Improvement in Vulnerability and Error Analysis: A Synthetic Measurement Approach

Surya Chandan Dhulipala, Cody Ruben, Arturo Bretas
Department of Electrical & Computer Engineering
University of Florida
Gainesville, FL
chandandhulipala@ufl.edu, cruben31@ufl.edu, arturo@ece.ufl.edu

Newton Bretas
Department of Electrical and Computer Engineering
University of Sao Paulo
Sao Carlos, SP, Brazil
ngbretas@sc.usp.br

***Abstract*—Gross error analysis of measurements is very challenging in power systems, even in transmission systems with relatively high measurement redundancy. It is especially challenging to detect gross errors in measurements which are most vulnerable to undetectable errors, considering the classical residual based approach. This paper aims to ameliorate gross error analysis by augmenting the current measurement set with synthetic measurements (SM). SM are allocated by performing vulnerability analysis on the power system. A novel index called Vulnerability Index (VI) has been proposed to evaluate which measurement or bus is vulnerable to undetectable errors for a given measurement scenario. This method was tested on a 14 bus system and the results show that augmentation using SM can improve the performance of bad data detection and identification in measurements. Analysis of the influence of SM on bus VI and exclusion of various types of measurements is also performed.**

***Index Terms*—Vulnerability index , Gross error analysis, synthetic measurements, Smart grid, Cyber-attack, State estimation.**

I. Introduction

An important issue for the power grid is the detection of gross errors in measurements. These measurements come from various devices, both old and new, throughout the grid and are used by Energy Management Systems (EMS) to provide real time monitoring and situational awareness of the grid. The State Estimation (SE) process uses all of these measurements to estimate the current state of the power grid, which is then used in many applications for real time monitoring and control [1]. Therefore, it is critical to correctly detect when a measurement is in error, as this will have cascading affects throughout an EMS. These gross errors can come in many forms, including broken devices, communication errors, and with the implementation of Smart Grid (SG) devices, cyber-attacks as well [2]. As more opportunities for gross errors are introduced to the grid, it becomes increasingly important to develop more accurate and robust bad data detection (BDD) solutions.

Since BDD is a critical application within an EMS, much research has been dedicated to this problem. The most common BDD solution that is used is the statistical Chi-square test which is based on the residual values of a Weighted Least Squares (WLS) SE [1]. There have been attempts to improve upon this by modifying the test [3], but it has been shown that the use of the residual values is an issue in itself. Rather, a geometrical interpretation of the WLS SE solution, known as the Innovation concept, shows that the error itself can be estimated and used in the Chi-squared test [4], [5]. This theory has been applied to BDD improvement [6], [7], including the development of a Vulnerability Index (VI) [8]. These strategies do not directly address the fact that measurement redundancy is an important factor in using the statistical Chi-squared test. The underlying assumption in using the Chi-squared distribution is that the measurements have normal Gaussian noise. Due to the nature of the Chi-squared distribution, this assumption holds more accurately with higher degrees of freedom, or added redundancy [1]. Psuedo-measurements have been used as an attempt to add redundancy to power systems, but since these are based on historical data, they are not always available and reliable for SE [9].

This paper proposes the use of Synthetic Measurements (SM) as a way to add reliable redundancy to the power grid measurement set. SM, unlike pseudo-measurements, are derived based on the current state of the system, making them robust to load changes over time. Furthermore, the VI will be used to locate SM strategically. The VI will quantify the weak points of the system in terms of BDD so that the SM are located where they will have the greatest impact on BDD.

The remainder of the paper is organized as follows. In Section II, we present the WLS SE algorithm along with the Innovation concept. In Section III, the Synthetic Measurement strategy is presented. In Section IV, simulation results of the proposed method are shown. Finally, some concluding remarks are presented in Section V.

II. Background

A. State Estimation and Geometrical Interpretation Theory

The power system with n buses and m measurements is modeled as a set of non-linear equations as follow [1]:

$$z = h(x) + e \tag{1}$$

Where $z \in \mathbb{R}^m$ is the measurement vector, $x \in \mathbb{R}^N$ is the state variables vector, $h(x) : \mathbb{R}^m \rightarrow \mathbb{R}^N, (m > N)$ is a non-linear differentiable function that relates the states to the measurements, e is the measurement error vector assumed with zero mean, standard deviation σ and having Gaussian

978-1-7281-1981-6/19/$31.00 ©2019 IEEE

Authorized licensed use limited to: University of Florida. Downloaded on May 18,2020 at 18:47:09 UTC from IEEE Xplore. Restrictions apply.

probability distribution, and N = 2n-1 is the number of unknown state variables.

Weighted Least Square (WLS) is a classical state estimator that searches for the best estimates of the states x of the well-known problem that minimizes the cost function as follow:

$$J(x) = \|z - h(x)\|^2_{R^{-1}} = [z - h(x)]^T R^{-1}[z - h(x)] \quad (2)$$

where R is the measurement covariance matrix. $J(x)$ index is a norm in the measurements vector space. Let $\hat{x}$ be the solution of the aforementioned minimization problem, then the estimated measurement vector is $\hat{z} = h(\hat{x})$. The residual is defined as the difference between $\hat{z}$ and z, which means $r = z - \hat{z}$. Linearizing (1) at a certain operating point x^* yields the following:

$$\triangle z = H \triangle x + e \quad (3)$$

where $H = \frac{\partial h}{\partial x}$ is the Jacobian matrix of h calculated at the point x^*. $\triangle z = z - h(x^*) = z - z^*$ and $\triangle x = x - x^*$ are the correction of measurement and state vector respectively. Under observability condition, i.e rank$(H) \geq N$, the vector space of measurements can be decomposed into two sub-spaces that are orthogonal to each other. Let P be a linear operator such that $\triangle \hat{z} = P \triangle z$ and the residual vector r be $\triangle z - \triangle \hat{z}$. Then, the vector $\triangle \hat{z} = H \triangle \hat{x}$ is orthogonal to the residual vector r, since P projects the measurement vector mismatch $\triangle z$ orthogonally in the range space of H. Equivalently,

$$< \triangle \hat{x}, r >= (H \triangle x)^T R^{-1}(\triangle z - H \triangle \hat{x}) = 0 \quad (4)$$

Solving (4) for $\triangle \hat{x}$, one can obtain the following:

$$\triangle \hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} \triangle z \quad (5)$$

In other words, the projection matrix P is the idempotent matrix that has the following expression:

$$P = H(H^T R^{-1} H)^{-1} H^T R^{-1} \quad (6)$$

So, it is possible to decompose the measurement vector to two component as follow:

$$e = \underbrace{Pe}_{e_U} + \underbrace{(I - P)e}_{e_D} \quad (7)$$

The component e_D is the detectable error which is the residual in the classical model while the component e_U is the undetectable error. e_D is in the orthogonal space to the range space of Jacobian whereas e_U is hidden in the Jacobian space. The geometric interpretation is illustrated in Fig.1.

$$\|e\|^2 = \|e_D\|^2 + \|e_U\|^2 \quad (8)$$

This error vector is called Composed Measurement Error (CME). In order to find the masked error and compose it, the Innovation Index (II) is introduced to quantify the undetectable error which is proposed by [4] and is presented in the following:

$$II_i = \frac{\|e^i_D\|}{\|e^i_U\|} = \frac{\sqrt{1 - P_{ii}}}{\sqrt{P_{ii}}} \quad (9)$$

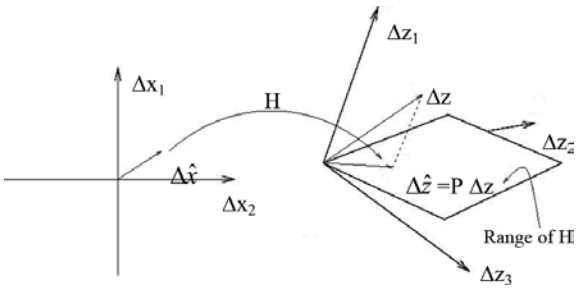

Fig. 1. Illustration of Geometrical Interpretation Method

Low Innovation index means there is a large component of error is not reflected from residual. Therefore, the residual will be very small even if there is a gross error. By using (8) and (9), the composed measurement error will be as follow:

$$CME_i = r_i \left(\sqrt{1 + \frac{1}{II_i{}^2}} \right) \quad (10)$$

If instead we work with normalized residual one can obtain Composed Normalized error (CNE) as follow:

$$CNE_i = r_i^N \left(\sqrt{1 + \frac{1}{II_i{}^2}} \right) \quad (11)$$

Where r^N is the normalized residual. Otherwise, error can be normalized by:

$$CME_i^N = \frac{r_i}{\sigma_i} \left(\sqrt{1 + \frac{1}{II_i{}^2}} \right) \quad (12)$$

where σ is the standard deviation for the measurement and one can use CNE to correct the measurement.

B. Vulnerability Index

The Vulnerability Index (VI), known as UI in [8], quantifies the vulnerability of a bus to undetectable gross errors in measurements associated with that bus. This metric is developed based on the Innovation concept presented in [4], [5]. Sensitivity Analysis is performed to determine the effect of certain input on the uncertainty of output. $S_{\triangle \hat{x}}$ is used to determine the impact $\triangle \hat{x}$ caused on the state estimate by an arbitrary perturbation Δz introduced in measurement vector. $S_{\triangle \hat{x}}$ ($\frac{\partial \hat{x}}{\partial z}$) is given by:

$$S_{\triangle \hat{x}} = \frac{\partial \hat{x}}{\partial z} = G^{-1} H^T W \quad (13)$$

where

$$G = H^T W H \text{ and } W = R^{-1}. \quad (14)$$

Similarly, the effect of Δz on the residual is given by S_r as:

$$S_r = I - H G^{-1} H^T W \quad (15)$$

where G and W are given by (14). Similar analysis can be performed to quantify the effect of Δz on C.M.E and this sensitivity matrix is denoted by S_{CME}. Since, residual is detectable component of error and CME is composed of both

This work was supported by NSF ECCS grants 1646229 and 1809739.

Authorized licensed use limited to: University of Florida. Downloaded on May 18,2020 at 18:47:09 UTC from IEEE Xplore. Restrictions apply.

the detectable and undetectable component of the error, we theorize the difference of S_{CME} and S_r [10], [11] should give a sensitivity matrix denoting the effect of Δz on undetectable component of error. We make use of aforementioned properties of sensitivity matrices to define a vulnerability index of a bus as shown in (16).

$$VI_i = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\sum_{j=1}^{J}(S_{CME}(k,j) - S_r(k,j))^2} \tag{16}$$

where K is the set of measurements associated with bus i, J is the set of all measurements, and S_{CME} and S_r are the sensitivities of the CME and residual, respectively, with respect to gross errors in measurements. The VI of a bus will quantify the vulnerability of that bus to gross errors with large undetectable components. The higher the VI of a bus, the more vulnerable that bus is to undetected errors considering the classical WLS solution.

C. Error Analysis

Assuming that the measurement errors e_i, $i = 1, ...m$, are normal and independent, each having mean zero and variance σ_i^2, $N(0, \sigma_i^2)$, then the performance index, $J(\hat{x})$, is

$$J(\hat{x}) = \sum_{i=1}^{m}\left(\frac{z_i^{meas} - \hat{z}_i}{\sigma_i}\right)^2 \tag{17}$$

and follows a χ^2_{m-N} distribution, i.e., a chi-square distribution with $m - N$ degrees of freedom, with m being the number of measurements and N the number of state variables, where $z_i^{meas} - \hat{z}_i = r$ (residual).

- If $J(\hat{x}) > C$ then reject the hypothesis H_O that there is no error.
- If $J(\hat{x}) \leq C$ then accept the hypothesis H_O.

where $C = \chi^2_{m-N,1-\alpha}$. A modified chi-square test for bad data detection has been proposed by [12]. [12] proposes standardizing residual (r) using standard deviation (SD) of residual from the diagonal of co-variance matrix of residual. A similar chi-square test can be formulated by replacing residual with the CME but with different degrees of freedom (m), while considering the minimization of the weighted norm of the error [13]. The CME should be normalized by the S.D of measurement and the proof that SD of CME is the same as SD of measurement is shown in the Appendix.

III. Synthetic Measurements using Vulnerability Index

A. Vulnerability Index

Errors in measurements can be introduced by various factors like cyber-attacks, faulty equipment and communication channel noise. The Vulnerability Index (VI) is not the property of the system but, given the measurement set for the system, the VI characterizes the vulnerability of the system to undetectable errors, considering the classical residual BBD. The VI of a bus characterizes the vulnerability of the bus to undetectable errors taking into account all the measurements associated with the bus. For a given measurement set, VI analysis can be performed on the system to determine where SM can be allocated to improve the detectability of errors.

B. Synthetic Measurements

In this paper, SM are defined as the measurements that are obtained form either the current state estimate or the states obtained from preceding scan of measurements. Adding redundancy via SM will cause the CMEs to have a more Gaussian behavior, thereby improving the reliability of the gross error detection and identification test. As opposed to the commonly used pseudo-measurements, SM are calculated from the current state of system, make them more appropriate for improving bad data error analysis. SM are calculated from the state of the system, while considering the weights associated to the precision of meters. Otherwise, the gross error analysis considers the two step state estimation approach [6], where the weights are given as a percentage of the measurement magnitude. SM are selected considering the bus VI values. All of the VI values are ordered and the buses with the highest VI values will have SM added to them. The number of buses and SM will depend on the system being analyzed.

IV. Results

The proposed method was tested using the 14-bus test system shown in Fig. 2. This system has 27 ($2*n-1$;where n is the number of buses) states. A measurement set consisting of both active and reactive power flows and injections is considered. A total of 81 measurements are considered making the redundancy equal to 3 which is typical for transmission systems. These are the measurements that are used directly in the measurement error detection. As described in Section III, the concept of SM is used to add measurement redundancy to the NLSE rather than the commonly used pseudo-measurements. In order to determine where on the system SM will be added, a VI analysis of the system is done. Based on the measurement set considered, the bus VI values for each of the fourteen buses on the system for various cases are shown in Table II. These pre-processing values are shown in the "Only SCADA" cases. Certain cases where a certain class (Pij and Pji) of measurements are removed were also presented. These cases were taken into account to depict lack of certain measurement channels in SCADA. Plots of VI for the aforementioned cases are illustrated in Fig.3 and Fig.4. Buses 2 and 5 have significantly higher VI when compared to other buses which is evident from the peaks in Fig.3 and Fig.4.

This analysis shows that bus 2 is the most vulnerable to gross errors in measurements going undetected by the standard residual based chi-square test followed by bus 5. We also see that buses 1 and 4 are also relatively vulnerable. The decision to introduce SM at bus 2 and bus 5 was made based on the fact that VI values for these buses were significantly higher than VI of other buses. It is proposed to generate SM for the real power flows from bus 5 to bus 2 and vice versa from the

This work was supported by NSF ECCS grants 1646229 and 1809739.

Authorized licensed use limited to: University of Florida. Downloaded on May 18,2020 at 18:47:09 UTC from IEEE Xplore. Restrictions apply.

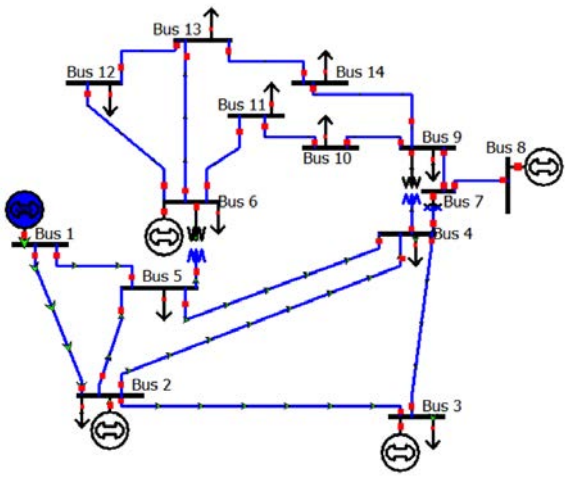

Fig. 2. 14 Bus Test System

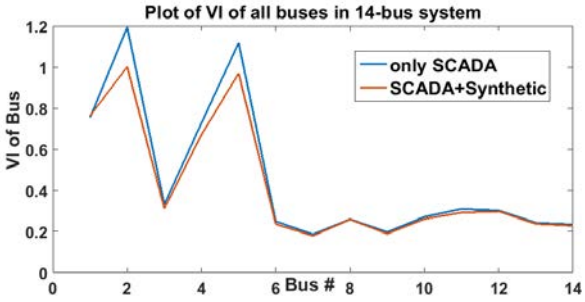

Fig. 3. Effect of addition of SM on VI of buses

above analysis. Additionally, reactive power injection at bus 2 and bus 5 were also considered as SM. This adds the most local redundancy to buses 2 and 5. The SM are created based on the results of the latest scan of measurements.

To evaluate the effectiveness of the introduction of SM in the error detection process, Monte Carlo analysis is performed in MATLAB considering 5000 scenarios of various loading conditions and random errors in the measurement. The various loading conditions are simulated in the MATPOWER package [14]. For this analysis, a scenario is correct only when the error is both detected and identified correctly. The improvement of accuracy in error detection was up to 7.74%. The improvement of accuracy of error analysis with introduction of SM is shown in Table I, III and IV. The introduction of SM also improves the VI of the bus, consequently the buses are less vulnerable to undetected gross errors.

TABLE I
ERROR ANALYSIS RESULTS FOR FULL SCADA CASE

	Detection	Identification	Total
Without SM	91.16%	65.93%	60.63%
With SM	91.62%	74.04%	67.84%

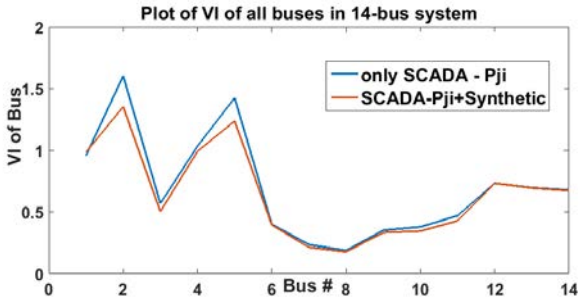

Fig. 4. Effect of addition of SM on VI of buses after power flow measurements are removed

TABLE II
VI VALUES OF ALL BUSES

Case	Bus VI values
Only SCADA	[0.753,1.193,0.331,0.728,1.117,0.249,0.186, 0.260,0.197,0.271,0.310,0.303,0.242,0.234]
SCADA+SM	[0.762,1.001,0.312,0.670,0.970,0.235,0.179, 0.258,0.188,0.261,0.292,0.298,0.237,0.228]
SCADA-Pij	[0.892,1.607,0.567,0.989,1.427,0.358,0.229, 0.261,0.287,0.444,0.507,0.439,0.419,0.432]
SCADA-Pij+SM	[0.911,1.361,0.502,0.949,1.241,0.344,0.217, 0.261,0.269,0.425,0.490,0.433,0.404,0.415]
SCADA-Pji	[0.957,1.603,0.568,1.033,1.424,0.400,0.235, 0.187,0.351,0.378,0.469,0.732,0.697,0.678]
SCADA-Pji+SM	[0.984,1.353,0.503,0.991,1.237,0.394,0.213, 0.177,0.332,0.344,0.428,0.733,0.694,0.672]

V. CONCLUSIONS

This paper presents a method to ameliorate gross error analysis by augmenting the current measurement set with synthetic measurements. Synthetic measurements are allocated by performing vulnerability analysis on the system. A novel index called Vulnerability Index (VI) has been proposed to evaluate which measurement or bus is vulnerable to undetectable errors for a given measurement scenario. Results show that augmentation using synthetic measurements can improve the performance of bad data detection and identification in measurements.

TABLE III
ERROR ANALYSIS RESULTS FOR SCADA - Pij CASE

	Detection	Identification	Total
Without SM	61.3%	51.83%	50.5%
With SM	81.9%	59.53%	53.6%

TABLE IV
ERROR ANALYSIS RESULTS FOR SCADA- Pji CASE

	Detection	Identification	Total
Without SM	60.86%	51.25%	50.7%
With SM	80.13%	69.78%	67.83%

This work was supported by NSF ECCS grants 1646229 and 1809739.

Authorized licensed use limited to: University of Florida. Downloaded on May 18,2020 at 18:47:09 UTC from IEEE Xplore. Restrictions apply.

APPENDIX

Sensitivity of CME:

$$S_r = \frac{r}{e}$$

$$CME_i = r_i \sqrt{1 + \frac{1}{II_i^2}}$$

Let:

$$IC_i = \sqrt{1 + \frac{1}{II_i^2}}$$

$$\mathbf{IC} = \begin{bmatrix} IC_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & IC_m \end{bmatrix}$$

So:

$$CME = \mathbf{IC} * r$$

$$S_{CME} = \frac{CME}{e} = \frac{\mathbf{IC} * r}{e} = \mathbf{IC} * S_r$$

Covariance and SD of CME:

$$\Omega_{CME} = Cov(CME) = E(CME * CME^T)$$

$$\Omega_{CME} = S_{CME} * E(e * e^T) * S_{CME}^T$$

$$\Omega_{CME} = S_{CME} * R * S_{CME}^T$$

$$\Omega_{CME} = \mathbf{IC} * S * R * S^T * \mathbf{IC}^T$$

$$\Omega_{CME} = \mathbf{IC} * S * R * \mathbf{IC}^T$$

$$\Omega_{CME,ii} = IC_i * S_{ii} * R_{ii} * IC_i$$

$$\Omega_{CME,ii} = IC_i^2 * S_{ii} * R_{ii}$$

$$IC_i^2 = 1 + \frac{1}{II_i^2} = 1 + \frac{1 - S_{ii}}{S_{ii}} = \frac{S_{ii} + 1 - S_{ii}}{S_{ii}} = \frac{1}{S_{ii}}$$

$$\Omega_{CME,ii} = \frac{1}{S_{ii}} * S_{ii} * R_{ii} = R_{ii}$$

$$\sqrt{\Omega_{CME,ii}} = \sigma_{CME,i} = \sqrt{R_{ii}} = \sigma_i$$

REFERENCES

[1] A. Monticelli, *State estimation in electric power systems: a generalized approach*. Springer Science & Business Media, 2012.

[2] M. Farag, M. Azab, and B. Mokhtar, "Cross-layer security framework for smart grid: Physical security layer," in *IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, pp. 1–7, IEEE, 2014.

[3] M. Göl and A. Abur, "A modified chi-squares test for improved bad data detection," in *2015 IEEE Eindhoven PowerTech*, pp. 1–5, June 2015.

[4] N. Bretas, A. Bretas, and S. Piereti, "Innovation concept for measurement gross error detection and identification in power system state estimation," *IET generation, transmission & distribution*, vol. 5, no. 6, pp. 603–608, 2011.

[5] N. G. Bretas, S. A. Piereti, A. S. Bretas, and A. C. P. Martins, "A geometrical view for multiple gross errors detection, identification, and correction in power system state estimation," *IEEE Transactions on Power Systems*, vol. 28, pp. 2128–2135, Aug 2013.

[6] N. G. Bretas and A. S. Bretas, "A two steps procedure in state estimation gross error detection, identification, and correction," *International Journal of Electrical Power & Energy Systems*, vol. 73, pp. 484 – 490, 2015.

[7] A. S. Bretas, N. G. Bretas, B. Carvalho, E. Baeyens, and P. P. Khargonekar, "Smart grids cyber-physical security as a malicious data attack: An innovation approach," *Electric Power Systems Research*, vol. 149, pp. 210 – 219, 2017.

[8] C. Ruben, S. Dhulipala, A. Bretas, and N. Bretas, "Optimal pmu allocation for enhanced gross error detection," in *2018 North American Power Symposium*, Sep 2018.

[9] K. A. Clements, "The impact of pseudo-measurements on state estimator accuracy," in *2011 IEEE Power and Energy Society General Meeting*, pp. 1–4, July 2011.

[10] N. Bretas, A. Bretas, and A. C. Martins, "Convergence property of the measurement gross error correction in power system state estimation, using geometrical background," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3729–3736, 2013.

[11] A. Bretas, N. Bretas, S. Braunstein, A. Rossoni, and R. Trevizan, "Multiple gross errors detection, identification and correction in three-phase distribution systems wls state estimation: A per-phase measurement error approach," *Electric Power Systems Research*, vol. 151, pp. 174–185, 2017.

[12] M. Göl and A. Abur, "A modified chi-squares test for improved bad data detection," in *PowerTech, 2015 IEEE Eindhoven*, pp. 1–5, IEEE, 2015.

[13] N. G. Bretas and A. S. Bretas, "The extension of the gauss approach for the solution of an overdetermined set of algebraic non linear equations," *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2018.

[14] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, pp. 12–19, Feb 2011.

This work was supported by NSF ECCS grants 1646229 and 1809739.

Authorized licensed use limited to: University of Florida. Downloaded on May 18,2020 at 18:47:09 UTC from IEEE Xplore. Restrictions apply.