M+1)\right\}. \tag{85}$$

Plugging (84) into (78) results in the expression for $E\left[\tilde{\boldsymbol{w}}_i\right]$

$$E\left[\tilde{\boldsymbol{w}}_i\right] = \mathcal{P}^T \{I_{MN} - \mathcal{M} diag\{\mathcal{Z}_1(i), \mathcal{Z}_2(i), \ldots, \mathcal{Z}_N(i)\}\} E\left[\tilde{\boldsymbol{w}}_{i-1}\right]. \tag{86}$$

B. MEAN SQUARE WEIGHT BEHAVIOR

Post-multiplying (77) by its transpose, taking expectation of the result and using Assumption A4 yield

$$\begin{aligned}\boldsymbol{K}(i) = {} & \mathcal{P}^T\boldsymbol{K}(i-1)\mathcal{P} + \mathcal{P}^T\mathcal{M}\mathcal{Y}_1\mathcal{M}\mathcal{P} \\ & - \mathcal{P}^T\mathcal{M}\mathcal{Y}_2\mathcal{P} - \mathcal{P}^T\mathcal{Y}_3\mathcal{M}\mathcal{P} \\ & + \mathcal{P}^T\mathcal{M}\mathcal{Y}_4\mathcal{M}\mathcal{P} + \mathcal{P}^T\mathcal{M}\mathcal{Y}_5\mathcal{M}\mathcal{P} \\ & - \mathcal{P}^T\mathcal{T}_4\mathbb{D}_i^T\mathcal{M}\mathcal{P} - \mathcal{P}^T\mathcal{M}\mathbb{D}_i\mathcal{T}_4{}^T\mathcal{P} + \mathcal{P}^T\mathcal{T}_4\mathcal{P}. \end{aligned} \tag{87}$$

where

$$\mathcal{Y}_1 = E\left[\mathcal{D}_i\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^T\mathcal{D}_i^T\right], \tag{88}$$

$$\mathcal{Y}_2 = E\left[\mathcal{D}_i\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^T\right], \tag{89}$$

$$\mathcal{Y}_3 = E\left[\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^T\mathcal{D}_i^T\right], \tag{90}$$

$$\mathcal{Y}_4 = E\left[\mathcal{G}_i\mathcal{G}_i^T\right]. \tag{91}$$

$$\mathcal{Y}_5 = E\left[\mathcal{D}_i\boldsymbol{Q}_i\boldsymbol{Q}_i^T\mathcal{D}_i^T\right]. \tag{92}$$

Then, we compute these expected values for cyclostationary input.

1) EXPECTATION VALUE $\mathcal{Y}_1$

Similar to the method of estimating $\mathcal{T}_1$, $\mathcal{Y}_1$ can be averagely partitioned into $N \times N$ blocks of size $M \times M$. Evidently, the $\{s, t\}$ *th* block of $\mathcal{Y}_1$ is given by

$$[\mathcal{Y}_1]_{s,t} = E\left\{\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\tilde{\boldsymbol{w}}_{s,i-1}\tilde{\boldsymbol{w}}_{t,i-1}^T\left[\operatorname{sgn}\left[\boldsymbol{u}_{t,i}^T\right]\boldsymbol{u}_{t,i}\right]^T\right\}. \tag{93}$$

Then, to calculate the expectation value, we consider the following two cases.

Case 1: $s = t$

Using Assumption A1 and algebraic theory, (93) implies (See Appendix A for detail)

$$\begin{aligned}[\mathcal{Y}_1]_{s,t} = {} & 2\mathcal{Z}_s(i)E\left[\tilde{\boldsymbol{w}}_{s,i-1}\tilde{\boldsymbol{w}}_{s,i-1}^T\right]\mathcal{Z}_s(i) \\ & + Tr\left[E\left[\tilde{\boldsymbol{w}}_{s,i-1}\tilde{\boldsymbol{w}}_{s,i-1}^T\right]\mathcal{Z}_s(i)\right]I_M. \end{aligned} \tag{94}$$

Case 2: $s \neq t$

Using Assumption A1 and A5 results in

$$\begin{aligned}[\mathcal{Y}_1]_{s,t} = {} & E\left[\operatorname{sgn}\left(\boldsymbol{u}_{s,i}^T\right)\boldsymbol{u}_{s,i}\right]E\left[\tilde{\boldsymbol{w}}_{s,i-1}\tilde{\boldsymbol{w}}_{t,i-1}^T\right] \\ & \times E\left[\operatorname{sgn}\left(\boldsymbol{u}_{t,i}^T\right)\boldsymbol{u}_{t,i}\right]^T. \end{aligned} \tag{95}$$

Substituting (83) into (95) yields

$$[\mathcal{Y}_1]_{s,t} = \mathcal{Z}_s(i)\mathcal{I}_s\boldsymbol{K}(i-1)\mathcal{I}_t^T\mathcal{Z}_t(i). \tag{96}$$

Combining the results of two cases, we obtain

$$\mathcal{Y}_1 = \mathbb{D}_i\boldsymbol{K}(i-1)\mathbb{D}_i + \mathbb{D}_i\left[\boldsymbol{K}(i-1)\circ\mathbb{I}\right]\mathbb{D}_i + T'(i) \tag{97}$$

where

$$\mathbb{I} = diag\underbrace{\left\{\mathbf{1}'_{M\times M}, \mathbf{1}'_{M\times M}, .., \mathbf{1}'_{M\times M}\right\}}_{N} \tag{98}$$

and

$$\begin{aligned}T'(i) = diag\{ & Tr\left[\mathcal{I}_1\boldsymbol{K}(i-1)\mathcal{I}_1\mathcal{Z}_1(i)\right], \\ & Tr\left[\mathcal{I}_2\boldsymbol{K}(i-1)\mathcal{I}_2\mathcal{Z}_2(i)\right], \ldots, \\ & Tr\left[\mathcal{I}_N\boldsymbol{K}(i-1)\mathcal{I}_N\mathcal{Z}_N(i)\right]\} \otimes I_M. \end{aligned} \tag{99}$$

2) EXPECTATION VALUES $\mathcal{Y}_2$ AND $\mathcal{Y}_3$

According to Assumption A5, (89) implies

$$\mathcal{Y}_2 = E\left[\mathcal{D}_i\right]E\left[\tilde{\boldsymbol{w}}_{i-1}\tilde{\boldsymbol{w}}_{i-1}^T\right]. \tag{100}$$

Substituting (84) into (100) results in

$$\mathcal{Y}_2 = diag\{\mathcal{Z}_1(i), \mathcal{Z}_2(i), \ldots, \mathcal{Z}_N(i)\}\boldsymbol{K}(i-1). \tag{101}$$

Comparing (89) and (90) yields

$$\mathcal{Y}_3 = \mathcal{Y}_2^T. \tag{102}$$

3) EXPECTATION VALUE $\mathcal{Y}_4$

Substituting (76) into (91) and using Assumption A2 yields

$$\mathcal{Y}_4 = diag\left\{\sigma_{1,v}^2, \sigma_{2,v}^2, \sigma_{3,v}^2, \ldots, \sigma_{N,v}^2\right\} \otimes I_M. \tag{103}$$

4) EXPECTATION VALUE $\mathcal{Y}_5$

$\mathcal{Y}_5$ can be averagely partitioned into $N \times N$ blocks which are square matrices with the length of M. The $\{s, t\}$ *th* block of $\mathcal{Y}_5$ is expressed as

$$[\mathcal{Y}_5]_{s,t} = E\left[\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\boldsymbol{q}_i\boldsymbol{q}_i^T\boldsymbol{u}_{t,i}^T\operatorname{sgn}\left[\boldsymbol{u}_{t,i}\right]\right]. \tag{104}$$

Then, to calculate the expectation value, we consider the following two cases.

Case 1: $s = t$

In this case, (104) implies

$$[\mathcal{Y}_5]_{s,t} = E\left[\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\boldsymbol{q}_i\boldsymbol{q}_i^T\boldsymbol{u}_{s,i}^T\operatorname{sgn}\left[\boldsymbol{u}_{s,i}\right]\right]. \tag{105}$$

$[\mathcal{Y}_5]_{s,t}$ can be calculated by conditioning the expectation on $u_{s,i}$, then averaging over q_i as follows

$$\begin{aligned}[\mathcal{Y}_5]_{s,t} & = E\left[E\left[\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\boldsymbol{q}_i\boldsymbol{q}_i^T\boldsymbol{u}_{s,i}^T\operatorname{sgn}\left[\boldsymbol{u}_{s,i}\right]|\boldsymbol{u}_{s,i}\right]\right] \\ & = E\left[\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}E\left[\boldsymbol{q}_i\boldsymbol{q}_i^T\right]\boldsymbol{u}_{s,i}^T\operatorname{sgn}\left[\boldsymbol{u}_{s,i}\right]\right]. \end{aligned} \tag{106}$$

Using assumption A3, (106) implies

$$[\mathcal{Y}_5]_{s,t} = \sigma_q^2 E\left[\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\boldsymbol{u}_{s,i}^T\operatorname{sgn}\left[\boldsymbol{u}_{s,i}\right]\right]. \tag{107}$$

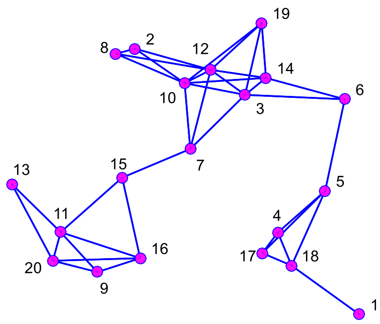


FIGURE 2. The topology of network.

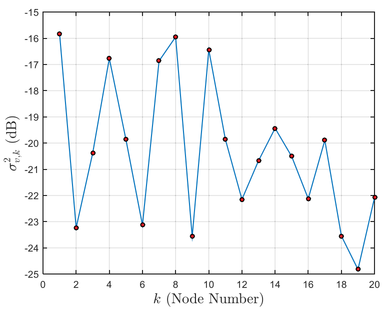
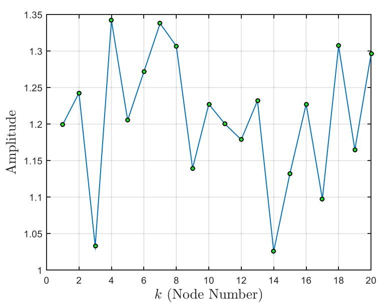
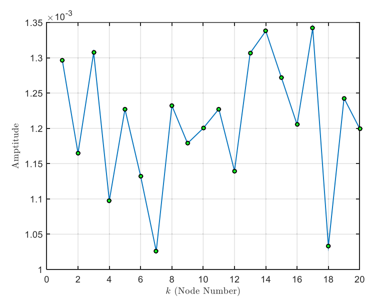


FIGURE 3. Noise variances $\sigma^2_{k,v}$ for 20 nodes.



(a)



(b)

FIGURE 4. The amplitudes of square roots of $\mathcal{H}_k$ and $\mathcal{L}_k$, (a) $\mathcal{H}_k$, (b) $\mathcal{L}_k$.

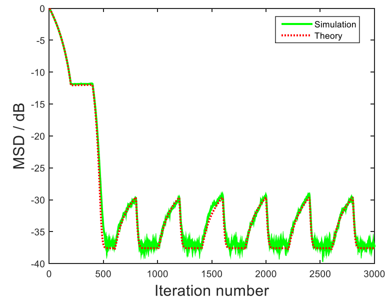
Using (83) and (A.11), we have (See Appendix B for detail)

$$[\mathcal{Y}_5]_{s,t} = \sigma_q^2 \frac{4}{\pi} R_{u,s}(i) + \sigma_q^2 Tr\left[R_{u,s}(i)\right] I_M \tag{108}$$

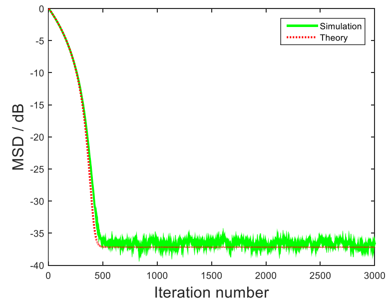
Case 2: $s \neq t$

Using Assumption A1 and A5 results in

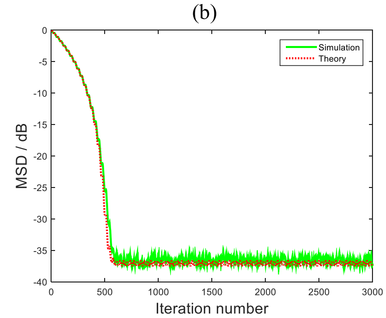
$$[\mathcal{Y}_5]_{s,t} = \sigma_q^2 E\left[\operatorname{sgn}\left(\boldsymbol{u}_{s,i}^T\right)\boldsymbol{u}_{s,i}\right] E\left[\operatorname{sgn}\left(\boldsymbol{u}_{t,i}^T\right)\boldsymbol{u}_{t,i}\right]^T. \tag{109}$$



(a)



(b)



(c)

FIGURE 5. The network MSD curves of DSEA for synchronous pulsed variations, (a) slow variation: $T = 400$ (b) fast variation: $T = 2$, (c) moderate variation: $T = 32$.

Substituting (83) and (85) into (109) yields

$$[\mathcal{Y}_5]_{s,t} = \sigma_q^2 \mathcal{Z}_s(i)\mathcal{Z}_t(i). \tag{110}$$

Combining the results of two cases, we obtain

$$\mathcal{Y}_5 = \sigma_q^2 \mathbb{D}_i \Upsilon \mathbb{D}_i + \sigma_q^2 \frac{4}{\pi} diag\left\{R_{u,1}(i),\ R_{u,2}(i), \ldots,\ R_{u,N}(i)\right\} + \sigma_q^2 S'(i) \tag{111}$$

where

$$\Upsilon = \begin{bmatrix} 0 & I & \cdots & I \\ I & 0 & \cdots & I \\ \vdots & \vdots & \ddots & \vdots \\ I & I & \cdots & 0 \end{bmatrix} \tag{112}$$

and

$$S'(i) = diag\left\{Tr\left[R_{u,1}(i)\right], Tr\left[R_{u,2}(i)\right], \ldots, Tr\left[R_{u,N}(i)\right]\right\} \otimes I_M \tag{113}$$

Combining (87), (97), (101), (102), (103) and (111), we can get the recursion of $\boldsymbol{K}(i)$ for the DSRA.

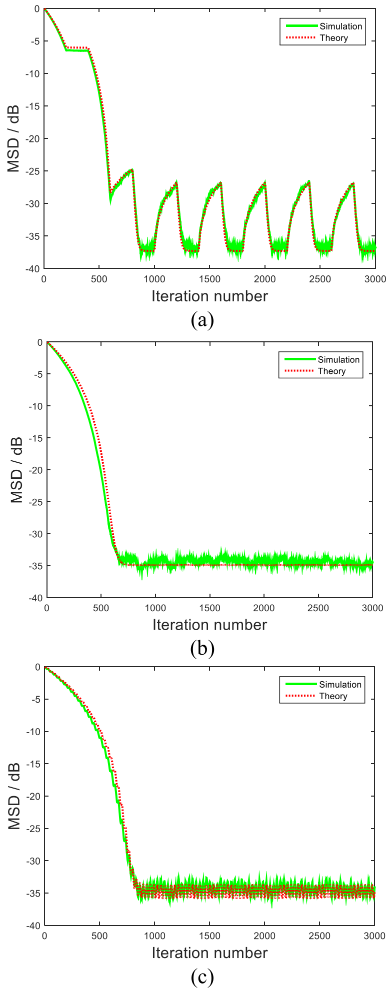


FIGURE 6. The network MSD curves of DSRA for synchronous pulsed variations, (a) slow variation: $T = 400$, (b) fast variation: $T = 2$, (c) moderate variation: $T = 32$.

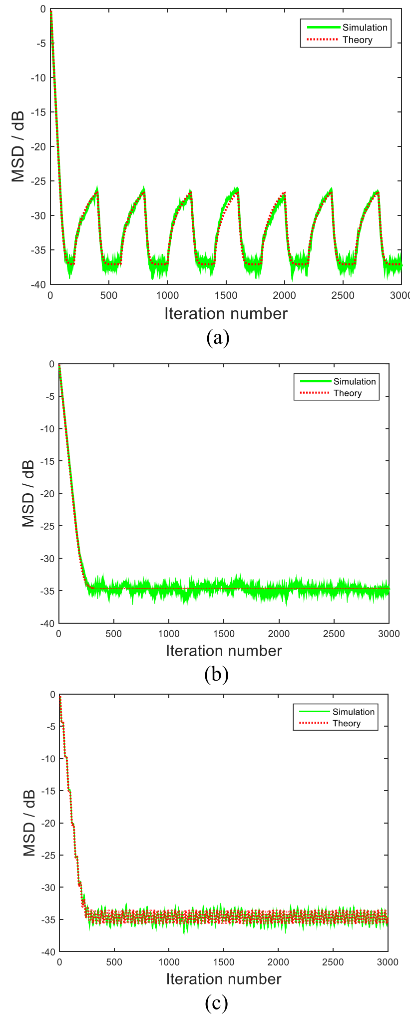


FIGURE 7. The network MSD curves of DSSA for synchronous pulsed variations, (a) slow variation: $T = 400$, (b) fast variation: $T = 2$, (c) moderate variation: $T = 32$.

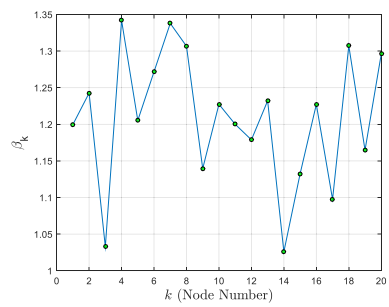


FIGURE 8. The magnitudes of square roots of β_k.

C. BEHAVIOR FOR SLOW VARYING INPUT POWER

According to [45], when cyclostationary input signals have slow varying input power, equation (67) holds. Then, substituting (67) into (86) yields

$$E\left[\tilde{\boldsymbol{w}}_i\right] = \mathcal{P}^T\left[I_{MN} - \sqrt{\frac{2}{\pi}}\mathcal{M}\left\{diag\left\{\sigma_{1,u}(i), \sigma_{2,u}(i), \ldots, \sigma_{N,u}(i)\right\} \otimes I_M\right\}\right]E\left[\tilde{\boldsymbol{w}}_{i-1}\right]. \tag{114}$$

Plugging (67) into (87) yields

$$\begin{aligned}\boldsymbol{K}(i) = {} & \mathcal{P}^T\boldsymbol{K}(i-1)\mathcal{P} + \mathcal{P}^T\mathcal{M}\mathcal{Y}_{sl,1}\mathcal{M}\mathcal{P} \\ & - \mathcal{P}^T\mathcal{M}\mathcal{Y}_{sl,2}\mathcal{P} - \mathcal{P}^T\mathcal{Y}_{sl,2}^T\mathcal{M}\mathcal{P} \\ & + \mathcal{P}^T\mathcal{M}\mathcal{Y}_4\mathcal{M}\mathcal{P} + \mathcal{P}^T\mathcal{M}\mathcal{Y}_{sl,5}\mathcal{M}\mathcal{P} \\ & - \mathcal{P}^T\mathcal{T}_4\mathcal{Y}_{sl,2}^T\mathcal{M}\mathcal{P} - \mathcal{P}^T\mathcal{M}\mathcal{Y}_{sl,2}\mathcal{T}_4{}^T\mathcal{P} + \mathcal{P}^T\mathcal{T}_4\mathcal{P}\end{aligned} \tag{115}$$

where

$$\mathcal{Y}_{sl,2} = \sqrt{\frac{2}{\pi}}\boldsymbol{h}(i)\boldsymbol{K}(i-1) \tag{116}$$

$$\begin{aligned}\mathcal{Y}_{sl,1} = {} & \frac{2}{\pi}\boldsymbol{h}(i)\boldsymbol{K}(i-1)\boldsymbol{h}(i) + \frac{2}{\pi}\boldsymbol{h}(i)\left[\boldsymbol{K}(i-1) \circ \mathbb{I}_{M^2N^2}\right]\boldsymbol{h}(i) \\ & + diag\left\{\mathbb{H}(i)\right\} \otimes \boldsymbol{I}_M,\end{aligned} \tag{117}$$

$$\mathcal{Y}_{sl,5} = \sigma_q^2\boldsymbol{h}(i)\Upsilon\boldsymbol{h}(i) + \sigma_q^2\frac{4}{\pi}\boldsymbol{h}^2(i) + \sigma_q^2 M\boldsymbol{h}^2(i) \tag{118}$$

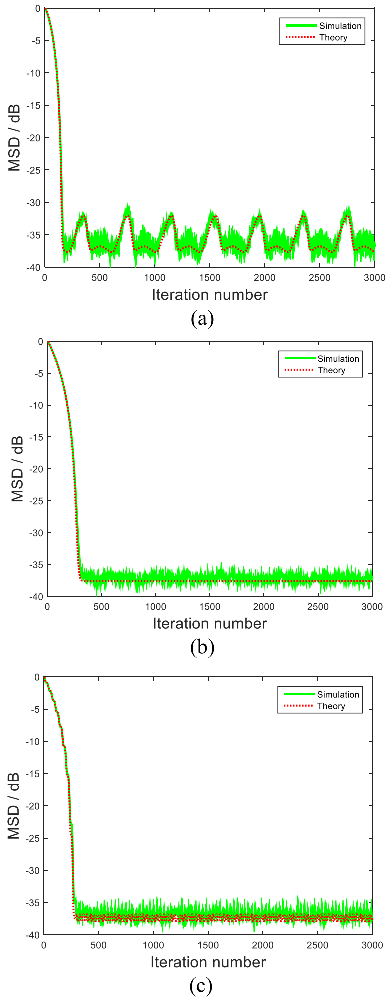


FIGURE 9. **The network MSD curves of DSEA for synchronous sinusoidal variation, (a) slow variation: $T = 400$ (b) fast variation: $T = 2$, (c) moderate variation: $T = 32$.**

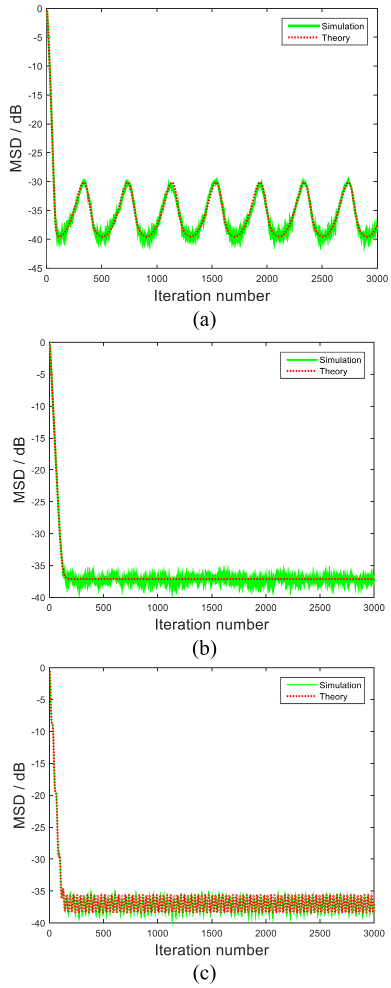


FIGURE 10. **The network MSD curves of DSRA for synchronous sinusoidal variations, (a) slow variation: $T = 400$ (b) fast variation: $T = 2$, (c) moderate variation: $T = 32$.**

with

$$\boldsymbol{h}(i) = diag\left\{\sigma_{1,u}(i), \sigma_{2,u}(i), \ldots, \sigma_{N,u}(i)\right\} \otimes I_M \quad (119)$$

and

$$\mathbb{H}(i) = \sqrt{\frac{2}{\pi}}\left\{\sigma_{1,u}(i)Tr\left[\mathcal{I}_1\boldsymbol{K}(i-1)\mathcal{I}_1\right], \sigma_{2,u}(i)Tr\left[\mathcal{I}_2\boldsymbol{K}(i-1)\mathcal{I}_2\right], \ldots, \sigma_{N,u}(i)Tr\left[\mathcal{I}_N\boldsymbol{K}(i-1)\mathcal{I}_N\right]\right\}. \quad (120)$$

Then, combining (64), (66) and (115), we can get the $MSD_k(i)$ and $MSD^{network}(i)$. From (115), we can also obtain that the diffusion algorithms can eliminate the influence of the cyclostationary signal.

VI. ANALYSIS OF THE DIFFUSION SIGN-SIGN ERROR ALGORITHM

Subtracting w_i^o from both sides of (12) yields

$$\begin{cases} \tilde{\boldsymbol{\psi}}_{k,i} = \tilde{\boldsymbol{w}}_{k,i-1} - \mu_k \operatorname{sgn}\left(\boldsymbol{u}_{k,i}^T\right) \operatorname{sgn}\left(\boldsymbol{u}_{k,i}\tilde{\boldsymbol{w}}_{k,i-1} - v_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{q}_i\right) - \boldsymbol{q}_i \\ \tilde{\boldsymbol{w}}_{k,i} = \sum_{l \in \mathcal{N}_k} a_{l,k}\tilde{\boldsymbol{\psi}}_{l,i}. \end{cases} \quad (121)$$

Using the definitions of global quantities, we obtain

$$\begin{cases} \tilde{\boldsymbol{\psi}}_i = \tilde{\boldsymbol{w}}_{i-1} - \mathcal{M}\boldsymbol{\mathcal{A}}_i\boldsymbol{B}_i - \boldsymbol{Q}_i \\ \tilde{\boldsymbol{w}}_i = \mathcal{P}^T\tilde{\boldsymbol{\psi}}_i \end{cases} \quad (122)$$

where

$$\boldsymbol{B}_i = \operatorname{sgn}\left[\boldsymbol{U}_i\right]. \quad (123)$$

Equivalently,

$$\tilde{\boldsymbol{w}}_i = \mathcal{P}^T\tilde{\boldsymbol{w}}_{i-1} - \mathcal{P}^T\mathcal{M}\boldsymbol{\mathcal{A}}_i\boldsymbol{B}_i - \mathcal{P}^T\boldsymbol{Q}_i. \quad (124)$$

A. MEAN WEIGHT BEHAVIOR

Taking expectation of both sides of (124) and using Assumption A4 yield

$$E\left[\tilde{\boldsymbol{w}}_i\right] = \mathcal{P}^T E\left[\tilde{\boldsymbol{w}}_{i-1}\right] - \mathcal{P}^T\mathcal{M}E\left[\boldsymbol{\mathcal{A}}_i\boldsymbol{B}_i\right]. \quad (125)$$

Let

$$\mathbb{Q}_i = E\left[\boldsymbol{\mathcal{A}}_i\boldsymbol{B}_i\right]. \quad (126)$$

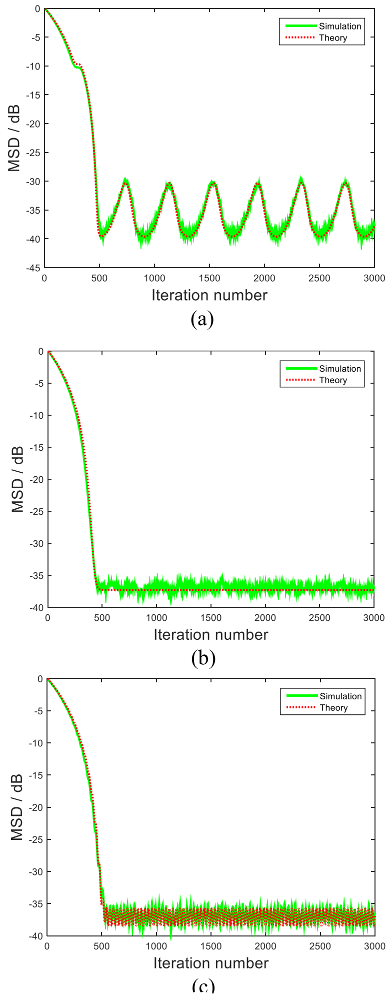


FIGURE 11. The network MSD curves of DSRA for synchronous sinusoidal variations, (a) slow variation: $T = 400$ (b) fast variation: $T = 2$, (c) moderate variation: $T = 32$.

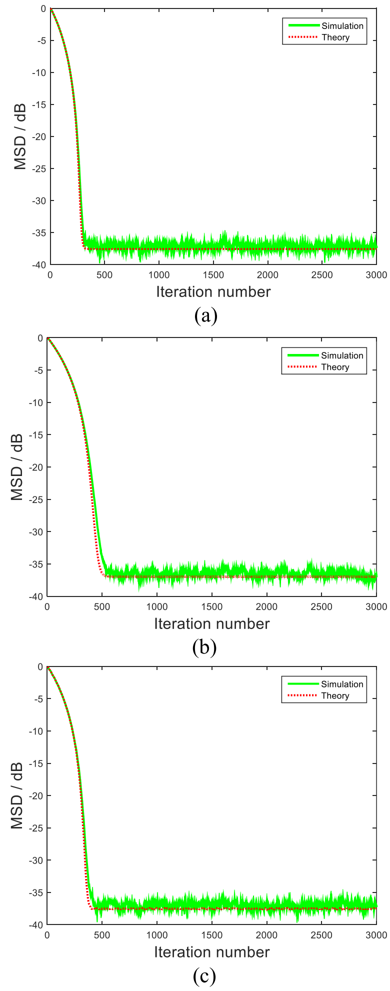


FIGURE 12. The network MSD curves of DSEA for asynchronous pulsed variations, (a) slow variation: $T = 400$ (b) fast variation: $T = 2$, (c) moderate variation: $T = 32$.

Then, we averagely partition $\mathbb{Q}_i$ into N blocks. Subsequently, the kth block of $\mathbb{Q}_i$ is given by

$$[\mathbb{Q}_i]_k = E\left[\text{sgn}\left(\boldsymbol{u}_{k,i}\tilde{\boldsymbol{w}}_{k,i-1} - v_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{q}_i\right)\text{sgn}\left(\boldsymbol{u}_{k,i}^T\right)\right]. \tag{127}$$

Evidently, the jth block element of $[\mathbb{Q}_i]_k$ is computed as

$$\{[\mathbb{Q}_i]_k\}_j = E\left[\text{sgn}\left(\boldsymbol{u}_{k,i}\tilde{\boldsymbol{w}}_{k,i-1} - v_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{q}_i\right)\text{sgn}\left(u_{k,i-j+1}\right)\right]. \tag{128}$$

In the sequel, conditioning the expectation on $\tilde{\boldsymbol{w}}_{k,i-1}$ and $\boldsymbol{q}_i$, then averaging over $\tilde{\boldsymbol{w}}_{k,i-1}$ for a given $\boldsymbol{q}_i$ and finally averaging over $\boldsymbol{q}_i$, we can get (129), as shown at the top of the next page. Using Assumption A3 and *Lemma*, and letting $a = \boldsymbol{u}_{k,i}\tilde{\boldsymbol{w}}_{k,i-1} - v_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{q}_i$ and $b = u_{k,i-j+1}$ in (18), equation (129) becomes (130), as shown at the top of the next page.

Since the terms $\tilde{\boldsymbol{w}}_{k,i-1}^T R_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1}$ and $\boldsymbol{q}_i^T R_{u,k}(i)\boldsymbol{q}_i$ are both the sum of the M terms which are on the order of the square of the numerator, the magnitude of the argument of the function $\sin^{-1}$ is on the order of $\sqrt{1/2M}$. Thereby, it is much less than 1. Assuming that the approximation $\sin^{-1}x \approx x$ holds, (130) implies (131), as shown at the top of the next page. In the sequel, using Assumption A4 and the approximations $\tilde{\boldsymbol{w}}_{k,i-1}^T R_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1} \approx E\left[\tilde{\boldsymbol{w}}_{k,i-1}^T R_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1}\right]$ and $\boldsymbol{q}_i^T R_{u,k}(i)\boldsymbol{q}_i \approx E\left[\boldsymbol{q}_i^T R_{u,k}(i)\boldsymbol{q}_i\right]$, (131) implies that

$$\{[\mathbb{Q}_i]_k\}_j = \frac{2}{\pi}\frac{\left[\tilde{\boldsymbol{w}}_{k,i-1}\right]_j \sigma_{k,u}(i-j+1)}{\sqrt{\sigma_{k,v}^2 + Tr\left[\mathcal{I}_k \boldsymbol{K}(i-1)\mathcal{I}_k^T R_{u,k}(i)\right] + Tr\left[\mathbb{O}R_{u,k}(i)\right]}}. \tag{132}$$

Using some algebraic manipulations leads to the expression for $[\mathbb{Q}_i]_k$

$$[\mathbb{Q}_i]_k = \frac{2}{\pi}\frac{\mathbb{A}_k(i)E\left[\tilde{\boldsymbol{w}}_{k,i-1}\right]}{\sqrt{\sigma_{k,v}^2 + Tr\left[\mathcal{I}_k \boldsymbol{K}(i-1)\mathcal{I}_k^T R_{u,k}(i)\right] + Tr\left[\mathbb{O}R_{u,k}(i)\right]}} \tag{133}$$

$$\{[\mathbb{Q}_i]_k\}_j = E\left\{E\left\{E\left[\operatorname{sgn}\left(\boldsymbol{u}_{k,i}\tilde{\boldsymbol{w}}_{k,i-1} - v_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{q}_i\right)\operatorname{sgn}\left(u_{k,i-j+1}\right)\middle|\tilde{\boldsymbol{w}}_{k,i-1},\boldsymbol{q}_i\right]\middle|\boldsymbol{q}_i\right\}\right\}. \tag{129}$$

$$\begin{aligned}\{[\mathbb{Q}_i]_k\}_j = &\ \frac{2}{\pi}E\left\{E\left\{\sin^{-1}\frac{[\tilde{\boldsymbol{w}}_{k,i-1}]_j E\left[u^2_{k,i-j+1}\right] + [\boldsymbol{q}_i]_j E\left[u^2_{k,i-j+1}\right]}{\sqrt{E\left[u^2_{k,i-j+1}\right]}\sqrt{\sigma^2_{k,v} + \tilde{\boldsymbol{w}}^T_{k,i-1}R_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1} + \boldsymbol{q}_i^T R_{u,k}(i)\boldsymbol{q}_i - \Gamma_k(i)}}\middle|\boldsymbol{q}_i\right\}\right\} \\ = &\ \frac{2}{\pi}E\left\{E\left\{\sin^{-1}\frac{[\tilde{\boldsymbol{w}}_{k,i-1}]_j\sigma^2_{k,u}(i-j+1) + [\boldsymbol{q}_i]_j\sigma^2_{k,u}(i-j+1)}{\sigma_{k,u}(i-j+1)\sqrt{\sigma^2_{k,v} + \tilde{\boldsymbol{w}}^T_{k,i-1}R_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1} + \boldsymbol{q}_i^T R_{u,k}(i)\boldsymbol{q}_i - \Gamma_k(i)}}\middle|\boldsymbol{q}_i\right\}\right\} \\ = &\ \frac{2}{\pi}E\left\{E\left\{\sin^{-1}\frac{[\tilde{\boldsymbol{w}}_{k,i-1}]_j\sigma_{k,u}(i-j+1) + [\boldsymbol{q}_i]_j\sigma_{k,u}(i-j+1)}{\sqrt{\sigma^2_{k,v} + \tilde{\boldsymbol{w}}^T_{k,i-1}R_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1} + \boldsymbol{q}_i^T R_{u,k}(i)\boldsymbol{q}_i - \Gamma_k(i)}}\middle|\boldsymbol{q}_i\right\}\right\}.\end{aligned} \tag{130}$$

$$\{[\mathbb{Q}_i]_k\}_j = \frac{2}{\pi}E\left[E\left[\frac{[\tilde{\boldsymbol{w}}_{k,i-1}]_j\sigma_{k,u}(i-j+1) + [\boldsymbol{q}_i]_j\sigma_{k,u}(i-j+1)}{\sqrt{\sigma^2_{k,v} + \tilde{\boldsymbol{w}}^T_{k,i-1}R_{u,k}(i)\tilde{\boldsymbol{w}}_{k,i-1} + \boldsymbol{q}_i^T R_{u,k}(i)\boldsymbol{q}_i - \Gamma_k(i)}}\middle|\boldsymbol{q}_i\right]\right]. \tag{131}$$

where

$$\mathbb{A}_k(i) = diag\left\{\sigma_{k,u}(i),\ \sigma_{k,u}(i-1),\ldots,\ \sigma_{k,u}(i-M+1)\right\}. \tag{134}$$

After some algebraic manipulations, (133) implies that

$$\mathbb{Q}_i = \frac{2}{\pi}\Theta(i)diag\left\{\mathbb{A}_1(i),\ \mathbb{A}_2(i),\ldots,\ \mathbb{A}_N(i)\right\}E\left[\tilde{\boldsymbol{w}}_{i-1}\right]. \tag{135}$$

Substituting (135) into (125) leads to the expression for $E\left[\tilde{\boldsymbol{w}}_{i-1}\right]$ of DSSA

$$\begin{aligned}E\left[\tilde{\boldsymbol{w}}_i\right] &= \mathcal{P}^T\left\{I_{MN} - \mathcal{M}\frac{2}{\pi}\Theta(i)diag\left\{\mathbb{A}_1(i),\ \mathbb{A}_2(i),\ldots,\ \mathbb{A}_N(i)\right\}\right\} \\ &\quad\times E\left[\tilde{\boldsymbol{w}}_{i-1}\right].\end{aligned} \tag{136}$$

B. MEAN SQUARE WEIGHT BEHAVIOR

Post-multiplying (124) by its transpose, taking expectation of the result and using Assumption A4, we have

$$\begin{aligned}\boldsymbol{K}(i) = &\ \mathcal{P}^T\boldsymbol{K}(i-1)\mathcal{P} - \mathcal{P}^T\mathcal{M}\mathbb{T}_1\mathcal{P} \\ &- \mathcal{P}^T\mathbb{T}_2\mathcal{M}\mathcal{P} + \mathcal{P}^T\mathcal{M}\mathbb{T}_3\mathcal{M}\mathcal{P} + \mathcal{P}^T\boldsymbol{\mathcal{T}}_4\mathcal{P}.\end{aligned} \tag{137}$$

where

$$\mathbb{T}_1 = E\left[\boldsymbol{\mathcal{A}}_i\boldsymbol{B}_i\tilde{\boldsymbol{w}}^T_{i-1}\right], \tag{138}$$

$$\mathbb{T}_2 = E\left[\tilde{\boldsymbol{w}}_{i-1}\boldsymbol{B}_i^T\boldsymbol{\mathcal{A}}_i^T\right], \tag{139}$$

$$\mathbb{T}_3 = E\left[\boldsymbol{\mathcal{A}}_i\boldsymbol{B}_i\boldsymbol{B}_i^T\boldsymbol{\mathcal{A}}_i^T\right]. \tag{140}$$

To analyze the mean square weight behavior, we calculate the expected values as follows

1) EXPECTATION VALUES $\mathbb{T}_1$ AND $\mathbb{T}_2$

$\mathbb{T}_1$ can be averagely partitioned into $N \times N$ blocks which are square matrices with the length of M. The $\{s,t\}$ *th* block of $\mathbb{T}_1$ is expressed as

$$(\mathbb{T}_1)_{s,t} = E\left\{\operatorname{sgn}\left(\boldsymbol{u}_{s,i}\tilde{\boldsymbol{w}}_{s,i-1} - v_s(i) - \boldsymbol{u}_{s,i}\boldsymbol{q}_i\right)\operatorname{sgn}\left(\boldsymbol{u}^T_{s,i}\right)\tilde{\boldsymbol{w}}^T_{t,i-1}\right\}. \tag{141}$$

The $\{j,k\}$ *th* entry of $(\mathbb{T}_1)_{s,t}$ is given by

$$\begin{aligned}&[(\mathbb{T}_1)_{s,t}]_{j,k} \\ &= E\left\{\operatorname{sgn}\left(\boldsymbol{u}_{s,i}\tilde{\boldsymbol{w}}_{s,i-1} - \boldsymbol{u}_{s,i}\boldsymbol{q}_i - v_s(i)\right)\operatorname{sgn}\left[u_{s,i-j+1}\right]\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k\right\}.\end{aligned} \tag{142}$$

Conditioning the expectation on $\tilde{\boldsymbol{w}}_{k,i-1}$ and $\boldsymbol{q}_i$, then averaging over $\tilde{\boldsymbol{w}}_{k,i-1}$ for a given $\boldsymbol{q}_i$ and finally averaging over $\boldsymbol{q}_i$ yield (143), as shown at the top of the next page. Using the *Lemma*, Assumption A4 and the approximations $\sin^{-1}x \approx x$, $\tilde{\boldsymbol{w}}^T_{s,i-1}R_{u,s}(i)\tilde{\boldsymbol{w}}_{s,i-1} \approx E\left[\tilde{\boldsymbol{w}}^T_{s,i-1}R_{u,s}(i)\tilde{\boldsymbol{w}}_{s,i-1}\right]$ and $\boldsymbol{q}_i^T R_{u,s}(i)\boldsymbol{q}_i \approx E\left[\boldsymbol{q}_i^T R_{u,s}(i)\boldsymbol{q}_i\right]$, we have (144), as shown at the top of the next page, where the first equality in (144) follows from the *Lemma,* the fourth equality in (144) follows from the approximation $\sin^{-1}x \approx x$ and the fifth equality in (144) follows from Assumption A4 and the approximations $\tilde{\boldsymbol{w}}^T_{s,i-1}R_{u,s}(i)\tilde{\boldsymbol{w}}_{s,i-1} \approx E\left[\tilde{\boldsymbol{w}}^T_{s,i-1}R_{u,s}(i)\tilde{\boldsymbol{w}}_{s,i-1}\right]$ and $\boldsymbol{q}_i^T R_{u,s}(i)\boldsymbol{q}_i \approx E\left[\boldsymbol{q}_i^T R_{u,s}(i)\boldsymbol{q}_i\right]$.

Then, combining the results of all nodes over network results in

$$\mathbb{T}_1 = \frac{2}{\pi}\boldsymbol{\mathcal{X}}_i \circ \boldsymbol{K}(i-1) \tag{145}$$

where

$$\boldsymbol{\mathcal{X}}_i = \Theta(i)col\left\{\mathcal{B}_1,\mathcal{B}_2,\ldots,\mathcal{B}_N\right\} \otimes \mathbf{1}_{MN} \tag{146}$$

with

$$\mathcal{B}_k(i) = col\left\{\sigma_{k,u}(i),\sigma_{k,u}(i-1),\ldots,\sigma_{k,u}(i-M+1)\right\} \tag{147}$$

Comparing equations (138) and (139), we easily have

$$\mathbb{T}_2 = \mathbb{T}_1^T. \tag{148}$$

$$\left[(\mathbb{T}_1)_{s,t}\right]_{j,k} = E\left\{E\left\{E\left\{\begin{array}{c}\mathrm{sgn}\left(u_{s,i}\tilde{\boldsymbol{w}}_{s,i-1} - v_s(i) - u_{s,i}\boldsymbol{q}_i\right)\\ \times\,\mathrm{sgn}\left[u_{s,i-j+1}\right]\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k\end{array}\middle|\tilde{\boldsymbol{w}}_{t,i-1}^T,\ \boldsymbol{q}_i\right\}\middle|\boldsymbol{q}_i\right\}\right\}. \tag{143}$$

$$\begin{aligned}\left[(\mathbb{T}_1)_{s,t}\right]_{j,k} &= \frac{2}{\pi}E\left\{E\left\{\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k \sin^{-1}\frac{\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_j E\left[u_{s,i-j+1}^2\right] + \left[\boldsymbol{q}_i\right]_j E\left[u_{s,i-j+1}^2\right]}{\sqrt{E\left[u_{s,i-j+1}^2\right]}\sqrt{\sigma_{s,v}^2 + \tilde{\boldsymbol{w}}_{s,i-1}^T R_{u,s}\tilde{\boldsymbol{w}}_{s,i-1} + \boldsymbol{q}_i^T R_{u,s}\boldsymbol{q}_i - \Gamma_s(i)}}\middle|\boldsymbol{q}_i\right\}\right\}\\
&= \frac{2}{\pi}E\left\{E\left\{\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k \sin^{-1}\frac{\sigma_{s,u}^2(i-j+1)\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_j + \sigma_{s,u}^2(i-j+1)\left[\boldsymbol{q}_i\right]_j}{\sigma_{s,u}(i-j+1)\sqrt{\sigma_{s,v}^2 + \tilde{\boldsymbol{w}}_{s,i-1}^T R_{u,s}\tilde{\boldsymbol{w}}_{s,i-1} + \boldsymbol{q}_i^T R_{u,s}\boldsymbol{q}_i - \Gamma_s(i)}}\middle|\boldsymbol{q}_i\right\}\right\}\\
&= \frac{2}{\pi}E\left\{E\left\{\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k \sin^{-1}\frac{\sigma_{s,u}(i-j+1)\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_j + \sigma_{s,u}(i-j+1)\left[\boldsymbol{q}_i\right]_j}{\sqrt{\sigma_{s,v}^2 + \tilde{\boldsymbol{w}}_{s,i-1}^T R_{u,s}\tilde{\boldsymbol{w}}_{s,i-1} + \boldsymbol{q}_i^T R_{u,s}\boldsymbol{q}_i - \Gamma_s(i)}}\middle|\boldsymbol{q}_i\right\}\right\}\\
&= \frac{2}{\pi}E\left\{E\left\{\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k \frac{\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_j\sigma_{s,u}(i-j+1) + \left[\boldsymbol{q}_i\right]_j\sigma_{s,u}(i-j+1)}{\sqrt{\sigma_{s,v}^2 + \tilde{\boldsymbol{w}}_{s,i-1}^T R_{u,s}\tilde{\boldsymbol{w}}_{s,i-1} + \boldsymbol{q}_i^T R_{u,s}\boldsymbol{q}_i - \Gamma_s(i)}}\middle|\boldsymbol{q}_i\right\}\right\}\\
&= \frac{2}{\pi}\frac{E\left\{\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_j\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k\right\}\sigma_{s,u}(i-j+1)}{\sqrt{\sigma_{s,v}^2 + E\left[\tilde{\boldsymbol{w}}_{s,i-1}^T R_{u,s}\tilde{\boldsymbol{w}}_{s,i-1}\right] + E\left[\boldsymbol{q}_i^T R_{u,s}\boldsymbol{q}_i\right]}} = \frac{2}{\pi}\frac{E\left\{\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_j\left[\tilde{\boldsymbol{w}}_{t,i-1}\right]_k\right\}\sigma_{s,u}(i-j+1)}{\sqrt{\sigma_{s,v}^2 + Tr(\mathcal{I}_k\boldsymbol{K}(i-1)\mathcal{I}_k^T R_{u,k}(i)) + Tr(\mathbb{O}R_{u,s})}}.\end{aligned} \tag{144}$$

2) EXPECTATION VALUE $\mathbb{T}_3$

Using the algebraic theory, (140) implies

$$\mathbb{T}_3 = E\left[\boldsymbol{B}_i\boldsymbol{B}_i^T\right]. \tag{149}$$

Using Assumption **A1**, we obtain

$$\mathbb{T}_3 = I_{MN}. \tag{150}$$

Substituting (145), (148) and (150) into (137), the recursion of $\boldsymbol{K}(i)$ of the DSSA can be obtained.

C. BEHAVIOR FOR SLOW VARYING INPUT POWER

When cyclostationary input signals have slow varying input power, equation (67) holds. Then, Substituting (67) into (136) leads to the recursion of $E\left[\tilde{\boldsymbol{w}}_i\right]$

$$E\left[\tilde{\boldsymbol{w}}_i\right] = \mathcal{P}^T\left(I - \frac{2}{\pi}\mathcal{M}\Theta_{sl}(i)\left(diag\{\sigma_{1,u}(i),\ \sigma_{2,u}(i),\ldots,\ \sigma_{N,u}(i)\}\otimes I_M\right)\right)E\left[\tilde{\boldsymbol{w}}_{i-1}\right]. \tag{151}$$

In the sequel, plugging (67) into (137) results in the expression for $\boldsymbol{K}(i)$

$$\begin{aligned}\boldsymbol{K}(i) &= \mathcal{P}^T\boldsymbol{K}(i-1)\mathcal{P} - \mathcal{P}^T\mathcal{M}\mathbb{T}_{sl,1}\mathcal{P}\\ &\quad - \mathcal{P}^T\mathbb{T}_{sl,1}^T\mathcal{M}\mathcal{P} + \mathcal{P}^T\mathcal{M}\mathbb{T}_{sl,3}\mathcal{M}\mathcal{P} + \mathcal{P}^T\boldsymbol{\mathcal{T}}_4\mathcal{P}\end{aligned} \tag{152}$$

where

$$\mathbb{T}_{sl,1} = \frac{2}{\pi}\left\{\left[\Theta_{sl}(i)\left\{col\{\sigma_{1,u}(i),\ \sigma_{2,u}(i),\ldots,\ \sigma_{N,u}(i)\}\otimes 1_M^T\right\}\right]\otimes \mathbf{1}_{NM}\right\}\circ\boldsymbol{K}(i-1), \tag{153}$$

$$\mathbb{T}_{sl,3} = diag\left\{\sigma_{1,u}(i),\ \sigma_{2,u}(i),\ldots,\ \sigma_{N,u}(i)\right\}\otimes I_M. \tag{154}$$

Then, combining (64), (66) and (152), the $MSD_k(i)$ and $MSD^{network}(i)$ can be obtained. From (152), we can also draw the conclusion that the diffusion algorithms can suppress the influence of the cyclostationary signal.

VII. SIMULATION

In this section, the computer simulations are carried out to demonstrate the performance of proposed algorithms and verify the theoretical results. It is assumed that the length of the unknown vector $\boldsymbol{w}_i^o$ is 5, which is the same as the length of the input vector. The parameter vector $\boldsymbol{w}_i^o$ to be estimated varies with time according to the first order random walk model with incremental variance $\sigma_q^2 = 10^{-6}$. In this paper, the simulated network topology contains 20 nodes, which is depicted in Fig. 2. The background noise of each node is the white Gaussian process with power $\sigma_{v,k}^2$ which is shown in Fig. 3. The initialization of $\boldsymbol{w}_i^o$ is generated by Gaussian process. The combination matrix P is computed according to uniform rule. The results of simulations are obtained by averaging over 100 independent runs.

A. RESULTS FOR INPUT POWER WITH SYNCHRONOUS PULSED VARIATIONS

In this subsection, the input signals have synchronous pulsed power time variations. The magnitudes of square roots of $\mathcal{H}_k$ and $\mathcal{L}_k$ are plotted in Fig.4 (a) and Fig.4 (b), respectively. $\{\omega_k\}$ are set to 0. Fig. 5 compares Monte Carlo (MC) simulations and theory of DSEA for slow, fast and moderate cases. The step sizes $\{\mu_k\}$ are all set to 0.004. It is found that there is

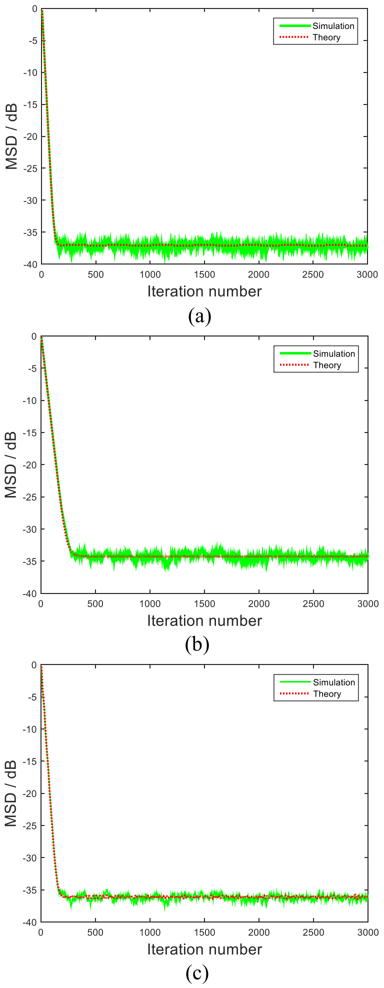


FIGURE 13. **The network MSD curves of DSRA for asynchronous pulsed variations, (a) slow variation: $T = 400$ (b) fast variation: $T = 2$, (c) moderate variation: $T = 32$.**

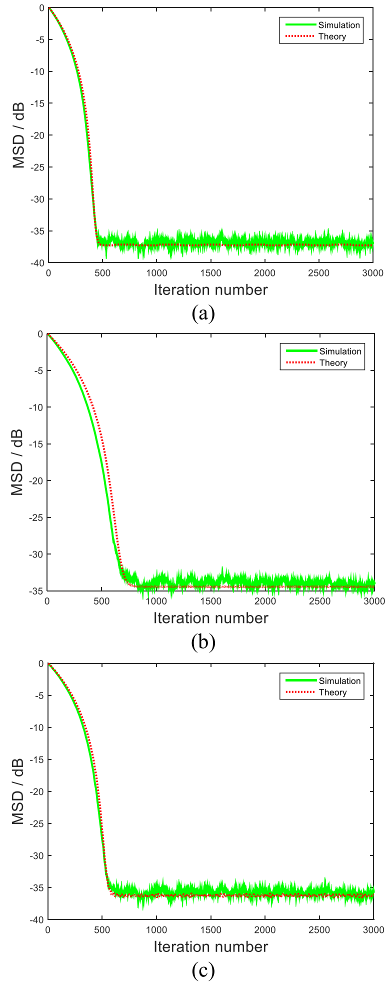


FIGURE 14. **The network MSD curves of DSSA for asynchronous pulsed variations, (a) slow variation: $T = 400$ (b) fast variation: $T = 2$, (c) moderate variation: $T = 32$.**

an excellent agreement between the simulation results and the theoretical results. There are several ripples in the curves for slow pulsed input power variations. However, the curves for fast and moderate pulsed input power variations have no ripples. Fig. 6 depicts the theoretical and simulated curves of the transient network MSD for DSRA. The step sizes $\{\mu_k\}$ are all set to 0.004. We observed that the curves of simulation are in line with their corresponding theoretical results. We also found that the curves have ripples for slow pulsed input power variations and have no ripples or little ripples for moderate and fast pulsed input power variations. Fig. 7 depicts the theoretical and simulated curves of the transient network MSD of DSSA for slow, fast and moderate sinusoidal input power variations. The step sizes $\{\mu_k\}$ are all set to 0.004. The results in Fig.7 illustrate an excellent match between simulations and theory.

B. RESULTS FOR INPUT POWER WITH SYNCHRONOUS SINUSOIDAL VARIATIONS

In this subsection, the input signals have the synchronous sinusoidal power variations. The magnitudes of square roots of β_k are plotted in Fig.8. $\{\varpi_k\}$ are set to 0. Fig. 9 gives the transient MSDs of simulations and theory of DSEA for slow, fast and moderate cases. The step sizes $\{\mu_k\}$ are all set to 0.004. It is seen that there is a good match between MC simulation and the theory. Fig. 10 illustrates the theoretical and simulated curves of the transient network MSD for DSRA. The step sizes $\{\mu_k\}$ are all set to 0.004. We can see the results of simulation match that of theory. We also present the results of simulation and theory of DSSA for slow, fast and moderate sinusoidal input power variations in Fig. 11. The step sizes $\{\mu_k\}$ are all set to 0.004. Once again, one notes a very good match between the results obtained through MC simulation and theory. From Figs. 9-11, we can also observe that there are ripples for the slow case and there are no or little ripples for moderate and fast cases.

C. RESULTS FOR INPUT POWER WITH ASYNCHRONOUS PULSED VARIATIONS

We study the performance of three signed variants of the DLMS algorithm over the network when the input signals

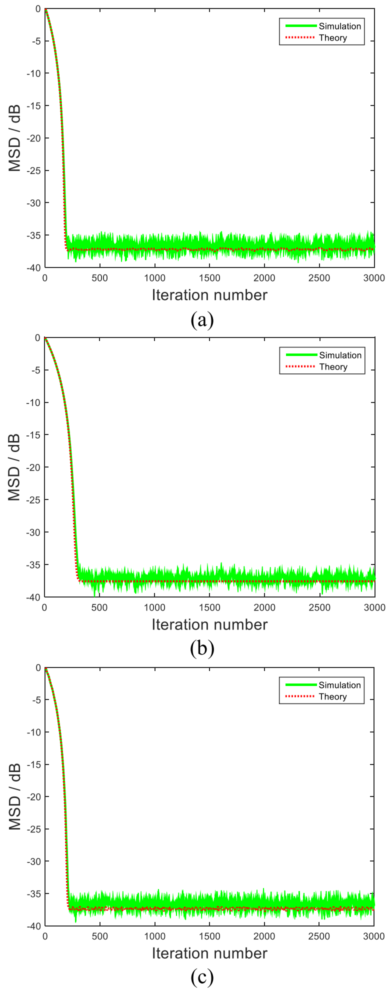


FIGURE 15. **The network MSD curves of DSEA for asynchronous sinusoidal variations, (a) slow variation: $T = 400$ (b) fast variation: $T = 2$, (c) moderate variation: $T = 32$.**

have asynchronous pulsed power variation. The magnitudes of square roots of $\mathcal{H}_k$ and $\mathcal{L}_k$ are plotted in Fig.4 (a) and Fig.4 (b), respectively. We set the delay time parameters ω_k for each node as follows

$$\omega_k = \omega_{k-1} + \tau \tag{155}$$

where τ is constant depending on the period of variations of the input power. In this simulation, τ is set to 1, 2 and 10 for slow, fast and moderate cases, respectively. The step sizes $\{\mu_k\}$ are all set to 0.004 for DSEA, DSRA and DSSA. Fig. 12 shows the results of simulation and theory of DSEA for slow, fast and moderate pulsed input power variations. Fig. 13 provides the theoretical and simulated curves of the transient network MSD for DSRA. Fig. 14 illustrates the theoretical and simulated curves of the transient network MSD for DSRA. From these figures, we can see that all these simulation results are in line with their corresponding theoretical results. In addition, we observed that the curves have no ripples regardless of the speed of the input power variations, which corresponds to the discussion of analytic theory.

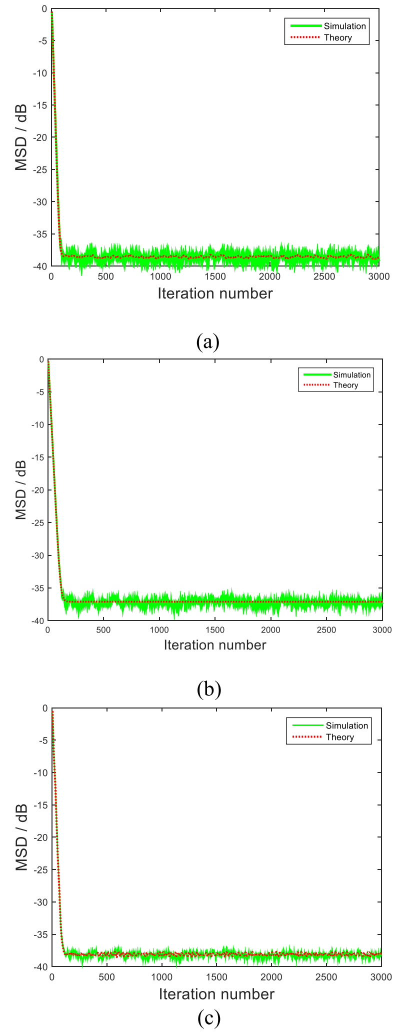


FIGURE 16. **The network MSD curves of DSRA for asynchronous sinusoidal variations, (a) slow variation: $T = 400$ (b) fast variation: $T = 2$, (c) moderate variation: $T = 32$.**

D. RESULTS FOR INPUT POWER WITH ASYNCHRONOUS SINUSOIDAL VARIATIONS

The performance of three signed variants of the DLMS algorithm over the network is considered for the input signals which have asynchronous sinusoidal power variation. The magnitudes of square roots of β_k are plotted in Fig.8. We set the delay time parameters ϖ_k for each node as follows

$$\varpi_k = \varpi_{k-1} + \tau' \tag{156}$$

where τ' is constant depending on the period of variations of the input power. In this simulation, τ' is set to 1, 2 and 10 for slow, fast and moderate cases, respectively. The step sizes $\{\mu_k\}$ are all set to 0.004 for DSEA, DSRA and DSSA. Fig. 15 compares simulations with theory of DSEA for slow, fast and moderate sinusoidal input power variations. Fig. 16 illustrates the theoretical and simulated curves of the transient network MSD for DSRA. Fig. 17 provides the theoretical and simulated curves of the transient network MSD for DSRA. It is also found that there an excellent agreement between the

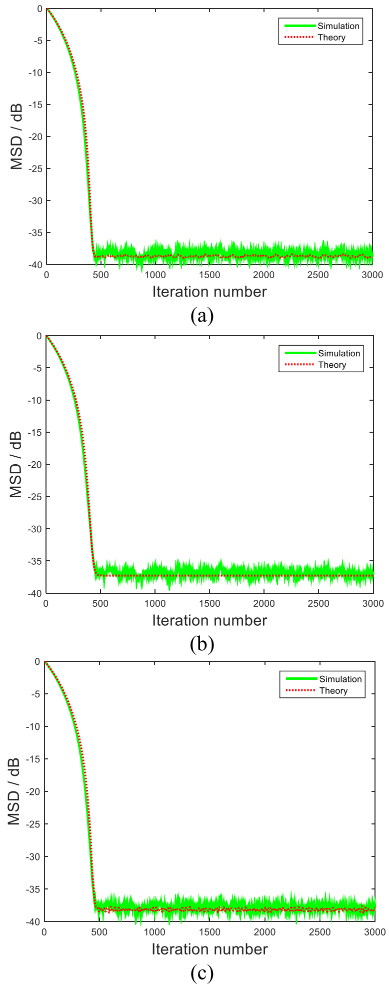


FIGURE 17. The network MSD curves of DSSA for asynchronous sinusoidal variations, (a) slow variation: $T = 400$ (b) fast variation: $T = 2$, (c) moderate variation: $T = 32$.

simulation results and the theoretical results. Also, there are no ripples in curves regardless of the speed of the input power variation.

VIII. CONCLUSION

As one of the signed variants of the DLMS algorithm, the DSEA has been presented in the previous reference. This paper presents two other signed variants of DLMS algorithm, i.e. DSRA and DSSA. Then, we study the performance of three signed variants of DLMS algorithm while the cyclostationary white Gaussian signals are selected as input signals and the unknown system is generated by the random walk model. Specifically, the mean weight behavior and mean square weight behavior of the algorithms are derived. Finally, the simulation experiments are carried out to verify the correctness of the analytic model. It is clearly seen that there is an excellent agreement between theory and the simulation results. It is also found that fast and moderate variation for both synchronous and the asynchronous case would cause very little magnitude ripples. However, when the input signals have synchronous slow variation power, the curves would have ripple. It is worth noted that when the input signals have asynchronous variation power, the curves don't have ripples or have little magnitude ripples. Thus, we obtain that the distributed algorithm can suppress the effects of the cyclostationary input.

APPENDIX A
PROOF OF (94)

For (93), conditioning the expectation on $\tilde{\boldsymbol{w}}_{s,i-1}$ and then averaging over $\tilde{\boldsymbol{w}}_{s,i-1}$ yield

$$[\mathcal{Y}_1]_{s,s} = E\left\{E\left\{\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\tilde{\boldsymbol{w}}_{s,i-1}\tilde{\boldsymbol{w}}_{s,i-1}^T \times\left[\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\right]^T\left|\tilde{\boldsymbol{w}}_{s,i-1}\right.\right\}\right\} \tag{A.1}$$

Let

$$C(i) = \operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\tilde{\boldsymbol{w}}_{s,i-1}\tilde{\boldsymbol{w}}_{s,i-1}^T\left[\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\right]^T \tag{A.2}$$

The $\{j, k\}$ *th* element of $C(i)$ is expressed as

$$c_{j,k}(i) = \sum_{l=1}^{M}\sum_{p=1}^{M}\left\{\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\right\}_{j,l}\left\{\tilde{\boldsymbol{w}}_{s,i-1}\tilde{\boldsymbol{w}}_{s,i-1}^T\right\}_{l,p} \times\left\{\left[\operatorname{sgn}\left[u_{s,i}^T\right]u_{s,i}\right]^T\right\}_{p,k} \tag{A.3}$$

where

$$\left\{\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\right\}_{j,l} = \operatorname{sgn}\left[u_{s,i-j+1}\right]u_{s,i-l+1} \tag{A.4}$$

$$\left\{\tilde{\boldsymbol{w}}_{s,i-1}\tilde{\boldsymbol{w}}_{s,i-1}^T\right\}_{l,p} = \left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_l\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_p \tag{A.5}$$

$$\left\{\left[\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\right]^T\right\}_{p,k} = u_{s,i-p+1}\operatorname{sgn}\left[u_{s,i-k+1}\right] \tag{A.6}$$

Using Assumption A1, we have

$$E\left\{c_{j,k}(i)\left|\tilde{\boldsymbol{w}}_{s,i-1}\right.\right\} = \sum_{l=1}^{M}\sum_{p=1}^{M}E\left\{\operatorname{sgn}\left[u_{s,i-j+1}\right]u_{s,i-l+1}u_{s,i-p+1}\operatorname{sgn}\left[u_{s,i-k+1}\right]\right\} \times\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_l\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_p \tag{A.7}$$

Using the moment factoring [51] results in

$$\begin{aligned} &E\left\{\operatorname{sgn}\left[u_{s,i-j+1}\right]u_{s,i-l+1}u_{s,i-p+1}\operatorname{sgn}\left[u_{s,i-k+1}\right]\right\} \\ &= E\left\{\operatorname{sgn}\left[u_{s,i-j+1}\right]u_{s,i-l+1}\right\}E\left\{u_{s,i-p+1}\operatorname{sgn}\left[u_{s,i-k+1}\right]\right\} \\ &\quad + E\left\{\operatorname{sgn}\left[u_{s,i-j+1}\right]\operatorname{sgn}\left[u_{s,i-k+1}\right]\right\}E\left\{u_{s,i-l+1}u_{s,i-p+1}\right\} \\ &\quad + E\left\{u_{s,i-l+1}\operatorname{sgn}\left[u_{s,i-k+1}\right]\right\}E\left\{\operatorname{sgn}\left[u_{s,i-j+1}\right]u_{s,i-p+1}\right\} \end{aligned} \tag{A.8}$$

Plugging back into (A.7) yields (A.9), as shown at the top of the next page. Equivalently,

$$\begin{aligned} E\left\{C(i)\left|\tilde{\boldsymbol{w}}_{s,i-1}\right.\right\} &= 2E\left\{\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\right\}\tilde{\boldsymbol{w}}_{s,i-1}\tilde{\boldsymbol{w}}_{s,i-1}^T E\left[\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\right]^T \\ &\quad + \tilde{\boldsymbol{w}}_{s,i-1}^T E\left[\boldsymbol{u}_{s,i}^T\boldsymbol{u}_{s,i}\right]\tilde{\boldsymbol{w}}_{s,i-1}E\left\{\operatorname{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\operatorname{sgn}\left[\boldsymbol{u}_{s,i}\right]\right\} \end{aligned} \tag{A.10}$$

$$\begin{aligned} E\left\{c_{j,k}(i)\,|\tilde{\boldsymbol{w}}_{s,i-1}\right\} =& \sum_{l=1}^{M}\sum_{p=1}^{M} E\left\{\mathrm{sgn}\left[u_{s,i-j+1}\right]u_{s,i-l+1}\right\}\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_l\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_p E\left\{u_{s,i-p+1}\,\mathrm{sgn}\left[u_{s,i-k+1}\right]\right\} \\ &+\sum_{p=1}^{M}\sum_{l=1}^{M} E\left\{\mathrm{sgn}\left[u_{s,i-j+1}\right]u_{s,i-p+1}\right\}\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_p\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_l E\left\{u_{s,i-l+1}\,\mathrm{sgn}\left[u_{s,i-k+1}\right]\right\} \\ &+E\left\{\mathrm{sgn}\left[u_{s,i-j+1}\right]\mathrm{sgn}\left[u_{s,i-k+1}\right]\right\}E\left\{\sum_{l=1}^{M}u_{s,i-l+1}\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_l\sum_{p=1}^{M}u_{s,i-p+1}\left[\tilde{\boldsymbol{w}}_{s,i-1}\right]_p\right\} \end{aligned} \quad (A.9)$$

For white Gaussian input,

$$E\left\{\mathrm{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\mathrm{sgn}\left[\boldsymbol{u}_{s,i}\right]\right\} = I_M \quad (A.11)$$

Then, after some algebraic manipulations, we can get

$$\begin{aligned}[\mathcal{Y}_1]_{s,s} =& 2\mathcal{Z}_s(i)E\left[\tilde{\boldsymbol{w}}_{s,i-1}\tilde{\boldsymbol{w}}_{s,i-1}^T\right]\mathcal{Z}_s(i) \\ &+Tr\left[E\left[\tilde{\boldsymbol{w}}_{s,i-1}\tilde{\boldsymbol{w}}_{s,i-1}^T\right]\mathcal{Z}_s(i)\right]I_M \end{aligned} \quad (A.12)$$

APPENDIX B
PROOF OF (108)

Obviously, according to (107), when $s = t$, the $\{j, k\}$ *th* element of $[\mathcal{Y}_5]_{s,t}$ is expressed as

$$\begin{aligned} &\left[[\mathcal{Y}_5]_{s,t}\right]_{j,k} \\ &= \sigma_q^2\sum_{l=1}^{M}\left\{\mathrm{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\right\}_{j,l}\left\{\left[\mathrm{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\right]^T\right\}_{l,k} \\ &= \sigma_q^2\sum_{l=1}^{M}E\left\{\mathrm{sgn}\left[u_{s,i-j+1}\right]u_{s,i-l+1}u_{s,i-l+1}\,\mathrm{sgn}\left[u_{s,i-k+1}\right]\right\} \end{aligned} \quad (B.1)$$

Using the moment factoring [51] results in

$$\begin{aligned} &E\left\{\mathrm{sgn}\left[u_{s,i-j+1}\right]u_{s,i-l+1}u_{s,i-l+1}\,\mathrm{sgn}\left[u_{s,i-k+1}\right]\right\} \\ &= E\left\{\mathrm{sgn}\left[u_{s,i-j+1}\right]u_{s,i-l+1}\right\}E\left\{u_{s,i-l+1}\,\mathrm{sgn}\left[u_{s,i-k+1}\right]\right\} \\ &\quad+E\left\{\mathrm{sgn}\left[u_{s,i-j+1}\right]\mathrm{sgn}\left[u_{s,i-k+1}\right]\right\}E\left\{u_{s,i-l+1}u_{s,i-l+1}\right\} \\ &\quad+E\left\{u_{s,i-l+1}\,\mathrm{sgn}\left[u_{s,i-k+1}\right]\right\}E\left\{\mathrm{sgn}\left[u_{s,i-j+1}\right]u_{s,i-l+1}\right\} \end{aligned} \quad (B.2)$$

Plugging back into (B.1) yields

$$\begin{aligned} \left[[\mathcal{Y}_5]_{s,t}\right]_{j,k} =& \sigma_q^2\sum_{l=1}^{M}E\left\{\mathrm{sgn}\left[u_{s,i-j+1}\right]u_{s,i-l+1}\right\} \\ &\times E\left\{u_{s,i-l+1}\,\mathrm{sgn}\left[u_{s,i-k+1}\right]\right\} \\ &+\sigma_q^2\sum_{l=1}^{M}E\left\{\mathrm{sgn}\left[u_{s,i-j+1}\right]u_{s,i-l+1}\right\} \\ &\times E\left\{u_{s,i-l+1}\,\mathrm{sgn}\left[u_{s,i-k+1}\right]\right\} \\ &+\sigma_q^2E\left\{\mathrm{sgn}\left[u_{s,i-j+1}\right]\mathrm{sgn}\left[u_{s,i-k+1}\right]\right\} \\ &\times E\left\{\sum_{l=1}^{M}u_{s,i-l+1}u_{s,i-l+1}\right\} \end{aligned} \quad (B.3)$$

Equivalently,

$$\begin{aligned}[\mathcal{Y}_5]_{s,t} =& 2\sigma_q^2E\left\{\mathrm{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\right\}E\left[\mathrm{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\boldsymbol{u}_{s,i}\right]^T \\ &+\sigma_q^2Tr\left\{E\left[\boldsymbol{u}_{s,i}^T\boldsymbol{u}_{s,i}\right]\right\}E\left\{\mathrm{sgn}\left[\boldsymbol{u}_{s,i}^T\right]\mathrm{sgn}\left[\boldsymbol{u}_{s,i}\right]\right\} \end{aligned} \quad (B.4)$$

Substituting (83) and (A.11) into (B.4), we have

$$[\mathcal{Y}_5]_{s,t} = \sigma_q^2\frac{4}{\pi}R_{u,s}(i) + \sigma_q^2Tr\left[R_{u,s}(i)\right]I_M \quad (B.5)$$

REFERENCES

[1] A. H. Sayed, "Diffusion adaptation over networks," in *Academic Press Library in Signal Processing*, R. Chellappa and S. Theodoridis, Eds. Amsterdam, The Netherlands: Elsevier, 2014, pp. 322–454.

[2] A. H. Sayed, "Adaptive networks," *Proc. IEEE*, vol. 102, no. 4, pp. 460–497, Apr. 2014.

[3] A. H. Sayed and C. G. Lopes, "Distributed recursive least-squares strategies over adaptive networks," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Monterey, CA, USA, Oct./Nov. 2006, pp. 233–237.

[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] 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.

[6] 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.

[7] 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.

[8] 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.

[9] A. Nedić and A. Ozdaglar, "Cooperative distributed multi-agent optimization," in *Convex Optimization in Signal Processing and Communications*, Y. C. Eldar and D. P. Palomar, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2009, pp. 340–386.

[10] I. D. Schizas, G. Mateos, and G. B. Giannakis, "Distributed LMS for consensus-based in-network adaptive processing," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2365–2382, Jun. 2009.

[11] S. Kar and J. M. F. Moura, "Distributed consensus algorithms in sensor networks with imperfect communication: Link failures and channel noise," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 355–369, Jan. 2009.

[12] S.-Y. Tu and A. H. Sayed, "Diffusion strategies outperform consensus strategies for distributed estimation over adaptive networks," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6217–6234, Dec. 2012.

[13] 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.

[14] J. Chen, C. Richard, and A. H. Sayed, "Multitask diffusion adaptation over networks," *IEEE Trans. Signal Process.*, vol. 62, no. 16, pp. 4129–4144, Aug. 2014.

[15] 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.

[16] X. Zhao and A. H. Sayed, "Clustering via diffusion adaptation over networks," in *Proc. 3rd Int. Workshop Cognit. Inf. Process. (CIP)*, Pontevedra, Spain, May 2012, pp. 1–6.
[17] J. Chen, C. Richard, and A. H. Sayed, "Diffusion LMS over multitask networks," *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2733–2748, Jun. 2015.
[18] J. Hua, C. Li, and H.-L. Shen, "Distributed learning of predictive structures from multiple tasks over networks," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4246–4256, May 2017, doi: 10.1109/TIE.2016.2588463.
[19] X. Zhao and A. H. Sayed, "Distributed clustering and learning over networks," *IEEE Trans. Signal Process.*, vol. 63, no. 13, pp. 3285–3300, Jul. 2015.
[20] J. Ni, J. Chen, and X. Chen, "Diffusion sign-error LMS algorithm: Formulation and stochastic behavior analysis," *Signal Process.*, vol. 128, pp. 142–149, Nov. 2016.
[21] F. S. Cattivelli and A. H. Sayed, "Modeling bird flight formations using diffusion adaptation," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2038–2051, May 2011.
[22] 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.
[23] R. Nassif, C. Richard, A. Ferrari, and A. H. Sayed, "Multitask diffusion adaptation over asynchronous networks," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2835–2850, Jun. 2016.
[24] R. Abdolee and B. Champagne, "Diffusion LMS strategies in sensor networks with noisy input data," *IEEE/ACM Trans. Netw.*, vol. 24, no. 1, pp. 3–14, Feb. 2016.
[25] S.-H. Yim, H.-S. Lee, and W.-J. Song, "A proportionate diffusion LMS algorithm for sparse distributed estimation," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 62, no. 10, pp. 992–996, Oct. 2015.
[26] F. Wen, "Diffusion least-mean P-power algorithms for distributed estimation in alpha-stable noise environments," *Electron. Lett.*, vol. 49, no. 21, pp. 1355–1356, Oct. 2013.
[27] E. Eweda, "Analysis and design of a signed regressor LMS algorithm for stationary and nonstationary adaptive filtering with correlated Gaussian data," *IEEE Trans. Circuits Syst.*, vol. CAS-37, no. 11, pp. 1367–1374, Nov. 1990.
[28] C. P. Kwong, "Control-theoretic design of the LMS and the sign algorithms in nonstationary environments," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 2, pp. 253–259, Feb. 1990.
[29] S. H. Cho and V. J. Mathews, "Tracking analysis of the sign algorithm in nonstationary environments," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 12, pp. 2046–2057, Dec. 1990.
[30] E. Eweda, "Comparison of RLS, LMS, and sign algorithms for tracking randomly time-varying channels," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 2937–2944, Nov. 1994.
[31] J. A. Bucklew, T. G. Kurtz, and W. A. Sethares, "Weak convergence and local stability properties of fixed step size recursive algorithms," *IEEE Trans. Inf. Theory*, vol. 30, no. 3, pp. 966–978, May 1993.
[32] S. Dasgupta, C. R. Johnson, Jr., and A. M. Baksho, "Characterizing persistent excitation for the sign-sign equation error identifier," *Automatica*, vol. 29, no. 6, pp. 1473–1489, Nov. 1993.
[33] E. Eweda, "Transient and tracking performance bounds of the sign-sign algorithm," *IEEE Trans. Signal Process.*, vol. 47, no. 8, pp. 2200–2210, Aug. 1999.
[34] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Process.*, vol. 86, pp. 639–697, Apr. 2006.
[35] W. A. Gardner, Ed., *Cyclostationarity in Communications and Signal Processing*. Piscataway, NJ, USA: IEEE Press, 1994.
[36] M. Nassar, J. Lin, Y. Mortazavi, A. Dabak, I. H. Kim, and B. L. Evans, "Local utility power line communications in the 3–500 kHz band: Channel impairments, noise, and standards," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 116–127, Sep. 2012.
[37] L. Castedo and A. R. Figueiras-Vidal, "An adaptive beamforming technique based on cyclostationary signal properties," *IEEE Trans. Signal Process.*, vol. 43, no. 7, pp. 1637–1650, Jul. 1995.
[38] R. Wang, C. Hou, and J. Yang, "Blind adaptive beamforming algorithm based on cyclostationary signals," in *Proc. Int. Conf. Image Anal. Signal Process.*, Apr. 2010, pp. 488–491.
[39] R. W. Heath, Jr., and G. B. Giannakis, "Exploiting input cyclostationarity for blind channel identification in OFDM systems," *IEEE Trans. Signal Process.*, vol. 47, no. 3, pp. 848–856, Mar. 1999.
[40] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Commun.*, vol. 21, no. 3, pp. 118–127, Jun. 2014.
[41] D. C. McLernon, "Analysis of LMS algorithm with inputs from cyclostationary random processes," *Electron. Lett.*, vol. 27, no. 2, pp. 136–138, Jan. 1991.
[42] E. Eweda, "Behavior of the least mean square algorithm with a periodically time-varying input power," *Int. J. Adapt. Control Signal Process.*, vol. 26, no. 12, pp. 1057–1075, Mar. 2012.
[43] E. Eweda, "Comparison of LMS and NLMS adaptive filters with a non-stationary input," in *Proc. 44th Annu. Asilomar Conf. Signals, Syst. Comput.*, Nov. 2010, pp. 1630–1634.
[44] N. J. Bershad and J. C. M. Bermudez, "Stochastic analysis of the LMS algorithm for non-stationary white Gaussian inputs," in *Proc. IEEE Statist. Signal Process. Workshop*, Jun. 2011, pp. 57–60.
[45] N. J. Bershad, E. Eweda, and J. C. M. Bermudez, "Stochastic analysis of the LMS and NLMS algorithms for cyclostationary white Gaussian inputs," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2238–2249, May 2014.
[46] E. Eweda, N. J. Bershad, and J. C. M. Bermudez, "Stochastic analysis of the least mean fourth algorithm for non-stationary white Gaussian inputs," *Signal, Image Video Process.*, vol. 8, no. 1, pp. 133–142, Jan. 2014.
[47] N. J. Bershad, E. Eweda, and J. C. M. Bermudez, "Stochastic analysis of an adaptive line enhancer/canceler with a cyclostationary input," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 104–119, Jan. 2016.
[48] E. Eweda and N. J. Bershad, "Stochastic analysis of the signed LMS algorithms for cyclostationary white Gaussian inputs," *IEEE Trans. Signal Process.*, vol. 65, no. 7, pp. 1673–1684, Apr. 2017.
[49] N. Shlezinger, K. Todros, and R. Dabora, "Adaptive filtering based on time-averaged MSE for cyclostationary signals," *IEEE Trans. Commun.*, vol. 65, no. 4, pp. 1746–1761, Apr. 2017, doi: 10.1109/TCOMM.2017.2655526.
[50] R. Price, "A useful theorem for nonlinear devices having Gaussian inputs," *IRE Trans. Inf. Theory*, vol. 4, no. 2, pp. 69–72, Jun. 1958.
[51] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd ed. New York, NY, USA: McGraw-Hill, 1991.
[52] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, 2nd ed. Reading, MA, USA: Addison-Wesley, 1994.
[53] R. Abdolee, V. Vakilian, and B. Champagne, "Tracking performance and optimal adaptation step-sizes of diffusion-LMS networks," *IEEE Trans. Control Netw. Syst.*, to be published, doi: 10.1109/TCNS.2016.2578044.
[54] S. O. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.
[55] A. H. Sayed, *Fundamentals of Adaptive Filtering*. New York, NY, USA: Wiley, 2003.

WENYUAN WANG received the B.E. degree from the Heilongjiang University of Science and Technology, China, in 2014. He is currently pursuing the Ph.D. degree in signal and information processing with the School of Electrical Engineering, Southwest Jiaotong University, China. His research interests include adaptive filtering algorithm and distributed filtering.

HAIQUAN ZHAO (M'11–SM'17) was born in Henan, China, in 1974. He received the B.S. degree in applied mathematics, and the M.S. degree and the Ph.D. degree in signal and information processing from Southwest Jiaotong University, Chengdu, China, in 1998, 2005, and 2011, respectively. Since 2012, he has been a Professor with the School of Electrical Engineering, Southwest Jiaotong University. He has authored or co-authored over 50 journal papers. He holds six invention patents. His current research interests include adaptive filtering algorithm, adaptive Volterra filter, nonlinear active noise control, nonlinear system identification, and chaotic signal processing. He is a member of the IEEE Computational Intelligence Society. He has served as an active reviewer for several IEEE Transactions, The Institution of Engineering and Technology, and other international journals.