2013 5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)

Dynamic Topology Adaptation for Distributed Estimation in Smart Grids

Songcen Xu
Communications Research Group
Department of Electronics
University of York, U.K.
Email: sx520@york.ac.uk

Rodrigo C. de Lamare
Communications Research Group
Department of Electronics
University of York, U.K.
CETUC, PUC–Rio, Brazil
Email: rcdl500@york.ac.uk

H. Vincent Poor
Dept. of Electrical Engineering
Princeton University
Princeton, NJ 08544, USA
Email: poor@princeton.edu

***Abstract*—This paper presents new dynamic topology adaptation strategies for distributed estimation in smart grids. A dynamic exhaustive search–based topology adaptation algorithm and a dynamic sparsity–inspired topology adaptation algorithm, which can exploit the topology of smart grids with poor–quality links and obtain performance gains, are proposed. An optimized combining rule, named the Hastings rule, is incorporated into the proposed dynamic topology adaptation algorithms. Compared with existing techniques for distributed estimation, the proposed algorithms have a better convergence rate and significantly improve the system performance. The performance of the proposed algorithms is compared with that of existing techniques in the IEEE 14–bus system.**

Keywords—Dynamic topology adaptation, distributed estimation, smart grids.

I. Introduction

The electric power industry is likely to involve many more fast information gathering and processing devices (e.g., phasor measurement units) in the future, enabled by advanced control, communication, and computation technologies [1]. As a result, the need for more decentralized estimation and control in smart grid systems will experience a high priority. Several studies have proposed strategies for distributed estimation [2], [3], [4]. With existing algorithms, the neighbors for each bus are fixed. When there are links that are more severely affected by noise or other disturbances, these approaches may not provide an optimized estimation performance for each specified bus. Moreover, with the number of neighbor buses increasing, each bus requires a large network bandwidth and transmit power. Therefore, a key problem with the strategies reported so far in the literature is that they do not exploit the topology of the smart grids and knowledge about poor links to improve the performance of distributed estimation techniques.

The objective of this paper is to propose fully distributed dynamic topology adaptation algorithms for distributed estimation in smart grids, in order to optimize the performance and minimize the mean-square error (MSE) associated with the estimates. We propose two dynamic topology adaptation strategies; the proposed algorithms exploit knowledge about poor links and the topology of the system to select a subset of links that results in improved estimation performance. For the first approach, we consider a dynamic exhaustive search–based topology adaptation (DESTA) strategy. For the DESTA algorithm, we consider all possible combinations for each bus with its neighbors. Then we choose the combination associated with the smallest MSE value.

In the second approach, we introduce the dynamic sparsity–inspired topology adaptation (DSITA) algorithm. A reweighted zero attraction (RZA) strategy is incorporated into the dynamic topology adaptation algorithm. The RZA approach is usually employed in applications dealing with sparse systems in such a way that it shrinks the small values in the parameter vector to zero, which results in better convergence rate and steady–state performance. Unlike prior work with sparsity–aware algorithms [5], [6], [7], the proposed DSITA algorithm exploits the possible sparsity of the MSE associated with each of the links in a different way and employs the Hastings rule [8]. The DSITA shrinks to zero the links that have poor performance. To implement DSITA, we introduce a convex penalty, i.e., an ℓ_1–norm term to adjust the combination coefficients for each bus with its neighbors, in order to select the neighbor buses that yield the smallest MSE values.

The dynamic topology adaptation is achieved as follows:

- For a specified bus, we calculate the MSE at all its neighbor buses including the specified bus itself through the previous estimate.
- For the bus with the maximum MSE, we impose a penalty and give a reward to the bus with the smallest MSE.

The proposed DSITA algorithm performs this process automatically. By using the DSITA algorithm, some buses with unsatisfactory performance will be eliminated and some poor buses will be taken into account when their contribution improves the overall performance, which means the system topology will change automatically as well. To further improve the performance of distributed estimation techniques, we consider the Hastings rule [8] to construct the initial combination coefficients and incorporate it into the proposed algorithms.

This paper is organized as follows. Section II describes the system model and the problem statement. In Section III, the proposed dynamic topology adaptation algorithms are introduced. The numerical simulation results are provide in Section IV. Finally, we conclude the paper in Section V.

Notation: We use boldface uppercase letters to denote matrices and boldface lowercase letters to denote vectors. We use $(\cdot)^{-1}$ to denote the inverse operator, $(\cdot)^H$ for conjugate transposition and $(\cdot)^*$ for complex conjugate transposition.

II. System Model and Problem Statement

We consider the IEEE 14–bus system [1], where 14 is the number of substations. At every time instant i, each bus $k, k = 1, 2, \ldots, 14$, takes a scalar measurement $z_k(i)$ according to

$$z_k(i) = H_k\big(\boldsymbol{x}(i)\big) + n_k(i), \quad k = 1, 2, \ldots, 14, \tag{1}$$

where $\boldsymbol{x}(i)$ is the state vector of the entire interconnected system, $H_k(\boldsymbol{x}(i))$ is a nonlinear measurement function for bus k. The quantity $n_k(i)$ is the measurement error with mean equal to zero, variance $\sigma_{n,k}^2$ and which corresponds to bus k. Fig. 1 shows a standard IEEE–14 bus system with four nonoverlapping control areas.

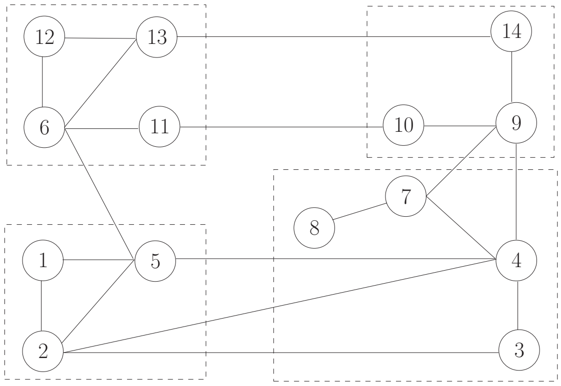

Fig. 1. IEEE 14–bus system

Initially, we focus on the linearized DC state estimation problem. The system is built with 1.0 per unit (p.u) voltage magnitudes at all buses and j1.0 p.u. branch impedance. Then, the state vector $\boldsymbol{x}(i)$ is taken as the voltage phase angle vector $\boldsymbol{\theta}$ for all buses. Therefore, the nonlinear measurement model for state estimation (1) is modified to

$$z_k(i) = \boldsymbol{h}_k^H(i)\boldsymbol{\theta} + n_k(i), \quad k = 1, 2, \ldots, 14, \tag{2}$$

where $\boldsymbol{h}_k(i)$ is the measurement Jacobian vector for bus k. Then, the aim for the distributed estimation algorithm is to compute an estimate of $\boldsymbol{\theta}$, which can minimize the cost function

$$J_x(\boldsymbol{x}) = \mathbb{E}\left|z_k(i) - \boldsymbol{h}_k^H(i)\boldsymbol{x}_k(i)\right|^2, \tag{3}$$

where $\mathbb{E}$ denotes the expectation operator, with respect to the random measurements $z_k(i)$.

A least–squares–type distributed algorithm, named Modified–Coordinated State Estimation ($\mathcal{M}$–$\mathcal{CSE}$), has been reported in the literature [4]. In this strategy, the system is decomposed into N areas. Based on the current state vector $\boldsymbol{x}_n(i)$, where n=1,2, …, N, the exchanged data $\{\boldsymbol{x}_l(i)\}_{l\in\Omega_n}$, and the measurement vector $\boldsymbol{z}_n$, the estimate of the state at the nth control area can be updated via the following formula

$$\begin{aligned}\boldsymbol{x}_n(i+1) = \boldsymbol{x}_n(i) - \Bigg[\beta(i)\sum_{l\in\Omega_n}(\boldsymbol{x}_n(i) - \boldsymbol{x}_l(i)) \\ - \alpha(i)\boldsymbol{H}_n^H\big(\boldsymbol{z}_n - \boldsymbol{H}_n\boldsymbol{x}_n(i)\big)\Bigg],\end{aligned} \tag{4}$$

where $\alpha(i)$, $\beta(i)$ are time–varying weight sequences.

For the existing strategies in the literature on smart grids, the system communication topology is fixed. This situation will cause a problem when some of the neighbor buses have poor performance, or the links between buses experience a disturbance. Also, there is no chance for the bus to discard the poorly performing neighbors rather than continue to use their information. In order to solve these problems and optimize the distributed estimation process, we need to provide the system with the ability to adapt the topology dynamically.

III. Proposed Dynamic Topology Adaptation Strategies

In this section, we introduce dynamic topology adaptation strategies for distributed estimation in smart grids. The aim of our proposed DESTA and DSITA algorithms is to optimize the distributed estimation procedure and improve the performance of the system. These two algorithmic strategies give the buses the ability to choose their neighbors based on their MSE performance. Note that other performance criteria are possible.

A. Hastings Rule

We first describe a combination rule, namely the Hastings rule which offers improved performance as compared to the Metropolis rule [8], and is incorporated into the proposed algorithms. The combination coefficient c_{kl} for a bus k and its neighbor bus l, can be calculated under the Hastings rule as follows:

$$c_{kl} = \begin{cases} \frac{\sigma_{n,k}^2}{max\{|\mathcal{N}_k|\sigma_{n,k}^2, |\mathcal{N}_l|\sigma_{n,l}^2\}}, & \text{if } k \neq l \text{ are linked} \\ 1 - \sum\limits_{l\in\mathcal{N}_k/k} c_{kl}, & \text{for } k = l \end{cases} \tag{5}$$

where $\mathcal{N}_k$ is defined as the set of all buses linking to bus k, including bus k itself, $|\mathcal{N}_k|$ denotes the cardinality of $\mathcal{N}_k$, and $\sigma_{n,k}^2$ stands for the noise variance on bus k. All c_{kl} should satisfy

$$\sum_{l\in\mathcal{N}_k \forall k} c_{kl} = 1. \tag{6}$$

The Hastings rule is a fully–distributed solution, as each bus k only needs to obtain the degree–variance product $(|\mathcal{N}_l|-1)\sigma_{n,l}^2$ from its neighbour l, to get the combination coefficient [9].

B. Dynamic Exhaustive Search–Based Topology Adaptation (DESTA)

In the proposed DESTA algorithm, we divide the distributed estimation procedure into two steps. The first step is the adaptation step and the second step is the combination step. For the proposed DESTA algorithm, we employ the adaptation strategy given by

$$\boldsymbol{\psi}_k(i) = \boldsymbol{x}_k(i-1) + \mu_k\boldsymbol{h}_k(i)\big[z_k(i) - \boldsymbol{h}_k^H(i)\boldsymbol{x}_k(i-1)\big]^*, \tag{7}$$

where μ_k is the step size for bus k. Following the adaptation step, we introduce the combination step for the DESTA algorithm, based on an exhaustive search strategy. At first, we introduce a tentative set Ω_s using a combinatorial approach described by

$$\Omega_s \triangleq {}^tC_T, \quad t = 1, 2, \ldots, T, \tag{8}$$

where ${}^tC_T$ stands for a combination operation, that refers to the combination of T elements taken t at a time without repetition. T is the total number of buses linked to bus k including bus k itself. This combinatorial strategy will cover all combination choices for each bus k with its neighbors. After the tentative set Ω_s is defined, we redefine the cost function (3) for each bus as

$$J_\psi(\boldsymbol{\psi}) \triangleq \mathbb{E}\left|z_k(i) - \boldsymbol{h}_k^H(i)\boldsymbol{\psi}\right|^2, \tag{9}$$

where

$$\boldsymbol{\psi} \triangleq \sum_{l\in\Omega_s} c_{kl}\boldsymbol{\psi}_l(i) \tag{10}$$

Then, we introduce the error pattern for each bus, which is defined as

$$e_{\Omega_s}(i) \triangleq z_k(i) - \boldsymbol{h}_k^H(i)\big[\sum_{l\in\Omega_s} c_{kl}\boldsymbol{\psi}_l(i)\big]. \tag{11}$$

For each bus k, the strategy that finds the best set Ω_s should solve the following optimization

$$\widehat{\Omega}_s = \arg\min_{\Omega_s} \big|e_{\Omega_s}(i)\big|. \tag{12}$$

After the adaptation step has been completed, the combination step is performed as given by

$$\boldsymbol{x}_k(i) = \sum_{l\in\widehat{\Omega}_s} c_{kl}\boldsymbol{\psi}_l(i). \tag{13}$$

The DESTA algorithm corresponds to equations (7)-(13) and the combination weights are obtained from (5).

C. Dynamic Sparsity-Inspired Topology Adaptation (DSITA)

The DESTA algorithm previously described needs to examine all possible sets to find a solution, which might result in unacceptable computational complexity for large systems such as the IEEE 118–bus system [1]. To solve this combinatorial problem with low complexity, we propose the sparsity–inspired based DSITA algorithm, which bears the simplicity of a standard diffusion least–mean–squares (LMS) algorithm and is suitable for adaptive implementations and scenarios where the parameters to be estimated are slowly time-varying.

The zero–attracting strategy (ZA), reweighted zero-attracting strategy (RZA) and zero-forcing (ZF) are reported in [5] and [10] as sparsity aware techniques. These approaches are usually employed in applications dealing with sparse systems in such a way that they shrink the small values in the parameter vector to zero, which results in better convergence and steady–state performance. Unlike existing methods that shrink the signal samples to zero, our proposed DSITA algorithm shrinks to zero the links that have poor performance [7].

We follow the same processing as in (7) for the adaptation step, then we redesign the combination step. First, we introduce the convex penalty term ℓ_1–norm into the combination step. Different penalty terms have been considered for this task. We have adopted the heuristic approach [5], [11] called reweighted zero–attracting strategy, into the combination step, because this strategy has shown an excellent performance and is simple to use. Then, we consider the log penalty function

$$f_1\big(e_{kl}(i)\big) = \log\big(1+\varepsilon|e_{kl}(i)|\big), \tag{14}$$

where the error pattern $e_{kl}(i)(l\in\mathcal{N}_k)$, which stands for the neighbour bus l of bus k including bus k itself, is defined as

$$e_{kl}(i) \triangleq z_k(i) - \boldsymbol{h}_k^H(i)\boldsymbol{\psi}_l(i) \tag{15}$$

and ε is the shrinkage magnitude. Then, we introduce the vector and matrix quantities required to describe the combination step. We first define a vector $\boldsymbol{c}_k$ that contains the combination coefficients for each neighbour of bus k including bus k itself as described by

$$\boldsymbol{c}_k \triangleq [c_{kl}] \qquad l\in\mathcal{N}_k. \tag{16}$$

Then, we introduce a matrix $\boldsymbol{\Psi}_k$ that includes all the estimated vectors, which are generated after the adaptation step in (7), for each neighbour of bus k and including bus k itself as given by

$$\boldsymbol{\Psi}_k \triangleq [\boldsymbol{\psi}_l(i)] \qquad l\in\mathcal{N}_k. \tag{17}$$

An error vector $\boldsymbol{e}_k$ that contains all error values calculated through (15) for each neighbour of bus k and including bus k itself is expressed by

$$\boldsymbol{e}_k \triangleq [e_{kl}(i)] \qquad l\in\mathcal{N}_k. \tag{18}$$

To devise the sparsity–inspired approach, we have modified the vector $\boldsymbol{e}_k$ in the following way: the maximum absolute value $|e_{kl}(i)|$ in $\boldsymbol{e}_k$ will be employed as $|e_{kl}(i)|$; the minimum absolute value will be set to $-|e_{kl}(i)|$, while the remaining entries will be set to zero. At this point, the combination step can be defined as

$$\boldsymbol{x}_k(i) = \sum_{j=1}^{|\mathcal{N}_k|}\left[\boldsymbol{c}_k[j] - \rho\frac{\partial f_1(\boldsymbol{e}_k[j])}{\partial \boldsymbol{e}_k[j]}\right]\boldsymbol{\Psi}_k[j], \tag{19}$$

where $\boldsymbol{c}_k[j], \boldsymbol{e}_k[j]$ and $\boldsymbol{\Psi}_k[j]$ stand for the jth element in the $\boldsymbol{c}_k, \boldsymbol{e}_k$ and $\boldsymbol{\Psi}_k$. ρ is used to control the shrinkage intensity of the algorithm. After that, we calculate the partial derivative of $\boldsymbol{e}_k[j]$:

$$\frac{\partial f_1(\boldsymbol{e}_k[j])}{\partial \boldsymbol{e}_k[j]} = \varepsilon\frac{\mathrm{sign}(\boldsymbol{e}_k[j])}{1+\varepsilon|\xi_{\min}|}. \tag{20}$$

In (20), the parameter $\xi_{\min}$ stands for the minimum absolute value of $e_{kl}(i)$ in $\boldsymbol{e}_k$. The function $\mathrm{sign}(a)$ is defined as

$$\mathrm{sign}(a) = \begin{cases} a/|a| & a\neq 0 \\ 0 & a=0. \end{cases} \tag{21}$$

Finally, by inserting (20) into (19), the proposed combination step is given by

$$\boldsymbol{x}_k(i) = \sum_{j=1}^{|\mathcal{N}_k|}\left[\boldsymbol{c}_k[j] - \rho\varepsilon\frac{\mathrm{sign}(\boldsymbol{e}_k[j])}{1+\varepsilon|\xi_{\min}|}\right]\boldsymbol{\Psi}_k[j]. \tag{22}$$

The proposed DSITA algorithm performs dynamic topology adaptation by the adjustment of the combination coefficients through $\boldsymbol{c}_k$ in (22). For the neighbor bus with the largest MSE value, after our modifications for $\boldsymbol{e}_k$, its $e_{kl}(i)$ value in $\boldsymbol{e}_k$ will be a positive number which will lead to the term $\rho\varepsilon\frac{\mathrm{sign}(\boldsymbol{e}_k[j])}{1+\varepsilon|\xi_{\min}|}$ in (22) being positive too. This means that the combining coefficient for this bus will be reduced and the weight for this bus to build the $\boldsymbol{x}_k(i)$ is reduced too. In contrast, for the neighbor bus with the minimum MSE, as its $e_{kl}(i)$ value in $\boldsymbol{e}_k$ will be a negative number, the term $\rho\varepsilon\frac{\mathrm{sign}(\boldsymbol{e}_k[j])}{1+\varepsilon|\xi_{\min}|}$ in (22) will be negative too. As a result, the weight for this node associated with the minimum MSE to build the $\boldsymbol{x}_k(i)$ is increased. For the remaining neighbor buses, the $e_{kl}(i)$ value in $\boldsymbol{e}_k$ is zero, which means the term $\rho\varepsilon\frac{\mathrm{sign}(\boldsymbol{e}_k[j])}{1+\varepsilon|\xi_{\min}|}$ in (22) is zero and there is no change for their weights to form $\boldsymbol{x}_k(i)$. The constraint on the combination of the coefficients in (6) is still satisfied. In conclusion, each bus k will first obtain a local estimate through (7). Then, each bus will employ (15)– (22) to perform the dynamic topology adaptation.

IV. Simulations

In this section, we compare our proposed dynamic topology adaptation LMS algorithms, DESTA and DSITA, with the $\mathcal{M}$–$\mathcal{CSE}$ [4], the single link selection LMS algorithm [9] and the standard diffusion Adapt–then–Combine (ATC) LMS algorithm [12] based on the MSE performance and the Phase Angle Gap. The MSE comparison is used to determine the accuracy of the algorithms, and the Phase Angle Gap is used to compare the convergence rate. In our scenario, 'Phase Angle Gap' stands for the phase angle difference between the target

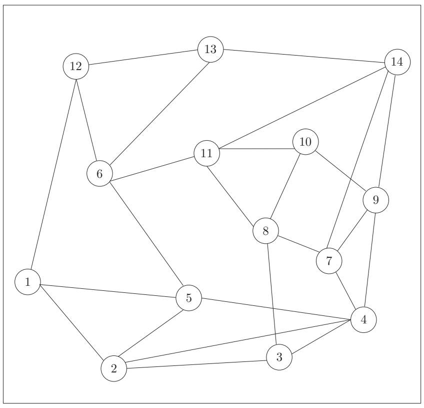

Fig. 2. IEEE 14–bus system for simulation

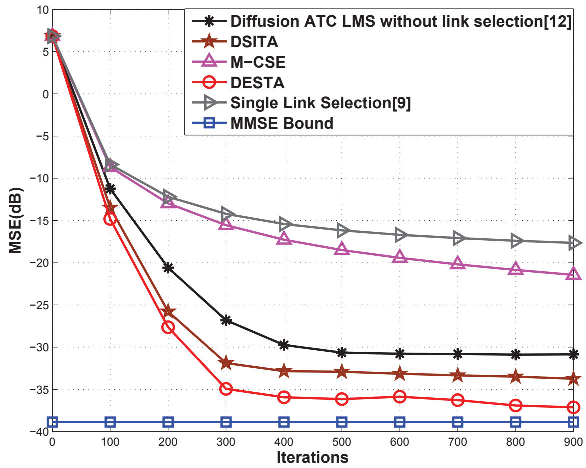

Fig. 3. MSE performance curves.

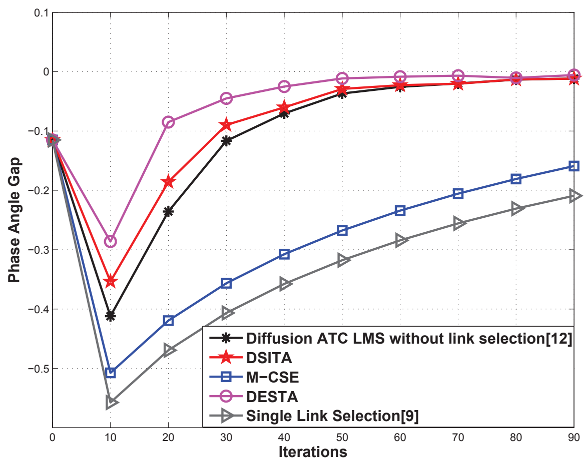

Fig. 4. Phase Angle Gap comparison for Bus 5.

$\boldsymbol{\theta}$ and the estimate $\boldsymbol{x}_k(i)$ for all buses. We define the IEEE–14 bus system as in Fig. 2.

All buses are corrupted by additive white Gaussian noise with equal variance $\sigma^2_{n,k} = 0.001$. The step size for the proposed DESTA and DSITA algorithms is 0.018. The parameter vector $\boldsymbol{\theta}$ is set to an all-one vector. The sparsity parameters of the DSITA algorithm are set to $\rho = 0.07$ and $\varepsilon = 10$. The results are averaged over 100 independent runs. From Fig. 3, it can be seen that our proposed DESTA algorithm has the best performance, and significantly outperforms the standard diffusion ATC algorithm and $\mathcal{M}$–$\mathcal{CSE}$ algorithm. DSITA is slightly worse than DESTA, which outperforms the remaining techniques.

To compare the convergence rate, we use the term 'Phase Angle Gap' to describe the results. We pick bus 5 and the first 90 iterations as an example to show our results. In Fig. 4, the DESTA algorithm still has the fastest convergence rate, while the DSITA algorithm is the second fastest. The estimates $\boldsymbol{x}_k(i)$ made from our proposed dynamic topology adaptation algorithms can quickly reach the target $\boldsymbol{\theta}$, which means the Phase Angle Gap will converge to zero.

V. Conclusion

In this paper, two dynamic topology adaptation strategies have been proposed for distributed estimation in smart grids. The DESTA algorithm uses an exhaustive search to perform the dynamic topology adaptation, and DSITA employs a sparsity–inspired approach with the ℓ_1–norm penalization. Numerical results have shown that the two proposed algorithms achieve a better convergence rate and lower MSE values than the existing distributed state estimation algorithms. These results hold also when employing other algorithms including RLS and distributed CG [13] techniques.

References

[1] A. Bose, "Smart transmission grid applications and their supporting infrastructure," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 11–19, Jun 2010.

[2] G. N. Korres, "A distributed multiarea state estimation," *IEEE Trans. Power Syst*, vol. 26, no. 1, pp. 73–84, Feb 2011.

[3] L. Xie, D.-H. Choi, and S. Kar, "Cooperative distributed state estimation: Local observability relaxed," *Proc. IEEE Power and Energy Society General Meeting*, 2011.

[4] L. Xie, D.-H. Choi, S. Kar, and H. V. Poor, "Fully distributed state estimation for wide-area monitoring systems," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1154–1169, September 2012.

[5] Y. Chen, Y. Gu, and A.O. Hero, "Sparse LMS for system identification," *Proc. IEEE ICASSP*, pp. 3125–3128, Taipei, Taiwan, May 2009.

[6] R. C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank processing based on joint and iterative interpolation, decimation and filtering," *IEEE Trans. Signal Proc.*, vol. 57, no. 7, pp. 2503 – 2514, July 2009.

[7] S. Xu, R. C. de Lamare, and H. V. Poor, "Adaptive link selection strategies for distributed estimation in diffusion wireless networks," *Proc. IEEE ICASSP*, pp. 3125–3128, Vancouver, Canada, May 2013.

[8] W. K. Hastings, "Monte carlo sampling methods using markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97109, Apr 1970.

[9] X. Zhao and A. H. Sayed, "Single-link diffusion strategies over adaptive networks," *Proc. IEEE ICASSP*, pp. 3749–3752, Kyoto, Japan, March 2012.

[10] R. Meng, R. C. de Lamare, and V. H. Nascimento, "Sparsity-aware affine projection adaptive algorithms for system identification," *Proc. Sensor Signal Processing for Defence Conference*, London, UK, 2011.

[11] Y. Chen, Y. Gu, and A. Hero, "Regularized least-mean-square algorithms," *Technical Report for AFOSR*, December 2010.

[12] 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, July 2008.

[13] S. Xu and R. C. de Lamare, "Distributed conjugate gradient sdtrategies for distributed estimation over sensor networks," *Proc. Sensor Signal Processing for Defence 2012*, London, UK, 2012.