Allocating PMUs to Facilitate Processing of Bad Data in SCADA-based State Estimation

Julio Cesar Stacchini de Souza, Milton Brown Do Coutto Filho, and Marcio Andre Ribeiro Guimaraens

Abstract—**Advanced network analysis and control tools in energy management systems depend on a reliable real-time database provided by power system state estimation (SE). The processing of gross errors—bad data (BD)—is considered an essential SE step in such a context. It is widely recognized that the adverse conditions with weak network observability hinder the identification and correction of multiple BD, a process notorious for its time-consuming nature. This paper proposes a method for facilitating the processing of BD in supervisory control and data acquisition (SCADA) measurements using phasor measurement units (PMUs). The main objective is to strategically allocate a limited number of PMUs to allow the phasor-aided state estimation (PHASE), which offers the advantage of effectively handling single and multiple BD. Tests with the IEEE 14-, 30-, and 118-bus systems are conducted to demonstrate the practical application of handling BD, including adverse conditions such as network spots with low redundant measurements.**

Index Terms—**Bad data, gross error, state estimation, energy management system, supervisory control and data acquisition (SCADA), phasor measurement unit (PMU).**

I. Introduction

In energy management systems (EMSs), power system state estimation (SE) provides essential real-time data for power grid analysis, control, and operation [1]-[3]. In a balanced load and generation operating regime, the system state is characterized by complex nodal voltages, represented by their magnitudes/phase angles, for a specific network configuration.

Redundancy of measurements is instrumental in SE to allow adequate state filtering and to safeguard against vulnerabilities that may compromise the effectiveness of estimation residual analysis [4] in handling gross errors—bad data (BD). However, searching for BD can be time-consuming, particularly when several of them occur [4], [5], which requires repeatedly applying the largest normalized residual (LNR) criterion for the elimination of suspect measurements, one-by-one or in groups [1]. Moreover, BD may arise when SE is executed under adverse conditions, i.e., when the network is close to unobservability. These conditions are characterized by measurement criticalities, which may manifest as a critical measurement or a critical group of measurements.

Criticality conveys the idea of decisive importance for a given purpose, supported by a careful judgment on the severity of consequences for SE, i.e., observability and capability of processing BD. SE input data are deemed critical if their unavailability, occurring by accident, through programmed maintenance, or by chance, renders the network unobservable. Gross errors impacting individual critical measurement or critical groups of measurements—tuples of cardinality k, abbreviated as Cks—are challenging (even impossible) [6] to handle by commonly available BD analysis techniques [5].

Phasor measurement units (PMUs) have been employed to monitor power networks, potentially benefiting the SE process [7]. Such benefits can be enhanced if PMUs are strategically located. Most studies regarding PMU allocation are devoted to ensuring the minimum network observability [8]-[15], whereas others aim to enhance the accuracy of hybrid estimators [16], [17], which are designed to make a concerted effort to process the supervisory control and data acquisition (SCADA) measurements (conventionally, branch active/reactive power flows, bus power injections, and voltage/current magnitudes) and synchrophasors (synchronized complex bus voltages and branch currents) [18]-[20]. Several methods for PMU allocation to enhance the processing of BD are reported in [21]-[23], aiming to allocate PMU measurements so that no critical measurements are present when using a hybrid estimator. However, most SE applications found in EMS can process only SCADA measurements. It is believed that such a situation will persist in the coming years, as challenges and implementation issues in using hybrid SE remain [24]-[26]. Nevertheless, the number of PMUs deployed in power grids is gradually increasing and demonstrating how synchrophasors can improve the reliability of SE results becomes opportune [27].

Since synchrophasors contribute to increasing data redundancy, their incorporation into existing SCADA-based estimators is desirable. However, it is necessary to conceive esti-

Manuscript received: April 2, 2025; revised: June 16, 2025; accepted: September 28, 2025. Date of CrossCheck: September 28, 2025. Date of online publication: November 25, 2025.

This work was supported by Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES), National Council for Scientific and Technological Development (CNPq), State of Rio de Janeiro Research Support Foundation (FAPERJ), and National Institute of Science and Technology in Electric Energy (INERGE).

J. C. S. de Souza (corresponding author) and M. A. R. Guimaraens are with the Electrical Engineering Department, Fluminense Federal University, Rio de Janeiro, Brazil (e-mail: julio@ic.uff.br; mguimaraens@id.uff.br).

M. B. D. C. Filho is with the Institute of Computing, Fluminense Federal University, Rio de Janeiro, Brazil (e-mail: mbrown@ic.uff.br).

DOI: 10.35833/MPCE.2025.000295

mation strategies that leverage both SCADA and PMU technologies. With the efficient processing of BD in mind, the estimation process, which takes the set of SCADA measurements and synchrophasors individually and processes them separately using two estimators, has been adopted in [28]. Synchrophasors can be submitted to an independent estimation process to produce the filtered state/measurements. Not only synchrophasors, but also SCADA measurements, are estimated from that filtered state. In such a case, the SCADA measurements and their estimates provided by the synchrophasor estimator (Sync estimator) are uncorrelated, implying that the differences between these values (residuals) do not experience the well-known effect of BD smearing [1], [2].

The phasor-aided state estimation (PHASE) is an advanced approach that leverages real-time synchrophasors [28]. These measurements include bus voltages (magnitudes and phase angles) and branch currents (usually in rectangular coordinates). They are highly accurate and synchronized via global positioning system (GPS), allowing for a more detailed and up-to-date view of the power grid. PHASE enables the construction of a standalone algorithm, independent of an existing one that processes only SCADA measurements. Synchrophasors are utilized as additional independent measurements, processed separately to enhance the data debugging in SCADA-based estimators, while concurrently identifying and eliminating BD when they occur in individual critical measurements or groups of measurements (i. e., Cks).

This paper demonstrates how to utilize PMUs more effectively to handle corrupted SCADA measurements, which are processed by conventional SE algorithms. For this purpose, this paper proposes allocating a limited number of PMUs, enabling the PHASE approach to identify multiple BD at once (in one iteration), which consumes less time than the conventional BD treatment (an iterative process requiring many residual analysis cycles). The construction of metering plans for SE is a complex and combinatorial optimization problem, often solved by a family of metaheuristic algorithms. This paper proposes a heuristic method for optimizing the PMU allocation to enhance the processing of BD through residual analysis of Sync estimator. The proposed method is simple and easy to implement, and does not require complex objective functions. Numerical simulations conducted on the IEEE 14-, 30-, and 118-bus test systems demonstrate the efficacy of the strategic PMU allocation in maximizing the benefits of BD processing schemes. The optimized allocation is obtained considering the availability of a limited number of PMUs for SE.

The main contributions of the paper are summarized as follows.

1) While traditional PMU allocation strategies are centered essentially on ensuring grid observability, the proposed method is directed to facilitate the processing of BD in power grids that are already observable by SCADA-based estimators.

2) Improvements in SE often emphasize the importance of utilizing all available data for processing. In this context, synchrophasors are recognized as a significant enhancement to SCADA measurements when addressing challenges in processing BD. The PHASE supports this assertion. Unlike the study in [28], which focuses on establishing an estimation process in areas observable by PMUs, this paper aims to allocate PMUs throughout the grid. The objectives are twofold: ① to maximize the scenarios in which PHASE can assist in overcoming an inefficient or unsuccessful processing of BD, as determined by the LNR criterion, and ② to identify BD involving measurement criticalities.

3) In general, only certain buses can have their state variables estimated exclusively through the processing of synchrophasors. This paper introduces and explores the concept of a PMU-observable bus within the proposed method. The measurements provided by the SCADA system, which can be estimated using only the state variables of PMU-observable buses, are defined as PMU-observable SCADA measurements.

4) The practical advantages of processing conventional spurious measurements with optimally allocated PMUs are demonstrated through simulations conducted on typical power grids, using three different strategies. By exploring the local nature of the estimation process with the concept of PMU-observable bus and PMU-observable SCADA measurement, the size of the power grids is not an issue.

5) The proposed method optimally allocates PMUs, allowing PHASE to be applied alongside state estimators that can process only SCADA measurements.

II. SE and Processing of BD

In this section, the fundamental aspects of the processing of BD and challenges that SE may face, derived from weak observability or the presence of criticalities, are addressed.

A. Fundamental Aspects

SE periodically processes a set of redundant measurements to determine the most likely system state. A fundamental feature of the SE process is its capability to detect and identify BD and suppress them if possible. For a given network topology, the relation between the collected measurements and the system state is given by:

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

where $\boldsymbol{x}$ and $\boldsymbol{z}$ are the n-dimensional state vector and m-dimensional measurement vector, respectively; $\boldsymbol{h}(\cdot)$ represents the measurement equations for the current network configuration; and $\boldsymbol{v}$ is a Gaussian noise vector with zero mean and diagonal covariance matrix $\boldsymbol{R}$. The state vector components are the bus voltage phase angles and magnitudes, while branch active/reactive power flows, bus power injections, and voltage magnitudes commonly constitute the measurement vector. Then, SE can be expressed as a weighted least-squares (WLS) problem with the following objective function:

$$\min J(\boldsymbol{x})=(\boldsymbol{z}-\boldsymbol{h}(\boldsymbol{x}))^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{z}-\boldsymbol{h}(\boldsymbol{x})) \tag{2}$$

The state estimate that minimizes $J(\boldsymbol{x})$ can be obtained using the WLS method by satisfying the optimality condition in (3), which leads to an iterative process [1], [2], where the state variables are updated at each iteration according to (4).

$$\frac{\partial J(\boldsymbol{x})}{\partial \boldsymbol{x}}=-\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{z}-\boldsymbol{h}(\boldsymbol{x}))=\boldsymbol{0} \tag{3}$$

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\boldsymbol{G}^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{z}-\boldsymbol{h}(\boldsymbol{x}_k)) \tag{4}$$

where $\boldsymbol{H}$ is the Jacobian matrix; $\boldsymbol{G}=\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H}$ is the gain matrix; and the subscript k is the iteration counter.

B. Challenges

In the residual analysis, the differences $\boldsymbol{r}(i)$ between measurements $\boldsymbol{z}$ and corresponding estimated values $\hat{\boldsymbol{z}}=\boldsymbol{h}(\hat{\boldsymbol{x}})$ are usually normalized and submitted to the following validation:

$$\boldsymbol{r}_{\mathrm{N}}(i)=|\boldsymbol{r}(i)|/\boldsymbol{\sigma}_{\boldsymbol{E}}(i)\le\varepsilon \tag{5}$$

$$\boldsymbol{E}=\boldsymbol{R}-\boldsymbol{H}\boldsymbol{G}^{-1}\boldsymbol{H}^{\mathrm{T}} \tag{6}$$

where i is the column/row index in a vector/matrix; $\boldsymbol{\sigma}_{\boldsymbol{E}}(i)=\sqrt{\boldsymbol{E}(i,i)}$; ε is the threshold; and the subscript N represents the normalized value. Measurements with $\boldsymbol{r}_{\mathrm{N}}$ violating the adopted threshold are declared suspects of having gross errors.

Given a single gross error, the LNR criterion correctly indicates the bad measurement, provided that sufficient local redundancy exists. However, this may not be the case if there are multiple BD, necessitating a time-consuming residual analysis to eliminate them through several identification-estimation cycles [2]. As it is impossible to know beforehand how many spurious measurements exist, a combinatorial trial-and-error procedure is employed to scan the list of suspect measurements, i.e., those with $\boldsymbol{r}_{\mathrm{N}}$ exceeding the threshold in (5). It is also noteworthy that when processing BD, the valid measurements may inadvertently be discarded, or critical redundancy conditions may be reached, leading one to believe that all BD have been eliminated, whereas some may have become undetectable.

The ability of SE to monitor the system state hinges on the observation capability of the measuring system. Critical data influence this capability—they represent an impending unobservability condition, rated on a graduated scale of severity established by the cardinality of Cks. As mentioned before, a Ck refers to a set of k elements that, when simultaneously unavailable, render the system unobservable. The presence of Cks undermines the SE ability to detect and identify the corrupted data. Gross errors cannot be detected in individual critical measurement (C1) nor identified in critical measurement pairs (C2s) by employing conventional residual analysis. The same holds for BD occurring in any $k-1$ measurement of Cks [5], [29], [30].

Positioning meters in strategic locations throughout the network is crucial to minimizing the exposure of SE to potential data unavailability. The PHASE, which is outlined below, can help alleviate the computational challenges involved in identifying multiple BD, including those related to Cks.

III. Processing of BD with PHASE

PHASE employs a WLS estimator, known as Sync estimator, which processes only synchrophasors ($\boldsymbol{z}_{sync}$). It provides the estimated SCADA measurements $\hat{\boldsymbol{z}}_{conv}^{(sync)}$ to enhance the residual analysis of estimators that process only SCADA measurements ($\boldsymbol{z}_{conv}$), referred to as conventional SCADA-based estimators (Conv estimators). PHASE allows a cross-validation test using the normalized residuals of Sync and Conv estimators [28]-[31].

The estimated state obtained with Sync estimator $\hat{\boldsymbol{x}}^{(sync)}$ can be used to estimate SCADA measurements via power flow equations, i.e., $\hat{\boldsymbol{z}}_{conv}^{(sync)}=\boldsymbol{h}_{conv}(\hat{\boldsymbol{x}}^{(sync)})$, enabling the computation of residuals of Sync estimator as:

$$\boldsymbol{r}_{conv}^{(sync)}(i)=\boldsymbol{z}_{conv}(i)-\hat{\boldsymbol{z}}_{conv}^{(sync)}(i) \tag{7}$$

where $\boldsymbol{r}_{conv}^{(sync)}(i)$ is the i^{th} element of the residual vector of SCADA measurements.

Note that $\hat{\boldsymbol{z}}_{conv}^{(sync)}(i)$ is independently computed by the Sync estimator, which processes only synchrophasors, thereby being uncorrelated with $\boldsymbol{z}_{conv}(i)$. Consequently, the $\boldsymbol{r}_{conv}^{(sync)}$ components are immune to the BD smearing effect.

The covariance matrix of $\boldsymbol{r}_{conv}^{(sync)}$, denoted as $\boldsymbol{V}_{conv}^{(sync)}$, is obtained using (7) and the expectation operator $E\{\cdot\}$, expressed as (8a). Define the covariance vector of the SCADA measurement errors as $\boldsymbol{R}_{conv}=E\{\boldsymbol{z}_{conv}\boldsymbol{z}_{conv}^{\mathrm{T}}\}$ and the covariance matrix of $\hat{\boldsymbol{z}}_{conv}^{(sync)}$ as $\boldsymbol{M}_{conv}^{(sync)}=E\{\hat{\boldsymbol{z}}_{conv}^{(sync)}(\hat{\boldsymbol{z}}_{conv}^{(sync)})^{\mathrm{T}}\}$. As $\boldsymbol{z}_{conv}$ and $\hat{\boldsymbol{z}}_{conv}^{(sync)}$ are uncorrelated, then $E\{\boldsymbol{z}_{conv}(\hat{\boldsymbol{z}}_{conv}^{(sync)})^{\mathrm{T}}\}=\boldsymbol{0}$. So, $\boldsymbol{V}_{conv}^{(sync)}$ is evaluated as (8b).

$$\begin{aligned}\boldsymbol{V}_{conv}^{(sync)}=E\{\boldsymbol{r}_{conv}^{(sync)}(\boldsymbol{r}_{conv}^{(sync)})^{\mathrm{T}}\}=&E\{\boldsymbol{z}_{conv}\boldsymbol{z}_{conv}^{\mathrm{T}}\}-E\{\boldsymbol{z}_{conv}(\hat{\boldsymbol{z}}_{conv}^{(sync)})^{\mathrm{T}}\}-\\&E\{\hat{\boldsymbol{z}}_{conv}^{(sync)}\boldsymbol{z}_{conv}^{\mathrm{T}}\}+E\{\hat{\boldsymbol{z}}_{conv}^{(sync)}(\hat{\boldsymbol{z}}_{conv}^{(sync)})^{\mathrm{T}}\}\end{aligned} \tag{8a}$$

$$\boldsymbol{V}_{conv}^{(sync)}=\boldsymbol{R}_{conv}+\boldsymbol{M}_{conv}^{(sync)} \tag{8b}$$

$$\boldsymbol{M}_{conv}^{(sync)}=\boldsymbol{H}_{conv}^{(sync)}\boldsymbol{S}^{(sync)}(\boldsymbol{H}_{conv}^{(sync)})^{\mathrm{T}} \tag{9}$$

where $\boldsymbol{H}_{conv}^{(sync)}=\partial\boldsymbol{h}_{conv}/\partial\boldsymbol{x}$ is the Jacobian matrix of the SCADA measurements, evaluated at $\boldsymbol{x}=\hat{\boldsymbol{x}}^{(sync)}$; and $\boldsymbol{S}^{(sync)}=\{(\boldsymbol{H}_{sync}^{(sync)})^{\mathrm{T}}\boldsymbol{R}_{sync}^{-1}\boldsymbol{H}_{sync}^{(sync)}\}^{-1}$ is the error covariance matrix of $\hat{\boldsymbol{x}}^{(sync)}$, $\boldsymbol{H}_{sync}^{(sync)}=\partial\boldsymbol{h}_{sync}/\partial\boldsymbol{x}$ is the Jacobian matrix of the synchrophasors, computed at $\boldsymbol{x}=\hat{\boldsymbol{x}}^{(sync)}$, and $\boldsymbol{R}_{sync}$ is the covariance vector of the synchrophasor error. The entry values in $\boldsymbol{R}_{conv}$ and $\boldsymbol{R}_{sync}$ are obtained from the variances of the SCADA and PMU measurements, such as in [7].

The elements of $\boldsymbol{r}_{conv}^{(sync)}$ can be normalized and subjected to a validation test:

$$\boldsymbol{r}_{\mathrm{N}conv}^{(sync)}(i)=\frac{|\boldsymbol{r}_{conv}^{(sync)}(i)|}{\boldsymbol{\sigma}_{\boldsymbol{V}_{conv}^{(sync)}}(i)}\le\varepsilon \tag{10}$$

where $\boldsymbol{\sigma}_{\boldsymbol{V}_{conv}^{(sync)}}(i)=\sqrt{\boldsymbol{V}_{conv}^{(sync)}(i,i)}$.

The Conv estimator processes exclusively SCADA measurements through the same type of calculations in (7)-(10). Therefore, using the notation that represents SCADA measurements ($\boldsymbol{z}_{conv},\boldsymbol{R}_{conv}$), the equations of the Conv estimator used in the residual analysis are given next, where the corresponding variables are defined in the same way as those in (7) - (10), and the only difference is that the superscript "($sync$)" is replaced with "($conv$)" to indicate the Conv estimator.

The vector of estimated SCADA measurements and its error covariance matrix are given by:

$$\hat{\boldsymbol{z}}_{conv}^{(conv)}=\boldsymbol{h}_{conv}(\hat{\boldsymbol{x}}^{(conv)}) \tag{11}$$

$$\boldsymbol{M}_{conv}^{(conv)}=\boldsymbol{H}_{conv}^{(conv)}\boldsymbol{S}^{(conv)}(\boldsymbol{H}_{conv}^{(conv)})^{\mathrm{T}} \tag{12}$$

The estimated residual of the i^{th} SCADA measurement and its covariance matrix are expressed as:

$$\boldsymbol{r}_{conv}^{(conv)}(i)=\boldsymbol{z}_{conv}(i)-\hat{\boldsymbol{z}}_{conv}^{(conv)}(i) \tag{13}$$

$$\boldsymbol{V}_{conv}^{(conv)}=\boldsymbol{R}_{conv}+\boldsymbol{M}_{conv}^{(conv)} \tag{14}$$

$\boldsymbol{r}_{conv}^{(conv)}(i)$ can be normalized as:

$$\boldsymbol{r}_{\mathrm{N}conv}^{(conv)}(i)=\frac{\left|\boldsymbol{r}_{conv}^{(conv)}(i)\right|}{\boldsymbol{\sigma}_{V_{conv}^{(conv)}}(i)}\le\varepsilon \tag{15}$$

where $\boldsymbol{\sigma}_{V_{conv}^{(conv)}}(i)=\sqrt{\boldsymbol{V}_{conv}^{(conv)}(i,i)}$.

With $\hat{\boldsymbol{x}}^{(conv)}$ and $\boldsymbol{S}^{(conv)}$, one can obtain $\hat{\boldsymbol{z}}_{sync}^{(conv)}$ and $\boldsymbol{M}_{sync}^{(conv)}$ and perform the normalized residual test, i.e., $\boldsymbol{r}_{\mathrm{N}sync}^{(conv)}$-test, for the i^{th} synchrophasor:

$$\hat{\boldsymbol{z}}_{sync}^{(conv)}=\boldsymbol{h}_{sync}(\hat{\boldsymbol{x}}^{(conv)}) \tag{16}$$

$$\boldsymbol{M}_{sync}^{(conv)}=(\boldsymbol{H}_{sync}^{(conv)})\boldsymbol{S}^{(conv)}(\boldsymbol{H}_{sync}^{(conv)})^{\mathrm{T}} \tag{17}$$

$$\boldsymbol{r}_{sync}^{(conv)}(i)=\boldsymbol{z}_{sync}(i)-\hat{\boldsymbol{z}}_{sync}^{(conv)}(i) \tag{18}$$

$$\boldsymbol{V}_{sync}^{(conv)}=\boldsymbol{R}_{sync}+\boldsymbol{M}_{sync}^{(conv)} \tag{19}$$

$$\boldsymbol{r}_{\mathrm{N}sync}^{(conv)}(i)=\frac{|\boldsymbol{r}_{sync}^{(conv)}(i)|}{\boldsymbol{\sigma}_{V_{sync}^{(conv)}}(i)}\le\varepsilon \tag{20}$$

where $\boldsymbol{\sigma}_{V_{sync}^{(conv)}}(i)=\sqrt{\boldsymbol{V}_{sync}^{(conv)}(i,i)}$; and the subscript *sync* represents the variables corresponding to the synchrophasor.

If $\boldsymbol{r}_{\mathrm{N}conv}^{(sync)}(i)$ in (10) exceeds the threshold but $\boldsymbol{r}_{\mathrm{N}conv}^{(conv)}$-test in (15) validates $\boldsymbol{z}_{conv}$, one can conclude that the estimated measurements $\hat{\boldsymbol{z}}_{conv}^{(sync)}$ are contaminated by the presence of gross errors in $\boldsymbol{z}_{sync}$. Considering this, $\hat{\boldsymbol{x}}^{(conv)}$ is validated and used to estimate the synchrophasors. Then, the $\boldsymbol{r}_{\mathrm{N}sync}^{(conv)}$-test in (20) can be applied to check violations and identify (all at once) the gross errors in PMU measurements.

Figure 1 presents a simplified block diagram of PHASE.

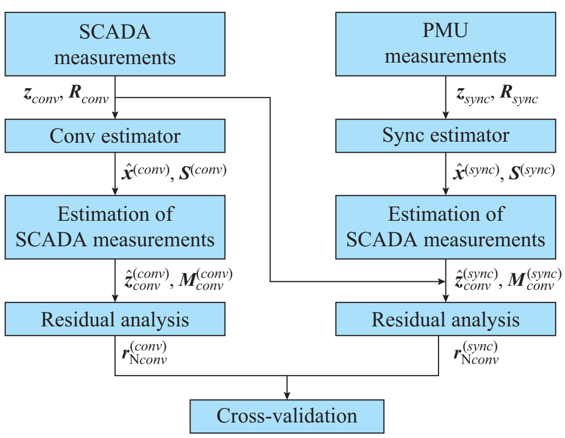

Fig. 1. Simplified block diagram of PHASE.

The following procedure is employed in the data validation to improve the processing of BD in Conv estimators after the normalized residual test has flagged some SCADA measurements as suspect.

Step 1: whenever possible, compute the normalized residuals of Sync estimator for the SCADA measurements flagged as suspect by the normalized residual test of Conv estimator.

Step 2: perform the test in (10) for each normalized residual computed in *Step 1*.

Step 3: perform the cross-validation test—SCADA measurements with normalized residuals of Conv and Sync estimators violating the corresponding thresholds are confirmed as BD, whereas those not flagged by Sync estimator are declared valid.

With the procedure described in *Steps 1-3*, the SCADA measurements flagged as BD by the cross-validation test can be block-eliminated/replaced, thus enhancing the residual analysis. Besides, the suspicious SCADA measurements not flagged as BD are declared valid, due to their large normalized residuals, which are caused by the well-known BD smearing effect. It is essential to note that the WLS method is applied to obtain the state estimates for both the Sync and Conv estimators. Readers can refer to recent studies [28], [31], and [32] to broaden the understanding of PHASE, where various experimental results, presented from different perspectives, corroborate the usefulness of PHASE.

It should be noted that the computing requirements for implementing the PHASE are already met by the infrastructure available in modern EMS. Regarding the communication requirements, numerous studies in the literature focus on establishing the communication infrastructure necessary to support various applications utilizing PMU data, including SE [33].

IV. Proposed Method

This section proposes allocating PMUs in a manner consistent with different objectives to enhance the capability of processing BD in SCADA-based SE. The local nature of SE is used in the proposed method to define the following concepts.

1) PMU-observable bus: the one whose state variables (voltage magnitude and phase angle) can be estimated using only synchrophasors.

2) PMU-observable SCADA measurement: the one that can be estimated using only the system state variables of PMU-observable buses.

3) Boundary measurement: SCADA measurement that is not PMU-observable and depends on state variables of at least one PMU-observable bus.

Three different strategies for processing the gross errors in SCADA measurements are proposed. ① Strategy 1: maximizing the number of bad SCADA measurements that can be eliminated all at once (block elimination), aiming at reducing repeated BD elimination/estimation cycles. ② Strategy 2: maximizing BD identification in SCADA measurements that constitute Cks. ③ Strategy 3: maximizing BD identification in SCADA measurements that constitute Cks and are associated only with a few remote terminal units (RTUs).

A. Strategy 1

As PHASE can perform block validation/elimination of BD, allocating PMUs to maximize the number of PMU-observable SCADA measurements is advantageous. In doing so, gross errors affecting these measurements can be readily identified and simultaneously eliminated by PHASE. Besides, having as many boundary measurements as possible is also desirable. Although boundary measurements cannot undergo the cross-validation test, gross errors involving them tend to contaminate the estimates of PMU-observable ones due to the smearing effect. Thus, the list of suspicious measurements can be substantially reduced by applying the cross-validation test to the PMU-observable ones. Such a feature will be illustrated in Section V. An experiment carried out with the IEEE 14-bus system shows that the effort to identify gross errors involving boundary measurements can also be alleviated by reducing the list of suspects with the cross-validation test.

Consider, for example, the six-bus system illustrated in Fig. 2(a)-(c) for different locations of PMUs. The PMU provides bus voltages and branch current synchrophasors (not depicted in the figure). Figure 2(a) and (b) shows three PMU-observable SCADA measurements if a PMU is allocated at bus 1 or bus 2. There are three boundary measurements (branch power flows $P_{4\text{-}3}$, $P_{4\text{-}5}$ and bus power injection P_4) in Fig. 2(a), compared with only two ($P_{4\text{-}3}$ and P_4) in Fig. 2(b). Figure 2(c) depicts a more advantageous situation with the PMU at bus 4, resulting in six PMU-observable SCADA measurements. In Fig. 2, the PMU-observable buses are marked in blue.

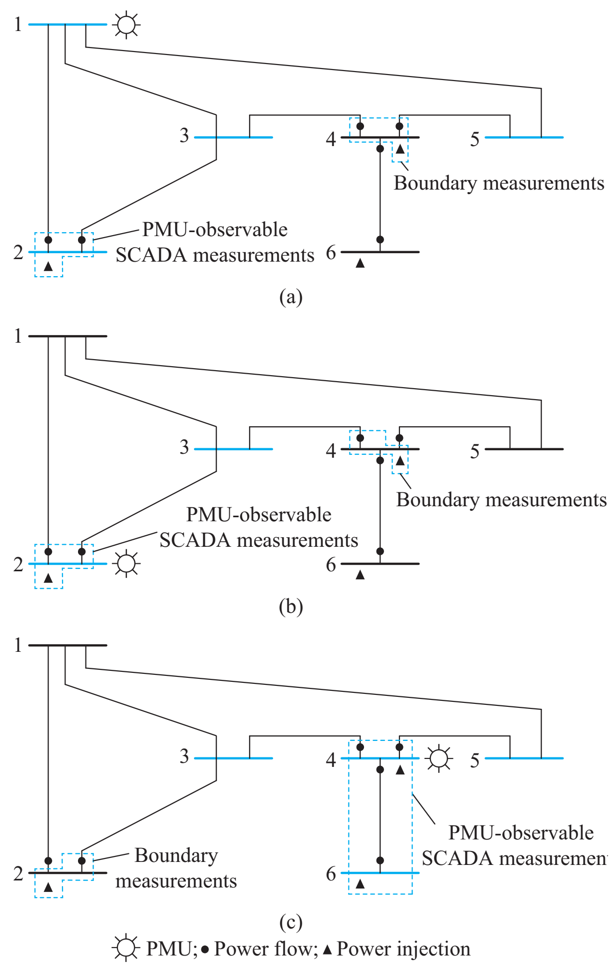

Fig. 2. Topology of six-bus system. (a) PMU is allocated at bus 1. (b) PMU is allocated at bus 2. (c) PMU is allocated at bus 4.

The proposed method is designed to allocate PMUs, considering the differences among various situations, as illustrated in Fig. 2(a)-(c). Considering the maximum coverage of SCADA measurements achieved with PHASE, an optimization problem can be stated as:

$$\max Obj_1 = \alpha m_{obs} + \beta m_{bound} \tag{21}$$

where m_{obs} is the number of PMU-observable SCADA measurements; m_{bound} is the number of boundary measurements; and α and β are the weights that are adjusted to maximize the number of PMU-observable SCADA measurements primarily (for block identification of BD), and among equally good solutions, the one with a larger m_{bound} is selected.

It is also important to emphasize that the PMU-observable SCADA measurements are calculated using the state estimates of the PMU-observable buses, i.e., those obtained by processing only synchrophasors during the cross-validation test. Whenever necessary, the PMU-observable SCADA measurements are suitable candidates for pseudo-measurements to replace SCADA measurements identified as BD [28].

B. Strategy 2

Suppose that a new objective, using PMU data to identify BD in Cks, is established. The measurements of a Ck should be submitted to the cross-validation test to cope with the impossibility of identifying BD affecting k or $k-1$ measurements of Cks [2]. This is possible if all measurements of the Ck are PMU-observable. Considering that a given set of PMUs is available, Strategy 2 detailed here aims to indicate where they can be allocated to maximize the number of Cks formed only by PMU-observable SCADA measurements. By doing that, the number of intractable BD is optimally reduced.

A different measuring system of the six-bus system is used here, as depicted in Fig. 3, to illustrate the capacity of dealing with BD in Cks according to the location of a PMU.

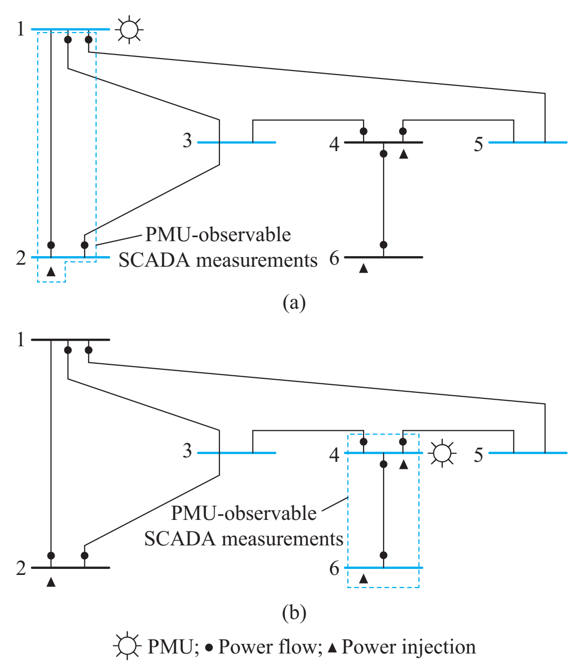

Fig. 3. A different measuring system of six-bus system. (a) PMU is allocated at bus 1. (b) PMU is allocated at bus 4.

The Cks identified by an extensive criticality analysis are shown in Table I, displayed by cardinality k. Identifying gross errors involving $k-1$ or k measurements of such Cks is challenging and cannot be handled by LNR test in classical SE [2]. However, PHASE can successfully process these BD [28], [31] if the PMUs are appropriately allocated.

TABLE I
CKS OBTAINED BY CRITICALITY ANALYSIS

k	Measurement
3	($P_{4\text{-}3}$, $P_{4\text{-}5}$, P_4); ($P_{2\text{-}1}$, $P_{2\text{-}3}$, P_2); ($P_{1\text{-}5}$, $P_{4\text{-}5}$, P_4); ($P_{1\text{-}5}$, $P_{4\text{-}3}$, P_4); ($P_{1\text{-}5}$, $P_{4\text{-}3}$, $P_{4\text{-}5}$)
4	($P_{4\text{-}6}$, P_4, $P_{6\text{-}4}$, P_6); ($P_{1\text{-}3}$, $P_{1\text{-}5}$, $P_{2\text{-}3}$, P_2); ($P_{1\text{-}3}$, $P_{1\text{-}5}$, $P_{2\text{-}1}$, P_2); ($P_{1\text{-}3}$, $P_{1\text{-}5}$, $P_{2\text{-}1}$, $P_{2\text{-}3}$)
5	($P_{4\text{-}3}$, $P_{4\text{-}5}$, $P_{4\text{-}6}$, $P_{6\text{-}4}$, P_6); ($P_{1\text{-}5}$, $P_{4\text{-}5}$, $P_{4\text{-}6}$, $P_{6\text{-}4}$, P_6); ($P_{1\text{-}5}$, $P_{4\text{-}3}$, $P_{4\text{-}6}$, $P_{6\text{-}4}$, P_6); ($P_{1\text{-}3}$, $P_{2\text{-}3}$, P_2, $P_{4\text{-}5}$, P_4); ($P_{1\text{-}3}$, $P_{2\text{-}3}$, P_2, $P_{4\text{-}3}$, P_4); ($P_{1\text{-}3}$, $P_{2\text{-}3}$, P_2, $P_{4\text{-}3}$, $P_{4\text{-}5}$); ($P_{1\text{-}3}$, $P_{2\text{-}1}$, P_2, $P_{4\text{-}5}$, P_4); ($P_{1\text{-}3}$, $P_{2\text{-}1}$, P_2, $P_{4\text{-}3}$, P_4); ($P_{1\text{-}3}$, $P_{2\text{-}1}$, P_2, $P_{4\text{-}3}$, $P_{4\text{-}5}$); ($P_{1\text{-}3}$, $P_{2\text{-}1}$, $P_{2\text{-}3}$, $P_{4\text{-}3}$, P_4); ($P_{1\text{-}3}$, $P_{2\text{-}1}$, $P_{2\text{-}3}$, $P_{4\text{-}5}$, P_4); ($P_{1\text{-}3}$, $P_{2\text{-}1}$, $P_{2\text{-}3}$, $P_{4\text{-}3}$, $P_{4\text{-}5}$)

Note: $P_{i\text{-}j}$ is the power flow between bus i and bus j; P_i is the power injection at bus i; and the measurements in a group like ($P_{4\text{-}3}$, $P_{4\text{-}5}$, P_4) are belonging to the same RTU.

Figure 3(a) shows five PMU-observable SCADA measurements with a PMU allocated at bus 1, whereas this number rises to six if the PMU is allocated at bus 4, as depicted in Fig. 3(b). However, in Fig. 3(a), four Cks are formed only by PMU-observable SCADA measurements against three Cks in the case of Fig. 3(b). Thus, if the capability of identifying gross errors in Cks is targeted, allocating the PMU at bus 1 would be preferable.

The following optimization problem can be stated, in which the number of Cks covered by PHASE is maximized:

$$\max Obj_2 = \mu N_{Ck} + \gamma m_{obs} \tag{22}$$

where N_{Ck} is the total number of Cks formed by all PMU-observable SCADA measurements; and μ and γ are the weights that are adjusted to prioritize the maximization of N_{Ck} (to cover as many Cks as possible), also concerning m_{obs}. Note that the term m_{bound} is not relevant in Strategy 2 and therefore not optimized in (22), since the secondary objective here is to maximize m_{obs}. The same will happen in Strategy 3.

C. Strategy 3

In Section IV-B, the objective is to maximize the number of cases in which PHASE can overcome the problems faced by the LNR method in processing multiple BD involving criticalities. The proposed method is flexible enough to incorporate practical aspects of the processing of BD, allowing for different trade-offs between the block elimination of gross errors and the identification of spurious measurements in Cks.

In some cases, errors affecting SCADA measurements of a single RTU are more likely to occur than those simultaneously compromising measurements coming from RTUs at different buses. This means that the gross errors in Cks formed by measurements distributed over a few RTUs are more likely to occur than those involving measurements distributed over many RTUs.

Consider, for example, using PHASE to tackle gross errors in Cks formed by measurements that belong to the same RTU. Allocating a PMU at bus 4, as in Fig. 3(b), is better than that at bus 1, as in Fig. 3(a). The reason for this is that in both cases, the criticality analysis has identified only one Ck formed by PMU-observable SCADA measurements belonging to the same RTU, i.e., ($P_{2\text{-}1}$, $P_{2\text{-}3}$, P_2) in Fig. 3(a) and ($P_{4\text{-}3}$, $P_{4\text{-}5}$, P_4) in Fig. 3(b). However, the number of PMU-observable SCADA measurements in Fig. 3(b) is greater than that in Fig. 3(a). While the number of Cks is prioritized for BD identification, the total number of PMU-observable SCADA measurements remains a concern for block elimination of BD in general, not just for Cks. Then, the optimization problem (22) can be modified to consider that only Cks formed by measurements distributed over a given number of RTUs are considered when allocating PMUs. Thus, the following objective function can be proposed:

$$\max Obj_3 = \mu N_{Ck}^{r} + \gamma m_{obs} \tag{23}$$

where N_{Ck}^{r} is the number of Cks formed only by PMU-observable SCADA measurements that belong to no more than r RTUs, which should be maximized. The weights μ and γ are the ones already presented in (22), with the maximization of N_{Ck}^{r} being prioritized over the maximization of m_{obs} in (23).

Strategy 3 exemplifies how the analyst can quickly adapt the proposed method to pursue other objectives, considering the practical experience and/or expert knowledge.

It should be noted that all the strategies presented above focus on allocating PMUs to improve the debugging capability of SE data. Traditional strategies, however, focus on enhancing the network observability and tend to favor the allocation of PMUs at buses that have more incident branches [15], i.e., those directly connected to more buses in the network. In doing so, more system states (bus voltage magnitudes and angles) are covered by the PMU measurements. This kind of strategies is not practical when aiming to enhance data debugging, as PMUs should be allocated to cover as many SCADA measurements as possible, or measurements that are critical or form critical k-tuples. For example, considering the illustrative six-bus system, buses 1, 3, and 4 would be regarded as equally suitable for PMU allocation according to the traditional strategies, giving no guidance on the best choice to improve data debugging capability.

Other strategies for allocating PMUs, suggested for future work, can be employed in the following cases.

1) Only m_{obs} and Cks in more relevant parts of the network, which are essential to be monitored, are considered.

2) Only Cks with a selected cardinality up to k are to be covered.

3) Historical data of measurement failures are considered.

4) The optimal trade-offs among different grid monitoring objectives are considered, which is a challenging task that involves solving a complex multi-objective optimization problem.

Regarding the use of PMU data, the temporary loss of the time-synchronization signal (LTSS) may occur due to GPS antenna failure, atmospheric disturbances, or electromagnetic interference [34], [35], which can affect the state estimated by the Sync estimator. The cross-validation test considered in this paper may be affected only if LTSS and bad SCADA measurements co-occur. This is a topic that warrants further investigation.

It is worth noting that the problem at hand can be classified as a combinatorial optimization, and the objective functions presented in (21)-(23) are well-suited for applying metaheuristics. Finding the best optimization technique to solve the problem is out of the scope of this paper.

V. Test Results

This section presents the numerical results of tests performed with the IEEE 14-, 30-, and 118-bus systems. The SCADA measurements are provided in [36] and [37].

This paper presents an enhanced method for utilizing PMUs to mitigate corrupted measurements from the SCADA system, which are processed through conventional SCADA-based SE algorithms. The proposed method involves strategically allocating a pre-defined number of PMUs, aiding PHASE to identify multiple BD points simultaneously in a single iteration of the LNR criterion. This method is more efficient compared with the traditional BD treatment, which requires numerous cycles of residual analysis. Several percentage metrics are adopted as a standardized and easily comparable way of presenting the results of the performance improvements achieved by the proposed method. The following percentage metrics are evaluated in relation to the number of SCADA measurements m: PMU-observable SCADA measurements (m_{obs}/m), boundary measurements (m_{bound}/m), Cks formed by all PMU-observable SCADA measurements (N_{Ck}/m), and Cks formed only by PMU-observable SCADA measurements that belong to no more than one RTU (N_{Ck}^{1}/m).

The proposed method is applied to allocate the available PMUs according to the strategies presented in Section IV. A genetic algorithm (GA) searches for an optimized solution to the PMU allocation problem. A binary string represents the chromosome (solution vector), and each gene is associated with a given network bus. The buses to which a PMU is allocated are those whose associated genes assume a value of 1, whereas all the others assume a value of 0. Crossover operations are performed so that the new individuals always have the pre-defined number of PMUs. Individuals that undergo mutation are randomly selected. A single PMU is randomly reallocated according to the mutation operator. An elitist strategy is also considered to guarantee that the best individuals of a population participate in the next generation. Before a new population is formed, the best individual from the current population undergoes a local search to find even better solutions. A PMU is randomly selected and reallocated to one of the buses directly connected to its current bus, which is then tested.

Regarding the weights adopted in the objective functions of (21)-(23), the weight α is ten times the weight β in Obj_1, whereas the weight γ is ten times the weight μ in Obj_2 and Obj_3.

A. Boundary Measurements

With PHASE, the cross-validation test presented in Section III can be carried out to perform block validation/elimination of bad SCADA measurements that are PMU-observable. In the case of multiple BD, PHASE is also expected to aid in data debugging of boundary measurements (defined in Section IV), i.e., those that depend on the state variables of both PMU-observable buses and those that are not observable by PMUs. Although the cross-validation test cannot be applied to the boundary measurements, the local nature of the smearing effect caused by gross errors involving boundary measurements can also raise suspicions about PMU-observable SCADA measurements, which can undergo the cross-validation test. Thus, when gross errors occur in boundary measurements, one can narrow down the list of suspicious measurements by using the cross-validation test, as in such cases, it is more likely that healthy measurements deemed suspect due to the smearing effect are PMU-observable. These healthy measurements can be promptly removed from the list of suspects after applying the cross-validation test. Consequently, this can also reduce the computational effort to eliminate BD using the LNR test. It is then desirable to maximize the number of PMU-observable SCADA measurements and to have as many boundary measurements as possible (see Section IV-A).

Tests are carried out to investigate the potential benefits of PHASE for data debugging when boundary measurements are involved, which is not addressed in [28]. The IEEE 14-bus system with only SCADA measurements is illustrated in Fig. 4, where the introduction of one PMU at each bus, one at a time, is simulated.

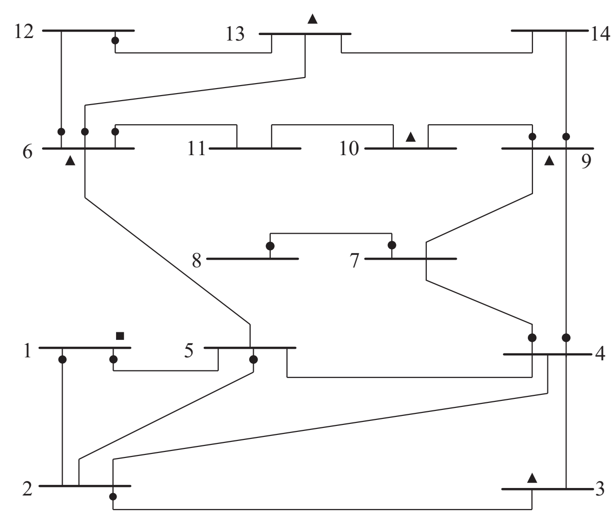

Fig. 4. IEEE 14-bus system with SCADA measurements.

Consider also that multiple different BD involving boundary measurements are inserted in each simulated case. Based on the normalized residuals, a list of suspicious measurements is formed. Those that can undergo cross-validation test are removed (all at once) from the list. The number of combinations performed in the identification-elimination cycles, to scan through the reduced list of suspects and eliminate remaining BD, is computed, which is compared with the number of combinations required without the block validation carried out by PHASE. It is observed that the list of suspects can be significantly reduced when there are gross errors involving boundary buses. Considering all tests performed with the IEEE 14-bus system, i.e., simulating multiple BD when a PMU is allocated at a single bus, the number of estimation-elimination cycles is, on average, reduced by 62%. The maximum and minimum reductions observed in the tests are 98% and 30%, respectively. As the inclusion of a PMU at each bus is tested, various local topology conditions and measurement redundancies are simulated. It is also worth mentioning that tests considering gross errors in measurements that are neither PMU-observable nor boundary measurements have also been performed. Still, only a slight reduction in the list of suspects is observed in such cases.

B. Tests with Strategy 1

Now, consider the IEEE 30- and 118-bus systems in the following. Table II shows the results with Strategy 1. The number of PMU-observable SCADA measurements (those that can be validated simultaneously) and the number of boundary buses (also targeted) are presented, considering the allocation of different numbers of PMUs (N_{PMU}). The percentage of PMU-observable SCADA measurements (those that can be block-validated) and the percentage of boundary measurements (also targeted) covered by PHASE are presented, considering different numbers of PMUs (N_{PMU}) to be allocated. These percentages correspond to the number of SCADA measurements covered by PHASE, out of a total of 44 SCADA measurements in the IEEE 30-bus system [36] and 176 SCADA measurements in the IEEE 118-bus system [37]. In Table II, the quantities that are optimized are boldfaced. N_{Ck} and N_{Ck}^1 that are covered by PHASE with the PMU allocations are also indicated. Although not directly related to the objective function of Strategy 1, these quantities are presented to facilitate further comparison with the results obtained by applying other optimization strategies.

TABLE II
RESULTS WITH STRATEGY 1

System	N_{PMU}	Percentage of different variables to m (%)			
		m_{obs}	m_{bound}	N_{Ck}	N_{Ck}^1
IEEE 30-bus	4	**54.55**	**40.91**	6.40	44.44
	6	**72.73**	**25.00**	23.02	66.67
	8	**86.36**	**13.64**	72.20	77.78
	10	**100.00**		100.00	100.00
IEEE 118-bus	5	**28.41**	**19.88**	8.53	12.50
	10	**51.14**	**29.55**	15.99	25.00
	15	**69.89**	**22.16**	31.97	37.50
	20	**83.52**	**14.77**	62.17	43.75

During the simulations, the pool solutions generated by the GA are successfully optimized to have as many PMU-observable SCADA measurements and boundary measurements as possible. The results presented in Table II are the best solutions found. Note that the allocation of four PMUs in the IEEE 30-bus system results in 95.45% of the SCADA measurements (42 out of 44) becoming either PMU-observable or boundary measurements. In the IEEE 118-bus system, 80.69% of the SCADA measurements (142 out of 176) are either PMU-observable or boundary measurements with the allocation of only ten PMUs. Most of them are PMU-observable, as this is the primary term of Obj_1 in (21).

C. Tests with Strategy 2

Table III shows the results with Strategy 2, i.e., to maximize the total number of Cks (cardinalities up to $k=5$) covered by PHASE with the PMU allocation. The results of each tested measuring system are obtained by considering different numbers of PMUs to be allocated. Although the total number of Cks associated with a single RTU is not the focus of Strategy 2, N_{Ck}^1 is also presented to allow a comparison with the results in Section IV-C. Again, the quantities that are optimized are boldfaced.

TABLE III
RESULTS WITH STRATEGY 2

System	N_{PMU}	Percentage of different variables to m (%)			
		m_{obs}	m_{bound}	N_{Ck}	N_{Ck}^1
IEEE 30-bus	4	**47.74**	34.09	**17.61**	33.33
	6	**68.18**	20.45	**41.05**	66.67
	8	**86.36**	13.64	**72.20**	77.78
	10	**100.00**		**100.00**	100.00
IEEE 118-bus	5	**17.61**	23.29	**17.23**	18.75
	10	**35.23**	34.09	**36.94**	18.75
	15	**52.27**	27.84	**58.08**	31.25
	20	**72.73**	18.18	**77.44**	37.50

The number of Cks covered (N_{Ck}) increases when Strategy 2 is adopted, as can be observed by comparing the results in Table III with those in Table II. On the other hand, the total number of observable measurements (m_{obs}) is reduced. This an expected result since, with Strategy 2, the maximization of m_{obs} is secondary to the maximization of N_{Ck}. Note that the values of m_{bound} shown in Table III are not optimized in the objective function presented in (22).

D. Tests with Strategy 3

When implementing Strategy 3, the maximization of N_{Ck}^1 is the primary focus. The results obtained for the allocation of different numbers of PMUs are presented in Table IV.

TABLE IV
RESULTS WITH STRATEGY 3

System	N_{PMU}	Percentage of different variables to m (%)			
		m_{obs}	m_{bound}	N_{Ck}	N_{Ck}^1
IEEE 30-bus	4	**52.27**	29.55	9.26	**55.56**
	6	**70.45**	27.27	16.82	**77.78**
	8	**81.82**	18.18	39.92	**100.00**
	10	**100.00**		100.00	**100.00**
IEEE 118-bus	5	**16.48**	25.55	5.15	**37.50**
	10	**30.11**	35.22	13.32	**68.75**
	15	**50.00**	27.27	22.38	**87.50**
	20	**71.02**	18.75	36.23	**100.00**

Table IV shows that N_{Ck}^1 is significantly larger when adopting Strategy 3, compared with the results obtained with Strategies 1 and 2. The number of PMU-observable SCADA measurements (a secondary objective with Strategy 3) is less than that obtained with Strategy 1 and similar to that obtained with Strategy 2. Note that in Table IV, the values of m_{bound} are not optimized in the objective function presented in (23).

The results presented in Tables II-IV show how the increase in the number of PMUs leads to an increase in the number of SCADA measurements covered by the PHASE. It can be observed that more measurements can be readily eliminated in case of gross errors affecting them, without the need to apply the LNR method and go through burdensome elimination-estimation cycles. Regarding the increase in the number of Ck-tuples covered by PHASE, the risk of failing to identify gross errors that affect $k-1$ measurements of a Ck-tuple is reduced. The trade-offs between the cost associated with the increment of PMUs and the improvement in BD debugging through PHASE can also be assessed if a multi-objective optimization approach is implemented. This will be the subject of future investigation.

To illustrate the effectiveness of the proposed method, Fig. 5 presents the results extracted from Tables II-IV, which correspond to the allocation of 15 PMUs in the IEEE 118-bus system with different strategies. It can be observed that m_{obs}, N_{Ck}, and N_{Ck}^1 are optimized and significantly larger when adopting Strategies 1, 2, and 3, respectively. It should be noted that the results shown in Fig. 5 correspond to the allocation of less than 50% PMUs needed to make all measurements PMU-observable. The results presented in this section demonstrate that the proposed method is flexible enough to accommodate various optimization strategies, which aim to enhance the processing of BD in SCADA-based estimators. Nevertheless, if available, a hybrid state estimator can also benefit from the PHASE and the proposed method. The hybrid state estimator can conveniently process only SCADA measurements, therefore playing the role of the Conv estimator shown in the flowchart of Fig. 1.

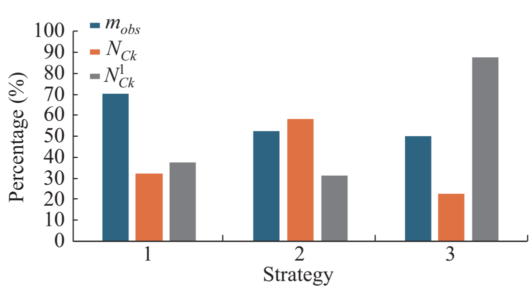

Fig. 5. Results corresponding to allocation of 15 PMUs in IEEE 118-bus system.

The computational burden of allocating PMUs with the three strategies is not a concern. The main reasons are: ① the intended application is to be carried out offline in planning studies of the metering system; ② the adopted objective functions are not complex ones, with the configurations in SCADA-based measurement system and their critical k-tuples known in advance. The tests performed so far indicate that the computational time to allocate 20 PMUs in the IEEE 118-bus system is expected to be no more than 17.5 s.

E. Final Remarks

The sensitivity analysis to the number of PMUs (4, 6, 8, and 10 for the IEEE 30-bus system, and 5, 10, 15, and 20 for the IEEE 118-bus system) depends on the strategy adopted in this paper, as shown in Tables II-IV. For instance, with 10 PMUs, all SCADA measurements in the IEEE 30-bus system are covered, regardless of strategy adopted. However, allocating 20 PMUs in the IEEE 118-bus system results in Strategy 1 covering 83.52% of SCADA measurements, Strategy 2 covering 72.73%, and Strategy 3 covering 71.02%. When considering the same allocation strategy, the increase in SCADA measurement coverage concerning the rise in the number of PMUs allocated is depicted in Tables II-IV.

PMUs provide direct bus voltage measurements, including magnitudes and phase angles [7], at much higher sampling rates (30 to 120 samples per second) than SCADA systems (1 sample every 2 to 10 s). To synchronize data for hybrid [38] and PHASE estimators [28], PMU readings are usually buffered for use in conjunction with SCADA measurements [39]. Aligning timestamps, resampling, and calibrating clocks are essential to prevent asynchronism between PMU and SCADA measurements, which can compromise the processing of BD.

PMUs are highly accurate, typically having very low total vector error (often less than 1%). However, their measurements are not immune to gross errors. Fortunately, bad PMU measurements are infrequent in well-maintained systems and typically arise from hardware faults, GPS issues, or communication errors. Nevertheless, tests showing the aptitude of PHASE to address gross errors in PMUs have been presented in [28].

The computation time to identify gross errors in SCADA measurements covered by PHASE comprises the time of running the Conv and Sync estimators, plus the time to cross-validate their results. This will depend on the efficiency of the implemented state estimators. Regarding the tests performed in this paper, the computational time spent in each step of PHASE to identify gross errors in four SCADA measurements of the IEEE 118-bus system with ten PMUs is: 0.52 s (Conv estimator), 0.12 s (Sync estimator), and 0.02 s (cross-validation test). The time spent by the Conv estimator is dominant, as a limited number of PMUs is assumed. No significant change in Sync estimator is observed when varying the number of PMUs allocated. Besides, as the Conv and Sync estimators are independent, they can be executed in parallel, lowering the total computational time even further.

VI. Conclusion

This paper presents a method to enhance the processing of BD in SE, particularly when gross errors occur in SCADA measurements. The proposed method maximizes the benefits of the PHASE served by PMUs allocated strategically. As many utilities employ state estimators that process only SCADA measurements, applying the proposed method is appealing. Different strategies are built to allow block validation of SCADA measurements, as well as the identification of BD in critical groups of measurements. Such features are optimized by using an evolutionary algorithm. Tests have been performed using the IEEE 14-, 30-, and 118-bus systems, considering the installation of different numbers of PMUs. The results show that the proposed method could successfully allocate the PMUs to achieve the goals defined in each allocation strategy, being flexible to accommodate different objectives.

References

[1] A. Monticelli, *Power System State Estimation: A Generalized Approach.* Norwell: Kluwer Academic Publishers, 1999.

[2] A. Abur and A. Gómez-Expósito, *Power System State Estimation: Theory and Implementation*. New York: Marcel Decker, 2004.

[3] M. B. D. C. Filho, *Power System State Estimation and Forecasting: Fundamentals and Advanced Topics*. Switzerland: Springer Nature, 2024.

[4] M. B. D. C. Filho, J. C. S. de Souza, and J. D. Glover, "Roots, achievements, and prospects of power system state estimation: a review on handling corrupted measurements," *International Transactions on Electrical Energy Systems*, vol. 29, no. 4, p. e2779, Apr. 2019.

[5] K. A. Clements and P. W. Davis, "Multiple bad data detectability and identifiability: a geometric approach," *IEEE Transactions on Power Delivery*, vol. 1, no. 3, pp. 355-360, Jul. 1986.

[6] V. B. B. Flôr, M. B. D. C. Filho, J. C. S. de Souza *et al*., "Critical data visualization to enhance protection schemes for state estimation," *IEEE Transactions on Smart Grid*, vol. 14, no. 2, pp. 1249-1261, Mar. 2023.

[7] A. G. Phadke and J. S. Thorp, *Synchronized Phasor Measurements and Their Applications*. New York: Springer, 2017.

[8] R. F. Nuqui and A. G. Phadke, "Phasor measurement unit placement techniques for complete and incomplete observability," *IEEE Transactions on Power Delivery*, vol. 20, no. 4, pp. 2381-2388, Oct. 2005.

[9] N. M. Manousakis, G. N. Korres, and P. S. Georgilakis, "Taxonomy of PMU placement methodologies," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 1070-1077, May 2012.

[10] N. M. Manousakis and G. N. Korres, "Optimal PMU placement for numerical observability considering fixed channel capacity: a semidefinite programming approach," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 3328-3329, Jul. 2016.

[12] P. M. Joshi and H. K. Verma, "Synchrophasor measurement applications and optimal PMU placement: a review," *Electric Power Systems Research*, vol. 199, p. 107428, Oct. 2021.

[13] M. K. Arpanahi, R. Torkzadeh, A. Safavizadeh *et al*., "A novel comprehensive optimal PMU placement considering practical issues in design and implementation of a wide-area measurement system," *Electric Power Systems Research*, vol. 214, p. 108940, Jan. 2023.

[14] A. A. Al-Qaralleh, K. Alawasa, and A. I. Al-Odienat, "Optimal location of PMUs to enhance state estimation of power systems," in *Proceedings of 2023 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology*, Amman, Jordan, May 2023, pp. 38-43.

[15] M. M. Ahmed, M. Amjad, M. A. Qureshi *et al*., "A critical review of state-of-the-art optimal PMU placement techniques," *Energies*, vol. 15, no. 6, p. 2125, Mar. 2022.

[16] J. Hazra, K. Das, B. K. S. Roy *et al*., "Multistage optimal PMU placement for hybrid state estimation," in *Proceedings of 2017 IEEE PES General Meeting*, Chicago, USA, Jul. 2017, pp. 1-5.

[17] M. Shahriar, I. Habiballah, and H. Hussein, "Optimization of phasor measurement unit (PMU) placement in supervisory control and data acquisition (SCADA)-based power system for better state-estimation performance," *Energies*, vol. 11, no. 3, p. 570, Mar. 2018.

[18] T. Bi, X. Qin, Q. Yang *et al*., "A novel hybrid state estimator for including synchronized phasor measurements," *Electric Power Systems Research*, vol. 78, no. 8, pp. 1343-1352, Aug. 2008.

[19] M. Göl and A. Abur, "A hybrid state estimator for systems with limited number of PMUs," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1511-1517, May 2015.

[20] G. N. Korres and N. M. Manousakis, "State estimation and bad data processing for systems including PMU and SCADA measurements," *Electric Power Systems Research*, vol. 81, no. 7, pp. 1514-1524, Jul. 2011.

[21] A. Gómez-Expósito, A. Abur, P. Rousseaux *et al*., "On the use of PMUs in power system state estimation," in *Proceedings Power System Computation Conferenc*e, Stockholm, Sweden, Aug. 2011, pp. 1-13.

[22] J. Chen and A. Abur, "Placement of PMUs to enable bad data detection in state estimation," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1608-1615, Nov. 2006.

[23] B. Gou and R. G. Kavasseri, "Unified PMU placement for observability and bad data detection in state estimation," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2573-2580, Nov. 2014.

[24] J. A. D. Massignan, J. B. A. London, C. D. Maciel *et al*., "PMUs and SCADA measurements in power system state estimation through Bayesian inference," in *Proceedings of 2019 IEEE Milan PowerTec*, Milan, Italy, Jun. 2019, pp. 1-6.

[25] O. Darmis and G. Korres, "A survey on hybrid SCADA/WAMS state estimation methodologies in electric power transmission systems," *Energies*, vol. 16, no. 2, p. 618, Jan. 2023.

[26] G. Cheng, Y. Lin, A. Abur *et al*., "A survey of power system state estimation using multiple data sources: PMUs, SCADA, AMI, and beyond," *IEEE Transactions on Smart Grid*, vol. 15, no. 1, pp. 1129-1151, Jan. 2024.

[27] A. S. Costa, A. Albuquerque, and D. Bez, "An estimation fusion method for including phasor measurements into power system real-time modeling," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1910-1920, May 2013.

[28] M. B. D. C. Filho and J. C. S, de Souza, "Enhanced bad data processing by phasor-aided state estimation," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2200-2209, Sept. 2014.

[29] K. C. Sou, H. Sandberg, and K. H. Johansson, "Computing critical *k*-tuples in power networks," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1511-1520, Aug. 2012.

[30] A. M. Coimbra, J. C. S. de Souza, M. B. D. C. Filho *et al*., "Network-based approach to identify criticalities in state estimation," *IEEE Transactions on Power Systems*, vol. 36, no. 4, pp. 3394-3405, Jul. 2021.

[31] T. Wu, C. Y. Chung, and I. Kamwa, "A fast state estimator for systems including limited number of PMUs," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4329-4339, Nov. 2017.

[32] J. Zhao, G. Zhang, M. L. Scala *et al*., "Enhanced robustness of state estimator to bad data processing through multi-innovation analysis," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 1610-1619, Aug. 2017.

[33] M. Shahraeini, M. H. Javidi, and M. S. Ghazizadeh, "Communication infrastructure planning for wide area measurement systems in power systems," *International Journal of Communication Networks and Distributed Systems*, vol. 10, no. 4, p. 319, Jun. 2013

[34] G. Feng and A. Abur, "Fault location using wide-area measurements and sparse estimation," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2938-2945, Jul. 2016.

[35] M. R. Jegarluei, P. Aristidou, W. Fernandes *et al*., "Wide-area backup protection using sparse synchronized/unsynchronized PMU measurements," *IEEE Transactions on Power Delivery*, vol. 38, no. 4, pp. 2630-2640, Aug. 2023.

[36] H. R. de O. Rocha, J. A. L. Silva, J. C. S. de Souza *et al*., "Fast and flexible design of optimal metering systems for power systems monitoring," *Journal of Control*, *Automation and Electrical Systems*, vol. 29, no. 2, pp. 209-218, Mar. 2018.

[37] M. B. D. C. Filho, J. C. S. de Souza, and J. E. V. Tafur, "Quantifying observability in state estimation," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2897-2906, Aug. 2013.

[38] J. Zhao and L. Mili, "A framework for robust hybrid state estimation with unknown measurement noise statistics," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 5, pp. 1866-1875, May 2018.

[39] V. Murugesan, Y. Chakhchoukh, V. Vittal *et al*., "PMU data buffering for power system state estimators," *IEEE Power and Energy Technology Systems Journal*, vol. 2, no. 3, pp. 94-102, Sept. 2015.

Julio Cesar Stacchini de Souza received the D.Sc. degree in electrical engineering from the Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, in 1996. Since 1992, he has been with Fluminense Federal University, Rio de Janeiro, Brazil, where he is currently a Full Professor. In 2014-2015, he was a Visiting Scholar at Imperial College London, London, UK. His research interests include computation method for power system and intelligent system application.

Milton Brown Do Coutto Filho received the D.Sc. degree in electrical engineering from the Federal University of Rio de Janeiro, Rio de Janeiro, Brazil, in 1983. His employment experience includes the position of an Associate Professor at Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, in 1977-1994, and a Visiting Scholar at Northeastern University, Boston, USA, in 1992-1993. Since 1994, he has been with Fluminense Federal University, Rio de Janeiro, Brazil, as a Full Professor. His research interests include computer application in power system.

Marcio Andre Ribeiro Guimaraens received the D. Sc. degree from the Institute of Computing, Fluminense Federal University, Rio de Janeiro, Brazil, in 2015. He is currently with Fluminense Federal University, as an Adjunct Professor. His research interests include computation method and intelligent system application.