The Extension of the Gauss Approach for the Solution of an Overdetermined Set of Algebraic Non-Linear Equations

Newton G. Bretas and Arturo S. Bretas

***Abstract*—In this brief is presented an extension of the Gauss approach for the solution of an overdetermined set of algebraic non-linear equations. Further, it is shown that the measurement error, in the Gauss approach, is in the measurement direction and that error has a unique decomposition: 1) the first component, which is orthogonal to the Jacobian range space, the residual and 2) the other which is on the Jacobian range space. The latter is hidden in the Jacobian space when one minimizes the residual. The extension of the Gauss approach is then in the sense to minimize the norm of the error. In engineering, the measurements may have gross errors, and then detection, identification, and correction of those errors are necessary. The Largest Normalized Error Test will be developed for that purpose. Considering the cyber-attack possibility, modeled as a malicious data attack, the error correction step is paramount. Applications on power networks will be used to show the hidden error component when using the Gauss minimization, and also to illustrate all the steps of the presented procedure as well as comparison to the current Gauss approach.**

***Index Terms*—Gauss minimization, orthogonal projections, recovering errors and composing errors.**

I. Introduction

The characteristic to deal with measurements in engineering and many other science's fields impose a necessity to filter those data, and the Gauss minimization is widely accepted as an appropriate approach. Many versions of the Gauss proposition have appeared and the most known one is the Weighted Least Squares Estimation (WLSE) [1]–[3]. The WLSE has several applications, as filter design [4], [5], cyber-physical security [6] and networks [7]. The Gauss approach proposes to minimize a functional given by the weighted residuals. Higher weights are attributed for the measurements of better qualities. However, if it is assumed the possibility of gross errors on the measurements, how could one attribute the weights to them? A second question that naturally arises is: (i) In case some measurements of the measurement set have gross errors will it be adequate to process them without a previous treatment? The intention of this brief is to provide answers to the previous questions and, at same time, propose solutions. In this brief, it will be shown that the error of the measurement equation has a unique decomposition: (i) one component orthogonal to the Jacobian range space and (ii) the other component is in the Jacobian range space. It will be shown also that the Gauss proposition minimizes only the first error component, known as the measurement residual; the other error component is hidden from the Gauss approach. Still, it will be shown how to estimate the other error component and then an extension of the Gauss proposition, where the minimization will be on the norm of error and not the residual. In order to estimate the hidden error component, the Innovation Index (II) will be introduced. The II of a measurement is the information it contains but not the other measurements of the measurement set. A two-bus power network will be used to show the hidden error component effect, when using the Gauss minimization, and two other networks will be used to illustrate all the steps of the presented procedure and used as comparison to the Gauss approach.

Manuscript received September 28, 2017; revised November 28, 2017; accepted January 18, 2018. This brief was recommended by Associate Editor H.-T. Zhang. *(Corresponding author: Newton G. Bretas.)*

N. G. Bretas is with the Department of Electrical and Computer Engineering, Engineering School of Sao Carlos, University of Sao Paulo, Sao Carlos 13566-590, Brazil (e-mail: ngbretas@sc.usp.br).

A. S. Bretas is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6130, USA.

Digital Object Identifier 10.1109/TCSII.2018.2796938

II. Gauss Approach: Theoretical Background

Given a set of non-linear equations, as described in the following:

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

where $z \in R^m$ is the measurement vector, $x \in R^N$ is the state variables vector, $h : R^m \to R^N$, $(m > N)$ is a continuously non linear differentiable function, $e \in R^m$ is the measurement error vector assumed having zero mean and Gaussian probability distribution. $N = 2n - 1$ is the number of unknown state variables to be estimated (n is the buses number of the network).

The Gauss approach proposes to minimize the functional [2]:

$$J(x) = [z - h(x)]^T W [z - h(x)] \tag{2}$$

where W is a symmetric and positive definite real matrix. At this stage of this brief, let us suppose $W = I$, that is, the weight matrix as being the identity matrix. The functional J is a norm in the measurement space R^m induced by the inner product $\langle u, v\rangle = u^T v$, that is:

$$\|e\|^2 = e^T e = \langle z - h(x), z - h(x)\rangle$$
$$\|e\|^2 = [z - h(x)]^T [z - h(x)] \tag{3}$$

with $\|e\|^2$ being the square norm of the vector e.

The solution of (1), minimizing (3), is the vector $\hat{x}$ obtained iteratively from the solution of the linearized equation:

$$z = Hx + e \tag{4}$$

where $H = \partial h/\partial x$ is the Jacobian of $h(x)$.

Let $\hat{x}$ be the estimated state, the solution of that minimization problem, then the estimated measurement vector will be given by $\hat{z} = h(\hat{x})$. The residual vector will be:

$$r = \left(z - \hat{z}\right) \in L\left(h_j\right)^{\perp}. \tag{5}$$

It happens that e, z and $H\hat{x}$ all belong to R^m, but $H\hat{x}$ belongs to the column space of H. That is, the space spanned by the columns h_j of H and is denoted by $L(h_j)$, or a linear combination of the column space of H. R^m can be written as the direct sum:

$$R^m = L\left(h_j\right) \oplus L\left(h_j\right)^{\perp}. \tag{6}$$

The measurement error, however, has a unique decomposition:

$$e = e_1 + e_2;\, e_1 \in L\left(h_j\right),\, e_2 \in L\left(h_j\right)^{\perp}. \tag{7}$$

Since z is given and $H\hat{x} \in L(H_j)$, the choice of x cannot affect e_1. As a direct conclusion, the least squares solution vector $\hat{x}$ is optimal only for case where $e_1 = 0$. This means that the projection of z on $L(h_j)$, call it $\hat{z}$ must equal $H\hat{x}$.

The important conclusion of this previous is that the WLSE is optimal only for the case where $e_1 = 0$. However, when one error component exists, the other one also will exist, then in real life, the measurement vector z has always an error vector, $e_1 \neq 0$, they are coupled.

Remark 1: One can induce that in case the measurement error e is contained inside an acceptable ball of radius r, that is, less than a threshold predefined value, and only in that case, the classical least mean squares will offer a reasonable solution. For the real life situations however, measurements may have large errors. Then, a generalization of the Least Squares approximation is required.

III. Extension of the Gauss Approach

Let us assume the possibility of the measurement error vector e having large magnitudes, that is, beyond a pre-defined acceptable value. If one wants to use the classical solution of the WLSE, one must first apply the Hypothesis Testing to check if the error vector magnitude is inside a pre-defined acceptable region. In case the answer to this question is yes, one could apply the classical WLS solution as previously presented, although that is not the optimal solution. Otherwise, one should identify the measurements having the error e larger than a chosen threshold value and then correct them.

To identify such measurements one needs to estimate the measurement error: the estimation of the measurement error can be obtained estimating the measurement residual as well as the hidden measurement error component, which is in the Jacobian range space. To do this the measurement Innovation Index (II) will be used, as presented in the following.

To estimate the measurement error (e), one need to estimate the error components e_1 and e_2, and then the measurement error will be composed, that is:

$$e^2 = e_1^2 + e_2^2. \tag{8}$$

Let $\hat{x}$ be the estimated state vector, result of the solution of the classical WLSE problem, which minimizes (3), and is given by:

$$\hat{x} = \left(H^T R^{-1} H\right)^{-1} H^T R^{-1} z = G^{-1} H^T R^{-1} z \tag{9}$$

with R being the weighting matrix. In this case, the estimated measurement vector $\hat{z}$ is given by:

$$\hat{z} = H\hat{x} = H\left(H^T R^{-1} H\right)^{-1} H^T R^{-1} z$$
$$\hat{z} = H G^{-1} H^T R^{-1} z = Kz \tag{10}$$

with G being equal to $H^T R^{-1} H$. The residual vector r is given by:

$$r = z - \hat{z} = z - Kz = (I - K)z = Se = e_2. \tag{11}$$

Since, for the i-th measurement, it is assumed $e^i \sim N(0, R_{ii}) = N(0, \sigma^2)$, then:

$$E\{r\} = E\{Se\} = SE\{e\} = 0 \tag{12}$$
$$cov\{r\} = E\left\{rr^T\right\} = SE\left\{ee^T\right\}S^T = SRS^T = SR = \Omega \tag{13}$$

where σ is the measurement standard deviation and:

$$\Omega = (I - K)R$$
$$\Omega = \begin{bmatrix} (1 - K_{11})\sigma_1^2 & K_{1m}\sigma_m^2 \\ K_{m1}\sigma_1^2 & (1 - K_{mm})\sigma_m^2 \end{bmatrix}. \tag{14}$$

Consequently, the normalized residual will be:

$$r_i^N = \frac{|r_i|}{\sqrt{\Omega_{ii}}} = \frac{|r_i|}{\sqrt{S_{ii} R_{ii}}} = \frac{|r_i|}{\sigma_i \sqrt{(1 - K_{ii})}}$$
$$r_i^N \sim N(0, 1) \tag{15}$$

and $e_1^i = Ke^i$ and $e_2^i = (I - K)e^i$, respectively.

To find the masked error component e_1, let us define the Innovation Index (II) for the i-th measurement as being:

$$II_i = \frac{\left\| e_2^{(i)} \right\|_{R^{-1}}}{\left\| e_1^{(i)} \right\|_{R^{-1}}} = \frac{\sqrt{1 - K_{ii}}}{\sqrt{K_{ii}}}. \tag{16}$$

Equation (16) is easily derived just using the definition of the norm and making some algebrism.

The II of a measurement is the ratio of new information it contains related to the other measurements of the measurement set [9]–[12]. Consequently, knowing the matrix K, the measurement II can be calculated; knowing this index, and the measurement residual, the e_1^i error component can be also estimated and the measurement error e composed by:

$$\|\hat{e}^i\|_{R^{-1}}^2 = \|e_2^i\|_{R^{-1}}^2 + \|e_1^i\|_{R^{-1}}^2 = \frac{1}{II_i^2}\|e_2^i\|_{R^{-1}}^2 + \|e_1^i\|_{R^{-1}}^2$$
$$\|\hat{e}^i\|_{R^{-1}}^2 = \left(1 + \frac{1}{II_i^2}\right)\|e_2^i\|_{R^{-1}}^2 = \left(1 + \frac{1}{II_i^2}\right) r_i^2. \tag{17}$$

Then, in this brief, the proposed Extension to the Gauss methodology is in the sense of by minimizing the error.

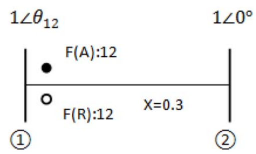
That is:

$$Min_x \sum_{i=1}^{m} (1 + \frac{1}{II_i^2}) r_i^2 \tag{18}$$

with r_i^2 given by (3).

The measurement error, estimated in this way, is known as Composed Measurement Error (CME). It can be normalized as:

$$CME^N = CME/\sigma. \tag{19}$$



AQ1 Fig. 1. Two-bus power network.

One can also obtain the measurement Composed Normalized Residual, that is, first normalizes the residual and then composes the error; consequently, the *CNE* is in the residual subspace:

$$CNE = \left(1 + \frac{1}{II^2}\right) r^N. \tag{20}$$

Remark 2: One should be aware that that CME^N and *CNE* are different quantities, the first pertaining to the error subspace and the other to the residual subspace.

For the gross error detection the χ^2 Hypothesis testing is applied to the CME^N vector, with the number of measurements m as degree of freedom; the reason for that being the error is not a correlated quantity.

For the gross error correction, (20) should be used since the CME^N, a vector of dimension m, is generated from the residual; this one being a vector of dimension $(m - n)$; as a consequence, the vector CME^N contains noises.

A two-bus power network will be used to show the application of the presented Gauss Extension formulation.

Consider the two-bus power network illustrated on Figure 1. F(A) : 12 = 1.0526 and F(R) = P : 12 = 0.120 p.u. are the active and reactive power flows measurements, respectively, and the phase angle difference between the two buses (θ_{12}) is the state variable to be estimated. Starting with an inicial phase angle difference estimate, the Jacobian matrix is calculated considering that state and the minimization of (18) is performed. The χ^2 Hypothesis testing is applied, and if that results in gross error detection, the Largest Normalized Error test is used to identify the measurement with error; the correction is made through the *CNE*.

Consider Figure 2, the plane formed by the F(A) : 12 and θ_{12} is represented.

Figure 2 shows clearly what the Gauss proposition minimizes is the residual, $(z - h(x))$, although what one needs to minimize is the error, pertaining to the measurement space, and that is the Extension of the Gauss approach this brief presents. The measurement F(R) will form a similar figure in another plane with θ_{12}.

A. Current Test for Detection and Identification of Gross Errors

With the normalized measurement residuals having normal distribution, it is assumed that the index $J(x)$, i.e., the function to be minimized in (3) has a Chi-square distribution (χ^2) with $m - N$ degrees of freedom. Then, choosing a probability "$1 - \alpha$" of false alarm (being α the significance level of the test), a number "C" is obtained (via Chi-square distribution table for $\chi^2_{m-N,\alpha}$) such that, in the presence of gross error: $J(x) > C$.

Remark 3: The previous assumption has the mistake of assuming the normalized residual as having a normal distribution. The reason for this affirmative is that m residuals has $m - N$ degrees of freedom and, consequently, cannot be an uncorrelated variable, contradicting then the assumption.

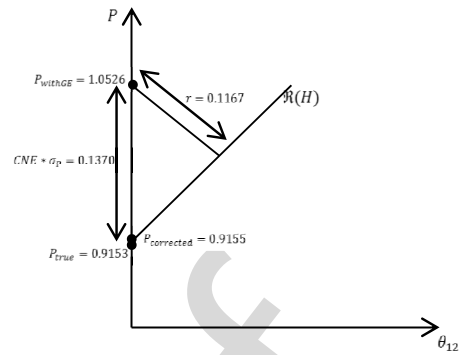


Fig. 2. The measurement error estimation.

This χ^2 test applied to the residual, however, is not adequately applied. Consequently, the tests results are not going to give the guarantee, with the chosen uncertainty degree, that at least one measurement may have the residual as superior to a chosen threshold value.

In the gross error detection test, this brief proposes to use the new $J(x)$ index with the measurement error and not the residual, as shown in (18).

For the purpose of gross error identification, the Largest Normalized Error test should then be used; after normalizing the error e it is submitted to the validation test:

Conjecture: Let be $e_k^N = |\frac{e_k}{\sigma_k}| \leq \lambda$ (threshold value), where e_k^N is the largest among all e_i^N, $i = 1, \ldots, m$; $\sigma_k = \sqrt{R_{kk}}$ is the *kth* measurement standard deviation, and R is the error covariance matrix. Then if $e_k^N > \lambda$, measurement gross error will be detected.

Remark 4: The current gross error detection test is in incorrect because: (i) it uses the residual as a metric for the error; (ii) it assumes a hyper-sphere for the normalized residual probability density function, assuming all the measurements are correct, except one of them, the one having error, what is not a valid real life situation. In what follows, it will be presented the property of the Largest Normalized Error Test (*LNET*), in order to detect and identify the measurements with errors as shown in the previous conjecture.

B. Generalization of the Largest Normalized Residual Test: The Largest Normalized Error Test

Theorem: Assuming all the measurements of a measurement set with limited random errors, and adding gross error only to one of the measurements, the measurement to which gross error was added will have the largest increment of error among all the measurements.

Proof: In [8] it was proved that the measurement error in its normalized form (CME^N) and the normalized residuals are formally equal, but numerically different from each other because the projection matrix K is different. Also, the measurement error pertains to the measurement sub-space, with m degrees of freedom, and the measurement residual pertains to the residual sub-space, of dimension $(m - N)$, and consequently with $(m-N)$ degrees of freedom and any comparison between them is meaningless.

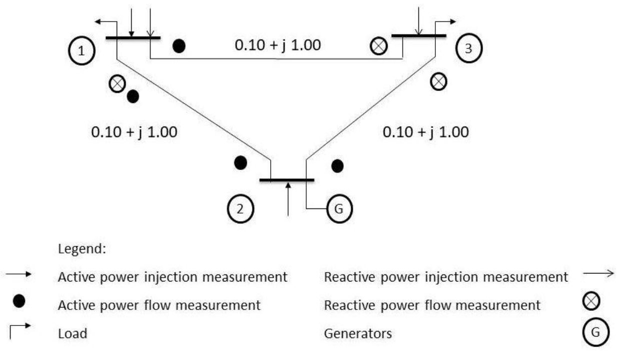


Fig. 3. The three-bus network.

Suppose, now, that the measurements are *perfect*, without any error, the measurement residual will be given by:

$$e_2 = r = (I - K)z = R_r W z \tag{21}$$

which in this situation are zeros for all the measurements. Let us suppose now that error is added only to the measurement z_i and, consequently, for this new measurement: $z_i^n = z_i + b_i\sigma_i$, and all the other measurements staying with the same magnitudes, n standing for *new* measurement. Following the same standard of demonstration of the classical largest normalized residual test, one obtains:

$$r^n - r = b_i \sigma_i^{-1} R_r u_i \tag{22}$$

where $R_r u_i$ is the i-th column of the residual covariance matrix R_r, as in the largest normalized residual test.

Consequently, and following the same standard of proof of the classical largest normalized residual test [2], one will obtain:

$$\left| r_i^{N,n} - r_i^N \right| \geq \left| r_j^{N,n} - r_j^N \right| \text{ for all } j \neq i. \tag{23}$$

Transforming the residuals in errors one will get:

$$\left| e_i^{N,n} - e_i^N \right| \geq \left| e_j^{N,n} - e_j^N \right|, j = 1, \ldots, m \tag{24}$$

because the measurement with error is the one with more new information. As a conclusion, the largest error increase will occur on the measurement with error. ■

If one generalizes the situation so that all measurements may have errors but are inside a ball of radius r and error is added to one measurement only, and using the previous theorem, one error magnitude will exist such that the largest error has to be in the measurement with error.

OBS: For the previous generalization the Implicit Function Theorem property was used, that is, if a property is valid for an initial condition it will be valid for any other since no bifurcation occurs going from one initial condition to the another. The previous demonstration gives support to use the largest normalized error test as a methodology to detect as well as to identify the measurements containing gross errors.

C. The Choice of the Weight Matrix to Detect Measurements With Gross Errors

To detect if there exist measurements containing gross errors, it must be assumed that all measurements may have error. This means one cannot use the measurement qualities as having effect on the measurement weights at the stage of gross error detection. To define an appropriate measurement weights at this SE stage let us suppose a simple problem where it does exist just one variable to estimate and many related measurements: suppose that all measurements have the same quality and do not have gross errors; then as a natural consequence, the weights for the measurements should be all equal. For example, the inverse of a fixed percentage of the measurement's magnitude, because in that way the variable estimated value will be an average of the measurements values. Suppose now a situation where some of the measurements may have error: if the measurement is such that, with the error, it has a value larger than the correct value, it will have less weight than a correct measurement and vice versa.

Conclusion: At the gross error detection stage, the measurement weights should be a fixed percentage of the measurement magnitude, independently of the measurement quality. At this stage, the measurement quality cannot be taken in account. Once the measurements with gross errors are identified and corrected, the weight matrix is as proposed in the classical state estimation, that is, the measurement quality should be now taken in account.

In the next section, two power networks will be simulated to show this brief proposition efficiency for the gross error analysis as well the fail of the classical gross error analysis proposal.

TABLE I
THREE BUS NETWORK SIMULATION

Measurement	*II*	Added GE	r^N	$CME^N_{Largest}$	*CNE*
I(A):1 = -1.62	1.03	-5.7σ	-3.04	-4.24	-5.09
GE identified with LNRT at I(A):1 → LNET: both correct					
I(A):2 = 2.59	1.88	-3.6σ	-3.04	-3.98	-3.44
GE identified with LNRT at I(A):2 → LNRT: both correct					
F(A):2-1 = 2.03	1.58	-3.7σ	-3.02	-3.25	-3.58
GE identified with LNRT at I(A):1 with r^N = -3.20 →failed LNET-Correct					
F(A):1-3 = 0.36	0.76	-5.2σ	-3.06	-3.08	-5.07
GE identified with LNRT at V:3 with r^N = -3.10 → failed LNET-correct					
F(A):2-3 = 0.56	1.71	-3.6σ	-3.06	-4.02	-3.54
GE identified with LNRT at F(A):2-3 → both correct					
I(R):1 = 0.37	0.70	-5.3σ	-3.01	-4.28	-5.22
GE identified with LNRT at F(A):1-3, r^N = -3.10 → failed LNET-correct					
I(R):3 = 0.40	1.49	-3.3σ	-3.02	-3.24	-3.64
GE identified with LNRT at F(R):3-2 , r^N = -3.10 → failed LNET-correct					
F(R):1-2 =0.37	0.72	-5.5σ	-3.04	-3.11	-5.21
GE identified with LNRTat V:3, r^N = -3.12 → failed LNET-correct					
F(R):3-1 = 0.14	0.97	-4.7σ	-3.03	-4.23	-4.35
GE identified with LNRT at F(R):3-2, r^N = -3.01 → failed					
LNET-correct					
F(R):3-2 = 0.26	1.47	-3.9σ	-3.07	-3.58	-3.71
GE identified with LNRT at F(R):3-2 → LNET: both correct					

IV. CASE STUDY

Consider Figure 3. Using a three-bus network with the measurements as shown in Figure 3, whose magnitudes are in Table I, after random errors are added to them, a gross error is added, one at a time, until the error is detected. Then

TABLE II
IEEE 14 BUS NETWORK SIMULATIONS

Measurement	***II***	**Added GE**	r^N	CME^N	***CNE***
Step 1					
P:2	0.242	3.9σ	1.946	3.52	4.35
LNRT failed		LNET correct			
Step 2					
Q:4	0.369	4.5σ	2.054	4.34	4.51
LNRT failed		LNET correct			
Step 3					
Q:3-2	0.628	3.6σ	2.752	3.27	4.17
LNRT failed		LNET correct			
Step 4					
P:3-4	0.707	3.2σ	2.402	3.14	3.87
LNRT failed		LNET correct			
Step 5					
P:12-13	1.053	3.7σ	2.845	3.74	3.92
LNRT failed		LNET correct			
Step 6					
Q:2-5	0.863	3.5σ	2.152	3.29	3.48
LNRT failed		LNET correct			
Step 7					
P:6	0.622	4.1σ	1.690	3.20	3.69
LNRT failed		LNET correct			
Step 8					
Q:1	1.483	3.3σ	3.009	3.40	3.61
LNRT correct		LNET correct			
Step 9					
Q:5	0.175	4.0σ	1.612	3.54	3.98
LNRT failed		LNET correct			
Step 10					
P:14-9	1.683	3.4σ	3.053	3.31	3.60
LNRT correct		LNET correct			

the measurement with error is identified using the current Largest Normalized Residual Test *(LNRT)* and the presented methodology of this brief, the *LNET*. For this case, the global redundancy level is equal to 2.2 ($GRL = 2.2$). As can be seen, using the classical approach, six wrong gross errors identification have occurred, but using this brief proposition, in all cases the gross error was correctly detected and the measurement which had gross error was correctly identified.

In Table I, the used nomenclature for the measurements are: $I(A)$: i means active power injection at bus i; $F(R)$: $i{-}j$ means reactive flow at line $i-j$.

As can be seen on Table I, the gross error detection and identification has worked correctly in all cases when using the *LNET* but in contrary, when using the *LNRT*, many detection and identification flaws have occurred. Another important point is the proximity between the added error and the corresponding *CNE* for the measurement identified as having the error. Consider now the IEEE14 bus system [4], and the same test conditions as previously applied. Test results are presented on Table II.

As can be seen on Table II, the gross error detection and identification has worked correctly in all simulated cases when using the *LNET* but in contrary, when using the residual *LNRT*, many fails occurred. In general, measurements with low Innovation Index tend to present a hidden error component of large magnitude, consequently leading the LNRT to fail in the error detection.

V. CONCLUSION

This brief presents an extension of the Gauss approach to solve overdetermined sets of algebraic non linear equations. It was shown that the measurement error is composed of two components: (i) one that is in the Jacobian range space (hidden component); (ii) other that is orthogonal to that space. The latter component comes out to be the residual. The WLSE should perform the minimization on the measurement error and not on the residual. The gross error detection is performed through the application of the χ^2 test on the composed error; the identification is performed using the *LNET*. In this stage of the SE the measurement weights should be a percentage of the measurements magnitude. The measurement correction is made using the measurement *CNE*. After the correction, a new Gauss minimization is performed, now taking in account the measurement quality to define the weights. This brief contribution when estimating the error and then the possibility of correcting the measurement having error is of paramount importance in these days due to cyber-attacks becoming every day more real. Two power networks have been used to show, in detail, each step of this brief proposition and at same time comparing the classical WLS solution. The added error magnitude and the estimated normalized error, *CNE*, are very close to each other, giving support to the idea of estimating the measurement errors and then correcting those measurements.

REFERENCES

[1] D. M. Falcao and M. A. Arias, "State estimation and observability analysis based on echelon forms of the linearized measurement models," *IEEE Trans. Power Syst.*, vol. 9, no. 2, pp. 979–987, May 1994.

[2] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, Feb. 2000.

[3] E. Handschin, F. C. Schweppe, J. Kohlas, and A. Fiechter, "Bad data analysis for power system state estimation," *IEEE Trans. Power App. Syst.*, vol. PAS-94, no. 2, pp. 329–337, Mar. 1975.

[4] A. Jiang, H. K. Kwan, Y. Zhu, N. Xu, and X. Liu, "Efficient WLS design of IIR digital filters using partial second-order factorization," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 7, pp. 703–707, Jul. 2016.

[5] S. Wang, W. Wang, S. Duan, and L. Wang, "Kernel Recursive least squares with multiple feedback and its convergence analysis," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 64, no. 10, pp. 1237–1241, Oct. 2017.

[6] S. Zonouz *et al.*, "SCPSE: Security-oriented cyber-physical state estimation for power grid critical infrastructures," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1790–1799, Dec. 2012.

[7] B. K. Das, M. Chakraborty, and J. Arenas-García, "Sparse distributed estimation via heterogeneous diffusion adaptive networks," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 11, pp. 1079–1083, Nov. 2016.

[8] N. G. Bretas, A. S. Bretas, and S. A. Piereti, "Innovation concept for measurement gross error detection and identification in power system state estimation," *IET Gener. Transm. Distrib.*, vol. 5, no. 6, pp. 603–608, Jun. 2011.

[9] 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 Trans. Power Syst.*, vol. 28, no. 3, pp. 2128–2135, Aug. 2013.

[10] N. G. Bretas and A. S. Bretas, "A two steps procedure in state estimation gross error detection, identification, and correction," *Int. J. Elect. Power Energy Syst.*, vol. 73, pp. 484–490, Dec. 2015.

[11] A. S. Bretas, N. G. Bretas, B. Carvalho, E. Bayens, and P. P. Khargonekar, "Smart grids cyber-physical security as a malicious data attack: An innovation approach," *Elect. Power Syst. Res.*, vol. 149, pp. 210–219, Aug. 2017.

[12] A. S. Bretas, N. G. Bretas, S. H. Braunstein, A. Rossoni, and R. D. Trevizan, "Multiple gross errors detection, identification and correction in three-phase distribution systems WLS State estimation: A per-phase measurement error approach," *Elect. Power Syst. Res.*, vol. 151, pp. 174–185, Oct. 2017.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

PLEASE NOTE: We cannot accept new source files as corrections for your paper. If possible, please annotate the PDF proof we have sent you with your corrections and upload it via the Author Gateway. Alternatively, you may send us your corrections in list format. You may also upload revised graphics via the Author Gateway.

AQ1: Please provide the better quality image for all the figures.

The Extension of the Gauss Approach for the Solution of an Overdetermined Set of Algebraic Non Linear Equations

Newton G. Bretas and Arturo S. Bretas

***Abstract*—In this brief is presented an extension of the Gauss approach for the solution of an overdetermined set of algebraic non linear equations. Further, it is shown that the measurement error, in the Gauss approach, is in the measurement direction and that error has a unique decomposition: 1) the first component, which is orthogonal to the Jacobian range space, the residual and 2) the other which is on the Jacobian range space. The latter is hidden in the Jacobian space when one minimizes the residual. The extension of the Gauss approach is then in the sense to minimize the norm of the error. In engineering, the measurements may have gross errors, and then detection, identification, and correction of those errors are necessary. The Largest Normalized Error Test will be developed for that purpose. Considering the cyber-attack possibility, modeled as a malicious data attack, the error correction step is paramount. Applications on power networks will be used to show the hidden error component when using the Gauss minimization, and also to illustrate all the steps of the presented procedure as well as comparison to the current Gauss approach.**

***Index Terms*—Gauss minimization, orthogonal projections, recovering errors and composing errors.**

I. Introduction

THE CHARACTERISTIC to deal with measurements in engineering and many other science's fields impose a necessity to filter those data, and the Gauss minimization is widely accepted as an appropriate approach. Many versions of the Gauss proposition have appeared and the most known one is the Weighted Least Squares Estimation (WLSE) [1]–[3]. The WLSE has several applications, as filter design [4], [5], cyber-physical security [6] and networks [7]. The Gauss approach proposes to minimize a functional given by the weighted residuals. Higher weights are attributed for the measurements of better qualities. However, if it is assumed the possibility of gross errors on the measurements, how could one attribute the weights to them? A second question that naturally arises is: (i) In case some measurements of the measurement set have gross errors will it be adequate to process them without a previous treatment? The intention of this brief is to provide answers to the previous questions and, at same time, propose solutions. In this brief, it will be shown that the error of the measurement equation has a unique decomposition: (i) one component orthogonal to the Jacobian range space and (ii) the other component is in the Jacobian range space. It will be shown also that the Gauss proposition minimizes only the first error component, known as the measurement residual; the other error component is hidden from the Gauss approach. Still, it will be shown how to estimate the other error component and then an extension of the Gauss proposition, where the minimization will be on the norm of error and not the residual. In order to estimate the hidden error component, the Innovation Index (*II*) will be introduced. The *II* of a measurement is the information it contains but not the other measurements of the measurement set. A two-bus power network will be used to show the hidden error component effect, when using the Gauss minimization, and two other networks will be used to illustrate all the steps of the presented procedure and used as comparison to the Gauss approach.

Manuscript received September 28, 2017; revised November 28, 2017; accepted January 18, 2018. This brief was recommended by Associate Editor H.-T. Zhang. *(Corresponding author: Newton G. Bretas.)*

N. G. Bretas is with the Department of Electrical and Computer Engineering, Engineering School of Sao Carlos, University of Sao Paulo, Sao Carlos 13566-590, Brazil (e-mail: ngbretas@sc.usp.br).

A. S. Bretas is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6130, USA.

Digital Object Identifier 10.1109/TCSII.2018.2796938

II. Gauss Approach: Theoretical Background

Given a set of non-linear equations, as described in the following:

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

where $z \in R^m$ is the measurement vector, $x \in R^N$ is the state variables vector, $h : R^m \to R^N$, $(m > N)$ is a continuously non linear differentiable function, $e \in R^m$ is the measurement error vector assumed having zero mean and Gaussian probability distribution. $N = 2n - 1$ is the number of unknown state variables to be estimated (n is the buses number of the network).

The Gauss approach proposes to minimize the functional [2]:

$$J(x) = [z - h(x)]^T W[z - h(x)] \tag{2}$$

where W is a symmetric and positive definite real matrix. At this stage of this brief, let us suppose $W = I$, that is, the weight matrix as being the identity matrix. The functional J is a norm in the measurement space R^m induced by the inner product $\langle u, v\rangle = u^T v$, that is:

$$\|e\|^2 = e^T e = \langle z - h(x), z - h(x)\rangle$$
$$\|e\|^2 = [z - h(x)]^T [z - h(x)] \tag{3}$$

with $\|e\|^2$ being the square norm of the vector e.

The solution of (1), minimizing (3), is the vector $\hat{x}$ obtained iteratively from the solution of the linearized equation:

$$z = Hx + e \tag{4}$$

where $H = \partial h/\partial x$ is the Jacobian of $h(x)$.

Let $\hat{x}$ be the estimated state, the solution of that minimization problem, then the estimated measurement vector will be given by $\hat{z} = h(\hat{x})$. The residual vector will be:

$$r = \left(z - \hat{z}\right) \in L\left(h_j\right)^{\perp}. \tag{5}$$

It happens that e, z and $H\hat{x}$ all belong to R^m, but $H\hat{x}$ belongs to the column space of H. That is, the space spanned by the columns h_j of H and is denoted by $L(h_j)$, or a linear combination of the column space of H. R^m can be written as the direct sum:

$$R^m = L\left(h_j\right) \oplus L\left(h_j\right)^{\perp}. \tag{6}$$

The measurement error, however, has a unique decomposition:

$$e = e_1 + e_2;\, e_1 \in L\left(h_j\right), e_2 \in L\left(h_j\right)^{\perp}. \tag{7}$$

Since z is given and $H\hat{x} \in L(H_j)$, the choice of x cannot affect e_1. As a direct conclusion, the least squares solution vector $\hat{x}$ is optimal only for case where $e_1 = 0$. This means that the projection of z on $L(h_j)$, call it $\hat{z}$ must equal $H\hat{x}$.

The important conclusion of this previous is that the WLSE is optimal only for the case where $e_1 = 0$. However, when one error component exists, the other one also will exist, then in real life, the measurement vector z has always an error vector, $e_1 \neq 0$, they are coupled.

Remark 1: One can induce that in case the measurement error e is contained inside an acceptable ball of radius r, that is, less than a threshold predefined value, and only in that case, the classical least mean squares will offer a reasonable solution. For the real life situations however, measurements may have large errors. Then, a generalization of the Least Squares approximation is required.

III. Extension of the Gauss Approach

Let us assume the possibility of the measurement error vector e having large magnitudes, that is, beyond a pre-defined acceptable value. If one wants to use the classical solution of the WLSE, one must first apply the Hypothesis Testing to check if the error vector magnitude is inside a pre-defined acceptable region. In case the answer to this question is yes, one could apply the classical WLS solution as previously presented, although that is not the optimal solution. Otherwise, one should identify the measurements having the error e larger than a chosen threshold value and then correct them.

To identify such measurements one needs to estimate the measurement error: the estimation of the measurement error can be obtained estimating the measurement residual as well as the hidden measurement error component, which is in the Jacobian range space. To do this the measurement Innovation Index (II) will be used, as presented in the following.

To estimate the measurement error (e), one need to estimate the error components e_1 and e_2, and then the measurement error will be composed, that is:

$$e^2 = e_1^2 + e_2^2. \tag{8}$$

Let $\hat{x}$ be the estimated state vector, result of the solution of the classical WLSE problem, which minimizes (3), and is given by:

$$\hat{x} = \left(H^T R^{-1} H\right)^{-1} H^T R^{-1} z = G^{-1} H^T R^{-1} z \tag{9}$$

with R being the weighting matrix. In this case, the estimated measurement vector $\hat{z}$ is given by:

$$\hat{z} = H\hat{x} = H\left(H^T R^{-1} H\right)^{-1} H^T R^{-1} z$$
$$\hat{z} = HG^{-1} H^T R^{-1} z = Kz \tag{10}$$

with G being equal to $H^T R^{-1} H$. The residual vector r is given by:

$$r = z - \hat{z} = z - Kz = (I - K)z = Se = e_2. \tag{11}$$

Since, for the i-th measurement, it is assumed $e^i \sim N(0, R_{ii}) = N(0, \sigma^2)$, then:

$$E\{r\} = E\{Se\} = SE\{e\} = 0 \tag{12}$$
$$cov\{r\} = E\left\{rr^T\right\} = SE\left\{ee^T\right\}S^T = SRS^T = SR = \Omega \tag{13}$$

where σ is the measurement standard deviation and:

$$\Omega = (I - K)R$$
$$\Omega = \begin{bmatrix} (1 - K_{11})\sigma_1^2 & K_{1m}\sigma_m^2 \\ K_{m1}\sigma_1^2 & (1 - K_{mm})\sigma_m^2 \end{bmatrix}. \tag{14}$$

Consequently, the normalized residual will be:

$$r_i^N = \frac{|r_i|}{\sqrt{\Omega_{ii}}} = \frac{|r_i|}{\sqrt{S_{ii} R_{ii}}} = \frac{|r_i|}{\sigma_i \sqrt{(1 - K_{ii})}}$$
$$r_i^N \sim N(0, 1) \tag{15}$$

and $e_1^i = Ke^i$ and $e_2^i = (I - K)e^i$, respectively.

To find the masked error component e_1, let us define the Innovation Index (II) for the i-th measurement as being:

$$II_i = \frac{\left\| e_2^{(i)} \right\|_{R^{-1}}}{\left\| e_1^{(i)} \right\|_{R^{-1}}} = \frac{\sqrt{1 - K_{ii}}}{\sqrt{K_{ii}}}. \tag{16}$$

Equation (16) is easily derived just using the definition of the norm and making some algebrism.

The II of a measurement is the ratio of new information it contains related to the other measurements of the measurement set [9]–[12]. Consequently, knowing the matrix K, the measurement II can be calculated; knowing this index, and the measurement residual, the e_1^i error component can be also estimated and the measurement error e composed by:

$$\|\hat{e}^i\|_{R^{-1}}^2 = \|e_2^i\|_{R^{-1}}^2 + \|e_1^i\|_{R^{-1}}^2 = \frac{1}{II_i^2}\|e_2^i\|_{R^{-1}}^2 + \|e_1^i\|_{R^{-1}}^2$$
$$\|\hat{e}^i\|_{R^{-1}}^2 = \left(1 + \frac{1}{II_i^2}\right)\|e_2^i\|_{R^{-1}}^2 = \left(1 + \frac{1}{II_i^2}\right) r_i^2. \tag{17}$$

Then, in this brief, the proposed Extension to the Gauss methodology is in the sense of by minimizing the error.

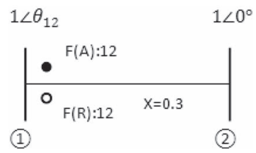
That is:

$$Min_x \sum_{i=1}^{m} (1 + \frac{1}{II_i^2}) r_i^2 \tag{18}$$

with r_i^2 given by (3).

The measurement error, estimated in this way, is known as Composed Measurement Error (CME). It can be normalized as:

$$CME^N = CME/\sigma. \tag{19}$$



AQ1 Fig. 1. Two-bus power network.

One can also obtain the measurement Composed Normalized Residual, that is, first normalizes the residual and then composes the error; consequently, the *CNE* is in the residual subspace:

$$CNE = \left(1 + \frac{1}{II^2}\right) r^N. \tag{20}$$

Remark 2: One should be aware that that CME^N and *CNE* are different quantities, the first pertaining to the error subspace and the other to the residual subspace.

For the gross error detection the χ^2 Hypothesis testing is applied to the CME^N vector, with the number of measurements m as degree of freedom; the reason for that being the error is not a correlated quantity.

For the gross error correction, (20) should be used since the CME^N, a vector of dimension m, is generated from the residual; this one being a vector of dimension $(m - n)$; as a consequence, the vector CME^N contains noises.

A two-bus power network will be used to show the application of the presented Gauss Extension formulation.

Consider the two-bus power network illustrated on Figure 1. F(A) : 12 = 1.0526 and F(R) = P : 12 = 0.120 p.u. are the active and reactive power flows measurements, respectively, and the phase angle difference between the two buses (θ_{12}) is the state variable to be estimated. Starting with an inicial phase angle difference estimate, the Jacobian matrix is calculated considering that state and the minimization of (18) is performed. The χ^2 Hypothesis testing is applied, and if that results in gross error detection, the Largest Normalized Error test is used to identify the measurement with error; the correction is made through the *CNE*.

Consider Figure 2, the plane formed by the F(A) : 12 and θ_{12} is represented.

Figure 2 shows clearly what the Gauss proposition minimizes is the residual, $(z - h(x))$, although what one needs to minimize is the error, pertaining to the measurement space, and that is the Extension of the Gauss approach this brief presents. The measurement F(R) will form a similar figure in another plane with θ_{12}.

A. Current Test for Detection and Identification of Gross Errors

With the normalized measurement residuals having normal distribution, it is assumed that the index $J(x)$, i.e., the function to be minimized in (3) has a Chi-square distribution (χ^2) with $m - N$ degrees of freedom. Then, choosing a probability "$1 - \alpha$" of false alarm (being α the significance level of the test), a number "C" is obtained (via Chi-square distribution table for $\chi^2_{m-N,\alpha}$) such that, in the presence of gross error: $J(x) > C$.

Remark 3: The previous assumption has the mistake of assuming the normalized residual as having a normal distribution. The reason for this affirmative is that m residuals has $m - N$ degrees of freedom and, consequently, cannot be an uncorrelated variable, contradicting then the assumption.

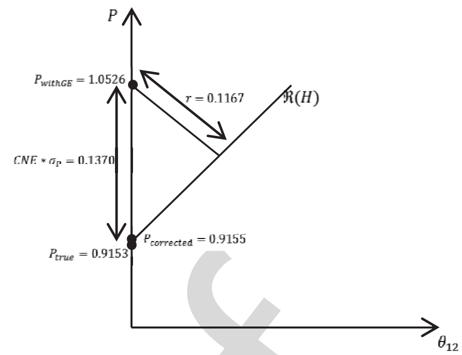


Fig. 2. The measurement error estimation.

This χ^2 test applied to the residual, however, is not adequately applied. Consequently, the tests results are not going to give the guarantee, with the chosen uncertainty degree, that at least one measurement may have the residual as superior to a chosen threshold value.

In the gross error detection test, this brief proposes to use the new $J(x)$ index with the measurement error and not the residual, as shown in (18).

For the purpose of gross error identification, the Largest Normalized Error test should then be used; after normalizing the error e it is submitted to the validation test:

Conjecture: Let be $e_k^N = |\frac{e_k}{\sigma_k}| \leq \lambda$ (threshold value), where e_k^N is the largest among all e_i^N, $i = 1, \ldots, m$; $\sigma_k = \sqrt{R_{kk}}$ is the *kth* measurement standard deviation, and R is the error covariance matrix. Then if $e_k^N > \lambda$, measurement gross error will be detected.

Remark 4: The current gross error detection test is in incorrect because: (i) it uses the residual as a metric for the error; (ii) it assumes a hyper-sphere for the normalized residual probability density function, assuming all the measurements are correct, except one of them, the one having error, what is not a valid real life situation. In what follows, it will be presented the property of the Largest Normalized Error Test (*LNET*), in order to detect and identify the measurements with errors as shown in the previous conjecture.

B. Generalization of the Largest Normalized Residual Test: The Largest Normalized Error Test

Theorem: Assuming all the measurements of a measurement set with limited random errors, and adding gross error only to one of the measurements, the measurement to which gross error was added will have the largest increment of error among all the measurements.

Proof: In [8] it was proved that the measurement error in its normalized form (CME^N) and the normalized residuals are formally equal, but numerically different from each other because the projection matrix K is different. Also, the measurement error pertains to the measurement sub-space, with m degrees of freedom, and the measurement residual pertains to the residual sub-space, of dimension $(m - N)$, and consequently with $(m-N)$ degrees of freedom and any comparison between them is meaningless.

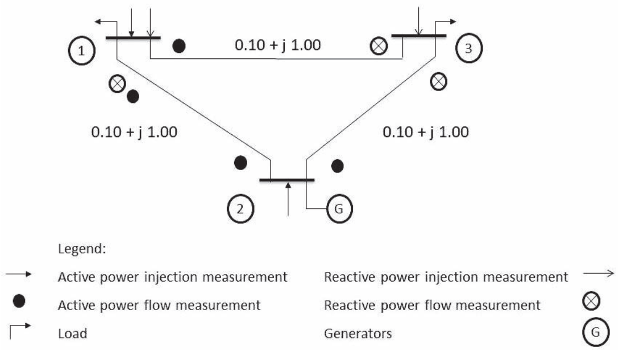


Fig. 3. The three-bus network.

Suppose, now, that the measurements are *perfect*, without any error, the measurement residual will be given by:

$$e_2 = r = (I - K)z = R_r W z \tag{21}$$

which in this situation are zeros for all the measurements. Let us suppose now that error is added only to the measurement z_i and, consequently, for this new measurement: $z_i^n = z_i + b_i\sigma_i$, and all the other measurements staying with the same magnitudes, n standing for *new* measurement. Following the same standard of demonstration of the classical largest normalized residual test, one obtains:

$$r^n - r = b_i\sigma_i^{-1} R_r u_i \tag{22}$$

where $R_r u_i$ is the i-th column of the residual covariance matrix R_r, as in the largest normalized residual test.

Consequently, and following the same standard of proof of the classical largest normalized residual test [2], one will obtain:

$$\left| r_i^{N,n} - r_i^N \right| \geq \left| r_j^{N,n} - r_j^N \right| \text{ for all } j \neq i. \tag{23}$$

Transforming the residuals in errors one will get:

$$\left| e_i^{N,n} - e_i^N \right| \geq \left| e_j^{N,n} - e_j^N \right|, j = 1, \ldots, m \tag{24}$$

because the measurement with error is the one with more new information. As a conclusion, the largest error increase will occur on the measurement with error. ■

If one generalizes the situation so that all measurements may have errors but are inside a ball of radius r and error is added to one measurement only, and using the previous theorem, one error magnitude will exist such that the largest error has to be in the measurement with error.

OBS: For the previous generalization the Implicit Function Theorem property was used, that is, if a property is valid for an initial condition it will be valid for any other since no bifurcation occurs going from one initial condition to the another. The previous demonstration gives support to use the largest normalized error test as a methodology to detect as well as to identify the measurements containing gross errors.

C. The Choice of the Weight Matrix to Detect Measurements With Gross Errors

To detect if there exist measurements containing gross errors, it must be assumed that all measurements may have error. This means one cannot use the measurement qualities as having effect on the measurement weights at the stage of gross error detection. To define an appropriate measurement weights at this SE stage let us suppose a simple problem where it does exist just one variable to estimate and many related measurements: suppose that all measurements have the same quality and do not have gross errors; then as a natural consequence, the weights for the measurements should be all equal. For example, the inverse of a fixed percentage of the measurement's magnitude, because in that way the variable estimated value will be an average of the measurements values. Suppose now a situation where some of the measurements may have error: if the measurement is such that, with the error, it has a value larger than the correct value, it will have less weight than a correct measurement and vice versa.

Conclusion: At the gross error detection stage, the measurement weights should be a fixed percentage of the measurement magnitude, independently of the measurement quality. At this stage, the measurement quality cannot be taken in account. Once the measurements with gross errors are identified and corrected, the weight matrix is as proposed in the classical state estimation, that is, the measurement quality should be now taken in account.

In the next section, two power networks will be simulated to show this brief proposition efficiency for the gross error analysis as well the fail of the classical gross error analysis proposal.

TABLE I
THREE BUS NETWORK SIMULATION

Measurement	***II***	**Added GE**	r^N	***CME**^{N} **Largest***	***CNE***
I(A):1 = -1.62	1.03	-5.7σ	-3.04	-4.24	-5.09
GE identified with LNRT at I(A):1 → LNET: both correct					
I(A):2 = 2.59	1.88	-3.6σ	-3.04	-3.98	-3.44
GE identified with LNRT at I(A):2 → LNRT: both correct					
F(A):2-1 = 2.03	1.58	-3.7σ	-3.02	-3.25	-3.58
GE identified with LNRT at I(A):1 with r^N = -3.20 →failed LNET-Correct					
F(A):1-3 = 0.36	0.76	-5.2σ	-3.06	-3.08	-5.07
GE identified with LNRT at V:3 with r^N = -3.10 → failed LNET-correct					
F(A):2-3 = 0.56	1.71	-3.6σ	-3.06	-4.02	-3.54
GE identified with LNRT at F(A):2-3 → both correct					
I(R):1 = 0.37	0.70	-5.3σ	-3.01	-4.28	-5.22
GE identified with LNRT at F(A):1-3, r^N = -3.10 → failed LNET-correct					
I(R):3 = 0.40	1.49	-3.3σ	-3.02	-3.24	-3.64
GE identified with LNRT at F(R):3-2 , r^N = -3.10 → failed LNET-correct					
F(R):1-2 =0.37	0.72	-5.5σ	-3.04	-3.11	-5.21
GE identified with LNRTat V:3, r^N = -3.12 → failed LNET-correct					
F(R):3-1 = 0.14	0.97	-4.7σ	-3.03	-4.23	-4.35
GE identified with LNRT at F(R):3-2, r^N = -3.01 → failed LNET-correct					
F(R):3-2 = 0.26	1.47	-3.9σ	-3.07	-3.58	-3.71
GE identified with LNRT at F(R):3-2 → LNET: both correct					

IV. CASE STUDY

Consider Figure 3. Using a three-bus network with the measurements as shown in Figure 3, whose magnitudes are in Table I, after random errors are added to them, a gross error is added, one at a time, until the error is detected. Then

TABLE II
IEEE 14 BUS NETWORK SIMULATIONS

Measurement	***II***	**Added GE**	r^N	CME^N	***CNE***
Step 1					
P:2	0.242	3.9σ	1.946	3.52	4.35
LNRT failed			LNET correct		
Step 2					
Q:4	0.369	4.5σ	2.054	4.34	4.51
LNRT failed			LNET correct		
Step 3					
Q:3-2	0.628	3.6σ	2.752	3.27	4.17
LNRT failed			LNET correct		
Step 4					
P:3-4	0.707	3.2σ	2.402	3.14	3.87
LNRT failed			LNET correct		
Step 5					
P:12-13	1.053	3.7σ	2.845	3.74	3.92
LNRT failed			LNET correct		
Step 6					
Q:2-5	0.863	3.5σ	2.152	3.29	3.48
LNRT failed			LNET correct		
Step 7					
P:6	0.622	4.1σ	1.690	3.20	3.69
LNRT failed			LNET correct		
Step 8					
Q:1	1.483	3.3σ	3.009	3.40	3.61
LNRT correct			LNET correct		
Step 9					
Q:5	0.175	4.0σ	1.612	3.54	3.98
LNRT failed			LNET correct		
Step 10					
P:14-9	1.683	3.4σ	3.053	3.31	3.60
LNRT correct			LNET correct		

the measurement with error is identified using the current Largest Normalized Residual Test *(LNRT)* and the presented methodology of this brief, the *LNET*. For this case, the global redundancy level is equal to 2.2 ($GRL = 2.2$). As can be seen, using the classical approach, six wrong gross errors identification have occurred, but using this brief proposition, in all cases the gross error was correctly detected and the measurement which had gross error was correctly identified.

In Table I, the used nomenclature for the measurements are: $I(A) : i$ means active power injection at bus i; $F(R) : i{-}j$ means reactive flow at line $i - j$.

As can be seen on Table I, the gross error detection and identification has worked correctly in all cases when using the *LNET* but in contrary, when using the *LNRT*, many detection and identification flaws have occurred. Another important point is the proximity between the added error and the corresponding *CNE* for the measurement identified as having the error. Consider now the IEEE14 bus system [4], and the same test conditions as previously applied. Test results are presented on Table II.

As can be seen on Table II, the gross error detection and identification has worked correctly in all simulated cases when using the *LNET* but in contrary, when using the residual *LNRT*, many fails occurred. In general, measurements with low Innovation Index tend to present a hidden error component of large magnitude, consequently leading the LNRT to fail in the error detection.

V. CONCLUSION

This brief presents an extension of the Gauss approach to solve overdetermined sets of algebraic non linear equations. It was shown that the measurement error is composed of two components: (i) one that is in the Jacobian range space (hidden component); (ii) other that is orthogonal to that space. The latter component comes out to be the residual. The WLSE should perform the minimization on the measurement error and not on the residual. The gross error detection is performed through the application of the χ^2 test on the composed error; the identification is performed using the *LNET*. In this stage of the SE the measurement weights should be a percentage of the measurements magnitude. The measurement correction is made using the measurement *CNE*. After the correction, a new Gauss minimization is performed, now taking in account the measurement quality to define the weights. This brief contribution when estimating the error and then the possibility of correcting the measurement having error is of paramount importance in these days due to cyber-attacks becoming every day more real. Two power networks have been used to show, in detail, each step of this brief proposition and at same time comparing the classical WLS solution. The added error magnitude and the estimated normalized error, *CNE*, are very close to each other, giving support to the idea of estimating the measurement errors and then correcting those measurements.

REFERENCES

[1] D. M. Falcao and M. A. Arias, "State estimation and observability analysis based on echelon forms of the linearized measurement models," *IEEE Trans. Power Syst.*, vol. 9, no. 2, pp. 979–987, May 1994.

[2] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, Feb. 2000.

[3] E. Handschin, F. C. Schweppe, J. Kohlas, and A. Fiechter, "Bad data analysis for power system state estimation," *IEEE Trans. Power App. Syst.*, vol. PAS-94, no. 2, pp. 329–337, Mar. 1975.

[4] A. Jiang, H. K. Kwan, Y. Zhu, N. Xu, and X. Liu, "Efficient WLS design of IIR digital filters using partial second-order factorization," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 7, pp. 703–707, Jul. 2016.

[5] S. Wang, W. Wang, S. Duan, and L. Wang, "Kernel Recursive least squares with multiple feedback and its convergence analysis," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 64, no. 10, pp. 1237–1241, Oct. 2017.

[6] S. Zonouz *et al.*, "SCPSE: Security-oriented cyber-physical state estimation for power grid critical infrastructures," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1790–1799, Dec. 2012.

[7] B. K. Das, M. Chakraborty, and J. Arenas-García, "Sparse distributed estimation via heterogeneous diffusion adaptive networks," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 11, pp. 1079–1083, Nov. 2016.

[8] N. G. Bretas, A. S. Bretas, and S. A. Piereti, "Innovation concept for measurement gross error detection and identification in power system state estimation," *IET Gener. Transm. Distrib.*, vol. 5, no. 6, pp. 603–608, Jun. 2011.

[9] 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 Trans. Power Syst.*, vol. 28, no. 3, pp. 2128–2135, Aug. 2013.

[10] N. G. Bretas and A. S. Bretas, "A two steps procedure in state estimation gross error detection, identification, and correction," *Int. J. Elect. Power Energy Syst.*, vol. 73, pp. 484–490, Dec. 2015.

[11] A. S. Bretas, N. G. Bretas, B. Carvalho, E. Bayens, and P. P. Khargonekar, "Smart grids cyber-physical security as a malicious data attack: An innovation approach," *Elect. Power Syst. Res.*, vol. 149, pp. 210–219, Aug. 2017.

[12] A. S. Bretas, N. G. Bretas, S. H. Braunstein, A. Rossoni, and R. D. Trevizan, "Multiple gross errors detection, identification and correction in three-phase distribution systems WLS State estimation: A per-phase measurement error approach," *Elect. Power Syst. Res.*, vol. 151, pp. 174–185, Oct. 2017.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

PLEASE NOTE: We cannot accept new source files as corrections for your paper. If possible, please annotate the PDF proof we have sent you with your corrections and upload it via the Author Gateway. Alternatively, you may send us your corrections in list format. You may also upload revised graphics via the Author Gateway.

AQ1: Please provide the better quality image for all the figures.