A Regularized State Estimation Scheme for a Robust Monitoring of the Distribution Grid

Achilleas Tsitsimelis, Carles Antón-Haro

Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
Parc Mediterrani de la Tecnologia - 08860 - Castelldefels - Barcelona (Spain)
email:{achilleas.tsitsimelis, carles.anton}@cttc.es

Abstract

In this paper, we propose a regularized state estimation (SE) scheme for the distribution grid. The ultimate goal is to track accurately the system state at a faster time scale according to the requirements of the new operational environment. The SE algorithm operates at two different time scales in which the set of available measurements are different. At the main time instants (every 15 minutes) the set of observations comprises SCADA measurements, pseudomeasurements and micro Phasor Measurement Units (µPMUs). In this case, we resort to a Regularized version of the Normal Equations based SE (R-NESE). In the intermediate time instants, only a reduced number of µPMU measurements is available. To circumvent observability issues, we exploit the fact that the voltage drop in adjacent buses is limited and, on that basis, a regularized weighted total variation estimation (WTVSE) problem is formulated. Then, the impact of in-line voltage regulators (IVLRs) is explicitly taken into consideration and that, forces us to decompose and solve the original SE problem for a number of smaller regions (D-WTVSE). The latter can be iteratively solved by resorting to the Alternating Direction Method of Multipliers (ADMM). Complementarily, we also present a µPMU placement method (µPP) in order to improve the conditioning of the R-NESE problem. This problem can be posed as a mixed integer semidefinite programming model (MISDP) and, thus, can be efficiently solved. The performance of the proposed scheme is numerically assessed on (mostly) a 95-bus distribution system for a number of realistic conditions of noise, load and photovoltaic generation profiles. A number of benchmarks are provided, as well.

Index Terms

Distribution grid, µPMU placement, robust, semidefinite programming, state estimation, Tikhonof regularization, two-time scale, weighted total variation.

I. INTRODUCTION

State Estimation (SE), namely, the determination of all the complex voltages in a power network from a set of measured variables from selected buses and branches, is a cornerstone for the monitoring of power systems. SE is a well-studied problem for the Transmission Grid (TG). In contrast, SE for the Distribution Grid (DG) has attracted lots of research interest over the last decade. However, DG SE turns out to be very challenging [1], [2].

To start with, the uniqueness of DG network topologies often renders the SE problem ill-conditioned [3]. Several factors, such as the radial topologies and the low reactance/resistance (X/R) ratio [4], degrade the efficiency of the conventional estimators operating at the transmission level. At the same time, these peculiarities make prohibitive the simplifications adopted in the case of TG SE, e.g., DC SE, and analytical models of the distribution network have to be considered [5].

Several works on DG SE [6-8] have proposed alternative SE schemes suitable for radial-shaped DGs. The underlying idea is to consider the branch currents as the state variables instead of bus voltage and angles. This modification can ensure increased simplicity of the non-linear functions that describe the relationship between the state variables and the AC load flows. For instance, a three-phase SE algorithm is presented in [9], where the power on the branches and the squared branch current magnitudes are estimated. Based on this concept, more recent works have taken into consideration the advent of active distribution networks. In [10] is introduced an efficient model capable of treating phasor measurements, which counts for radial and weakly meshed topologies. From another point of view, a number of schemes have leveraged on the aforementioned characteristics in order to built heuristic methods. A monitoring algorithm based on a new formulation of the AC radial load flows can be found in [11]. The proposed routine relies on the voltage measurements of a crucial set of buses and the neighboring state variables are adjusted iteratively to fit the measured values. In the same context, a model order reduction method is performed on the DG in [12]. The algorithm selects a number of nodes and branches and reduces them to an equivalent one. Consequently, the SE is capable of providing an estimate relying on a reduced number of PMUs.

This work was supported by the EU Seventh Framework Program under grants 607774 and 646469 (ADVANTAGE, P2P Smart Test).
One more recent heuristic approach is presented in [13]. The method relies on an ad-hoc regularized parameter in order to deal with possible numerical instability introduced by the DG specificities. Interestingly, the scheme is implemented with slight modifications to the classical weighted least squares (WLS) estimator.

Another challenge is the limited availability of measurements. In this case, substantial efforts have been devoted to strategies aimed to deal with the lack of system observability. Resorting to data mining techniques on daily load patterns and smart meter data, has been instrumental towards the definition of efficient strategies for the construction of pseudomeasurement sets [14]. By using such pseudomeasurements in conjunction with actual SCADA readings, (slightly) redundant measurement sets can be constructed. In [15] for instance, the authors investigate a two-time scale SE based on WLS. In particular, separating the information (SCADA, pseudomeasurements) according to their availability in time, the authors formulate the SE problem in a two-time scale fashion and investigate the efficiency of the resulting scheme under a diversity of load patterns. Nonetheless, these observation sets consist of a large number of (pseudo) power injection measurements, that in turn affects significantly the numerical conditioning of the normal equations [16] and deteriorates the SE accuracy [17]. In [18], it has been shown that the correlation among these pseudomeasurements reduces substantially the estimation performance. Consequently, an advanced processing of the high uncertainty load injections becomes a necessity [19]. A prior statistical treatment in order to overcome the nonsynchronized nature of the smart meter data is proposed in [20]. Based on the observation that the short-time load variations follow a normal distribution, the authors present a modified WLS-based SE. Further, non-technical restrictions also affect the accuracy of the pseudo loads. A comprehensive study on the DG SE and the relevant challenges is presented in [21]. The privacy restrictions with respect to the data aggregation from smart meters are emphasized. These restrictions affect significantly the accuracy of the pseudomeasurements, degrading the voltage profile accuracy. In order to overcome this barrier, two strategies are proposed. The first is based on the consideration of the correlations among errors, while the second leverages on the power loss estimation at the LV system.

Furthermore, the increasing penetration of Distributed Energy Resources (DER) along with the introduction of new components, such as electric vehicles and distributed energy storage systems, have transformed the DG into an active infrastructure [22]. Besides, the power production profile of DERs (e.g., photovoltaic plants) exhibits variability at a shorter time scale. This results into an increasing number of voltage fluctuations and reverse power flows, which in turn may result in: i) the disconnection of the renewable resources and ii) financial losses for the utilities, respectively [23]. For that reason, the distribution system operators attempt to keep an appropriate (within operational constraints) voltage profile at the medium voltage (MV) level system leveraging on active elements, such as the on-load tap changers at the primary distribution substations and the in-line voltage regulators (ILVRs) across the feeders [24]. However, under specific operational situations these control devices may fail [25]. Consequently, the interest in monitoring (and optimizing) DG’s MV systems with an increased temporal granularity and the maximum reliability has become a necessity.

To that aim, distribution management systems may resort to low-cost, high-resolution micro Phasor Measurement Units ($\mu$PMUs) [26]. These devices are exclusively designed to provide visibility for the MV feeders, where, unlike high voltage systems, the phase angle differences between two measured locations lie usually at the level of small fractions of a degree. Thus, on the one hand, $\mu$PMUs typically operate at faster timescale, as opposed to legacy SCADA and pseudomeasurements (millisecond vs. second or minute timescales). On the other, the precision of their (linear and synchronized) measurements is higher, compared to the classical PMUs deployed at the TG. Indeed, a low number of $\mu$PMUs is able to improve significantly the estimated voltage profile [27], but also to enhance other applications for the DG [28]. However, a number of challenges, mostly related with the optimal number and position of the phasor meters, have to be also addressed. In [29] is presented a mathematical analysis on the impact of synchrophasor technology on the voltage profile estimation. The main sources of uncertainty introduced from the PMU’s are analyzed and interestingly, the authors prove that the number and location of the meters is of quite importance. As a matter of fact, several studies in the literature deal with these issues, following different rationales. For instance, an incremental placement approach can be found in [30], [31]. In the latter, the authors present an integer linear programming model aimed to reach a predefined reliability level of system observability using a minimum
number of phasor meters.

A. Contribution

In this paper, we propose a regularized SE scheme for the DG robust monitoring, which operates in a two-time scale fashion. The ultimate goal is to track accurately the system state at a faster time scale with increased reliability, according to the needs of the new operational environment. In contrast with the previous schemes, the presented algorithm is able to provide monitoring with an increased temporal granularity and thus, to track short-term off-limit system conditions. At the same time, the scheme presents robustness against the high uncertainty of the available information. In details, extending our prior work [32], here is presented a SE algorithm which consists of the following parts:

- A state estimator that operates on the main time instants $kT$, i.e., every 15 minutes (Fig. 1). The observations comprise SCADA measurements and pseudomeasurements. On top of these, a limited number of precise synchrophasors is added. This low redundant measurement set, characterized by a noise variation diversity, results to an ill-conditioned optimization problem and affects significantly the SE accuracy [16]. In order to overcome this barrier and elaborating [32], the SE is posed as a regularized non-linear least squares optimization problem. Resorting to the well-known Tikhonof regularization [33], the problem formulation points out to the regularized version of the normal equations based SE (R-NESE). In addition, prior system information is introduced to the regularization term, that is, the last computed state estimate. Specifically, the proposed R-NESE presents the following properties:
  - Improves the conditioning of the initial SE problem. Due to that reason, the quadratic convergence characteristics of the normal equations are preserved, leading to a fast convergence rate.
  - Presents robustness against uncertainties raised by the low redundancy measurement sets, the high error variance of pseudomeasurements and the specificities of the DG, e.g., low X/R ratio.
  - The proposed scheme is practical for implementation since slight modifications are needed to the existing normal equations based estimators.

- A state estimation scheme that operates on the intermediate time instants $kT + nt$, i.e., between the main computed estimates (Fig. 1). In this case, we exclusively leverage on the positioned $\mu$PMUs at a subset of buses. As unavoidably the number of measurements is low, observability is not guaranteed. To circumvent that, we again resort to regularization techniques [33] leveraging on prior (expert) information on selected DG features. Specifically, the SE is formulated as a weighted total variation estimation (WTVSE) problem which limits variations in voltage estimates in adjacent buses. This stems from the fact that, in DGs, branch impedances (electrical lengths) and current flows tend to be lower than in TGs and so are voltage drops in adjacent buses [34]. On that basis, we propose a rule to define branch-specific weights for the regularizer. Further, possible zero power injection buses are exploited and introduced to the problem as constraints. Finally, extending our prior work [32], the presence of ILVRs across the feeders has been taken into consideration. This, in consequence of the total variation method features, lead as to re-formulate the problem as a constrained decomposed WTVSE (D-WTVSE). Then the D-WTVSE problem is effectively solved by finding an iterative solution based on the Alternating Direction Method of Multipliers (ADMM) [35].

In a glance, the D-WTVSE routine presents the following features:

- An efficient SE model that relies on a limited number of $\mu$PMUs and the zero power injection buses. The model counts for feeder structural differences (e.g., rural or urban), the intermittent nature of DER power production and active components installed within the feeder.
- The SE numerical treatment leverages on the ADMM which offers advantages over the $\ell_1$-norm (WTV regularization) solution and model simplicity compared to traditional methods for ILVRs incorporation.
- Based on the superior properties of the ADMM, the algorithm presents satisfactory convergence rate and a low computational time is needed for full convergence.

Complementarily, here is proposed an ad-hoc $\mu$PMU placement method ($\mu$PP), which has been designed according to the demands of the proposed two-time scale SE scheme. To do so, following a similar rationale with already developed algorithms for sensor scheduling purposes [36], we pose a mixed integer semidefinite (MISDP) optimization problem. In particular, the objectives of the $\mu$PP problem are: (i) to ensure system observability (if needed) for the R-NESE scheme, while taking into consideration the existing measurements; (ii) to optimize the conditioning of the R-NESE problem [17], that in turn will result into more accurate state estimates and (iii) to allocate a sufficient subset of the predefined number of $\mu$PMUs in each DG region, according to the requirements raised from D-WTVSE.

The rest of the paper is organized as follows. In the next section we provide the basic DG SE system model. In subsections III-A and III-B we present the routines for the main (R-NESE) and intermediate time instants (D-WTVSE) SE schemes, respectively. The solution of D-WTVSE via ADMM is presented in Section IV, while Section V introduces the $\mu$PMU...
placement method. Finally, the proposed algorithms are numerically assessed in Section VI and we draw our conclusions in Section VII.

II. SYSTEM MODEL

Consider a balanced radial distribution grid (see Fig.2) represented by the graph \( G = (V,B) \), where \( V \) denotes the set of buses with cardinality \( |V| = N \); and \( B \) stands for the set of edges that describes their interconnections (branches), with cardinality \( |B| = N-1 \). The complex current injections at the buses, i.e., \( I = [I_1, \ldots, I_N]^T \), satisfy \( I = YV \), where \( Y \in \mathbb{C}^{N \times N} \) is the nodal admittance matrix [38] and \( v = [V_1, V_2, \ldots, V_N]^T \in \mathbb{C}^{N} \) stands for the complex voltages at all the buses that determines the state of the DG. The ultimate goal of power system state estimator (SE) is to estimate \( v \) from a set of measurements at selected buses and branches.

Two types of measurements are considered available, namely, legacy and \( \mu \)PMU measurements. As depicted in Fig. 1, their availability is divided at two different time scales and, moreover, observations are of different nature. The first category comprises readings from the SCADA system as well as pseudo-measurements. Legacy measurements are typically obtained at time intervals of \( T = 15 \) minutes\(^1\) [15] and include: squared voltage magnitudes, \( |V_m|^2 \) for \( m \in \mathcal{V} \); power flows at the branches, given by \( S_{m,t} = P_{m,t} + jQ_{m,t} = V_m I_{m,t}^* \) for \( (m,t) \in B \); power injections, given by \( S_m = P_m + jQ_m = V_m I_m^* \) for \( m \in \mathcal{V} \) and squared branch current magnitudes, \( |I_{m,l}|^2 \) for \( (m,l) \in B \). In contrast, the sampling period of \( \mu \)PMUs \( t = T/q \) is smaller, typically on the order of few milliseconds [26], and provide the DG operator with updated snapshots of complex bus voltages, \( V_m = \Re\{V_m\} + j\Im\{V_m\} \) for \( m \in \mathcal{V} \). Notice that legacy measurements are non-linear functions of the system state, whereas \( \mu \)PMU measurements turn out to be linear ones.

Let vector \( \mathbf{x} = \left[ \Re\{V_1\}, \Im\{V_1\}, \ldots, \Re\{V_N\}, \Im\{V_N\} \right]^T \in \mathbb{R}^{2N} \) denote the Cartesian representation of the system state, and \( \mathbf{z} = [r^T, y^T]^T \in \mathbb{R}^M \) stand for the corresponding measurement vector, where \( r \in \mathbb{R}^L \) and \( y \in \mathbb{R}^K \) are the vectors of legacy and \( \mu \)PMU measurements, respectively. It is worth noting that, typically, applies \( M = L + K > 2N \) but \( K < 2N \) due to deployment and operational costs. Finally, we can write:

\[
\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{e} \tag{1}
\]

where \( \mathbf{h}(\mathbf{x}) \) denotes a non-linear function of \( \mathbf{z} \) on \( \mathbf{x} \) in compliance with the AC power flow model [38]; and \( \mathbf{e} \in \mathbb{R}^M \) stands for zero-mean Gaussian noise with known covariance matrix \( \mathbf{R}_e \).

III. A TWO-TIME SCALE STATE ESTIMATION SCHEME

Bearing all the above in mind, we propose the following SE algorithm. At time instants \( kT, (k + 1)T, \ldots, (k + n)T \) for \( k \in \mathbb{N} \), where both legacy and \( \mu \)PMU measurements are available, a SE algorithm based on the regularized normal equations is adopted (see Section III-A next). On the contrary, for the intermediate time instants between \( kT \) and \( (k+1)T \) we adopt a state estimator formulated as a constrained decomposed weighted total variation regularization problem (subsection III-B). Note that the latter can only rely on the observations coming from a reduced number of \( \mu \)PMUs and the possibly available zero power injections.

\(^1\)With respect to system observability, SCADA observations and pseudomeasurements are sufficiently available every 15 minutes [15].
A. Regularized Normal Equations-based State Estimator (R-NESE)

According to (1), the conventional state estimator is given by the solution to the following non-convex optimization problem:

$$\hat{x} = \arg\min_{x} \frac{1}{2} \|z - h(x)\|^2$$

(2)

where, in the above (and following) expression(s), the time index has been dropped for the ease of notation. To solve (2), one can resort to the so-called Normal Equations [16]:

$$\mathcal{G}^{(\nu)} \Delta x^{(\nu)} = J_x^{(\nu)T} R_e^{-1} [z - h(x^{(\nu)})]$$

(3)

where \(\nu\) stands for the algorithm iteration index; matrix \(J_x^{(\nu)} \in \mathbb{R}^{M \times 2N}\) denotes the Jacobian matrix of \(h(x)\) evaluated at \(\nu^{th}\) iteration (i.e., the matrix of all first order partial derivatives of the non-linear equations vector \(h(x^{(\nu)})\)); \(R_e \in \mathbb{R}^{M \times M}\) is the known covariance matrix based on the standard deviation \(\sigma\) of the measurements \((\sigma_1, \ldots, \sigma_M)\); \(\mathcal{G}^{(\nu)} = J_x^{(\nu)T} R_e^{-1} J_x^{(\nu)}\) stands for the gain matrix and \(\Delta x^{(\nu)} = x^{(\nu+1)} - x^{(\nu)}\) stands for the vector of increments of the state variables.

However, in the case of DG SE the numerical accuracy is dramatically influenced by the low redundant measurement set which consists of observations characterized from accuracy diversity. The latter is translated to a near-singular gain matrix \(\mathcal{G}\) [16] which in turn results into an ill-conditioned system of equations in (3).

For that reason, we adopt the following regularized non-linear least-squares optimization problem [39]:

$$\hat{x} = \arg\min_{x} \frac{1}{2} \left[ (z - h(x))^{T} R_e^{-1} (z - h(x)) + \tau \|x - x_p\|^2 \right]$$

(4)

that refers to the Tikhonof regularization method [33], also known as ridge regression [40] in the field of statistics. In (4), \(\tau\) is the parameter that controls the amount of regularization establishing the trade-off between the regularization term \(\|x - x_p\|^2\) and data fidelity [41]. In addition, \(x_p\) is an application dependent vector, here, the previous (in time) estimated state.

The Tikhonof regularization method is commonly applied to non-well posed optimization models [42]. These problems present numerical instability, that means, small perturbations to the input data results to large variances to the output. In other words, the solution lies within an excessively large set. Thus, the regularization method exploits prior information in order to restrict this set. More precisely, based on the assumption that the exact solution is described from a smooth function, the algorithm returns a result that approximates efficiently the real one. From a mathematical perspective, the underlying rationale is to extract the source of ill-conditioning from the respective system of equations. With respect to the SE problem (3), this entails the filtering of the contributions to the state estimate \(\hat{x}\) which correspond to the lower singular values of \(\mathcal{G}\). The latter is linked with the value of the control parameter \(\tau\) and consequently, an appropriate choice is needed [41].

Now, let us call \(f(x)\) the non-linear function in problem (4). Expanding \(f(x)\) into its Taylor series around \(x^{(\nu)}\) and neglecting the higher order terms, gives us:

$$\tilde{f}(x) = f(x^{(\nu)}) + \nabla f(x^{(\nu)})(x - x^{(\nu)}) + \frac{1}{2} \nabla^2 f(x^{(\nu)})(x - x^{(\nu)})^2$$

(5)

and the first order optimality condition for (5) can be computed as:

$$\nabla \tilde{f}(x) = \nabla f(x^{(\nu)}) + \nabla^2 f(x^{(\nu)})(x - x^{(\nu)}) = 0$$

(6)

That, lead us to the iterative Gauss-Newton method:

$$x^{(\nu+1)} = x^{(\nu)} - [\nabla^2 f(x^{(\nu)})]^{-1} \nabla f(x^{(\nu)})$$

(7)

with:

$$\nabla f(x^{(\nu)}) = - J_x^{(\nu)T} R_e^{-1} (z - h(x^{(\nu)})) + \tau(x^{(\nu)} - x_p)$$

(8)

and

$$\nabla^2 f(x^{(\nu)}) = J_x^{(\nu)T} R_e^{-1} J_x^{(\nu)} + \tau I$$

(9)

One can easily observe that equation (7) points out to the regularized form of the normal equations (3):

$$\left(J_x^{(\nu)T} R_e^{-1} J_x^{(\nu)} + \tau(\nu) I\right) \Delta x^{(\nu)} = J_x^{(\nu)T} R_e^{-1} [z - h(x^{(\nu)})] - \tau(\nu)(x^{(\nu)} - x_p)$$

(10)

where \(\tau(\nu)\) has to be computed in each iteration \(\nu\) (see Section VI).

Algorithm 1 summarizes the corresponding iterative procedure.
from ILVRs is equal. That given, the ILVRs (when active) will break down (in contrast to the case of substation transformers), the voltage base (used for the per-unit calculations) in both areas aside regulation if the respective limits have been exceeded (e.g., 0.95 - 1.05 per-unit). However, we have to take into account that there is considered the presence of the so-called in-line voltage regulators (ILVRs), which automatically provide voltage levels, leading the total variation regularization term in (12) to possible inefficiency. In order to deal with this specificity in [32].

Accordingly, matrix $L_m$ in the regularization term generates the vector with the voltage differences for each pair of adjacent buses; and $\nu$ is the regularization parameter, which controls the tradeoff between data fidelity and prior knowledge. The matrix $R$ denotes zero-mean Gaussian noise with known covariance matrix $R_w$. In comparison with the full measurement vector $z$, the number of measurements provided by the $\mu$PMUs is typically lower than the dimension of the system state vector, i.e, $K \leq 2N$. This possibly renders the SE problem unsolvable since the system of equations (11) is underdetermined. This can be alleviated by introducing a regularization term accounting for the fact that, in DGs, the voltage difference between adjacent buses is small.

From all the above, the state estimation problem can be posed as a constrained weighted total variation regularization problem (WTVSE) [43]:

$$\hat{x} = \arg \min_{\{x\}} \frac{1}{2} \|y - Hx\|_2^2 + \lambda \|WLx\|_1$$  

s.t. $\hat{y}x = 0$  

where, in the score function above, the term $\frac{1}{2} \|y - Hx\|_2^2$ encourages fidelity in the solution to the vector of observations $y$; $U(x) = \|WLx\|_1$ is the regularization term, which introduces a prior model of the state vector $x$; and $\lambda$ denotes the regularization parameter, which controls the tradeoff between data fidelity and prior knowledge. The matrix $L \in \mathbb{R}^{(2N-2) \times 2N}$ in the regularization term generates the vector with the voltage differences for each pair of adjacent buses $m$ and $l$. Accordingly, matrix $W \in \mathbb{R}^{(2N-2) \times (2N-2)}$ is the diagonal weighting matrix which reflects the specific characteristics (i.e., electrical lengths) of the various branches in MV feeders\footnote{These may be substantially different in e.g., rural or urban areas.}. Finally, the equality constraint $\hat{y}x = 0$ in problem (12) accounts for the $d$ zero current (power) injection buses. A more detailed description of the previous terms has been given in [32].

Then, the WTVSE problem (12) is re-formulated to account for active elements installed within the DG. Specifically, there is considered the presence of the so-called in-line voltage regulators (ILVRs), which automatically provide voltage regulation if the respective limits have been exceeded (e.g., 0.95 - 1.05 per-unit). However, we have to take into account that (in contrast to the case of substation transformers), the voltage base (used for the per-unit calculations) in both areas aside from ILVRs is equal. That given, the ILVRs (when active) will break down the feeders in areas of different per-unit voltage levels, leading the total variation regularization term in (12) to possible inefficiency. In order to deal with this specificity of WTVSE, for the numerical treatment of the ILVRs we will follow a decomposed formulation based on their location. Nevertheless, this offers the advantage of increased model simplicity compared to the commonly implemented methods, which are based on the augmented state vector and admittance matrix [34]. As an illustrative example, Fig. 3 presents the 95-bus UK DG. The system is divided in three areas according to the location of the ILVRs and in each case, the buses connected to the ILVR branch, i.e., \{23-24\} and \{54-75\}, become the border nodes of the respective adjacent regions.

Now, consider a DG feeder that includes $P$ ILVRs. Let assume that each regulator divides the system in different areas, namely $r$ and $s$, of different voltage levels (Fig. 4) and the total number of areas is $A$. Accordingly, we define $A_r$ as the set of adjacent areas to area $r$ and $S_{rs}$ as the set of adjacent buses $m$ and $l$ between areas $r$ and $s$. Then, the constraints

\begin{algorithm}
\caption{Regularized NESE state estimation at an arbitrary time instant $kT$}
1: Initialize $\nu = 0$, $x^{(0)}$, $x_p$, $J_x^{(0)}$, $G^{(0)}$, $\tau$.
2: repeat
3: \hspace{1em} $\nu = \nu + 1$
4: \hspace{1em} Compute $\tau^{(\nu)}$.
5: \hspace{1em} Compute $\Delta x_k^{(\nu)}$ with (10).
6: \hspace{1em} Update $\hat{x}_k^{(\nu)} = \Delta x_k^{(\nu)} + \hat{x}_k^{(\nu-1)}$.
7: until $\Delta x_k^{(\nu)} < \varepsilon$
\end{algorithm}
Fig. 3: Decomposition of the 95-bus UK DG according to the location of the ILVRs.

Fig. 4: WTVSE decomposition according to the ILVR.

\[ a_{rs} V_{mR} = V_{lR} \quad \text{and} \quad a_{rs} V_{mI} = V_{lI} \]

have to be satisfied, where \( a_{rs} \) stands for the tap position indicator of the corresponding ILVR, usually displayed in the distribution management system.

Taking into account all the above, (12) can be re-formulated as a decomposed WTVSE (D-WTVSE) optimization problem based on the positions of the ILVRs (Fig. 4):

\[
\min_{\{x_r\}} \sum_{r=1}^{A} \frac{1}{2} \| y_r - H_r x_r \|_2^2 + \lambda \| W_r L_r x_r \|_1
\]

s.t. \( \tilde{Y}_r x_r = 0_r \quad \text{for} \quad s \in A_r, (m,l) \in S_{rs}, \forall r \) (13)

where \( x_r \in \mathbb{R}^{2N_r} \), \( y_r \in \mathbb{R}^{K_r} \), \( H_r \in \mathbb{R}^{K_r \times 2N_r} \) refer to area \( r \) and follow the rationale explained above. The same applies for \( W_r \in \mathbb{R}^{(2N_r - 2) \times (2N_r - 2)} \), \( L_r \in \mathbb{R}^{(2N_r - 2) \times 2N_r} \), and \( \tilde{Y}_r \). The vector \( e \) is the singleton vector with all elements zero except one with 1 as its entry. For instance, \( e_{mR}^{(r)} \) stands for a vector with 1 at the position \( 2m-1 \) \((2m)\) that refers to the real (imaginary) part of the voltage at the bus \( m \). In addition, the superscript \( r \) denotes that the vector size is that of \( x_r \). Finally, as already mentioned, \( a_{rs} \) refers to the real value ratio of the voltage at the two sides of the ILVR (buses \( m \) and \( l \)).

IV. Solving the D-WTVSE Problem via ADMM

In order to solve the D-WTVSE problem (13), we resort to the Alternating Direction Method of Multipliers (ADMM) [35]. Among other candidates, ADMM presents computational advantages over the non-smooth \( \ell_1 \)-norm treatment and adapts

3The tap position practically refers to the number of winding turns at the primary and secondary side of the ILVR (amount of voltage regulation provided) and is translated to the real value ratio (e.g. 1.05) between the respective primary and secondary side voltage.
well to decomposed optimization problems. To that aim, however, problem (13) must be re-written as:

$$
\begin{aligned}
& \text{arg} \min_{\{x_r\}, \{c_{rs}\}, \{\theta_r\}} \sum_{r=1}^{A} \frac{1}{2} \|y_t - H_t x_r\|^2 + \lambda \|\theta_r\|_1 \\
& \text{s.t.} \begin{bmatrix}
W_t L_t \\
Y_t \\
M_{rs}
\end{bmatrix} x_r - \begin{bmatrix}
\theta_r \\
c_{rs}
\end{bmatrix} = 0_r,
\end{aligned}
$$

(14)

where $M_{rs} \in \mathbb{R}^{2 \times 2N_r}$ for all $s \in A_r$ is defined as:

$$
|M_{rs}|_{i,j} = \begin{cases}
    a_{rs}, & i = 1, \quad j = 2m - 1; \\
    a_{rs}, & i = 2, \quad j = 2m; \\
    0, & \text{otherwise}
\end{cases}
$$

(15)

and $c_{rs} = [c_{rsml(I)}; c_{rsml(J)}]$ for all $s \in A_r$ stands for an introduced auxiliary vector per each pair of connected areas via ILVR, in order to fully decompose the problem. Moreover, $\theta_r$ plays the role of an auxiliary variable and, hence, it introduces an additional constraint. Accordingly, the augmented Lagrangian reads:

$$
\mathcal{L}(x_r, c_{rs}, \theta_r, \mu_r, \pi_r, \kappa_{rs}) = \sum_{r=1}^{A} \left[ \frac{1}{2} \| y_t - H_t x_r \|^2 + \lambda \| \theta_r \|_1 \\
+ \mu_r^T (W_t L_t x_r - \theta_r) + \frac{c_1}{2} \| W_t L_t x_r - \theta_r \|^2 + \pi_r^T (Y_t x_r) \\
+ \frac{c_2}{2} \| Y_t x_r \|^2 \right] + \sum_{s \in A_r} \frac{c_3}{2} \| M_{rs} - c_{rs} \|^2 + \kappa_{rs}^T (M_{rs} - c_{rs})
$$

(16)

with $\mu_r$, $\pi_r$ and $\kappa_{rs}$ standing for the vectors of Lagrange multipliers, and $c_1$, $c_2$, $c_3$ denoting pre-defined positive constants. The ADMM iterates for the updates of the primal and dual variables yield:

$$
\begin{align*}
    x_r^{(\nu+1)} &= \text{arg} \min_{x_r} \mathcal{L}(x_r, c_{rs}^{(\nu+1)}, \theta_r^{(\nu)}, \mu_r^{(\nu)}, \pi_r^{(\nu)}, \kappa_{rs}) \\
    c_{rs}^{(\nu+1)} &= \text{arg} \min_{c_{rs}} \mathcal{L}(x_r^{(\nu+1)}, c_{rs}, \theta_r^{(\nu)}, \mu_r^{(\nu)}, \pi_r^{(\nu)}, \kappa_{rs}) \\
    \theta_r^{(\nu+1)} &= \text{arg} \min_{\theta_r} \mathcal{L}(x_r^{(\nu+1)}, c_{rs}^{(\nu+1)}, \theta_r, \mu_r^{(\nu)}, \pi_r^{(\nu)}, \kappa_{rs}) \\
    \mu_r^{(\nu+1)} &= \mu_r^{(\nu)} + c_1 (W_t L_t x_r^{(\nu+1)} - \theta_r^{(\nu+1)}) \\
    \pi_r^{(\nu+1)} &= \pi_r^{(\nu)} + c_2 (Y_t x_r^{(\nu+1)}). \\
    \kappa_{rs}^{(\nu+1)} &= \kappa_{rs}^{(\nu)} + c_3 (M_{rs} - c_{rs}^{(\nu+1)}).
\end{align*}
$$

(17-22)

One can easily prove that the closed form of the sequential updates of the primal variables at iteration $\nu$ can be written as:

$$
\begin{aligned}
    x_r^{(\nu+1)} &= \left( H_t^T H_t + c_1 L_t^T W_t^T W_t L_t + c_2 Y_t^T Y_t \right)^{-1} \\
&\quad + \sum_{s \in A_r} \left( M_t^T M_{rs} \right)^{-1} \\
&\quad \left( H_t^T Y_t - L_t W_t^T \mu_r^{(\nu)} + c_1 L_t^T W_t^T \theta_r^{(\nu)} - Y_t \pi_r^{(\nu)} \right) \\
&\quad + \sum_{s \in A_r} \left( M_{rs} c_{rs} - \kappa_{rs} M_{rs}^{(\nu)} \right). \\
    c_{rs}^{(\nu+1)} &= \frac{M_{rs} x_r^{(\nu+1)} + M_{sr} x_s^{(\nu+1)} + \frac{\kappa_{rs}^{(\nu)} + \kappa_{sr}^{(\nu)}}{2c_3}}{2} \\
    \theta_r^{(\nu+1)} &= S_{\lambda/c_1} (W_t L_t x_r^{(\nu+1)} + c_1^2 \mu_r^{(\nu)}).
\end{aligned}
$$

(23-25)
with $S_a(\chi)$ standing for the soft thresholding operator, that is,

$$
S_a(\chi) = \begin{cases} 
\chi + a & \text{if } \chi < -a; \\
0 & \text{if } |\chi| \leq a; \\
\chi - a & \text{if } \chi > a.
\end{cases}
$$

(26)

V. \(\mu\)PMU PLACEMENT ($\mu$PP)

As already mentioned (Sections I and III), in the case of DG SE the system observability is achieved using a large number of power injection measurements. These values are either pseudomeasurements with a noise standard deviation $\sigma$ at the level of $10^{-1}$ [21], or virtual measurements, such as zero power injections, with $\sigma$ at the level of $10^{-5}$. Unfortunately, these particular measurements generate an ill-conditioned gain matrix $G$ [16] and affect remarkably the SE accuracy [17].

The degree to which $G$ is ill-conditioned is given by the condition number $\kappa$, that is:

$$
\kappa = \|G\| \|G^{-1}\| \tag{27}
$$

or equivalently:

$$
\kappa = \lambda_{\text{max}}(G)/\lambda_{\text{min}}(G) \tag{28}
$$

where $\lambda_{\text{max}}$ and $\lambda_{\text{min}}$ stand for the maximum and minimum eigenvalue of $G$, respectively.

Bearing the above in mind, the purpose of this section is to propose a $\mu$PMU placement method that enhances robustness to the R-NESE estimator. Specifically, the target is, firstly to construct a full rank $G$ (if needed); and more importantly to optimize the condition number $\kappa$. Moreover, any requirements raised from the intermediate state estimator D-WTVSE are taken into account explicitly.

First, let us consider a number of $o$ phasor meters for installation and a binary decision vector $q \in \mathbb{R}^N$, where each element $q_i$ for $i = 1 \ldots N$ denotes the presence or absence of a $\mu$PMU at the respective bus $i$. With respect to the decision vector $q$, the gain matrix $G$ can be written as follows [44]:

$$
G(q) = [G_0 + \sum_{i=1}^N q_i G_i] \tag{29}
$$

where $G_i = J_i^T R_i^{-1} J_i$ is the corresponding gain matrix for each candidate $\mu$PMU for $i = 1 \ldots N$ and $G_0 = J_0^T R_0^{-1} J_0$ is the initial gain matrix generated from the already existed (legacy) measurements. Both $G_0$ and $G_i$ for $i = 1 \ldots N$ are evaluated at $x_0$, that is, the initial point for (10). The power system is said to be numerically observable when $G(q) \succ 0$.

Based on all the previous, the $\mu$PMU placement problem ($\mu$PP) can be posed as the follow mixed integer semidefinite programming (MISDP) optimization program:

\[
\begin{align*}
\min_{q_i, \xi, \omega} \quad & \xi - \omega \\
\text{s.t.} \quad & \omega I \preceq G(q) \preceq \xi I \tag{30a} \\
& \omega_0 \leq \omega \leq \xi \leq \xi_0 \tag{30b} \\
& \sum_{i=1}^{|N|} q_i = o \tag{30c} \\
& \forall r \in A, \forall i \in N_r, \sum_{i=1}^{|N_r|} q_i = o_r \quad N_r \subseteq N \tag{30d} \\
& \forall i \in N \setminus \{1\} \quad (q_i + q_{i-1}) \leq 1 \tag{30e} \\
& \forall i \in \{0, 1\}^N \tag{30f}
\end{align*}
\]

The first constraint in the $\mu$PP problem, (30b), can be interpreted as two separate linear matrix inequalities. The first, reads $G(q) - \omega I \succeq 0$ and guarantees a well-invertible positive definite $G(q)$ where $\omega$ stands for an auxiliary (real value) variable [33]. The role of $\omega$ is to dominate the minimum eigenvalue of $G(q)$ and satisfies $\omega > \omega_0$ where $\omega_0 >> 0$ denotes an

\(^4\)Commonly, $x_0$ is chosen as the flat voltage profile, i.e., all voltage magnitudes are set to 1 and all angles to 0.
appropriately chosen constant. The second linear matrix inequality, $\xi I - \mathcal{G}(q) \succeq 0$, follows the same rationale where the auxiliary variable $\xi$ dominates the maximum eigenvalue of $\mathcal{G}(q)$. The final goal is to minimize the difference $\xi - \omega$ with respect to the binary decision vector $q$. The computer simulations have shown that both approaches conclude to the same optimal solution.

The rest of the constraints are related with the number and the location of the $\mu$PMUs. The constraint (30d) refers to the total constant number $o$ of $\mu$PMUs to be placed based on different criteria, such as economic budget and communication feasibility study [45]. The respective constraint (30e) satisfies the requirement to allocate a sufficient number of meters $o_r$ at the set of buses $\mathcal{N}_r$ of each area $r$ (see subsection III-B). The last, (30f), is in charge of preventing the placement of two meters in adjacent buses.

Comparing the $\mu$PP method with other works from the literature, in terms of qualitative characteristics, a similar approach can be found in [46]. However, the authors emphasize on the placement of (non-synchronized) voltage and power flow meters. The same applies for [47], where a circuit representation-based mixed integer linear programming model is presented. With respect to the synchrophasor technology, a work of similar rationale can be found in [30]. The authors present a PMU placement method that improves the numerical stability of the SE problem. More specifically, the target of the proposed approach is to decrease the variances of the state estimates in order to improve the algorithmic accuracy, and to avoid the correlation among measurements aiming to optimize the bad data analysis. Other interesting works can be found in [48], [49], where the number of meters (to be placed) is optimized in relevance with financial criteria. In contrast, the presented method is able to allocate a predefined number of $\mu$PMUs according to possible budget availability and communication constraints [45]. Besides, based on the constraints (30d) and (30e), the operator is able to control the number of meters located to the different feeder areas. For instance, this can be exploited for increased monitoring in a specific subset of buses where industrial loads or high PV penetration are located.

VI. NUMERICAL RESULTS

The performance of the presented regularized SE scheme has been numerically assessed on the 95-bus UK distribution grid, Fig. 2 [37]. This system has a nominal voltage of 11 kV and the installed active/reactive load is 3.83 MW/0.95 MVar, respectively. As Fig. 2 illustrates, we have considered four photovoltaic (PV) plants as DER. Their nominal capacity is 0.8, 0.8, 0.5 and 0.4 MW and they are connected to buses 15, 36, 68 and 91, respectively. Concerning observations, we have considered a set of legacy measurements with redundancy $L/N \approx 1.1$. This includes: (i) 5 voltage measurements located at the substation and the DER buses; (ii) 3 pairs of branch power flows $P_{i,j}$, $Q_{i,j}$ located at the substation; (iii) 4 squared current magnitudes at the feeders $|I_{i,j}|^2$; and (iv) 90 pairs of power injections $P_i, Q_i$, out of which 40 are pseudomeasurements, and 50 are zero power injections (virtual measurements). Following the rationale of similar works [50], each type of the measurement subsets is corrupted with independent and identical distributed (i.i.d.) Gaussian noise with standard deviations ($\sigma$) $10^{-3}$, $2 \cdot 10^{-3}$, $4 \cdot 10^{-3}$, $3 \cdot 10^{-2}$, and $2 \cdot 10^{-6}$ per-unit (p.u.) [38], respectively. On top of these observations, 12 $\mu$PMU measurements are incorporated, located according to the $\mu$PP method proposed in Section V, with $\sigma = 10^{-4}$ p.u. and $2 \cdot 10^{-4}$ rad for voltage magnitude and angle, respectively [26]. Unless otherwise stated, computer simulation results are averaged over 5000 realizations of noise, load and PV generation profiles. That means, 5000 exact power flow simulations have been considered, where in each realization the loads have been assumed varying as Gaussian-distributed with $P_m \sim \mathcal{N}(P_m, 0.06)$, $Q_m \sim \mathcal{N}(Q_m, 0.06)$ for $m = 1, 2, \ldots, N \in \mathcal{V}$. From these load flows, 5000 measurement sets have been generated with varying noise where in each measurement it has been considered the actual value as mean deviated according to the above standard deviations. In order to compute the error of the associated state estimate of each realization we leverage on the Root-Mean-Square Error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{2N} (\hat{x}_i - x_i)^2}{2N}}$$  \hspace{1cm} (31)

where $\hat{x}_i$ is the estimated element of the state vector $\hat{x}$ and $x_i$ is the exact value. For the power flow analysis, we have used MATPOWER 5.1 [51]. Results have been verified through CVX [52]. For the MISDP $\mu$PP problem solution, we have used the optimization software YALMIP [53] and specifically the solver SEDUMI.

The constants $\xi_0$ and $\xi_1$ can be chosen with criteria based on the $\lambda_{\min}$ and $\lambda_{\max}$ of the matrices $\mathcal{G}_0$ and $\mathcal{G}_i$ for $i = 1 \ldots N$. A further brute-force search is needed to lead to the appropriate constant values.

The binary decision vector $q$ renders the $\mu$PP problem computational complex. However, the applied solvers, using a branch and bound framework, provide, if not the global optimal, an efficient local optimal solution for small and medium size applications (as in the case of DG feeders). Another approach is to solve the convex counterpart of the $\mu$PP problem where the binary constraint $q \in \{0, 1\}$ is relaxed to $q \in [0, 1]$. In that case, the decision vector will consist of real value elements; for that reason a rounding procedure and a local search should follow. The computer simulations have shown that both approaches conclude to the same optimal solution.
A. R-NESE

First of all, we focus on the performance of the R-NESE scheme (10). To recall, it operates at time instants $kT$, when both legacy and $\mu$PMU measurements are available. The algorithmic performance of the proposed scheme has been numerically assessed under a diversity of scenarios. Primarily, there have been considered two cases with different error level for the pseudomeasurements, i.e., $\sigma_1 = 0.03$ and $\sigma_2 = 0.15$ p.u.. For each case, we present two scenarios, namely, with and without including $\mu$PMUs. For the scheme with $\mu$PMUs, the measurement configuration with 12 meters proposed in subsection VI-C has been incorporated. Further, in order to demonstrate explicitly the advantages provided from the regularization term in (4), the following benchmarks have been used: (i) the NESE scheme solving (3) and (ii) the R-NESE scheme solving (10) with $x_p = 0$. Finally, in all realizations $x_p$ refers to the last estimated state from (D)-WTVSE scheme at the time instant $kT + nt$ (see Fig. 1 and subsection VI-B) with an additional error of $\sigma = 0.01$ p.u. in order to encapsulate further uncertainty.

Figs. 5 and 6 show the probability density functions (PDFs) and the average value of the RMSE (p.u.) over the 5000 realizations, respectively. With respect to the PDFs, by definition the area below the curves is equal to 1. As the figures present, NESE reaches high values of error because of the high percentage of pseudomeasurements ($100/205 = 48.8\%$) among the low redundant measurement set. In approach (ii), the accuracy has been drastically increased, however, the gain is moderate compared to R-NESE. This is clearly due to the introduction of the previous estimated state $x_p$, which erases possible uncertainty raised from the heuristic nature of the regularization method. The proposed approach attains the highest accuracy, i.e., $5.7 \times 10^{-3}$ p.u. (Table I), and two orders of magnitude lower compared to the benchmarks. In contrast with other works [14], the scheme presents robustness against the high uncertainty of the pseudo-load injections. The simulations reveal that for all buses the absolute error for the voltage magnitude stays below $1 \times 10^{-2}$ p.u.. Interestingly, we also observe a significant improvement on the condition number $\kappa$ of the gain matrix $G$ that, as it will become clear later, is related with the accuracy and convergence rate of the algorithm. Concerning the second scenario (Figs. 5 and 6, right side), where $\mu$PMUs have been included in the measurement set, the proposed approach, R-NESE, presents the similar superior numerical behavior. However, with regard to these graphs, two further observations can be mentioned: (i) as expected, synchrophasor measurements improve the accuracy for all approaches (Table I); and (ii), interestingly, the figures show that benchmark (ii) presents a slight performance degradation compared to (i). The latter can be explained by the SVD analysis of the Tikhonof solution (10) [40]. More precisely, according to the chosen value of $\tau$, specific contributions to $\hat{x}$ that correspond to the lower eigenvalues of $G$ will be filtered out. However, for this specific scenario, as the conditioning of $G$ has been improved (due to the $\mu$PMUs), the optimal value of $\tau$ (see below) dominates further eigenvalues that compose profitable information.

Additionally, for the consistent presentation of the results, Fig. 7 shows the RMSE for each state variable averaged over the realizations, that is:

$$RMSE_i = \sqrt{\frac{\sum_{n=1}^{nor} (\hat{x}_i - x_i)^2}{nor}}$$

(32)

where $\hat{x}_i$ is the estimated element of the state vector $\hat{x}$, $x_i$ is the exact value and nor is the number of realizations. To mention that, in this case the state variables are demonstrated using the polar form (and not the real and imaginary part of each complex voltage), that is, $\theta_i$ in rad and $V_i$ in p.u. values for $i = 1...N$. As we can observe at the (left) graph, the R-NESE presents high accuracy for the angle estimates. The maximum $RMSE_i$ is below the value of $1 \times 10^{-2}$ rad. At the same time, the RMSE of the majority of the bus phase angles lies within the interval of $3 \times 10^{-3} - 5 \times 10^{-3}$ rad. In contrast, for the second scenario where $\mu$PMUs have been incorporated, the majority of voltage angles RMSE stays below $1 \times 10^{-3}$ rad and the maximum case reaches the value of $6 \times 10^{-3}$ rad. Important to note that, in the buses (and the respective neighboring ones) where the phasor meters have been positioned the RMSE is decreased. The same behavior is observed for both scenarios in the case of voltage magnitudes.

TABLE I: RMSE and condition number $\kappa$ of $G$ for each approach, without (top) and with (bottom) $\mu$PMUs, $\sigma_1 = 0.03$ for pseudomeasurements, subsection VI-A.

| Approach | RMSE (p.u.) | $\kappa$ |
|----------|-------------|----------|
| NESE     | $1.84 \times 10^{-1}$ | $8.8 \times 10^{4}$ |
| $\tau \|x\|^2$ | $8.86 \times 10^{-2}$ | $3.62 \times 10^{9}$ |
| $\tau \|x - x_p\|^2$ | $5.7 \times 10^{-4}$ | $4.41 \times 10^{9}$ |

| Approach | RMSE (p.u.) | $\kappa$ |
|----------|-------------|----------|
| NESE     | $1.15 \times 10^{-2}$ | $6 \times 10^{7}$ |
| $\tau \|x\|^2$ | $1.18 \times 10^{-4}$ | $5.9 \times 10^{7}$ |
| $\tau \|x - x_p\|^2$ | $3.9 \times 10^{-3}$ | $4.1 \times 10^{8}$ |

In addition, Fig. 8 shows the results for the case where an increased error level is considered for the pseudomeasurements, i.e., $\sigma_2 = 0.15$ p.u.. Firstly, it is observed an accuracy degradation for all approaches, Table II. It is worth mentioning
that, for the two benchmarks the RMSE reaches high values (e.g., 3.28 p.u. for NESE). As it was expected, the high error deviation measurements lead to an ill-conditioned matrix $G$ ($\kappa = 5.39 \times 10^{16}$), that in turn, leads many of the realizations to converge at non-efficient local minima or not converge at all, Fig. 8. On the other hand, R-NESE performs well, presenting a satisfactory average RMSE value, i.e. $1.1 \times 10^{-2}$ p.u. without $\mu$PMUs and $8.4 \times 10^{-3}$ p.u. with $\mu$PMUs.

Again, Fig. 9 presents the RMSE for each state variable. As it was expected, compared to the previous case, the results are slightly scaled to higher estimation error. For instance, the majority of the RMSE for the bus phase angles lies within the interval $4 \times 10^{-3} - 8 \times 10^{-3}$ rad. For the case with $\mu$PMUs this error interval is $1 \times 10^{-5} - 5 \times 10^{-5}$ rad. However, in both angles and voltage magnitudes the R-NESE scheme, still, attains high accuracy despite the higher error of the pseudomeasurements. The computer simulations have shown that, the impact of the increased uncertainty seems to be insignificant due to the enhanced regularization term, i.e., $\|x - x_p\|^2_2$. 

Fig. 5: PDF of the RMSE with $\sigma_1 = 0.03$ p.u.. Without (left graph) and with (right graph) $\mu$PMUs.

Fig. 6: Average RMSE for the 95-bus DG and $\sigma_1 = 0.03$ p.u.. Without (left graph) and with (right graph) $\mu$PMUs.

Fig. 7: Averaged $RMSE_i$ over the realizations for each state variable $i = 1...2N$ with $\sigma_1 = 0.03$ p.u., with and without $\mu$PMUs. The left graph illustrates the voltage phase angles and the right graph the voltage magnitudes.
Fig. 8: Average RMSE for the 95-bus DG and $\sigma_2 = 0.15$ p.u. Without (left graph) and with (right graph) $\mu$PMUs.

Fig. 9: Averaged $RMSE_i$ over the realizations for each state variable $i = 1...2N$ with $\sigma_1 = 0.15$ p.u., with and without $\mu$PMUs. The left graph illustrates the voltage phase angles and the right graph the voltage magnitudes.

TABLE II: RMSE and condition number $\kappa$ of $G$ for each approach, without (top) and with (bottom) $\mu$PMUs, $\sigma_2 = 0.15$ for pseudomeasurements, subsection VI-A.

| Approach | RMSE (p.u.) | $\kappa$          |
|----------|-------------|------------------|
| NESE     | 3.28        | $5.39 \times 10^6$ |
| $\tau | | 0.21        | $4 \times 10^7$    |
| $\tau | | $1.1 \times 10^{-2}$ | $4.42 \times 10^7$ |

| Approach | RMSE (p.u.) | $\kappa$          |
|----------|-------------|------------------|
| NESE     | $8.08 \times 10^{-3}$ | $9.7 \times 10^7$ |
| $\tau | | $5.96 \times 10^{-2}$ | $4.65 \times 10^8$ |
| $\tau | | $8.4 \times 10^{-3}$ | $4.4 \times 10^8$ |

Additionally, for the consistent presentation of the results, we provide a comparison between the proposed scheme R-NESE and NESE, following an alternative approach on the measurements (in-)accuracy. In contrast with the method described above [50], here, the error standard deviation is expressed as percentage of the actual measurement value [10]. Specifically, for each type of measurement we have considered the following accuracy: (i) 1% for voltage measurements, (ii) 3% for branch power flows and current magnitudes, (iii) 50% and 0.02% for pseudomeasurements and zero power injections, respectively, and (iv) 1% for the magnitude and 0.01 rad for the angles in the case of $\mu$PMUs. The standard deviation $\sigma$ has been assumed as equal to a third of the accuracy value in each case. Fig. 10 illustrates the RMSE for each state element $\hat{x}_i$ averaged, as before, over 5000 realizations, considering the same measurement set. The (left) graph reveals that, NESE reaches a maximum $RMSE_i$ value of $3 \times 10^{-3}$ rad, presenting also additional error peaks in a number of buses. In contrast, the maximum $RMSE_i$ for R-NESE is $2 \times 10^{-3}$ rad, while the rest of the estimated angles present low error. In addition, with regard to the voltage magnitudes, R-NESE outperforms significantly in terms of accuracy. On the one hand, NESE reaches an RMSE at the level of $3 - 4 \times 10^{-3}$ p.u. with a maximum value of $8 \times 10^{-3}$ p.u.. On the other, R-NESE presents a maximum $RMSE_i$ of $1.5 \times 10^{-3}$ p.u. while the majority of the estimated voltage magnitudes error stays below $5 \times 10^{-4}$ p.u.. To conclude, independently from the adopted rationale on the measurement noise, the computer simulations have shown the same trend. That is, the proposed routine R-NESE attains high accuracy even in scenarios with low redundant measurement sets which include a large number of high uncertainty pseudomeasurements.

Next, we focus on the computational behavior of R-NESE with emphasis on the number of iterations and absolute time (s) for full convergence. In this computer simulation we have considered the 33-bus and 56-bus distribution systems with a respective low redundancy measurement set, i.e., $L/N \approx 1.1$. For complementary purposes, the 95-bus DG is also presented. Beyond algorithmic accuracy, in Table III it is observed for all test systems a decrease of the number of iterations and in accordance, of the absolute time. For instance, for the 56-bus test system and the case of NESE the algorithm has converged after 7 iterations. In contrast, R-NESE is able to converge in 3 iterations. Consequently, the absolute time has decreased from
0.322 s to 0.121 s. This optimized convergence behavior is linked to the improved $\kappa$. More specifically, a well conditioned system of equations can lead the normal equations framework to a global convergence with quadratic characteristics [54] (see Appendix A).

Relatively with the above analysis, it is provided a comparison between the proposed R-NESE routine and the NESE scheme. With respect to the latter, instead of the commonly used flat start, i.e., $V_i = 1$ p.u. and $\theta = 0$ rad for $i = 1...N$, as initial condition is exploited the last available estimated state, that is, $\hat{x}_0 = x_p$. The computer simulation scenario considers the 95-bus DG and the measurement set already described, including $\mu$PMUs. Fig. 11 shows the $RMSE_i$ of each state variable for both cases. As the graph illustrates, the R-NESE outperforms in terms of accuracy. More specifically, the left graph shows that the RMSE for all voltage phase angles stays below $1 \times 10^{-2}$ rad. In contrast, in the case of NESE the maximum $RMSE_i$ reaches the value of $5 \times 10^{-2}$ rad and, a large number of estimated angles reach a similar level of error. The same applies for the voltage magnitudes (right graph) where the results are quite similar. Based on the above numerical behavior of the compared algorithms, the following conclusion can be drawn. Although the NESE routine initiates from a point close to the optimal one (i.e., $x_p$), the bad conditioning of the optimization problem leads with high probability the algorithm to a non-efficient local minima.

Finally, we examine the computation of the control parameter $\tau$, which establishes the trade-off between regularization and data fidelity. Different methods exist in the literature, such as the generalized cross validation and the L-curve method [42]. The latter appears to provide robustness and consequently, we have adapted this approach. In Fig. 12 we provide the L-curve for an instance of R-NESE (with $\mu$PMUs and $\sigma_1 = 0.03$ p.u.). This method constitutes a graphical tool depicting the error of the regularization term $\|x - x_p\|_2^2$ versus the error of the residual term $R^{-1}_e \|z - h(x)\|_2^2$. In particular, in every iteration $\nu$ of Algorithm 1, the two terms are computed for a number of different $\tau'$ values. As Fig. 12 illustrates, these values correspond to different points of the graph and the common method is to choose the one that falls within the maximum curvature, e.g., $10^1$. 
TABLE III: RMSE, κ of G, Iterations and Absolute Time for the 33-bus (top), 56-bus (middle) and 95-bus (bottom) distribution systems. Scenario without μPMUs and σT = 0.03 for pseudomeasurements.

|          | Approach | RMSE (p.u.) | κ      | Iterations | Time (s) |
|----------|----------|-------------|--------|------------|----------|
| 33-bus   | NESE     | 1.2 × 10^-4 | 1.19×10^9 | 5          | 0.115    |
|          | τ ‖x - x0‖_2^2 | 5.3 × 10^-4 | 1.24×10^9 | 3          | 0.081    |
| 56-bus   | NESE     | 9.3 × 10^-4 | 8.86×10^9 | 7          | 0.3221   |
|          | τ ‖x - x0‖_2^2 | 6.39 × 10^-4 | 2.87×10^9 | 5          | 0.1871   |
| 95-bus   | NESE     | 1.84 × 10^-1 | 8.8 ×10^9 | 10         | 0.5391   |
|          | τ ‖x - x0‖_2^2 | 8.86 × 10^-2 | 3.62×10^9 | 7          | 0.4641   |
|          | τ ‖x - x0‖_2^2 | 4.1 × 10^-3 | 4.41×10^9 | 3          | 0.2031   |

B. D-WTVSE

Here, we investigate the numerical behavior of the D-WTVSE scheme, which operates at the intermediate time instants, that is, between kT and (k+1)T. To recall, all simulations have been averaged over 5000 realizations of different noise and load injections. First, the impact of the control parameter λ on the algorithmic performance of (D)-WTVSE is illustrated. Fig. 13 presents the convergence and the RMSE for a number of different parameter values within the range 10^-7 – 10^-1. The graph shows that lower values are related with slower convergence rate. For instance, for λ = 10^-6 the algorithm attains the highest accuracy. However, more than 50 iterations are needed to converge. In contrast, at the case of λ = 10^-4 the scheme attains the same accuracy within 15-20 iterations. Exhaustive simulations with the trial-end-error method have shown that a parameter value within the range of 10^-5 – 10^-4 balances the trade-off between accuracy and convergence speed. For a further investigation on the control parameter choice a number of works exist in the literature. Similarly with subsection VI.A, the most robust approaches are based on the L-curve method [41] and the generalized cross validation [55]. Interestingly, the former method provides the same results with the trial-end-error simulations. Other approaches can also be found, such the variational Bayes approach [56].

Further, Figs. 14 and 15 depict the convergence behavior and accuracy for the D-WTVSE optimization problem solved via ADMM. Here, just K = 24 measurements (voltage real and imaginary parts) are available from the μPMU configuration proposed in subsection VI-C (Table VIII). In the examined scenario, we have considered the presence of two ILVRs, providing 5% voltage regulation and positioned at the branches {23 – 24} and {54 – 75} (Fig. 3). Fig. 14 indicates that, for this scenario, the program has converged for all areas within maximum 10-15 iterations reaching at the same time high accuracy. Clearly, as the next figure shows (Fig. 15), the estimated magnitudes and angles approach with high precision the actual values, presenting a slight variance in a small number of buses. In both cases of ILVRs, {23 – 24} and {54 – 75}, the estimated voltages present high accuracy with a slight deviation with respect to the first. However, with proper adjustment of the positive constants c1 and c3 (16), a trade-off can be established between ILVR-bus variable consensus and voltage variance limitation.

In addition, for the same scenario the proposed D-WTVSE is compared with WTSE. To do so, WTSE has been modified in order to incorporate the ILVR models based on the augmented admittance matrix approach [34]. Further, the regularization term ||WLx||_1 had to be modified accordingly. The estimation error and the absolute time needed for full convergence for the two methods are presented in Table IV. Interestingly, the proposed D-WTVSE presents better accuracy.
Fig. 13: RMSE vs. iterations for the D-WTVSE ($\lambda = 10^{-7} - 10^{-1}$, $c_1 = 5 \times 10^{-4}$, $c_2 = 1 \times 10^{-5}$, $c_3 = 1 \times 10^{-4}$).

Fig. 14: RMSE vs. iterations for the D-WTVSE scenario ($\lambda = 10^{-4}$, $c_1 = 5 \times 10^{-4}$, $c_2 = 1 \times 10^{-5}$, $c_3 = 1 \times 10^{-4}$).

Fig. 15: Actual and estimate voltage (magnitude and angle) for each bus with the D-WTVSE ($\lambda = 10^{-4}$, $c_1 = 5 \times 10^{-4}$, $c_2 = 1 \times 10^{-5}$, $c_3 = 1 \times 10^{-4}$).
This is illustrated also in Fig. 16 which shows the magnitude error for both cases. More specifically, we observe that the augmented WTVSE attains slightly higher accuracy at the corresponding ILVR buses. However, this is not the case for the rest of the nodal estimated voltages. In addition to that, D-WTVSE performs better in terms of absolute time convergence. This is due to the fact that the algorithm is solved in a distributed manner. To clarify here that, the time illustrated in Table IV refers to the maximum time needed among the three areas (area 1). The decreased time complexity of the proposed scheme offers the opportunity for an increased number of intermediate state estimates (see subsection VI-D). In general, the above analysis reveals that, exploiting the proposed D-WTVSE routine instead of the augmented WTVSE, performance degradation is avoided.

Finally, the efficiency of the regularization term \( \|WLx\|_1 \), which renders the intermediate SE problem solvable, is numerically assessed under three different scenarios: (i) the base case, where we consider the nominal data of the 95-bus UK DG [37]; (ii) a case with PV generation (i.e., PV connected to the buses 15, 36, 68 and 91 with nominal capacity is 0.8, 0.8, 0.5 and 0.4 MW, respectively) added to the load profile; and (iii) a scenario without PV generation and with a number of additional high loads. These are located distant from phasor measured nodes (Table V). The last case can be considered as a worst case scenario where adjacent (non-measured) buses with low/high load (e.g., small/large number of households) absorb low/high current. Fig. 17 presents the estimated and real voltage magnitude profile (upper graph) and the absolute voltage magnitude error (bottom graph) for the case (i). Here, the estimated voltage profile attains the exact one with high accuracy. The bottom graph shows that all absolute errors are below the value of \( 5.5 \times 10^{-3} \) p.u.. In case (ii), Fig. 18, the error stays below \( 1 \times 10^{-3} \) p.u. for the majority of voltage estimates. As it was expected, the high PV penetration has resulted to a smooth voltage profile that consequently, leads the total variation regularization term to superior performance. Last, Fig. 19 shows the results for scenario (iii). The diversity between loads has resulted to a non-smooth real voltage profile. Accordingly, at the bottom graph we detect that the highest deviations with respect to the real voltage magnitudes are located to the buses where the high loads have been positioned. Still, in this worst-case scenario, D-WTVSE is capable of providing a satisfactory state estimate where the absolute error of all the estimate voltages remain below \( 1 \times 10^{-2} \) p.u.. Above all, the simulations have indicated that, the exploitation of the zero power injections and the introduced weights on the branches [32] aid significantly to the D-WTVSE accuracy.

| Approach | RMSE (p.u.) | Time (s) |
|----------|-------------|----------|
| D-WTVSE  | \( 4.8 \times 10^{-3} \) | 0.0393   |
| WTVSE    | \( 5.6 \times 10^{-3} \) | 0.0395   |

| Bus | P (MW) | Q (MVar) |
|-----|--------|----------|
| 12  | 0.6    | 0.05     |
| 27  | 0.7    | 0.05     |
| 46  | 0.9    | 0.1      |
| 59  | 0.7    | 0.08     |
| 27  | 0.65   | 0.05     |
| 46  | 0.95   | 0.15     |
Fig. 17: Actual vs estimate voltage magnitude (upper graph) and absolute voltage magnitude error (bottom graph) for scenario (i) ($\lambda = 10^{-4}$, $c_1 = 5 \times 10^{-4}$, $c_2 = 1 \times 10^{-5}$, $c_3 = 1 \times 10^{-4}$).

Fig. 18: Actual vs estimate voltage magnitude (upper graph) and absolute voltage magnitude error (bottom graph) for scenario (ii) ($\lambda = 10^{-4}$, $c_1 = 5 \times 10^{-4}$, $c_2 = 1 \times 10^{-5}$, $c_3 = 1 \times 10^{-4}$).
Fig. 19: Actual vs estimate voltage magnitude (upper graph) and absolute voltage magnitude error (bottom graph) for scenario (iii) \((\lambda = 10^{-4}, c_1 = 5 \times 10^{-4}, c_2 = 1 \times 10^{-5}, c_3 = 1 \times 10^{-4})\).

C. Numerical Assessment of \(\mu\)PP Algorithm

In this subsection, we evaluate the \(\mu\)PP method presented in Section V. To do so, we consider three different MV DG systems, namely, a 33-bus, a 56-bus and the 95-bus UK DG that has been already considered for the SE computer simulations. In addition, in order to measure the impact of different \(\mu\)PMU configurations, the following metrics are adapted: (i) the condition number \(\kappa\) of the gain matrix \(G_{\text{sol}}\) after the convergence of the program \((30)\); (ii) the \(\log(\kappa)\) and (iii) the averaged RMSE of the SE algorithm \((3)\).

First, the proposed algorithm is evaluated considering the 33-bus DG, Table VI. For this case, we have assumed a singular gain matrix \(G_0 \in \mathbb{R}^{66 \times 66}\) (non-observable system) and a limited number of 4/6/8 \(\mu\)PMUs to be placed. The table depicts the results for \(G_0\), two random \(\mu\)PMU placement configurations for comparison reasons \((G_1, G_2)\) and the solution \(G_{\text{sol}}\) according to Section V. The results have shown that the proposed method has managed to provide a full-rank gain matrix. More importantly, the provided solution is the optimal one in terms of conditioning \((\kappa)\) and consequently in terms of the desired SE accuracy. For all cases of different \(\mu\)PMU numbers and compared to \(G_1\) and \(G_2\), \(G_{\text{sol}}\) presents a decreased condition number \(\kappa\) at the level of one order of magnitude. For instance, comparing \(G_{\text{sol}}\) and \(G_1\) for the scenario of 6 \(\mu\)PMUs, Table VI shows that the former presents an improved \(\kappa\), i.e., \(1.3 \times 10^6\) compared to \(1.5 \times 10^7\), that corresponds to a decreased estimation error, i.e., \(1.9 \times 10^{-2}\) p.u. compared to \(3.3 \times 10^{-2}\) p.u.. A similar improvement of the numerical accuracy can been also observed for the other cases of \(\mu\)PMU numbers.

Further, we investigate the \(\mu\)PP algorithmic performance on the 56-bus DG, which constitutes a rural feeder [57]. In this case, the system includes a number of short, as well as long branches, something that deteriorates the conditioning of \(G\) [16]. Table VII indicates that the addition of a moderate number of synchrophasor meters, improves significantly the condition number \(\kappa\) of \(G_0\) \((1.2 \times 10^{10})\). In particular, in the case of random trials \(G_1\) and \(G_2\), \(\kappa\) has been improved by two orders of magnitude. In contrast, the proposed method has reduced the condition number by three orders of magnitude, i.e., \(3.6 \times 10^7\) compared to \(1.2 \times 10^{10}\), something that is translated into a lower estimation error, i.e., from \(9.5 \times 10^{-2}\) p.u. to \(1.6 \times 10^{-2}\) p.u..

Finally, for completeness, in Table VIII is presented the optimal solution for the 95-bus DG, which has been considered as the \(\mu\)PMU configuration in subsections VI-B and VI-D. Again, the \(\mu\)PP method provides a gain matrix \(G_{\text{sol}}\) with an improved \(\kappa\) compared to \(G_0\), \(G_1\) and \(G_2\) (three and one orders of magnitude lower, respectively). Therefore, this results into a decreased estimation error, e.g., from \(1.9 \times 10^{-1}\) p.u. to \(9 \times 10^{-3}\) p.u.
TABLE VI: Impact assessment of the μPP algorithm presented in Section V for the 33-bus test system and different number of μPMUs.

| Gain matrix | Number of μPMUs | Buses | \( \kappa \) | \( \log(\kappa) \) | RMSE (p.u.) |
|-------------|----------------|-------|-----------|----------------|-------------|
| \( g_0 \)  | 0              | –     | –         | \( 1n \)       | \( 1n \)     |
| \( g_1 \)  | 4              | 1, 9, 19, 29 | 2.9\times10^1 | 17.2 | 4.8\times10^{-2} |
| \( g_2 \)  | 4              | 2, 10, 19, 31 | 2.4\times10^1 | 17   | 4.4\times10^{-2} |
| \( g_{out} \) | 4            | 2, 14, 26, 32 | 2.3\times10^0  | 14.8 | 2.9\times10^{-2} |

TABLE VII: Impact assessment of the μPP algorithm presented in Section V for the 56-bus test system.

| Gain matrix | Number of μPMUs | Buses | \( \kappa \) | \( \log(\kappa) \) | RMSE (p.u.) |
|-------------|----------------|-------|-----------|----------------|-------------|
| \( g_0 \)  | 0              | –     | –         | \( 1n \)       | \( 1n \)     |
| \( g_1 \)  | 8              | 1, 6, 11, 19, 25, 33 | 1.5\times10^1 | 16.5 | 3.4\times10^{-2} |
| \( g_2 \)  | 8              | 1, 7, 12, 19, 26, 31 | 6.4\times10^1 | 15.7 | 2.5\times10^{-2} |
| \( g_{out} \) | 8            | 2, 7, 11, 16, 24, 32 | 1.3\times10^0  | 14.1 | 1.9\times10^{-2} |

TABLE VIII: Impact assessment of the μPP algorithm presented in Section V for the 95-bus test system.

| Gain matrix | Number of μPMUs | Buses | \( \kappa \) | \( \log(\kappa) \) | RMSE (p.u.) |
|-------------|----------------|-------|-----------|----------------|-------------|
| \( g_0 \)  | 0              | –     | –         | \( 1n \)       | \( 1n \)     |
| \( g_1 \)  | 8              | 5, 12, 20, 25, 34, 42, 50, 54 | 2.9\times10^1 | 19.5 | 1.9\times10^{-2} |
| \( g_2 \)  | 8              | 1, 10, 19, 26, 44, 43, 57, 56 | 9.4\times10^0 | 18.4 | 1.8\times10^{-2} |
| \( g_{out} \) | 8            | 6, 13, 21, 25, 28, 49, 46, 50 | 3.6\times10^1  | 17.4 | 1.6\times10^{-2} |

D. Full scheme

Finally, we evaluate the ability of the proposed SE scheme to track short time-scale voltage variations introduced from the stochastic behavior of the PV plants. The time-sequence diagram in Fig. 20 illustrates the operation of the two-time scale SE scheme based on the proposed R-NESE and D-WTVSE algorithms. A sample of the aggregated power generation profile under consideration for the set of 4 PV plants is shown in Fig. 21. The profile refers to the PV generation, acting on June 27, 2018, between the hours of 08:20 and 08:50 am, at a MV feeder located to the central part of Greece. The active power measurements have a resolution of 6 seconds. Further, for each active/reactive load-bus of the 95-bus UK DG has been adapted a real load pattern which refers to the same day and time period, Fig. 22. The performance of the proposed SE scheme (along with that of two benchmarks) is shown in Fig. 23 next. In all cases, the estimation error is defined again as the RMSE (31), computed by either the R-NESE or D-WTVSE schemes.

Specifically, the 'R-NESE & D-WTVSE' curve presents the estimation error associated to a SE running (i) the R-NESE scheme with both legacy and μPMU measurements at \( kT \) time instants (we let \( T = 15 \)); and (ii) the proposed D-WTVSE scheme at the intermediate ones (we let \( q = T/t = 8 \), for illustrative purposes). The 'R-NESE-last' curve provides a benchmark for the case where no SE scheme is used at intermediate time instants. Hence, we compute the estimation error between the actual system state at time \( kT + nt \) and the latest available R-NESE estimate computed at \( kT \) (on the basis of both legacy and μPMU measurements). The 'R-NESE-updated PMUs' curve depicts the achievable error at intermediate time instants with the outdated legacy measurements collected at \( kT \) (to preserve observability) and the updated μPMU readings gathered at \( kT + nt \). Besides, a set of 12 μPMUs has been incorporated to the legacy measurement set. According to the configuration proposed from the μPP algorithm, the meters have been placed at buses 1, 8, 21, 32, 37, 50, 64, 66, 73, 75, 84 and 93.

Fig. 23 indicates that the estimation error for 'R-NESE-last' is the highest. This follows from the fact that, the last available state estimate is outdated. At the same time the system voltage profile has been differentiated according to the PV power variations. Indeed, the larger the deviation of the power generated by the PV plants at \( kT + nt \) (see Fig. 21), the larger the
estimation error (e.g., at $T + 4t$). Instead, by replacing the outdated $\mu$PMU measurements by updated ones (‘R-NESE-updated $\mu$PMUs’ curve), performance significantly improves. This proves the large impact that $\mu$PMUs have on the computation of the system state. However, the proposed ‘R-NESE & D-WTVSE’ scheme clearly outperforms both benchmarks. As the graph shows, the proposed scheme reaches high accuracy in all time instants independently of the rapid variations of the PV power generation. Besides, the graph shows that the main time instants solution, i.e., R-NESE, presents lower estimation error compared to D-WTVSE. The latter comes from the exploitation of the regularization term incorporating prior information. As a general remark, computer simulations show that the proposed SE routine can constitute a reliable monitoring tool, capable of tracking short-time voltage variations at the MV system.

VII. CONCLUSIONS

In this paper, we have proposed a regularized SE scheme for the robust monitoring of the distribution grid. The algorithm consists of two different time-scale parts; a robust state estimator that operates on the regular time instants (every 15 minutes) and a regularized SE scheme for the intermediate time period. The final goal was to accurately track the system state at a faster time scale with increased reliability, according to the needs of the new operational environment. With respect to the main time instants, the estimator has been formulated as the regularized version of the normal-equations based SE solution (R-NESE). The computer simulations (on the 95-bus UK DG) have shown that R-NESE presents robustness against the high uncertainty in the active power profile.
raised by the low redundancy measurement sets and the pseudomeasurements. Moreover, the enhanced regularization term led the algorithm to increased accuracy and improved convergence rate compared with other approaches. For the intermediate time instants, we have presented a decomposed weighted total variation state estimation (D-WTVSE) model which is solved by means of the Alternating Direction Method of Multipliers (ADMM). The problem has been decomposed according to the ILVR locations across the feeder, exploiting further information from the tap changer indicators. As expected, the scheme presented a slightly diminished performance compared to the main time instants estimator. Nonetheless, the extended simulation results have proven that D-WTVSE is able to provide an accurate state estimate under a diversity of scenarios. Further, the low computational complexity of the algorithm provides the opportunity for multiple state estimates within the intermediate time interval. Complementarily, we have presented a $\mu$PMU placement ($\mu$PP) method, which has been posed as a mixed integer semidefinite programming (MISDP) model. We have observed that the proposed approach is able to provide a full-rank gain matrix while taking into consideration the existing measurements. More importantly, the presented solution improved significantly the conditioning of the SE problem and the estimation accuracy. Additional constraints related with the number and location of the $\mu$PMUs were also satisfied.

**Appendix A**

**Computational Complexity Analysis**

**A. R-NESE**

Let us first consider the computational complexity of the regularized normal equations (R-NESE) algorithm (10), presented in subsection III-A. To do so, we will leverage on $\mathcal{O}(\cdot)$ notation, that from now on corresponds to per iteration complexity. In addition, for the ease of notation, let us assume $n$ state variables and $k$ measurements. The main computational burden in:

$$\Delta x^{(\nu)} = \left( J_x^{(\nu)} R_e^{-1} J_x^{(\nu)} + \tau^{(\nu)} I \right)^{-1} J_x^{(\nu)} R_e^{-1} [z - h(x^{(\nu)})] - \tau^{(\nu)} (x^{(\nu)} - x_p)$$

entails to a number of matrix multiplications and the matrix inversion. The former can be formed at a cost of $\mathcal{O}(kn^2)$. On the other hand, the matrix inversion usually takes place with a Cholesky factorization at a cost of $\mathcal{O}(n^3)$. Taking into account that in distribution grid SE $n \approx k$, the overall computational complexity of R-NESE scales at $\mathcal{O}(n^3)$. Nevertheless, exploiting the inherent sparsity of $G^{(\nu)}$ this cost can be reduced at $\mathcal{O}(n^2)$ [58], [59].

With regard to the convergence rate of the algorithm, the normal equations based schemes commonly present quadratic convergence characteristics [54]. However, this is strongly dependent from the condition number $\kappa$ of the gain matrix $G^{(\nu)}$ in each iteration. According to the literature [60], in the case of a well-conditioned system of equations a few number of iterations are sufficient for the full convergence of the algorithm.

**B. D-WTVSE**

The computational complexity analysis for the D-WTVSE algorithm follows the same rationale. First, consider the ADMM-based solution presented in subsection IV. One can observe that the update step of $x^{\nu+1}$ (28) dominates the computational
burden of each ADMM iteration. More specifically, the computation of:

\[
x_r^{\nu+1} = \left( H_r^T H_r + c_1 L_r^T W_r L_r + c_2 \bar{Y}_r^T \bar{Y}_r + \sum_{s \in A_r} M_{rs}^T M_{rs} \right)^{-1} \left( H_r^T \bar{Y}_r - L_r W_r \mu_r + c_1 L_r^T W_r \bar{\theta}_r - \bar{Y}_r \pi_r + \sum_{s \in A_r} (M_{rs}^T c_{rs} - \kappa_r M_{rs}^T) \right) \]

(34)

consists of a number of matrix multiplications and mainly a matrix inversion. As all matrices involved in the computations are characterized from sparsity, again the overall computational complexity has a final cost of \(O(n^3)\). Important to note here that, in contrast with R-NESE where the gain matrix \(G^{(\nu)}\) has to be re-computed in each iteration, the implementation of D-WTVSE can exploit one additional advantage. In the case where the control parameters \(c_1, c_2, c_3\) have constant values during the ADMM iterations, the factorization of the matrix inversion can be cached in order to reduce significantly the computational complexity [35].

Concerning the convergence rate of the algorithm, a number of previous works have proved that ADMM presents a linear convergence behavior [61]. That means, the number of iterations needed for a full convergence grows linearly with the size of the problem. However, in the medium voltage SE case, where the number of the state variables can be characterized as moderate, a few decades of iterations are needed. Again, the convergence behavior of this method strongly depends from the conditioning of the problem on the update of \(x_r^{\nu+1}\). That given, an improved condition number of the first part in (34) will lead to a decreased number of iterations.

REFERENCES

[1] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan and F. Bu, "A Survey on State Estimation Techniques and Challenges in Smart Distribution Systems," in IEEE Transactions on Smart Grid, vol. 10, no. 2, pp. 2312-2322, March 2019.

[2] A. Primidianto and C. Lu, "A Review on Distribution System State Estimation," in IEEE Transactions on Power Systems, vol. 32, no. 5, pp. 3875-3883, Sept. 2017.

[3] C. Muscas, M. Pau, P. A. Pegoraro and S. Sulis, "Smart electric energy measurements in power distribution grids," in IEEE Instrumentation & Measurement Magazine, vol. 18, no. 1, pp. 17-21, February 2015.

[4] D. Della Giustina, M. Pau, P. A. Pegoraro, F. Ponci and S. Sulis, "Electrical distribution system state estimation: measurement issues and challenges," in IEEE Instrumentation & Measurement Magazine, vol. 17, no. 6, pp. 36-42, December 2014.

[5] A. Gómez-Expósito, C. Gómez-Quiles and I. Džafić, "State Estimation in Two Time Scales for Smart Distribution Systems," in IEEE Transactions on Power Systems, vol. 29, no. 3, pp. 1638-1647, Aug. 2014.

[6] M. E. Baran and A. W. Kelley, "A branch-current-based state estimation method for distribution systems," in IEEE Transactions on Power Systems, vol. 10, no. 1, pp. 483-491, Feb. 1995.

[7] Wenchuan Wu, Yuntao Ju, Boming Zhang, Hongbin Sun, A distribution system state estimator accommodating large number of ampere measurements, International Journal of Electrical Power & Energy Systems, Volume 76, 2015, Pages 498-507, ISSN 0142-0615.

[8] A. Gómez-Expósito, C. Gómez-Quiles and I. Džafić, "State Estimation in Two Time Scales for Smart Distribution Systems," in IEEE Transactions on Power Systems, vol. 29, no. 3, pp. 1638-1647, Aug. 2014.

[9] E. Manitsas, R. Singh, B. C. Pal and G. Strbac, "Distribution System State Estimation Using an Artificial Neural Network Approach for Pseudo Measurement Modeling," in IEEE Transactions on Power Delivery, vol. 29, no. 1, pp. 310-319, Feb. 2014.

[10] A. Gómez-Expósito, C. Gómez-Quiles and I. Džafić, "State Estimation in Two Time Scales for Smart Distribution Systems," in IEEE Transactions on Power Systems, vol. 29, no. 3, pp. 1638-1647, Aug. 2014.

[11] A. Gómez-Expósito, C. Gómez-Quiles and I. Džafić, "State Estimation in Two Time Scales for Smart Distribution Systems," in IEEE Transactions on Power Systems, vol. 29, no. 3, pp. 1638-1647, Aug. 2014.
