Distributed Widely Linear Kalman Filtering for Frequency Estimation in Power Networks

Sithan Kanna, *Student Member, IEEE*, Dahir H. Dini, Yili Xia, *Member, IEEE*, S. Y. Hui, *Fellow, IEEE*, and Danilo P. Mandic, *Fellow, IEEE*

***Abstract*—Motivated by the growing need for robust and accurate frequency estimators at the low- and medium-voltage distribution levels and the emergence of ubiquitous sensors networks for the smart grid, we introduce a distributed Kalman filtering scheme for frequency estimation. This is achieved by using widely linear state space models, which are capable of estimating the frequency under both balanced and unbalanced operating conditions. The proposed distributed augmented extended Kalman filter (D-ACEKF) exploits multiple measurements without imposing any constraints on the operating conditions at different parts of the network, while also accounting for the correlated and noncircular natures of real-world nodal disturbances. Case studies over a range of power system conditions illustrate the theoretical and practical advantages of the proposed methodology.**

***Index Terms*—Adaptive networks, frequency estimation, Kalman filters, sensor fusion, smart grid.**

I. Introduction

THE MODERN power network, commonly known as the "smart grid," aims to be more reliable by incorporating distributed generation, typically from renewable sources (wind) and often in a local fashion (solar, micro-wind turbines). The rapid growth in renewable energy sources in the smart grid also gives rise to challenges in maintaining the stability of the network, as these renewable sources are connected to low-voltage distribution networks through grid-interfaced power electronics inverters.

The balance between power generation and consumption is a pre-requisite for stable operation of the power network. An imbalance between generation and load causes a frequency deviation and grid operators require fast and accurate estimates of the frequency to control the stability of the network [1], [2]. In addition, power inverters themselves require accurate frequency estimates for stable grid-interfacing under noisy conditions. As inverters are switching circuits, they themselves introduce impulsive noise to the voltage signal [3], which, in turn, requires frequency estimation algorithms with enhanced robustness against noise and voltage spikes.

Currently used frequency estimation techniques include: i) Fourier transform approaches [4]–[7], ii) gradient decent and least squares adaptive estimation [8], and iii) state space methods and Kalman filters [9]–[12]. However, these are typically designed for single-phase systems and often assume balanced operating conditions (equal voltage amplitudes and equally spaced phases) and are therefore inadequate for the demands of modern three–phase and dynamically optimised power systems.

To avoid the loss of information and the compromise in accuracy by estimating the frequency from only one of the three phases, the Clarke transform is used to jointly represent a three-phase voltage signal as a complex-valued scalar variable known as the $\alpha\beta$ voltage [13]–[15]. Our earlier work showed that this $\alpha\beta$ voltage admits a widely linear autoregressive model under both balanced and unbalanced system conditions [16]. It was also shown that standard, strictly linear, complex-valued estimators applied to the complex $\alpha\beta$ voltage introduce biased estimates for unbalanced system conditions, widely linear estimators (also known as "augmented" estimators) are able to provide optimal and consistent estimates of the system frequency over a range of operating conditions [16]–[19].

The aim of this work is to extend the single-node widely linear state space frequency estimator in [18] to the distributed scenario to suit the requirements of the smart grid. Distributed estimation has already found application in both military and civilian scenarios [20]–[24], as cooperation between the nodes (sensors) provides more accurate and robust estimation over the independent nodes, while approaching the performance of centralised systems at much reduced communication overheads. Recent distributed approaches include diffusion least-mean-square estimation [25]–[27] and Kalman filtering [22], [24], [28], however, these consider circular measurement noise without cross-nodal correlations, which is not realistic in real-world power systems.

To this end, we propose the diffusion augmented (widely linear) complex Kalman filter (D-ACKF) and the diffusion augmented extended Kalman filter (D-ACEKF) by adapting the diffusion scheme in [29] and the Kalman filtering models in [18], [30] to suit our problem where the system frequency is identical over a certain geographical area at the distribution level while the voltage imbalances and cross-nodal correlations can be different. In particular, extending the widely linear frequency estimator to the distributed case is non-trivial as the states that are to be estimated are not-identical in the network, as required by the classical diffusion scheme [29].

Manuscript received December 10, 2014; revised April 10, 2015; accepted May 22, 2015. Date of publication June 23, 2015; date of current version June 29, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Isao Yamada.

S. Kanna, D. H. Dini, and D. P. Mandic are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: ssk08@ic.ac.uk; d.mandic@ic.ac.uk).

Y. Xia is with the School of Infirmation Science and Engineering, Southeast University, Nanjing 210096, China.

S. Y. (Ron) Hui is with Hong Kong University, Hong Kong, and also with Imperial College London, London SW7 2AZ, U.K. (e-mail: ronhui@eee.hku.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSIPN.2015.2442834

The proposed D-ACEKF is able improve the frequency estimates at the low- and medium-voltage distribution levels of the electricity grid by exploiting a diversity of measurements without imposing any constrains on the operating conditions. Moreover, the proposed D-ACEKF accounts for the correlation between the observation noises at neighbouring nodes, typically encountered when node signals are exposed to common sources of interference (harmonics, fluctuations of reactive power), which is not catered for in current estimation algorithms. Case studies using synthetic and real world signals support the claims.

II. Background on Widely Linear Modelling

Statistical signal processing in the complex-domain underpins a number of disciplines, including wireless communications [31], [32] and power systems [16]. Although it may be convenient to process complex-valued data by representing the real and imaginary parts as a bivariate signal in the real domain, any intuition and physical meaning inherent in processing in the complex domain would be obscured. However, the statistical tools for complex random variables were derived by assuming (often implicitly) that these complex random variables were second order circular (proper)[1] [33]–[35].

Consider the problem of estimating a real-valued vector $\mathbf{x} \in \mathbb{R}^L$ from the data vector $\mathbf{y} \in \mathbb{R}^K$. According the linear minimum mean square error (LMMSE) estimation theory, the optimal estimate (in terms of second order statistics) is[2]

$$\hat{\mathbf{x}} = \mathsf{E}\{\mathbf{x}|\mathbf{y}\} = \mathbf{R}_{\mathbf{xy}}\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{y} \tag{1}$$

where $\mathbf{R}_{\mathbf{xy}} = \mathsf{E}\{\mathbf{x}\mathbf{y}^T\}$ is the cross-covariance matrix, $\mathbf{R}_{\mathbf{y}} = \mathsf{E}\{\mathbf{y}\mathbf{y}^T\}$ is the autocovariance matrix, and $(\cdot)^T$ denotes the transpose operator. When the vectors $\mathbf{x}$ and $\mathbf{y}$ are complex-valued, the so-called *strictly linear* estimation assumes the same model in (1) with complex valued correlation matrices $\mathbf{R}_{\mathbf{xy}} = \mathsf{E}\{\mathbf{x}\mathbf{y}^H\}$ and $\mathbf{R}_{\mathbf{y}} = \mathsf{E}\{\mathbf{y}\mathbf{y}^H\}$ where the transpose operator is replaced with the Hermitian operator $(\cdot)^H$ [30], [36].

It has been shown that the strictly linear model in (1) is sub-optimal for a general class of complex-valued signals. The widely linear minimum mean square error (WLMMSE) solution exploits the full second order statistics of the signal contained within the augmented autocovariance matrix

$$\mathbf{R}_{\mathbf{y}}^a = \mathsf{E}\begin{bmatrix}\mathbf{y}\\ \mathbf{y}^*\end{bmatrix}\begin{bmatrix}\mathbf{y}^H & \mathbf{y}^T\end{bmatrix} = \begin{bmatrix}\mathbf{R}_{\mathbf{y}} & \mathbf{P}_{\mathbf{y}}\\ \mathbf{P}_{\mathbf{y}}^* & \mathbf{R}_{\mathbf{y}}^*\end{bmatrix} \tag{2}$$

and augmented cross-covariance matrix

$$\mathbf{R}_{\mathbf{xy}}^a = \mathsf{E}\begin{bmatrix}\mathbf{x}\mathbf{y}^H & \mathbf{x}\mathbf{y}^T\end{bmatrix} = \begin{bmatrix}\mathbf{R}_{\mathbf{xy}} & \mathbf{P}_{\mathbf{xy}}\end{bmatrix} \tag{3}$$

[1]Proper signals have real and imaginary parts with equal variance and are uncorrelated. Circular signals have rotationally invariant probability density functions (pdf). We use the terms "circular" and "proper" interchangeably for Gaussian signals, as they have pdfs that can be fully characterised by the first and second order statistics.

[2]This problem is sometimes posed as that of estimating vector $\mathbf{y}$ from observations $\mathbf{x}$. To be consistent with the Kalman filtering literature, however, we chose to represent the state with variable $\mathbf{x}$ and observation with $\mathbf{y}$.

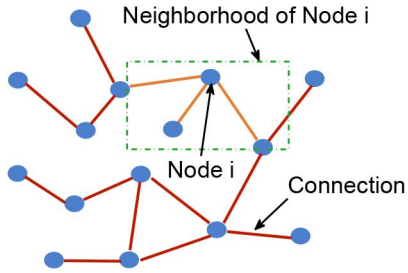


Fig. 1. An example of a distributed network topology.

which not only contain the autocovariance and cross-covariance matrices $\mathbf{R}_{\mathbf{yy}}$ and $\mathbf{R}_{\mathbf{xy}}$ but also the pseudocovariance matrices $\mathbf{P}_{\mathbf{xy}} = \mathsf{E}\{\mathbf{x}\mathbf{y}^T\}$ and $\mathbf{P}_{\mathbf{y}} = \mathsf{E}\{\mathbf{y}\mathbf{y}^T\}$. The WLMMSE solution is given by

$$\hat{\mathbf{x}} = \mathsf{E}\{\mathbf{x}|\mathbf{y}, \mathbf{y}^*\} = \mathbf{B}\mathbf{y} + \mathbf{C}\mathbf{y}^* \tag{4}$$

where the $\mathbf{B} = \left[\mathbf{R}_{\mathbf{xy}} - \mathbf{P}_{\mathbf{xy}}\mathbf{R}_{\mathbf{y}}^{-*}\mathbf{P}_{\mathbf{y}}^*\right]\mathbf{D}_{\mathbf{y}}^{-1}$, $\mathbf{C} = [\mathbf{P}_{\mathbf{xy}} - \mathbf{R}_{\mathbf{xy}}\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{P}_{\mathbf{y}}]\,\mathbf{D}_{\mathbf{y}}^{-*}$ and $\mathbf{D}_{\mathbf{y}} = \left[\mathbf{R}_{\mathbf{y}} - \mathbf{P}_{\mathbf{y}}\mathbf{R}_{\mathbf{y}}^{-*}\mathbf{P}_{\mathbf{y}}^*\right]$ is the Schur's complement of the augmented auto-covariance matrix $\mathbf{R}_{\mathbf{y}}^a$ in (2). The WLMMSE solution in (4) is equivalent to the LMMSE estimator derived from treating the real and imaginary parts of the signals as bivariate real-valued vectors [37] but provides much more physical insight in the analysis. For jointly-circular signals that have vanishing pseudocovariance matrices $\mathbf{P}_{\mathbf{xy}} = \mathbf{0}$ and $\mathbf{P}_{\mathbf{y}} = \mathbf{0}$ (referred to as circular or proper signals), the widely linear model in (4) degenerates to the strictly linear model in (1).

III. Diffusion Kalman Filtering

For a standard linear state space model, every node i in a distributed system (see Fig. 1) is given by [38]

$$\begin{aligned}\mathbf{x}_n &= \mathbf{F}_{n-1}\mathbf{x}_{n-1} + \mathbf{w}_n\\ \mathbf{y}_{i,n} &= \mathbf{H}_{i,n}\mathbf{x}_n + \mathbf{v}_{i,n}\end{aligned} \tag{5}$$

where $\mathbf{x}_n \in \mathbb{C}^L$ and $\mathbf{y}_{i,n} \in \mathbb{C}^K$ are respectively the state vector at time instant n and observation (measurement) vector at node i. The symbol $\mathbf{F}_{n-1}$ denotes the state transition matrix, $\mathbf{w}_n \in \mathbb{C}^L$ white state noise, $\mathbf{H}_{i,n}$ the observation matrix, and $\mathbf{v}_{i,n} \in \mathbb{C}^K$ white measurement measurement noise (both at node i).

Standard state space models assume the noises $\mathbf{w}_n$ and $\mathbf{v}_{i,n}$ to be uncorrelated and zero-mean, with their covariance and pseudocovariance matrices described by

$$\begin{aligned}\mathsf{E}\begin{bmatrix}\mathbf{w}_n\\ \mathbf{v}_{i,n}\end{bmatrix}\begin{bmatrix}\mathbf{w}_k^H & \mathbf{v}_{i,k}^H\end{bmatrix} &= \begin{bmatrix}\mathbf{Q}_n & \mathbf{0}\\ \mathbf{0} & \mathbf{R}_{i,n}\end{bmatrix}\delta_{nk}\\ \mathsf{E}\begin{bmatrix}\mathbf{w}_n\\ \mathbf{v}_{i,n}\end{bmatrix}\begin{bmatrix}\mathbf{w}_k^T & \mathbf{v}_{i,k}^T\end{bmatrix} &= \begin{bmatrix}\mathbf{P}_n & \mathbf{0}\\ \mathbf{0} & \mathbf{U}_{i,n}\end{bmatrix}\delta_{nk}\end{aligned} \tag{6}$$

where δ_{nk} is the standard Kronecker delta function.

A. Distributed Complex Kalman Filter

The distinguishing feature of the proposed class of distributed Kalman filters is that we generalise the diffusion

strategy in [22] by equipping it with state and noise models that do not impose any restrictions on: i) the correlation properties of the cross-nodal observation noises, or ii) the signal and noise circularity at different nodes. This also allows distributed Kalman filtering algorithms [22], [24], [39] to be used in wider application scenarios. Fig. 1 illustrates the distributed estimation scenario; the highlighted neighbourhood of node i compromises the set of nodes, denoted by $\mathcal{N}_i$, that communicate with the node i (including Node i itself). The state estimate at node i with a complex Kalman filter (CKF) is then based on all the data from the neighbourhood $\mathcal{N}_i$ consisting of $M = |\mathcal{N}_i|$ nodes, and is denoted by $\underline{\widehat{\mathbf{x}}}_{i,n|n}$, where the symbol $|\mathcal{N}_i|$ denotes the number of nodes in the neighbourhood $\mathcal{N}_i$. Finally, the collective neighbourhood observation equation at node i is given by

$$\underline{\mathbf{y}}_{i,n} = \underline{\mathbf{H}}_{i,n}\mathbf{x}_n + \underline{\mathbf{v}}_{i,n} \tag{7}$$

while the collective (neighbourhood) variables are defined as

$$\begin{aligned}
\underline{\mathbf{y}}_{i,n} &= \left[\mathbf{y}_{i_1,n}^T, \mathbf{y}_{i_2,n}^T, \ldots, \mathbf{y}_{i_M,n}^T\right]^T \\
\underline{\mathbf{H}}_{i,n} &= \left[\mathbf{H}_{i_1,n}^T, \mathbf{H}_{i_2,n}^T, \ldots, \mathbf{H}_{i_M,n}^T\right]^T \\
\underline{\mathbf{v}}_{i,n} &= \left[\mathbf{v}_{i_1,n}^T, \mathbf{v}_{i_2,n}^T, \ldots, \mathbf{v}_{i_M,n}^T\right]^T
\end{aligned}$$

where $\{i_1, i_2, \ldots, i_M\}$ are the nodes in the neighbourhood $\mathcal{N}_i$. The covariance and pseudocovariance matrices of the collective observation noise vector are given by

$$\underline{\mathbf{R}}_{i,n} = \mathsf{E}\left\{\underline{\mathbf{v}}_{i,n}\underline{\mathbf{v}}_{i,n}^H\right\} = \begin{bmatrix} \mathbf{R}_{i_1,n} & \mathbf{R}_{i_1 i_2,n} & \cdots & \mathbf{R}_{i_1 i_M,n} \\ \mathbf{R}_{i_2 i_1,n} & \mathbf{R}_{i_2,n} & \cdots & \mathbf{R}_{i_2 i_M,n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{i_M i_1,n} & \mathbf{R}_{i_M i_2,n} & \cdots & \mathbf{R}_{i_M,n} \end{bmatrix}$$

$$\underline{\mathbf{U}}_{i,n} = \mathsf{E}\left\{\underline{\mathbf{v}}_{i,n}\underline{\mathbf{v}}_{i,n}^T\right\} = \begin{bmatrix} \mathbf{U}_{i_1,n} & \mathbf{U}_{i_1 i_2,n} & \cdots & \mathbf{U}_{i_1 i_M,n} \\ \mathbf{U}_{i_2 i_1,n} & \mathbf{U}_{i_2,n} & \cdots & \mathbf{U}_{i_2 i_M,n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{U}_{i_M i_1,n} & \mathbf{U}_{i_M i_2,n} & \cdots & \mathbf{U}_{i_M,n} \end{bmatrix}$$

where $\mathbf{R}_{i_a,n} = \mathsf{E}\left\{\mathbf{v}_{i_a,n}\mathbf{v}_{i_a,n}^H\right\}$, $\mathbf{R}_{i_a i_b,n} = \mathsf{E}\left\{\mathbf{v}_{i_a,n}\mathbf{v}_{i_b,n}^H\right\}$, $\mathbf{U}_{i_a,n} = \mathsf{E}\left\{\mathbf{v}_{i_a,n}\mathbf{v}_{i_a,n}^T\right\}$ and $\mathbf{U}_{i_a i_b,n} = \mathsf{E}\left\{\mathbf{v}_{i_a,n}\mathbf{v}_{i_b,n}^T\right\}$, for $a, b \in \{1, 2, \ldots, M\}$.

Diffusion step. The local neighbourhood state estimates are followed by the diffusion (combination) step, given by

$$\widehat{\mathbf{x}}_{i,n|n} = \sum\nolimits_{\ell \in \mathcal{N}_i} c_{\ell i}\underline{\widehat{\mathbf{x}}}_{\ell,n|n} \tag{8}$$

which produces the diffused state estimates $\widehat{\mathbf{x}}_{i,n|n}$ as a weighted sum of the estimates from the neighbourhood $\mathcal{N}_i$, where $c_{\ell i} \geq 0$ are the weighting coefficients satisfying $\sum_{\ell \in \mathcal{N}_i} c_{\ell i} = 1$. The combination weights $c_{\ell i}$ used by the diffusion step in (8) can obey a number of rules, including the Metropolis [25], Laplacian [40] or the nearest neighbour [22] rules, however, finding the set of optimal weights remains an open issue though progress has been made in important cases[29], [41], [42].

The distributed complex Kalman filter (D-CKF) aims to approach the performance of a centralised Kalman filter (access to data from all the nodes) via neighbourhood collaborations and diffusion, and is summarised in Algorithm 1. Each node within D-CKF forms a collective observation, as in (7), using information from its neighbours; thereafter, each node computes a neighbourhood state estimate which is again shared with neighbours in order to be used for the diffusion step.

Algorithm 1. Diffusion Complex Kalman Filter (D-CKF)

Initialisation: for Nodes $i = \{1, 2, \ldots, N\}$

1: $\widehat{\mathbf{x}}_{i,0|0} = \mathsf{E}\left\{\mathbf{x}_0^T\right\}$

2: $\mathbf{M}_{i,0|0} = \mathsf{E}\left(\mathbf{x}_0 - \widehat{\mathbf{x}}_{i,0|0}\right)\left(\mathbf{x}_0 - \hat{\mathbf{x}}_{i,0|0}\right)^H$

At each time instant $n > 0$:

1: **for** Nodes $i = \{1, 2, \ldots, N\}$ **do**

2: $\widehat{\mathbf{x}}_{i,n|n-1} = \mathbf{F}_{n-1}\widehat{\mathbf{x}}_{i,n-1|n-1}$

3: $\mathbf{M}_{i,n|n-1} = \mathbf{F}_{n-1}\mathbf{M}_{i,n-1|n-1}\mathbf{F}_{n-1}^H + \mathbf{Q}_n$

4: $\mathbf{G}_{i,n} = \mathbf{M}_{i,n|n-1}\underline{\mathbf{H}}_{i,n}^H\left(\underline{\mathbf{H}}_{i,n}\mathbf{M}_{i,n|n-1}\underline{\mathbf{H}}_{i,n}^{aH} + \underline{\mathbf{R}}_{i,n}^a\right)^{-1}$

5. $\underline{\widehat{\mathbf{x}}}_{i,n|n} = \widehat{\mathbf{x}}_{i,n|n-1} + \mathbf{G}_{i,n}\left(\underline{\mathbf{y}}_{i,n} - \underline{\mathbf{H}}_{i,n}\widehat{\mathbf{x}}_{i,n|n-1}\right)$

6: $\mathbf{M}_{i,n|n} = (\mathbf{I} - \mathbf{G}_{i,n}\underline{\mathbf{H}}_{i,n})\mathbf{M}_{i,n|n-1}$

7: Diffuse the states from the network:

$\widehat{\mathbf{x}}_{i,n|n} = \sum_{\ell \in \mathcal{N}_i} c_{\ell i}\underline{\widehat{\mathbf{x}}}_{\ell,n|n}$

8: **end for**

Remark 1: The D-CKF algorithm[3] given in Algorithm 1 is a variant of that proposed in [22]. It employs the standard (strictly linear) state space model (5), and thus does not cater for widely linear complex state space models or noncircular state and observation noises ($\mathbf{P}_n \neq \mathbf{0}$ and $\mathbf{U}_{i,n} \neq \mathbf{0}$ for $i = 1, \ldots, N$). Unlike existing distributed complex Kalman filters, the D-CKF presented in Algorithm 1 caters for the correlations between the neighbourhood observation noises. When no such correlations exits, the D-CKF is identical to Kalman filter given in [22].

B. Distributed Augmented Complex Kalman Filter

We next employ the widely linear model in (4) in conjunction with D-CKF in Algorithm 1 to cater for widely linear state and observation models, and for improper measurements, states, and state and observation noises. The widely linear version of the distributed state space model in (5) (see also [30], [43]) is given by

$$\begin{aligned}
\mathbf{x}_n &= \mathbf{F}_{n-1}\mathbf{x}_{n-1} + \mathbf{A}_{n-1}\mathbf{x}_{n-1}^* + \mathbf{w}_n \\
\mathbf{y}_{i,n} &= \mathbf{H}_{i,n}\mathbf{x}_n + \mathbf{B}_{i,n}\mathbf{x}_n^* + \mathbf{v}_{i,n}.
\end{aligned} \tag{9}$$

The compact, "augmented" representation, of this model is

$$\begin{aligned}
\mathbf{x}_n^a &= \mathbf{F}_{n-1}^a\mathbf{x}_{n-1}^a + \mathbf{w}_n^a \\
\mathbf{y}_{i,n}^a &= \mathbf{H}_{i,n}^a\mathbf{x}_n^a + \mathbf{v}_{i,n}^a
\end{aligned} \tag{10}$$

where $\mathbf{x}_n^a = [\mathbf{x}_n^T, \mathbf{x}_n^H]^T$ and $\mathbf{y}_n^a = [\mathbf{y}_n^T, \mathbf{y}_n^H]^T$, while

$$\mathbf{F}_n^a = \begin{bmatrix} \mathbf{F}_n & \mathbf{A}_n \\ \mathbf{A}_n^* & \mathbf{F}_n^* \end{bmatrix} \quad \text{and} \quad \mathbf{H}_{i,n}^a = \begin{bmatrix} \mathbf{H}_{i,n} & \mathbf{B}_{i,n} \\ \mathbf{B}_{i,n}^* & \mathbf{H}_{i,n}^* \end{bmatrix}.$$

For *strictly linear systems*, $\mathbf{A}_n = \mathbf{0}$ and $\mathbf{B}_{i,n} = \mathbf{0}$, so that the widely linear (augmented) state space model degenerates

[3]The matrices $\mathbf{M}_{i,n|n}$ and $\mathbf{M}_{i,n|n-1}$ do not represent the covariances of $\widehat{\mathbf{x}}_{i,n|n}$ and $\widehat{\mathbf{x}}_{i,n|n-1}$, as is the case for the standard Kalman filter operating on linear Gaussian systems. This is due to the use of the suboptimal diffusion step, which updates the state estimate and not the covariance matrix $\mathbf{M}_{i,n|n}$.

into a strictly linear one. However, the augmented state space representation is still preferred in order to account for the pseudocovariances of the noise vectors which reflect the *impropriety of the noise*.

The augmented covariance matrices of the process noise vector $\mathbf{w}_n^a = [\mathbf{w}_n^T, \mathbf{w}_n^H]^T$ and observation noise vector $\mathbf{v}_{i,n}^a = [\mathbf{v}_{i,n}^T, \mathbf{v}_{i,n}^H]^T$ are then given by

$$\begin{aligned} \mathbf{Q}_n^a &= \mathsf{E}\left\{\mathbf{w}_n^a\, \mathbf{w}_n^{aH}\right\} = \begin{bmatrix} \mathbf{Q}_n & \mathbf{P}_n \\ \mathbf{P}_n^* & \mathbf{Q}_n^* \end{bmatrix} \\ \mathbf{R}_{i,n}^a &= \mathsf{E}\left\{\mathbf{v}_{i,n}^a \mathbf{v}_{i,n}^{aH}\right\} = \begin{bmatrix} \mathbf{R}_{i,n} & \mathbf{U}_{i,n} \\ \mathbf{U}_{i,n}^* & \mathbf{R}_{i,n}^* \end{bmatrix}. \end{aligned} \tag{11}$$

Neighbourhood variables. To perform collaborative estimation of the state within distributed networks, neighbourhood observation equations use all available neighbourhood observation data, to give

$$\underline{\mathbf{y}}_{i,n} = \underline{\mathbf{H}}_{i,n}\mathbf{x}_n + \underline{\mathbf{B}}_{i,n}\mathbf{x}_n^* + \underline{\mathbf{v}}_{i,n} \tag{12}$$

where the symbol $\underline{\mathbf{B}}_{i,n} = \left[\mathbf{B}_{i_1,n}^T, \mathbf{B}_{i_2,n}^T, \ldots, \mathbf{B}_{i_M,n}^T\right]^T$ denotes the conjugate observation matrix, and $\{i_1, i_2, \ldots, i_M\} \in \mathcal{N}_i$. The augmented neighbourhood observation equations now become

$$\underline{\mathbf{y}}_{i,n}^a = \underline{\mathbf{H}}_{i,n}^a \mathbf{x}_n^a + \underline{\mathbf{v}}_{i,n}^a \tag{13}$$

with the augmented neighbourhood variables defined as

$$\underline{\mathbf{y}}_{i,n}^a = \begin{bmatrix} \underline{\mathbf{y}}_{i,n} \\ \underline{\mathbf{y}}_{i,n}^* \end{bmatrix},\ \underline{\mathbf{H}}_{i,n}^a = \begin{bmatrix} \underline{\mathbf{H}}_{i,n} & \underline{\mathbf{B}}_{i,n} \\ \underline{\mathbf{B}}_{i,n}^* & \underline{\mathbf{H}}_{i,n}^* \end{bmatrix},\ \underline{\mathbf{v}}_{i,n}^a = \begin{bmatrix} \underline{\mathbf{v}}_{i,n} \\ \underline{\mathbf{v}}_{i,n}^* \end{bmatrix}. \tag{14}$$

Consequently, the covariance of the augmented neighbourhood observation noise $\underline{\mathbf{v}}_{i,n}^a$ takes the form

$$\underline{\mathbf{R}}_{i,n}^a = \mathsf{E}\left\{\underline{\mathbf{v}}_{i,n}^a \underline{\mathbf{v}}_{i,n}^{aH}\right\} = \begin{bmatrix} \underline{\mathbf{R}}_{i,n} & \underline{\mathbf{U}}_{i,n} \\ \underline{\mathbf{U}}_{i,n}^* & \underline{\mathbf{R}}_{i,n}^* \end{bmatrix}. \tag{15}$$

Remark 2: The augmented second order statistics in (15) caters for both the covariances $\mathsf{E}\left\{\mathbf{v}_{i,n}\mathbf{v}_{i,n}^H\right\}$ and cross-covariances $\mathsf{E}\{\mathbf{v}_{i,n}\mathbf{v}_{\ell,n}^H\}$, $i \neq \ell$ between the nodal observation noises. This is achieved through the standard covariance matrix $\underline{\mathbf{R}}_{i,n}$ and the pseudocovariances $\mathsf{E}\left\{\mathbf{v}_{i,n}\mathbf{v}_{i,n}^T\right\}$, while the cross-pseudocovariances $\mathsf{E}\{\mathbf{v}_{i,n}\mathbf{v}_{\ell,n}^T\}$ are accounted for through the pseudocovariance matrix $\underline{\mathbf{U}}_{i,n}$. Finally, the augmented diffused state estimate becomes

$$\widehat{\mathbf{x}}_{i,n|n}^a = \sum\nolimits_{\ell \in \mathcal{N}_i} c_{\ell i}\, \widehat{\underline{\mathbf{x}}}_{\ell,n|n}^a \tag{16}$$

and represents a weighted average of the augmented (neighbourhood) state estimates. The proposed distributed augmented complex Kalman filter (D-ACKF), employs the widely linear state space model in (9), and is given in Algorithm 2.

For strictly linear systems ($\mathbf{A}_n = \mathbf{0}$, $\mathbf{B}_{i,n} = \mathbf{0}$, $\forall\ n, i$) with circular state and observation noises ($\mathbf{P}_n = \mathbf{0}$, $\mathbf{U}_{i,n} = \mathbf{0}$, $\forall\ n, i$), the D-ACKF and D-CKF algorithms yield identical state estimates for all time instants n. Notice that the D-ACKF is more general than the D-CKF, since it also caters for the non-circular natures of data and noise, together with correlated state and observation noises.

Algorithm 2. Diffusion Augmented Complex Kalman Filter (D-ACKF)

Initialisation: for Nodes $i = \{1, 2, \ldots, N\}$

$\widehat{\mathbf{x}}_{i,0|0}^a = \mathsf{E}\left[\mathbf{x}_0^T,\quad \mathbf{x}_0^H\right]^T$

$\mathbf{M}_{i,0|0}^a = \mathsf{E}\left(\mathbf{x}_0^a - \widehat{\mathbf{x}}_{i,0|0}^a\right)\left(\mathbf{x}_0^a - \widehat{\mathbf{x}}_{i,0|0}^a\right)^H$

At each time instant $n > 0$:

for Nodes $i = \{1, 2, \ldots, N\}$ **do**

$\widehat{\mathbf{x}}_{i,n|n-1}^a = \mathbf{F}_{n-1}^a \widehat{\mathbf{x}}_{i,n-1|n-1}^a$

$\mathbf{M}_{i,n|n-1}^a = \mathbf{F}_{n-1}^a \mathbf{M}_{i,n-1|n-1}^a \mathbf{F}_{n-1}^{aH} + \mathbf{Q}_n^a$

$\mathbf{G}_{i,n}^a = \mathbf{M}_{i,n|n-1}^a \underline{\mathbf{H}}_{i,n}^{aH}\left(\underline{\mathbf{H}}_{i,n}^a \mathbf{M}_{i,n|n-1}^a \underline{\mathbf{H}}_{i,n}^{aH} + \underline{\mathbf{R}}_{i,n}^a\right)^{-1}$

$\widehat{\underline{\mathbf{x}}}_{i,n|n}^a = \widehat{\mathbf{x}}_{i,n|n-1}^a + \mathbf{G}_{i,n}^a\left(\underline{\mathbf{y}}_{i,n}^a - \underline{\mathbf{H}}_{i,n}^a \widehat{\mathbf{x}}_{i,n|n-1}^a\right)$

$\mathbf{M}_{i,n|n}^a = (\mathbf{I} - \mathbf{G}_{i,n}^a \underline{\mathbf{H}}_{i,n}^a)\mathbf{M}_{i,n|n-1}^a$

Diffuse the states from the network:

$\widehat{\mathbf{x}}_{i,n|n}^a = \sum_{\ell \in \mathcal{N}_i} c_{\ell i}\, \widehat{\underline{\mathbf{x}}}_{\ell,n|n}^a$

end for

IV. Distributed Augmented Complex Extended Kalman Filter

Consider a nonlinear state space model of the form

$$\begin{aligned} \mathbf{x}_n &= \mathbf{f}\left[\mathbf{x}_{n-1}\right] + \mathbf{w}_n \\ \mathbf{y}_{i,n} &= \mathbf{h}_i\left[\mathbf{x}_n\right] + \mathbf{v}_{i,n} \end{aligned} \tag{17}$$

where the nonlinear functions $\mathbf{f}[\cdot]$ and $\mathbf{h}_i[\cdot]$ are respectively the (possibly time varying) process model and observation model at node i, the remaining variables are as defined above. Within the extended Kalman filter (EKF) framework, the nonlinear state and observation functions are approximated by their first order Taylor series expansions (TSE) about the state estimates $\widehat{\mathbf{x}}_{i,n-1|n-1}$ and $\widehat{\mathbf{x}}_{i,n|n-1}$ for each node i, so that [44]

$$\begin{aligned} \mathbf{x}_n &\approx \mathbf{F}_{i,n-1}\mathbf{x}_{n-1} + \mathbf{A}_{i,n-1}\mathbf{x}_{n-1}^* + \mathbf{r}_{i,n-1} + \mathbf{w}_n \\ \mathbf{y}_{i,n} &\approx \mathbf{H}_{i,n}\mathbf{x}_n + \mathbf{B}_{i,n}\mathbf{x}_n^* + \mathbf{u}_{i,n} + \mathbf{v}_{i,n} \end{aligned} \tag{18}$$

where the Jacobians of functions $\mathbf{f}[\cdot]$ and $\mathbf{h}_i[\cdot]$ are defined as

$$\mathbf{F}_{i,n} = \frac{\partial \mathbf{f}[\mathbf{x}]}{\partial \mathbf{x}}\Big|_{\mathbf{x}=\widehat{\mathbf{x}}_{i,n|n}},\qquad \mathbf{A}_{i,n} = \frac{\partial \mathbf{f}[\mathbf{x}]}{\partial \mathbf{x}^*}\Big|_{\mathbf{x}^*=\widehat{\mathbf{x}}_{i,n|n}^*},$$

$$\mathbf{H}_{i,n} = \frac{\partial \mathbf{h}_i[\mathbf{x}]}{\partial \mathbf{x}}\Big|_{\mathbf{x}=\widehat{\mathbf{x}}_{i,n|n-1}} \quad \text{and} \quad \mathbf{B}_{i,n} = \frac{\partial \mathbf{h}_i[\mathbf{x}]}{\partial \mathbf{x}^*}\Big|_{\mathbf{x}^*=\widehat{\mathbf{x}}_{i,n|n-1}^*}$$

and the vectors

$$\begin{aligned} \mathbf{r}_{i,n} &= \mathbf{f}[\widehat{\mathbf{x}}_{i,n-1|n-1}] - \mathbf{F}_{i,n-1}\widehat{\mathbf{x}}_{i,n-1|n-1} - \mathbf{A}_{i,n-1}\widehat{\mathbf{x}}_{i,n-1|n-1}^* \\ \mathbf{u}_{i,n} &= \mathbf{h}_i[\widehat{\mathbf{x}}_{i,n|n-1}] - \mathbf{H}_{i,n}\widehat{\mathbf{x}}_{i,n|n-1} - \mathbf{B}_{i,n}\widehat{\mathbf{x}}_{i,n|n-1}^* \end{aligned}$$

are deterministic inputs calculated from the state space model and state estimate. The full second order statistics in the linearised state space in (18) is accounted for by its widely linear version, for which the augmented form is given by

$$\begin{aligned} \mathbf{x}_n^a &\approx \mathbf{F}_{i,n-1}^a \mathbf{x}_{n-1}^a + \mathbf{r}_{i,n-1}^a + \mathbf{w}_n^a \\ \mathbf{y}_{i,n}^a &\approx \mathbf{H}_{i,n}^a \mathbf{x}_n^a + \mathbf{u}_{i,n}^a + \mathbf{v}_{i,n}^a \end{aligned} \tag{19}$$

where $\mathbf{r}^a_{i,n} = [\mathbf{r}^T_{i,n}, \mathbf{r}^H_{i,n}]^T$, $\mathbf{u}^a_{i,n} = [\mathbf{u}^T_{i,n}, \mathbf{u}^H_{i,n}]^T$, while

$$\mathbf{F}^a_{i,n} = \begin{bmatrix} \mathbf{F}_{i,n} & \mathbf{A}_{i,n} \\ \mathbf{A}^*_{i,n} & \mathbf{F}^*_{i,n} \end{bmatrix} \quad \text{and} \quad \mathbf{H}^a_{i,n} = \begin{bmatrix} \mathbf{H}_{i,n} & \mathbf{B}_{i,n} \\ \mathbf{B}^*_{i,n} & \mathbf{H}^*_{i,n} \end{bmatrix}.$$

The collective neighbourhood augmented observation equation for node i now takes the form

$$\underline{\mathbf{y}}^a_{i,n} = \underline{\mathbf{h}}^a_i[\mathbf{x}_n] + \underline{\mathbf{v}}^a_{i,n} \tag{20}$$

while the collective observation function is defined as

$$\underline{\mathbf{h}}^a_i[\mathbf{x}_n] = \left[\underline{\mathbf{h}}^T_i[\mathbf{x}_n], \underline{\mathbf{h}}^H_i[\mathbf{x}_n]\right]^T$$
$$\underline{\mathbf{h}}_i[\mathbf{x}_n] = \left[\mathbf{h}^T_{i_1}[\mathbf{x}_n], \mathbf{h}^T_{i_2}[\mathbf{x}_n], \ldots, \mathbf{h}^T_{i_M}[\mathbf{x}_n]\right]^T$$

where $i \in \{i_1, i_2, \ldots, i_M\}$ spans all the nodes in the neighbourhood $\mathcal{N}_i$. The first order approximation of (20) is then

$$\underline{\mathbf{y}}^a_{i,n} \approx \underline{\mathbf{H}}^a_{i,n}\mathbf{x}^a_n + \underline{\mathbf{u}}^a_{i,n} + \underline{\mathbf{v}}^a_{i,n} \tag{21}$$

while the Jacobian of the collective observation becomes

$$\underline{\mathbf{H}}^a_{i,n} = \begin{bmatrix} \underline{\mathbf{H}}_{i,n} & \underline{\mathbf{B}}_{i,n} \\ \underline{\mathbf{B}}^*_{i,n} & \underline{\mathbf{H}}^*_{i,n} \end{bmatrix}$$

with $\underline{\mathbf{H}}_{i,n} = [\mathbf{H}^T_{i_1,n}, \mathbf{H}^T_{i_2,n}, \ldots, \mathbf{H}^T_{i_M,n}]^T$ and $\underline{\mathbf{B}}_{i,n} = [\mathbf{B}^T_{i_1,n}, \mathbf{B}^T_{i_2,n}, \ldots, \mathbf{B}^T_{i_M,n}]^T$, calculated as

$$\mathbf{H}_{i_k,n} = \left.\frac{\partial \mathbf{h}_{i_k}[\mathbf{x}]}{\partial \mathbf{x}}\right|_{\mathbf{x}=\widehat{\mathbf{x}}_{i,n|n-1}} \text{ and } \mathbf{B}_{i_k,n} = \frac{\partial \mathbf{h}_{i_k}[\mathbf{x}]}{\partial \mathbf{x}^*}|_{\mathbf{x}^*=\widehat{\mathbf{x}}^*_{i,n|n-1}}.$$

Algorithm 3 summarises the proposed distributed augmented complex extended Kalman filter, where each node i shares its (nonlinear) observation model $\mathbf{h}_i[\cdot]$ with its neighbours. The function $\mathsf{Jacobian}\,(\mathbf{f}, \mathbf{x})$ for steps 3 and 5 in Algorithm 3 computes the Jacobian matrix of the function $\mathbf{f}$ evaluated at the point $\mathbf{x}$.

Algorithm 3. Diffusion Augmented Complex Extended Kalman Filter (D-ACEKF)

Initialisation: for Nodes $i = \{1, 2, \ldots, N\}$

1: $\widehat{\mathbf{x}}^a_{i,0|0} = \mathsf{E}\left[\mathbf{x}^T_0, \quad \mathbf{x}^H_0\right]^T$

2: $\mathbf{M}^a_{i,0|0} = \mathsf{E}\left(\mathbf{x}^a_0 - \widehat{\mathbf{x}}^a_{i,0|0}\right)\left(\mathbf{x}^a_0 - \hat{\mathbf{x}}^a_{i,0|0}\right)^H$

At each time instant $n > 0$:

1: **for** Nodes $i = \{1, 2, \ldots, N\}$ **do**

2: $\widehat{\mathbf{x}}^a_{i,n|n-1} = \left[\mathbf{f}^T\left[\widehat{\mathbf{x}}_{i,n-1|n-1}\right], \mathbf{f}^H\left[\widehat{\mathbf{x}}_{i,n-1|n-1}\right]\right]^T$

3: $\mathbf{F}^a_{i,n-1} = \mathsf{Jacobian}\left(\mathbf{f}^a,\ \widehat{\mathbf{x}}^a_{i,n-1|n-1}\right)$

4: $\mathbf{M}^a_{i,n|n-1} = \mathbf{F}^a_{i,n-1}\mathbf{M}^a_{i,n-1|n-1}\mathbf{F}^{aH}_{i,n-1} + \mathbf{Q}^a_n$

5: $\mathbf{H}^a_{i,n} = \mathsf{Jacobian}\left(\mathbf{h}^a_i,\ \widehat{\mathbf{x}}^a_{i,n|n-1}\right)$

6: $\mathbf{G}^a_{i,n} = \mathbf{M}^a_{i,n|n-1}\underline{\mathbf{H}}^{aH}_{i,n}\left(\underline{\mathbf{H}}^a_{i,n}\mathbf{M}^a_{i,n|n-1}\underline{\mathbf{H}}^{aH}_{i,n} + \underline{\mathbf{R}}^a_{i,n}\right)^{-1}$

7: $\underline{\widehat{\mathbf{x}}}^a_{i,n|n} = \widehat{\mathbf{x}}^a_{i,n|n-1} + \mathbf{G}^a_{i,n}\left(\underline{\mathbf{y}}^a_{i,n} - \mathbf{h}^a_i\left[\widehat{\mathbf{x}}_{i,n|n-1}\right]\right)$

8: $\mathbf{M}^a_{i,n|n} = (\mathbf{I} - \mathbf{G}^a_{i,n}\underline{\mathbf{H}}^a_{i,n})\mathbf{M}^a_{i,n|n-1}$

9: Diffuse the states from the network:
$\widehat{\mathbf{x}}^a_{i,n|n} = \sum_{\ell \in \mathcal{N}_i} c_{\ell i}\underline{\widehat{\mathbf{x}}}^a_{\ell,n|n}$

10: **end for**

Remark 3: The D-ACEKF algorithm in Algorithm 3 extends the Distributed Extended Kalman filter in [45] to widely linear state spaces, and caters for the improper second order statistical moments of the state and noise models, together with the correlations present between the nodal observation noises. This is a perfect match for distributed estimation in unbalanced smart grids.

V. Distributed Widely Linear Frequency Estimation

The proposed augmented state space models are particularly suited for frequency estimation in power grid, as due to system inertia, the frequency can be assumed identical over the network of measurement nodes, while unbalanced systems and distributed nodes generate noncircular and noisy measurements [17], [19]. For a three phase system, the instantaneous voltages at a node i are given by

$$\underbrace{\begin{bmatrix} v_{a,i,n} \\ v_{b,i,n} \\ v_{c,i,n} \end{bmatrix}}_{\mathbf{v}_{i,n}} = \underbrace{\begin{bmatrix} V_{a,i,n}\cos(\omega nT + \phi_{a,i}) \\ V_{b,i,n}\cos(\omega nT + \phi_{b,i} - \frac{2\pi}{3}) \\ V_{c,i,n}\cos(\omega nT + \phi_{c,i} + \frac{2\pi}{3}) \end{bmatrix}}_{\mathbf{s}_{i,n}} + \underbrace{\begin{bmatrix} z_{a,i,n} \\ z_{b,i,n} \\ z_{c,i,n} \end{bmatrix}}_{\mathbf{z}_{i,n}} \tag{22}$$

where $V_{a,i,n}$, $V_{b,i,n}$, and $V_{c,i,n}$ are the amplitudes of the three-phase voltages at time instant n, $\omega = 2\pi f$ the angular frequency, f the system frequency, T the sampling interval, while $z_{a,i,n}$, $z_{a,i,n}$, and $z_{a,i,n}$ are zero-mean observation noise processes. The term $\phi_{a,i}$ is used to denote the phase of the fundamental component, while $\phi_{b,i} = \phi_{a,i} + \Delta_{b,i}$ and $\phi_{c,i} = \phi_{a,i} + \Delta_{c,i}$ are used to indicate the phase distortions relative to a balanced three-phase system. Notice that the only common parameter for all the nodes in the neighbourhood of node i is the digital frequency $2\pi fT$.

A. Clarke Transform for Dimensionality Reduction

Although, the frequency of the system can be estimated directly from any one of the three-phases, utilising the information from all three phases to estimate the frequency is more robust to noise [46]. We would like to estimate the frequency of the signal in (22) by first reducing the dimensionality of the signal from $\mathbb{R}^3$ to $\mathbb{C}$. Instead of using any single-phase voltage, we jointly-represent the three phases by firstly projecting the signal on to an orthogonal basis using the Clarke (or $\alpha\beta$) transform

$$\begin{bmatrix} v_{\alpha,i,n} \\ v_{\beta,i,n} \end{bmatrix} = \underbrace{\sqrt{\frac{2}{3}}\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}}_{\text{ClarkeMatrix}} \begin{bmatrix} v_{a,i,n} \\ v_{b,i,n} \\ v_{c,i,n} \end{bmatrix} \tag{23}$$

thus converting the $\alpha\beta$ voltage vector on the left side of (23) into a complex-valued scalar $v_{i,n} = v_{\alpha,i,n} + jv_{\beta,i,n}$. The transformation of the three-phase voltage vector $\mathbf{v}_{i,n}$ in (22) into a complex-valued scalar can be summarized as

$$\begin{aligned} v_{i,n} &= \mathbf{e}^H\mathbf{v}_{i,n} = \mathbf{e}^H\mathbf{s}_{i,n} + \mathbf{e}^H\mathbf{z}_{i,n} \\ &= s_{i,n} + z_{i,n} \end{aligned} \tag{24}$$

where $\mathbf{e}^H = \sqrt{\frac{2}{3}}\begin{bmatrix}1 & e^{j\frac{2\pi}{3}} & e^{-j\frac{2\pi}{3}}\end{bmatrix}$ and $s_{i,n}$ and $z_{i,n}$ are complex-valued scalars corresponding to the signal and the noise respectively. Using Euler's formula, the noise-free signal $\mathbf{s}_{i,n}$ in (22) can be expressed as

$$\mathbf{s}_{i,n} = \frac{1}{2}\begin{bmatrix}\bar{V}_{a,i,n} & \bar{V}^*_{a,i,n}\\ \bar{V}_{b,i,n} & \bar{V}^*_{b,i,n}\\ \bar{V}_{c,i,n} & \bar{V}^*_{c,i,n}\end{bmatrix}\begin{bmatrix}e^{j\omega nT}\\ e^{-j\omega nT}\end{bmatrix} \tag{25}$$

where the phasors are given by $\bar{V}_{a,i,n} = V_{a,i,n}e^{j\phi_{a,i}}$,

$\bar{V}_{b,i,n} = V_{b,i,n}e^{j\left(\phi_{b,i}-\frac{2\pi}{3}\right)}$, and $\bar{V}_{c,i,n} = V_{c,i,n}e^{j\left(\phi_{c,i}+\frac{2\pi}{3}\right)}$.

Substituting (25) into (24) gives

$$s_{i,n} = \mathbf{e}^H\mathbf{s}_{i,n} = A_{i,n}e^{j\omega nT} + B_{i,n}e^{-j\omega nT} \tag{26}$$

where the scalar phasors $A_{i,n}$ and $B_{i,n}$ are

$$\begin{aligned}A_{i,n} &= \frac{\sqrt{6}}{6}\left[V_{a,i,n}e^{j\phi_{a,i}} + V_{b,i,n}e^{j\phi_{b,i}} + V_{c,i,n}e^{j\phi_{c,i}}\right]\\ B_{i,n} &= \frac{\sqrt{6}}{6}\Big[V_{a,i,n}e^{-j\phi_{a,i}} + V_{b,i,n}e^{-j\left(\phi_{b,i}+\frac{2\pi}{3}\right)}\\ &\quad + V_{c,i,n}e^{-j\left(\phi_{c,i}-\frac{2\pi}{3}\right)}\Big]\end{aligned} \tag{27}$$

and the complex-valued observation noise is

$$z_{i,n} = \mathbf{e}^H\mathbf{z}_{i,n} = \sqrt{\frac{2}{3}}\left[z_{a,i,n} + e^{j\frac{2\pi}{3}}z_{b,i,n} + e^{-j\frac{2\pi}{3}}z_{c,i,n}\right].$$

Therefore, under general operating conditions, the complex-valued voltage in (24) takes the form

$$v_{i,n} = A_{i,n}e^{j\omega nT} + B_{i,n}e^{-j\omega nT} + z_{i,n}. \tag{28}$$

For a balanced system under nominal conditions, $V_{a,i,n} = V_{b,i,n} = V_{c,i,n}$ and $\phi_{a,i} = \phi_{b,i} = \phi_{c,i}$, so that the coefficient $B_{i,n}$ vanishes, see (27), and the complex-valued voltage is

$$v_{i,n} = s_{i,n} + z_{i,n} = A_{i,n}e^{j\omega nT} + z_{i,n}. \tag{29}$$

Moreover, the signal $s_{i,n}$ in (29) can be expressed as an autoregressive (AR) model[4] $s_n = e^{j\omega T}s_{n-1}$ such that at each node i, the voltage measurement is

$$v_{i,n} = e^{j\omega T}s_{i,n-1} + z_{i,n}. \tag{30}$$

Fig. 2 illustrates that for a balanced system, for which $V_{a,i,n} = V_{b,i,n} = V_{c,i,n}$ and $\phi_{a,i} = \phi_{b,i} = \phi_{c,i}$, the Clarke's voltage $v_{i,n}$ in (30) has a circular trajectory, thus making the signal circular. The *system* in (30) that generates the circular signal is *strictly linear* because the output of the system is only a function of the input $s_{i,n-1}$ and not its conjugate $s^*_{i,n-1}$.

Strictly Linear State Space Model 1. It is based on the AR(1) voltage evolution model in (30) whose strictly linear state space model for the balanced system system at a node i is given in (32a) and (32b), where the state variables are $h_n \overset{\text{def}}{=} e^{j\omega T}$ and

[4]The usual assumption in this type of estimation is that for a sampling frequency $>>$ 50Hz, we have $A_n \approx A_{n-1}$.

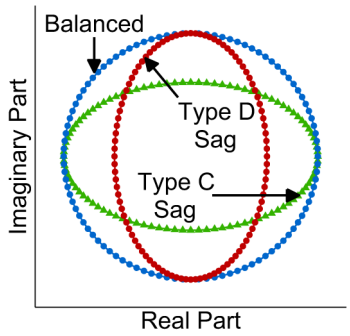


Fig. 2. For a balanced system, characterised by $V_{a,i,n} = V_{b,i,n} = V_{c,i,n}$ and $\Delta_{b,i} = \Delta_{c,i} = 0$, the trajectory of Clarke's voltage $v_{i,n}$ is circular (blue line). For unbalanced systems, the voltage trajectories are noncircular (red and green lines).

the noise-free signal $s_{i,n}$, while $\mathbf{w}_n$ and $z_{i,n}$ are respectively the state and observation noises. The system frequency is then derived from the state variable h_n as

$$\hat{f}_n = \frac{1}{2\pi T}\arctan\left(\frac{\mathsf{Im}\{h_n\}}{\mathsf{Re}\{h_n\}}\right) \tag{31}$$

where $\mathsf{Re}\{\cdot\}$ and $\mathsf{Im}\{\cdot\}$ are respectively the real and imaginary parts of a complex variable.

State Space Model 1. Strictly Linear State Space (StSp-SL)

1: State Equation

$$\begin{bmatrix}h_n\\ s_{i,n}\end{bmatrix} = \begin{bmatrix}h_{n-1}\\ s_{i,n-1}h_{n-1}\end{bmatrix} + \mathbf{w}_n \tag{32a}$$

2: Observation Equation

$$v_{i,n} = \begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix}h_n\\ s_{i,n}\end{bmatrix} + z_{i,n} \tag{32b}$$

The strictly linear system model in (30) is inaccurate when the system is operating under unbalanced conditions, such as when the voltage amplitudes $V_{a,i,n}$, $V_{b,i,n}$ and $V_{c,i,n}$ are no longer equal or if the condition $\Delta_{b,i} = \Delta_{c,i} = 0$ is not satisfied. In these cases, $B_{i,n}$ in (27) is not zero and the signal becomes noncircular (ellipse in Fig. 2). Therefore, this standard strictly linear estimator introduces biased estimates that oscillate at twice the system frequency for unbalanced system conditions [17].

Strictly Linear AR(2) based State Space Model. Without assuming a balanced operating condition, the $\alpha\beta$ voltage in (26) can be expressed as a *strictly linear* AR(2) model [47]–[50]

$$s_{i,n} = 2\cos(\omega T)s_{i,n-1} - s_{i,n-2} \tag{33}$$

The system frequency can be derived by defining the state as $2\cos(\omega T)$ and estimating it using a complex Kalman filter. This strictly linear AR(2) model is referred to as the strictly linear "three-point" (SL3PT) model and serves as a benchmark for our proposed method, as it can be applied

to both any single-phase voltage and the complex-valued $\alpha\beta$ voltage. A major disadvantage to the three-point model in (33) is that it produces biased estimates in the presence of noise [51].

Widely Linear AR(1) based State Space Model 2. Observe that the $\alpha\beta$ voltage in (26) can be interpreted as the sum of two phasors, one rotating clockwise and the other rotating counter clockwise at the same frequency.

If we were to express the signal (26) in an autoregressive form, it is only natural and intuitive to consider the previous value $s_{i,n-1}$ *and* its conjugate $s^*_{i,n-1}$, where the conjugate represents the phasor rotating in the opposite direction. The widely linear model for the complex-valued $\alpha\beta$ voltage at any node i is therefore given by [16]–[18]

$$s_{i,n} = h_{i,n-1}s_{i,n-1} + g_{i,n-1}s^*_{i,n-1}. \tag{34}$$

This is a first-order widely linear autoregressive model with coefficients $h_{i,n}$ and $g_{i,n}$. By substituting the values of $s_{i,n-1}$ and $s^*_{i,n-1}$ from (26) into (34) and some algebraic manipulation, we obtain [17]

$$h_{i,n} = e^{j\omega T} - \frac{B^*_{i,n}}{A_{i,n}}g_{i,n}, \quad g_{i,n} = \frac{B_{i,n}}{A^*_{i,n}}\left(e^{-j\omega T} - h_{i,n}\right). \tag{35}$$

Solving the two equations in (35) for the system frequency yields

$$\hat{f}_n = \frac{1}{2\pi T}\arctan\left(\frac{\sqrt{\mathsf{Im}^2\{h_{i,n}\} - |g_{i,n}|^2}}{\mathsf{Re}\{h_{i,n}\}}\right). \tag{36}$$

From (35), notice that when the system is balanced ($B_{i,n} = 0$), the coefficient $g_{i,n} = 0$, and the widely linear frequency estimate in (36) is identical to its strictly linear counterpart in (31). While the strictly linear AR(2) model in (33) is identical for both balanced and unbalanced voltages, the widely linear model provides an intuitive advantage as the coefficient $g_{i,n}$ represents the negative sequence which characterises the imbalance of the system voltage.

The state space model in (37a) and (37b) provides a realistic and robust characterisation of real world power systems, as it represents both balanced and unbalanced systems, while its nonlinear state equation also models the coupling between state variables. State Space Model 2 can be implemented using the proposed distributed augmented complex extended Kalman filter in Section IV.

State Space Model 2 Widely Linear State Space (StSp-WL)

1: State Equation

$$\begin{bmatrix} h_{i,n} \\ g_{i,n} \\ s_{i,n} \\ h^*_{i,n} \\ g^*_{i,n} \\ s^*_{i,n} \end{bmatrix} = \begin{bmatrix} h_{i,n-1} \\ g_{i,n-1} \\ s_{i,n-1}h_{i,n-1} + s^*_{i,n-1}g_{i,n-1} \\ h^*_{i,n-1} \\ g^*_{i,n-1} \\ s^*_{i,n-1}h^*_{i,n-1} + s_{i,n-1}g^*_{i,n-1} \end{bmatrix} + \mathbf{w}_n \tag{37a}$$

2: Observation Equation

$$\begin{bmatrix} v_{i,n} \\ v^*_{i,n} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_{i,n} \\ g_{i,n} \\ s_{i,n} \\ h^*_{i,n} \\ g^*_{i,n} \\ s^*_{i,n} \end{bmatrix} + \begin{bmatrix} z_{i,n} \\ z^*_{i,n} \end{bmatrix} \tag{37b}$$

B. Diffusion-step

The classical diffusion scheme in (16) obtains the weighted average of the estimated states $\widehat{\mathbf{x}}^a_{i,n|n}$ in the neighbourhood $\mathcal{N}_i$ under the assumption that each node is estimating the same optimal state. However, the state estimate of the State Space Model 2, which is given by

$$\widehat{\mathbf{x}}^a_{i,n|n} = \left[\widehat{h}_{i,n|n},\, \widehat{g}_{i,n|n},\, \widehat{s}_{i,n|n},\, \widehat{h}^*_{i,n|n},\, \widehat{g}^*_{i,n|n},\, \widehat{s}^*_{i,n|n}\right]^T \tag{38}$$

is unique at each node since the coefficients $h_{i,n}$ and $g_{i,n}$ are functions of both system frequency ω (which is identical in the network) and the level of imbalance, $\frac{B_{i,n}}{A_{i,n}}$ (which is not necessarily identical in the network) – see (35). Applying the diffusion scheme without considering this phenomenon will therefore result in spurious frequency estimates when the level of system imbalance is different at each node.

This problem is circumvented by the fact that the state $\widehat{\mathbf{x}}^a_{i,n|n}$ also includes the noise-free signal estimate $\widehat{s}_{i,n|n}$. When all the nodes are connected and weighed equally, each node effectively estimates the frequency from the *diffused* signal estimate $\widehat{s}_{i,n|n}$ formed by diffusing the signal estimates $\underline{\widehat{s}}_{\ell,n|n}$ from each node

$$\widehat{s}_{i,n|n} = \frac{1}{|\mathcal{N}_i|}\sum_{\ell\in\mathcal{N}_i}\underline{\widehat{s}}_{\ell,n|n} = \widehat{A}_{i,n}e^{j\widehat{\omega}nT} + \widehat{B}_{i,n}e^{-j\widehat{\omega}nT} \tag{39}$$

where $\widehat{A}_{i,n} = \frac{1}{|\mathcal{N}_i|}\sum_{\ell\in\mathcal{N}_i}\underline{\widehat{A}}_{\ell,n}$ and $\widehat{B}_{i,n} = \frac{1}{|\mathcal{N}_i|}\sum_{\ell\in\mathcal{N}_i}\underline{\widehat{B}}_{\ell,n}$ are the diffused phasors in the neighbourhood $\mathcal{N}_i$.

Remark 4: Since the frequency is derived from the *diffused* signal estimate in (39), which is now identical throughout the network, our proposed D-ACEKF yields the correct frequency estimate even with different levels of imbalance at different nodes.

VI. Frequency Estimation Examples

A. Experiment (Synthetic Data) Set-Up

The simulations were based on a network of 5 substations (nodes) where each substation has access to three-phase voltage measurements via transformers with metering capabilities. Without loss in generality, we used the distributed network topology shown in Fig. 3. The power system under consideration had a nominal frequency of 50Hz, and was sampled at a rate of 5kHz while the signal to noise ratio (SNR) was determined by the metering accuracy class of the potential transformer. The BS EN 61869-1:2009 standard for the metering accuracy of potential transformers classifies six separate

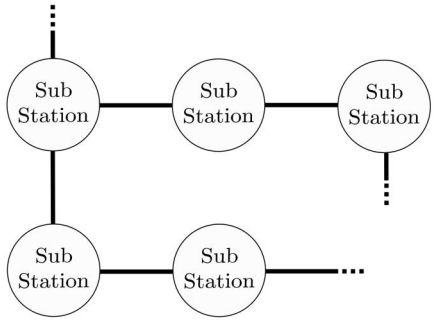


Fig. 3. A distributed power network with $N = 5$ nodes (Sub-stations) used in the simulations.

classes for metering requirements, which translates to an SNR range of 30 dB to 60 dB [52]. To illuminate the robustness of our proposed augmented diffusion Kalman filters, we have chosen an SNR level of 25 dB in all our simulations, unless stated otherwise.

B. Algorithms

The proposed D-ACEKF (Algorithm 3) was benchmarked against the D-CEKF which uses State Space Model 1 and the distributed strictly linear three point (D-SL3PT) algorithm which uses the signal model in (33) with the D-CKF in Algorithm 2. The D-SL3PT was chosen since it is a well-know classical algorithm which employs a similar principle to the D-ACEKF (i.e. estimating the frequency with an autoregressive model of the signal). Single node (uncooperative) estimates of the algorithms are also presented for completeness.

Case Study #1: Voltage sags. In the first set of simulations, the performances of the algorithms were evaluated for an initially balanced system which became unbalanced after undergoing a Type C voltage sag starting at 0.1s, characterised by a 20% voltage drop and 10^o phase offset on both the v_b and v_c channels, followed by a Type D sag starting at 0.3s, characterised by a 20% voltage drop at line v_a and a 10% voltage drop on both v_b and v_c with a 5^o phase angle offset. The phasor trajectories and degrees of noncircularity of these system imbalances are illustrated in Fig. 4.

Fig. 5 shows that, conforming with the analysis, the widely linear algorithms, ACEKF and D-ACEKF, were able to converge to the correct system frequency for both balanced and unbalanced operating conditions, while the strictly linear algorithms, CEKF and D-CEKF, were unable to accurately estimate the frequency during the voltage sag due to under-modeling of the system (not accounting for its widely linear nature) – see (34). As expected, the widely linear and strictly linear algorithms had similar performances under balanced conditions, as illustrated in the time interval 0–0.1 s. The distributed algorithms, D-CEKF and D-ACEKF, outperformed their uncooperative counterparts, CEKF and ACEKF, owing to the sharing of information between neighbouring nodes.

Case Study #2: Presence of switching noise. Fig. 6 illustrates frequency estimation in the presence of random spike noise, which models the switching noise from the inverters which interface the renewable sources to the grid. The

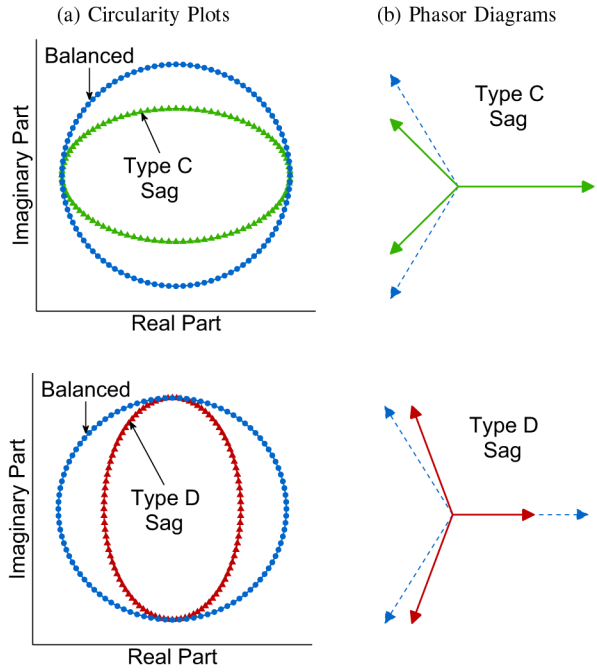


Fig. 4. Circularity (left) and phasor (right) views of Type C and D unbalanced voltage sags. The real-imaginary phasor plots illustrate the noncircularity of Clarke's voltage in unbalanced conditions, indicated by the elliptical shapes of circularity plots. The eccentricity of this ellipse (degree of noncircularity) serve to identify the type of fault (in this case a voltage sag).

distributed algorithms, D-CEKF and D-ACEKF, outperformed their uncooperative counterparts, CEKF and ACEKF, because neighbouring nodes were able to share information to facilitate better estimation performances.

Case Study #3: Bias and variance of the proposed estimators. Fig. 7 illustrates the bias and variance for the proposed distributed frequency estimators. The steady state frequency estimate at a node i for the trial m is given by $\hat{f}_{i,ss}[m]$. The bias and variance of the frequency estimators were calculated over 2000 independent trials using

$$\begin{aligned} \text{Bias} &= \frac{1}{2000 \cdot N} \sum_{m=1}^{2000} \sum_{i=1}^{N} \hat{f}_{i,ss}[m] - f_0 \\ \text{Variance} &= \frac{1}{2000 \cdot N} \sum_{m=1}^{2000} \sum_{i=1}^{N} \left(\hat{f}_{i,ss}[m] - \bar{f}_{i,ss} \right)^2 \end{aligned} \tag{40}$$

where $N = \{1, 5\}$ for single-node and distributed cases respectively, while $f_0 = 50$Hz is the fundamental frequency. The sample mean was evaluated as $\bar{f}_{i,ss} = \frac{1}{2000} \sum_{m=1}^{2000} \hat{f}_{i,ss}[m]$. The algorithms were evaluated at different SNR levels for an unbalanced system undergoing a Type D voltage sag. Observe that both the single- and multiple-node widely linear algorithms, ACEKF and D-ACEKF, were asymptotically unbiased (left panel) while both the single- and multiple-node strictly linear algorithms were biased. In terms of the variance of the estimators (right panel), both the distributed estimation algorithms outperformed their non-cooperative counterparts, while the only consistent estimator was the proposed D-ACEKF.

Case Study #4: Benchmark against classical methods. The D-ACEKF was compared to the single-node (uncooperative)

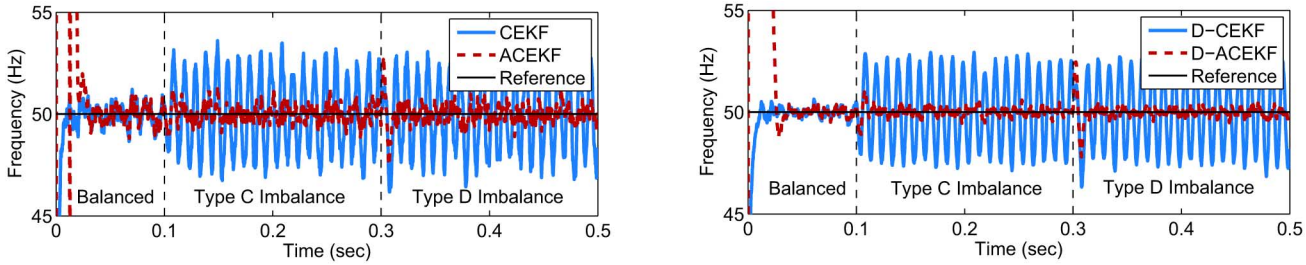


Fig. 5. Frequency estimation performance of single node (CEKF and ACEKF) and distributed (D-CEKF and D-ACEKF) algorithms for a system at 25dB SNR. The system is balanced up to 0.1s, it then undergoes a Type C voltage imbalance followed by a Type D voltage imbalance at 0.3s.

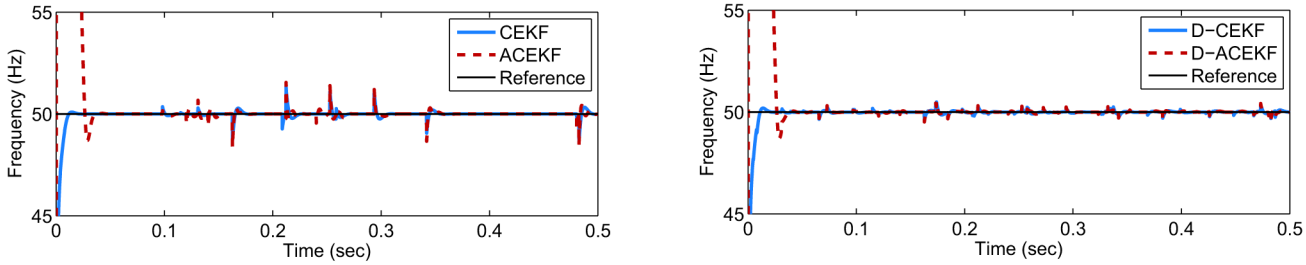


Fig. 6. Frequency estimation performance of single node (CEKF and ACEKF) and distributed (D-CEKF and D-ACEKF) algorithms when the phase voltages of three nodes in the network are contaminated with random spike noise.

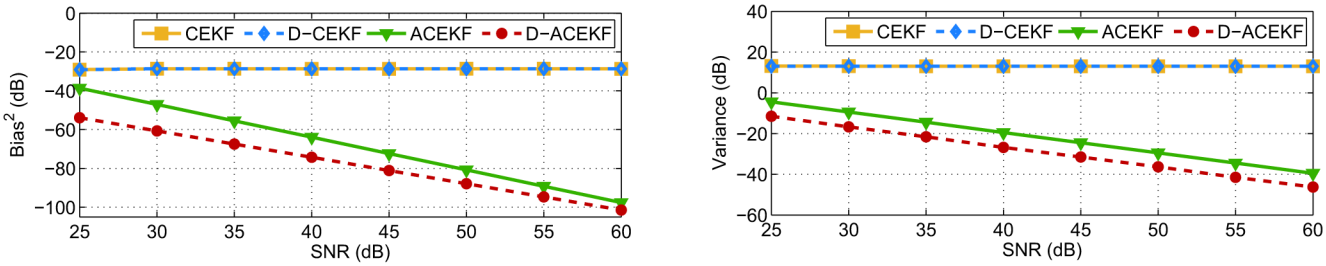


Fig. 7. Bias and variance analysis of the proposed distributed state space frequency estimators for an unbalanced system undergoing a Type D voltage sag. *Left:* Estimation bias. *Right:* Estimation variance.

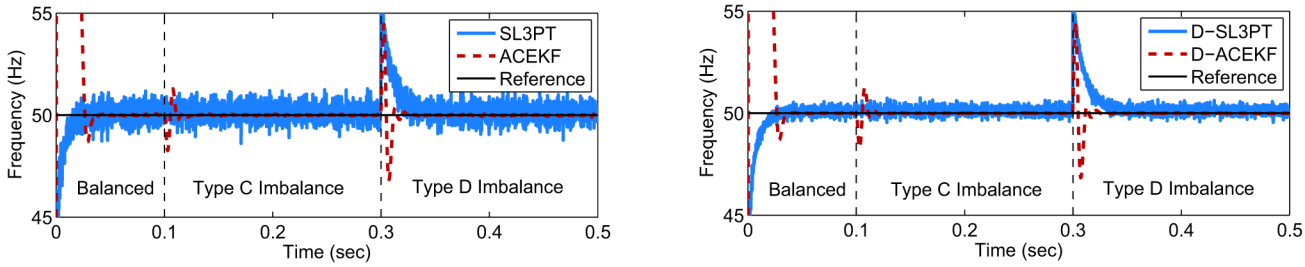


Fig. 8. Frequency estimation performance of single node (SL3PT and ACEKF) and distributed (D-SL3PT and D-ACEKF) algorithms for a system at 50dB SNR. The system is balanced up to 0.1s, it then undergoes a Type C voltage imbalance followed by a Type D voltage imbalance at 0.3s.

and distributed implementations of the classical strictly linear three-point (SL3PT) model given in (33). The set-up in Case Study #1 was used with the exception that the SNR was increased to 50 dB as the SL3PT was performing poorly at lower SNRs. For a fair comparison, the D-SL3PT was configured to have the same convergence speed as the D-ACEKF. From Fig. 8, we can observe that the D-SL3PT produced noisier estimates compared to the D-ACEKF when their adaptation speeds were equalised.

Next, the bias and variance of the D-ACEKF were benchmarked against the D-SL3PT using the methodology in (40). As seen in Fig. 9, while the variance of the frequency estimates decreased with the diffusion scheme, the bias of the SL3PT model did not. As mentioned in Section V, the SL3PT and D-SL3PT both exhibit high bias due to modelling inaccuracies in the presence of noise [51]. This shows the markedly increased robustness to noise of the widely linear model compared to the strictly linear AR(2) model.

Case Study #5: Different Voltage Sags. Fig. 10 shows the profile of the voltages at different nodes in the network. Each substation underwent different faults, which is reflected by the different relative amplitudes and phase shifts of the $\alpha\beta$

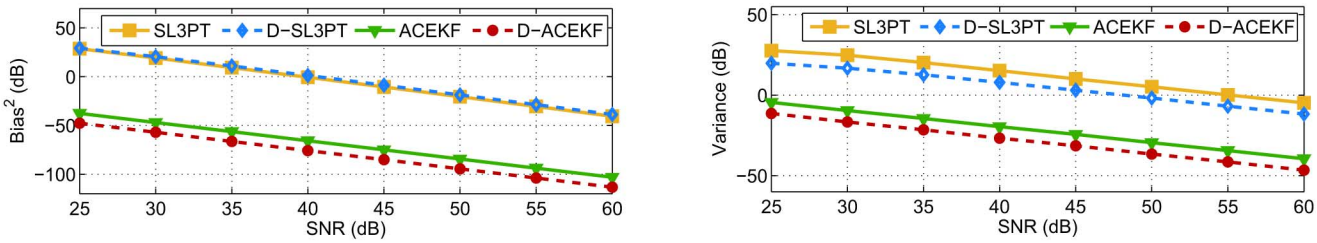


Fig. 9. Bias and variance analysis of the proposed distributed state space frequency estimators against classical strictly linear three-point (SL3PT) models for an unbalanced system undergoing a Type D voltage sag. *Left:* Estimation bias. *Right:* Estimation variance.

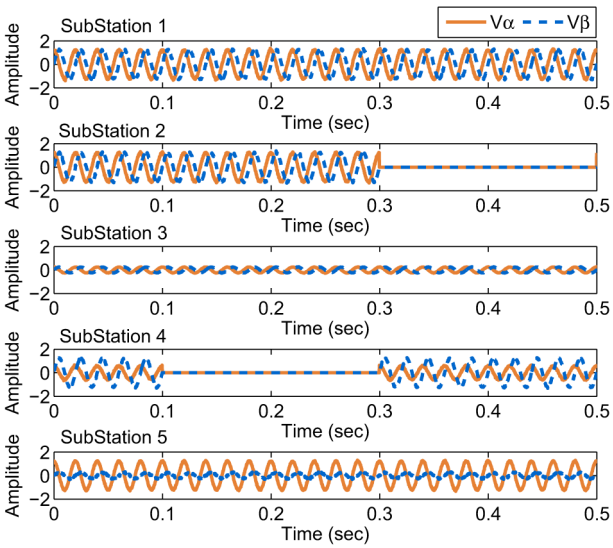


Fig. 10. Each substation (node) has a different $\alpha\beta$ voltages including cases where the voltage drops to zero (line cut).

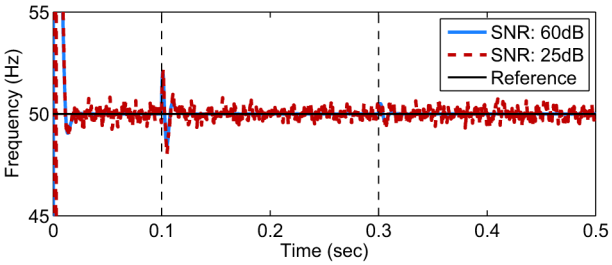


Fig. 11. The D-ACEKF is able to estimate the frequency in a distributed setting with different types of faults at each node, see Fig. 10 for voltage profiles.

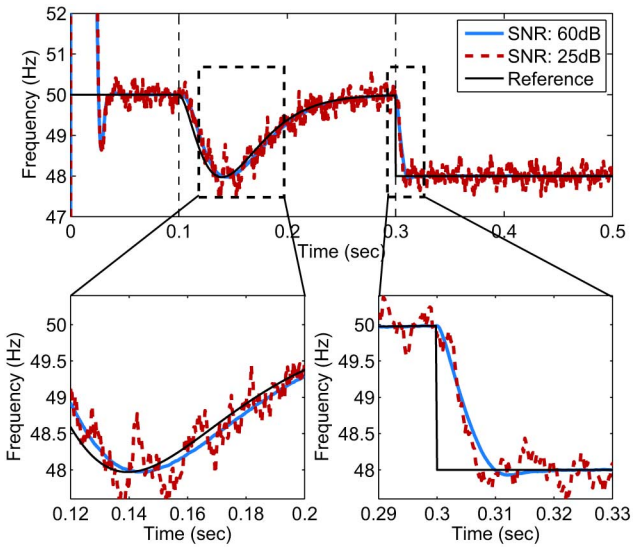


Fig. 12. Frequency tracking performance of D-ACEKF at 25dB and 60dB SNR, which experiences a gradual change fro, 0.1s to 0.3s and a step change in system frequency to 48Hz at 0.3s. The D-ACEKF is able to track both slow and rapid changes in frequency.

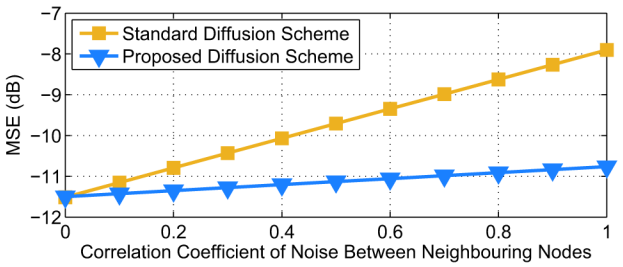


Fig. 13. The mean square error (MSE) of the D-ACEKF which shares only the state estimates (standard diffusion scheme) and full D-ACEKF which shares both state estimates and observations (proposed diffusion scheme) with various degrees of correlation between three nodes in the network.

voltages. In addition Substations 2 and 4 underwent total line failures from 0.3s to 0.5s and 0.1s to 0.3s respectively.

Fig. 11 shows that due to its unique state space structure, the proposed D-ACEKF was able to estimate the frequency of the network even with different levels of imbalance at each node, see Remark #4 in Section V-B.

Case Study #6: Frequency tracking. Fig. 12 illustrates the performance the D-ACEKF when a power network was contaminated with white noise at 25dB and 60dB SNR and underwent a gradual drop and increase in frequency from 0.1s to 0.3s. This is a typical scenario when generation does not match the load and system inertia keeps the frequency from changing too quickly. From 0.3s to 0.5s the system undergoes a step-change in frequency; this is the scenario when there is not enough inertial response in the system caused by the influx of electronic inverters that are replacing synchronous machines. The D-ACEKF was able to track the frequency in both cases illustrating that its suitability for both the current electricity grid and future smart grids.

Case Study #7: Noise Correlation. Next, the D-ACEKF was evaluated for the case where the observation noises between three nodes were correlated, exhibiting various degrees of correlation. The proposed D-ACEKF was compared to the D-ACEKF with the classical diffusion scheme where only the

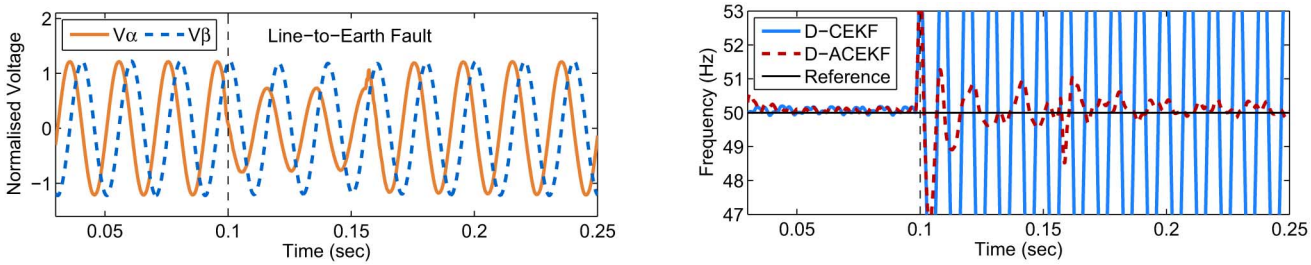


Fig. 14. Real world case study. *Left*: The $\alpha\beta$ voltages at Sub-station 1 before and during the fault event. *Right*: Frequency estimation using the proposed algorithms before and during the fault event.

states were shared between the nodes. As seen in (20), the proposed diffusion scheme shares both observations $v_{i,n}$ and state estimates $\widehat{\mathbf{x}}_{i,n|n}$, which allows the proposed diffusion scheme to exploit cross-nodal correlations of the observation noise. Fig. 13 shows the mean square error (MSE) of the frequency estimates where $\mathrm{MSE} = \mathrm{Bias}^2 + \mathrm{Variance}$, and the Bias and Variance terms are defined in (40). The proposed D-ACEKF outperformed the classical diffusion scheme especially when the cross-nodal correlations of the noise were large.

Real World Case Study. We next assessed the performance of the proposed algorithms on a real world case study using three-phase voltage measurements from two adjacent sub-stations in Malaysia during a brief line-to-earth fault on the 29th June 2014. This caused voltage sags, similar to those in Case Study #1. The three-phase measurements were sampled at 5 kHz and the voltage values were normalized. The left panel in Fig. 14 shows the normalized $\alpha\beta$ voltages at one of the sub-stations. The fault that occurs in phase A around 0.1s is reflected in the voltage dip in v_α. The right panel in Fig. 14 shows the frequency estimate from the D-ACEKF and D-CEKF. Conforming with the analysis and the single node scenario in Fig. 5, the collaborative widely linear D-ACEKF was able to track the real world frequency of a power network under both balanced and unbalanced conditions, whereas the strictly linear D-CEKF was unable to track the frequency after 0.1s when the line-to-earth fault (non-circularity) occurred.

VII. Conclusion

We have proposed a novel class of diffusion based distributed complex-valued Kalman filters for cooperative frequency estimation in power networks. To cater for the general case of improper states, observations, and state and observation noises, we have introduced the distributed (widely linear) augmented complex Kalman filter (D-ACKF) and its nonlinear version, the distributed augmented complex Kalman filter (D-ACEKF). These have been shown to provide sequential state estimation of the generality of complex signals, both circular and noncircular, within a general and unifying framework which also caters for correlated nodal observation noises. This novel widely linear framework has been applied for distributed state space based frequency estimation in the context of three-phase power systems, and has been shown to be optimal for both balanced and unbalanced operating conditions. Simulations over a range of balanced and unbalanced power system conditions and for both synthetic and real world measurements have illustrated that the proposed distributed state space algorithms are consistent estimators, offering accurate and fast frequency estimation in both balanced an unbalanced system conditions.

Appendix

A. Bias Analysis of the D-ACKF Estimates

To analyse the mean performance of the D-ACKF (Algorithm 2), we employ the methodologies in [22][30]. First consider the local (non-diffused) error at node $i \in [1, N]$

$$\underline{\mathbf{e}}^a_{i,n|n} \overset{\text{def}}{=} \mathbf{x}^a_n - \underline{\widehat{\mathbf{x}}}^a_{i,n|n}. \tag{41}$$

Substituting Step 2.5 in Algorithm 2, into the non-diffused error (41) gives

$$\underline{\mathbf{e}}^a_{i,n|n} = \mathbf{x}^a_n - \widehat{\mathbf{x}}^a_{i,n|n-1} - \mathbf{G}^a_{i,n}\left(\underline{\mathbf{y}}^a_{i,n} - \underline{\mathbf{H}}^a_{i,n}\widehat{\mathbf{x}}^a_{i,n|n-1}\right). \tag{42}$$

Using the optimal observation equation $\mathbf{y}^a_n$ in (10), this non-diffused error can be expressed as

$$\underline{\mathbf{e}}^a_{i,n|n} = \left(\mathbf{I} - \mathbf{G}^a_{i,n}\underline{\mathbf{H}}^a_{i,n}\right)\mathbf{e}^a_{i,n|n-1} - \mathbf{G}^a_{i,n}\underline{\mathbf{v}}^a_{i,n}. \tag{43}$$

Now consider the diffused *a priori* error $\mathbf{e}^a_{i,n|n-1} \overset{\text{def}}{=} \mathbf{x}^a_n - \widehat{\mathbf{x}}^a_{i,n|n-1}$. Subtracting $\mathbf{x}^a_n$ in (10) from Step 2.2 in Algorithm 2 yields the expression for the diffused *a priori* error as

$$\mathbf{e}^a_{i,n|n-1} = \mathbf{F}^a_{n-1}\mathbf{e}^a_{i,n-1|n-1} + \mathbf{w}^a_n. \tag{44}$$

Finally, consider the diffused *a posteriori* error $\mathbf{e}^a_{i,n|n} \overset{\text{def}}{=} \mathbf{x}^a_n - \widehat{\mathbf{x}}^a_{i,n|n}$ which can be expressed in terms of the local (non-diffused) errors in (41) using the diffusion step in (16) as

$$\mathbf{e}^a_{i,n|n} = \sum\nolimits_{\ell\in\mathcal{N}_i} c_{\ell i}\underline{\mathbf{e}}^a_{\ell,n|n}. \tag{45}$$

Substituting (43) and (44) into (45) and using $\mathbf{M}^a_{\ell,n|n}(\mathbf{M}^a_{\ell,n|n-1})^{-1} = \mathbf{I} - \mathbf{G}^a_{\ell,n}\underline{\mathbf{H}}^a_{\ell,n}$ from the covariance update in Step 2.6 in Algorithm 2, we have

$$\mathbf{e}^a_{i,n|n} = \sum_{\ell\in\mathcal{N}_i} c_{\ell i}\Big[\mathbf{M}^a_{\ell,n|n}(\mathbf{M}^a_{\ell,n|n-1})^{-1}\mathbf{F}^a_{n-1}\mathbf{e}^a_{\ell,n-1|n-1} + \mathbf{M}^a_{\ell,n|n}(\mathbf{M}^a_{\ell,n|n-1})^{-1}\mathbf{w}^a_n - \mathbf{G}^a_{\ell,n}\underline{\mathbf{v}}^a_{\ell,n}\Big]. \tag{46}$$

Upon taking the statistical expectation $\mathsf{E}\{\cdot\}$, the recursion in (46) leads to a closed form expression for the mean error of the D-ACKF algorithm, given by

$$\mathsf{E}\left\{\mathbf{e}^a_{i,n|n}\right\} = \sum_{\ell\in\mathcal{N}_i} c_{\ell i}\mathbf{M}^a_{\ell,n|n}(\mathbf{M}^a_{\ell,n|n-1})^{-1}\mathbf{F}^a_{n-1}\mathsf{E}\left\{\mathbf{e}^a_{\ell,n-1|n-1}\right\} = \mathbf{0}. \tag{47}$$

Remark 5: The D-ACKF is an unbiased estimator of both proper and improper complex random signals.

ACKNOWLEDGMENT

The authors would like to thank G. K. Supramaniam from TNB Malaysia for his insights and for providing us with the real-world measurements.

REFERENCES

[1] R. Piwko *et al.*, "Interconnection requirements for variable generation," North American Electric Reliability Corporation (NERC), Atlanta, GA, USA, Tech. Rep., 2012.

[2] A. Junyent Ferre, Y. Pipelzadeh, and T. Green, "Blending HVDC-link energy storage and offshore wind turbine inertia for fast frequency response," *IEEE Trans. Sustain. Energy*, vol. 6, no. 3, pp. 1059–1066, Jul. 2015.

[3] M. H. J. Bollen *et al.*, "Bridging the gap between signal and power," *IEEE Signal Process. Mag.*, vol. 26, no. 4, pp. 11–31, Jul. 2009.

[4] M. Wang and Y. Sun, "A practical, precise method for frequency tracking and phasor estimation," *IEEE Trans. Power Del.*, vol. 19, no. 4, pp. 1547–1552, Oct. 2004.

[5] T. Lobos and J. Rezmer, "Real-time determination of power system frequency," *IEEE Trans. Instrum. Meas.*, vol. 46, no. 4, pp. 877–881, Aug. 1997.

[6] T. Routtenberg and L. Tong, "Joint frequency and phasor estimation under the KCL constraint," *IEEE Signal Process. Lett.*, vol. 20, no. 6, pp. 575–578, Jun. 2013.

[7] T. Routtenberg and L. Tong, "Joint frequency and phasor estimation in unbalanced three-phase power systems," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2014, pp. 2982–2986.

[8] Y. Xiao, R. K. Ward, L. Ma, and A. Ikuta, "A new LMS-based Fourier analyzer in the presence of frequency mismatch and applications," *IEEE Trans. Circuits Syst.*, vol. 52, no. 1, pp. 230–245, Jan. 2005.

[9] C. H. Huang, C. H. Lee, K. J. Shih, and Y. J. Wang, "Frequency estimation of distorted power system signals using a robust algorithm," *IEEE Trans. Power Del.*, vol. 23, no. 1, pp. 41–51, Jan. 2008.

[10] P. K. Dash, A. K. Pradhan, and G. Panda, "Frequency estimation of distorted power system signals using extended complex Kalman filter," *IEEE Trans. Power Del.*, vol. 14, no. 3, pp. 761–766, Jul. 1999.

[11] Ying Hu, A. Kuh, A. Kavcic, and D. Nakafuji, "Real-time state estimation on micro-grids," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2011, pp. 1378–1385.

[12] H. Mangesius, S. Hirche, and D. Obradovic, "Quasi-stationarity of electric power grid dynamics based on a spatially embedded Kuramoto model," in *Proc. Amer. Control Conf. (ACC)*, 2012, pp. 2159–2164.

[13] W. C. Duesterhoeft, M. W. Schulz, and E. Clarke, "Determination of instantaneous currents and voltages by means of alpha, beta, and zero components," *Trans. Amer. Inst. Elect. Eng.*, vol. 70, no. 2, pp. 1248–1255, 1951.

[14] G. C. Paap, "Symmetrical components in the time domain and their application to power network calculations," *IEEE Trans. Power Syst.*, vol. 15, no. 2, pp. 522–528, May 2000.

[15] E. Clarke, *Circuit Analysis of A.C. Power Systems*. Hoboken, NJ, USA: Wiley, 1943.

[16] Y. Xia and D. P. Mandic, "Widely linear adaptive frequency estimation of unbalanced three-phase power systems," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 1, pp. 74–83, Jan. 2012.

[17] Y. Xia, S. C. Douglas, and D. P. Mandic, "Adaptive frequency estimation in smart grid applications: Exploiting noncircularity and widely linear adaptive estimators," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 44–54, Sep. 2012.

[18] D. H. Dini and D. P. Mandic, "Widely linear modeling for frequency estimation in unbalanced three-phase power systems," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 2, pp. 353–363, Feb. 2013.

[19] D. Mandic, Y. Xia, and D. Dini, "WO2014053610-frequency estimation," WO Patent App. PCT/EP2013/070,654, Apr. 2014.

[20] P. A. Stadter *et al.*, "Confluence of navigation, communication, and control in distributed spacecraft systems," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 17, no. 5, pp. 26–32, May 2002.

[21] D. P. Mandic *et al.*, *Signal Processing Techniques for Knowledge Extraction and Information Fusion*. New York, NY, USA: Springer, 2008.

[22] F. S. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Trans. Autom. Control*, vol. 55, no. 9, pp. 2069–2084, Sep. 2010.

[23] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 401–420, Mar. 2006.

[24] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proc. 46th IEEE Conf. Decis. Control*, Dec. 2007, pp. 5492–5498.

[25] 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.

[26] 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, Nov. 2011.

[27] S. Kanna, S. Talebi, and D. Mandic, "Diffusion widely linear adaptive estimation of system frequency in distributed power grids," in *Proc. IEEE Int. Energy Conf. (ENERGYCON)*, May 2014, pp. 772–778.

[28] U. A. Khan and J. M. F. Moura, "Distributing the Kalman filter for large-scale systems," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4919–4935, Oct. 2008.

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

[30] D. H. Dini and D. P. Mandic, "A class of widely linear complex Kalman filters," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 5, pp. 775–786, May 2012.

[31] Z. Gao, H.-Q. Lai, and K. J. R. Liu, "Differential space-time network coding for multi-source cooperative communications," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3146–3157, Nov. 2011.

[32] Y. Mao and M. Wu, "Tracing malicious relays in cooperative wireless communications," *IEEE Trans. Inf. Forensics Secur.*, vol. 2, no. 2, pp. 198–212, Jun. 2007.

[33] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 411–420, Feb. 1997.

[34] J. Navarro-Moreno, "ARMA prediction of widely linear systems by using the innovations algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3061–3068, Jul. 2008.

[35] J. Eriksson and V. Koivunen, "Complex random vectors and ICA models: Identifiability, uniqueness, and separability," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1017–1029, Mar. 2006.

[36] J. Navarro-Moreno, J. Moreno-Kayser, R. M. Fernandez-Alcala, and J. C. Ruiz-Molina, "Widely linear estimation algorithms for second-order stationary signals," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4930–4935, Dec. 2009.

[37] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge, U.K.: Cambridge Univ. Press, 2010.

[38] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. Hoboken, NJ, USA: Wiley, 1996.

[39] S. Kar and J. M. F. Moura, "Gossip and distributed Kalman filtering: Weak consensus under weak detectability," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1766–1784, Apr. 2011.

[40] R. Bru, L. Elsner, and M. Neumann, "Convergence of infinite products of matrices and inner-outer iteration schemes," *Electron. Trans. Numer. Anal.*, vol. 2, pp. 183–193, Dec. 1994.

[41] X. R. Li, Y. Zhu, J. Wang, and C. Han, "Optimal linear estimation fusion: Part I: Unified fusion rules," *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2192–2208, Sep. 2003.

[42] A. H. Sayed, "Diffusion Adaptation over Networks," in *Academic Press Library in Signal Processing*. MA, USA: Elsevier, 2014, vol. 3, pp. 323–454.

[43] D. H. Dini and D. P. Mandic, "Cooperative adaptive estimation of distributed noncircular complex signals," in *Proc. 46th Asilomar Conf. Signals Syst. Comput. (ASILOMAR)*, 2012, pp. 1518–1522.

[44] D. H. Dini and D. P. Mandic, "Widely linear complex extended Kalman filters," in *Proc. Sensor Signal Process. Defence Conf.*, 2011, pp. 1–5.

[45] F. Cattivelli and A. Sayed, "Distributed nonlinear Kalman filtering with applications to wireless localization," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Mar. 2010, pp. 3522–3525.

[46] M. Canteli, A. Fernandez, L. Eguiluz, and C. Estebanez, "Three-phase adaptive frequency measurement based on Clarke's transformation," *IEEE Trans. Power Del.*, vol. 21, no. 3, pp. 1101–1105, Jul. 2006.

[47] L. Li, W. Xia, D. Shi, and J. Li, "Frequency estimation on power system using recursive-least-squares approach," in *Proc. Int. Conf. Inf. Technol. Softw. Eng.*, 2013 vol. 211, pp. 11–18.

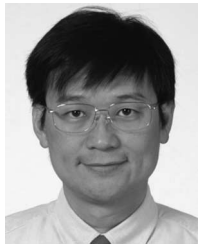
[48] P. Dash, D. Swain, A. Routray, and A. Liew, "An adaptive neural network approach for the estimation of power system frequency," *Elect. Power Syst. Res.*, vol. 41, no. 3, pp. 203–210, 1997.
[49] M. K. Mahmood, J. E. Allos, and M. A. Abdul-Karim, "Microprocessor implementation of a fast and simultaneous amplitude and frequency detector for sinusoidal signals," *IEEE Trans. Instrum. Meas.*, vol. IM-34, no. 3, pp. 413–417, Sep. 1985.
[50] Y. Xia, Z. Blazic, and D. Mandic, "Complex-valued least squares frequency estimation for unbalanced power systems," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 3, pp. 638–648, Mar. 2015.
[51] H. So and P. Ching, "Adaptive algorithm for direct frequency estimation," *IEE Proc. Radar Sonar Navig.*, vol. 151, no. 6, pp. 359–364, Dec. 2004.
[52] *Instrument Transformers. General Requirements*, BS EN 61869-1:2009, 2009, pp. 1–70.

Sithan Kanna (S'13) received the M.Eng. degree in electrical and electronic engineering with management from Imperial College London, London, U.K., in 2012. His research interests include complex-valued statistical signal processing and frequency estimation in the smart grid. He was awarded a full scholarship by The Rector's Fund at Imperial College London to pursue the Ph.D. degree in adaptive signal processing.

Dahir H. Dini received the M.Eng. degree in electrical and electronic engineering and the Ph.D. degree in adaptive signal processing from Imperial College London, London, U.K., in 2009, and 2014, respectively. His work on frequency estimation for the smart grid has been the subject of several papers and patents.

Yili Xia (M'11) received the B.Eng. degree in information engineering from Southeast University, Nanjing, China, in 2006, the M.Sc. (Hons.) degree in communications and signal processing from the Department of Electrical and Electronic Engineering, Imperial College London, London, U.K., in 2007, and the Ph.D. degree in adaptive signal processing from the same college in 2011.

He has been a Research Associate with Imperial College London since graduation, and is currently an Associate Professor with the School of Information and Engineering, Southeast University, Nanjing, China. His research interests include linear and nonlinear adaptive filters, and complex valued and quaternion valued statistical analysis.



S. Y. (Ron) Hui (M'87–SM'94–F'03) received the B.Sc. (Hons.) degree in engineering from the University of Birmingham, Birmingham, U.K., in 1984, and the D.I.C. and Ph.D. degree from Imperial College London, London, U.K., in 1987. Currently, he is the Philip Wong Wilson Wong Chair Professor with the University of Hong Kong, Hong Kong, and a part-time Chair Professor with Imperial College London. He has authored over 300 technical papers, including more than 190 refereed journal publications. Over 55 of his patents have been adopted by industry. Dr. Hui is an Associate Editor of the IEEE TRANSACTIONS ON POWER ELECTRONICS and the IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, and an Editor of the IEEE JOURNAL OF EMERGING AND SELECTED TOPICS IN POWER ELECTRONICS. He is a Fellow of the Australian Academy of Technological Sciences and Engineering. He was the recipient of the IEEE Rudolf Chope R&D Award from the IEEE Industrial Electronics Society and the IET Achievement Medal (The Crompton Medal) in November 2010 and the 2015 IEEE William E. Newell Power Electronics Award.

Danilo P. Mandic (M'99–SM'03–F'12) is a Professor of Signal Processing with Imperial College London, London, U.K., where he has been involved in nonlinear adaptive signal processing and nonlinear dynamics. He has been a Guest Professor with Katholieke Universiteit Leuven, Leuven, Belgium, the Tokyo University of Agriculture and Technology, Tokyo, Japan, and Westminster University, London, U.K., and a Frontier Researcher with RIKEN, Wako, Japan. He has two research monographs titled *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability* (West Sussex, U.K.: Wiley, 2001) and *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models* (West Sussex, U.K.: Wiley, 2009), an edited book titled *Signal Processing Techniques for Knowledge Extraction and Information Fusion* (New York, NY, USA: Springer, 2008), and more than 200 publications on signal and image processing. Prof. Mandic has been a member of the IEEE Technical Committee on Signal Processing Theory and Methods, and an Associate Editor of the IEEE SIGNAL PROCESSING MAGAZINE, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the IEEE TRANSACTIONS ON NEURAL NETWORKS. He has produced award winning papers and products resulting from his collaboration with the industry.