Electric Power Systems Research 212 (2022) 108377

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

# Multi-area parameter error identification for large power systems ☆

Ramtin Khalili, Ali Abur *

*Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA*

ARTICLE INFO

*Keywords:*
Bad data
Multi-area state estimation
Network parameter error
Outlier detection

ABSTRACT

Power grid model parameters may contain errors due to various reasons. Detecting and correcting parameter errors typically requires significant computational effort due to the size and complexity of the parameter database. While the normalized Lagrange multiplier (NLM) method can effectively detect, identify and correct parameter errors, its computational burden could rapidly grow with increasing system size. This paper addresses this issue by proposing a multi-area parameter error identification method. Each area has its own outlier detection tool for detecting the incorrect parameters and measurements within the area. On the other hand, due to the reduced redundancy at area boundaries, parameter errors on branches incident to boundary buses may not be detected. Such errors are subsequently detected by a coordination level estimator completing the system-wide parameter detection procedure. Performance of the developed method is demonstrated using the IEEE 118-bus and 2000-bus Texas synthetic systems.

## 1. Introduction

State estimation (SE) is a critical part of the energy management systems for monitoring the power system and detecting the errors in the measurements or the network model. Following the introduction of power system SE by Schweppe in 1970, the computational issues were recognized in applying SE in large utility systems. Multi-area state estimation (MASE) formulations were thus proposed to improve the scalability and numerical issues. The MASE proved to be promising by reducing CPU times especially when implemented in parallel processors. Given a large interconnected power grid with multiple control areas, MASE formulation facilitates efficient implementation of system-wide SE where individual regional control centers can monitor the local states and a coordination level SE will not only compute the global solution but also process bad data missed by individual area SEs. Thus, state estimators can be implemented to monitor the local states across multiple voltage levels in the energy management systems.

Each individual system operator (ISO) has its own state estimator for monitoring the local network based on the local measurements in the multi-area and distributed SE. The primary challenge, however, is how to obtain the global solution for the whole system. The literature on the topic is quite rich in the past several decades. A two-level SE was proposed in [1,2], where the main motivation was to reduce computational burden. Several investigators contributed to the development of efficient ways to utilize a MASE algorithm as discussed in [3–7]. In the first level (lower level), ISOs make use of the existing SE algorithms to estimate the states based on their local substation measurements. Then, the boundary estimated states and measurements are transferred to a coordination level (aggregation level, second level, upper level) for the final adjustments and determination of the global states. Phase angles are updated based on the estimated phase angle differences for each area in the coordination level.

The benefits of incorporating phasor measurement unit (PMU) measurements into MASE are illustrated in [4]. In [5], PMUs are systematically placed in each area to obtain the phase angle differences between the zones. The MASE concept is further exploited in [6] using non-overlapping areas to carry out observability and bad data analyses. One advantage of using MASE is that sensor failures, bad data or cyber attacks in one region will not deteriorate the SE solution of other areas. In fact, the work reported in [7] is inspired by this observation. A fixed Jacobian matrix is used in [8] as done in the fast second-order load flow method. Sensitivity functions of areas are being exchanged in [9] instead of boundary measurements and states.

Distributed SE methods aim to address the scalability issue without any central coordinator [10–14]. An iterative distributed SE algorithm relying on exchange of states among areas is presented in [10]. A distributed SE based on Huber's robust M-estimator is developed in [11]. The bilinear SE is extended to MASE in [12,13] with minimum data exchange between the adjacent areas. A distributed multi-agent SE method with asynchronous communicating agents is proposed in [14]. A comprehensive review of multi-area and distributed SE methods can be found in [15].

☆ This work made use of ERC shared facilities supported by the ERC Program of the NSF and the DOE under NSF Award Number EEC-1041877.
* Corresponding author.
*E-mail addresses:* rkhalili@ece.neu.edu (R. Khalili), abur@ece.neu.edu (A. Abur).

https://doi.org/10.1016/j.epsr.2022.108377
Received 2 October 2021; Received in revised form 24 March 2022; Accepted 2 July 2022
Available online 22 July 2022

All of the MASE algorithms proposed so far assume that the network parameters are perfectly known. This assumption however is seldom true due to various reasons including human data entry errors, unreported changes in line configurations or conductors, changes due to variations in ambient temperature, etc. Early parameter error identification methods are primarily based on residual sensitivity analysis as described in [16,17]. Another approach is based on augmenting a suspect set of parameters with the states and then estimating both simultaneously [18]. Both approaches have obvious limitations as they cannot differentiate between measurement and parameter errors. Moreover, the scalability of these methods to large power systems is difficult especially for the state augmentation methods.

An alternative approach which is based on the normalized Lagrange multiplier (NLM) test is introduced in [19] for detecting, identifying and correcting parameter errors. The salient point of the NLM method is its compatibility with the largest normalized residuals (LNR) test for bad data detection which enables simultaneous detection and identification of parameter and measurement errors. The NLM test is revisited and further improved in [20,21]. The first step to perform the NLM test is to calculate the Jacobian with respect to the network parameters ($H_p$) which is typically quite large for large scale power grids.

In this paper, the multi-area framework is extended to the simultaneous measurement and parameter error identification problem formulation. From the computational point of view, the bottleneck in implementation and execution of the LNR and NLM methods in large systems is the calculation of the covariance matrices for the residual and Lagrange multipliers. So, it is possible to exploit the computational framework of MASE for parameter error identification as well. This is accomplished by detecting parameter and measurement errors in individual areas by the local state estimators and corresponding NLM/LNR tests. Assuming that each area measurement configuration is well designed with enough redundancy, undetected parameter/measurement errors will mainly be associated with the boundary branches where the measurement redundancy is inadvertently reduced due to the network partitioning. Such errors will be detected at the coordination level, where the shared boundary information from the individual areas will be used for state estimation and error processing. This framework enables accurate monitoring of energy transactions across area boundaries in a deregulated power grid. If a parameter error exists at the boundary of two areas, it cannot be detected by individual areas, but will be identified at the coordination level.

It is shown that the CPU time can be drastically improved by avoiding the full system integrated SE and bad data processing. It is shown earlier in [22] that sparsity methods can be used to significantly reduce the computational burden of the NLM method. The presented multi-area approach of this paper will further improve the performance by reducing the size of the covariance matrices. Also, in cases where a centralized SE cannot be implemented due to the reluctance of individual areas to share measurement and network data, this approach can facilitate implementation of the NLM/LNR based parameter and measurement error detection method for the entire system.

## 2. Multi-Area State Estimation Solution

The MASE is motivated by two objectives: first, increasing the SE computational efficiency by executing the SE on parallel processors and second, the need to obtain a system-wide solution considering a deregulated multi-area power grid. First, a brief review of the MASE formulation will be presented below.

### 2.1. Area specification and individual area state estimation

Consider an 8-bus example system shown in Fig. 1. Tie lines (B4–B5) separate different zones and three types of buses are identified for the given network partitioning:

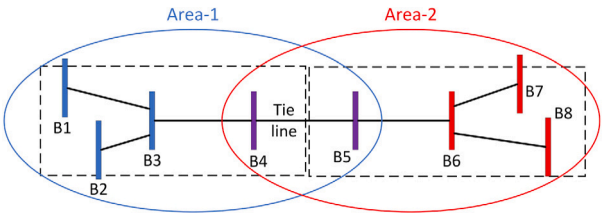


**Fig. 1.** Area configuration and bus assignment.

- *Internal* buses ($\mathcal{B}^{int}$): buses which are within individual areas and not connected to any of the buses from other areas. B1, B2, and B3 are internal buses for area-1.
- *Boundary* buses ($\mathcal{B}^{b}$): buses which are directly connected to other areas by tie lines. B4 and B5 are boundary buses for area-1 and area-2, respectively.
- *External* buses ($\mathcal{B}^{ext,\kappa}$): buses which are within the $\kappa$-tier out of the boundary buses. B5 and B4 are external buses with $\kappa = 1$ for area-1 and area-2, respectively.

Each area includes buses from the internal, boundary, and external ($\kappa$-tier-out) sets. Thus, areas overlap with each other, and boundary buses are included in at least two different areas. Choosing $\kappa \in \mathbb{N}$ for external buses is optional and a bigger value leads to a more accurate solution. However, $\kappa > 1$ means that the ISOs should share more information with the neighbors which increases the size of each area and, more importantly, may not be allowed by the area privacy policies. So, a reasonable choice would be $\kappa = 1$, where only boundary buses of neighboring zones are included in each area.

Considering the SCADA measurements, measurement equations for the area $i = \{1, 2, \dots, l\}$ is given by:

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

where, $\boldsymbol{z}_i$ and $\boldsymbol{e}_i$ are $[m_i \times 1]$ measurement and measurement error vectors for the area $i$, respectively. $\boldsymbol{h}(.)$ denote the nonlinear measurement vector function. $\boldsymbol{x}_i$ is the state vector for the area $i$ which is defined as follows:

$$\boldsymbol{x}_i = \left[[\boldsymbol{x}_i^{int}]^T \quad [\boldsymbol{x}_i^{b}]^T \quad [\boldsymbol{x}_i^{ext,\kappa}]^T\right]^T \tag{2}$$

where, $\boldsymbol{x}_i^{int}$, $\boldsymbol{x}_i^{b}$, and $\boldsymbol{x}_i^{ext,\kappa}$ consist of the phase angles and voltage magnitudes in $\mathcal{B}_i^{int}$, $\mathcal{B}_i^{b}$, and $\mathcal{B}_i^{ext,\kappa}$, respectively. Typically active and reactive power injections and flows and also voltage magnitudes at some buses are measured by SCADA system. So, $\boldsymbol{z}_i$ can be formed as follows:

$$\boldsymbol{z}_i^T = \left[[P_k^i]^T \quad [P_{km}^i]^T \quad [Q_k^i]^T \quad [Q_{km}^i]^T \quad [|V_k^i|]^T\right] \tag{3}$$

where:

$$\begin{aligned} &\text{power flows:} && k, m \in \{\mathcal{B}_i^{int} \cup \mathcal{B}_i^{b} \cup \mathcal{B}_i^{ext,\kappa}\} \\ &\text{power injections:} && k \in \{\mathcal{B}_i^{int} \cup \mathcal{B}_i^{b} \cup \mathcal{B}_i^{ext,\kappa-1}\} \end{aligned}$$

Note that $\kappa - 1$ is used in $\mathcal{B}_i^{ext,\kappa-1}$ for the injections since expression for an injection measurement includes states of one-tier-out from that bus. As there is no more information beyond the external bus(es), their injections will be excluded from the area injection set.

Given the non-linearity of (1), the weighted least squares (WLS) can be formulated as follows for each area:

$$\begin{aligned} &min\ \boldsymbol{J}(\boldsymbol{x}_i) = \boldsymbol{r}_i^T \boldsymbol{R}_i^{-1} \boldsymbol{r}_i \\ &s.t. \quad \boldsymbol{r}_i = \boldsymbol{z}_i - \boldsymbol{h}_i(\boldsymbol{x}_i) \end{aligned} \tag{4}$$

where, $\boldsymbol{R}_i$ is the measurement error covariance matrix which is defined as follows for area $i$:

$$Cov(\boldsymbol{e}_i) = \boldsymbol{R}_i = diag(\sigma_1^2, \sigma_2^2, \dots, \sigma_{m_i}^2) \tag{5}$$

where, $\sigma$ denotes the standard deviation representing the accuracy of the corresponding measure. Then, states for each area are estimated independently by recursively solving the following equation which is known as the normal equation:

$$[\boldsymbol{G}_i(\hat{\boldsymbol{x}}_i)]\Delta\hat{\boldsymbol{x}}_i = \boldsymbol{H}_i^T(\hat{\boldsymbol{x}}_i)\boldsymbol{R}_i^{-1}[\boldsymbol{z}_i - \boldsymbol{h}_i(\hat{\boldsymbol{x}}_i)] \tag{6}$$

where, $\boldsymbol{G}_i$ is the gain matrix for the area $i$:

$$\boldsymbol{G}_i(\hat{\boldsymbol{x}}_i) = [\boldsymbol{H}_i^T(\hat{\boldsymbol{x}}_i)]\boldsymbol{R}_i^{-1}[\boldsymbol{H}_i(\hat{\boldsymbol{x}}_i)] \tag{7}$$

### 2.2. Coordination level state estimation

Individual area SE solutions are obtained by performing the local estimators on the extended areas as explained in the previous subsection. In order to obtain the system-wide solution, each area transfers the measurements, states, and parameters pertaining to the boundary, $\kappa$-tier-out, and $\kappa$-tier-in buses. Then, with the transferred information from the first level, the same WLS SE problem of (4) will be solved:

$$\begin{aligned} &min\ J(\boldsymbol{x}_c) = \boldsymbol{r}_c^T\boldsymbol{R}_c^{-1}\boldsymbol{r}_c \\ &s.t. \quad \boldsymbol{r}_c = \boldsymbol{z}_c - \boldsymbol{h}_c(\boldsymbol{x}_c) \end{aligned} \tag{8}$$

where, subscript $c$ denotes the coordination level. Let us define $\mathcal{B}_i^{int,\kappa}$ as the set of $\kappa$-tier-in buses from boundary buses in the area $i$. Then, the coordination level state vector can be defined as follows:

$$\boldsymbol{x}_c = \left[[\boldsymbol{x}_i^{int,\kappa}]^T \quad [\boldsymbol{x}_i^b]^T \quad [\boldsymbol{x}_i^{ext,\kappa}]^T \quad \boldsymbol{u}^T\right]^T \tag{9}$$

for all $i = \{1, 2, \dots, l\}$. $\boldsymbol{x}_i^{int,\kappa}$, $\boldsymbol{x}_i^b$, and $\boldsymbol{x}_i^{ext,\kappa}$ consist of the phase angles and voltage magnitudes in $\mathcal{B}_i^{int,\kappa}$, $\mathcal{B}_i^b$, and $\mathcal{B}_i^{ext,\kappa}$, respectively. $\boldsymbol{u} = [u_1, u_2, \dots, u_l]^T$ is a $[l\times1]$ vector determining the phase angle adjustment of area $i \in \{2, \dots, l\}$ with respect to the area 1 reference bus, due to the selection of local references by the individual area state estimators.

$$\boldsymbol{z}_c^T = \left[[\boldsymbol{P}_k^i]^T\ [\boldsymbol{P}_{km}^i]^T\ [\boldsymbol{Q}_k^i]^T\ [\boldsymbol{Q}_{km}^i]^T\ [\hat{\theta}_k^i]^T\ [|\hat{V}_k^i|]^T\right] \tag{10}$$

where:

power flows: $k, m \in \{\mathcal{B}_i^{int,\kappa} \cup \mathcal{B}_i^b \cup \mathcal{B}_i^{ext,\kappa}\}$
power injections: $k \in \{\mathcal{B}_i^{int,\kappa-1} \cup \mathcal{B}_i^b \cup \mathcal{B}_i^{ext,\kappa-1}\}$
Pseudo measurements: $k \in \{\mathcal{B}_i^{int,\kappa} \cup \mathcal{B}_i^b \cup \mathcal{B}_i^{ext,\kappa}\}$

$|\hat{V}_k^i|$ and $\hat{\theta}_k^i$ are voltage magnitudes and phase angles which are estimated in the first stage by individual area estimators. So, these estimations can be used in the coordination level as pseudo measurements. In order to assign appropriate weights for these pseudo measurements, state covariance matrix should be calculated for each area as follows:

$$\begin{aligned} \boldsymbol{R}_{\hat{\boldsymbol{x}}_i} &= \mathbb{E}(\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_i^T) = \mathbb{E}(\boldsymbol{G}_i^{-1}\boldsymbol{H}_i^T\boldsymbol{R}_i^{-1}\boldsymbol{e}\boldsymbol{e}^T\boldsymbol{R}_i^{-1}\boldsymbol{H}_i\boldsymbol{G}_i^{-1}) \\ &= \boldsymbol{G}_i^{-1}\boldsymbol{H}_i^T\boldsymbol{R}_i^{-1}\mathbb{E}(\boldsymbol{e}\boldsymbol{e}^T)\boldsymbol{R}_i^{-1}\boldsymbol{H}_i\boldsymbol{G}_i^{-1} \\ &= \boldsymbol{G}_i^{-1} \end{aligned} \tag{11}$$

Using (2), (3), (9), and (10), the structure of the measurement Jacobian $\boldsymbol{H}$ will take the following form:

$$\boldsymbol{H} = \begin{array}{c} \\ \partial\{P_k, P_{km}\} \\ \partial\{Q_k, Q_{km}\} \\ \partial|V_k| \\ \partial\hat{\theta}_k \\ \partial\left|\hat{V}_k\right| \end{array} \begin{array}{c} \begin{array}{cccc} \partial\boldsymbol{x}^{int} & \partial\boldsymbol{x}^{b} & \partial\boldsymbol{x}^{ext} & \partial\boldsymbol{u} \end{array} \\ \left[\begin{array}{cccc} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ & & & \\ & & & \end{array}\right] \end{array} \tag{12}$$

The given matrix structure refers to the coordination level Jacobian where the blue rows correspond to the pseudo measurements, and the column to the vector $\boldsymbol{u}$. Removing the blue rows and columns will yield the Jacobian structure for individual areas.

### 2.3. Bad data detection and correction

Bad data processing follows the largest normalized residual (LNR) method. The residual sensitivity matrix will be obtained first as:

$$\boldsymbol{S} = [\boldsymbol{I}]_{m_i\times m_i} - \boldsymbol{K} \tag{13}$$

where:

$$\boldsymbol{K} = \boldsymbol{H}\boldsymbol{G}^{-1}\boldsymbol{H}^T\boldsymbol{R}^{-1} \tag{14}$$

Then, the residual covariance matrix for the first and second (coordination) levels will be computed as follows:

$$\boldsymbol{\Omega} = cov(\boldsymbol{r}) = \mathbb{E}(\boldsymbol{r}\boldsymbol{r}^T) = \boldsymbol{S}\boldsymbol{R} \tag{15}$$

Obtaining the $m\times m$ matrices $\boldsymbol{K}$, $\boldsymbol{S}$, and $\boldsymbol{\Omega}$ requires significant CPU time for very large systems. Addressing this bottleneck will be discussed in Section 4.

The normalized residuals are obtained by dividing the measurement residuals by the square root of the diagonal entries of the residual covariance matrix as follows:

$$\boldsymbol{r}^N(q) = \frac{\boldsymbol{r}(q)}{\sqrt{\boldsymbol{\Omega}(q,q)}} \tag{16}$$

Based on the LNR test, if the largest value among $\boldsymbol{r}^N$ entries is greater than 3, then the corresponding measure will be asserted as erroneous. The LNR test is implemented in all the individual areas as well as in the coordination level to identify bad data within each area and boundaries. Given the smaller size matrices especially in (14)–(15) significantly simplifies and accelerates the computations. If the $q$th measurement is detected as bad data, it will be corrected as follows:

$$\boldsymbol{z}^{cor}(q) = \boldsymbol{z}(q) - \frac{\boldsymbol{r}(q)}{\boldsymbol{S}(q,q)} \tag{17}$$

## 3. Multi-area parameter error identification

The measurements can be written as a function of the states and parameter errors vector $\boldsymbol{p}_e$ ($\boldsymbol{p}_e = \boldsymbol{p} - \boldsymbol{p}^{true}$) which can be expressed as follows for individual areas $i = \{1, 2, \dots, l\}$ and also for the coordination level [19]:

$$\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{p}_e) + \boldsymbol{e} \tag{18}$$

Initially assuming no parameter errors but incorporating the parameter error vector into the WLS SE formulation as a constraint:

$$\begin{aligned} &min\ J(\boldsymbol{x}) = \boldsymbol{r}^T\boldsymbol{R}^{-1}\boldsymbol{r} \\ &s.t. \quad \boldsymbol{r} = \boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x}) \\ &\qquad \boldsymbol{p}_e = 0 \end{aligned} \tag{19}$$

The equality constraints are usually incorporated in the Lagrangian using appropriate Lagrange multipliers as follows:

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{p}_e, \lambda) = \frac{1}{2}\boldsymbol{r}^T\boldsymbol{R}^{-1}\boldsymbol{r} - \lambda^T\boldsymbol{p}_e \tag{20}$$

So, there are three optimality conditions with respect to $x$, $\lambda$, and $p$ that should be satisfied. The first one $\partial\mathcal{L}/\partial x = 0$ yields the normal Eq. (6) as explained before. $\partial\mathcal{L}/\partial\lambda = 0$ gives the equality constraint $\boldsymbol{p}_{ei} = 0$ itself. Solving the third one, $\partial\mathcal{L}/\partial p = 0$ yields the Lagrange multipliers vector as follows:

$$\lambda = -[\boldsymbol{H}_p]^T\boldsymbol{R}^{-1}\boldsymbol{r} \tag{21}$$

where, $\boldsymbol{H}_p$ is the Jacobian with respect to the network parameters. Supposing the $\Pi$ model for the transmission lines, there are three parameters associated with each line, i.e., resistance, reactance, and

line charging susceptance. So, the structure of $\boldsymbol{H}_p$ can be represented as follows:

$$\boldsymbol{H}_p = \begin{array}{c} \\ \partial\{P_k, P_{km}\} \\ \partial\{Q_k, Q_{km}\} \\ \partial|V_k| \\ \partial\hat{\theta}_k \\ \partial|\hat{V}_k| \end{array} \begin{array}{c} \begin{array}{ccc} \partial\boldsymbol{R} & \partial\boldsymbol{X} & \partial\boldsymbol{B} \end{array} \\ \begin{bmatrix} \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ & & \end{bmatrix} \end{array} \tag{22}$$

where, the blue areas are related to the $\boldsymbol{H}_p$ in the coordination level when the pseudo measurements are also incorporated. Supposing $n_{br}$ as the number of transmission lines, $\boldsymbol{R}$, $\boldsymbol{X}$, and $\boldsymbol{B}$ are $[n_{br} \times 1]$ vectors of all branch resistance, reactance, and line charging susceptances, respectively. The covariance matrix of the Lagrange multipliers is defined as follows:

$$\begin{aligned} \boldsymbol{\Lambda} &= cov(\lambda) = \mathbb{E}(\lambda\lambda^T) \\ &= \boldsymbol{H}_p^T \boldsymbol{R}^{-1} \mathbb{E}(\boldsymbol{r}\boldsymbol{r}^T) \boldsymbol{R}^{-1} \boldsymbol{H}_p \\ &= [\boldsymbol{H}_p]^T \boldsymbol{R}^{-1} \boldsymbol{S} \boldsymbol{H}_p \end{aligned} \tag{23}$$

$\boldsymbol{\Lambda}$ is a $[3n_{br} \times 3n_{br}]$ matrix whose dimension will be extremely large even for a typical utility system. As will be explained later multi-area approach is an efficient method for addressing this computational bottleneck.

Normalizing the Lagrange multipliers with respect to the square root of the diagonal entries of its covariance matrix:

$$\lambda^N(k) = \frac{\lambda(k)}{\sqrt{\boldsymbol{\Lambda}(k,k)}} \tag{24}$$

According to the normalized Lagrange multiplier (NLM) test (similar to the LNR test), if the largest $\lambda^N$ is greater than 3, then the corresponding parameter will be selected as a suspect parameter. This test is implemented in the individual areas in order to secure the local estimators against parameter errors. However, it is possible to miss some erroneous cases due to the lack of measurement redundancy. As such, the NLM test is also implemented in the coordination level to detect undetected parameter errors by the local area estimators. This is feasible at the coordination level due to measurement consolidation and regaining the reduced redundancy at the boundaries.

Similar to the measurement correction, if the $q$th parameter is detected as erroneous by the NLM test, this can be corrected using the corresponding Lagrange multiplier and its covariance, as follows [23]:

$$\boldsymbol{p}^{cor}(q) = \boldsymbol{p}(q) - \frac{\lambda(q)}{\boldsymbol{\Lambda}(q,q)} \tag{25}$$

The comprehensive block diagram of the proposed method is shown in Fig. 2.

## 4. Discussion and simulation results

In order to test and validate the robustness of the proposed multi-area approach, several simulations are performed on the IEEE 118-bus system, and 2000-bus Texas synthetic system [24]. Unless otherwise stated, power injection pairs (active and reactive) at all buses, power flow pairs at *from* end of the branches, and voltage magnitudes at 10% of the buses constitute the measurement configuration for all the test cases. Furthermore, although it is reasonable to assume at least one PMU in each area, no such assumption is made for the simulations. If there is at least one PMU in each area, then the SE results will be improved as the phase angle differences between the areas will be known. Features of test systems, including the number of buses, number of branches, and number of areas are given in Table 1.

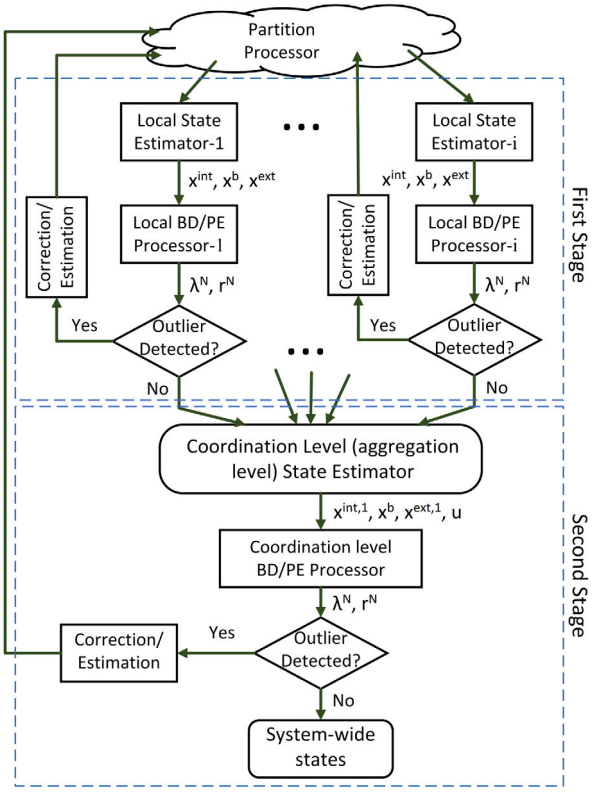


**Fig. 2.** Block diagram of the proposed method.

**Table 1**
Test systems features.

| | IEEE 118-bus | 2000-bus Texas |
|---|---|---|
| No. of Buses | 118 | 2000 |
| No. of Branches | 179 | 2667 |
| No. of Areas | 3 | 8 |

### 4.1. Identifying parameter error in internal branches

In this subsection, parameter errors in the internal branches are studied. Internal branches are those whose terminal buses are in $\mathcal{B}_i^{int}$. The main purpose is to show how the local areas are identifying and correcting the parameter errors which occur within the areas. To this end, the resistance of the transmission line 31 – 32 in area-1 and the line charging susceptance of the line 56 – 59 in area-2 in the IEEE-118 bus system are increased by 30% and 40%, respectively. Results of parameter error identification and correction are shown in Table 2. Note that the choice of these and other branches for the rest of the case studies is totally random and used as examples to illustrate the approach. The NLM approach is a well-established method for identifying the parameter errors (please see [19–22]).

Similar to bad data detection using NR method, even a single parameter error will cause increases in multiple residuals. As a result, observing a large number of parameters with significant $\lambda^N$s is expected. However, based on the NLM method only the largest $\lambda^N$ corresponds to the erroneous parameter in each identification cycle. The three largest values of $\lambda^N$ and the associated parameters are shown in Table 2 for the area-1, 2, and the coordination level in the second column. Results of area-3 are not shown due to negligible $\lambda^N$ values. The largest $\lambda^N$ in area-1 is 55.60 which corresponds to $X_{31-32}$. So, area-1 SE promptly detects $X_{31-32}$ as erroneous based on the largest NLM. $B_{56-59}$ is also detected and corrected locally within area-2. Since the coordination level has not detected an error in the boundaries, only the

**Table 2**
Results of identifying multiple parameter errors in internal branches: IEEE 118-bus.

| | Par | $\lambda^N$ | Detected?<br>Type<br>Parameter | Err. value<br>Est. value<br>True value |
|---|---|---|---|---|
| Area-1 | $X_{31-32}$ | 55.60 | Yes | 0.12805 |
| | $X_{28-29}$ | 36.05 | PE | 0.0984 |
| | $R_{29-31}$ | 35.06 | $X_{31-32}$ | 0.0985 |
| Area-2 | $B_{56-59}$ | 26.26 | Yes | 0.1547 |
| | $R_{56-59}$ | 15.34 | PE | 0.1104 |
| | $B_{54-56}$ | 9.89 | $B_{56-59}$ | 0.1105 |
| Coord. | – | $< 10^{-4}$ | No | – |

**Table 3**
Results of identifying multiple parameter errors in internal branches: 2000-bus Texas.

| | Par | $\lambda^N$ | Detected?<br>Type<br>Parameter | Err. value<br>Est. value<br>True value |
|---|---|---|---|---|
| Area-2 | $B_{112-93}$ | 108.75 | Yes | 1.0144 |
| | $X_{112-93}$ | 36.94 | PE | 0.7801 |
| | $X_{156-135}$ | 36.93 | $B_{112-93}$ | 0.7803 |
| Area-3 | $R_{345-369}$ | 53.19 | Yes | 0.0382 |
| | $R_{231-369}$ | 31.72 | PE | 0.0293 |
| | $B_{345-369}$ | 24.56 | $R_{345-369}$ | 0.0294 |
| Area-6 | $X_{1080-1238}$ | 89.92 | Yes | 0.1028 |
| | $X_{1080-1682}$ | 73.64 | PE | 0.0791 |
| | $X_{1238-1126}$ | 73.28 | $X_{1080-1238}$ | 0.0791 |
| Coord. | – | $< 10^{-4}$ | No | – |

largest $\lambda^N$ is shown in the last row of Table 2 which is insignificant. The status of parameter error identification is shown in the third column. Finally, the incorrect parameter is estimated using (25). The erroneous (Err.), estimated (Est.), and true values of the parameters are listed in the fourth column. While multiple parameter errors are identified and corrected by individual areas, this process would have been performed sequentially if a central estimator were used.

The same scenario is repeated for the 2000-bus system where the resistance of the transmission line $345 - 369$ in area-3, the reactance of the line $1080 - 1238$ in area-6, and the line charging susceptance of the line $112 - 93$ in area-2 are all increased by 30%, simultaneously. The results of parameter error identification and correction are shown is Table 3. The results are similar to the ones in Table 2 and all the parameter errors are detected and corrected within areas 2, 3, and 6.

### 4.2. Computation time analysis

In order to compare the performance of the multi-area and integrated implementation, a central SE, bad data identification, and parameter error identification is also implemented. All the matrices in both multi-area and central applications are implemented using sparse matrix methods in MATLAB. Power flows from both ends of the branches are considered in this subsection in order to increase the number of measurements and hence the size of system matrices. So, the total number of measurements and states for the 2000-bus system are 14,868 and 4,000, respectively. Then, the central SE along with the bad data and parameter error detection tools are executed, and the CPU time profile is obtained. Fig. 3 is a pie chart showing the percentage of the CPU time spent in calculating the system matrices for integrated SE. As evident from Fig. 3, roughly 74% of the total CPU time is spent calculating $K$. Furthermore, more than 95% is related to computation of $K$, $\Lambda$, $S$, and $\Omega$ which are utilized in bad data and parameter error identification routines.

The performance of the central and multi-area formulation for the 2000-bus system is also illustrated in Fig. 4. The CPU time for all areas, as well as the coordination (second) level and the central estimator, are shown in Fig. 4. The results reflect the CPU time spent for the SE,

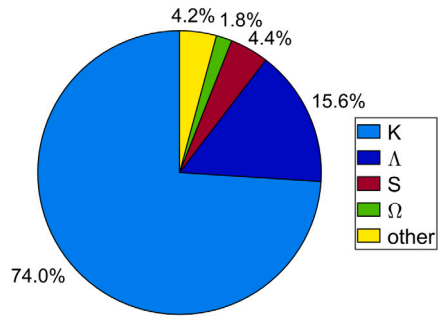


**Fig. 3.** Percent of CPU time for calculating important matrices.

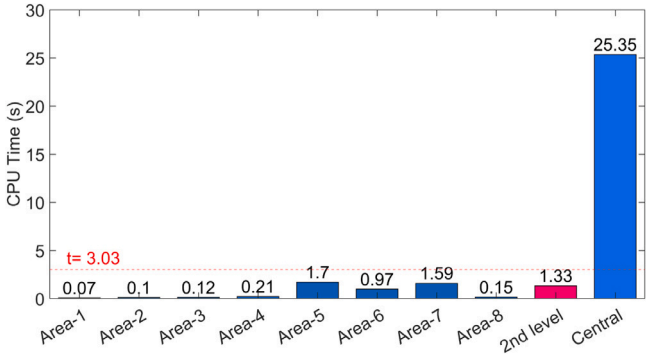


**Fig. 4.** CPU time for the individual areas, coordination, and central estimators, bad data, and parameter error identification.

bad data detection, and parameter error identification tasks in seconds. If the individual estimators are assigned to different processors, then the total CPU time for the first level will be limited by the slowest area. The dashed red line in Fig. 4 shows the total CPU time for the first and coordination level as they are running sequentially. These simulation results validate the computational benefits of using the proposed multi-area approach over the integrated solution.

### 4.3. Parameter error in boundary branches

In this subsection, parameter error is identified in a boundary branch using the multi-area method. Boundary buses and areas for the IEEE 118-bus system are shown in Fig. 5. The reactance $X_{38-65}$ is corrupted by increasing the true value by 25%. The line $38 - 65$ is on the boundary between area-1 and 2. With the existing measurement configuration, the error is still detectable. So, some of the flow measurements are deleted in order to simulate the case where there is not enough measurement redundancy in the boundary. Results of the parameter error detection for adjacent areas, i.e., area-1 and 2 and also the coordination level, are tabulated in Table 4.

As can be seen from Table 4, the incorrect parameter could not be detected or identified by local estimators. Even though this error is not detected, its impact on the estimated states, especially in the adjacent buses, is observed. This is the main reason why the coordination level SE is needed. Since the power injections at both terminals are used in the coordination level, the redundancy is improved, and the NLM method manages to detect and identify the erroneous parameter.

### 4.4. Simultaneous parameter and measurement errors

One of the distinguishing features of the NLM/LNR method is its ability to detect and identify measurement and parameter errors simultaneously. In this subsection, the performance of the method is

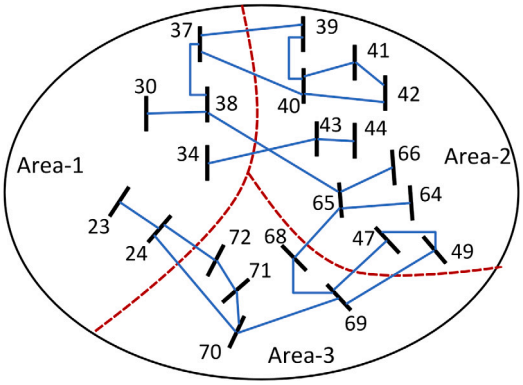


**Fig. 5.** IEEE 118-bus zones and boundary construction.

**Table 4**
Results of identifying multiple parameter errors in boundary branches: IEEE 118-bus.

| | Par | $\lambda^N$ | Detected? type parameter | Err. value Est. value True value |
|---|---|---|---|---|
| Area-1 | – | $< 10^{-3}$ | No | – |
| Area-2 | $X_{65-68}$ | 22.42 | Yes | – |
| | $X_{38-65}$ | 22.41 | PE | – |
| | $R_{65-68}$ | 19.68 | Not Identified | – |
| Coord. | $X_{38-65}$ | 25.87 | Yes | 0.1232 |
| | $X_{65-68}$ | 21.14 | PE | 0.0987 |
| | $R_{38-37}$ | 18.52 | $B_{56-59}$ | 0.0986 |

**Table 5**
Results of simultaneous measurement and parameter errors: 2000-bus, Area-1.

| Par/ Meas | Cycle 1 | Cycle 2 | Cycle 3 | Cycle 4 |
|---|---|---|---|---|
| | $P_{17}^{inj}$ | $X_{17-29}$ | $P_{17}^{inj}$ | $X_{17-29}$ |
| $r_{max}^N$ | 213.62 | 64.75 | 9.41 | 2.45 |
| $\lambda_{max}^N$ | 98.13 | 35.69 | 5.49 | 4.48 |
| Est. Value | −0.1995 | 0.1401 | **−0.1828** | **0.1419** |
| True Value | −0.1834 | 0.1419 | −0.1834 | 0.1419 |

tested when measurement and parameter errors coexist in the 2000-bus system. To this end, the reactance $X_{17-29}$ in area-1 is corrupted by a 20% increase. Furthermore, a sign error i intentionally introduced in the active power injection at bus 17 ($P_{17}^{inj}$). The results of measurement and parameter error detection for this scenario are shown in Table 5. Since these errors do not have a significant impact on the neighboring buses, Table 5 reflects the results of Area-1.

As indicated in Table 5 the NLM and LNR methods are cyclic, which means one error will be detected and corrected in each cycle and the SE execution will be repeated. For each cycle the maximum of $\boldsymbol{r}_{max}^N$ and $\lambda_{max}^N$ are reported. Then, the suspect measurement or parameter is determined based on these values. Once the error is detected, it will be corrected before going to the next cycle. So, the estimated and true values of the erroneous parameter/measurement are presented in the last two rows. Since these two errors are conforming and the magnitude of the introduced errors is substantial, it takes four iterations for both errors to be cleared. After the fourth iteration, the $\boldsymbol{r}^N$ and $\lambda^N$ values drop off to insignificant values, below the detection threshold. The final estimated values for $X_{17-29}$ and $P_{17}^{inj}$ are shown in blue in Table 5.

## 5. Conclusion

This paper proposes a multi-area parameter error identification which improves the efficiency especially in large power systems. This is achieved by implementing bad data and parameter error detection in the local areas if enough measurement redundancy is present. Then, the boundary information between all areas is carried to the next level. In this level, measurement redundancy will be enhanced and the errors that are missed earlier, will be identified. It is shown that this has a great impact on the efficiency of the overall bad data and parameter error identification. Furthermore the scalability of the method is investigated using a large synthetic system. The CPU time analysis reveals that the proposed multi-area approach provides significant speed-up compared to the integrated solution.

## CRediT authorship contribution statement

**Ramtin Khalili:** Conceptualization, Methodology, Software. **Ali Abur:** Investigation, Methodology, Supervision.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

[1] T. Van Cutsem, J. Horward, M. Ribbens-Pavella, Y. El-Fattah, Hierarchical state estimation, Int. J. Electr. Power Energy Syst. 2 (2) (1980) 70–80.
[2] T. Van Cutsem, J.-L. Horward, M. Ribbens-Pavella, A two-level static state estimator for electric power systems, IEEE Trans. Power Appl. Syst. PAS-100 (8) (1981) 3722–3732.
[3] R. Ebrahimian, R. Baldick, State estimation distributed processing [for power systems], IEEE Trans. Power Syst. 15 (4) (2000) 1240–1246.
[4] L. Zhao, A. Abur, Multi area state estimation using synchronized phasor measurements, IEEE Trans. Power Syst. 20 (2) (2005) 611–617.
[5] W. Jiang, V. Vittal, G.T. Heydt, A distributed state estimator utilizing synchronized phasor measurements, IEEE Trans. Power Syst. 22 (2) (2007) 563–571.
[6] G.N. Korres, A distributed multiarea state estimation, IEEE Trans. Power Syst. 26 (1) (2011) 73–84.
[7] P. Ren, A. Abur, Avoiding divergence in multi-area state estimation, IEEE Trans. Power Syst. 34 (4) (2019) 3178–3187.
[8] S. Iwamoto, M. Kusano, V. Quintana, Hierarchical state estimation using a fast rectangular-coordinate method, IEEE Trans. Power Syst. 4 (3) (1989) 870–880.
[9] Y. Guo, L. Tong, W. Wu, H. Sun, B. Zhang, Hierarchical multi-area state estimation via sensitivity function exchanges, IEEE Trans. Power Syst. 32 (1) (2016) 442–453.
[10] A.J. Conejo, S. de la Torre, M. Canas, An optimization approach to multiarea state estimation, IEEE Trans. Power Syst. 22 (1) (2007) 213–221.
[11] V. Kekatos, G.B. Giannakis, Distributed robust power system state estimation, IEEE Trans. Power Syst. 28 (2) (2013) 1617–1626.
[12] W. Zheng, W. Wu, A. Gomez-Exposito, B. Zhang, Y. Guo, Distributed robust bilinear state estimation for power systems with nonlinear measurements, IEEE Trans. Power Syst. 32 (1) (2017) 499–509.
[13] W. Zheng, W. Wu, An adaptive distributed quasi-Newton method for power system state estimation, IEEE Trans. Smart Grid 10 (5) (2019) 5114–5124.
[14] K. Saxena, A.R. Abhyankar, Agent-based distributed computing for power system state estimation, IEEE Trans. Smart Grid 11 (6) (2020) 5193–5202.
[15] A. Gómez-Expósito, A. de la Villa Jaén, C. Gómez-Quiles, P. Rousseaux, T. Van Cutsem, A taxonomy of multi-area state estimation methods, Electr. Power Syst. Res. 81 (4) (2011) 1060–1069.
[16] T. Van Cutsem, V. Quintana, Network parameter estimation using online data with application to transformer tap position estimation, in: IEE Proc. C (Gener., Transmiss. and Distrib.), Vol. 135, no. 1, IET, 1988, pp. 31–40.
[17] W.-H. Liu, F.F. Wu, S.-M. Lun, Estimation of parameter errors from measurement residuals in state estimation (power systems), IEEE Trans. Power Syst. 7 (1) (1992) 81–89.
[18] W.-H. Liu, S.-L. Lim, Parameter error identification and estimation in power system state estimation, IEEE Trans. Power Syst. 10 (1) (1995) 200–209.
[19] J. Zhu, A. Abur, Identification of network parameter errors, IEEE Trans. Power Syst. 21 (2) (2006) 586–592.
[20] R. Khalili, A. Abur, Transmission line parameter error identification and estimation in three-phase networks, IEEE Trans. Power Syst. (2021) http://dx.doi.org/10.1109/TPWRS.2021.3118007, early access.

[21] R. Khalili, A. Abur, Detection of errors in three-phase line models using synchronized phasor measurements, in: Porc. IEEE PES ISGT-Europe Conference, 2020, pp. 1131–1135.

[22] Y. Lin, A. Abur, Highly efficient implementation for parameter error identification method exploiting sparsity, IEEE Trans. Power Syst. 32 (1) (2017) 734–742.

[23] Y. Lin, A. Abur, Fast correction of network parameter errors, IEEE Trans. Power Syst. 33 (1) (2018) 1095–1096.

[24] A.B. Birchfield, T. Xu, K.M. Gegner, K.S. Shetye, T.J. Overbye, Grid structural characteristics as validation criteria for synthetic networks, IEEE Trans. Power Syst. 32 (4) (2017) 3258–3265.