

Observability Analysis and Restoration for Distribution Grids: A Robust Greedy Numerical Algorithm

Natalia Alguacil^{id}, Senior Member, IEEE, Henar Herrero^{id}, and Cristina Solares^{id}

Abstract—This paper presents a robust and efficient algorithm for analyzing and restoring the observability of distribution grids. A linearized load flow model is used to characterize three- and multi-phase distribution grids under balanced and unbalanced conditions. Assuming that metering devices are installed to measure nodal magnitudes at some buses, the rank of the matrix associated with the non-metered magnitudes is determined by using a greedy strategy combined with a Gram-Schmidt orthonormalization procedure. Unlike other numerical-based techniques, the proposed algorithm avoids common errors caused by denominators close to zero. As an additional salient feature, a mathematically sound stopping criterion is provided to guarantee the robustness of the proposed algorithm. Furthermore, for unobservable grids, the algorithm identifies the linearly dependent rows in the aforementioned matrix, which correspond to the buses that should be metered to restore grid observability. The effective performance of the proposed algorithm is illustrated using several medium- and low-voltage benchmarks with up to 906 buses.

Index Terms—Distribution grids, multi-phase networks, observability analysis, observability restoration, phasor measurement unit (PMU).

I. INTRODUCTION

OBSERVABILITY analysis and restoration have become essential for the reliable and cost-efficient operation of the increasingly active distribution grids [1]. In the past, due to their passive behavior and the scarcity of metering devices, these networks were typically operated with limited observability [2], [3]. However, with the recent deployment of smart meters [4] and phasor measurement units (PMUs) [5], full observability of distribution grids can be achieved or restored

if such devices are conveniently located [6]. Furthermore, the observability of distribution grids can be improved if the complementary strengths of smart meters and PMUs are exploited [7].

Similarly to the observability analysis carried out for transmission systems, accurate and updated information on their topology, network parameters, and metering locations is also required in distribution systems [3], [8]. However, since distribution grids are reconfigured more frequently than transmission networks [9], computing limitations are tighter for these grids. As a consequence, dynamic observability analysis should be efficiently implemented to capture topological changes during grid operation. In addition to the reconfiguration issue, distribution grids may operate under unbalanced conditions caused by both untransposed distribution lines and unbalanced loads, which requires a multi-phase modeling [10]. In such cases, the phases in which single- or multi-phase equipment is connected should be clearly identified [11]. Note that phase identification is part of topology identification [12], [13], which in turn is required for observability analysis.

This problem has been thoroughly studied for transmission systems [14], [15], [16], [17], [18], [19], [20], [21], [22], wherein single-phase equivalents for balanced conditions are adopted. However, few works are available for distribution grids [1]. Existing approaches are broadly classified as topological [23], [24] and numerical algorithms [25], [26], [27], [28], [29].

In [23], grid observability was studied using a topological algorithm in which the radial nature of the network was exploited. Likewise, a topological algorithm was proposed in [24], in which the rank of the system matrix is determined by analyzing the topology of the system graph. A numerical criterion was also provided to analyze the rank deficiency of the system matrix.

The numerical algorithm described in [25] was based on the inverse function theory and the Jacobian matrix of the system. Moreover, an orthogonal linear rotational transformation was used to address the large r/x ratio in distribution systems. In [26], a multi-phase numerical observability analysis was developed using an orthogonal factorization of the three-phase measurement Jacobian matrix. To that end, a weak coupling $P-v$ and $Q-\theta$, as in transmission systems, was considered. The work presented in [27] proposed a numerical algorithm for observability analysis based on the transformation of the Gain matrix using Cholesky decomposition.

Manuscript received 10 June 2022; revised 18 October 2022 and 30 January 2023; accepted 8 March 2023. Date of publication 14 March 2023; date of current version 23 October 2023. This work was supported in part by Grant PID2021-126566OB-I00 and Grant PID2021-122579OB-I00, funded by the Spanish Ministry of Science and Innovation MCIN/AEI/10.13039/501100011033 and by the ERDF, EU “A Way of Making Europe”; and in part by Grant SBPLY/21/180501/000154, funded by the Junta de Comunidades de Castilla-La Mancha, Spain, and by the ERDF. Paper no. TSG-00822-2022. (Corresponding author: Natalia Alguacil.)

Natalia Alguacil is with the Departamento de Ingeniería Eléctrica, Electrónica, Automática y Comunicaciones, E.T.S.I. Industrial, Universidad de Castilla-La Mancha, 13071 Ciudad Real, Spain (e-mail: Natalia.Alguacil@uclm.es).

Henar Herrero and Cristina Solares are with the Departamento de Matemáticas, Universidad de Castilla-La Mancha, 13071 Ciudad Real, Spain (e-mail: Henar.Herrero@uclm.es; Cristina.Solares@uclm.es).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TSG.2023.3256488>.

Digital Object Identifier 10.1109/TSG.2023.3256488

A decoupled PQ grid model typically used in transmission systems was revisited to identify unobservable branches and critical data for unbalanced three-phase distribution systems. In addition, a triangular factorization of the Gain matrix was performed considering branch current magnitudes and angles as the state variables. In [28], a methodology to generate pseudo-measurements along the distribution network was proposed to guarantee full observability of the grid. Recently, in [29], correlations between measurements were taken into account to improve system observability and an observability metric useful for state estimation was provided.

Note that in many studies the effect of the distribution network is modeled by linear load flow equations such as those presented in [30]. It is also worth emphasizing that the computational burden of existing techniques [23], [24], [25], [26], [27], [28], [29] grows with the size of the required vectors and matrices, which may be up to three times larger for unbalanced operation as compared to a balanced setting. Thus, one of the yet unresolved challenges for both topological and numerical algorithms is related to their scalability when addressing large distribution grids. Moreover, observability restoration is typically neglected in the related literature, [27] being a relevant exception. Unfortunately, the approach described in [27] is prone to issues associated with its reliance on the method presented in [14], which may involve dividing by numbers close to zero and/or pivoting positions.

As an alternative, a novel robust greedy algorithm is proposed here for both observability analysis and observability restoration of three- and multi-phase distribution grids under balanced and unbalanced conditions. Based on a customarily used linear network model [30], the proposed algorithm starts by conveniently partitioning the system matrix, as done in [18], in which the problem of observability in linear systems of equations was addressed. Bearing in mind that measurements of active/reactive power injections, voltage magnitudes, and voltage angles are only available for a subset of buses, the algorithm determines whether these measurements are sufficient or not to guarantee grid observability. That is, if the corresponding nodal magnitudes in the non-metered buses can be inferred from the available measurements. Furthermore, in case of unobservability, the algorithm identifies the buses at which meters should be located to restore observability while considering practical multi-phase metering devices. Several case studies, including multi-phase grids [31], are examined to comprehensively illustrate the performance in small instances and the computational efficiency of the algorithm in larger instances.

This paper significantly departs from the existing literature in the following methodological aspects. First, the proposed algorithm calculates the rank of the system matrix, an optimal basis of the row space, and the coordinates of the remaining rows in that basis. Interestingly, the methodology requires neither inverting matrices nor selecting pivots, while avoiding the errors caused when dividing by values close to zero [25], [26], [27], [28], [29]. Moreover, unlike the algorithms considered in [25], [26], [27], [28], [29], the proposed approach is numerically stable since it does not require setting any tolerance to distinguish between zero and nonzero values. Additionally,

in contrast to [23], [24], [25], [26], [28], and [29] observability restoration is considered. Furthermore, the proposed approach differs from the algorithm presented in [22] for transmission grids in two crucial aspects: 1) the algorithm relies on a mathematically sound stopping criterion that does not require tuning tolerances, and 2) the restoration of observability is carried out by identifying the best candidate buses to be metered using a practical criterion suitable for distribution grids. Finally, the above distinctive features allow analyzing and restoring observability for instances that are far larger than those addressed in [23], [24], [25], [26], [27], [28], [29].

The main contributions of this work are as follows:

- Based on an existing numerical algorithm for observability analysis in transmission systems and a widely used linear load flow model, a novel observability approach is developed to accommodate the singularities of both balanced and unbalanced distribution grids. As a major novelty within the context of distribution grids, the proposed methodology avoids the errors caused by dividing by values close to zero.
- Unlike its transmission-related counterpart and most approaches for distribution systems, the restoration of grid observability is handled by identifying the non-metered buses at which smart meters or PMUs should be located. Furthermore, consistently with practical metering devices that meter all the phases of a bus, observability is restored by including the best candidate bus in the subset of metered buses with all its phases as a single block.
- A mathematically sound stopping criterion is provided to guarantee the robustness of the algorithm, thus avoiding tuning tolerances.
- The efficiency of the algorithm is empirically proven using two IEEE benchmarks and a 906-bus feeder under different scenarios, which are generated by modifying both the number and the location of metered buses. Thus, for the first time in the literature, distribution system operators are provided with a practical tool to analyze and restore observability.

The rest of this paper is organized as follows. In Section II, the observability matrices for distribution systems are derived. In Section III, the proposed algorithm for observability analysis and restoration is presented. In Section IV, numerical results from several case studies are provided and discussed. In Section V, the main conclusions are summarized. Finally, a theorem is included in the Appendix to justify the suitability of the stopping criterion of the proposed algorithm.

II. PROBLEM STATEMENT

This section presents the matrices required for the proposed algorithm for distribution grid observability analysis and restoration. Based on [24] and [30], the following linear approximation of the load flow equations is considered:

$$\mathbf{v} \approx \mathbf{R}\mathbf{p} + \mathbf{X}\mathbf{q} \quad (1)$$

$$\boldsymbol{\theta} \approx \mathbf{X}\mathbf{p} - \mathbf{R}\mathbf{q} \quad (2)$$

where \mathbf{v} is the vector of voltage magnitudes, $\boldsymbol{\theta}$ is the vector of voltage angles, \mathbf{p} is the vector of active power nodal

injections, and \mathbf{q} is the vector of reactive power nodal injections, whereas, for a given grid topology, \mathbf{R} and \mathbf{X} are the real-valued matrices computed from the line resistances and reactances, respectively, as done in [24].

A. Single-Phase Formulation With Smart Meter Data

Under balanced conditions, three-phase distribution grids can be characterized using an equivalent single-phase model. Thus, the vectors and matrices in (1) and (2) are n - and $n \times n$ -dimensional, respectively. When smart meters are deployed, the nodal vectors \mathbf{v} , \mathbf{p} , and \mathbf{q} are available. As done in [24], it is assumed that smart meters are located at a subset of buses $\mathcal{M} \subset \mathcal{N}$, where $\mathcal{N} = \{1, 2, \dots, n\}$ is the set of all buses in the distribution grid. Thus, the above nodal vectors can be partitioned as $\mathbf{v} = [\mathbf{v}_{\mathcal{M}}, \mathbf{v}_{\hat{\mathcal{M}}}]$, $\mathbf{p} = [\mathbf{p}_{\mathcal{M}}, \mathbf{p}_{\hat{\mathcal{M}}}]$, and $\mathbf{q} = [\mathbf{q}_{\mathcal{M}}, \mathbf{q}_{\hat{\mathcal{M}}}]$, where $\hat{\mathcal{M}} = \mathcal{N} \setminus \mathcal{M}$ denotes the subset of non-metered buses.

Using the notation introduced above, the system of equations in (1) can be partitioned as follows:

$$\begin{bmatrix} \mathbf{v}_{\mathcal{M}} \\ \mathbf{v}_{\hat{\mathcal{M}}} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{\mathcal{M},\mathcal{M}} & \mathbf{R}_{\mathcal{M},\hat{\mathcal{M}}} \\ \mathbf{R}_{\hat{\mathcal{M}},\mathcal{M}} & \mathbf{R}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \end{bmatrix} \begin{bmatrix} \mathbf{p}_{\mathcal{M}} \\ \mathbf{p}_{\hat{\mathcal{M}}} \end{bmatrix} + \begin{bmatrix} \mathbf{X}_{\mathcal{M},\mathcal{M}} & \mathbf{X}_{\mathcal{M},\hat{\mathcal{M}}} \\ \mathbf{X}_{\hat{\mathcal{M}},\mathcal{M}} & \mathbf{X}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \end{bmatrix} \begin{bmatrix} \mathbf{q}_{\mathcal{M}} \\ \mathbf{q}_{\hat{\mathcal{M}}} \end{bmatrix} \quad (3)$$

where subindices \mathcal{M} and $\hat{\mathcal{M}}$ refer to the rows/columns of the matrices and vectors associated with the metered and non-metered buses, respectively.

System (3) can be rearranged to place the non-metered magnitudes in the left-hand side and the metered magnitudes in the right-hand side. That is,

$$\begin{bmatrix} \mathbf{R}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{X}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{0}_{\mathcal{M},\hat{\mathcal{M}}} \\ \mathbf{R}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} & \mathbf{X}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} & -\mathbf{I}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \end{bmatrix} \begin{bmatrix} \mathbf{p}_{\hat{\mathcal{M}}} \\ \mathbf{q}_{\hat{\mathcal{M}}} \\ \mathbf{v}_{\hat{\mathcal{M}}} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_{\mathcal{M}} - \mathbf{R}_{\mathcal{M},\mathcal{M}}\mathbf{p}_{\mathcal{M}} - \mathbf{X}_{\mathcal{M},\mathcal{M}}\mathbf{q}_{\mathcal{M}} \\ -\mathbf{R}_{\hat{\mathcal{M}},\mathcal{M}}\mathbf{p}_{\mathcal{M}} - \mathbf{X}_{\hat{\mathcal{M}},\mathcal{M}}\mathbf{q}_{\mathcal{M}} \end{bmatrix} \quad (4)$$

where $\mathbf{0}$ denotes a matrix of zeroes and \mathbf{I} is the identity matrix.

To check if the system is observable, i.e., if \mathbf{v} , \mathbf{p} , and \mathbf{q} can be determined for the non-metered buses, the transpose of the matrix in the left-hand side of (4) is required [22], which is formulated as follows:

$$\mathbf{H} = \begin{bmatrix} \mathbf{R}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{R}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \\ \mathbf{X}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{X}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \\ \mathbf{0}_{\mathcal{M},\hat{\mathcal{M}}} & -\mathbf{I}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \end{bmatrix} \quad (5)$$

Matrix \mathbf{H} is analyzed using the iterative procedure described in Section III.

B. Single-Phase Formulation With PMU Data

If PMUs are installed in the grid, expressions (1) and (2) should be jointly considered since \mathbf{v} , \mathbf{p} , \mathbf{q} , and θ are metered by the available PMUs. To that end, both systems of equations are partitioned as described above. Thus, the partitioning of (1)

yields (3) whereas the system of equations (2) is partitioned analogously giving rise to:

$$\begin{bmatrix} \theta_{\mathcal{M}} \\ \theta_{\hat{\mathcal{M}}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{\mathcal{M},\mathcal{M}} & \mathbf{X}_{\mathcal{M},\hat{\mathcal{M}}} \\ \mathbf{X}_{\hat{\mathcal{M}},\mathcal{M}} & \mathbf{X}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \end{bmatrix} \begin{bmatrix} \mathbf{p}_{\mathcal{M}} \\ \mathbf{p}_{\hat{\mathcal{M}}} \end{bmatrix} - \begin{bmatrix} \mathbf{R}_{\mathcal{M},\mathcal{M}} & \mathbf{R}_{\mathcal{M},\hat{\mathcal{M}}} \\ \mathbf{R}_{\hat{\mathcal{M}},\mathcal{M}} & \mathbf{R}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \end{bmatrix} \begin{bmatrix} \mathbf{q}_{\mathcal{M}} \\ \mathbf{q}_{\hat{\mathcal{M}}} \end{bmatrix} \quad (6)$$

Next, equations (3) and (6) are conveniently reordered to obtain the following system of equations:

$$\begin{bmatrix} \mathbf{R}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{X}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{0}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{0}_{\mathcal{M},\hat{\mathcal{M}}} \\ \mathbf{R}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} & \mathbf{X}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} & -\mathbf{I}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} & \mathbf{0}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \\ \mathbf{X}_{\mathcal{M},\hat{\mathcal{M}}} & -\mathbf{R}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{0}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{0}_{\mathcal{M},\hat{\mathcal{M}}} \\ \mathbf{X}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} & -\mathbf{R}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} & \mathbf{0}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} & -\mathbf{I}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \end{bmatrix} \begin{bmatrix} \mathbf{p}_{\hat{\mathcal{M}}} \\ \mathbf{q}_{\hat{\mathcal{M}}} \\ \mathbf{v}_{\hat{\mathcal{M}}} \\ \theta_{\hat{\mathcal{M}}} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_{\mathcal{M}} - \mathbf{R}_{\mathcal{M},\mathcal{M}}\mathbf{p}_{\mathcal{M}} - \mathbf{X}_{\mathcal{M},\mathcal{M}}\mathbf{q}_{\mathcal{M}} \\ -\mathbf{R}_{\hat{\mathcal{M}},\mathcal{M}}\mathbf{p}_{\mathcal{M}} - \mathbf{X}_{\hat{\mathcal{M}},\mathcal{M}}\mathbf{q}_{\mathcal{M}} \\ \theta_{\mathcal{M}} - \mathbf{X}_{\mathcal{M},\mathcal{M}}\mathbf{p}_{\mathcal{M}} + \mathbf{R}_{\mathcal{M},\mathcal{M}}\mathbf{q}_{\mathcal{M}} \\ -\mathbf{X}_{\hat{\mathcal{M}},\mathcal{M}}\mathbf{p}_{\mathcal{M}} + \mathbf{R}_{\hat{\mathcal{M}},\mathcal{M}}\mathbf{q}_{\mathcal{M}} \end{bmatrix} \quad (7)$$

Similar to (5), matrix \mathbf{H} can be obtained by transposing the matrix in the left-hand side of (7):

$$\mathbf{H} = \begin{bmatrix} \mathbf{R}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{R}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} & \mathbf{X}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{X}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \\ \mathbf{X}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{X}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} & -\mathbf{R}_{\mathcal{M},\hat{\mathcal{M}}} & -\mathbf{R}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \\ \mathbf{0}_{\mathcal{M},\hat{\mathcal{M}}} & -\mathbf{I}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} & \mathbf{0}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{0}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \\ \mathbf{0}_{\mathcal{M},\hat{\mathcal{M}}} & \mathbf{0}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} & \mathbf{0}_{\mathcal{M},\hat{\mathcal{M}}} & -\mathbf{I}_{\hat{\mathcal{M}},\hat{\mathcal{M}}} \end{bmatrix} \quad (8)$$

C. Multi-Phase Formulation

The \mathbf{H} matrices in (5) and (8) corresponding to the equivalent single-phase formulations can be readily extended to the general multi-phase setting. To that end, the dimensions of the vectors and matrices for unbalanced three-phase systems should be increased as follows: \mathbf{v} is the $3n$ -dimensional vector of voltage magnitudes, θ is the $3n$ -dimensional vector of voltage angles, \mathbf{p} is the $3n$ -dimensional vector of active power nodal injections, and \mathbf{q} is the $3n$ -dimensional vector of reactive power nodal injections. Analogously, \mathbf{R} and \mathbf{X} are real-valued $3n \times 3n$ -dimensional matrices [30]. Note that, for multi-phase systems with components with either 1, 2, or 3 phases, the dimensions of the above vectors and matrices should be adjusted accordingly. These modifications pave the way for the development of a novel algorithm to analyze and restore the observability of multi-phase distribution grids.

III. PROPOSED ALGORITHM

As sketched in the pseudocode below, the proposed algorithm, which is referred to as Algorithm 1, involves two phases, namely, *Analysis Phase*, and *Restoration Phase*. Note that the main steps of the algorithm are labeled at the left hand-side of the pseudo-codes. Both phases are described as in Algorithms 2 and 3.

In the *Analysis Phase* the maximum number of row vectors in matrix \mathbf{H} that are linearly independent are selected, so that a basis of the subspace of \mathbf{R}^l generated by the selected row vectors of \mathbf{H} is obtained. The selection process requires orthonormalizing the basis to calculate distances as explained next.

Algorithm 1: Proposed Algorithm

Input: The $m \times l$ matrix \mathbf{H} and parameter ϵ
 $N \leftarrow 1$;
while $N < m$ **do**
M1 $(N, \mathbf{s}, \mathcal{S}_N, \mathbf{A}) \leftarrow \text{analysis}(\mathbf{H})$;
 if $N < m$ **then**
M2 $\mathbf{H} \leftarrow \text{restoration}(N, \mathbf{s}, \mathbf{A}, \epsilon)$;
 else
M3 **stop**;
 end
end

Algorithm 2: Analysis Phase

Procedure analysis(\mathbf{H})
A1 $I \leftarrow 1$ and arbitrarily select a value for the row index k ;
 $s_1 \leftarrow k, \mathbf{w}_1 \leftarrow \mathbf{h}_k / \|\mathbf{h}_k\|, \mathcal{W}_1 \leftarrow \text{span}\{\mathbf{w}_1\}$;
A2 **for** $j = 1, \dots, m$ **do**
 $a(1, j) \leftarrow (\mathbf{w}_1 \bullet \mathbf{h}_j)$;
 $\text{pr}_{\mathcal{W}_1}(\mathbf{h}_j) \leftarrow a(1, j)\mathbf{w}_1$;
 $E_j^{(1)} \leftarrow \|\mathbf{h}_j - \text{pr}_{\mathcal{W}_1}(\mathbf{h}_j)\|$;
 end
 $e_1 \leftarrow \max_{j=1, \dots, m} E_j^{(1)}$;
A3 $I \leftarrow 2, s_2 \leftarrow \underset{j=1, \dots, m}{\text{argmax}} E_j^{(1)}$;
A4 Orthonormalize the vectors $\{\mathbf{w}_1, \mathbf{h}_{s_2}\}$ by Gram-Schmidt
 and obtain vector \mathbf{w}_2 ;
 $\mathcal{W}_2 \leftarrow \text{span}\{\mathbf{w}_1, \mathbf{w}_2\}, \mathcal{S}_2 \leftarrow \{\mathbf{w}_1, \mathbf{w}_2\}$;
A5 **for** $j = 1, \dots, m$ **do**
 $a(2, j) \leftarrow (\mathbf{w}_2 \bullet \mathbf{h}_j)$;
 $\text{pr}_{\mathcal{W}_2}(\mathbf{h}_j) \leftarrow a(1, j)\mathbf{w}_1 + a(2, j)\mathbf{w}_2$;
 $E_j^{(2)} \leftarrow \|\mathbf{h}_j - \text{pr}_{\mathcal{W}_2}(\mathbf{h}_j)\|$;
 end
 $e_2 \leftarrow \max_{j=1, \dots, m} E_j^{(2)}$;
A6 **while** $I < m$ & $I < l$ & $e_I < e_{I-1}$ **do**
A7 $I \leftarrow I + 1, s_I \leftarrow \underset{j=1, \dots, m}{\text{argmax}} E_j^{(I-1)}$;
A8 Orthonormalize the vectors $\{\mathbf{w}_1, \dots, \mathbf{w}_{I-1}, \mathbf{h}_{s_I}\}$ by
 Gram-Schmidt and obtain vector \mathbf{w}_I ;
 $\mathcal{W}_I \leftarrow \text{span}\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_I\}$;
 $\mathcal{S}_I \leftarrow \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_I\}$;
A9 **for** $j = 1, \dots, m$ **do**
 $a(I, j) \leftarrow (\mathbf{w}_I \bullet \mathbf{h}_j)$;
 $\text{pr}_{\mathcal{W}_I}(\mathbf{h}_j) \leftarrow a(1, j)\mathbf{w}_1 + \dots + a(I, j)\mathbf{w}_I$;
 $E_j^{(I)} \leftarrow \|\mathbf{h}_j - \text{pr}_{\mathcal{W}_I}(\mathbf{h}_j)\|$;
 end
 $e_I \leftarrow \max_{j=1, \dots, m} E_j^{(I)}$;
 end
A10 **if** $e_I \geq e_{I-1}$ **then**
 $N \leftarrow I - 1$;
 else
 $N \leftarrow I$;
 end
 return $N, \mathbf{s}, \mathcal{S}_N, \mathbf{A}$;

An arbitrary row vector in \mathbf{H} , \mathbf{h}_k , is first selected and normalized as $\mathbf{w}_1 = \mathbf{h}_k / \|\mathbf{h}_k\|$. The next vector selected is the farthest from the subspace generated by vector \mathbf{w}_1 , denoted by \mathcal{W}_1 . To identify which is the farthest vector, the distance of a generic row vector \mathbf{h}_j to the subspace \mathcal{W}_1 is calculated as

Algorithm 3: Restoration Phase

Procedure restoration($N, \mathbf{s}, \mathbf{A}, \epsilon$)
R1 $\mathbf{t} \leftarrow (1 : m)$ excluding \mathbf{s} ;
 $c(1 : N, 1 : (m - N)) \leftarrow a(1 : N, \mathbf{t})$;
R2 **for** $j = 1, \dots, (m - N)$ **do**
 $z_j \leftarrow \text{length}(\text{find}(\text{abs}(c(1 : N, j)) < \epsilon))$;
 end
 $z_{\max} \leftarrow \underset{j=1, \dots, m-N}{\text{argmax}} z_j$;
R3 Add the bus corresponding to z_{\max} and all its phases to
 the subset of metered buses \mathcal{M} ;
R4 Update matrix \mathbf{H} and the corresponding values of m and
 l accordingly;
 return \mathbf{H} ;

$d(\mathbf{h}_j, \mathcal{W}_1) = \|\mathbf{h}_j - \text{pr}_{\mathcal{W}_1}(\mathbf{h}_j)\|$, where $\text{pr}_{\mathcal{W}_1}(\mathbf{h}_j) = (\mathbf{w}_1 \bullet \mathbf{h}_j)\mathbf{w}_1$. That is, the distance of vector \mathbf{h}_j to subspace \mathcal{W}_1 requires the projection of this vector on the subspace, which, in turn, is defined through the orthonormal basis of \mathcal{W}_1 . Once the farthest vector is found, the subspace, \mathcal{W}_2 , generated by this newly identified vector and vector \mathbf{w}_1 is considered. Similarly, the farthest vector from this new subspace is selected by first orthonormalizing the basis formed by vectors \mathbf{w}_1 and \mathbf{w}_2 , and then projecting the other row vectors on \mathcal{W}_2 so that the corresponding distances to this subspace \mathcal{W}_2 are calculated. This process is repeated by selecting new vectors. Note that the maximum distance decreases iteratively. The process stops when either the matrix is full rank or the maximum distance is greater than that of the previous iteration. For a theoretical justification of this stopping criterion, the interested reader is referred to *Theorem 1* in the Appendix.

By considering the farthest row vector at each iteration, the row dependencies are efficiently identified. If the distance of the farthest row vector to a subspace is 0, this vector is linearly dependent on the vectors already included in the subspace and so are the remaining row vectors outside the subspace. Note that a distance equal to 0 can only be attained once all linearly independent vectors have been identified. As backed by *Theorem 1*, this situation occurs when the maximum distance at one iteration is greater than that of the previous iteration. Since this stopping criterion does not require setting an arbitrary tolerance to distinguish between zero and nonzero values, we claim that the proposed procedure is numerically stable as opposed to existing ones.

The output of the *Analysis Phase* is:

- The rank of matrix \mathbf{H} , i.e., N .
- Vector \mathbf{s} , whose entries are integers indexing the rows in \mathbf{H} that are linearly independent.
- The basis \mathcal{S}_N , which is a set formed by the orthonormal vectors obtained throughout the iterative process.
- Matrix $\mathbf{A} = (a(i, j))_{\substack{1 \leq i \leq N \\ 1 \leq j \leq m}}$, whose columns are the coordinates of the rows of \mathbf{H} expressed in the basis \mathcal{S}_N .

If $N = m$, i.e., if the number of components in \mathbf{s} is equal to m , the system is observable. Otherwise, if $N < m$, matrix \mathbf{A} is the starting point to restore grid observability.

As for the proposed *Restoration Phase*, unlike [22], where row dependencies were not exploited, observability is restored by analyzing the columns of matrix \mathbf{C} , where

$\mathbf{C} = (c(i, j))_{\substack{1 \leq i \leq N \\ 1 \leq j \leq (m-N)}}$ is a submatrix of \mathbf{A} . The entries of vector \mathbf{t} are the indexes of the rows in \mathbf{H} that are linearly dependent, so that these entries identify the columns in \mathbf{A} that are selected to be included in matrix \mathbf{C} . Note that each row in \mathbf{C} is associated with a vector in the set \mathcal{S}_N while each column corresponds to a linearly dependent row in \mathbf{H} expressed in the basis \mathcal{S}_N . In particular, the original criterion proposed in this work to decide which is the best candidate bus to be metered consists in finding the column vector of \mathbf{C} with the largest number of coordinates in \mathcal{S}_N less than ϵ . Once the algorithm identifies a bus as the best candidate, all its phases are included in the subset of metered buses \mathcal{M} . In case that none of the column vectors of \mathbf{C} satisfies the ϵ -criterion, any bus associated with the columns of \mathbf{C} can be arbitrarily selected to be included in \mathcal{M} . It is worth emphasizing that the selection of the bus to be added to \mathcal{M} takes into consideration that practical metering devices in distribution networks are multi-phase. It is remarkable that observability is successfully restored regardless of the selected value for ϵ , which only impacts on the selection of buses, and, hence, on the number of iterations in the *Restoration Phase*.

The main differences of the analysis procedure with respect to the algorithm presented in [22] lie in Step A10 and in the *Restoration Phase*. As per Step A10, the *Analysis Phase* terminates when the proposed stopping criterion $e_l \geq e_{l-1}$ is satisfied. Note that this condition was not used in [22]. It should be also emphasized that this work improves upon [22] by providing the aforementioned theoretical justification for this novel criterion, namely, *Theorem 1* in the Appendix.

It should be highlighted that the proposed algorithm is related to a QR factorization, but the rows of matrix \mathbf{H} are rearranged in a different order by using an optimal orthonormalization procedure. The proposed technique is based on the Gram-Schmidt orthonormalization of a selected basis using a greedy strategy over the space generated by the rows of matrix \mathbf{H} . The algorithm is greedy as it makes the best choice at each step [32]. Note that rows are not sequentially chosen as done under the classical Gram-Schmidt orthonormalization. Rather, at each iteration, the row with the largest distance to the corresponding subspace is selected. Interestingly, despite its greedy nature, no approximation is involved.

Finally, it should be noted that the computational complexity of the proposed algorithm is $O(N^3)$.

IV. CASE STUDIES

To illustrate the effective performance of the proposed algorithm for analyzing and restoring grid observability, several benchmarks are examined under balanced and unbalanced conditions, assuming that multi-phase metering devices are available for the metered buses. First, simulations for the IEEE 4-bus feeder [33] are run for didactical purposes. Then, observability is checked and restored for the IEEE 13-bus feeder [34], which is a medium-voltage multi-phase distribution grid. Finally, the computational efficiency of the algorithm is illustrated on the European 906-bus low-voltage feeder [10]. For all instances, ϵ is set to 10^{-7} according to the order of magnitude of the involved numerical values.

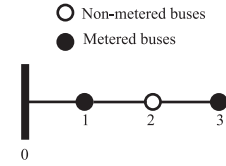


Fig. 1. IEEE 4-bus feeder – One-line diagram for metering scenario 1ϕ -a.

Simulations were run using MATLAB [35] on a workstation SIE Ladon with two Intel Xeon Gold 6248R processors at 3 GHz and 768 GB of RAM.

A. IEEE 4-Bus Feeder

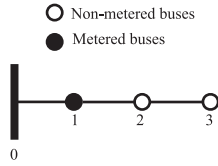
The first case study considers a four-wire line configuration and a substation transformer rated 12.47 kV/4.16 kV and 6 MVA. For expository purposes, different operating conditions and metering scenarios are used to analyze the observability of this grid.

1) *Balanced Operation*: Balanced conditions are examined first. Thus, the single-phase (1ϕ) formulations presented in Sections II-A and II-B are used. Fig. 1 shows the one-line diagram for a representative metering scenario wherein smart meters are located at buses 1 and 3. Under this scenario, which is referred to as scenario 1ϕ -a, the magnitudes \mathbf{p}_1 , \mathbf{q}_1 , \mathbf{v}_1 , \mathbf{p}_3 , \mathbf{q}_3 , and \mathbf{v}_3 are metered and the corresponding matrix \mathbf{H} is presented in (9). Note that labels \mathbf{p}_2 , \mathbf{q}_2 , and \mathbf{v}_2 in (9) denote the correspondence between the non-metered magnitudes and the rows of this matrix.

$$\mathbf{H} = \begin{matrix} \mathbf{p}_2 \\ \mathbf{q}_2 \\ \mathbf{v}_2 \end{matrix} \begin{bmatrix} 0.0045 & 0.0145 & 0.0145 \\ 0.0092 & 0.0692 & 0.0692 \\ 0.0000 & 0.0000 & -1.0000 \end{bmatrix} \quad (9)$$

Considering the rows of matrix \mathbf{H} in (9), the *Analysis Phase* is applied as follows. In Step A1, the first row of \mathbf{H} is arbitrarily selected. Once this row is normalized, vector $\mathbf{w}_1 = [0.2143, 0.6907, 0.6907]$ is obtained. In Step A2, the projections of rows 2 and 3 on the space $\mathcal{W}_1 = \text{span}\{\mathbf{w}_1\}$ are calculated. Then, the distances of each projection are computed as $E_2^{(1)} = 0.0120$ and $E_3^{(1)} = 0.7232$ whereas $e_1 = \max\{0.0120, 0.7232\} = 0.7232$. In Step A3, the third row of \mathbf{H} is selected since it yields the largest distance. Applying the Gram-Schmidt method, vector $\mathbf{w}_2 = [0.2047, 0.6596, -0.7232]$ is obtained in Step A4. In Step A5, the projection of row 2 of matrix \mathbf{H} on the space $\mathcal{W}_2 = \text{span}\{\mathbf{w}_1, \mathbf{w}_2\}$ is calculated and $e_2 = 0.0117$. Once Step A6 is applied, vector $\mathbf{w}_3 = [-0.9551, 0.2964, 0.0000]$ is calculated, the three rows of matrix \mathbf{H} being all linearly independent vectors ($I = m = 3$). As $N = 3$, the algorithm stops in Step M3. Since \mathbf{p}_2 , \mathbf{q}_2 , and \mathbf{v}_2 are uniquely determined, the grid is deemed as observable.

It should be emphasized that, for this benchmark, non-observability is identified for any single-phase scenario in which bus 1 is non-metered, no matter whether the remaining buses are metered or not. Fig. 2 shows a different scenario, namely, scenario 1ϕ -b, in which the magnitudes \mathbf{p}_2 , \mathbf{q}_2 , \mathbf{v}_2 , \mathbf{p}_3 , \mathbf{q}_3 , and \mathbf{v}_3 are non-metered. For this scenario, matrix \mathbf{H} is given in (10).

Fig. 2. IEEE 4-bus feeder – One-line diagram for metering scenario 1 ϕ -b.

$$\mathbf{H} = \begin{matrix} \mathbf{p}_2 \\ \mathbf{p}_3 \\ \mathbf{q}_2 \\ \mathbf{q}_3 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \end{matrix} \begin{bmatrix} 0.0045 & 0.0145 & 0.0145 \\ 0.0045 & 0.0145 & 0.0647 \\ 0.0092 & 0.0692 & 0.0692 \\ 0.0092 & 0.0692 & 0.1721 \\ 0.0000 & -1.0000 & 0.0000 \\ 0.0000 & 0.0000 & -1.0000 \end{bmatrix} \quad (10)$$

At the third iteration of the *Analysis Phase* ($I = 3$), the largest distance is equal to $e_3 = 1.415 \times 10^{-16}$. Moreover, three linearly independent vectors are obtained and, hence, matrix \mathbf{H} has no full rank. The output of the *Analysis Phase* at this iteration is:

- The rank of \mathbf{H} , N , which is equal to 3, i.e., $N = l$.
- Vector $\mathbf{s} = (1, 5, 6)$, whose entries are the indices of the rows of matrix \mathbf{H} that are linearly independent.
- The basis \mathcal{S}_3 given in (11), whose rows are the orthonormal vectors that are obtained along the algorithm.

$$\mathcal{S}_3 = \begin{bmatrix} 0.2143 & 0.6907 & 0.6907 \\ 0.2047 & -0.7232 & 0.6596 \\ 0.9551 & 0.0000 & -0.2964 \end{bmatrix} \quad (11)$$

Once Step R1 is applied, vector $\mathbf{t} = (2, 3, 4)$ is obtained thus yielding matrix \mathbf{C} in (12). This matrix includes the coordinates of the linearly dependent rows in \mathbf{H} expressed in the basis \mathcal{S}_3 .

$$\mathbf{C} = \begin{matrix} \mathbf{p}_2 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \end{matrix} \begin{matrix} \mathbf{p}_3 & \mathbf{q}_2 & \mathbf{q}_3 \end{matrix} \begin{bmatrix} 0.0557 & 0.0976 & 0.1686 \\ 0.0331 & -0.0025 & 0.0654 \\ -0.0149 & -0.0117 & -0.0422 \end{bmatrix} \quad (12)$$

Thus, under scenario 1 ϕ -b, the grid is deemed as non-observable since matrix \mathbf{H} in (10) is rank-deficient.

To restore observability, matrix \mathbf{C} in (12) is explored. Each vector in \mathcal{S}_3 , which is a row in the matrix given in (11), is associated with a linearly independent row in \mathbf{H} . Then, as denoted by the labels in \mathbf{C} , the rows in this matrix correspond to the linearly independent magnitudes in (10), i.e., \mathbf{p}_2 , \mathbf{v}_2 , and \mathbf{v}_3 , whereas the columns in \mathbf{C} are related to the linearly dependent magnitudes in \mathbf{H} , namely, \mathbf{p}_3 , \mathbf{q}_2 , and \mathbf{q}_3 . In this case, the ϵ -criterion in Step R2 is not met since all the values in \mathbf{C} are larger than 10^{-7} . That is, it is indifferent to select either bus 2 or bus 3. Let us consider that a smart meter can be placed at bus 2. Adding bus 2 to the subset of metered buses results in a new matrix \mathbf{H} with all rows linearly independent, thereby identifying the system as observable. Analogously, if a smart meter is placed at bus 3, the system would also be deemed as observable.

This benchmark is also useful to illustrate the single-phase analysis when PMUs are available. To that end, scenario 1 ϕ -c is considered, wherein buses 1 and 3 are non-metered. For this scenario, the proposed algorithm is initially fed with the matrix

\mathbf{H} given in (13). As a result, matrix \mathbf{H} is identified as rank-deficient and six linearly independent vectors are obtained.

$$\mathbf{H} = \begin{matrix} \mathbf{p}_1 \\ \mathbf{p}_3 \\ \mathbf{q}_1 \\ \mathbf{q}_3 \\ \mathbf{v}_1 \\ \mathbf{v}_3 \\ \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_3 \end{matrix} \begin{bmatrix} 0.0045 & 0.0045 & 0.0045 & 0.0092 & 0.0092 & 0.0092 \\ 0.0145 & 0.0045 & 0.0647 & 0.0692 & 0.0092 & 0.1721 \\ 0.0092 & 0.0092 & 0.0092 & -0.0045 & -0.0045 & -0.0045 \\ 0.0692 & 0.0092 & 0.1721 & -0.0145 & -0.0045 & -0.0647 \\ 0.0000 & -1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & -1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & -1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & -1.0000 \end{bmatrix} \quad (13)$$

To restore observability, the application of Step R1 of the algorithm yields the matrix \mathbf{C} given in (14).

$$\mathbf{C} = \begin{matrix} \mathbf{p}_1 \\ \mathbf{v}_1 \\ \mathbf{v}_3 \\ \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_3 \\ \mathbf{q}_3 \end{matrix} \begin{matrix} \mathbf{p}_3 & \mathbf{q}_1 \end{matrix} \begin{bmatrix} 0.1512 & 0.0000 \\ 0.0350 & -0.0095 \\ -0.0249 & -0.0102 \\ 0.0846 & 0.0021 \\ -0.0822 & 0.0033 \\ -0.0174 & 0.0102 \end{bmatrix} \quad (14)$$

According to the ϵ -criterion, the algorithm selects the second column of matrix \mathbf{C} in (14) since it features the maximum number of entries with an absolute value less than 10^{-7} , namely one. Note that this column is associated with the reactive power injected at bus 1. Thus, bus 1 is added to the subset of metered buses and matrix \mathbf{H} is updated accordingly. Since the new matrix \mathbf{H} is of full rank, observability can be restored if a PMU is placed at bus 1.

2) *Unbalanced Operation*: The use of the multi-phase formulation presented in Section II-C is illustrated with a scenario considering unbalanced operation. Under this scenario, it is assumed that no smart meter is placed at bus 2, i.e., the magnitudes \mathbf{p}_2 , \mathbf{q}_2 , and \mathbf{v}_2 are not metered in all phases.

For this case, matrix \mathbf{H} is given in (15), shown at the bottom of the next page. The application of the proposed approach to this multi-phase instance gives rise to the following outcome:

- The rank of \mathbf{H} is 9.
- Vector $\mathbf{s} = (1, 9, 8, 7, 6, 5, 4, 2, 3)$.
- The basis \mathcal{S}_9 in (16), shown at the bottom of the next page.

Thus, the IEEE 4-bus feeder for this particular multi-phase scenario is deemed as observable since the unknown magnitudes \mathbf{p}_2 , \mathbf{q}_2 , and \mathbf{v}_2 in all phases can be uniquely determined.

B. IEEE 13-Bus Feeder

The IEEE 13-bus feeder has laterals with one, two, and three phases. The delta-wye substation transformer is rated 115 kV/4.16 kV and 5 kVA.

Considering balanced operation and smart meters, the algorithm has been first applied to all possible instances wherein a single bus is non-metered. As done in [36], the single-phase equivalent feeder is obtained by assuming that distribution lines are symmetrically transposed and every bus is supplied by all three phases. From this study, it can be concluded that

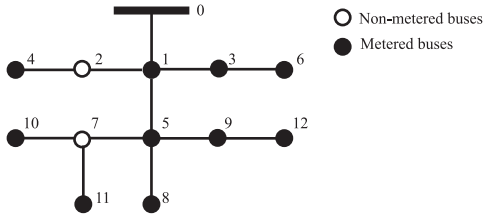


Fig. 3. IEEE 13-bus feeder – One-line diagram for metering scenario 1φ-a.

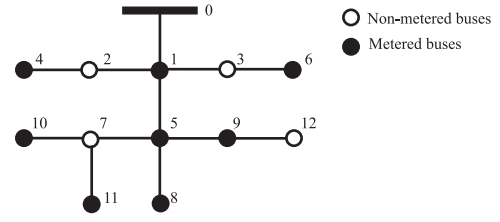


Fig. 4. IEEE 13-bus feeder – One-line diagram for metering scenario 1φ-b.

the IEEE 13-bus feeder is unobservable when buses 1, 5, and 8 are the only non-metered buses.

Next, the metering scenario 1φ-a depicted in Fig. 3 is evaluated. For this scenario, it is assumed that magnitudes \mathbf{p}_i , \mathbf{q}_i , and \mathbf{v}_i are known for all buses except buses 2 and 7. The *Analysis Phase* requires six iterations to find six linearly independent rows in the corresponding matrix \mathbf{H} , with the largest distance being $e_6 = 1.0281 \times 10^{-2}$. Thus, the grid is identified as observable under this scenario since matrix \mathbf{H} is of full rank ($N = m = 6$), and, hence, \mathbf{p}_2 , \mathbf{q}_2 , \mathbf{v}_2 , \mathbf{p}_7 , \mathbf{q}_7 , and \mathbf{v}_7 can be uniquely determined.

The proposed approach has also been applied to the metering scenario 1φ-b shown in Fig. 4, with \mathbf{p}_i , \mathbf{q}_i , and \mathbf{v}_i , for $i = 2, 3, 7, 12$, being non-metered. For this scenario, the stopping criterion is satisfied for $I = 11$, with $e_{11} = 2.9302 \times 10^{-5} > e_{10} = 1.669 \times 10^{-12}$ and $N = 10$. Thus, the grid is deemed as non-observable since only 10 linearly independent rows are found in \mathbf{H} , which is a rank-deficient matrix.

To restore observability, matrix \mathbf{C} in (17) is computed. Note that the rows/columns in \mathbf{C} are associated with the linearly independent/dependent rows in \mathbf{H} .

The second column of \mathbf{C} in (17), which is associated with the active power injected at bus 7, features the largest number of zeroes. Thus, bus 7 is added to the subset of metered buses and matrix \mathbf{H} is updated accordingly. For this new \mathbf{H} , the grid

is still identified as non-observable. As a consequence, bus 3 is added to the subset of metered buses in the next iteration to eventually restore grid observability.

$$\mathbf{C} = \begin{matrix} & \mathbf{p}_3 & \mathbf{p}_7 \\ \mathbf{p}_2 & 0.4111 & 0.6131 \\ \mathbf{v}_3 & -0.0713 & 0.0454 \\ \mathbf{v}_7 & -0.0128 & -0.1128 \\ \mathbf{v}_{12} & -0.0138 & -0.0763 \\ \mathbf{v}_2 & 0.0555 & 0.1434 \\ \mathbf{q}_7 & 0.0353 & 0.2487 \\ \mathbf{q}_3 & 0.0853 & -0.0049 \\ \mathbf{q}_{12} & -0.0021 & -0.0302 \\ \mathbf{q}_2 & -0.0125 & 0.0000 \\ \mathbf{p}_{12} & 0.0000 & 0.0000 \end{matrix} \quad (17)$$

Moreover, a new scenario is generated assuming that PMUs are available at all buses except for buses 2, 3, 7, and 12, for which the corresponding magnitudes \mathbf{p}_i , \mathbf{q}_i , \mathbf{v}_i , and θ_i are unknown. After 16 iterations of the analysis procedure, the proposed approach determines that this matrix has full rank with \mathbf{p}_i , \mathbf{q}_i , \mathbf{v}_i , and θ_i being uniquely determined for $i = 2, 3, 7, 12$. Therefore, the grid is observable for this scenario.

$$\mathbf{H} = \begin{matrix} \mathbf{p}_2^a \\ \mathbf{p}_2^b \\ \mathbf{p}_2^c \\ \mathbf{q}_2^a \\ \mathbf{q}_2^b \\ \mathbf{q}_2^c \\ \mathbf{v}_2^a \\ \mathbf{v}_2^b \\ \mathbf{v}_2^c \end{matrix} \begin{bmatrix} 0.0067 & 0.0023 & 0.0022 & 0.0167 & 0.0023 & 0.0022 & 0.0167 & 0.0023 & 0.0022 \\ 0.0023 & 0.0068 & 0.0023 & 0.0023 & 0.0168 & 0.0023 & 0.0023 & 0.0168 & 0.0023 \\ 0.0022 & 0.0023 & 0.0067 & 0.0022 & 0.0023 & 0.0167 & 0.0022 & 0.0023 & 0.0167 \\ 0.0158 & 0.0073 & 0.0056 & 0.0758 & 0.0073 & 0.0056 & 0.0758 & 0.0073 & 0.0056 \\ 0.0073 & 0.0153 & 0.0062 & 0.0073 & 0.0753 & 0.0062 & 0.0073 & 0.0753 & 0.0062 \\ 0.0056 & 0.0062 & 0.0156 & 0.0056 & 0.0062 & 0.0756 & 0.0056 & 0.0062 & 0.0756 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & -1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & -1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & -1.0000 \end{bmatrix} \quad (15)$$

$$\mathbf{S}_9 = \begin{bmatrix} 0.2663 & 0.0914 & 0.0874 & 0.6637 & 0.0914 & 0.0874 & 0.6637 & 0.0914 & 0.0874 \\ 0.0234 & 0.0080 & 0.0077 & 0.0583 & 0.0080 & 0.0077 & 0.0583 & 0.0080 & -0.9962 \\ 0.0246 & 0.0085 & 0.0081 & 0.0614 & 0.0085 & 0.0081 & 0.0614 & -0.9958 & 0.0000 \\ 0.2417 & 0.0831 & 0.0794 & 0.6024 & 0.0831 & 0.0794 & -0.7432 & 0.0000 & 0.0000 \\ -0.0187 & 0.0501 & 0.1758 & -0.1571 & 0.0501 & 0.9690 & 0.0000 & 0.0000 & 0.0000 \\ -0.0057 & 0.1627 & 0.0240 & -0.1457 & 0.9717 & -0.0868 & 0.0000 & 0.0000 & 0.0000 \\ -0.8931 & -0.1090 & -0.1893 & 0.3766 & 0.0813 & 0.0796 & 0.0000 & 0.0000 & 0.0000 \\ -0.1823 & 0.9558 & 0.1223 & -0.0407 & -0.1767 & -0.0726 & 0.0000 & 0.0000 & 0.0000 \\ -0.1958 & -0.1735 & 0.9506 & 0.0006 & -0.0104 & -0.1666 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix} \quad (16)$$

TABLE I
IEEE 13-BUS FEEDER – RESULTS

Single-phase with smart meters				
Number of smart meters	Non-metered buses	Number of linearly dependent rows in \mathbf{H}	Computing time (s)	Buses to restore observability
10	2, 9	0	0.0022	–
8	2, 3, 7, 9	2	0.0036	7, 3
7	4, 6, 7, 8, 11	5	0.0037	8, 7, 6
Single-phase with PMUs				
Number of PMUs	Non-metered buses	Number of linearly dependent rows in \mathbf{H}	Computing time (s)	Buses to restore observability
8	2, 3, 7, 9	0	0.0401	–
7	4, 6, 7, 8, 11	2	0.0053	7
Multi-phase with smart meters				
Number of smart meters	Non-metered buses-phases	Number of linearly dependent rows in \mathbf{H}	Computing time (s)	Buses-phases to restore observability
10	2-b, 2-c, 7-a, 7-c	0	0.0039	–
9	3-a, 3-b, 3-c, 9-a, 9-b, 9-c, 11-a	0	0.1027	–
9	3-a, 3-b, 3-c, 9-a, 9-b, 9-c, 12-a, 12-b, 12-c	9	0.0202	9-a, 9-b, 9-c
Multi-phase with PMUs				
Number of PMUs	Non-metered buses-phases	Number of linearly dependent rows in \mathbf{H}	Computing time (s)	Buses-phases to restore observability
10	3-a, 3-b, 3-c, 9-a, 9-b, 9-c	0	0.0084	–
7	1-a, 1-b, 1-c, 2-b, 2-c, 7-a, 7-c, 9-a, 9-b, 9-c, 12-a, 12-b, 12-c	6	0.0542	9-a, 9-b, 9-c

Additional representative metering scenarios with either smart meters or PMUs have been specifically designed to further illustrate the performance of the proposed approach. In Table I, the results for the single- and multi-phase models are summarized. It should be noted that the computing times reported in this table correspond to the observability analysis, as observability restoration involves the installation of extra devices, which is to be decided in the future.

As evidenced in Table I, the grid is observable for scenarios in which the percentage of non-metered buses(-phases) ranges between 11% and 33%. Moreover, the percentage of non-metered buses(-phases) that should be metered to restore observability ranges between 20% and 60%. Computing times are similar in all scenarios, ranging between $O(10^{-3})$ s and $O(10^{-1})$ s.

Finally, in order to validate the proposed methodology, all metering scenarios considered in [24] have been analyzed under balanced conditions and identical results are obtained, as expected.

C. European 906-Bus Low-Voltage Feeder

In the European 906-bus low-voltage feeder (ELV), all laterals have three phases and the substation transformer, rated 11 kV/416 V and 800 kVA, has a delta-wye configuration. For this benchmark, both single- and multi-phase models are analyzed. For the sake of simplicity, the metering scenarios are randomly generated including only smart meters or PMUs.

Table II lists the size of matrix \mathbf{H} and the required computing time per iteration for metering scenarios with full observability. Note that the computational burden grows with the matrix size.

TABLE II
ELV – RESULTS FOR OBSERVABLE INSTANCES

Size of \mathbf{H}	Time per iteration (s)
273×905	0.16
364×1810	1.55
408×2715	3.92
544×5430	19.00
1088×5430	72.99

Table III summarizes the results for metering scenarios requiring observability restoration. It should be noted that scenarios 1–4 are associated with the single-phase formulation while scenario 5 corresponds to the multi-phase model.

In scenario 1, with 45 non-metered buses (5% of the total number of buses), the algorithm finds that 26 rows of \mathbf{H} are linearly dependent and that 25 buses, i.e., 56% of the non-metered buses, are required to restore observability. In scenario 2, with the same number of non-metered buses as scenario 1 but with PMUs, the algorithm finds that only 2 rows of \mathbf{H} are linearly dependent and that one bus (50% of the non-metered buses) is required to restore observability. In scenario 3, for which the number of non-metered buses is increased to 181 (20% of the total number of buses), the algorithm finds that 166 rows of \mathbf{H} are linearly dependent and that 147 buses, i.e., 81% of the non-metered buses, are required to restore observability. In scenario 4, with the same number of non-metered buses as scenario 3 but with PMUs, the algorithm finds that 26 rows of \mathbf{H} are linearly dependent and that 11 buses (6% of the non-metered buses) are required to restore observability. Finally, in the multi-phase scenario 5 with 91 non-metered buses (10% of the total number of buses), the

TABLE III
ELV – RESULTS FOR NON-OBSERVABLE INSTANCES

Scenario #	Number of non-metered buses	Size of \mathbf{H}	Number of linearly dependent rows in \mathbf{H}	Computing time (s)	Number of buses to restore observability
1	45	135×905	26	4	25
2	45	180×1810	2	24	1
3	181	543×905	166	203	147
4	181	724×1810	26	6404	11
5	91	819×2715	192	9005	32

algorithm finds that 192 rows of \mathbf{H} are linearly dependent and that 32 buses (35% of the non-metered buses) are required to restore observability.

From the results summarized in Table III, it can be concluded that the number of non-metered buses required to restore observability is scenario-dependent. It is also worth mentioning that the percentage of metered buses is not correlated with the percentage of buses in which metering devices are required to restore observability. Additionally, it should be emphasized that, for the scenarios under consideration, the size of the matrices \mathbf{H} ranges between 135×905 and 819×2715 , while the corresponding computing times are between 4 s and 9005 s.

Finally, the application of the stopping criterion in Step A10 is illustrated using scenario 1 in Table III. This instance, which is not observable, has 109 linearly independent rows in \mathbf{H} . The maximum distances e_I at iterations $I = 108, 109,$ and 110 are, respectively, 4×10^{-4} , 2×10^{-11} , and 6×10^{-4} . It should be noted that e_{109} is not exactly zero due to rounding errors. As can be noted, the maximum distance decreases down to a value near zero at iteration $I = 109$, whereas it increases at iteration $I = 110$. Additionally, according to *Theorem 1* in the Appendix, $pr_{S_{109}} \mathbf{h}_j = \mathbf{h}_j$, $j = 1, \dots, 135$, which in turn implies that $E_j^{(109)} = \|\mathbf{h}_j - pr_{S_{109}} \mathbf{h}_j\| = 0$ and $\mathbf{h}_j \in \mathcal{W}_{109}$, $j = 1, \dots, 135$. Thus, the maximum number of linearly independent rows is $N = 109$. In other words, as the maximum distance increases at iteration $I = 110$, the 110^{th} vector is linearly dependent on previous vectors. Thus, at iteration $I = 110$ the *Analysis Phase* of the algorithm terminates with $N = 109$ providing the information required to start the *Restoration Phase*.

The stopping criterion proposed in [22], i.e., $e_I \leq 10^{-7}$, has been alternatively considered, yielding identical results for this scenario. However, if a different tolerance is selected, the algorithm may require a different number of iterations and even provide an incorrect output. In particular, for $e_I \leq 10^{-15}$, the *Analysis Phase* terminates for $N = 135$ with matrix \mathbf{H} being of full rank, thereby identifying the grid as observable. This behavior is depicted in Fig. 5, in which the logarithm with base 10 of the largest distance e_I is plotted for each iteration I . From this figure, it is inferred that the vectors between iterations 110 and 135 are included in the basis, thus requiring 25 additional iterations as compared to the proposed stopping criterion and leading to a wrong observability analysis.

Therefore, checking if the maximum distance decreases down to a minimum value and then stops decreasing at the

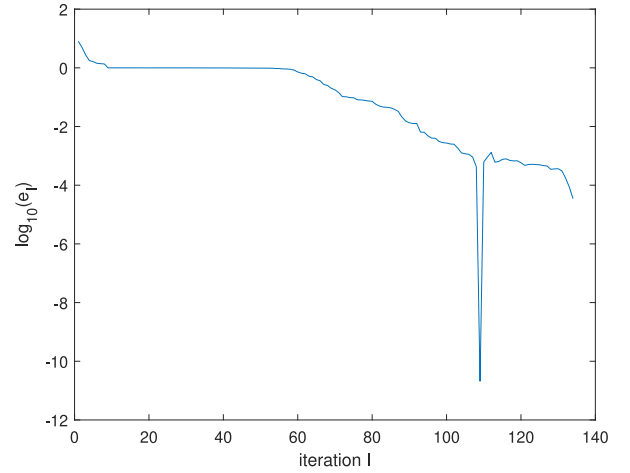


Fig. 5. ELV – Evolution of \log_{10} of the maximum distance along the iterative process using an alternative stopping criterion.

next iteration is an effective stopping criterion. This finding is theoretically backed by *Theorem 1* in the Appendix.

V. CONCLUSION

This paper has presented a robust greedy algorithm for analyzing and restoring observability in distribution grids. Unlike previous works, the algorithm proposed to determine the rank of the system matrix is numerically stable. The observability analysis phase relies on the iterative determination of the number of linearly independent rows in the matrix representing system operation. If needed, observability restoration is subsequently incorporated as a relevant distinctive feature while using an effective practical criterion for selecting the best candidate buses to be metered. Moreover, the robustness of the algorithm is guaranteed by a mathematically sound stopping criterion that does not require setting any tolerance to distinguish between zero and nonzero values. Thus, the proposed numerical algorithm overcomes the limitations of existing approaches by avoiding divisions by values close to zero.

Furthermore, the proposed algorithm is suitable for three-phase and multi-phase distribution grids under both balanced and unbalanced conditions. Results from several case studies based on medium- and low-voltage distribution grids with up to 906 buses demonstrate the effective performance and the practical applicability of the proposed algorithm.

Further research will consider the generalization of the proposed formulation to accommodate other metering devices.

Additionally, the algorithm will be extended to identify observable islands in case of grid unobservability, and to consider a more accurate network model.

APPENDIX

The stopping criterion in Step A10 is justified by the following theorem, which is a direct consequence of the orthogonal projection theorem.

Theorem 1: Let us consider the vectorial space \mathcal{R}^l , endowed with the usual scalar product and its subspace \mathcal{W}_N , such that $\dim(\mathcal{W}_N) = N$, with an orthonormal basis $\mathcal{S}_N = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_N\}$.

If $\mathbf{w}_{N+1} \in \mathcal{W}_N$ and $\mathcal{S}_{N+1} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_N, \mathbf{w}_{N+1}\}$, for any $\mathbf{u} \in \mathcal{W}_N$, we define

$$\begin{aligned} pr_{\mathcal{S}_N} \mathbf{u} &= (\mathbf{w}_1 \bullet \mathbf{u})\mathbf{w}_1 + (\mathbf{w}_2 \bullet \mathbf{u})\mathbf{w}_2 + \dots + \\ &\quad (\mathbf{w}_N \bullet \mathbf{u})\mathbf{w}_N, \\ pr_{\mathcal{S}_{N+1}} \mathbf{u} &= (\mathbf{w}_1 \bullet \mathbf{u})\mathbf{w}_1 + (\mathbf{w}_2 \bullet \mathbf{u})\mathbf{w}_2 + \dots + \\ &\quad (\mathbf{w}_N \bullet \mathbf{u})\mathbf{w}_N + (\mathbf{w}_{N+1} \bullet \mathbf{u})\mathbf{w}_{N+1}, \end{aligned}$$

and, hence:

$$\|\mathbf{u} - pr_{\mathcal{S}_{N+1}} \mathbf{u}\| \geq \|\mathbf{u} - pr_{\mathcal{S}_N} \mathbf{u}\|.$$

Proof: Let \mathcal{S}_N be an orthonormal basis with $pr_{\mathcal{S}_N} \mathbf{u} = pr_{\mathcal{W}_N} \mathbf{u} = \mathbf{u}$. Consequently:

$$\|\mathbf{u} - pr_{\mathcal{S}_N} \mathbf{u}\| = 0.$$

Subtracting the distances, the following inequality holds:

$$\|\mathbf{u} - pr_{\mathcal{S}_{N+1}} \mathbf{u}\| - \|\mathbf{u} - pr_{\mathcal{S}_N} \mathbf{u}\| = \|-(\mathbf{w}_{N+1} \bullet \mathbf{u})\mathbf{w}_{N+1}\| \geq 0.$$

Therefore,

$$\|\mathbf{u} - pr_{\mathcal{S}_{N+1}} \mathbf{u}\| \geq \|\mathbf{u} - pr_{\mathcal{S}_N} \mathbf{u}\|. \quad \blacksquare$$

REFERENCES

- [1] A. Primadianto and C.-N. Lu, "A review on distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3875–3883, Sep. 2017.
- [2] M. E. Baran, "Challenges in state estimation on distribution systems," in *Proc. Power Eng. Soc. Summer Meeting*, Vancouver, BC, Canada, 2001, pp. 429–433.
- [3] B. R. Pokhrel, N. Karthikeyan, B. Bak-Jensen, J. R. Pillai, S. M. Mazhari, and C. Y. Chung, "An intelligent approach to observability of distribution networks," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Portland, OR, USA, 2018, pp. 1–5.
- [4] M. U. Hashmi, A. Koirala, R. Lundholm, H. Ergun, and D. Van Hertem, "Evaluation of voltage magnitude based unbalance metric for low voltage distribution networks," 2022, *arXiv:2202.09669*.
- [5] R. S. Biswas, B. Azimian, and A. Pal, "A micro-PMU placement scheme for distribution systems considering practical constraints," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Montreal, QC, Canada, 2020, pp. 1–5.
- [6] S. Bhela, V. Kekatos, and S. Veeramachaneni, "Enhancing observability in distribution grids using smart meter data," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5953–5961, Nov. 2018.
- [7] S. Lin and H. Zhu, "Enhancing the spatio-temporal observability of grid-edge resources in distribution grids," *IEEE Trans. Smart Grid*, vol. 12, no. 6, pp. 5434–5443, Nov. 2021.
- [8] G. Cavarro, R. Arghandeh, K. Poolla, and A. von Meier, "Data-driven approach for distribution network topology detection," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Denver, CO, USA, 2015, pp. 1–5.
- [9] R. Dutta, S. Chakrabarti, and A. Sharma, "Topology tracking for active distribution networks," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 2855–2865, Jul. 2021.
- [10] M. Bazrafshan and N. Gatsis, "Comprehensive modeling of three-phase distribution systems via the bus admittance matrix," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 2015–2029, Mar. 2018.
- [11] Y. Ma et al., "Phase identification of smart meters by spectral clustering," in *Proc. 2nd IEEE Conf. Energy Internet Energy Syst. Integr. (EI2)*, Beijing, China, 2018, pp. 1–5.
- [12] G. Cavarro, V. Kekatos, and S. Veeramachaneni, "Voltage analytics for power distribution network topology verification," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 1058–1067, Jan. 2019.
- [13] M. Bariya, D. Deka, and A. von Meier, "Guaranteed phase & topology identification in three phase distribution grids," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3605–3612, Jul. 2021.
- [14] B. Gou and A. Abur, "An improved measurement placement algorithm for network observability," *IEEE Trans. Power Syst.*, vol. 16, no. 4, pp. 819–824, Nov. 2001.
- [15] A. Abur and A. Gómez-Expósito, *Power System State Estimation—Theory and Implementations*. New York, NY, USA: Marcel Dekker, Inc., 2004.
- [16] E. Castillo, A. J. Conejo, R. E. Pruneda, and C. Solares, "State estimation observability based on the null space of the measurement Jacobian matrix," *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1656–1658, Aug. 2005.
- [17] E. Castillo, A. J. Conejo, R. E. Pruneda, and C. Solares, "Observability analysis in state estimation: A unified numerical approach," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 877–886, May 2006.
- [18] E. Castillo, A. J. Conejo, R. E. Pruneda, and C. Solares, "Observability in linear systems of equations and inequalities: Applications," *Comput. Oper. Res.*, vol. 34, no. 6, pp. 1708–1720, Jun. 2007.
- [19] R. E. Pruneda, C. Solares, A. J. Conejo, and E. Castillo, "An efficient algebraic approach to observability analysis in state estimation," *Electr. Power Syst. Res.*, vol. 80, no. 3, pp. 277–286, Mar. 2010.
- [20] N. M. Manousakis and G. N. Korres, "Observability analysis and restoration for systems with conventional and phasor measurements," *Int. Trans. Electr. Energy Syst.*, vol. 23, no. 8, pp. 1548–1566, Nov. 2013.
- [21] A. S. Costa, A. Albuquerque, and D. Bez, "An estimation fusion method for including phasor measurements into power system real-time modeling," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1910–1920, May 2013.
- [22] H. Herrero and C. Solares, "A greedy algorithm for observability analysis," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1638–1641, Mar. 2020.
- [23] D. Thukaram, J. Jerome, and C. Surapong, "A robust three-phase state estimation algorithm for distribution networks," *Electr. Power Syst. Res.*, vol. 55, no. 3, pp. 191–200, Sep. 2000.
- [24] S. Bhela, V. Kekatos, and S. Veeramachaneni, "Power distribution system observability with smart meter data," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Montreal, QC, Canada, 2017, pp. 1070–1074.
- [25] R. Gelagaev, P. Vermeyen, J. Vandewalle, and J. Driesen, "Numerical observability analysis of distribution systems," in *Proc. 14th Int. Conf. Harmonics Qual. Power*, Bergamo, Italy, 2010, pp. 1–6.
- [26] F. Magnago, L. Zhang, and M. K. Celik, "Multiphase observability analysis in distribution systems state estimation," in *Proc. Power Syst. Comput. Conf.*, Genoa, Italy, 2016, pp. 1–7.
- [27] S. Wang, D. Liang, L. Ge, and L. Wu, "PQ decoupled 3-phase numerical observability analysis and critical data identification for distribution systems," *Int. Trans. Electr. Energy Syst.*, vol. 27, no. 6, p. e2317, Jun. 2017.
- [28] T. M. Soares, U. H. Bezerra, and M. E. D. L. Tostes, "Full-observable three-phase state estimation algorithm applied to electric distribution grids," *Energies*, vol. 12, no. 7, p. 1327, Apr. 2019.
- [29] Y. John, V. Venkataramanan, A. M. Annaswamy, A. Dubey, and A. K. Srivastava, "Improved observability in distribution grids using correlational measurements," *IEEE Access*, vol. 10, pp. 27320–27329, 2022.
- [30] S. Bolognani and F. Dörfler, "Fast power system analysis via implicit linearization of the power flow manifold," in *Proc. 53rd Annu. Allerton Conf. Commun. Control Comput. (Allerton)*, Monticello, IL, USA, 2015, pp. 402–409.
- [31] W. H. Kersting, *Distribution System Modeling and Analysis*, 4th ed. Boca Raton, FL, USA: CRC Press, 2017.
- [32] V. N. Temlyakov, "Greedy approximation," *Acta Numerica*, vol. 17, pp. 235–409, May 2008.

- [33] P. M. D. O.-D. Jesus, M. A. Alvarez, and J. M. Yusta, "Distribution power flow method based on a real quasi-symmetric matrix," *Electr. Power Syst. Res.*, vol. 95, pp. 148–159, Feb. 2013.
- [34] W. H. Kersting, "Radial distribution test feeders," in *Proc. IEEE Power Eng. Soc. Winter Meeting*, Columbus, OH, USA, 2001, pp. 908–912.
- [35] *MATLAB, Version 9.12.0.1927505 (R2022a)*, MathWorks, Inc., Natick, MA, USA, 2022.
- [36] V. Kekatos, L. Zhang, G. B. Giannakis, and R. Baldick, "Fast localized voltage regulation in single-phase distribution grids," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Miami, FL, USA, 2015, pp. 725–730.

Natalia Alguacil (Senior Member, IEEE) received the Ingeniero en Informática degree from the Universidad de Málaga, Málaga, Spain, in 1995, and the Ph.D. degree in power systems operations and planning from the Universidad de Castilla-La Mancha, Ciudad Real, Spain, in 2001, where she is currently a Full Professor of Electrical Engineering. Her research interests include operations, planning, and economics of power systems, as well as optimization.

Henar Herrero received the B.S. degree in mathematics from the Universidad de Valladolid, Valladolid, Spain, in 1989, and the Ph.D. degree in sciences from the Universidad de Navarra, Pamplona, Spain, in 1994. She is currently a Full Professor of Mathematics with the Universidad de Castilla-La Mancha, Ciudad Real, Spain. Her research interests are numerical modeling of fluid dynamics problems and observability of power systems.

Cristina Solares received the B.S. and Ph.D. degrees in mathematics from the Universidad de Cantabria, Santander, Spain, in 1994 and 1997, respectively. She is currently an Associate Professor with the Universidad de Castilla-La Mancha, Ciudad Real, Spain. Her research interests include linear algebra applications, observability of power systems, and new mathematic teaching strategies.