Sparse Distribution System State Estimation: An Approximate Solution Against Low Observability

Alireza Akrami, M. Salman Asif, and Hamed Mohsenian-Rad†
Department of Electrical and Computer Engineering, University of California, Riverside, CA, USA
e-mails: {aakrami, sasif, hamed}@ece.ucr.edu

***Abstract*—When it comes to distribution system state estimation (DSSE), limited network observability is a major concern, due to limited sensor deployment at practical power distribution systems. To address this issue, this paper proposes a novel DSSE approach, based on *sparse recovery* for distribution networks with low-observability, where the measurements come from only a handful of distribution-level phasor measurement units. Here, the DSSE problem is formulated over *differential synchrophasors*, and in form of a *least absolute shrinkage and selection operator* (Lasso), which is solved using the *alternating direction method of multipliers* (ADMM). Importantly, our solution method is *dynamic*, because it uses the state estimation results from the previous time slots in order to update the weights in the instances of the Lasso problem so as to enhance the DSSE performance.**

***Keywords*—Distribution system state estimation, low-observability, sparse recovery, differential synchrophasors.**

I. Introduction

A. The Issue

Traditionally, measurements in distribution networks are unsynchronized, have relatively low reporting rates, and are often limited to SCADA and smart meters. Recently, distribution-level phasor measurement units, a.k.a. micro-PMUs, have been deployed in some utility feeders. Micro-PMUs provide precise and time-synchronized phasor measurements for nodal voltage and branch current at high reporting rates, such as once every 8 milliseconds [1]. However, due to the cost and labor, only a small number of micro-PMUs are installed in each feeder in practice. Such few number of installations has some real-world applications, such as asset monitoring [2], phase identification [3], and fault detection [4]; also see [5] for a related survey.

Distribution system state estimation (DSSE) is a critical task to monitor power distribution systems. However, obtaining full observability through micro-PMU measurements is a major practical obstacle. Thus, a method is needed in practice to run the DSSE under *low-observability* conditions, i.e. only from a limited number of micro-PMU measurement points.

B. Intuitions Based on Sparsity

The proposed DSSE approach in this paper is based on two intuitions. First, by examining a typical micro-PMU data stream, one can see that micro-PMU measurements are often in *quasi-steady-state* conditions. That is, the changes in the loads and other operating conditions in distribution circuits are *not* frequent, when compared to the high reading rate of

†The corresponding author is H. Mohsenian-Rad. This work is supported in part by UCOP grant LFR-18-548175 and DoE grant EE 0008001.

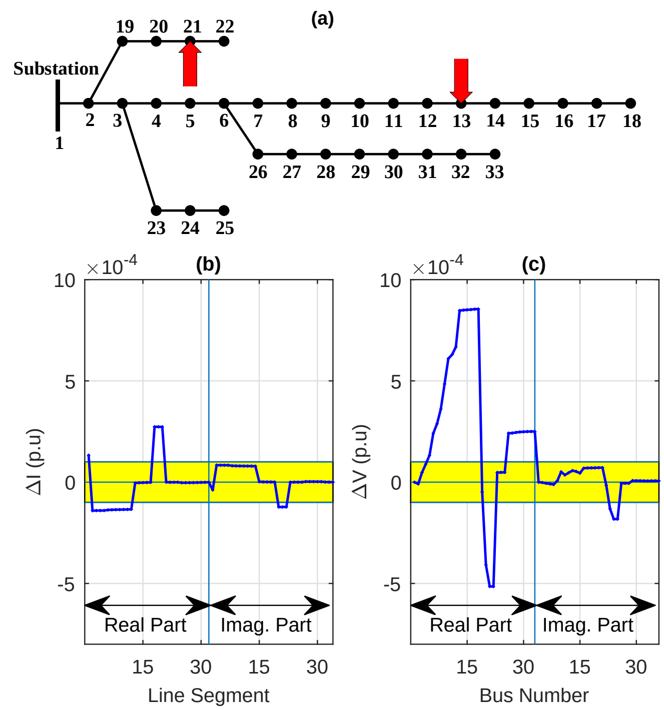

Fig. 1. An example to illustrate approximate sparsity: (a) An instance of the IEEE 33-bus test system with changes in load at buses 13 and 21; (b) The changes in line currents, i.e., ΔI; (c) The changes in bus voltages, i.e., ΔV.

micro-PMUs. Therefore, there is often useful information to carry forward from one instance of micro-PMU measurements to the next. Hence, we should focus on estimating *differential synchrophasors*, c.f. [4]. That is, we should define the state variables in the DSSE problem to be the *changes* in voltage phasors, denoted by ΔV, instead of voltage phasor itself.

The second intuition is that, the DSSE problem that is defined based on ΔV, can be seen as a *sparse recovery* problem in signal processing, c.f. [6]. This is particularly the case if the unknowns in the DSSE problem are defined as not only the changes in nodal voltages but also the changes in branch currents. This is because a typical change in a distribution circuit, such as a switching event, results in considerable changes only in a few states of the system.

An example is shown in Fig. 1, where the change in the states of the distribution system are due to major changes in the loads at bus 13 and 21, as marked on Fig. 1(a). The changes in branch current phasors, denoted by ΔI, are considerable only for those line segments which connect these load buses to the substation; and negligible elsewhere, see Fig. 1(b). As for the changes in voltage phasors, the changes are remarkable only for a subset of buses, in particular at those closer to the buses with load change. Also, the changes are more significant for the real parts rather than the imaginary parts, see Fig. 1(c). We can see that if we approximate the smaller state values that reside within the *zero approximation region*, i.e. the yellow ribbon in the figures, many of the entries in the overall vector of system states will be zero, making it a *sparse* vector.

C. Summary of Contributions

The main contributions in this paper are as follows:

1) Developing a new formulation for the DSSE problem in low-observable distribution networks, per the intuitions that we listed in Section I.B, such that we can treat the under-determined DSSE problem as a sparse recovery problem.

2) Developing a *dynamic reweighting* algorithm to solve the formulated DSSE problem. It comprises a basic *alternating direction method of multipliers* (ADMM), plus a mechanism to update the coefficients in the objective function based on the state estimation solutions in the previous time slots.

3) Conducting numerical case studies and comparing the performance of the proposed method with that of a conventional weighted least square (WLS) method aided by pseudo-measurements. We also conduct sensitivity analysis to investigate the impact of choosing the design parameters.

II. Related Work

A common approach to address low-observability conditions is to use pseudo-measurements, such as historical data from smart meters, in order to turn the under-determined system of equations in the DSSE problem into an over-determined type. However, for any DSSE method that uses pseudo-measurements, the accuracy of the DSSE solution directly depends on the accuracy and the overall data quality of the pseudo-measurements that are being used. For instance, in [7], the deviations in load profiles are used as pseudo-measurements for the purpose of solving the DSSE problem. A key issue is to construct more realistic load profiles to improve the accuracy of the corresponding DSSE solution.

Recently, machine learning is also used to resolve the observability issue in DSSE problems. The data-driven techniques in this context are again mostly utilized to generate high fidelity pseudo-measurement. For example, in [8], a Bayesian state estimator is developed via deep learning, wherein the system states and measurements are modeled as random variables with joint probability distributions. However, the challenge remains in practice on how to obtain the unknown joint probability distribution among the random variables through a learning mechanism, based on limited measurements.

Compressed sensing and sparse signal recovery has also recently attracted some attention to deal with low-observability in DSSE, such as in [9]. However there is a fundamental difference between this paper and the work in [9], and that is the fact that here we define the DSSE problem based on *differential synchrophasors*, as we pointed out earlier. This results in immediately branching away from [9] in our problem formulation, and subsequently our proposed solution methods. This is an important distinction, because one can make a stronger and more practical case for sparsity in the DSSE problem when the states of the system are defined based on the differential current and voltage synchrophasaors. Another drastic difference with [9] is that our proposed method is in a *dynamic* approach, where it uses the state estimation results from the previous time slots so as to update the weights of objective function of the state estimation problem to improve its performance, while approach in [9] is static.

III. Sparse DSSE Method

A. Standard DSSE Problem

Consider a typical linear DSSE problem [10]:

$$z = Hx + e. \tag{1}$$

As in [10], we assume that the measurements come from micro-PMUs. However, given the scope of this paper, suppose the number of micro-PMUs is very limited, which results in a low-observability condition. Accordingly, we define $z \in \mathbb{R}^m$ as the vector of micro-PMU measurements that includes m distinct nodal voltage and branch current phasors at all micro-PMU locations, $x \in \mathbb{R}^n$ is the vector of state variables and H is the measurement function matrix of size $m \times n$.

If Problem (1) is an under-determined system of equations, then the DSSE problem has low-observability, and it has infinite number of solutions. Thus, in order to find the unique solution of the system states, Problem (1) should be modified.

B. Reformulated DSSE Problem

To deal with the low-observability conditions, we select the system state variables as the changes in the voltage phasors of all buses at time slot t in comparison with time slot $t-1$ and denote the resulting vector by

$$x^t = \left[\ \Delta V^t_{j,r} \ \vdots \ \Delta V^t_{j,i}\ \right]^T, \tag{2}$$

where

$$\begin{cases} \Delta V^t_{j,r} = V^t_{j,r} - V^{t-1}_{j,r} \\ \Delta V^t_{j,i} = V^t_{j,i} - V^{t-1}_{j,i} \end{cases} \quad \forall j \in \mathcal{N}_{nodes}, \quad \forall t \in \mathcal{T}. \tag{3}$$

Here, $\Delta V^t_{j,r}$ and $\Delta V^t_{j,i}$ denote the real part and imaginary part of the voltage phasor difference at node *j*, respectively, and $\mathcal{N}_{nodes}$ denotes for the set of all buses on the network.

Similarly, at each time slot t, we define the changes in the current phasor on the line between bus k and bus l as:

$$\begin{aligned} \Delta I^t_{kl,r} &= G_{kl}(\Delta V^t_{k,r} - \Delta V^t_{l,r}) - B_{kl}(\Delta V^t_{k,i} - \Delta V^t_{l,i}) \\ \Delta I^t_{kl,i} &= B_{kl}(\Delta V^t_{k,r} - \Delta V^t_{l,r}) + G_{kl}(\Delta V^t_{k,i} - \Delta V^t_{l,i}) \end{aligned} \tag{4}$$

where G_{kl} and B_{kl} denote the conductance and the susceptance of line *kl*, respectively. The equations of the form in (4) can be defined for $\forall kl \in \mathcal{N}_{branch}$, where $\mathcal{N}_{branch}$ denotes the set of all line segments in the network.

Next, we define the vector of the measurements as follows:

$$\begin{gathered} z^t = \left[\ \Delta V^t_{j,r} \ \vdots \ \Delta V^t_{j,i} \ \vdots \ \Delta I^t_{kl,r} \ \vdots \ \Delta I^t_{kl,i}\ \right]^T \\ \forall j \in \mathcal{M}_{node}, \quad \forall kl \in \mathcal{M}_{branch}, \end{gathered} \tag{5}$$

where $\mathcal{M}_{node}$ and $\mathcal{M}_{branch}$ denote the set of nodes and lines whose voltage and currents are measured by micro-PMUs. By replacing x^t from (2) and z^t from (5) in (1) we can obtain a reformulated version of the standard DSSE problem.

C. DSSE as a Sparse Recovery Problem

From Section I.B, we know that the vector of changes in voltage phasors is a somewhat sparse vector. If we now incorporate the changes in current phasors as auxiliary variables, within the state variables, then the resulting vector of unknowns x^t becomes even more sparse. In this regard, we can further adjust the system states as follows:

$$\begin{aligned} x^t = \left[\ \Delta V^t_{j,r} \ \vdots\ \Delta V^t_{j,i} \ \vdots\ \Delta I^t_{kl,r} \ \vdots\ \Delta I^t_{kl,i}\ \right]^T \\ \forall j \in \mathcal{N}_{node}, \quad \forall kl \notin \mathcal{M}_{branch}. \end{aligned} \tag{6}$$

The reformulated DSSE problem in Section III.B can now be seen as a sparse recovery problem, which is formulated as a *least absolute shrinkage and selection operator* (Lasso) [6]:

$$\min_{x^t} \frac{1}{2}\left\|z^t - Hx^t\right\|_2^2 + \lambda\left\|x^t\right\|_1. \tag{7}$$

where the first term is the estimation error and the second term with ℓ_1-norm is the *sparsity regularizer*, which limits the number of non-zero entries of x^t through the penalty factor λ. Importantly, we can determine the *zero-approximation region* in Fig. 1 by adjusting the value of λ. It should also be noted that corresponding values in z^t to newly added unknowns of x^t (i.e. changes in current phasor for segments without micro-PMU measurement) are zero. In fact, we move ΔI^t_{kl} in (4) to the right side of equation and incorporate the constraints in (4) in the formulation in (7) as part of matrix H. Therefore, we do not need to include them as separate constraints.

IV. Dynamic Solution Approach

A. Basic Iterations

The optimization (7) is convex but non-smooth, c.f. [11]. A computationally convenient way to solve (7) is to use ADMM [11]. In this regard, we first rewrite problem (7) as:

$$\begin{aligned} &\min_{x^t, y} \frac{1}{2}\left\|z^t - Hx^t\right\|_2^2 + \lambda\|y\|_1 \\ &\text{s.t.} \qquad x^t = y. \end{aligned} \tag{8}$$

The augmented Lagrangian of (8) with penalty factor $\rho > 0$ has got the form of:

$$\begin{aligned} \mathcal{L}_\rho(x^t, y, u) = \frac{1}{2}\left\|z^t - Hx^t\right\|_2^2 + \lambda\|y\|_1 + \rho\langle u, x^t - y\rangle \\ + \frac{\rho}{2}\left\|x^t - y\right\|_2^2 \end{aligned} \tag{9}$$

Next, we can derive the following system of iterative equations by minimizing $\mathcal{L}_\rho(x^t, y_{k-1}, u_{k-1})$ over x^t and $\mathcal{L}_\rho(x^t_k, y, u_{k-1})$ over y to solve the above reformulated optimization problem [11]:

$$\begin{aligned} x^t_{k+1} &:= (H^\top H + \rho I)^{-1}(H^\top z^t + \rho(y_k - u_k)) \\ y_{k+1} &:= S_{\lambda/\rho}(x^t_{k+1} + u_k) \\ u_{k+1} &:= u_k + x^t_{k+1} - y_{k+1} \end{aligned} \tag{10}$$

where subscript k denotes for the k-th iteration, u is the dual variable corresponding to the constraint in (8), and operator $S_{\lambda/\rho}$ is the proximal operator [11].

The ADMM iterations in (10) determine the values of the very small entries, i.e. those that can be approximated by zero in the sparse recovery process, based on the values of λ and ρ, and the pre-determined maximum number of iterations. This provides an advantage in implementing the DSSE algorithm because it provides a knob to control the extent of the zero-approximation for the states that seem small.

B. Dynamic Enhancement

Given the differential nature of the state variables in this study, the proposed DSSE problem is inherently a *dynamic* framework and its performance can be further enhanced through learning its sparse characteristics, i.e., learning which unknowns can be zero-approximated. This can be done at each time slot by using a *dynamic reweighting* based on the largest obtained values from the previous time slot. The changes in the system states are due to changes in loads or grid components. The buses that experience the largest changes in their voltage are often those that are close to such loads or grid components. In this regard, we propose to rewrite (8) as:

$$\begin{aligned} &\min_{x^t, y} \frac{1}{2}\left\|z^t - Hx^t\right\|_2^2 + \lambda\|y\|_1 \\ &s.t. \qquad F^t x^t = y \end{aligned} \tag{11}$$

where F is a diagonal $n \times n$ reweighting matrix. Suppose we are at time slot t. Let Γ_{t-1} denote the set of the half largest differential states that were non-zero at time slot $t-1$, i.e., during the previous time slot. We update the entries of F^t as:

$$F^t_{ii} = \begin{cases} 1 & ;\text{if} \qquad i \in \Gamma_{t-1} \\ 1/|x^{t-1}_i + \epsilon| & ;\text{otherwise} \end{cases} \tag{12}$$

where ϵ is a small positive value. From (12), if the previous differential state was among the half of largest non-zero differential states, i.e., the voltage or current phasor corresponding to that state variable had a considerable change, then the weight corresponding to that differential state remains 1, which is the default value as in (8). However, if the previous differential state was zero, then the weight is set in an inverse proportional relationship with its value during the previous time slot. This will relax the constraint of sparsity on these entries. Note that, inverse proportional relationship assigns large penalty coefficients to the approximately zero elements; this helps the algorithm to improve its accuracy in the next time slots by ignoring the corresponding state variables.

Given F^t, the iterative solutions in (10) are updated as:

$$\begin{aligned} x^t_{k+1} &:= (H^\top H + \rho(F^t)^\top F^t)^{-1}(H^\top z^t + \rho F^t(y_k - u_k)) \\ y_{k+1} &:= S_{\lambda/\rho}(F^t x^t_{k+1} + u_k) \\ u_{k+1} &:= u_k + F^t x^t_{k+1} - y_{k+1}. \end{aligned} \tag{13}$$

C. Recoverability Analysis

Consider the problem formulation in (7). Based on the vector of unknowns defined in (6), we can rewrite H as:

$$H = \left[\ W \ \vdots\ -I\ \right], \tag{14}$$

where W is the sub-matrix corresponding to changes in voltage and I is the identity sub-matrix corresponding to changes in current. We can write the vector of measurements as:

$$\begin{aligned} z^t = \left[\ W \ \vdots\ -I\ \right]\left[\ \Delta V^t_j \ \vdots\ \Delta I^t_{kl}\ \right]^T = W\Delta V^t - \Delta I^t \\ \forall j \in \mathcal{N}_{node}, \quad \forall kl \notin \mathcal{M}_{branch}. \end{aligned} \tag{15}$$

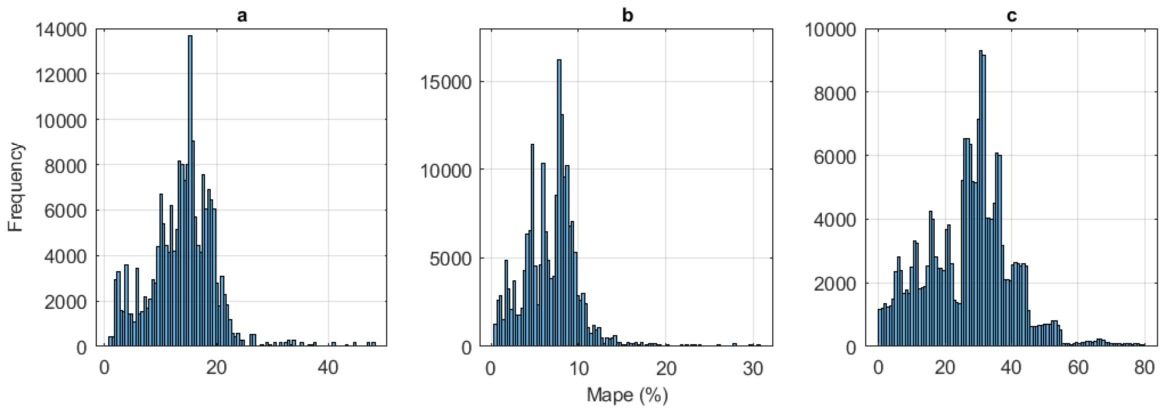

Fig. 2. Distribution of MAPE for: (a) sparse DSSE without reweight; (b) sparse DSSE with reweight; (c) WLS DSSE aided by pseudo-measurements.

From the illustrative example in Fig. 1 in Section I, we know that vector ΔI^t is *more sparse* than vector ΔV^t. Therefore, we can treat ΔI^t as outliers in $W \Delta V^t$, which resembles the problem of ℓ_1 *decoding* [12]. If W is *over-determined* and has *full rank*, then under certain conditions, we can guarantee the recovery of ΔV^t by minimizing $|z^t - W\Delta V^t|_1$ as long as the outliers ΔI^t are sufficiently sparse. In this regard, whether and to what extent we can guarantee the true recoverability of state variables depends on the number and location of micro-PMUs, which affect matrix W. For the example in our base case studies, W has about 50 useful singular vectors in a 74-dimensional space, which implies that we can resolve approximately 12 arbitrary outliers (equivalent to 12 differences in current of branches without micro-PMU measurement) in the best case and none of them in the worst case. These bounds rely on sufficient conditions for recoverability; empirical performance can be better. As we will see in the numerical case studies in the next section, the proposed DSSE method is very successful in accurately estimating the states of the system.

V. Numerical Case Studies

The performance of the proposed sparse DSSE is evaluated on IEEE 33-bus distribution test system. 2,000 Monte Carlo scenarios are constructed via MATPOWER toolbox in MATLAB R2017b. Each scenario includes 100 time slots which simulates transition of a scenario from initial state towards the final state. In the base cases, it is considered that the net load might change for a limited number of buses (up to three buses) toward different values, from 50% decrease to 50% increase across two successive micro-PMU reporting intervals.

In base case study, four micro-PMUs are placed on nodes 8, 20, 24, and 28. All 200,000 data samples are solved through proposed sparse DSSE without and with the weight update to investigate the effect of dynamic reweighting on the method. Also, to show the efficiency of proposed method in comparison with other existing DSSE which deal with low-observability, all data samples are run into a conventional WLS which uses the initial values of smart meters as pseudo-measurements over the whole time intervals. We have considered the mean absolute percentage error as defined in (16) to evaluate the performance of proposed method, since method assessment based on mean absolute error might be misleading due to relatively small values of voltage difference:

$$e_{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \frac{|\hat{x}_i - x_i|}{|x_i^{max}|} \times 100 \tag{16}$$

where $|x_i^{max}|$ is the largest voltage difference in each scenario.

Distribution of MAPE for each of three DSSE methods is shown in Fig. 2, which illustrates that proposed method has got a better performance over the conventional WLS aided by pseudo-measurements. The average and standard deviation of MAPE over all data samples is outlined in Tables I, which demonstrates that sparse DSSE with weight update has got the lowest average and standard deviation error.

TABLE I
Statistical Charachteristic of proposed methods vs. conventional WLS aided by pseudo-measurements for MAPE

Method	Average MAPE	STD of MAPE	Max. MAPE
Sparse DSSE without reweight	13.91%	5.75	48.21%
Sparse DSSE with with reweight	6.83%	3.23	30.75%
Conventional WLS aided by pseudo-measurements	25.33%	12.96	79.96%

To get a better insight on how inaccurate might be the performance of proposed method, distribution of worst case of estimation error (i.e. maximum absolute percentage voltage estimation error) over all scenarios is plotted in Fig. 3. Based on Fig. 3, the average absolute percentage error for worst case of scenarios is about 35%, which means that for a typical scenario the highest error of estimation for the voltage difference is 35% in average for a network with low-observability.

To better investigate the performance of proposed method, a sensitivity analysis is performed, wherein mean absolute percentage value of estimation error with respect to node number is investigated for three distinct micro-PMU installations (Fig. 4). In the base case (case1), wherein only four micro-PMUs are installed, the lower error along each of four laterals of 33-bus test feeder, belongs to the nodes that are equipped with micro-PMUs. As we get closer to the dead-ends, the estimation error is being increased. To investigate the effect of adding micro-PMUs at the end of laterals, a new micro-PMU is installed at node 18 (case 2) and 33 (case 3), respectively in two separate case studies. Regarding Fig.4, adding a single micro-PMU in

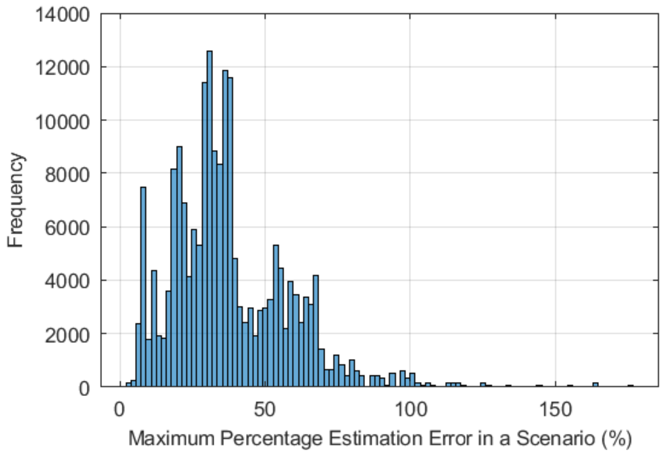

Fig. 3. Distribution of the worst case in estimation error over all scenarios.

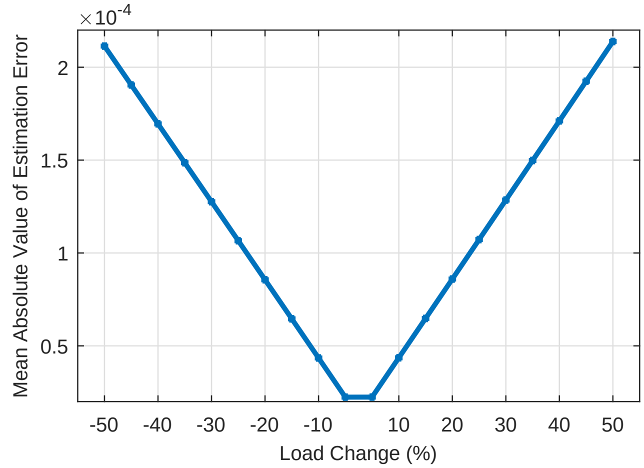

Fig. 5. Mean absolute value of estimation error vs. change in load.

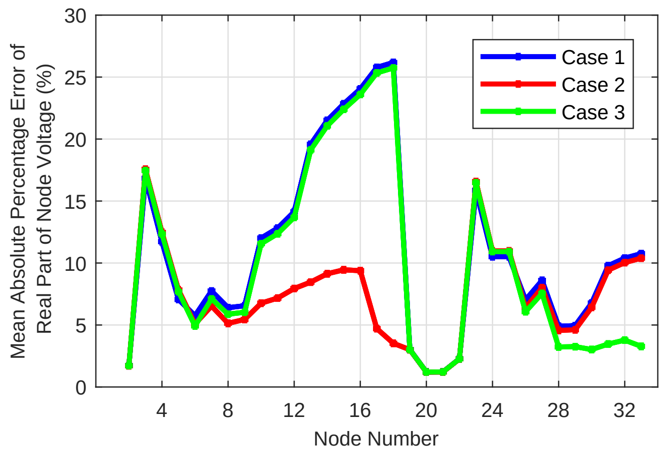

Fig. 4. Mean absolute percentage value of estimation error at each bus.

the dead-end nodes, causes a remarkable improvement in the accuracy of proposed sparse DSSE for estimation of voltage of nodes which are along the lateral of node equipped with new sensor, even though full-observability condition is not still obtained. But, it does not affect the performance of estimation for nodes which are not along that particular lateral which mainly originates from the radiality of distribution grids.

To investigate the effect of load change on the performance of proposed sparse DSSE, mean absolute error with respect to load change is shown in Fig. 5. In this sensitivity analysis, only scenarios wherein a single event happens are investigated. As magnitude of load change is increased, it causes the estimation error to incline. It should be noted that, the symmetrical shape of Fig. 5 is due to the low penetration of distributed energy resources in the test cases, which causes the system to make up for the difference of current only from the substation node. So, based on the linear equations in (4), load reduction or increment in a same amount will have the same error.

VI. Conclusions

A novel DSSE approach is proposed for low-observability conditions based on two key concepts: defining the state variables based on differential synchrophasors; and formulating and solving the resulting DSSE problem as a sparse signal recovery problem. Furthermore, a dynamic framework is proposed wherein the space DSSE approach uses the results from the previous time slots to update the sparsity weights in the objective function; to enhance the performance. Several case studies are examined and sensitivity analysis are performed which demonstrate the effectiveness of proposed method.

References

[1] H. Mohsenian-Rad, E. Stewart, and E. Cortez, "Distribution synchrophasors: Pairing big data with analytics to create actionable information," *IEEE Power Energy Mag.*, vol. 16, no. 6, pp. 26–34, Apr. 2018.

[2] A. Shahsavari *et al.*, "A data-driven analysis of capacitor bank operation at a distribution feeder using micro-pmu data," in *Proc. of the IEEE PES ISGT*, Apr. 2017.

[3] M. Wen, R. Arghandeh, A. von Meier, K. Poolla, and V. O. K. Li, "Phase identification in distribution networks with micro-synchrophasors," in *Proc. of the IEEE PES General Meeting*, Jul. 2015.

[4] M. Farajollahi, A. Shahsavari, E. M. Stewart, and H. Mohsenian-Rad, "Locating the source of events in power distribution systems using micro-pmu data," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6343–6354, Nov 2018.

[5] A. von Meier, E. Stewart, A. McEachern, M. Andersen, and L. Mehrmanesh, "Precision micro-synchrophasors for distribution systems: A summary of applications," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2926–2936, Nov 2017.

[6] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer, 2010.

[7] A. Angioni, T. Schlosser, F. Ponci, and A. Monti, "Impact of pseudo-measurements from new power profiles on state estimation in low-voltage grids," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 1, pp. 70–77, Jan. 2016.

[8] K. R. Mestav, J. Luengo-Rozas, and L. Tong, "State estimation for unobservable distribution systems via deep neural networks," in *Proc. of the IEEE PES General Meeting*, Aug. 2018, pp. 1–5.

[9] M. Majidi, M. Etezadi-Amoli, and H. Livani, "Distribution system state estimation using compressive sensing," *Int. J. Elect. Power Energy Syst.*, vol. 88, pp. 175–186, 2017.

[10] M. Gol and A. Abur, "Lav based robust state estimation for systems measured by pmus," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1808–1814, Jul. 2014.

[11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn*, vol. 3, no. 1, pp. 1–122, 2011.

[12] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.