# Diffusion Widely Linear Adaptive Estimation of System Frequency in Distributed Power Grids

Sithan Kanna[1], Sayed Pouria Talebi[2], Danilo P. Mandic[3]

*Department of Electrical and Electronic Engineering, Imperial College London*
*Exhibition Road, London, SW7 2BT, United Kingdom*
[1] ssk08@ic.ac.uk
[2] s.talebi12@imperial.ac.uk
[3] d.mandic@imperial.ac.uk

***Abstract*—Motivated by the need for decentralized estimation in the electricity grid, we apply the diffusion augmented complex least mean square (D-ACLMS) algorithm to estimate the frequency of a three-phase power system under balanced and unbalanced voltage conditions in a distributed setting. In addition to using local measurements, the diffusion adaptation strategy allows nodes to share their frequency estimates with neighbouring nodes which enables the network to exploit multiple measurements of the voltage signal. The analysis provides convergence conditions of the D-ACLMS and also gives the bounds on parameter values which guarantee that the diffusion type of adaptation stabilizes the network. Simulations under both balanced and unbalanced conditions support the analysis.**

***Index Terms*—Complex least mean square (CLMS), widely linear modelling, augmented statistics, diffusion adaptation, three-phase power system.**

## I. Introduction

The rapid growth in the penetration of intermittent power sources, distributed generation and the abundance of data from sensor networks call for the decentralization of parameter estimation algorithms [1]. Decentralized estimation algorithms also have the benefit of requiring less power for communication (since there is no fusion centre to send the data to) and are more robust to sensor failures and changes in network topology [2]. An important parameter that is routinely estimated in the electricity grid is the system or mains frequency.

Utility companies and electricity grid operators rely on the estimate of the mains frequency as it indicates the balance between generation and load. If demand is greater than generation, the frequency drops, whereas if generation is greater than demand, the frequency rises. The need for fast and accurate frequency estimates for the control of the grid is also important because grid operators have a legal requirement to keep the nominal frequency between a certain range; in the U.K., for example, the range is between 49.5 Hz and 50.5 Hz.

The system frequency can be estimated using one of the three-phase voltage signals or a projection of the three-phase voltage onto two orthogonal axes (known as $\alpha$-$\beta$ axes) [3]. The latter method is preferred as it utilizes the information present in all three phases [4]. Therefore, we shall restrict our discussion to $\alpha$-$\beta$ voltages. Under balanced operating conditions, the $\alpha$-$\beta$ voltage can be represented as a strictly linear complex-valued autoregressive (AR) process and standard strictly linear estimators like the complex least mean square (CLMS) [5], recursive least squares (RLS) [6], [7] and complex Kalman filter [8] can be used to estimate the system frequency.

The estimators mentioned thus far assume a balanced operating condition and consequently form a biased estimate of the frequency when there is an imbalance (e.g. voltage sag). Our previous work in employing widely linear modelling to estimate the system frequency resulted in the augmented CLMS (ACLMS) [9], augmented complex Kalman filter [10], and augmented minimum variance distortionless response (AMVDR) spectrum [11] algorithms. These widely linear algorithms are able to estimate the system frequency under balanced and unbalanced voltage conditions by using the input vector and its conjugate.

In this work, we extend the ACLMS based frequency estimation method in [9] to a distributed setting where nodes in a network (each using the ACLMS locally) collaborate to estimate the system frequency using a diffusion adaptation strategy [12], [13]. The distributed version of the ACLMS, referred to the diffusion ACLMS (D-ACLMS) [14], is suitable for exploiting the diversity in having multiple measurements of the system voltage. In the diffusion adaptation setting, each node uses an adaptive algorithm (in this case, the ACLMS) to estimate a parameter of interest and combines it with the estimates it receives from neighbouring nodes. The application of the D-ACLMS for a frequency estimation task is desirable because of its

1) adaptivity, making it suitable for real-time estimation of the system frequency;
2) ability to estimate the frequency under balanced and unbalanced operating conditions;
3) implementation in a distributed manner, which eliminates the need for a fusion centre [15].

The rest of this paper is organized as follows: in Section II, the D-ACLMS is reviewed for a general setting. In Section III, we show how the D-ACLMS is applied to estimate the frequency. Section IV presents the convergence analysis for the algorithm and finally Section V contains simulations of the D-ACLMS performing a frequency estimation task in a distributed sensor network.

978-1-4799-2449-3/14/$31.00 ©2014 IEEE

## NOTATION

| | |
|---|---|
| $a, \mathbf{a}, \mathbf{A}$ | Scalar, column vector, matrix; |
| $(\cdot)^T, (\cdot)^H$ | Transpose, Hermitian transpose |
| $\otimes$ | Kronecker product; |
| $\mathbf{I}_M$ | $M \times M$ identity Matrix; |
| $\mathbf{1}_N$ | $N \times 1$ vector of ones; |
| $y_{i,k}$ | Local output of node $i$ at time $k$; |
| $\{d_{i,k},\ \mathbf{z}_{i,k}\}$ | Local desired signal and input vector of node $i$ at time $k$; |
| $j$ | Complex number, $\sqrt{-1}$; |
| $(\cdot)^*$ | Complex conjugation; |
| $\mathsf{E}\{\cdot\}$ | Statistical expectation operator; |
| $\rho(\mathbf{A})$ | Spectral radius (maximum eigenvalue) of a matrix $\mathbf{A}$; |
| $\mathsf{Re}\{\cdot\}, \mathsf{Im}\{\cdot\}$ | Real, Imaginary parts of a complex number; |

## II. BACKGROUND: DIFFUSION AUGMENTED COMPLEX LEAST MEAN SQUARE

We consider a collaborative estimation task that is carried out by several nodes in a network of $N$ nodes. As shown in Figure 1, a node $i$ has communication links with other nodes in its neighbourhood, $\mathcal{N}_i$. The cardinality of the neighbourhood, denoted by $|\mathcal{N}_i|$, is the number of nodes the node $i$ is connected to, including itself.

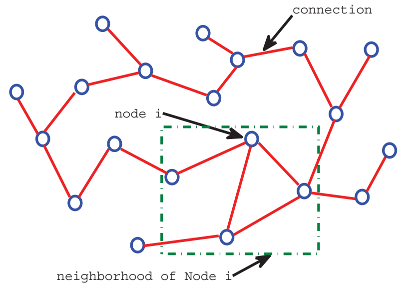



Fig. 1. A general distributed network showing a node $i$ and its neighbourhood.

Each node receives measurement data $\{d_{i,k},\ \mathbf{x}_{i,k}\}$ at every time instant $k$, where $d_{i,k}$ and $\mathbf{x}_{i,k}$ are respectively the desired signal and input vector at node $i$. The desired signal $d_{i,k} \in \mathbb{C}$ and input vector $\mathbf{x}_{i,k} \in \mathbb{C}^M$ are assumed to be related through a widely linear model, given by

$$d_{i,k} = \mathbf{h}_o^H \mathbf{x}_{i,k} + \mathbf{g}_o^H \mathbf{x}_{i,k}^* + \eta_{i,k}. \tag{1}$$

The measurement noise, represented by $\eta_{i,k}$, is assumed to be zero-mean, and temporally and spatially uncorrelated, so that

$$\mathsf{E}\left\{\eta_{i,k}\eta_{j,m}^*\right\} = \sigma_\eta^2 \delta(i-j)\delta(k-m) \tag{2}$$

where $\sigma_\eta^2$ denotes the variance of $\eta_{i,k}$ and $\delta(\cdot)$ is the Kronecker delta function. For convenience, the weight vectors $\mathbf{h}_o$, $\mathbf{g}_o$ and input vectors $\mathbf{x}_{i,k}$, $\mathbf{x}_{i,k}^*$ shall be represented using the compact augmented notation as

$$\mathbf{w}_o = \begin{bmatrix}\mathbf{h}_o \\ \mathbf{g}_o\end{bmatrix} \qquad \mathbf{z}_{i,k} = \begin{bmatrix}\mathbf{x}_{i,k} \\ \mathbf{x}_{i,k}^*\end{bmatrix}$$

so that the signal model in (1) can be written as

$$d_{i,k} = \mathbf{w}_o^H \mathbf{z}_{i,k} + \eta_{i,k}. \tag{3}$$

where $\mathbf{z}_{i,k}$ = At node $i$, the estimate of the local desired signal, $d_{i,k}$, is therefore given by a widely linear model

$$y_{i,k} = \mathbf{w}_{i,k}^H \mathbf{z}_{i,k} \tag{4}$$

where the augmented weight vector $\mathbf{w}_{i,k}$ and augmented input vector $\mathbf{z}_{i,k}$ are complex-valued vectors of length $2M$. The error $e_{i,k}$ between the desired signal and the estimate is given by

$$e_{i,k} = d_{i,k} - y_{i,k}. \tag{5}$$

At each node, the diffusion augmented complex least mean square (D-ACLMS) algorithm performs a stochastic gradient descent minimization procedure on two cost functions in succession. Firstly, all the nodes minimize the squared error (SE) cost function

$$\mathcal{J}_{i,k}^{(\mathrm{SE})} = |e_{i,k}|^2 = |d_{i,k} - y_{i,k}|^2 \tag{6}$$

using their local measurements $d_{i,k}$ and input vectors $\mathbf{z}_{i,k}$. Note that the squared error cost function is used to approximate the mean square error (MSE), $\mathsf{E}\left\{|e_{i,k}|^2\right\}$, at each node. The SE cost function is minimized using a stochastic gradient descent framework [14]

$$\boldsymbol{\psi}_{i,k+1} = \mathbf{w}_{i,k} - \mu \frac{\partial \mathcal{J}_{i,k}^{(\mathrm{SE})}}{\partial \mathbf{w}^*}\bigg|_{\mathbf{w}=\mathbf{w}_{i,k}} \tag{7}$$

where the step-size $\mu$ controls the speed of convergence along the error performance surface. The term $\boldsymbol{\psi}_{i,k+1}$ is used instead of $\mathbf{w}_{i,k+1}$ to indicate that the estimate in the iteration in (7) is an intermediate one and will be succeeded by another step. The partial derivative of $\mathcal{J}_{i,k}^{(\mathrm{SE})}$ is then found [16], [17]

$$\frac{\partial \mathcal{J}_{i,k}^{(\mathrm{SE})}}{\partial \mathbf{w}^*} = \frac{\partial [e_{i,k} e_{i,k}^*]}{\partial \mathbf{w}^*} = e_{i,k}^* \frac{\partial e_{i,k}}{\partial \mathbf{w}^*} + e_{i,k} \frac{\partial e_{i,k}^*}{\partial \mathbf{w}^*} \tag{8}$$

Since $e_{i,k} = d_{i,k} - \mathbf{w}_{i,k}^H \mathbf{z}_{i,k}$,

$$\frac{\partial e_{i,k}}{\partial \mathbf{w}^*} = -\mathbf{z}_{i,k} \tag{9}$$

which makes the adaptation step

$$\boldsymbol{\psi}_{i,k+1} = \mathbf{w}_{i,k} + \mu e_{i,k}^* \mathbf{z}_{i,k}. \tag{10}$$

The nodes then transmit their local estimates $\boldsymbol{\psi}_{i,k+1}$ to neighbouring nodes so that each local node performs the second step of the adaptation process, whereby every node minimizes the difference between its own weight estimate in (10) and the weights in its neighbourhood $\mathcal{N}_i$, by minimizing the cost function

$$\mathcal{J}_{i,k}^{(\mathrm{AV})} = \sum_{\ell \in \mathcal{N}_i} b_{\ell i} \|\boldsymbol{\psi}_{\ell,k+1} - \boldsymbol{\psi}_{i,k+1}\|^2 \tag{11}$$

which is the weighted norm (using weights $b_{\ell i}$) of the difference between the estimates in the neighbourhood. The minimization of this cost function is also carried out using the stochastic gradient descent method where

$$\mathbf{w}_{i,k+1} = \boldsymbol{\psi}_{i,k+1} - \mu \frac{\partial \mathcal{J}_{i,k}^{(\mathrm{AV})}}{\partial \boldsymbol{\psi}^*}\bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_{i,k+1}} \tag{12}$$

$$= \boldsymbol{\psi}_{i,k+1} + \sum_{\ell=1}^{N} \mu b_{\ell i}(\boldsymbol{\psi}_{\ell,k+1} - \boldsymbol{\psi}_{i,k+1}) \tag{13}$$

Rearranging the equation gives us

$$\mathbf{w}_{i,k+1} = (1 - \sum_{\ell \neq i} \mu b_{\ell i})\boldsymbol{\psi}_{i,k+1} + \sum_{\ell \neq i} \mu b_{\ell i} \boldsymbol{\psi}_{\ell,k+1} \tag{14}$$

Defining $a_{\ell i} = \mu b_{\ell i}$ we can see that $a_{ii} = 1 - \sum_{\ell \neq i} a_{\ell i}$, which gives us

$$\mathbf{w}_{i,k+1} = \sum_{\ell=1}^{N} a_{\ell i} \boldsymbol{\psi}_{\ell,k+1} \tag{15}$$

The weighting coefficients (sometimes referred to as the "trust coefficients") are zero, $a_{\ell i} = 0$, if the node $i$ and $\ell$ are not connected to each other. For sufficiently small step-sizes, the combiners are also positive, $a_{\ell i} \geq 0$, and are chosen so that they sum up to one [18], that is

$$\sum_{\ell=1}^{N} a_{\ell i} = 1 \implies a_{ii} = 1 - \sum_{\ell \neq i} a_{\ell i} \tag{16}$$

Although the determination of the optimal weights for an arbitrary network of nodes is challenging without accurate knowledge of the data statistics for every node [19], it is possible to set combination rules based on information like the cardinality of the neighbourhood. An example of combination methodology is the Metropolis rule, where [18]

$$a_{\ell i} = \begin{cases} 1/\max\{|\mathcal{N}_i|, |\mathcal{N}_\ell|\} & \text{if } i \neq \ell \text{ are connected} \\ 1 - \sum_{i \neq \ell} a_{\ell i} & \text{if } i = \ell \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

A discussion about the choice of combiners is out of the scope of this paper and we refer the reader to [18].

## III. D-ACLMS for Distributed Frequency Estimation

### A. Modelling a Three Phase System

The three-phase voltages of a power system measured at a single node (in a noise-less environment) can be represented with a vector

$$\begin{bmatrix} v_{a,k} \\ v_{b,k} \\ v_{c,k} \end{bmatrix} = \begin{bmatrix} V_{a,k}\cos(\omega k \Delta T + \phi_a) \\ V_{b,k}\cos(\omega k \Delta T + \phi_b - 2\pi/3) \\ V_{c,k}\cos(\omega k \Delta T + \phi_c + 2\pi/3) \end{bmatrix} \tag{18}$$

where the magnitudes $V_{a,k}$, $V_{b,k}$, and $V_{c,k}$ are the peak values of these sinusoids, $\Delta T = \frac{1}{f_{\mathrm{samp}}}$ is the sampling interval, $\phi$ is the phase shift, $k$ is the time index and $\omega = 2\pi f_0$, where $f_0$ is the system frequency we are interested in. A power system is said to be in a balanced operating condition when: (a) All the three signals have the same peak voltage, i.e. $V_{a,k} = V_{b,k} = V_{c,k} = V$. (b) The three voltages have a phase shift of exactly 120 degrees, i.e. $\phi_a = \phi_b = \phi_c = \phi$.

The three-phase signal can be projected onto two orthogonal axes, forming $v_\alpha$ and $v_\beta$ with the Clarke's Transform (also known as the $\alpha\beta$-transform)

$$\begin{bmatrix} v_{0,k} \\ v_{\alpha,k} \\ v_{\beta,k} \end{bmatrix} = \underbrace{\sqrt{\frac{2}{3}} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}}_{\mathbf{C}=\text{Clarke's Matrix}} \begin{bmatrix} v_{a,k} \\ v_{b,k} \\ v_{c,k} \end{bmatrix} \tag{19}$$

A third component, known as the zero-sequence voltage, $v_{0,k}$, is needed to make the transformation invertible. Under a balanced condition, the transformation is designed to make the third component zero.

The $\alpha$ and $\beta$ components can be conveniently represented as a complex number by setting

$$v_k = v_{\alpha,k} + j v_{\beta,k} \tag{20}$$

Under balanced conditions, the zero-sequence voltage $v_{0,k}$ is equal to zero and the complex voltage is a complex exponential

$$v_k = A e^{j(\omega k \Delta T + \phi)} \tag{21}$$

When the system is in a temporary imbalance, like the ones classified by [20], the complex-valued representation in (21) is incorrect. The general form of the complex-valued voltage that arises from the Clarke's Transform is

$$v_k = A_k e^{j(\omega k \Delta T + \phi)} + B_k e^{-j(\omega k \Delta T + \phi)} \tag{22}$$

where

$$A_k = \frac{\sqrt{6}(V_{a,k} + V_{b,k} + V_{c,k})}{6} \tag{23}$$

$$B_k = \frac{\sqrt{6}(2V_{a,k} - V_{b,k} - V_{c,k})}{12} - j\frac{\sqrt{2}(V_{b,k} - V_{c,k})}{4}. \tag{24}$$

The coefficient $B_k$ becomes zero when the system is balanced and this model collapses into the one shown in (21).

### B. Adaptive Estimation of Frequency

When there is an imbalance in the system, the coefficient $B_k$ is non-zero and the complex signal is non-circular [21]. We will assume that at a node $i$, the complex-valued voltage shown in (22) is corrupted by measurement noise $\eta_{i,k}$, so that

$$v_{i,k} = A_{i,k} e^{j(\omega k \Delta T + \phi)} + B_{i,k} e^{-j(\omega k \Delta T + \phi)} + \eta_{i,k}. \tag{25}$$

We re-introduce the subscript $i$ to indicate that we are moving from a single node setting to a distributed setting. Under a balanced operating condition (when the term $B_{i,k} = 0$), the complex voltage in (25) can be represented by a strictly linear Auto-Regressive (AR) model of order one,

$$v_{i,k+1} = w_o^* v_{i,k} + \eta_{i,k} \tag{26}$$

where $w_o = e^{-j\omega\Delta T}$. Any strictly linear adaptive filter can be used to estimate $e^{-j\omega\Delta T}$ where the system frequency can then be obtained from the phase of the coefficient [5]. When there is an imbalance in the system, the term $B_k$ in equation (22) is non-zero and this makes the strictly linear AR(1) model in (26) inadequate to describe the relationship between $v_k$ and $v_{k+1}$.

The widely linear AR(1) model for the complex voltage, however, can be used to model the data for balanced and unbalanced operating conditions and is given by

$$v_{i,k+1} = h_o^* v_{i,k} + g_o^* v_{i,k}^* + \eta_{i,k}. \tag{27}$$

Assuming that $A_{i,k+1} \approx A_{i,k}$ and $B_{i,k+1} \approx B_{i,k}$ for a high sampling frequency [9]

$$e^{j\omega\Delta T} = h_o^* + g_o^* \frac{B_{i,k}^*}{A_{i,k}}, \qquad e^{-j\omega\Delta T} = h_o^* + g_o^* \frac{A_{i,k}}{B_{i,k}}. \tag{28}$$

Using a widely linear estimator for the signal model (27), $\hat{v}_{i,k+1} = h_k^* v_{i,k} + g_k^* v_{i,k}^*$, the frequency can then be estimated as

$$\begin{aligned} f_{i,k} &= \frac{1}{2\pi\Delta T} \tan^{-1}\left(\frac{-\mathsf{Im}\left\{h_{i,k} + u_{i,k}^* g_{i,k}\right\}}{\mathsf{Re}\left\{h_{i,k} + u_{i,k}^* g_{i,k}\right\}}\right) \\ u_{i,k} &= \frac{j}{g_k^*}\left(\mathsf{Im}\left\{h_{i,k}\right\} + \sqrt{[\mathsf{Im}\left\{h_{i,k}\right\}]^2 - |g_{i,k}|^2}\right) \end{aligned} \tag{29}$$

For the full derivation, we refer to [9], [22]. The frequency estimation given by expressions (29) can be accomplished with any widely linear estimator. We use the D-ACLMS to track this frequency due to the benefits it provides as described in Section I. The D-ACLMS algorithm is summarized in Algorithm 1.

**Algorithm 1** Diffusion Augmented Complex Least Mean Square (D-ACLMS) for Frequency Estimation

**Input data**: $\mathbf{v}_{i,k} = [v_{i,k},\ v_{i,k}^*]^T$

**Desired sequence**: $v_{i,k+1}$

**for** Nodes $i = \{1, \ldots, N\}$ **do**

$\hat{v}_{i,k+1} = \mathbf{w}_{i,k}^H \mathbf{v}_{i,k}$ where $\mathbf{w}_{i,k} = [h_{i,k},\ g_{i,k}]^T$

$e_{i,k} = v_{i,k+1} - \hat{v}_{i,k+1}$

$\boldsymbol{\psi}_{i,k+1} = \mathbf{w}_{i,k} + \mu e_{i,k}^* \mathbf{v}_{i,k}$

$\mathbf{w}_{i,k+1} = \sum_{\ell=1}^{N} a_{\ell i} \boldsymbol{\psi}_{i,k+1}$

$f_{i,k} = \frac{1}{2\pi\Delta T} \tan^{-1}\left(\frac{-\mathsf{Im}\left\{h_{i,k}+u_{i,k}^* g_{i,k}\right\}}{\mathsf{Re}\left\{h_{i,k}+u_{i,k}^* g_{i,k}\right\}}\right)$

where $u_{i,k} = \frac{j}{g_k^*}\left(\mathsf{Im}\left\{h_{i,k}\right\} + \sqrt{[\mathsf{Im}\left\{h_{i,k}\right\}]^2 - |g_{i,k}|^2}\right)$

**end for**

## IV. Convergence of the D-ACLMS

Consider a network with $N$ adaptive filters, each with a filter length of $2M$. The factor of two is present due to the use of augmented input and weight vectors. Consider a block-diagonal input matrix where each augmented input vector is placed in a separate column

$$\mathbf{Z}_k = \begin{bmatrix} \mathbf{z}_{1,k} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{z}_{2,k} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{z}_{N,k} \end{bmatrix}, \quad \mathbf{Z}_k \in \mathbb{C}^{2NM \times N} \tag{30}$$

and $\mathbf{0}$ is a $2M \times 1$ column of zeros. The other variables in the network are

$$\mathbf{y}_k = \left[y_{1,k}, \ \ldots \ , y_{N,k}\right]^T, \qquad \mathbf{y}_k \in \mathbb{C}^{N\times 1} \tag{31}$$

$$\mathbf{d}_k = \left[d_{1,k}, \ \ldots \ , d_{N,k}\right]^T, \qquad \mathbf{d}_k \in \mathbb{C}^{N\times 1} \tag{32}$$

$$\mathbf{w}_k = \left[\mathbf{w}_{1,k}^T, \ \ldots \ , \mathbf{w}_{N,k}^T\right]^T, \qquad \mathbf{w}_k \in \mathbb{C}^{2NM\times 1} \tag{33}$$

$$\boldsymbol{\psi}_k = \left[\boldsymbol{\psi}_{1,k}^T, \ \ldots \ , \boldsymbol{\psi}_{N,k}^T\right]^T, \qquad \boldsymbol{\psi}_k \in \mathbb{C}^{2NM\times 1} \tag{34}$$

The weights $a_{\ell i}$ are placed in an $N \times N$ matrix, where the weights that node $i$ assigns to nodes $\ell \in \mathcal{N}_i$ are placed in the $i$-th column of $\mathcal{A}$

$$\mathcal{A} = \begin{bmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2i} & \cdots & a_{2N} \\ \vdots & \cdots & \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{Ni} & \cdots & a_{NN} \end{bmatrix}, \quad \mathcal{A} \in \mathbb{R}^{N\times N}. \tag{35}$$

For convenience we also assume that the matrix $\mathcal{A}$ is symmetric $\mathcal{A}^T = \mathcal{A}$. This condition makes $\mathcal{A}$ a doubly stochastic matrix for which the rows and columns add up to one, that is

$$\mathbf{1}_N^T \mathcal{A} = \mathbf{1}_N^T, \qquad \mathcal{A}\mathbf{1}_N = \mathbf{1}_N \tag{36}$$

Finally, we construct a matrix $\mathbf{A}$

$$\mathbf{A} = \mathcal{A} \otimes \mathbf{I}_{2M} \qquad \mathbf{A} \in \mathbb{R}^{2NM\times 2NM} \tag{37}$$

where $\mathbf{I}_{2M}$ is a $2M$-by-$2M$ identity matrix. Consequently, the D-ACLMS algorithm for the whole network can be represented by

$$\mathbf{y}_k = \mathbf{Z}_k^T \mathbf{w}_k^* \tag{38}$$

$$\mathbf{e}_k = \mathbf{d}_k - \mathbf{y}_k \tag{39}$$

$$\boldsymbol{\psi}_{k+1} = \mathbf{w}_k + \mu \mathbf{Z}_k \mathbf{e}_k^* \tag{40}$$

$$\mathbf{w}_{k+1} = \mathbf{A}\boldsymbol{\psi}_{k+1} \tag{41}$$

assuming all the nodes have the same step-size $\mu$. The weight update from (40) and (41) can be combined into a single step as

$$\mathbf{w}_{k+1} = \mathbf{A}\mathbf{w}_k + \mu \mathbf{A}\mathbf{Z}_k \mathbf{e}_k^* \tag{42}$$

Assuming that all the nodes are estimating the same optimal weight vector $\mathbf{w}_{o,\text{ind}} \in \mathbf{C}^{2M\times 1}$, we are able to construct the optimal weight vector for the network by stacking $N$ copies of $\mathbf{w}_{o,\text{ind}}$ on top of each other to form $\mathbf{w}_o \in \mathbf{C}^{2NM\times 1}$. Defining the weight error vector $\tilde{\mathbf{w}}_k \triangleq \mathbf{w}_o - \mathbf{w}_k$ and recognizing that $\mathbf{A}\mathbf{w}_o = \mathbf{w}_o$ allows us to formulate the weight error recursion as

$$\tilde{\mathbf{w}}_{k+1} = \mathbf{A}\tilde{\mathbf{w}}_k - \mu \mathbf{A}\mathbf{Z}_k \left(\mathbf{Z}_k^H \tilde{\mathbf{w}}_k + \boldsymbol{\eta}_k^*\right) \tag{43}$$

where $\boldsymbol{\eta}_k$ is the vector containing measurement noise for all the nodes. Applying the statistical expectation operator and assuming that the measurement noise is independent of the input

$$\mathsf{E}\left\{\tilde{\mathbf{w}}_{k+1}\right\} = \mathbf{A}\underbrace{\left[\mathbf{I}_{2MN} - \mu\mathsf{E}\left\{\mathbf{Z}_k\mathbf{Z}_k^H\right\}\right]}_{\triangleq \mathbf{B}} \mathsf{E}\left\{\tilde{\mathbf{w}}_k\right\} \tag{44}$$

The matrix $\mathsf{E}\left\{\mathbf{Z}_k\mathbf{Z}_k^H\right\}$ is a block-diagonal matrix with the augmented covariance matrices of the input at each node in the diagonal

$$\mathsf{E}\left\{\mathbf{Z}_k\mathbf{Z}_k^H\right\} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{0}_{2M} & \cdots & \mathbf{0}_{2M} \\ \mathbf{0}_{2M} & \mathbf{R}_2 & \cdots & \mathbf{0}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{2M} & \mathbf{0}_{2M} & \cdots & \mathbf{R}_N \end{bmatrix} \tag{45}$$

where $\mathbf{R}_i = \mathsf{E}\left\{\mathbf{z}_{i,k}\mathbf{z}_{i,k}^H\right\}$ is the augmented covariance matrix of the input at node $i$ and $\mathbf{0}_{2M}$ is a $2M$-by-$2M$ matrix of zeros. Therefore, the weight error recursion can be written as

$$\mathsf{E}\left\{\tilde{\mathbf{w}}_{k+1}\right\} = \mathbf{AB}\mathsf{E}\left\{\tilde{\mathbf{w}}_k\right\} \tag{46}$$

For this recursion to be a contraction, the spectral radius of the matrix $\mathbf{AB}$ has to be less than unity. The spectral radius of a matrix, is bounded by any induced matrix norm [23]

$$\rho(\mathbf{AB}) \leq \|\mathbf{AB}\| \leq \|\mathbf{A}\|\|\mathbf{B}\| \tag{47}$$

Since the norm of a doubly stochastic matrix is unity, $\rho(\mathbf{AB}) \leq \|\mathbf{B}\|$. As $\mathbf{B}$ is a Hermitian block-diagonal matrix, its block maximum norm is equal to its spectral radius $\|\mathbf{B}\|_{b,\infty} = \rho(\mathbf{B})$ which implies that

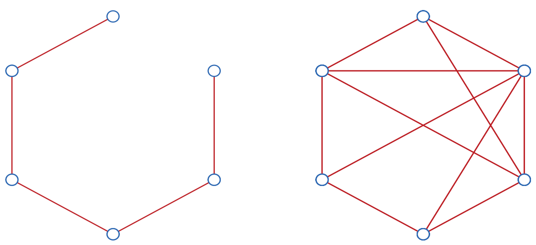
$$\rho(\mathbf{AB}) \leq \rho(\mathbf{B}) \tag{48}$$

for any doubly stochastic matrix $\mathbf{A}$ [18].

**Remark:** The diffusion strategy enhances the stability range of the network by reducing the spectral radius of $\mathbf{AB}$.

## V. Simulations

The simulations for the D-ACLMS was based on a network of $N = 6$ nodes. Figure 2(a) shows a sparsely connected network, where each node only has access to a maximum of two other nodes, whereas Figure 2(b) shows a network with a larger number of connections between the nodes. Unless stated otherwise, the network with the least number of connections is chosen (Figure 2(a)) because it serves as a worst-case scenario for the performance gain as increasing the number of connections only improves the diffusion algorithm.



(a) Sparsely connected (b) Densely connected

Fig. 2. A six-node network with two different levels of connectivity.

Each node in the network has access to noisy measurements of a common complex-valued $\alpha$-$\beta$ voltage.

$$v_{i,k} = v_{\alpha,k} + jv_{\beta,k} + \eta_{i,k} \quad \eta_i \sim \mathcal{N}(0,\sigma_\eta^2) \tag{49}$$

where the measurement noise, $\eta_{i,k}$, is modelled as complex-valued circular white Gaussian noise with the same variance, $\sigma_\eta^2$, at each node. The complex-valued $\alpha$-$\beta$ voltage is generated by applying the Clarke's transformation on a three-phase voltage signal that oscillates at a constant 50 Hz.

$$\begin{bmatrix} v_{\alpha,k} \\ v_{\beta,k} \end{bmatrix} = \sqrt{\frac{2}{3}}\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \mathrm{Re}\left\{\begin{bmatrix} \bar{V}_a \\ \bar{V}_b \\ \bar{V}_c \end{bmatrix} e^{j\omega k\Delta T}\right\} \tag{50}$$

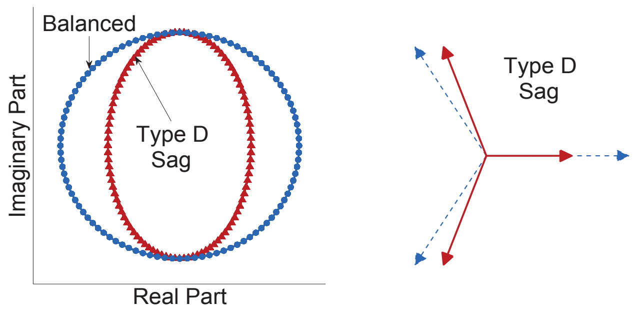
where $\bar{V}_a$, $\bar{V}_b$ and $\bar{V}_c$ are phasor representations of the three-phase voltage. The sampling frequency $f_{\text{samp}} = 1/\Delta T$ is chosen to be 2 kHz. Under a balanced operating condition, $\bar{V}_a = 1$, $\bar{V}_b = -\frac{1}{2} - j\frac{\sqrt{3}}{2}$ and $\bar{V}_c = -\frac{1}{2} + j\frac{\sqrt{3}}{2}$. An imbalance in the system causes voltage sags that can be represented by changing the phasors as shown in Table I [20].

| **Voltage Sag** | $\bar{V}_a$ | $\bar{V}_b$ | $\bar{V}_c$ |
|---|---|---|---|
| Type A | $\gamma$ | $-\frac{\gamma}{2} - j\frac{\sqrt{3}\gamma}{2}$ | $-\frac{\gamma}{2} + j\frac{\sqrt{3}\gamma}{2}$ |
| Type B | $\gamma$ | $-\frac{1}{2} - j\frac{\sqrt{3}}{2}$ | $-\frac{1}{2} + j\frac{\sqrt{3}}{2}$ |
| Type C | 1 | $-\frac{1}{2} - j\frac{\sqrt{3}\gamma}{2}$ | $-\frac{1}{2} + j\frac{\sqrt{3}\gamma}{2}$ |
| Type D | $\gamma$ | $-\frac{\gamma}{2} - j\frac{\sqrt{3}}{2}$ | $-\frac{\gamma}{2} + j\frac{\sqrt{3}}{2}$ |

TABLE I

Voltage sags and their phasor represntations.

The D-ACLMS was benchmarked against the diffusion-CLMS (D-CLMS) which uses a strictly linear model for the data under a distributed setting. Both the D-CLMS and D-ACLMS were used with the same input and desired signals and were configured for the same network topology. The signal was generated under a Type D voltage sag, with $\gamma = 0.9$.




(a) Circularity Plot (b) Phasor Representation

Fig. 3. Geometric and phasor diagrams of the voltage under a Type D imbalance.

The D-CLMS and D-ACLMS weights were initialized randomly and the same learning rate, $\mu = 0.2$ was used for each filter across the network. Although having different learning rates introduces some diversity in the network, we restrict ourselves to the basic case of using the same learning rate.

The top panel of Figure 4 shows the average mean square error (MSE) of the frequency estimate of across the nodes using the D-CLMS and D-ACLMS at different levels of measurement noise power. The average MSE value was calculated by averaging the MSE of the frequency estimate from all the nodes obtained using 100 independent trials. The formula is given by

$$\mathrm{MSE_{av}} = 10\log_{10}\left[\frac{1}{100\cdot 6}\sum_{m=1}^{100}\sum_{i=1}^{6}(f_{i,k}^{(\mathrm{m})} - f_0)^2\right] \tag{51}$$

where $f_0 = 50$. The top panel of Figure 4 shows that the D-CLMS has a higher average MSE level than the D-ACLMS when there is an imbalance in the system. This is expected as the D-CLMS lacks the modelling capability of D-ACLMS to model voltage imbalances in the system.

To observe the performance of each node in the six node network, the individual MSEs of the nodes (at SNR = 60 dB) are plotted on the bottom panel of Figure 4. The MSE values at the nodes with and without the diffusion strategy shows that sharing the frequency estimates reduces the steady state MSE for all the nodes in the network.

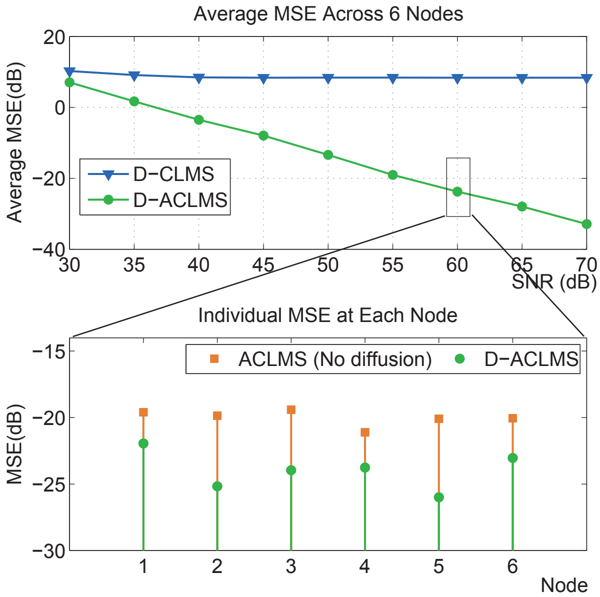



Fig. 4. *Top panel*: The average MSE of the frequency estimate of D-ACLMS is lower than that of the D-CLMS under a Type D sag. *Bottom panel*: MSE for each node in the network with and without cooperation at SNR = 60 dB shows that the diffusion strategy reduces the steady state error for all the nodes.

Figure 5 shows the average MSE across the nodes for various levels of connectivity in the network. As the connectivity between the nodes increases from a sparse to a dense level (where all the nodes are allowed to communicate with each other), the average steady-state MSE decreases. This is because as the number of connections between the nodes are increased, the network behaves like a centralized one where the data from the nodes are aggregated and processed at a fusion center.

Figure 6 shows the frequency estimate given by the the D-ACLMS and D-CLMS at a randomly selected node in the network. The system was simulated with an SNR level of 60 dB and with a balanced set of three-phase voltages with a

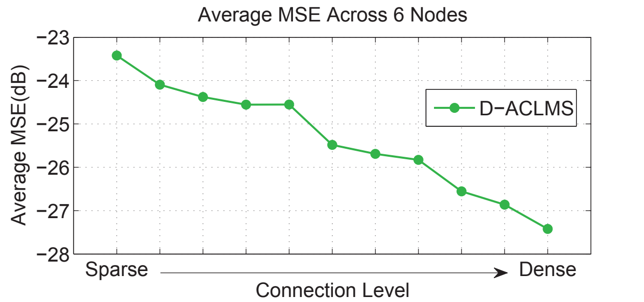



Fig. 5. MSE for the D-ACLMS decreases with increasing connectivity at an SNR level of 60 dB.

temporary Type D imbalance between 1 s and 2 s. The D-CLMS oscillates during the Type D imbalance at twice the nominal frequency due to under-modelling errors [22]. The D-ACLMS is able to estimate the 50 Hz frequency accurately except during the transition period between the operating conditions (1 s to 1.08 s and 2 s to 2.08 s).

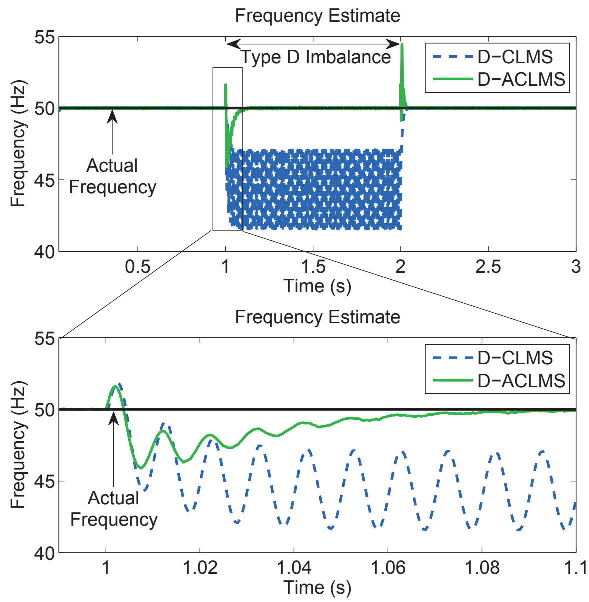



Fig. 6. Frequency estimate at a randomly selected node with a temporary Type D imbalance between 1 s and 2 s.

## VI. Conclusion

We have applied the diffusion augmented complex least mean square (D-ACLMS) algorithm for a frequency estimation task in a distributed sensor network. The D-ACLMS algorithm benefits from the widely linear modelling from the ACLMS and has the ability to reduce steady-state MSE by exploiting multiple measurements in a distributed setting. The convergence analysis shows that the doubly stochastic matrix used in the diffusion strategy is able to stabilize the network. Simulations show that the D-ACLMS has a lower steady-state MSE compared to the D-CLMS for different levels of SNRs. The diffusion strategy also is shown to reduce the steady-state MSE of all the nodes in a network.

## REFERENCES


[1] S. Kar and J. Moura, “Real-time adaptive distributed state estimation in smart grids,” in *Proc. of the Asia Pacific Signal Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 1–7, 2012.

[2] J. Tsitsiklis, D. Bertsekas, and M. Athans, “Distributed asynchronous deterministic and stochastic gradient optimization algorithms,” *IEEE Transactions on Automatic Control*, vol. 31, pp. 803 – 812, sep 1986.

[3] M. H. Bollen and I. Gu, *Signal processing of power quality disturbances*, vol. 30. John Wiley & Sons, 2006.

[4] M. Canteli, A. Fernandez, L. Eguiluz, and C. Estebanez, “Three-phase adaptive frequency measurement based on Clarke’s transformation,” *IEEE Transactions on Power Delivery*, vol. 21, no. 3, pp. 1101–1105, 2006.

[5] A. Pradhan, A. Routray, and A. Basak, “Power system frequency estimation using least mean square technique,” *IEEE Transactions on Power Delivery*, vol. 20, no. 3, pp. 1812–1816, 2005.

[6] L. Li, W. Xia, D. Shi, and J. Li, “Frequency estimation on power system using recursive-least-squares approach,” in *Proc. of the 2012 International Conference on Information Technology and Software Engineering*, vol. 211 of *Lecture Notes in Electrical Engineering*, pp. 11–18, Springer Berlin Heidelberg, 2013.

[7] Z. Chen, Z. Sahinoglu, and H. Li, “Fast frequency and phase estimation in three phase power systems,” in *Proc. of the IEEE Power and Energy Society General Meeting (PES)*, pp. 1–5, 2013.

[8] P. Dash, A. Pradhan, and G. Panda, “Frequency estimation of distorted power system signals using extended complex kalman filter,” *IEEE Transactions on Power Delivery*, vol. 14, no. 3, pp. 761–766, 1999.

[9] Y. Xia and D. Mandic, “Widely linear adaptive frequency estimation of unbalanced three-phase power systems,” *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 1, pp. 74–83, 2012.

[10] D. Dini, Y. Xia, S. Douglas, and D. Mandic, “Widely linear state space models for frequency estimation in unbalanced three-phase systems,” in *Proc. of the 7th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pp. 9–12, 2012.

[11] Y. Xia and D. Mandic, “Augmented MVDR spectrum-based frequency estimation for unbalanced power systems,” *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 7, pp. 1917–1926, 2013.

[12] C. Lopes and A. Sayed, “Diffusion least-mean squares over adaptive networks,” in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 3, pp. III–917 –III–920, Apr. 2007.

[13] C. Lopes and A. Sayed, “Diffusion least-mean squares over adaptive networks: Formulation and performance analysis,” *IEEE Transactions on Signal Processing*, vol. 56, pp. 3122 –3136, Jul. 2008.

[14] Y. Xia, D. P. Mandic, and A. H. Sayed, “An adaptive diffusion augmented CLMS algorithm for distributed filtering of noncircular complex signals,” *IEEE Signal Process. Lett.*, vol. 18, no. 11, pp. 659–662, 2011.

[15] A. Gomez-Exposito, A. Abur, A. de la Villa Jaen, and C. Gomez-Quiles, “A multilevel state estimation paradigm for smart grids,” *Proc. of the IEEE*, vol. 99, no. 6, pp. 952–976, 2011.

[16] B. D. H. Brandwood, “A complex gradient operator and its applications in adaptive array theory,” *IEE Proceedings: Communications, Radar and Signal Proceedings*, vol. 130, no. 1, pp. 11–16, 1983.

[17] A. van den Bos, “Complex gradient and Hessian,” *IEE Proc.: Vision, Image and Signal Processing*, vol. 141, no. 6, pp. 380–383, 1994.

[18] A. H. Sayed, “Diffusion Adaptation over Networks,” *ArXiv e-prints*, May 2012.

[19] X.R. Li, Y. Zhu, J. Wang, and C. Han, “Optimal linear estimation fusion : Part I: Unified fusion rules,” *IEEE Transactions on Information Theory*, vol. 49, pp. 2192 – 2208, Sep. 2003.

[20] M. Bollen, “Characterisation of voltage sags experienced by three-phase adjustable-speed drives,” *IEEE Transactions on Power Delivery*, vol. 12, no. 4, pp. 1666–1671, 1997.

[21] Y. Xia and D. P. Mandic, “A widely linear least mean phase algorithm for adaptive frequency estimation of unbalanced power systems,” *International Journal of Electrical Power & Energy Systems*, vol. 54, no. 0, pp. 367 – 375, 2014.

[22] Y. Xia, S. Douglas, and D. Mandic, “Adaptive frequency estimation in smart grid applications: Exploiting noncircularity and widely linear adaptive estimators,” *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 44–54, 2012.

[23] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, 2012.