Factorized Solution of Power System State Estimation

DOCTORAL THESIS

by

Catalina Gómez Quiles

Supervisors:

Antonio de la Villa Jaén

Antonio Gómez Expósito

Departamento de Ingeniería Eléctrica

Escuela Superior de Ingeniería

Universidad de Sevilla

Sevilla, España

March 2012

A mi familia, por su apoyo incondicional.

Acknowledgments

I would like to thank each and everyone who contributed to make this work possible and those who have taken part in this step in my life. In particular, to my family, and especially my parents, for their unconditional support, deep affection and guidance.

To my supervisors Antonio de la Villa Jaén for his consideration and perseverance and Antonio Gómez Expósito for his optimism and determination.

To my fellow research and lecturing colleagues, to our professors and administrative staff who make our Department an enjoyable place.

Finally, to the Pan-European Grid Advanced Simulation and State Estimation (PEGASE) project of the European Union and to the TALENTIA Scholarship Program (Programa de Becas Talentia, Junta de Andalucía - Consejería de Economía, Innovación y Ciencia), government of Andalusia, Spain, and the research project funds 'Plan Nacional' ENE2007-62997 and ENE2010-18867 for the financial support.

Contents

List of Tables

List of Figures

Chapter 1

Introduction, Context and Objectives

1.1 Energy management systems and state estimators

State Estimators (SE) determine the most likely state of a power system from sets of remotely captured measurements which are collected periodically by SCADA systems via Remote Terminal Units (RTU). The role of the SE is crucial in modern Energy Management Systems (EMS), where a diversity of applications rely on accurate system snapshots [58]. The regulatory wave of the last decades has stressed the importance of the SE tool, in an open-access context in which many more transactions on much more congested networks have to be properly tracked in real time and also recorded for off-line engineering studies.

Currently, the scope of SEs is mostly limited to the transmission level, where each Transmission System Operator (TSO) continuously tracks its own grid from a centralized EMS (see Fig. 1.1). However, in some power systems, trading and pricing may also take place at lower voltage levels, which are typically not closely monitored by a state estimator. This is mainly due to the unavailability of sufficiently redundant set of measurements at these voltage levels. Hence, there is a need to improve the

monitoring capability for these parts of the systems and to facilitate reliable and effective operation of power markets at these levels as well.

At the uppermost voltage levels, where interconnections for energy trading exist, tie-line power flows have to be properly monitored, for which TSOs must get access to both the electrical model and real-time measurements of its neighbors, at least those in the adjacent buses. For this purpose, the external grid is usually represented by a reduced equivalent circuit [45]. Therefore, existing SEs, being tailored to the needs of a single TSO, with very neat borders, are not designed to significantly interact with its neighbors or subordinate networks. As explained later, this is not a satisfactory state of affairs, in view of both the needs and possibilities offered by smart grids, but commercial software evolves at a lower pace than theoretical developments.

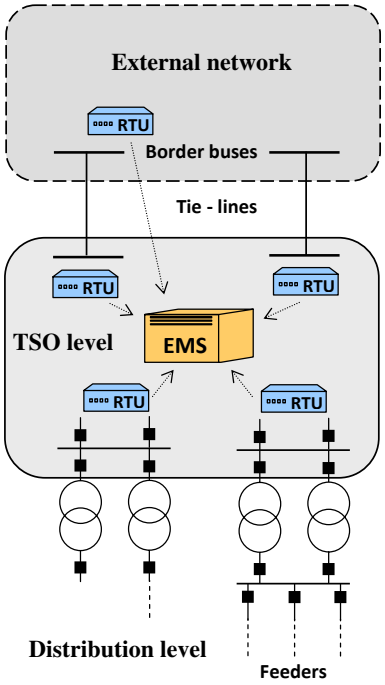


Figure 1.1: Conventional TSO-level State Estimation paradigm.

1.2 Future smart grid context

The SE paradigm described will have to drastically change with the advent of the smart grid. On the one hand, new generations of digital devices, such as Phasor Measurement Units (PMU) or Intelligent Electronic Devices (IED), intended for measurement, protection and control, being less expensive and more flexible than existing analog equipment, will invade virtually every corner of future networks [48, 37]. This will provide a more accurate, complex and highly redundant information system, allowing SEs to extend their scope well beyond presently observable areas and also to incorporate advanced functions that have not yet reached the industrial stage, in spite of being conceptually mature on the researcher blackboard. On the other hand, smart transmission grids should further promote the development of regional energy markets, involving distant energy transactions. This also implies wide-area physical interactions, possibly of catastrophic consequences in case of cascaded failures [20, 18].

In the following subsections, outstanding technological innovations associated with the smart grid concept will be succinctly analyzed from the point of view of their influence in the conceptual design of future SE architectures.

1.2.1 Digital substations

The protection, metering and control functions in substations are naturally distributed by the role and location of each device, being designed in general to provide primary protection or monitoring of an individual substation equipment. These functions may be performed by smart multi-functional and communicative units, so-called IEDs. They are broadly defined in [28] as “devices incorporating one or more processors with the capability to receive or send data/control signals from or to an external

source (e.g., electronic multifunction meters, digital relays, controllers)".

The IEDs, employing efficient signal processing techniques, are becoming the source of much more information in real time than the one existing in old substations. Apart from implementing specific protection or control algorithms, they can provide externally electrical magnitudes measured by protection transformers as well as phase differences among them [37]. Those measurements can be synchronized, both at the substation and wide-area levels, by means of the Global Positioning System (GPS) satellite clock time reference.

The quality of the SE process strongly relies on the redundancy of the measurement set. For this reason, the possibility of incorporating all the information provided by IEDs, including the ones whose primary function is not measurement but protection, is very attractive.

New system architectures need to be devised for pre-processing the ever-increasing amount of information gathered by the IEDs. This goal is achieved by replacing the conventional centralized systems, based on RTUs and numerous protection and control devices, with local area network based systems and advanced multifunctional protection and control IEDs [10].

IEC 61850, the global communication standard for Substation Automation System, defines the communication between IEDs and not only solves the interoperation problem but also specifies other system requirements, like message performance and information security in Substation Automation System network [10]. IEC 61850 allows interoperability of IEDs from different manufacturers without the use of protocol converters.

The standard defines two communication buses between the different subsystems

within a substation (see Fig. 1.2). The process bus is devoted to gathering information about electrical magnitudes, such as voltage or current, as well as switching status information, from the transformers and transducers connected to the primary power system process. The station bus is aimed at allowing primary communications between the Logical Nodes, which provide the various station protection, control, monitoring, and logging functions. The communication technologies involved in these buses include: Ethernet on fiber optic, TCP/IP and MMS (ISO9506). This architecture supports remote network access for all types of data reads and writes.

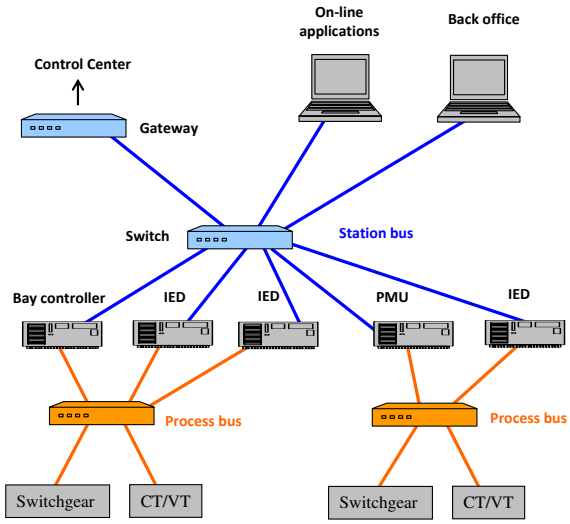


Figure 1.2: Substation-level hardware platform, according to IEC 61850.

1.2.2 Synchrophasors

Most of the energy management system applications assume that the system is in a pseudo steady state where a.c. circuit analysis can be carried out using phasors. Network model and the voltage phasors at all system buses are used to determine the state of the system. While the bus voltage phasors can be estimated based on

a redundant set of measurements, direct measurement of phasors will be possible only if measurements are synchronized. Phasor measurement units are devices which take advantage of the GPS Satellites in order to time synchronize the measurements. Voltage and current signals are collected at secondaries of instrument transformers and are sampled via analog to digital converters at 48 samples/cycle. These samples are then processed and synchronized with Universal Time Coordinated (UTC) time from a GPS receiver within 1 microsecond accuracy [48]. Time synchronized samples are processed to obtain time-stamped voltage and current phasors, which are then transmitted over Ethernet to Phasor Data Concentrators (PDC), which will then send data to control center SCADA server. IEEE Standard C37.118 describes the requirements, format and communication protocol for data provided by PMUs [29].

PMUs may have several channels, each of which will record one phase of a voltage or current signal. Two sets of three channels are typically used for three phase voltage measurements at the substation and several sets of three channels will be used for measurement of three phase currents along incident lines or transformers. Positive sequence components rather than individual phase signals are typically used by network applications, hence three phase signals are processed to compute the positive sequence components. The string of positive sequence phasors that are computed by the PMU will then be communicated at 30 samples/sec to the PDC.

1.2.3 Active distribution systems

Distribution automation is a mature concept whose real potential was never realized due to the lack of reasonably priced infrastructures. In fact, unlike in transmission systems, most functions in this field (fault detection, service restoration, network

reconfiguration, etc.) have been traditionally performed with the help of mobile service teams on call.

Regarding the information that can be found at the distribution level, virtually no measuring devices have been installed until recently to monitor the operating condition of medium voltage (MV) feeders. Typically, the current (or sometimes the power flow) at each feeder head, along with the voltage magnitude at the MV bus, are telemetered and gathered at dedicated Distribution Management Systems (DMS). But no real-time information is obtained on what happens downstream, unless a fault occurs. The substation bus voltage is kept almost constant by the use of automatic under-load tap changers, in the hope that customer voltages remain acceptable for nearly all operating points.

This situation is rapidly changing for several reasons, including:

- Distributed generators (DG) connected at this level frequently reverse the sign of power flows, creating overvoltage problems that should be properly monitored and prevented. On the other hand, the energy they inject, frequently at premium prices, should also be carefully monitored and recorded.

- Smart meters, currently being deployed, provide hourly customer demands via Power Line Carrier (PLC), regular cellular phone technology or alternative means. There is a trend to concentrate all this information at the secondary transformer centers, from where it will then be submitted upstream to the distribution substation or DMS.

- Cheap fault current detectors, along with automatic or remotely operated reclosers, are being installed at strategically selected points to speed up the service

restoration process. These devices are capable of providing less accurate current values that could also be attached to the remaining information sent to the substation.

Consequently, the massive introduction of DG and a plethora of distribution automation devices, at network levels which are not currently supervised by TSOs, will also contribute to the development of ubiquitous monitoring systems [60].

1.2.4 Interconnected transmission systems

Energy markets outcomes may be significantly affected by unique information regarding the present and likely states of the grid. But gathering such information is a real challenge when the energy transactions take place over networks that cross national or regional market borders. For this and other strategic reasons international regulatory entities are promoting worldwide the creation of regional-level system operators, in an attempt to eliminate barriers and better coordinate multi-TSO transactions.

In the USA, for instance, realizing that competition was hindered because only a handful of utilities owned and controlled a large portion of the region's transmission, the Federal Energy Regulatory Commission issued the Order 890 "Preventing Undue Discrimination and Preference in Transmission Service" and the 2010 Notice of Proposed Rulemaking on "Transmission Planning and Cost Allocation by Transmission Owning and Operating Public Utilities". These orders amended its regulations in order to remedy opportunities for this undue discrimination and address deficiencies in the pro forma open access transmission tariff. Enabling interconnection-wide operation via these rules that extend the local, regional and inter-regional planning processes, will be greatly facilitated by the proposed multi-level state estimation

scheme.

Across the Atlantic, the European Regulators' Group for electricity and gas launched in 2007 an initiative to create a series of Regional Energy Market projects within the EU, in order to remove barriers to cross-border trade between countries as a first step towards the completion of a single EU market for electricity [19]. Other long-term ambitious projects for long-distance bulk transmission of renewable energy, such as the Desertec initiative [17] will further stress the need for transnational cooperation in network monitoring.

1.3 Objectives of the thesis

The initial objective of this work was to develop, implement and test a computationally efficient algorithm for multi-area state estimation, capable of providing optimal estimates of very large-scale interconnected grids, such as the European Extra-High Voltage (EHV) network, based on existing TSO-level estimation results and requiring a minimum amount of information exchange. The algorithm was intended for integration into the FP7 PEGASE project prototypes.

This goal counted on the vast experience of the Power Engineering group at the University of Sevilla in substation-level estimation [25, 65, 64]. The starting point was the, then recently introduced, two-step state estimator aimed at linearly preprocessing raw measurements at the substation level followed by the conventional Weighted Least Squares (WLS) nonlinear estimator at the TSO level [26]. That work provided for the first time an alternative and rigorous formulation of the two-level SE by factorizing the solution of the nonlinear WLS problem into two subproblems, the first of them linear and geographically distributed.

It was soon realized that the structure of the problem faced in the PEGASE project was, at a higher hierarchical level, very similar to that of [26], except for the fact that the first stage was nonlinear and the second one linear (just the opposite to [26]). This motivated the development of a general framework by which the nonlinear WLS problem is factorized into two sequential subproblems, both of them nonlinear. The nice feature of this formulation is that existing or future hierarchical SE problems constitute particular cases of this general decomposition scheme, in which one of the subproblems can be linear. In addition to the two motivating cases (the substation-level followed by the TSO-level SE, on the one hand, and the PEGASE multi-area problem on the other), a hierarchically lower scenario was also identified and addressed, namely the one arising in distribution substations, where the smart grid technology will render the distribution feeders observable. Overall, the factorization-based scheme devised in this thesis allows seamless integration of multilevel SE, accomplishing hierarchical monitoring of very large-scale interconnected power systems. Incorporation of PMU outputs (complex voltages and currents measurements) has been considered in the multi-area context.

Almost at the same time, another fruitful line of research arose from this theoretical framework, namely the possibility of factorizing, with the help of suitable auxiliary variables, the solution of the conventional WLS SE into two subproblems, the first of them linear. Unlike in the hierarchical SE case, the objective was not to perform a geographical decomposition, but to reduce a nonlinear problem into a sequence of simpler and less expensive problems, potentially enjoying better convergence rates.

The last major development of this work has to do with a simple modification of the initialization procedure associated with the nonlinear stage of the factorized

SE, leading to an extremely fast three-stage SE procedure (two linear WLS problems with a nonlinear transformation in between) which attains convergence for practical purposes in a single execution. Generalization to the equality constrained formulation has been also addressed.

1.4 Structure of the thesis

This dissertation has been prepared in accordance with the new regulations of the University of Sevilla, allowing theses which are backed up by several journal papers to be reduced to an extended summary, stressing the main results and contributions of the research, followed by an annex with all related papers. Therefore, unlike in the past, the reader will not find in the body of this document an exhaustive account of all the developments and results obtained during the course of this work. Instead, he/she is referred to each of the accompanying papers for all the details.

For the sake of self-sufficiency and clarity, however, the thesis is organized in six chapters, each summarizing what is usually covered by one or several chapters in the conventional format. Chapter 2 briefly reviews the state of the art on state estimation. In Chapter 3, a two-stage generalized factorization scheme for state estimation is presented. This approach, in combination with geographical decomposition techniques, is applied in Chapter 4 to hierarchical SE for multilevel environments (from distribution feeders to huge regional systems). Chapter 5 introduces a special set of auxiliary state variables allowing the two-stage WLS factorization to be applied to conventional SE in such a way that the first step reduces to a linear problem. This is further pursued to obtain a very efficient three-stage bilinear decomposition. Both the unconstrained and equality-constrained SE are considered. Chapter 6 concludes

the body of this document, summarizing the main contributions of the thesis. The ideas to be explored in future works are also suggested. The Appendix collects the set of published or submitted papers directly related with this work [T1] - [T7], where the different methods discussed in the thesis are presented in full detail, including numerical examples illustrating the performance of the proposed schemes.

Chapter 2

Power System State Estimation: State of the Art

2.1 A brief historical perspective

Since its introduction by F. Schweppe, in the late sixties [54], the SE tool has benefited from a large number of theoretical developments and practical improvements, which are well documented in [53, 42, 3].

In addition to the pioneering work of Tinney and others on sparsity methods [59], there have been computational improvements such as the so-called Fast Decoupled State Estimator (FDSE), proposed by Monticelli et. al. [43], based on the successful Fast Decoupled Load Flow concept [57]. Another significant improvement was introduced in order to address the issues of numerical stability and ill-conditioning of the conventional WLS approach [67]. It was observed that the use of artificially high weights for very accurate measurements such as zero injections and rather low weights for much less accurate pseudo-measurements would lead to poor convergence of the so-called Normal equations. It was shown that employing a computationally more expensive method of orthogonal or QR factorization would significantly improve the solution accuracy compared to the less stable Cholesky factorization scheme [63].

Alternatively, it was shown that these very accurate or exact measurements, such as zero injections, could be incorporated as explicit equality constraints into the optimization problem formulation, effectively eliminating the need to use artificially high weights [21]. One such formulation used the so called Hachtel's augmented matrix approach, which could be implemented either directly [22] or using 2×2 pivots in combination with block arithmetic for improved computational efficiency [47, 5].

Network observability and bad data processing constitute two important functions related to the SE problem. Observability analysis is performed in advance in order to determine if the entire state vector is observable and, if not, to identify observable islands. Both numerical [46] and topological [34] algorithms, have been proposed and implemented [15]. Strongly related to the network observability analysis was the optimal measurement design which would ensure full network observability under credible loss of RTUs or communication channels [36]. In many instances, the measurement set might be corrupted with gross errors (outliers) and thus, the assumption that all measurement errors were Normally distributed would no longer be true. In such cases, if those gross errors (bad data) were not detected and removed by simple plausibility tests before the execution of the SE, the solution would be biased or even the algorithm could fail to converge. Hence, statistical tests such as the chi-squares test and largest normalized residual test, based on chi-squares and standard Normal distributions respectively, were developed in order to detect and identify bad data. Both tests relied on calculated measurement residuals once the SE algorithm converged [44]. More elaborate techniques, such as hypothesis testing identification (HTI), have also been proposed to handle cases involving multiple interacting bad data for which other methods were less effective [39].

There were numerous other developments which addressed a wide spectrum of issues ranging from statistical robustness of estimation [12, 38], hierarchical multi-area estimation [61], inclusion of Ampere measurements at the subtransmission level [52, 2, 23], incorporation of inequality constraints [16, 55], detection and identification of network parameter and topology errors [56, 41].

In the last decade there has been an increased interest in the so-called Generalized State Estimator (GSE), aimed at developing circuit breaker models to improve the SE capability for topology error processing [42]. This involved a detailed physical-level modeling of bus sections, that should be used in combination with zooming techniques in order to cope with the huge size of the resulting model [4, 40]. An implicit GSE model has been recently developed. The model maintained the capability to identify topology errors, while using a slightly-augmented state vector [65]. The same idea could be used to detect and identify network parameter errors [31].

More recently, PMU devices are expected to introduce major improvements in SE performance and capabilities [48, 49].

The most important developments, from the point of view of this work, will be reviewed in more detail in what follows.

2.2 Maximum likelihood estimation and the Normal equations

Given the following measurement equation [3]:

$$z = h(x) + e \tag{2.1}$$

where:

x is the state vector to be estimated (size $n = 2N - 1$, N being the number of buses),

z is the known measurement vector (size $m > n$),

h is the vector of functions, usually nonlinear, relating error free measurements to the state variables,

e is the vector of measurement errors, customarily assumed to have a Normal distribution with zero mean and known covariance matrix R,

the WLS estimator provides the maximum likelihood estimation by minimizing the following scalar function:

$$J = r^T W r = \sum_{i=1}^{m} W_i r_i^2 \tag{2.2}$$

where:

$r = z - h(\hat{x})$ is the measurement residual,

$\hat{x}$ is the estimated state vector, and

$W = R^{-1}$ is the weighting matrix.

When errors are independent R is a diagonal matrix with values σ_i^2, where σ_i is the standard deviation of the error associated with measurement i.

The minimum of the scalar J can be obtained by iteratively solving the so-called Normal equations:

$$G_k \Delta x_k = H_k^T W [z - h(x_k)] \tag{2.3}$$

where:

$H_k = \partial h/\partial x$ is the Jacobian evaluated at $x = x_k$,

$G_k = H_k^T W H_k$ is the gain matrix,

$\Delta x_k = x_{k+1} - x_k$, k being the iteration counter.

Iterations finish when an appropriate tolerance is reached on Δx_k. The covariance of the estimate is given by:

$$\text{cov}(\hat{x}) = G_k^{-1}$$

2.3 Observability analysis

Observability analysis determines if a state of the entire system can be estimated given the set of available measurements. For a given network topology, the observability of the system does not only depend on the number of measurements, but also on the type and location. In an initially fully observable system, telecomunnication errors, meter failures or topology changes may lead to isolated observable islands with independent phase angle references. Therefore, this analysis must be carried out each time the structure of the measurement set or the network topology get modified. When the system is not fully observable, additional meters or pseudo-measurements may have to be used at specific locations in order to estimate the operating state for the complete system.

Both numerical and topological algorithms, which are out of the scope of this work, have been developed to analyze observability (see [15]).

2.4 Bad data detection and identification

The assumption that measurement errors have a Gaussian (Normal) distribution is not always acceptable, as some of the available measurements may contain large errors due to various reasons such as telemetry failures, interferences in the communication systems, biased meters, etc. Although some of these errors are obvious and can be eliminated prior to the estimation process, it cannot be discarded that non Gaussian errors remain in the measurement set, leading to a biased solution or even to convergence failure.

Statistical tests have been developed in order to detect and identify bad data. In WLS state estimator, the largest normalized residual test is by far the most common. For a non-critical measurement z_i [3], the corresponding normalized residual r_i^N (with Standard Normal distribution, $N(0,1)$) is defined as follows:

$$r_i^N = \frac{\mid r_i \mid}{\sqrt{\Omega_{ii}}} \tag{2.4}$$

where Ω_{ii} is the diagonal entry of the residual covariance matrix Ω [44]. It is proved that if there is a single bad datum within the set of measurements, the largest normalized residual will correspond to that measurement. If the largest normalized residual exceeds a defined threshold c (e.g., $c = 3$) then the associated measurement is considered bad datum. This measurement is eliminated from the measurement set and the WLS estimation must be run again, until the largest normalized residual does not exceed the threshold.

2.5 Fast decoupled state estimation

The calculation of the gain matrix and its triangular decomposition constitute the heaviest part, in terms of computational efficiency, of the WLS estimator. The fast decoupled formulation [43] reduces this burden by means of the following approximations:

1. Keep the gain matrix constant along the different iterations.

2. Decouple the active and reactive problems making use of the low sensitivity of the real (reactive) power equation to the magnitude (phase angle) of the voltage measurements in high transmission systems.

A more simplified version is the DC State Estimator, which assumes the voltage equal to 1.0 per unit, and neglects the shunt elements and branch resistances. This formulation only considers the active problem, modeling the power flows as a linear function of the phase angles.

2.6 Equality-constrained state estimation

In practice, exactly known magnitudes, such as zero injections, must be accommodated by WLS estimators, which is sometimes done by considering them as very accurate measurements with arbitrarily large weights. However, the coexistence of extremely uneven weights may lead to convergence problems due to ill-conditioning of the gain matrix [3]. The equality-constrained formulation avoids the use of high weights by explicitly adding the necessary constraints to the WLS estimation model

[8, 21], which becomes:

$$z = h(x) + e \tag{2.5}$$

$$h_e(x) = 0 \tag{2.6}$$

where $h_e(x) = 0$ is the set of equality constraints representing the virtual measurements (which are excluded therefore from $h(x)$ and z).

Then, the associated optimization problem can be formulated as:

$$\begin{aligned} \text{minimize} \quad & J(x) = \frac{1}{2}[z - h(x)]^T W [z - h(x)] \\ \text{subject to} \quad & h_e(x) = 0 \end{aligned} \tag{2.7}$$

By applying the Gauss-Newton method to the first-order optimality conditions of the resulting Lagrangian function, the solution of the constrained optimization problem is obtained by repeatedly solving the following system:

$$\begin{bmatrix} \beta H^T W H & H_e^T \\ H_e & 0 \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \lambda \end{bmatrix} = \begin{bmatrix} \beta H^T W \Delta z_k \\ -h_e(x_k) \end{bmatrix} \tag{2.8}$$

where λ is the vector of Lagrange multipliers and β is a scaling factor which can be introduced to improve the condition number of the coefficient matrix. According to [3], Chapter 3, a reasonable value for β is,

$$\beta^{-1} = \max W_{ii} \tag{2.9}$$

The main shortcoming of the above system is that the coefficient matrix is indefinite. However, there exist efficient solvers fully exploiting the resulting matrix sparsity [14, 27, 6].

2.7 Multi-area state estimation

Research on Multi-Area State Estimation (MASE) can be traced back to the late 70's, shortly after state estimators started being put into service [61]. This addresses the problem of performing efficient state estimation on huge power systems, with the twofold motivation of reducing computing time (under the then-limited computational resources) and exploiting the fact that real-time measurements are gathered within areas by the various control centers which are distributed over the grid.

Various methodologies can be used to solve the decomposition-coordination problem involved in MASE. A large majority of methods resort to the classical WLS formulation [62], but the Weighted Least Absolute Value has also been adopted [1]. Most hierarchical schemes rely, at both the local and the central levels, on an iterative scheme to solve the Normal equations of concern. In decentralized schemes, with possible coordination at the iteration level, Lagrangian relaxation-based algorithms are usually adopted. Some works also introduce certain heuristics intended to simplify the optimality conditions of the coordination problem. A category of MASE approaches formulate the WLS equations as an optimization problem, usually involving a Lagrangian function explicitly handling constraints imposed by network equations and/or boundary conditions.

MASE is one of the key applications of the factorized SE framework presented in this work. For this reason, a thorough taxonomy and state of the art review on

MASE has been preliminarily performed [T1].

Chapter 3

Factorization of Nonlinear WLS Problems

This chapter presents the two-stage factorized solution of the nonlinear WLS problem arising in SE, including two particular cases in which one of the resulting subproblems is linear. This material constitutes the theoretical basis for the several SE applications discussed in subsequent chapters [T2].

3.1 Two-Stage factorization of nonlinear WLS problems

The factorized approach to solve the WLS problem arises when the nonlinear measurement model (2.1) is "unfolded" into two sequential WLS problems, as follows:

$$z = f_1(y) + e \tag{3.1}$$

$$y = f_2(x) + e_y \tag{3.2}$$

where y is a vector of intermediate variables, selected in such a way that the solution of the pair (3.1)-(3.2) offers any advantage over that of (2.1). For the resulting factorized

model to be equivalent to the original one the following condition must be satisfied,

$$h(x) = f_1[f_2(x)] \quad \Rightarrow \quad H(x) = F_1(y)F_2(x) \tag{3.3}$$

where H, F_1 and F_2 represent the Jacobian matrices of h, f_1 and f_2 respectively.

As shown in the Appendix of [T2] the optimal estimate $\hat{x}$ provided by the conventional iterative process (2.3) can be alternatively obtained by successively solving the following pair of equations:

$$[F_1^T W F_1]\Delta y_k = F_1^T W[z - f_1(y_k)] \tag{3.4}$$

$$[F_2^T G_1 F_2]\Delta x_k = F_2^T G_1[\tilde{y} - f_2(x_k)] \tag{3.5}$$

where $\tilde{y}$ in (3.5) is the estimate of the intermediate vector provided by (3.4), and the weighting matrix G_1 satisfies,

$$G_1 = [\mathrm{cov}(y)]^{-1} = F_1^T W F_1 \tag{3.6}$$

In other words, the weighting matrix of the second WLS problem is the gain matrix of the first one.

Full equivalence between the original and the factorized models requires that the linearization of f_1 and f_2 be performed at the same point in the n-dimensional space represented by x. This leads in the general case to an outer iterative process in which the two WLS subproblems are successively solved.

Accordingly, the factorized procedure can be formally decomposed into the following sequence of steps, where the first run is separately considered for clarity of

presentation [T2].

First run:

- *Stage 1*: Find $\tilde{y}$ by repeatedly solving (3.4) until convergence. As a byproduct, $\tilde{G}_1$ is available.
- *Stage 2*: Find $\hat{x}$ by repeatedly solving (3.5) with $G_1 = \tilde{G}_1$.

Subsequent runs (if needed):

- *Stage 1*: Update the Jacobian, $\hat{F}_1$, and the gain matrix, $\hat{G}_1$, for $\hat{y} = f_2(\hat{x})$. Keeping these matrices constant, find $\tilde{y}$ by repeatedly solving (3.4).
- *Stage 2*: Find $\hat{x}$ by repeatedly solving (3.5) with $G_1 = \hat{G}_1$.

The process is repeated until two consecutive runs of stage 2 provide close enough values of $\hat{x}$.

Figure 3.1 represents schematically the two-stage procedure outlined above. The auxiliary vector y plays the role of a state vector when solving Stage 1 and that of a pseudomeasurement vector when solving stage 2. Note that the weighting matrix of stage 2 is no longer diagonal, but its sparsity can be fully exploited. The fact that subsequent runs of stage 1 involve only constant matrices can also be exploited.

In practice, unless the measurement vector is very noisy and/or contains key bad data, the first run of stages 1 and 2 will provide sufficiently accurate results.

For the sake of clarity, the above description assumes that all raw measurements z can be used in stage 1. If this is not the case, then stage 2 can be easily redesigned to handle simultaneously the pseudomeasurement $\tilde{y}$ and the components of z not yet used during Stage 1. This is explained in detail in [26],[T3].

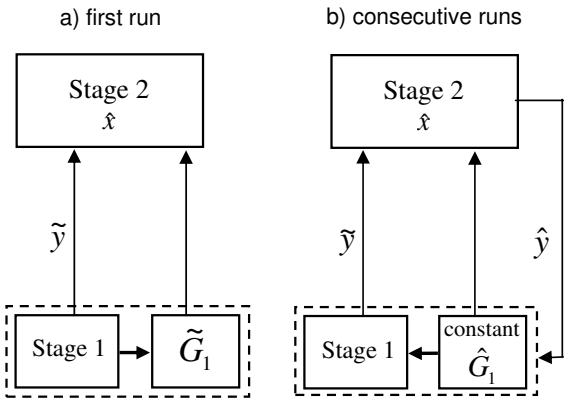


Figure 3.1: Schematic flowchart of the two-stage WLS factorized procedure.

3.2 Particular cases involving a linear WLS subproblem

The fully nonlinear factorized model presented above reduces to any of the particular cases discussed in this section when one of the WLS subproblems becomes linear.

3.2.1 First stage is linear

This case, by far the most interesting in practice, arises when the intermediate vector y can be chosen in such a way that the measurement model of Stage 1 becomes linear [26],[T3]. Then, the nonlinear systems (3.1) and (3.2) reduce to:

$$z = Ay + e \tag{3.7}$$

$$y = f_2(x) + e_y \tag{3.8}$$

The general nonlinear model can be easily particularized to this case by systematically replacing $f_1(y)$ and F_1 in the expressions above by Ay and A, respectively. The most relevant implication is that the iterative system (3.4) reduces in this case

to the following linear one:

$$[A^T W A]\tilde{y} = A^T W z \tag{3.9}$$

Accordingly, the gain matrix

$$G_1 = A^T W A \tag{3.10}$$

remains constant, so long as the network topology and measurement set structure are unaltered. Therefore, changes in the measurement values and state variables originated by the daily load evolution do not alter the linear model of stage 1, which is one of the sources of computational saving associated with the factorized approach.

The factorized procedure reduces in this case just to a single run of stages 1 and 2. Stage 1 constitutes a linear prefilter of the raw measurement vector, which can be a valuable tool by itself (for instance to perform preliminary bad data analysis when the redundancy if sufficiently high). Accuracy of the linear estimate delivered by Stage 1 could be subsequently checked by comparing $\tilde{y}$ with $\hat{y} = f_2(\hat{x})$.

3.2.2 Second stage is linear

In this case, the WLS problem of stage 2 becomes linear. This may arise when all raw measurements are processed during Stage 1 and no lossy network components are involved in stage 2, which reduces typically to a trivial model relating state variables with their estimates provided by stage 1 [50]. Then, the nonlinear systems (3.1) and (3.2) reduce to:

$$z = f_1(y) + e \tag{3.11}$$

$$y = Bx + e_y \tag{3.12}$$

The general model can be easily particularized to this case by systematically replacing $f_2(x)$ and F_2 by Bx and B, respectively. The most relevant implication is that the iterative system (3.5) reduces to the following linear one:

$$[B^T G_1 B]\hat{x} = B^T G_1 \tilde{y} \tag{3.13}$$

Note that stage 2 involves a constant, usually quite sparse and trivial Jacobian, B, but a weighting matrix G_1 which compactly embeds most of the information associated with the original SE problem (raw measurement covariance, network topology and parameters).

Like in the general nonlinear case, the factorized scheme may theoretically involve several executions of stages 1 and 2. In practice, however, it is seldom needed to repeat the two-stage process, unless the required accuracy is extremely high or the raw measurement set is abnormally noisy.

3.3 Selecting the intermediate variables

The auxiliary vector y can be defined in several ways, depending on the application of interest. Each of the following two chapters describes a different strategy for choosing y. In Chapter 4 the main goal is to geographically decompose stage 1 so that decoupled subsystems arise which can be solved in a distributed manner. For this purpose, y can be usually obtained by simply duplicating a small subset of entries of the conventional state vector x, as explained below. Then, stage 2 is aimed at unifying the slightly discordant estimates of those duplicated entries, which are provided by each decoupled subsystem.

In Chapter 5 the driver is to enhance the convergence speed and computational

performance of the conventional Gauss-Newton iterative scheme. In this case, the components of vector y include each and every nonlinear term in the power flow equations, so that stage 1 becomes linear while the size of y is kept as small as possible.

Chapter 4

Multilevel Hierarchical State Estimation

This chapter discusses the application of the factorized solution approach, presented in the previous chapter, to a hierarchical context in which large-scale multilevel systems, ranging from distribution feeders and associated substations to the regional multi-TSO interconnected grid, are to be estimated [T2]. The multilevel hierarchy is dealt with by applying the two-stage factorization procedure to each pair of adjacent levels. What all applications have in common is that stage 1 is performed in a geographically distributed manner while stage 2 coordinates the solutions provided by stage 1. The main driver in this case is to process the raw information as close as possible to the level (feeder, substation, ...) in which it is captured. This way, stage 1 can be performed in parallel and a minimum amount of information has to be submitted to a central processor for coordination of solutions during stage 2.

Before dealing with the particular applications, the next section describes in detail how the generic two-stage factorization procedure can be tailored to those cases in which stage 1 is geographically distributed.

4.1 Geographically decomposed factorized solution

Let us assume that a large set of geographically scattered measurements can be naturally grouped in clusters, weakly coupled with each other. Figure 4.1.(a) illustrates this common situation for the case of four areas or clusters. Within each area two types of variables can be distinguished: a) internal variables, not involved in the measurement models of neighbor areas; b) border variables, appearing in the measurement model of at least an adjacent area. Accordingly, the measurement model (2.1) can be decomposed as follows for the case of j clusters,

$$\begin{aligned} z_1 &= h_1(x_{i1}, x_b) + e_1 \\ z_2 &= h_2(x_{i2}, x_b) + e_2 \\ &\vdots \\ z_j &= h_j(x_{ij}, x_b) + e_j \end{aligned} \tag{4.1}$$

where x_{ik} represents the set of internal variables for area k and x_b comprises the union of all border variables. Ideally, the number of interior variables should be much larger than that of border variables, but in practice this may not be always the case. Notice that the set x_b constitutes the coupling term among all measurement submodels.

In order for stage 1 to be geographically distributed, the intermediate vector y should be selected in such a way that stage 1 reduces to a set of fully decoupled SE subproblems, as shown in Fig. 4.1.(b). Unlike x_b, whose components are shared by two or more areas, the vector of border variables y_b is composed of disjoint components, each involved in the measurement model of a single cluster. Depending on the

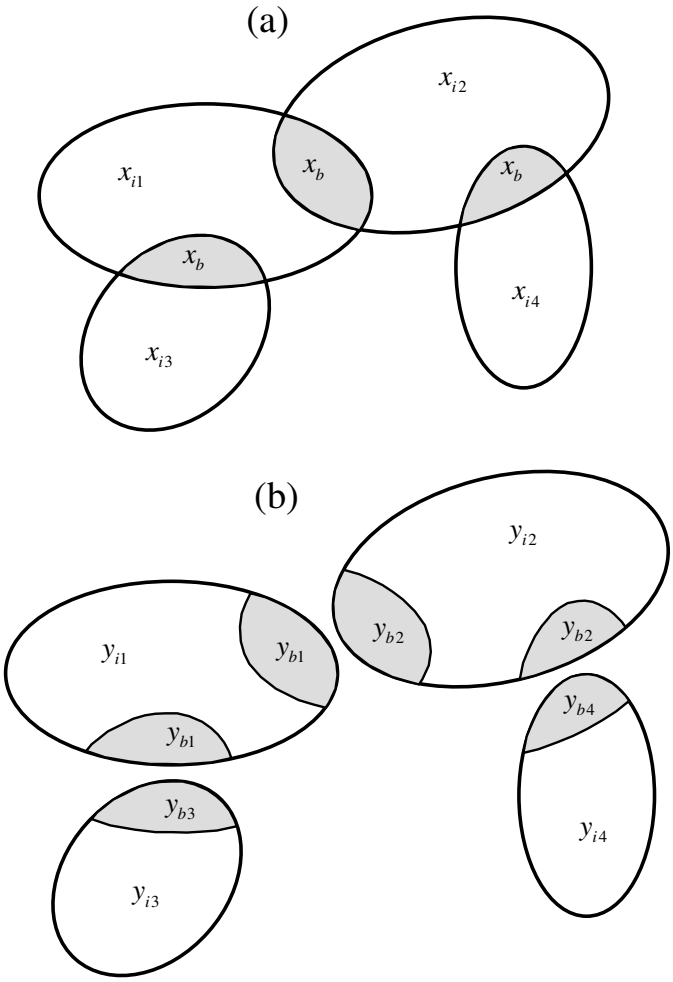


Figure 4.1: Network model composed of four natural clusters.

measurement model and spatial structure of the particular SE problem being decomposed, the augmented vector y_b can be obtained by simply replicating border state variables of the original problem, adding certain measured magnitudes to the state vector, etc.

According to the geographical decomposition achieved, the measurement vector z is split into j components, each one exclusively related to the respective components

of y. Therefore, the decoupled measurement models of stage 1 can be mathematically formulated as follows,

$$\begin{aligned} z_1 &= f_{11}(y_{i1}, y_{b1}) + e_1 \\ z_2 &= f_{12}(y_{i2}, y_{b2}) + e_2 \\ &\vdots \\ z_j &= f_{1j}(y_{ij}, y_{bj}) + e_j \end{aligned} \tag{4.2}$$

Figure 4.2 schematically illustrates the interactions between both stages in case stage 1 is performed in a distributed manner. The solution corresponding to each cluster can be iteratively obtained, in the nonlinear case, by solving the associated Normal equation system (3.4).

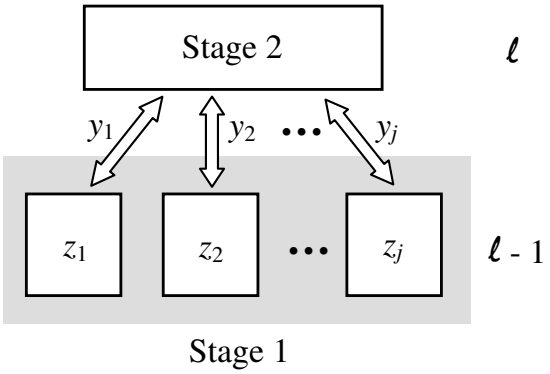


Figure 4.2: Distributed implementation of Stage 1.

Regarding stage 2, consider for simplicity the nonlinear-linear case (extending the distributed formulation to the fully nonlinear context is straightforward). Let $\tilde{y}_b$ and $\tilde{y}_i$ be the estimates provided by stage 1 of the border and interior variables, respectively, for all clusters $1, 2, \ldots, j$. Then the mathematical model of stage 2, for

the linear case, can be written as follows,

$$\tilde{y}_b = Bx_b + e_b \tag{4.3}$$

$$\tilde{y}_i = x_i + e_i \tag{4.4}$$

The weighting matrix arising in stage 2 is a block diagonal matrix formed by simply juxtaposing the individual gain matrices corresponding to each cluster of stage 1, yielding,

$$G_1 = \begin{bmatrix} G_{11} & & & \\ & G_{12} & & \\ & & \ddots & \\ & & & G_{1j} \end{bmatrix} = \begin{bmatrix} \blacktriangle & \diamond & & & & & & \\ \diamond & \blacktriangledown & & & & & & \\ & & \blacktriangle & \diamond & & & & \\ & & \diamond & \blacktriangledown & & & & \\ & & & & \ddots & & & \\ & & & & & & \blacktriangle & \diamond \\ & & & & & & \diamond & \blacktriangledown \end{bmatrix} \tag{4.5}$$

where, as suggested by the rightmost matrix structure, each diagonal block is composed of four submatrices (self and mutual covariance terms between interior an border variables).

Symmetrically reordering the rows/columns so that border and interior variables

are grouped together, leads to the following blocked structure,

$$G_1 = \begin{bmatrix} G_{bb} & G_{ib} \\ G_{ib}^T & G_{ii} \end{bmatrix} = \left[\begin{array}{cc} \blacktriangle & \diamond \\ & \blacktriangle \quad \diamond \\ \ddots \quad\quad \ddots & \\ & \blacktriangle \quad\quad \diamond \\ \hline \diamond & \blacktriangledown \\ & \diamond \quad\quad \blacktriangledown \\ \ddots \quad\quad \ddots & \\ & \diamond \quad\quad \blacktriangledown \end{array}\right] \tag{4.6}$$

Each and every major block above is composed of j decoupled blocks. Furthermore, the block sizes in G_{ii} (bottom right) are frequently much larger than those of G_{bb} (top left), which is an important feature from the computational point of view (e.g., partly distributed solution of stage 2).

Based on the above notation, the Normal equations to be solved at stage 2 can be formulated as,

$$\begin{bmatrix} B^T & \\ & I \end{bmatrix} \begin{bmatrix} G_{bb} & G_{ib} \\ G_{ib}^T & G_{ii} \end{bmatrix} \begin{bmatrix} B & \\ & I \end{bmatrix} \begin{bmatrix} x_b \\ x_i \end{bmatrix} = \begin{bmatrix} B^T & \\ & I \end{bmatrix} \begin{bmatrix} G_{bb} & G_{ib} \\ G_{ib}^T & G_{ii} \end{bmatrix} \begin{bmatrix} \tilde{y}_b \\ \tilde{y}_i \end{bmatrix} \tag{4.7}$$

and, rearranging,

$$\begin{bmatrix} B^T G_{bb} B & B^T G_{ib} \\ G_{ib}^T B & G_{ii} \end{bmatrix} \begin{bmatrix} x_b \\ x_i \end{bmatrix} = \begin{bmatrix} B^T G_{bb} & B^T G_{ib} \\ G_{ib}^T & G_{ii} \end{bmatrix} \begin{bmatrix} \tilde{y}_b \\ \tilde{y}_i \end{bmatrix} \tag{4.8}$$

The above linear system could be directly solved by any of the generic techniques developed for symmetric systems (e.g. Cholesky or orthogonal factorization). However, this would be wasteful for two main reasons: 1) no advantage is taken of the distributed nature of the problem, reflected by the blocked structure of the coefficient matrix; 2) no benefits are obtained from the existing factorization of G_1, performed during stage 1.

Many published procedures on multi-area SE ignore the mutual covariance between internal and border variables ($G_{ib} \approx 0$), as a means of quickly obtaining a suboptimal solution. In fact, crude diagonal approximations of the covariance matrix components can also be found when formulating the coordination problem [62],[T1]. As shown by the examples presented in [50], arbitrarily neglecting mutual covariance values may lead to poor solutions. Therefore, future SE implementations should be redesigned to account for non-diagonal weighting matrices, which is the price paid for the strengthened interactions of existing SEs with their neighborhood.

There are two alternatives to more efficiently solve the system (4.8), taking into account the specific structure of the problem, which are discussed in detail in [T2].

4.2 Application to hierarchical multilevel state estimation

The technological developments brought about by the smart grid context, along with more aggressive regulatory schemes intended to promote efficient trading of clean

energy, allow to envision a future in which SEs will spread from MV distribution feeders to the bulk EHV transmission network, spanning several interconnected areas [60, 9].

In this thesis a multi-level, hierarchical SE paradigm is envisioned, which we believe is the natural choice to deal with the explosion of heterogeneous information arising in this multi-agent distributed environment. At least three major levels are identified, as shown schematically in Fig. 4.3 (a fourth level associated with distribution feeders can also be considered).

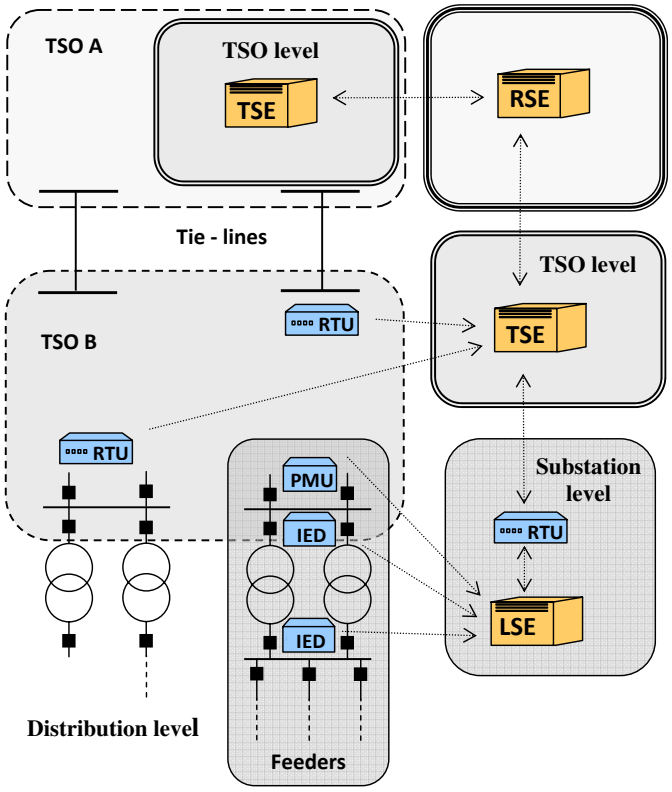


Figure 4.3: Smart grid-oriented multi-level state estimation paradigm.
[LSE: Local State Estimator; TSE: TSO-level SE; RSE: Regional multi-TSO SE]

At the lowest level, a Local SE (LSE) can be implemented to preliminarily deal with the information collected within a substation or small set of adjacent substations. A great majority of raw measurements will be processed at this distributed level, where a modest but sufficient computing power already exists. Distribution substations, delivering power to a large number of secondary transformers through a set of radial feeders, constitute a particular relevant case. In those substations, it is advantageous and makes sense to process each radial feeder in a decoupled manner, leading to a fourth level of information processing.

The results provided by the LSE have to be transmitted through existing RTUs and communication channels to the TSO-level SE (TSE). At this intermediate level, commercially available software can be adopted with minor modifications, the major difference with respect to a conventional SE being that pre-filtered rather than raw measurements are handled.

At the uppermost level, a Regional SE (RSE) will be needed to synchronize and refine the results separately provided by each TSO affiliated with the interconnected system, particularly near the border nodes. The RSE will be a customized tool, designed in such a way that the amount of information exchanged with subordinate TSEs is kept to a minimum. This SE level will significantly benefit from wide-area measurements provided by PMUs.

Figure 4.4 shows the resulting hierarchy, including the feeder level arising in distribution substations. The three resulting bi-levels of application are represented in blue shadow.

The interactions between adjacent levels can be mathematically formulated and justified as particular customized cases of the factorized framework presented in the

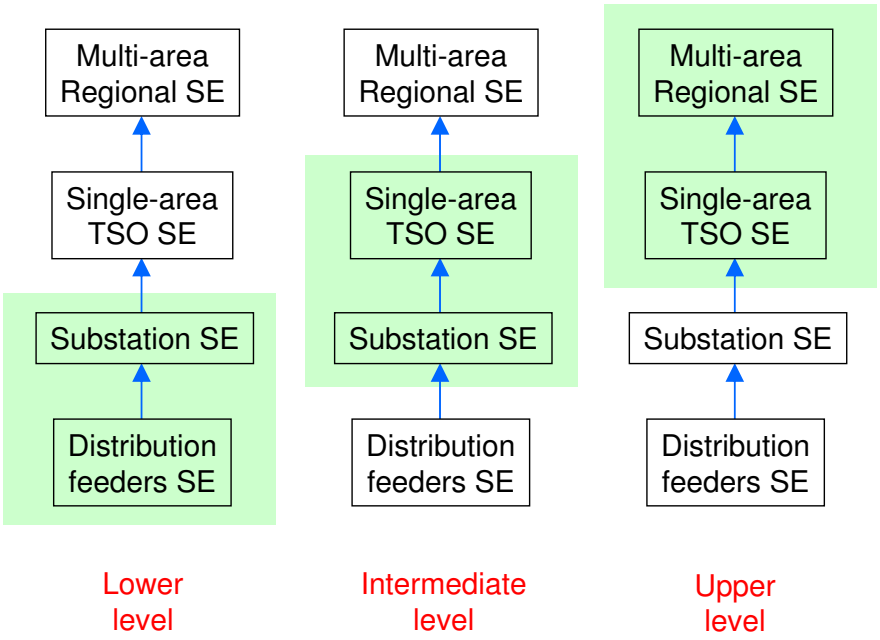


Figure 4.4: Hierarchical multi-level architecture.

previous chapter, in which the first stage is geographically distributed. Each level will be briefly addressed in what follows (the reader is referred to [T2] for more details).

4.2.1 Lower level: feeder system & distribution substation

Distribution substations, delivering power to a large number of secondary transformers through a set of radial feeders, constitute a relevant particular case, in which the size of the resulting SE problem may be discouraging. In fact, a single distribution feeder may comprise hundreds of electrical buses (i.e distribution transformers). Although a simple laptop has enough computing power nowadays to face the conventional SE model, encompassing in this case both the substation and the associated feeders, the factorized SE provides an alternative approach by which the weak electrical coupling among the involved subsystems can be exploited.

Typically, the entire set of radial feeders within a substation is connected to one or at most two busbar sections. Therefore, unlike in transmission or subtransmission systems, what all radial feeders have in common is just an electrical bus, where the electrical interaction between the substation itself and the set of feeders, or between any couple of feeders, takes place (see Figure 4.5).

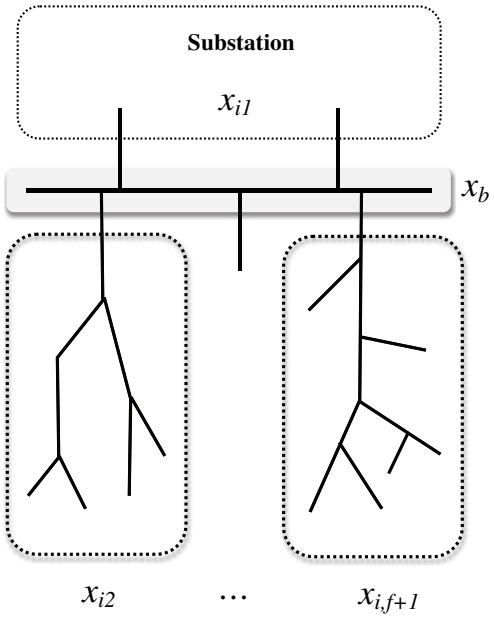


Figure 4.5: Subsystems connected to a common bus in a distribution substation.

This suggests the use of a decomposition-coordination technique to solve the resulting SE problem, like the one described above. In the first stage the overall SE problem is decomposed into $f+1$ decoupled problems (f being the number of feeders) simply by replicating the common bus from which the feeders hang. Then, the SE corresponding to each subsystem is solved in a decoupled manner. The second stage consists of coordinating the solution provided by each decoupled SE. This involves a linear WLS problem, which reduces typically to a trivial model relating state variables

with their estimates provided by the first stage.

As shown by the experimental results in [T4], main advantages of the proposed solution approach are:

- Potential reduction of the computational effort (simpler or constant Jacobian components, smaller size of the state vector during the iterative process, etc.).
- Early bad data processing capability within each subsystem.
- Possibility of tailoring the solution technique in accordance with the peculiarities of each subsystem (ill-conditioning typically arises when solving distribution feeders).
- No need to adopt the same modeling details for each subsystem (e.g. three-phase feeder model versus single-phase substation model).
- Less risk of convergence problems, as a problematic feeder will be identified during the first stage and replaced by equivalent injections.

To illustrate these advantages, representative tests results taken from [T4] are summarized next. The case study used for the simulations is composed of 4 radial feeders, comprising two replicas of the 69- and 85-bus benchmark systems. The difference between the respective pair of replicas consists in the redundancy levels, defined as high ('H') and low ('L'). The feeders are connected to a substation in which two transformers are connected in parallel. One thousand sets of measurements are generated by adding Gaussian noise to the exact values of the measurements provided by a load flow solution. Two error levels have been considered for the whole set of measurements, corresponding to different measurement standard deviations: $\sigma = 0.01$ and $\sigma = 0.025$.

The iterations required by the proposed method to converge are compared to those of a conventional estimator. Table 4.1 collects the average number of iterations required by the conventional method and by the nonlinear decoupled stage of the proposed scheme (stage 1 of the factorized solution, according to the notation introduced above). In the latter case, each subsystem (substation and feeders) is separately shown. On the average, the convergence of the decoupled subsystems is better than that of the entire system. Note that the substation, not as ill-conditioned as the feeders, involves typically one less iteration to converge. The last row refers to a modified case in which the active power of all buses in the 69L feeder is multiplied by 3.2 (for the scenarios with $\sigma = 0.01$). This significantly increases the number of iterations required by the standard method and, to a lesser extent, the affected feeder. However, the substation and the remaining feeders are obviously unaffected, which is a clear advantage of the proposed methodology.

Table 4.1: Number of iterations

	Conv.	Subs.	69H	69L	85H	85L
Exact	5	4	5	5	5	5
$\sigma = 0.01$	5.64	3.90	5.00	5.00	4.99	5.00
$\sigma = 0.025$	5.59	4.05	5.00	5.00	4.98	5.01
$3.2P(69L)$	10.02	3.90	5.00	7.00	4.99	5.00

Experiments including bad data have also been performed. In most cases, when a single bad data is included, the proposed method is able to detect the incorporated bad data after the first stage, allowing the elimination of the bad measurement before performing the second stage. However, main advantages over the conventional estimator are found when dealing with multiple interacting bad data in different feeders. In the experiment carried out in [T4], the conventional estimator wrongly identifies

the bad data, while the proposed method reduces the interacting bad data to single bad data when separately solving each feeder during the first stage and, therefore, is able to detect and identify both of them.

4.2.2 Intermediate level: substation & TSO

When the spotlight is moved a step upwards, a two-level SE arises in which the set of individual substations constitutes the first distributed level while the grid interconnecting them becomes the second.

Chronologically, the application of the factorized solution approach to this geographical scope preceded this thesis [26]. In fact, from the perspective of the general framework proposed in this chapter, such an application constitutes a valuable particular case in which the first distributed stage is linear and the second coordination phase reduces essentially to a conventional TSO-level SE.

Note however that, in addition to the linear substation model considered in [26], the theory developed in this thesis allows also nonlinear substation models to be considered, leading to a factorized solution approach in which both stages are nonlinear [T2].

The LSE at the substation level is useful not only to pre-filter and reduce the size of the measurement set that should be sent to the TSO level, but in many instances also for early detection of model and network inconsistencies at the substation level. For the redundancy ratios expected in next-generation substations, most topology errors and bad data could be handled at this level [64].

4.2.3 Upper level: TSO & Regional SO

The factorization-based approach described above provides also the theoretical basis for an upper SE hierarchy, in which the set of interconnected TSEs constitutes the first level and the RSE is the second level.

The geographical decomposition is achieved in a natural manner by replicating the state variables corresponding to tie-line terminal buses, which are shared by adjacent TSOs. In this context, both SE levels lead in general to nonlinear models, unless all raw measurements are processed at least once at the TSO level, which is generally the case. When this happens, the RSE solution reduces to a linear system in which the nonlinearity of the overall SE model is embedded into the weighting matrix, composed of the gain matrices provided by each TSE.

Main features of the RSE are:

- Each TSO submits to the RSE its estimate of the border variables ($\tilde{y}_b$), along with the statistical information associated with the interior components (see [T2]). The overlapping degree determines the amount of border variables to be exchanged, but this is usually very small in relative terms. This is a welcome feature in nowadays deregulated systems, in which confidentiality of data is a major concern.

- The state vector at the RSE level comprises the border variables shared by the interconnected areas, as well as an auxiliary vector u composed of relative phase angle references, intended to synchronize all areas with respect to a global phase origin. Note that this was not an issue for the other applications of the two-stage procedure considered in this thesis. In the distribution substation case this is so because the common bus is chosen as the phase origin. At the

TSO level phase angles are simply not included in the estimates provided by substations, unless PMUs are used.

- For each border component of the state vector two or more estimates are available in general, which increases the redundancy to estimate the auxiliary vector, u.

A majority of multi-area SE procedures published up to date, combining heuristics and/or relaxation techniques with two-level WLS-based optimization, can be considered simplified particular cases of the general framework discussed herein, frequently providing suboptimal solutions [62],[T1].

To evaluate the advantages of the proposed method, experiments on a real 3-TSO system with a total of 2948 buses have been carried out [50]. Realistic measurements have been generated by incorporating Gaussian noise to the exact magnitudes calculated from a load flow solution. Two sets with different redundancy levels have been generated, being 2.52 and 5.05 the low ('L') and high ('H') levels respectively. Standard deviations associated to the measurements are: 0.01 for voltages; 0.015 for power flows; and 0.02 for power injection measurements.

In order to assess the accuracy of the solution provided by the two-stage estimator, the averages of the absolute errors associated with the estimates of state variables have been computed, as follows:

$$S_V = \frac{1}{N}\sum_{i=1}^{N}|\hat{V}_i - V_i^{\text{ex}}| \quad ; \quad S_\theta = \frac{1}{N-1}\sum_{i=1}^{N-1}|\hat{\theta}_i - \theta_i^{\text{ex}}|$$

where V^{ex} and θ^{ex} represent the exact values, $\hat{V}_i$ and $\hat{\theta}_i$ represent the estimated values and N is the number of buses. Those indexes are calculated, for both the proposed

factorized solution (single run) and the conventional estimator, and the average values for a convergence threshold equal to 0.0001 are shown in Table 4.2.

Table 4.2: Average values of estimation errors

Redundancy	Conventional		Proposed	
	S_v	S_θ	S_v	S_θ
L	1.73E-04	3.40E-04	2.40E-04	4.17E-04
H	1.08E-04	1.49E-04	1.29E-04	1.39E-04

The table data show that the estimation errors of the proposed two-stage method are within the same order of magnitude as those provided by the conventional method. Simulation times have also been compared for the two methods. To better understand these results, it must be clarified that the algorithms have been coded in Matlab (version R2008a) and run under Windows 7 on a 64-bit i5 Intel Core laptop (2.27GHz, 4GB of RAM). Every effort has been made to fully optimize all codes. In addition to the builtin Matlab capability to handle sparse matrices, other aspects related with the efficient implementation of the SE code are detailed in [T6]. Execution times for the tests are shown in Table 4.3. To calculate the total time of the proposed method, the estimations performed at the different TSOs during the first step (non-linear SE), are considered to run in parallel.

As seen in Table 4.3, for the 3-TSO system, the proposed factorized method is three times faster than the conventional one.

Table 4.3: Execution times (in seconds) for the proposed method and the conventional estimator in the real 3-TSO network

Redundancy	Proposed					Conventional
	Step 1			Step 2	Total	
	TSO1	TSO2	TSO3			
L	1.73	1.68	0.22	0.36	2.09	6.08
H	3.23	3.89	0.41	0.23	4.13	12.63

4.2.4 Incorporation of Phasor Measurement Units

Phasor measurements provide phase angles of bus voltage and branch current phasors which have long been missing from measurement sets since the introduction of state estimation function in the control centers [48]. Availability of direct measurement of these phase angles necessitates some changes in the state estimation formulation and also opens up new opportunities of enhancement for existing state estimators.

One simple yet fundamental difference in the formulation of the state estimation problem is the disappearance of the user assigned reference angle [71]. When collecting phasor measurements from PMUs at a PDC, one of the PMUs is typically chosen as a reference and all other measurements are reported with respect to this reference angle. Note that, in a regional multi-TSO estimation, the auxiliary vector u defined above is only necessary to synchronize the areas without PMUs to the UTC time reference. Therefore, vector u will only have entries for these specific areas (TSOs) without PMUs, being it unnecessary in case PMUs exist at all TSOs.

Regarding state estimation, the accuracy of PMU data is a very important issue. It is recognized that synchrophasor measurements are usually more precise than conventional SCADA ones. Conceptually, PMU data are time tagged with precision

better than 1 microsecond and magnitude accuracy that is better than 0.1%. Benefits of incorporating PMUs in existing and future state estimators are discussed in [T5].

Improvements brought by the addition of PMUs to a hierarchical state estimation have been experimentally analyzed for the Paneuropean network. This multi-TSO system is made up of the transmission networks of 27 European TSOs, including Turkey, with a total of 9020 nodes [51]. For the simulations, a base set of measurements is defined, to which different number of PMUs are incorporated, generating different scenarios. The measurements are generated by adding Gaussian noise to the exact measurements calculated from a load flow solution, being the noise proportional to the standard deviation associated to the measurements.

The base set of measurements includes, at the tie-lines, voltage magnitudes measured at both ends and power flows and injections (active and reactive) at only one end of the line. For the rest of the system: voltage magnitudes at all the nodes; power flows at one side of the branches; and power injections at half of the nodes randomly chosen. The standard deviations associated with the measurements of the base set are: 0.004 for the voltages, 0.008 for the power flows (active and reactive) and 0.01 for the power injections. The convergence threshold used for the simulations is 10-6.

Five different scenarios, in which increasing numbers of PMUs are incorporated to the base set of measurements, have been generated. The buses in which PMUs have been incorporated are randomly chosen among the buses whose voltage level is 220kV or higher. The amount of PMUs considered goes from 100 to 500. The standard deviation associated with the PMU measurements is 0.001, which implies four times more accuracy than that of voltage magnitude measurements. Measurements incorporated by each PMU are: voltage phasor at the specific node where the PMU is

located, current phasor injection at the node, and current phasor flows at all the lines incident to the node. The voltage phasor measurements are incorporated in its polar representation, which is trivial for conventional state vectors, also in polar form. The current phasors are used in rectangular form. This way, ill-conditioning problems can be avoided, and Jacobian entries associated to the current phasor measurements are well defined.

Some results of the tests reported in [51] are summarized here, for which scenarios incorporating 100 to 500 PMUs have been generated. In this case, a new index S_v is used, but this time referred to the difference of the exact and estimated voltage phasors (magnitude of the complex difference) instead of voltage magnitudes.

Figure 4.6 shows that the index Sv decreases as the number of PMUs increases. However, the reduction is not linear, and tends somewhat to saturate [51].

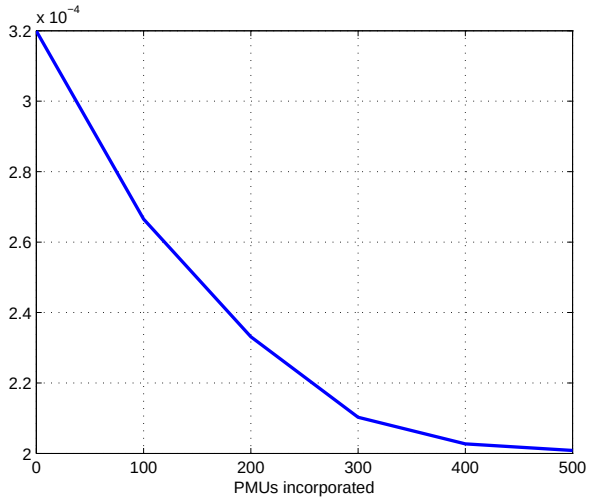


Figure 4.6: S_v index for scenarios with different number of PMUs in the Paneuropean network

Chapter 5

Multi-Stage State Estimation

This chapter focuses on applying more elaborated versions of the factorized solution approach presented in Chapter 3 to conventional WLS-based SEs, in an attempt to reduce the computational cost (simpler or constant Jacobian components, smaller size of the state vector during the iterative process), enhance the speed of convergence (less iterations) or even assure convergence in difficult ill-conditioned cases.

The first section introduces a minimum set of auxiliary variables y so that the conventional SE procedure is decomposed into a linear pre-filter followed by a nonlinear WLS problem [T3].

The second section goes a step further by decomposing the nonlinear phase into an explicit nonlinear transform and a second linear filter, leading to the so-called bilinear SE [T6]. Finally, the equality-constrained version of the bilinear estimator is considered in the last section [T7].

5.1 Two-stage factorization of the conventional state estimation

The two-stage factorized approach presented above can be applied to enhance the performance of conventional TSO-level state estimators, provided a suitable set of

intermediate variables y can be found.

For this purpose, two variables per branch are introduced in such a way that any power (flow or injection) measurement becomes strictly a linear function of those variables, along with squared voltage magnitudes (this set of variables was first introduced in [24], in the context of radial load flow solutions).

For each branch connecting buses i and j the following pair of variables is defined:

$$K_{ij} = V_i V_j \cos\theta_{ij} \tag{5.1}$$

$$L_{ij} = V_i V_j \sin\theta_{ij} \tag{5.2}$$

where $V_i \underline{\lfloor \theta_i}$ is the complex voltage at bus i. In addition, squared voltage magnitudes,

$$U_i = V_i^2 \tag{5.3}$$

replace voltage magnitudes, both in the measurement and in the intermediate state vector, which is composed of $2b + N$ variables for a system with b branches and N buses:

$$y = \{U_i, K_{ij}, L_{ij}\} \tag{5.4}$$

The first linear step is aimed at reducing the size of the raw measurement vector to the maximum extent, without losing any relevant statistical information. The estimate of y, along with the associated covariance matrix, are then used in the second WLS nonlinear estimator, according to the methodology described above.

Exhaustive test results on IEEE benchmark networks presented in [T3] show that the proposed two-stage approach outperforms existing SEs, in terms of both convergence rate and computational cost. The benefits achieved are more noticeable in the

presence of highly redundant and accurate measurement systems, particularly under peak loading conditions. For accurate enough measurement sets the results provided by the linear prefilter can be acceptable for many practical purposes, irrespective of the network loading level.

To illustrate the improvements brought by this approach, some significant results from the tests performed in [T3] are presented at this point. Three test networks have been used: IEEE 118- and 298-bus systems, and a 118-bus modified network, referred to as 118-2P, in which the load level has been duplicated (see [T3] for details). Two redundancy levels have been considered, referred to as 'High' (H) and 'Low' (L). One hundred sets of measurements have been randomly generated for each scenario.

In order to assess the accuracy of the linear prefilter, the indexes S_v and S_θ, defined in section 4.2.3, have been computed for the estimate preliminary provided by this filter. These indexes are compared in Table 5.1 with those arising after the first iteration of the conventional SE solution, for the particular error scenario with $\sigma_V = 0.002$, for voltage measurements, and $\sigma_P = 2\sigma_V$ for power measurements. As can be seen, the accuracy of the linear estimator is roughly an order of magnitude better than that obtained after the first conventional iteration. The accuracy gain is much higher when the doubly loaded case (118-2P) is considered.

Table 5.2 shows the average number of kflops[1] for the sets of scenarios with $\sigma_V = 0.002$ and $\sigma_P = 2\sigma_V$, including both redundancy levels. The first stage of the proposed method and the first iteration of the conventional one are accounted for separately. The cost of the initialization procedure for the second stage is also shown. In parentheses, the average number of iterations needed by each method is provided,

[1] floating-point operations $\times 10^3$

Table 5.1: Average values of estimation errors after the first iteration (conventional) and first linear stage (proposed) for $\sigma_V = 0.002$ and $\sigma_P = 2\sigma_V$

Test Case	Conventional		Proposed	
	S_v	S_θ	S_v	S_θ
118-L	0.0040	0.0045	0.0011	0.0007
118-H	0.0058	0.0045	0.0002	0.0005
298-L	0.0108	0.0156	0.0011	0.0013
298-H	0.0105	0.0154	0.0002	0.0007
118-2P-L	0.0160	0.0636	0.0011	0.0010
118-2P-H	0.0214	0.0638	0.0002	0.0006

excluding the first stage and the first iteration respectively.

Table 5.2: No. of kflops associated with different stages of the solution processes for $\sigma_V = 0.002$ and $\sigma_P = 2\sigma_V$

Test Case	Conventional		Proposed			% Sav.
	First	Rest (it.)	First	Init.	Sec. (it.)	
118-L	76	275 (3)	53	13	177 (2)	31
118-H	162	575 (3)	154	13	357 (2)	29
298-L	167	851 (4)	110	30	600 (2.9)	27
298-H	407	1414 (3.1)	311	30	1168 (2.7)	17
118-2P-L	76	366 (4)	53	13	191 (2.2)	42
118-2P-H	162	767 (4)	154	13	357 (2)	44

The following comments are in order regarding the results obtained:

- The average computational saving of the proposed method for the 600 cases collected in table 5.2 is about 32% (rightmost column of the table).

- The total saving stems from a combination of two factors: 1) the enhanced convergence rate (typically the two-stage scheme requires at least one iteration

less); 2) the lower computational effort required by each iteration.

- Both the first stage of the proposed method and the first iteration of the conventional scheme are cheaper, as no trigonometric functions are involved.

- The computational effort reduction originates mainly in the computation of the Jacobian matrix and the residual vector (both entities are obtained simultaneously to avoid duplicated computations).

- The proposed method is more advantageous under peak loading conditions, owing to the extra iteration typically saved.

5.2 Three-stage bilinear state estimation

The two-stage factorized approach presented in the previous section can be enhanced by further factorizing the second nonlinear stage, leading to a three-stage state estimation methodology.

Both the unconstrained and the equality constrained versions of the three-stage state estimation methodology are succinctly presented here. The reader is referred to the attached papers [T6] and [T7] for the complete details.

In the unconstrained formulation [T6] the nonlinear measurement model (2.1) is replaced by two linear models, which become coupled through a nonlinear change of

variables, as follows[2]:

$$z = By + e \tag{5.5}$$

$$u = f_u(y) \tag{5.6}$$

$$u = Cx + e_u \tag{5.7}$$

In the above factorized model the following auxiliary vectors are introduced:

- Intermediate state vector y,

$$y = \{U_i, K_{ij}, L_{ij}\} \tag{5.8}$$

  whose $2b + N$ components are defined by (5.1)-(5.3) in the previous section.

- Pseudo-measurement vector u, composed also of $2b + N$ variables,

$$u = \{\alpha_i, \alpha_{ij}, \theta_{ij}\} \tag{5.9}$$

  where,

  $\alpha_i = \ln U_i = 2 \ln V_i$

  $\alpha_{ij} = \alpha_i + \alpha_j$

  $\theta_{ij} = \theta_i - \theta_j$.

The nonlinear functions $f_u(\cdot)$, as well as the constant matrices B and C, are easily obtained from the above definitions (see [T6] for the details).

The solution process comprises the following three steps:

[2]To facilitate cross-referencing, the notation adopted in this document is in full agreement with the original papers, even though this leads sometimes to certain inconsistencies between different chapters. For instance, matrix B below is denoted as matrix A in earlier chapters.

1) Compute $\tilde{y}$ by solving the linear system,

$$G_B \tilde{y} = B^T W z \tag{5.10}$$

where

$$G_B = B^T W B \tag{5.11}$$

is the associated gain matrix.

2) Obtain $\tilde{u}$ from

$$\tilde{u} = f_u(\tilde{y}) \tag{5.12}$$

and the weighting matrix,

$$W_u = \text{cov}^{-1}(\tilde{u}) = \tilde{F}_u^{-T} G_B \tilde{F}_u^{-1} \tag{5.13}$$

where $\tilde{F}_u$ is the Jacobian of $f_u(\cdot)$, composed of scalar or 2×2 diagonal blocks with trivial inverses. In fact, as shown in [T6], $\tilde{F}_u^{-1}$ is directly obtained at virtually no cost.

3) Compute $\hat{x}$ by solving the linear system,

$$G_C \hat{x} = C^T W_u \tilde{u} \tag{5.14}$$

where:

$$G_C = C^T W_u C \tag{5.15}$$

is the associated gain matrix.

In summary, in the first stage, raw measurements are processed by an augmented-state linear estimator, for which convergence is never an issue. Then, a nonlinear but explicit transform is applied to this initial estimate, the result of which is finally used by another reduced-order linear estimator. It turns out that the Jacobian of the nonlinear transform, which is needed to compute the weighting matrix of the final linear stage, has a trivial inverse. Moreover, this latter filter involves only the well-known incidence matrix, topologically relating branch variables to the nodal ones.

Figure 5.1 schematically illustrates the mathematical decomposition track followed in this thesis, starting from the conventional state estimator (left), continuing with the two-stage factorized approach (center) and finishing with the three-stage bilinear methodology (right).

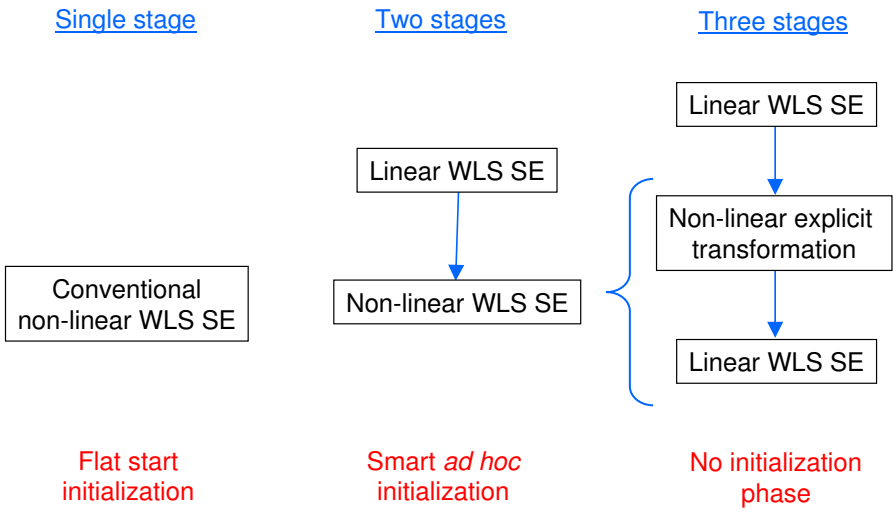


Figure 5.1: Mathematical decomposition: multistage factorized paradigm

Experimental results on IEEE benchmark networks show that the proposed three-stage approach yields, in a single run, virtually the same solution as that provided

by the Gauss-Newton iterative scheme. Only in the presence of extremely high measurement errors and very low redundancy levels might a second run of the last filter be justified, for which the Jacobian of the nonlinear transform should be previously updated. Accordingly, the resulting solution times are between three to four times smaller than those of the conventional iterative method. In tracking mode, over an order of magnitude savings have been achieved.

Test results corresponding to the IEEE 118- and 298-bus benchmark systems are analyzed in [T6]. Experiments corresponding to the 118-bus system are summarized next.

From exact load flow solutions, several measurement sets with varying noise and redundancy levels are simulated. For the results to be more representative, one thousand measurement sets are randomly generated for each scenario. Different error levels are simulated as follows: for voltage measurements, two values of σ_V are considered (0.002 and 0.01). In turn, for each value of σ_V, two increasing values of s.d. for power measurements are tested ($\sigma_P = \rho\sigma_V$, with $\rho = 2$ and $\rho = 5$). For each simulated error two measurement sets are considered, defined as low (L) and high (H) redundancy respectively. The standard WLS methodology, based on the Gauss-Newton iterative method, is applied to each scenario, yielding a number of iterations ranging typically from 4 to 6 for a convergence threshold of $1E-5$ (applied to the largest absolute value change in any state vector component). Then, a single run of the proposed three-stage procedure is also performed for each test case. In order to compare the accuracy of the solution provided by the proposed procedure (single run) with that of the conventional solution approach, the indexes S_V and S_θ are calculated. Only the results for the low redundancy scenario are presented here. Figure 5.2 and 5.3 present,

for the 118-bus network, the probability density functions (pdf) corresponding to the one thousand S_V and S_θ values, respectively, after full convergence of the conventional method (gray continuous line) and after a single execution of the bi-linear estimator (dashed black line). Close to each pdf, representing a unique combination of measurement noise and redundancy level, the minimum number of iterations required by the standard Gauss-Newton approach to converge is shown (the average number of iterations for the entire set of one thousand simulations is slightly higher, as there is always a small percentage of anomalous cases requiring more iterations).

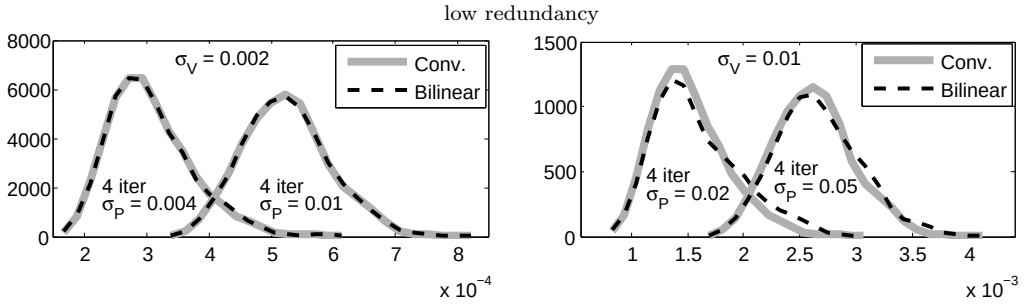


Figure 5.2: Pdf of S_V values for the 118-bus system.

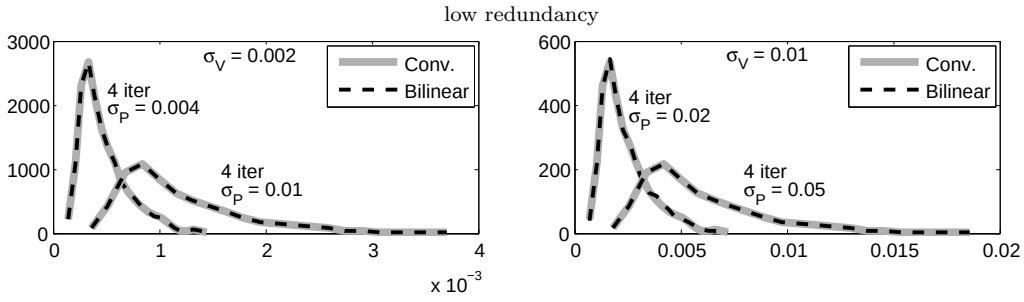


Figure 5.3: Pdf of S_θ values for the 118-bus system.

The speedup, defined as T_c/T_a, where T_c and T_a represent the solution times of the conventional and the alternative factorized methodologies, respectively, has been

computed for each case. For a single run of the proposed method (corresponding to results shown in Figures 5.2 and 5.3), a 2.98 speedup is reached. Since the first linear stage is computationally less expensive for repeated solutions, in which only the measurement values get modified, a 13.37 speedup is achieved when this 'hot' start scenarios are considered.

5.3 Equality-constrained bilinear estimation

Like the unconstrained bilinear SE procedure, the constrained version proposed in this work is composed of three main steps, which are discussed in the logical sequence below.

5.3.1 First linear filter

It is assumed that the equality constraints become linear in terms of the auxiliary vector y, which is indeed the case when null injections are considered. Therefore, the original SE model (2.5)-(2.6), which is nonlinear in terms of x, can be written as follows:

$$z = By + e \tag{5.16}$$

$$Ey = b \tag{5.17}$$

where $b = 0$ for null-injection constraints. The structure and values of matrix E are identical to those in B for regular injection measurements.

The WLS estimate $\tilde{y}$, given the set of measurements (5.16) and the constraints

(5.17), is obtained by solving the augmented system,

$$\begin{bmatrix} B^T W B & E^T \\ E & 0 \end{bmatrix} \begin{bmatrix} \tilde{y} \\ \tilde{\lambda} \end{bmatrix} = \begin{bmatrix} B^T W z \\ b \end{bmatrix} \tag{5.18}$$

which also provides an estimate of the vector of Lagrange multipliers, $\tilde{\lambda}$. Like in the original nonlinear formulation (2.8), a scaling factor β can be introduced in the above system, which is omitted for simplicity of notation.

The gain matrix in (5.18) is the inverse of the associated covariance matrix,

$$G_{Ba} = \begin{bmatrix} B^T W B & E^T \\ E & 0 \end{bmatrix} = \mathrm{cov}^{-1} \begin{bmatrix} \tilde{y} \\ \tilde{\lambda} \end{bmatrix} \tag{5.19}$$

where the subscript a denotes 'augmented'. This matrix will appear in subsequent steps as a weighting matrix.

5.3.2 Intermediate nonlinear transformation

In order to get equation systems with homogeneous sizes, the nonlinear transform (5.6) is trivially augmented as follows:

$$\begin{bmatrix} \tilde{u} \\ \tilde{\lambda} \end{bmatrix} = \begin{bmatrix} f_u(\tilde{y}) \\ \tilde{\lambda} \end{bmatrix} \tag{5.20}$$

which, in incremental form, yields,

$$\begin{bmatrix} \Delta u \\ \Delta\lambda \end{bmatrix} = \underbrace{\begin{bmatrix} \tilde{F}_u & 0 \\ 0 & I \end{bmatrix}}_{\tilde{F}_{ua}} \begin{bmatrix} \Delta y \\ \Delta\lambda \end{bmatrix} \tag{5.21}$$

This allows the covariance matrix of the vector $[\tilde{u}, \tilde{\lambda}]$ to be related with that of $[\tilde{y}, \tilde{\lambda}]$, given by (5.19). In terms of inverses,

$$\mathrm{cov}^{-1} \begin{bmatrix} \tilde{u} \\ \tilde{\lambda} \end{bmatrix} = \tilde{F}_{ua}^{-T} G_{Ba} \tilde{F}_{ua}^{-1} = W_{ua} \tag{5.22}$$

where the inverse of $\tilde{F}_{ua}$ is trivially obtained.

5.3.3 Second linear filter

The last stage reduces to the following WLS linear model:

$$\tilde{u} = Cx + e_u \tag{5.23}$$

$$\tilde{\lambda} = \lambda + e_\lambda \tag{5.24}$$

where the two subsystems become coupled by the noise covariance (5.22). The corresponding estimates are provided by solving the system,

$$\left(C_a^T W_{ua} C_a\right) \begin{bmatrix} \hat{x} \\ \hat{\lambda} \end{bmatrix} = C_a^T W_{ua} \begin{bmatrix} \tilde{u} \\ \tilde{\lambda} \end{bmatrix} \tag{5.25}$$

where the following augmented matrix is introduced,

$$C_a = \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix}$$

In summary, the equality-constrained bilinear procedure, involving (5.18), (5.20) and (5.25), is simply obtained by properly augmenting the unconstrained version, (5.10), (5.12) and (5.14), with the vector of Lagrange multipliers.

The main and foremost advantage of this approach is to do away with the use of high weights for exact measurements, a strategy that deteriorates the condition of gain matrices in state estimation methods that do not account for equality constraints explicitly. In [T7], experiments are carried out for two standard network benchmarks as well as different sets of equality constraints and low and extremely high noise scenarios. Results show important improvements in matrix conditioning with respect to previous methods (with all the numerical advantages behind it) yet the solutions reached after just the first run lie substantially close to those provided by traditional methods after four or five iterations.

Results corresponding to the 118-bus benchmark network are summarized next. Table 5.3 presents the condition numbers of the two linear steps for both, the equality-constrained bilinear procedure and the unconstrained bilinear approach, for a scenario with 20% of buses with injections set as exact measurements. These measurements are modeled as equality constraints in the proposed method, while in the unconstrained bilinear implementation they are modeled as high weighted measurements. Parameter ρ represents the factor by which the weights applied to the exact measurements are multiplied with respect to the average weight associated to regular measurements.

Table 5.3: Log of condition numbers of different gain matrices for the 118-bus systems

Method	$Log(\rho)$	118-bus	
		Step 1	Step 2
Equality Constrained Bilinear	-	5.81	5.78
Unconstrained Bilinear	3	8.86	8.75
	4	9.85	9.75
	5	10.9	10.8

As can be seen, condition numbers increase with ρ in both linear steps of the three-stage bilinear procedure. The impact of the condition number on the accuracy of the solution of linear equation systems is analyzed in [T7].

Chapter 6

Conclusions and Future Work

6.1 Conclusions

This work has first provided an account of the current context and challenges faced by EMS, including the incorporation of technological breakthroughs (digital substations, smart distribution feeders, PMUs, etc.) and the need to cope with energy transactions which are flourishing across national borders. Then, the most relevant aspects of the now well-established SE tool have been succinctly reviewed.

Next, a mathematical framework has been developed allowing nonlinear WLS problems to be factorized into two sequential subproblems, which are also nonlinear in the general case. Several particular applications of this two-stage decomposition scheme have been considered and implemented, which can be grouped in two categories:

1. Geographically distributed hierarchical SE environments, in which several two-stage processes, one of them linear, are concatenated from the distribution system up to the multi-area regional system, leading to a multilevel SE scheme. In this category, the most outstanding development is the Paneuropean SE,

which has been successfully tested on the whole European EHV network and integrated into the FP7 PEGASE project prototypes.

2. Conventional single-area SEs, in which the introduction of auxiliary variables allows the nonlinear WLS problem to be factorized into simpler subproblems. In this category, the bilinear formulation (two linear WLS problems with a nonlinear transformation in between) has outperformed any other existing formulation of the WLS problem, including the others proposed in this work, achieving convergence for practical matters in a single execution.

Figure 6.1 graphically represents the major developments of this work, and their causal relationships.

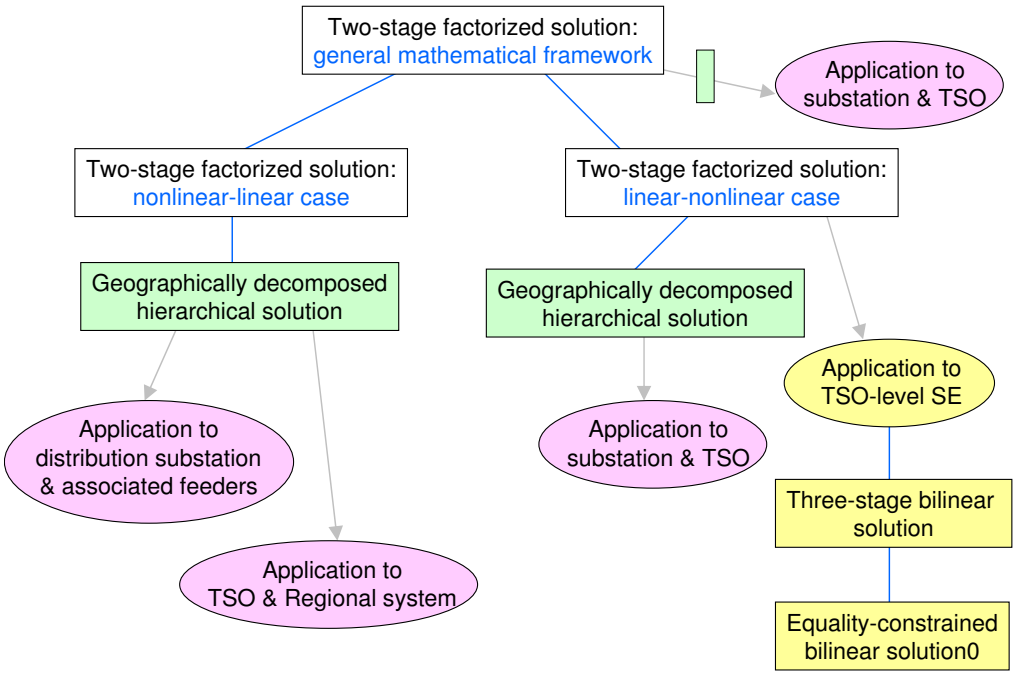


Figure 6.1: General overview of the main thesis developments and their relationships

6.2 Thesis contributions

The contributions of this thesis are summarized below, according to their theoretical or practical relevance.

- Theoretical contributions:

    1. A general two-stage factorization procedure for nonlinear WLS problems. This approach is subsequently used to implement a multilevel hierarchical SE scheme, by which very large-scale power systems can be efficiently and accurately monitored. The algorithm can be particularized for different geographical decomposition levels: distribution feeders, substations, transmission and multi-area systems [T2].

    2. A factorized SE approach in which the first stage reduces to a linear filter of raw measurements [T3]. This can be also deemed as a particular case of the general two-stage factorization procedure developed in [T2]. Simulation results show that the proposed approach converges faster than the conventional SE and is computationally more efficient.

    3. The factorized SE concept has been applied also to smart distribution substations [T4], taking advantage of the weak electrical coupling existing among the set of feeders connected to the same substation busbar. Advantages of this solution technique lie in early bad data processing capability as well as the possibility of tailoring both the network model and the solution technique according to the peculiarities of each subsystem (e.g. unbalanced or ill-conditioned feeders).

4. A bilinear three-stage SE factorization has been proposed [T6] which provides, in a non iterative manner, virtually the same solution as that provided by the Gauss-Newton iterative scheme arising in the WLS conventional method.

5. The equality-constrained version of the bilinear SE has been also elaborated [T7]. This work can be considered an extension of [T6], which is advantageous in the prevention of the ill-conditioning typically arising when equality constraints are handled as virtual measurements with huge weights.

- Practical contributions:

1. A taxonomy of multi-area state estimation methods, offering a unifying description of a relatively large number of works devoted to the subject [T1].

2. A survey on the use of PMUs in power system SE [T5], presenting the benefits that existing and future SE can achieve by incorporating these devices in the monitoring process.

3. An efficient implementation of the two-step multi-area SE, which has been tested on the UCTE interconnected system. This prototype software includes PMU devices.

6.3 Future work

Improvements of the proposed methods and future research lines have been identified, among which the more significant are outlined next:

- Development of observability techniques for the state variables arising in the first linear stage of the proposed factorized methods.
- In this work, the bilinear factorization (unconstrained and equality constrained) has been applied to the TSO-level SE. Applications to other levels (substation, feeder...) should be explored.
- Extension of the proposed bilinear scheme to incorporate non-conventional measurements, such as current measurements and/or PMUs, to cope with undergoing technological improvements in this field.
- Application of the hierarchical decomposition and factorized approaches to non-quadratic (robust) estimators.
- Exploring the advantages that the proposed factorized schemes could bring to other power system analysis problems.

Bibliography

[1] A. Abur and M.K. Celik, "Multi-area Linear Programming. State Estimator using Dantzig-Wolfe Decomposition", Proc. Tenth Power Syst. Comput. Conf., pp.1038 - 1044 , 1990.

[2] A. Abur, A. Gómez-Expósito , "Detecting Multiple Solutions in State Estimation in the Presence of Current Magnitude Measurements," IEEE Transactions on Power Systems, vol. 12, no. 1, pp. 370- 375, February 1997.

[3] A. Abur, A. Gómez-Expósito, "Power System State Estimation: Theory and Implementation," Marcel Dekker, 2004.

[4] O. Alsac, N. Vempati, B. Stott, A. Monticelli, "Generalized State Estimation," IEEE Transactions on Power Systems, vol. 13, no. 3, pp. 1069-1075, August 1998.

[5] F. Alvarado, W. Tinney, "State Estimation Using Augmented Blocked Matrices," IEEE Transactions on Power Systems, vol. 5, no. 3, pp. 911-921, August 1990.

[6] E. Anderson et. al., "Lapack User's Guide", http://www.netlib.org/lapack/lug/lapack_lug.html

[7] R. N. Anderson, "The Distributed Storage-Generation Smart Electric Grid of the Future," in Proceedings of the Workshop, The 10-50 Solution: Technologies and

Policies for a Low-Carbon Future. The Pew Center on Global Climate Change and the National Commission on Energy Policy, 2004.

[8] Aschmoneit F., Peterson N., Adrian E., "State Estimation with Equality Constraints", 10th PICA Conference Proceedings, Toronto, pp. 427-430, May 1977.

[9] A. Bose, "Smart Transmission Grid Applications and Their Supporting Infrastructure," IEEE Transactions on Smart Grid, vol.1, no.1, pp. 11-19, June 2010.

[10] K. Brand, V. Lohmann, W. Wimmer, "Substation Automation Handbook," Utility Automation Consulting Lohmann, Bremgarten, Switzerland, 2003.

[11] E. Caro, A. Conejo, R. Mínguez, "Power system state estimation considering measurement dependencies," IEEE Transactions on Power Systems, vol. 24, no. 4, pp. 1875-1885, November 2009.

[12] M. K. Celik, A. Abur, "A robust WLAV state estimator using transformations," IEEE Transactions on Power Systems, vol. 7, no. 1, pp. 106-113, February 1992.

[13] J. Chen, A. Abur, "Placement of PMUs to enable bad data detection in state estimation," IEEE Transactions on Power System, vol. 21, no. 4, pp. 1608-1915, Nov 2006.

[14] Y. Chen, T. A. Davis, W. W. Hager, S. Rajamanickam, "Algorithm 887: CHOLMOD, Supernodal Sparse Cholesky Factorization and Update/Downdate," ACM Transactions on Mathematical Software (TOMS) Journal, Volume 35 Issue 3, October 2008.

[15] K. A. Clements, "Observability methods and optimal meter placement," International Journal of Electrical Power and Energy Systems, vol. 12, no. 2, pp. 88-93, April 1990.

[16] K. A. Clements, P. W. Davis, K. D. Frey, "Treatment of inequality constraints in power system state estimation," IEEE Transactions on Power Systems, vol. 10, no. 2, 567-574, May 1995.

[17] http://www.desertec.org/

[18] www.energy-regulators.eu/portal/page/portal/eer_home/eer_publications /ceer_ergeg_papers/electricity/2007/e06-bag-01-06_blackout-finalreport _2007-02-06.pdf.

[19] http://www.energy-regulators.eu/portal/page/portal/EER_HOME /EER_INITIATIVES

[20] www.energie-schweiz.ch/imperia/md/content/energiemrkteetrgertechnik en/elektrizitt/strompanne03/12.pdf.

[21] A. Gjelsvik, "The significance of the Lagrange Multipliers in WLS State Estimation with Equality Constraints," in Proceedings of the 11th Power Systems Computation Conference, Avignon, pp. 619-625, August 1993.

[22] A. Gjelsvik ,S. Aam, L. Holten , "Hachtel's Augmented Matrix Method - A Rapid Method Improving Numerical Stability in Power System Static State Estimation," IEEE Transactions on Power Apparatus and Systems, vol. PAS-104, pp. 2987-2993, November 1985.

[23] A. Gómez-Expósito, A. Abur, "Generalized Observability Analysis and Measurement Classification," IEEE Transactions on Power Systems, vol.13, no.3, pp. 1090-1096, August 1998.

[24] A. Gómez-Expósito and E. Romero, "Reliable load flow technique for radial distribution networks", *IEEE Transactions on Power Systems*, Vol. 14, No. 3, pp. 1063-1069, August 1999.

[25] A. Gómez Expósito, A. de la Villa Jaén, "Reduced Substation Models for Generalized State Estimation", IEEE Transactions on Power Systems, vol. 16, no. 4, pp. 839-846, November 2001.

[26] A. Gómez-Expósito, A. de la Villa Jaén, "Two-Level State Estimation With Local Measurement Pre-Processing," IEEE Transactions on Power Systems, vol. 24, no. 2, pp. 676-684, May 2009.

[27] J.D. Hogg and J.A. Scott, "An indefinite sparse direct solver for multicore machines", Technical Report TR-RAL-2010-011. Available from http://www.numerical.rl.ac.uk/reports/reports.shtml

[28] IEEE Standard Definition, Specification and Analysis of Systems Used for Supervisory Control, Data Acquisition, and Automatic Control, IEEE Std. C37.1.1994, 1994.

[29] IEEE Standard for Synchrophasors for Power Systems (IEEE Std C37.118-2005)

[30] W. Jiang, V. Vittal, G.T. Heydt, "Diakoptic State Estimation Using Phasor Measurement Units", *IEEE Transactions on Power Systems*, vol. 23, No. 4, pp. 1580-1589, Nov. 2008.

[31] Z. Jun, A. Abur;, "Identification of network parameter errors," IEEE Transactions on Power Systems, vol. 21, no. 2, pp. 586-592, May 2006.

[32] H. M. Kim, J. J. Lee, D. J. Kang, "A Platform for Smart Substations," fgcn, vol. 1, pp.579-582, Future Generation Communication and Networking (FGCN 2007), vol. 2, 2007.

[33] M. Korkali, A. Abur, "Placement of PMUs with Channel Limits," Proceedings of the IEEE Power and Energy Society General Meeting, Calgary, CA, July 26-30, 2009.

[34] G.R. Krumpholz, K.A. Clements, P.W. Davis, "Power System Observability: A Practical Algorithm Using Network Topology," IEEE Transactions on Power Apparatus and Systems, vol. PAS-99, no. 4, pp. 1534-1542, July 1980.

[35] C.N. Lu, J.H. Teng, W.-H.E. Liu , "Distribution state estimation", IEEE Transactions on Power Systems, vol.10, no. 1, pp. 229-240, February 1995.

[36] F. H. Magnago, A. Abur, "Unified Approach to Robust Meter Placement Against Bad Data and Branch Outages," IEEE Transactions on Power Systems, vol.15, no.3, pp.945-949, August 2000.

[37] J. D. McDonald, "Electric Power Substation Engineering," CRC Press, Boca Raton, FL, 2003.

[38] L. Mili, M. G. Cheniae, N. S. Vichare, P. J. Rousseeuw, "Robust state estimation based on projection statistics of power systems," IEEE Transactions on Power Systems, vol. 11, no. 2, pp. 1118-1127, May 1996.

[39] L. Mili, T. Van Cutsem, M. Ribbens-Pavella, "Hypothesis Testing Identification: A New Method for Bad Data Analysis in Power System State Estimation," IEEE Transactions on Power Apparatus and Systems, vol. PAS-103, no. ll, pp. 3239-3252, 1984.

[40] Monticelli, A., "Electric power system state estimation," Proceedings of the IEEE , vol. 88, no.2, pp.262-282, Feb 2000.

[41] A. Monticelli,"Modeling Circuit Breakers in Weighted Least Squares State Estimation," IEEE Transactions on Power Systems, vol. 8, no. 3, pp. 1143-1149, August 1993.

[42] A. Monticelli, "State Estimation in Electric Power System. A generalized Approach," Kluwer Academic Publishers, 1999.

[43] A. Monticelli, A. Garcia, "Fast decoupled state estimators," IEEE Transactions on Power Systems, vol. 5, no. 2, pp. 556-564, 1990.

[44] A. Monticelli, A. Garcia, "Reliable Bad Data Processing for Real-Time State Estimation," IEEE Transactions on Power Apparatus and Systems, vol. PAS-102, no. 5, pp. 1126-1139, May 1983.

[45] A. Monticelli, F. Wu, "A Method That Combines Internal State Estimation and External Network Modeling," IEEE Transactions Power Apparatus and Systems, vol. PAS-104, no.1, pp. 91-103, January 1985

[46] A. Monticelli, F. F. Wu, "Network Observability: Theory," IEEE Transactions on Power Apparatus and Systems, vol. PAS-104, no. 5, pp. 1042-1048, May 1985.

[47] R. Nucera, M. Gilles , "A Blocked Sparse Matrix Formulation for the Solution of Equality-Constrained State Estimation," IEEE Transactions on Power Systems, vol. 6, no.1, pp. 214-224, February 1991.

[48] A. G. Phadke, "Synchronized Phasor Measurements in Power Systems," IEEE Computer Applications in Power, vol. 6, Issue 2, pp. 10-15, April 1993.

[49] A. G. Phadke, J. S. Thorp, "Synchronized Phasor Measurements and Their Applications," Springer, 2008.

[50] Deliverable 2.1 part 2, "Algorithms for state estimation of ETN," European Community's 7th Framework Programme, PEGASE project (Pan European Grid Advanced Simulation and state Estimation), 2011, http://fp7-pegase.eu/

[51] Deliverable 2.2, "Improvement of SE performances by PMUs and IEDs," European Community's 7th Framework Programme, PEGASE project (Pan European Grid Advanced Simulation and state Estimation), 2012, http://fp7-pegase.eu/

[52] J. M. Ruiz Muñoz, A. Gómez-Expósito, "A Line-current Measurement Based State Estimator," IEEE Transactions on Power Systems, vol. 7, pp. 513-519, May 1992.

[53] F. C. Schweppe , E. J. Handschin, "Static State Estimation in Electric Power Systems,", Proceedings of the IEEE vol. 62 , no. 7, pp. 972-982, July 1974.

[54] F. C. Schweppe, J. Wildes, D. B. Rom, "Power system static state estimation, Part I, II, III," IEEE Transactions on Power Apparatus and Systems, vol. PAS-89, no. 1, pp. 120-135, January 1970.

[55] H. Singh, F. L. Alvarado, W. Liu,"Constrained LAV State Estimation Using Penalty Functions," IEEE Transactions on Power Systems, vol. 12, no. 1, pp. 383-388, February 1997.

[56] I.W. Slutsker, S. Mokhtari, K.A. Clements, "Real Time Recursive Parameter Estimation in Energy Management Systems," IEEE Transactions on Power Systems, vol. 11, no. 3, pp. 1393-1399, August 1996.

[57] B. Stott, "Fast Decoupled Load Flow," IEEE Transactions on Power Apparatus and Systems, vol. PAS-93, no. 3, pp. 859 - 869, May 1974.

[58] B. Stott, O. Alsac, A. J. Monticelli, "Security Analysis and Optimization," Proceedings of the IEEE , vol. 75, no. 12, pp. 1623-1644, December 1987.

[59] W. Tinney, R. Walker, "Direct Solutions of Sparse Network Equations by Optimally Ordered Triangular Factorization," Proceedings of the IEEE, vol. 55, pp. 1801-1809, 1967.

[60] I. H. Valenzo, "Information Architecture Design for the Electricity Distribution Network," Master thesis, Delft University of Technology, 2009.

[61] T. Van Cutsem, J. L. Horward, M. Ribbens-Pavella, "A Two-level Static State Estimator for Electric Power Systems," IEEE Transactions Power Apparatus and Systems, vol. PAS-100, no. 8, pp. 3722-3732, August 1981.

[62] T. Van Cutsem, M. Ribbens-Pavella, "Critical Survey of Hierarchical Methods for State Estimation of Electric Power Systems," IEEE Transactions on Power Apparatus and Systems, vol. 102, pp. 3415-3424, 1983.

[63] M. Vempati, I. Slutsker, W. Tinney, "Enhancements to Givens Rotations for Power System State Estimation," IEEE Transactions on Power Systems, vol. 6, no.2, pp. 842-849, May 1991.

[64] A. de la Villa Jaén, P. Cruz Romero, A. Gómez-Expósito, "Substation Data Validation by Local Three-phase Generalized State Estimators," IEEE Transactions on Power Systems, vol. 20, no. 1, pp. 264-271, February 2005.

[65] A. de la Villa Jaén, A. Gómez-Expósito, "Implicitly Constrained Substation Model for State Estimation," IEEE Transactions on Power Systems, vol. 17, no. 3, pp. 850-856, August 2002.

[66] J. Weiqing , V. Vittal, G.T. Heydt, "A Distributed State Estimator Utilizing Synchronized Phasor Measurements", *IEEE Transactions on Power Systems*, vol. 22, No. 2, pp. 563-571, May 2007.

[67] F. F. Wu, "Power System State Estimation: A Survey," International Journal of Electrical Power and Energy Systems, vol. 12, no. 2, pp. 80-87, April 1990.

[68] B. Xu, A. Abur, "Observability Analysis and Measurement Placement for System with PMUs," IEEE PES Power Systems Conference & Exposition, New York, Oct.10-13, 2004.

[69] B. Xu, J. Y. Yeo, A. Abur, "Optimal Placement and Utilization of Phasor Measurements for State Estimation", Power Systems Computation Conference, Liege (Belgium), August 22-26, 2005.

[70] L. Zhao , A. Abur, "Multiarea State Estimation Using Synchronized Phasor Measurements", *IEEE Transactions on Power Systems*, vol. 20, No. 2, pp. 611-617, May 2005.

[71] Jun Zhu, A. Abur, "Effect of Phasor Measurements on the Choice of Reference Bus for State Estimation," Proceedings of the IEEE PES General Meeting, June 24-28, 2004, Tampa, FL.

[72] Jun Zhu, A. Abur, "Identification of Network Parameter Errors Using Phasor Measurements," Proceedings of the IEEE Power and Energy Society General Meeting, Calgary, CA, July 26-30, 2009.

Thesis Publications

[**T1**] A. Gómez-Expósito, A. de la Villa Jaén, C. Gómez-Quiles, P. Rousseaux and T. Van Cutsem, "A Taxonomy of Multi-Area State Estimation Methods", **Electric Power Systems Research**, vol. 81, pp. 1060-1069, 2011.

[**T2**] A. Gómez-Expósito; Abur, A.; A. de la Villa Jaén, A.; C. Gómez-Quiles, "A Multilevel State Estimation Paradigm for Smart Grids", **Proceedings of the IEEE**, vol.99, no.6, pp.952-976, June 2011.

[**T3**] C. Gómez-Quiles, A. de la Villa Jaén, A. Gómez-Expósito, "A Factorized Approach to WLS State Estimation", **IEEE Transactions on Power Systems**, vol.26, no.3, pp.1724-1732, Aug. 2011.

[**T4**] C. Gómez-Quiles, A. Gómez-Expósito, A. de la Villa Jaén, "State Estimation for Smart Distribution Substations", submitted to **IEEE Transactions on Smart Grids**.

[**T5**] A. Gómez-Expósito, A. Abur, P. Rousseaux, A. de la Villa Jaén, C. Gómez-Quiles, "On the Use of PMUs in Power System State Estimation", Proc. of the Power System Computation Conference (PSCC), Stockholm, Sweden, 2011.

[**T6**] A. Gómez-Expósito, C. Gómez-Quiles, A. de la Villa Jaén, "Bilinear Power System State Estimation", to appear in **IEEE Transactions on Power Systems**; available in IEEE Xplore.

[**T7**] C. Gómez-Quiles, H. Gil, A. de la Villa Jaén, A. Gómez-Expósito, "Equality-Constrained Bilinear State Estimation", submitted to **IEEE Transactions on Power Systems**.

Electric Power Systems Research 81 (2011) 1060–1069

Contents lists available at ScienceDirect

Electric Power Systems Research

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

Short survey

A taxonomy of multi-area state estimation methods

Antonio Gómez-Expósito [a,*], Antonio de la Villa Jaén [a], Catalina Gómez-Quiles [a], Patricia Rousseaux [b], Thierry Van Cutsem [b,1]

[a] Department of Electrical Engineering, University of Seville, Camino de los Descubrimientos s/n, 41092 Seville, Spain
[b] Department of Electrical Engineering and Computer Science (Montefiore Institute) of the University of Liège, Sart Tilman B37, B-4000 Liège, Belgium

ARTICLE INFO

Article history:
Received 21 July 2010
Received in revised form 27 October 2010
Accepted 16 November 2010
Available online 19 December 2010

Keywords:
Multi-area state estimation
Hierarchical state estimation
Distributed state estimation
Factorized state estimation

ABSTRACT

This paper presents a critical review of the state of the art in multi-area state estimation (MASE) methods, which are currently gaining renewed interest due to their capability of properly tracking multi-TSO transactions and accommodating highly redundant information systems. Based on several classification criteria, a taxonomy of MASE methods is first proposed. Two main categories, namely two-step or hierarchical versus decentralized, are identified. Then, for each class of methods, the resulting model structure and area interactions are discussed and a brief presentation is made of a selected subset of references.

1. Introduction

State estimators (SE) determine the most likely state of a power system from sets of measurements which are captured remotely at substations and collected periodically by SCADA systems via remote terminal units (RTU). The information provided by the SE is crucial in nowadays energy management systems (EMS), where a diversity of applications dealing with the economic and secure operation of transmission networks rely on accurate and continuously updated snapshots of the system. The new regulatory paradigm arisen in the last decade has even stressed the importance of the SE tool, in an open-access context in which many more transactions on much more congested networks have to be properly tracked.

Research on multi-area state estimation (MASE) can be traced back to the late 1970s, shortly after state estimators started being put into service. It addressed the problem of performing efficient state estimation on large power systems, with the twofold motivation of gaining computing time (under the then-limited computational resources) and exploiting the fact that real-time measurements are gathered within areas by the various control centers distributed over the grid. MASE relies on some kinds of decomposition–coordination scheme, taking advantage of the usually weaker geographical or measurement coupling among areas, in combination with well-established solution methods.

MASE has regained interest over the last decade owing to projects of having central entities monitoring large interconnected systems. This is the case, for instance, of Independent System or Regional Transmission Operators in the U.S. [1], while projects of supra-national centers monitoring the European grid start becoming a reality (e.g. Coreso [2]). Furthermore, large-scale incidents experienced over the recent years have stressed the need for a better real-time visibility of the operating state of the grid well beyond the extent covered by the state estimator of a single country or company. For reliability, computational efficiency and model maintenance reasons, it does not sound reasonable to collect and process the huge set of data of those large grids at a single place. Hence, the idea of MASE.

This paper proposes a taxonomy of MASE methods, offering a unifying description of a relatively large number of works devoted to the subject, and probably a few more to come, with the advances in phasor measurement technology. It significantly enhances and updates the only known survey on the topic, published early in the eighties [3]. Due to space limitations, it has not been possible to exhaustively include all publications in the comparative analysis. Neither does the paper deal with the interesting related problems of bad data identification and observability analysis.

Other applications of parallel and distributed processing to power systems are quoted in [4].

The structure of this paper is as follows. Section 2 briefly reviews the conventional nonlinear SE formulation, Section 3 lists the most

* Corresponding author.
E-mail address: age@us.es (A. Gómez-Expósito).
[1] He is with the Fund for Scientific Research - FNRS.

0378-7796/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.epsr.2010.11.012

important desirable features of MASE and Section 4 defines the notation adopted in the paper. The criteria used to classify MASE methods are defined in Section 5. A unified description of the hierarchical and the decentralized approaches is presented in Sections 6 and 8, respectively, while the corresponding literature surveys are given in Sections 7 and 9. Section 10 offers the conclusions.

2. Background on state estimation

The SE relies on the following measurement equation [5]:

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

where:

x is the state vector to be estimated (size n),
z is the known measurement vector (size $m > n$),
h is the vector of functions, usually nonlinear, relating error free measurements to the state variables,
e is the vector of measurement errors, customarily assumed to have a Normal distribution with zero mean and known covariance matrix R. When errors are independent R is a diagonal matrix with values σ_i^2, where σ_i is the standard deviation of the error associated with measurement i.

In conventional bus-branch SE models the state vector is composed of voltage magnitudes and phase angles, whereas the measurement vector typically comprises power injections, branch power flows and voltage magnitudes. Recently, the availability of synchro-phasors (PMUs) has made it possible to incorporate phase angle measurements into the SE process.

The weighted least squares (WLS) estimator minimizes:

$$J = \sum_{i=1}^{m} W_i r_i^2$$

where:

$r_i = z_i - h_i(\hat{x})$ is the measurement residual,
$\hat{x}$ is the estimated state vector, and
W_i is the respective weighting coefficient.

The state estimate can be obtained by iteratively solving the so-called Normal equations:

$$G_k \Delta x_k = H_k^T W [z - h(x_k)] \tag{2}$$

where:

$H_k = \partial h / \partial x$ is the Jacobian evaluated at $x = x_k$,
$G_k = H_k^T W H_k$ is the gain matrix,
$W = \mathrm{diag}(W_i)$ is the weighting matrix,
$\Delta x_k = x_{k+1} - x_k$, k being the iteration counter.

Iterations are terminated when an appropriate tolerance is reached on Δx_k. The covariance of the estimate is:

$$\mathrm{Cov}(\hat{x}) = G_k^{-1}$$

provided the covariance σ_i^2 is used as weight W_i.

Then, the bad data processing function, aimed at detecting, identifying and eliminating bad analog measurements, is activated. This usually relies on the normalized residual test [6].

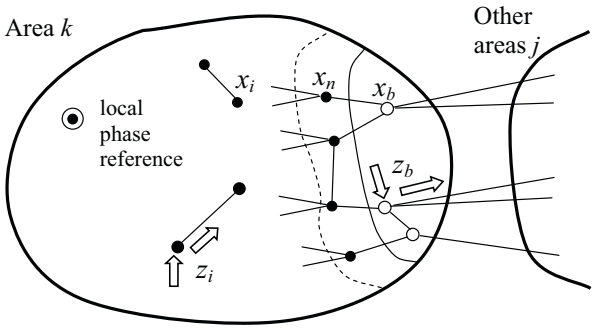


Fig. 1. Classification of measurements and state variables.

3. Desirable features of MASE

In principle, processing the whole measurement set in a single WLS estimator provides the "optimal" estimate. This conventional scheme is referred to as "integrated" in the sequel. This section enumerates the most desirable features that a MASE should exhibit compared to the integrated scheme.

- *Robustness*: the capability to converge to an acceptable solution under a wide range of circumstances (topology, measurement configuration, the presence of large bad data, etc.). Problem decomposition may affect convergence properties of MASEs.
- Accuracy: the estimate has to be accurate enough for operating purposes. Ideally, it should be the same as that with the integrated scheme. In practice, optimality is not a finality per se: other advantages gained out from decomposed approaches can be considered as more important, provided accuracy remains acceptable of course.
- High computational efficiency: increased speed is one of the expected attractive features of MASE, provided that the algorithmic complexity linked to the decomposition–coordination remains limited.
- Amount of data exchange: information exchange between processors should be kept as low as possible in order to avoid possible time delays and thus an increase in the estimator response time and/or require a larger communication bandwidth. Attention should also be paid to not having many measurements shared by more than one computer, which increases the complexity of the data acquisition process.
- Bad data analysis: the bad data rejection capability should be preserved, more specifically in the proximity of area boundaries. The possibility to easily compute the normalized residuals is another important issue.

4. Nomenclature

This section provides the common definitions and notation that will serve to understand the diverse MASE procedures described below, some of them differing only in subtle details, which are frequently hidden by the particular jargon used in the original publication.

Consider an arbitrary area k, directly connected to other areas through tie-lines, as shown in Fig. 1. The state variables of area k can be classified as follows:

- x_{ik}: internal variables, associated with buses which are not terminals of tie-lines;
- x_{bk}: border variables, associated with terminal buses of tie-lines;
- x_{nk}: subset of variables in x_{ik} associated with first neighbors of border buses. In certain cases, this subset may contain second and, eventually, deeper neighbors of border buses, depending

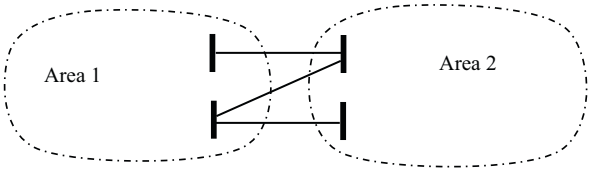


Fig. 2. Non-overlapping areas.

on the amount of information that adjacent areas are willing to exchange.

The phase angles of buses in area k (internal and border buses) are referred for convenience to that of an arbitrary local reference bus. Such relative phase angles are involved in the expression of any power flowing in internal branches. However, power flows through tie-lines, as well as power injections at tie-line terminal buses, involving buses of different areas, require that absolute phase angles be used. Therefore, for a system comprising A interconnected areas, an additional set of variables, termed u, is introduced containing the phase angle differences of $A-1$ local references with respect to the remaining one, chosen as the absolute reference.

To distinguish between both references, a vector x will denote state variables referred to the global phase reference, whereas a vector y will be used when voltage phase angles refer to their respective local reference.

A similar classification of available measurements can be made, again for any given area k, as illustrated also in Fig. 1:

- z_{ik}: internal measurements, exclusively related with variables x_{ik} (or y_{ik}) and x_{bk} (or y_{bk}) in area k;
- z_{bk}: border measurements, functions of x_{bk} and the vectors x_{bj} of other areas ($j \neq k$), and possibly x_{nk}.

5. Classification criteria of MASE

MASE methods can be classified and discussed according to a variety of criteria detailed hereafter.

5.1. Area overlapping level

MASE procedures are based on system decomposition into interconnected areas. Depending on the coordination strategy adopted to reconcile local estimates, the degree of overlapping, i.e. the number of buses and/or branches in common between two adjacent areas, may range from zero to several layers of border buses and associated branches, as explained below:

(1) *Non-overlapping areas* have no bus and no branch in common; they are connected by tie-lines ending at border buses (see Fig. 2). Those tie-lines define the *interconnection* area.
(2) *Border-bus overlapping areas* are adjacent areas overlapping over just one layer of border buses (see Fig. 3); there is no tie-line connecting two areas. This situation can be artificially created from the previous case, defining a virtual border bus at the mid-point of each tie-line and extending each area up to this new bus (see Fig. 4). This particular decomposition will be referred to as "mid-point virtual bus overlapping areas". By so doing, one gets rid automatically of boundary injection measurements.
(3) *Tie-line overlapping areas* share tie-lines and the corresponding border buses (see Fig. 5).
(4) *Extended (or deep) overlapping areas* share several layers of neighbors of border buses (see Fig. 6).

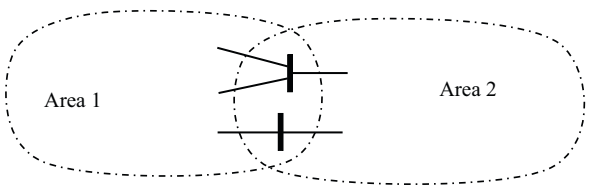


Fig. 3. Border-bus overlapping areas.

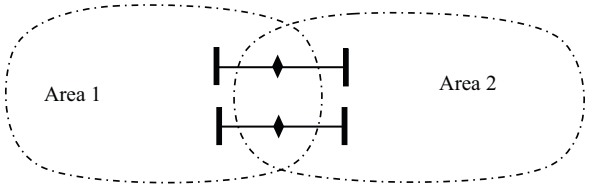


Fig. 4. Mid-point virtual bus overlapping areas.

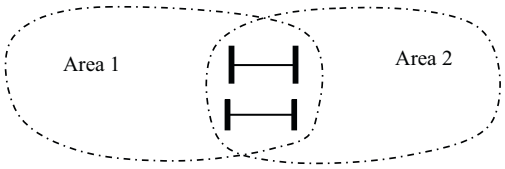


Fig. 5. Tie-line overlapping areas.

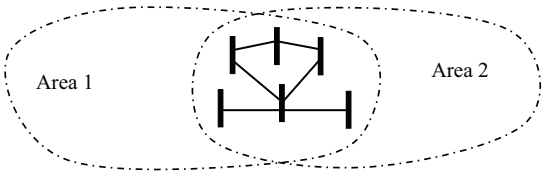


Fig. 6. Extended overlapping areas.

For classification purposes, the following simplified terminology will be adopted: non-overlapping areas, minimally overlapping areas (item 2) and fully overlapping areas (items 3 and 4).

5.2. Computing architecture

Two computer architectures and hardware environments are relevant:

(1) In a *hierarchical scheme* (see Fig. 7), a *master* processor distributes the work among *slave* computers performing local area SE and, subsequently, coordinates the local estimates. In this scheme, slave processors, which can be located remotely (distributed architecture) or at the same physical place (parallel architecture), communicate only with the central computer.
(2) in a *decentralized architecture* (see Fig. 8), there is no central computer; each local processor communicates only with those processors in charge of neighboring areas, exchanging border

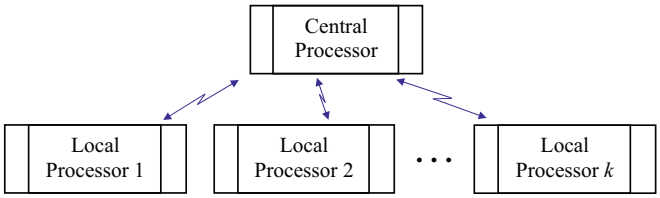


Fig. 7. Hierarchical architecture.

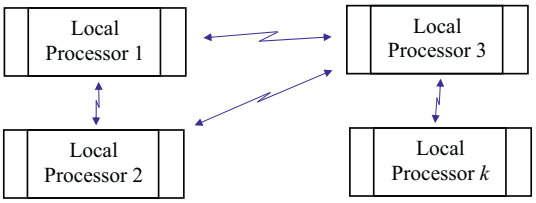


Fig. 8. Decentralized distributed architecture.

information. Usually, the iterative estimation algorithm relying on this architecture does not involve any coordination phase.

5.3. Coordination scheme

Depending on the stage at which the local estimates get coordinated, three coordination schemes can be distinguished.

(1) *Coordination at the SE level*: each area submits its results to the central processor after full convergence of its local SE. Methods relying on this coordination type presents the following essential features:
The central level refines and coordinates local estimates by computing the u variables. When resorting to a single coordination, this strategy leads to suboptimal estimates, but usually of acceptable accuracy; communication bandwidth requirements are kept low by limiting the data exchanges between the central and the local processors; independent SEs allow simple implementation at the local level with minor adaptations to existing SE software packages.
(2) *Coordination at the iteration level*: results are submitted for coordination after each iteration of the local SE, with the following main features:
Repeated coordinations allow faster convergence to the optimal solution than in the previous scheme; the price to pay is an important increase in data exchange between the central and the local computers or between neighboring areas in the decentralized architecture; generally, it requires more complex code implementations, and does not allow to keep existing SE algorithms.
(3) *Hybrid coordination* is a compromise between the above two schemes, in which several local iterations are performed before coordination.

5.4. Measurement synchronization

When considering only conventional power and voltage magnitude measurements, synchronization between areas may be an issue. Either we look for similar accuracy to that of an integrated scheme or we accept some suboptimality. In the former case, it is important to synchronize the various measurement gatherings so that the hierarchical scheme does not add to the time skew present in conventional data acquisition (e.g., a synchronizing signal could be sent to the various local computers). If suboptimality is accepted, this synchronization requirement can be somewhat relaxed.

Incorporating information provided by PMU provides new possibilities for measurement synchronization. In theory, using a whole set of measurements coming from PMU would allow full synchronization of snapshots captured by all areas. In practice, however, this may be too cumbersome or even impossible.

5.5. Process synchronization

Local processors may run in a synchronous or asynchronous manner. In the decentralized architecture, the process is asynchronous by nature: each local processor performs iterations at its own pace, using the best information available from its neighbors. In the hierarchical scheme, two synchronization modes are possible:

- when coordination takes place at the SE level, there is no need to impose synchronous local estimations so that each local processor usually runs its SE at its own rate. The central processor always uses the last estimate provided by each area, no matter if those local estimates do not refer to the same time instant. This time skew among local estimates may affect the accuracy of the final solution;
- when coordination takes place at the iteration level, the central computer may (synchronous) or may not (asynchronous) wait for the slowest local processor to complete the current iteration, before coordinating the solution.

5.6. Solution methodology

Various general methodologies can be used to solve the decomposition–coordination problem involved in MASE. A large majority of methods rely on the classical WLS formulation. Most hierarchical schemes rely, at both the local and the central levels, on an iterative scheme to solve the Normal equations of concern. In decentralized schemes, with possible coordination at the iteration level, Lagrangian relaxation-based algorithms are usually adopted. Some works also introduce certain heuristics intended to simplify the optimality conditions of the coordination problem. A category of MASE approaches formulate the WLS equations as an optimization problem, usually involving a Lagrangian function explicitly handling constraints imposed by network equations and/or boundary conditions.

Among the above criteria, the one related to computing architecture is adopted to further classify existing MASE methods in two main categories: two-step hierarchical vs. decentralized MASE, respectively. In the following sections, for each category, a short presentation is followed by a literature review.

6. Two-step hierarchical MASE

In this section, we consider hierarchical SE methods comprising two main steps: (1) local solutions are obtained at the area level, ignoring totally or partly the information and constraints stemming from neighboring areas; (2) local solutions are coordinated by a central processor in order to cope with the interactions among areas ignored during the first stage. Each step is separately discussed below.

6.1. First step: local solutions

During the first stage each area k independently solves a local WLS problem based on the measurement equation:

$$z_k = h_k(y_k) + e_k \tag{3}$$

where phase angles in y_k refer to the local phase reference of area k. The components of y_k and z_k are determined by the overlapping level adopted, as follows:

- Non overlapping:

$$y_k = \begin{bmatrix} y_{ik} \\ y_{bk} \end{bmatrix}; \qquad z_k = z_{ik}$$

By definition, each state variable is estimated only once, within its respective area. Note that tie-line power flows and border injec-

tions cannot be individually used in this scheme, as they involve external state variables. However, z_k may additionally include certain pseudomeasurements, obtained by previously processing raw border measurements. For instance, if the injection at a given border bus, along with all power flows leaving the incident tie-lines, are measured, then a pseudomeasurement representing the net power injection can be added to z_k.

- Border-bus overlapping:

$$y_k = \begin{bmatrix} y_{ik} \\ y_{bk} \end{bmatrix}; \qquad z_k = z_{ik}$$

The tie-line notion vanishes in this case, as every line fully lies within a single area. Accordingly, all power flow measurements can be handled by their respective local estimator. However, border injections cannot be handled at this level, as they involve external variables. Note that several (at least two) estimates result for the same border variables, each one referred to its own phase reference.

- Tie-line overlapping:

$$y_k = \begin{bmatrix} y_{ik} \\ y_{bk} \\ y_{bj} \end{bmatrix}; \qquad z_k = \begin{bmatrix} z_{ik} \\ z_{bk} \end{bmatrix}$$

In this case, border injection measurements can be used within their respective area and the possibility exists for tie-line power flow measurements to be used twice. Like in the former case, border state variables are estimated at least twice, each one referred to its own phase reference.

- Extended overlapping:

$$y_k = \begin{bmatrix} y_{ik} \\ y_{bk} \\ y_{bj} \\ y_{nj} \end{bmatrix}; \qquad z_k = \begin{bmatrix} z_{ik} \\ z_{bk} \\ z_{bj} \end{bmatrix}$$

In this case, the state vector includes the subset y_{nj}, which means that border injections of neighboring areas can be also added to z_k.

In summary, the geographical scope of each local problem determines the set of variables included in the state vector and, consequently, the set of measurements that can be locally handled.

Once the iterative process has converged, the local solution $\tilde{y}_k$ is obtained, along with its covariance matrix:

$$\mathrm{Cov}(\tilde{y}_k) = G_k^{-1} = (H_k^T W_k H_k)^{-1}$$

where H_k is the Jacobian of h_k, computed during the last iteration, and W_k the respective weighting submatrix.

Note that the estimates $\tilde{y}_{ik}$ of variables sufficiently distant from the border are generally optimal for practical purposes after the local solution process, as the influence of border measurements is negligible. However, the estimates $\tilde{y}_{bk}$ and, to a lesser extent, $\tilde{y}_{nk}$ are suboptimal, unless the border measurements are very accurate.

6.2. Second step: centralized coordination

The geographical scope of this step is determined by the overlapping level adopted and the amount of information each area is willing to exchange with its neighbors.

The state vector at this stage comprises at least the border variables corresponding to the A areas, along with the phase reference vector u introduced above, i.e.

$$x_c = \begin{bmatrix} x_b \\ u \end{bmatrix} \quad \text{with } x_b = [x_{b1}, \ldots, x_{bk}, \ldots, x_{bA}]^T$$

while the measurement vector at this step includes two components:

$$z_c = \begin{bmatrix} \tilde{y}_b \\ z_b \end{bmatrix}$$

where $\tilde{y}_b$ is the local estimate of x_b, with phase angles referred to local buses, and

$$z_b = [z_{b1}, \ldots, z_{bk}, \ldots, z_{bA}]^T$$

represents the set of border measurements not yet used at the first step. Note that, for each component of the state vector x_b two or more estimates may be available in $\tilde{y}_b$, which increases the redundancy to estimate the vector u.

In the extended overlapping case, or when border injections are to be handled at this stage, first neighbors of border buses are also involved, leading to the augmented state vector:

$$x_c = \begin{bmatrix} x_b \\ x_n \\ u \end{bmatrix} \quad \text{with } x_n = [x_{n1}, \ldots, x_{nk}, \ldots, x_{nA}]^T$$

In this case, z_c should also contain the component $\tilde{y}_n$ of the local estimates, to assure observability.

Including second and further neighbors of border buses in the state vector improves the optimality of the solution after the second step, but leads to heavier information exchange and increased computational effort.

The measurement model of the coordination phase is composed of a nonlinear system:

$$z_{bk} = h_{ck}(x_c) + e(z_{bk}), \quad k = 1, \ldots, A \tag{4}$$

together with a linear one:

$$\tilde{y}_b = x_b - Bu + e(y_b) \tag{5}$$

where the entries of B are equal to zero for the voltage magnitude components of x_b and equal to one for phase angle components.

In the above measurement model the covariance of $e(z_{bk})$ is known and that of $e(y_b)$ is contained in the respective submatrix of $\mathrm{Cov}(\tilde{y})$. As this is a dense matrix, it is customary to obtain and retain only its diagonal elements, leading to suboptimal estimates at the coordination step.

The iterative solution of the WLS problem arising from (4) and (5) provides the estimate $\hat{x}_c$. Based on this estimate it is theoretically possible to further refine the values of other internal variables, particularly those close enough to the border, but this is seldom considered in the literature.

7. Literature survey on hierarchical MASE

Table 1 provides a classification of hierarchical MASE methods, according to the different criteria identified in Section 5.

Due to space limitations only the most representative proposals are briefly presented. For this purpose they are grouped in two separate subsections, depending on whether the coordination phase takes place only once, at the end of the local SE process, or the local solutions are repeatedly coordinated after each iteration.

7.1. Coordination at the state estimation level

In 1972, Clements et al. [7] developed a technique relying on border-bus overlapping. After the local estimation processes finish, the coordinating phase reconciles local estimates, which differ at

Table 1
Classification of hierarchical MASE methods.

Reference, author	Area overl. [a]	Solut. meth. [b]	Estim. state [c]	Meas. type [d]	Coord. scheme [e]
[7], Clements	MO	NE	Sub	C	SE
[8], Kobayashi	NO	NE	Opt	C	It
[9], Irving	NO	NE–H	Sub	C	SE
[10], Marsh	NO	NE–O	Opt	C	It
[11], Van Cutsem	NO	NE	Sub	C	SE
[12], Wallach	NO	NE–H	Sub	C	SE
[13], Brice	NO	NE–R	Opt	C	It
[14], Mukai	NO	NE–R	Opt	C	It
[15], Kurzin	NO	NE–H	Sub	C	SE
[16], Lo	NO	NE–H	Sub	C	SE
[17], Iwamoto	NO	NE–O	Sub	C	It
[18], El-Keib	NO	NE–H	Sub	C	SE
[19], Ahmed	NO	NE–H	Sub	C	It
[20], El-Keib	NO	O	Opt	C	It
[21], Falcao	MO	NE–H	Sub	C	SE–It
[22], Korres	NO	NE	Sub	C	SE
[23], Ebrahimian	MO	NE–O	Opt	C	It
[24], Aguado	MO	NE–O	Opt	C	It
[25], Zhao	FO	NE	Sub	P	SE
[27], Yan	NO	NE–O	Sub	P	SE
[28], Jeffers	FO	NE–H	Sub	P	SE
[29], Jiang	FO	NE–H	Sub	P	SE
[30], Patel	FO	NE	Sub	P	SE
[31], Jiang	NO	NE–O	Opt	P	It
[32], Gómez	NO	NE	Opt	C	SE
[33], Korres	NO	NE–O	Opt.	C	It

[a] NO: non-overlap; MO: minimally overlap; FO: fully overlap.
[b] NE: Normal equations; R: relaxation; O: optimization; H: heuristic.
[c] Opt: optimal; Sub: suboptimal.
[d] C: conventional only; P: considers PMU.
[e] SE: SE level; It: iteration level.

common buses. Local state vectors are then re-estimated through a non-iterative WLS procedure, using as pseudo-measurements the previous local estimates along with the estimates at common buses obtained from neighboring areas. Power injection measurements at boundary nodes are not taken into account.

In 1977 and 1979, Irving and Sterling [9] proposed to extend the technique of "diakoptics" or network tearing to the state estimation problem. This technique, previously applied to the load flow problem, exploits the nearly block diagonal structure of the Jacobian matrix, diagonal blocks corresponding to non overlapping regions. Links between blocks come from interconnection measurements. Initial local estimates are corrected according to the diakoptics formulae in order to take into account information brought by interconnection measurements. The estimate is supposed to be obtained in only one iteration. As the exact application of diakoptics leads to cumbersome calculations, the authors proposed mathematical simplifications without providing some physical interpretation.

In 1981, Wallach et al. [12] set up a method intended for parallel computation. The original network is divided into disjoint areas, upon the condition that each resulting area is observable and power flow measurements exist at both ends of tie lines. Such power flows are equivalent to power injections from the point of view of the internal area. A slave processor performs the local WLS solution for each area, in a decoupled manner. Then, the master processor simply adjusts the area phase angle references so that all phase angles refer to a single phase origin. For this purpose, a simplistic technique is adopted, consisting of computing the phase angle difference across one of the tie lines linking adjacent areas.

In 1981 also, Van Cutsem et al. [11] developed a two-level estimator based on non-overlapping areas. Only internal measurements are used at the area level, while the coordination level is aimed at estimating tie-line power flows through vector u. In addition to tie-line power flow measurements, local estimates at the boundary buses are added as pseudo-measurements to the coordination phase, for which approximate covariances are adopted (diagonal elements of the respective matrix).

In 1983, Kurzin [15] proposed a method conceptually very similar to that of Van Cutsem. The difference lies in the way coordination is carried out. A heuristic approach using arithmetic means is used. The objective is to reduce the size and as a consequence the computing time, of the coordination estimation. From the active power flow measurements through tie lines, the phase angle differences across those lines are estimated. Then, the average of the computed values, in case there are two or more tie lines, is taken as the final estimate.

In the first part of the 1988 paper by Lo et al. [16], six approaches to coordinate the solution of local state estimators (the first five directly derived from the previous works by Van Cutsem et al. and Kurzin) are analytically formulated and reviewed. In the second part, the performance of those algorithms is experimentally tested. A fast decoupled SE formulation is assumed.

In the work of El-Keib et al. in 1990 [18], an extensive review of hierarchical SE methods formerly published is first presented. Then, an improvement to the method of Van Cutsem et al. is proposed regarding the coordination phase. The use of modified injections at boundary buses is suggested, by invoking Kirchhoff's law at those buses. This requires extending the domain of the measurement vector to include the estimates of power flows in lines connecting the boundary buses with internal buses, increasing the redundancy at the upper level and improving the overall results of the state estimator. No tests are provided of this idea.

In 1995, Falcao et al. [21] proposed a two-step solution methodology combining conventional state estimation algorithms with an optimization technique to cope with coupling constraints at the centralized level. Boundary buses belong to adjacent overlapping areas. The number of boundary buses may be kept to a minimum,

or may incorporate a few extra (internal) buses in order to facilitate bad data processing. It is assumed that there are no boundary injections (fictitious null-injection buses can be artificially created to achieve this). A decoupled estimator is used incorporating a set of linear constraints, which are introduced to force state variables in overlapping areas to assume the same values. Then, certain off-diagonal blocks are ignored in the matrices arising in the normal equations, in such a way that only boundary state variables are involved when updating the state variables within each area.

Several coordination schemes are tested: (1) apply the coupling constraint terms after every iteration of the local estimators (this algorithm actually belongs to the category of methods described in the following section); (2) first allow the local estimators to converge to the desired tolerance, and then apply the coupling constraint corrections without any further local estimation iterations; (3) allow the local estimators to converge to a tolerance relatively close to the desired one, and from then on, use an alternating iterative scheme between the local estimations and the coupling constraints corrections (this is a hybrid algorithm between those described in this and the next section). Under scheme 1, each area SE is performed in a synchronous manner, requiring coordination at each iteration. This requires a lot of information exchange. On the other hand, scheme 2 is of an asynchronous nature, as computations can continue even in the absence of information from other areas.

In the work of Korres and Contaxis, in 2000 [22], the entire system is divided into non-overlapping areas. Each area is individually solved, disregarding boundary measurements so that the state variables of adjacent areas do not show up. Then, at the second level, a reduced model involving tie lines and boundary measurements is handled. In this reduced model, the areas become "supernodes" linked by tie lines, while "equivalent" power flow measurements, representing the sum of power flows on the tie lines incident to each boundary bus, are used. The results of the second step are not used subsequently to update the estimates of internal nearby buses.

In [1] the use of Internet is suggested to exchange the necessary information during the coordination phase. In order to obtain phase angles of each area with respect to a global reference, phase differences across tie-lines are computed, and then the average is taken for the set of tie-lines connecting each couple of areas. The authors state that "the average scan time to get SE outputs may be of the order of 4 minutes, and therefore the Internet is well suited as the communication medium". This may not be true, however, in all cases.

All of the above methodologies provide suboptimal estimates as they imply neglecting in local SEs some nearby external and/or boundary measurements. However, for most of them, test results show acceptable accuracy in normal conditions. Besides, bad data close to boundaries, if not properly identified and removed, can degrade the accuracy of boundary and nearby variables and also, in some cases, affect convergence of the distributed SE process.

7.2. Coordination at the iteration level

In 1974, Kobayashi et al. [8] applied the model coordination principle of Mezarovic [34] from hierarchical system theory to the problem of state estimation. System decomposition into non-overlapping areas, connected by tie-lines, is considered. One iteration of the overall process consists of three steps: (1) using internal measurements solve local WLS problems to update internal variables, with u and border state variables fixed to their most recent values; (2) using interconnection measurements solve a WLS problem for the interconnection area to update u, with all state variables fixed to their current values; (3) using the whole set of measurements, coordinate previous steps by computing border state variables, with u and internal state variables fixed to their current values.

In 1979, Marsh and Cristi [10] formulated the WLS state estimation as an optimization problem with equality constraints. The algorithm is developed while assuming the system composed of several interconnected areas. The equality constraints come on one hand, from the power balance equations at each node, and on the other hand, from the diakoptical formulation of the interconnections between subsystems. They are taken into account by introducing Lagrange multipliers. Minimization is performed by means of a steepest descent algorithm. Each iteration requires a high amount of data transfer while proper convergence may lead to a relatively large number of iterations. Bad data analysis is not considered and appears to be a challenging problem.

In 1981, Brice and Cavin [13] developed a hierarchical distributed algorithm as follows: at each iteration, the central processor broadcasts the current value of the state vector. Then, each satellite processor computes its portion of both the Jacobian matrix and the right hand side vector of Eq. (2), which are sent back to the central processor. Next, all matrix and vector components are gathered at the central level, where the Normal equations are built and solved, and the process is repeated. Therefore, only a small fraction of the computations is parallelized, but the optimal solution is obtained in the same number of iterations as in the conventional formulation. In order to increase the amount of computations performed in parallel, the solution of the Normal equations can be also distributed among existing processors by applying a Gauss-Seidel iterative method, instead of using a direct solution approach based on Cholesky factorization.

The outstanding feature of the work by Iwamoto et al., in 1989 [17], is that rectangular coordinates, along with an extension of the popular second-order load flow algorithm, previously developed by Iwamoto and Tamura, are resorted to. The main advantage is that a constant Jacobian matrix naturally arises, significantly reducing the computational burden. In the problem formulation the interaction among subsystems is taken into account through the tie-line bus voltages. The hierarchical structure of the proposed method consists of two levels: the upper level, where the optimal tie-line bus voltages are evaluated; and the lower level, where the optimal states of each subsystem are determined by minimizing a cost function that involves the entire system. This constitutes a significant difference from a majority of published methods, first solving the local levels and then coordinating the solution at the upper level.

El-Keib et al. [20] presented in 1992 a multi-area approach for solving the Weighted Least Absolute Value (WLAV) state estimation problem. The proposal is based on the Dantzig-Wolfe Decomposition Principle, applied to the resulting LP problem. Each subproblem is solved independently by the revised simplex method and sends the solution to a master problem. The master problem incorporates this information and its solution is used to modify the objective functions of the subproblems. The process is continued till optimality is reached.

In 1995, Falcao et al. [21] proposed the three coordination schemes described in the previous section, one of them exchanging information at the iteration level.

Ebrahimian and Baldick [23] applied in 2000 the so-called "Auxiliary Problem Principle", a well-known technique in large-scale optimization, to develop a two-step procedure based on border-bus overlapping areas. The overall WLS objective function is expressed as the sum of partial functions, each one corresponding to a subsystem. Additional equality constraints are considered expressing the fact that the estimates of boundary buses in adjacent areas should be identical. The problem is decomposed by linearizing the augmented Lagrangian, yielding an iterative procedure in which subsystem solutions (involving several inner iterations) are alternated with Lagrange multiplier updates (outer loop). When solving

each subsystem, the border information is interpreted as a set of pseudo-measurements. The optimal solution is approached as outer iterations progress, but several constants should be tuned for a particular system to guarantee convergence.

A similar approach was adopted by Aguado et al. in 2001, who also tested a Lagrangian Relaxation-based iterative scheme [24].

In general, other SE related issues, such as observability analysis and bad data handling are not fully addressed in papers dealing with MASE methods. A notable exception is the recent work [33], where distributed solutions to those problems are proposed.

7.3. Incorporation of PMUs

The incorporation of PMUs in a hierarchical SE process was first considered by Zhao and Abur in 2005 [25]. The resulting areas overlap over tie-lines allowing boundary measurements (injections and tie-line power flows) to be used by adjacent areas [26]. This requires transferring to the central entity the topology of those lines internally connected to the boundary buses, yielding two or more estimates for shared buses. Then, the second level receives the boundary estimates from each area, the boundary measurements (including PMUs) and the necessary values of covariance matrices corresponding to locally estimated boundary variables. Boundary variables are re-estimated along with area slack phase angles.

Yan et al., in 2006 [27], developed a decomposition procedure based on the bordered-block diagonal form (BBDF) of the gain matrix. It is assumed that PMUs provide real-time boundary states, and hence power flows between subsystems, exactly. Then, using the data of PMUs as boundary conditions, the problem of distributed WLS decoupled state estimation turns into a multiarea optimization problem with equality constraints.

In the work of Jiang et al., 2007 [29], the overall system is decomposed into a certain number of non-overlapping subsystems based on a geographical basis. Each subsystem conducts its local SE with respect to its own slack bus, disregarding boundary power injections. Also each subsystem has a PMU installed at the slack bus. Measurements from PMUs coordinate the SE solution of each subsystem. A sensitivity analysis is performed for each bus to assess the degrees of impact from the neighboring subsystems. Those buses with a sensitivity exceeding a threshold are re-estimated at the central level along with boundary buses. The central SE incorporates the following information: all tie-line power flows, boundary injections, other internal measurements related to the sensitive internal buses and estimates of both boundary and sensitive internal buses, which are considered as pseudo measurements.

Jiang et al. [31] have recently (2008) presented a diakoptic-based SE algorithm. In the proposed approach, the SE problem is partitioned into a number of subproblems, obtained by removing tie-line measurements. Intermediate subsystem solutions are sent to a central computer for completing the state estimation process by taking tie-line measurements into consideration. The capability of PMUs to provide accurate synchronization of measurements is used to make each subproblem solvable and to coordinate the voltage angles of each subsystem SE solution. The optimal solution is reached.

7.4. Factorized SE

A different MASE perspective arises when each individual substation is regarded as an area. This allows a majority of raw measurements to be pre-processed by a linear estimator at the substation level [32]. The information provided by this stage, essentially composed of power flows and voltage magnitudes, is then integrated within the framework of a conventional SE. The main advantage of this hierarchical procedure, particularly in future substations with highly redundant measurement sets, is that the linear prefiltering phase can be run in a decoupled and geographically distributed manner, significantly reducing the bandwidth requirements. As a byproduct, a reduction in the computational effort is also achieved.

8. Decentralized MASE

In this category of methods the central coordinator is missing, neighboring areas directly exchanging border information with each other.

Each area proceeds in a similar fashion to the local solution phase of the hierarchical scheme. Internal measurements of area k, z_{ik}, can be readily used, as they are functions exclusively of state variables within the same area. On the contrary, border measurements involve state variables of neighboring areas j according to:

$$z_{bk} = h_{bk}(x_{ik}, x_{bk}, x_{bj}) + e(z_{bk}) \tag{6}$$

When iteratively solving the WLS problem in area k, the state vector components in (6) corresponding to adjacent areas, x_{bj}, must be replaced by the best available estimates, $\tilde{x}_{bj}$, submitted through the bilateral communication channel. This constitutes a relaxation-based distributed implementation of the global WLS problem. The information exchange can take place in a synchronous manner, provided it is feasible and convenient for all processors to run at the same pace, or asynchronously.

Typically, decentralized procedures take longer to converge than hierarchical ones, as they neglect some of the information handled by the central processor. Furthermore, bad data detection issues are more involved.

9. Literature survey on decentralized MASE

Table 2 provides a classification of fully distributed MASE methods, according to the different criteria identified in Section 5.

Due to space limitations only the most representative proposals are briefly reviewed.

In 1970, in his fundamental paper [35], Schweppe briefly described two techniques to reduce the computational burden of WLS estimation. By adding a fictitious bus at the mid-point of each tie-line, the system is supposed to be composed of border-bus overlapping areas. The first technique, termed "spatial quantization" consists of performing independent local estimations using within each region the corresponding local measurements. Major drawbacks of this very preliminary idea are the following: (i) no coordination between local estimates is performed, which induces loss of accuracy in tie-line power flow estimates (one mid-tie line appears as an antenna in each area); (ii) difficulty to identify bad data on tie-lines where local redundancy is poor.

In the second technique, denoted "spatial sweep", regions are estimated sequentially using, in addition to local measurements, the estimates of fictitious buses in tie-lines of neighboring areas already estimated. Relaxation with several iterations of this sweep procedure is needed, processing regions in different orders. This iterative formulation can reduce drastically the computational efficiency.

In the work by Brice and Cavin [13], reviewed in the previous section, the authors outlined also a relaxation-based approach, by which each processor solves asynchronously the non-linear WLS problem corresponding to its area of influence. The relaxation algorithm tries to obtain the state variables that minimize the objective function of its own area under the assumption that the state variables of adjacent areas have reached the optimal values, which is not actually the case. This slows down the overall convergence, compared with the iterative solution of the Normal equations applied to the entire network, but the computations are

Table 2
Classification of decentralized MASE methods.

Reference, author	Area overl. [a]	Solut. meth. [b]	Estim. state [c]	Meas. type [d]	Coord. scheme [e]
[35], Schweppe	MO	NE–R	Sub	C	SE–It
[13], Brice	NO	NE–R	Opt	C	It
[36], Lin	MO	O	Opt	C	It
[37,38], Lin	MO	NE–R	Opt	C	It
[39], Carvalho	MO	NE–H	Sub	C	It
[40], Huang	MO	NE–R	Opt	C	It
[41], Conejo	MO	NE–R	Sub	C	It

[a] NO: non-overlap; MO: minimally overlap; FO: fully overlap.
[b] NE: Normal equations; R: relaxation; O: optimization; H: heuristic.
[c] Opt: optimal; Sub: suboptimal.
[d] C: conventional only; P: considers PMU.
[e] SE: SE level; It: iteration level.

fully distributed. There is also a need for each processor to receive the current values of nearby state variables, significantly increasing the communication requirements. Given appropriate convexity assumptions, convergence of this asynchronous process is always guaranteed, no matter in which order and at which rate each subproblem is solved. Implementation of bad data analysis on a decentralized way is difficult.

The work by Lin, in 1990 and 1992 [36,37], combines Recursive Quadratic Programming with the Dual method. The distributed SE is aimed at its utilization in a decentralized control system, assuming a high speed data communication network, which is topologically the same and physically in parallel with the power network (a local processor is assigned to each bus). Global convergence of the distributed scheme to the optimal solution is shown.

In a subsequent work by Lin and Lin [38], recognizing the complexity and practical difficulties of this approach, in terms of both software and hardware, the authors proposed the following modifications: (a) each processor is in charge of an entire area, not a single bus, leading to a small-scale distributed architecture interconnected by a tree-shaped communication system; (b) a much simpler partially asynchronous block-Jacobi method is adopted.

Like in the previous work, Carvalho and Barbosa, 1998 [39], assume that the computer systems of adjacent areas are connected by fast data communication links, forming a computer network. That is, there is no centralized coordinating computer. The entire system is divided into border-bus overlapping areas. Fictitious, null-injection border buses are created if necessary so that boundary injections do not appear. When updating the state variables of a given area, the values of state variables corresponding to border buses are replaced by weighted averages of those computed in neighbor areas, where the weighting coefficients are obtained from the diagonals of the inverse gain matrices. Two algorithms are considered: (1) the SEs within each area are run in a synchronous manner, information being exchanged at the iteration level. This leads to inefficiencies because of processors being idle part of the time; (2) each area fully performs the SE in an asynchronous manner, correction terms being exchanged only at the end. Some gain in computation time is obtained at the cost of accuracy deterioration. The results are anyway acceptable if redundancy is high enough.

Conejo et al. [41] proposed in 2007 a multi-area decentralized SE procedure, based on optimization concepts. The resulting algorithm, closely resembling the relaxation-based approaches of Brice–Cavin [13] and Carvalho–Barbosa [39], proceeds as follows: (1) initialize state variables for all areas; (2) each area solves its corresponding problem, using available values of boundary state variables as pseudo-measurements; (3) if state variables do not change significantly within two iterations, stop; the solution has been reached. Otherwise, neighboring areas interchange their estimates for state variables corresponding to border buses, and the procedure continues in (2). No indication is provided on how to select appropriate weights for border pseudo-measurements.

10. Conclusion

Multi-area state estimation methods were introduced nearly forty years ago, partly to circumvent the limitations of by then available computers. The tremendous increase in computational power has not decreased the attractiveness of MASE. On the contrary, it has got renewed interest owing to the need of properly monitoring energy transactions across TSO borders in large interconnections, while at the same time processing the real-time data at the most appropriate place, and possibly preserving their confidentiality. The expected increase of smart grid applications also calls for the MASE techniques, in order to prevent communication infrastructures from being unnecessarily burdened with the resulting information explosion.

Over thirty references on MASE, including both journals and conference proceedings, are analyzed in this work. As a consequence, a taxonomy of MASE is established with the help of previously identified classification criteria. Such criteria include: area overlapping degree, computer architecture, coordination scheme, process and measurement synchronization and solution methodology. For each major category of methods (hierarchical and decentralized) an effort has been made to identify the relevant state and measurement vector components, as well as their interactions during the solution phases, all this under a common notation. Then, a brief presentation is made of a relatively large selection of references, trying to point out their distinguishing features and limitations.

A somewhat expected but noteworthy conclusion of this work is that there exists a compromise solution between optimality and computational cost (or complexity) of the resulting procedure. A majority of hierarchical procedures lead to suboptimal, yet accurate enough solutions in just two steps, particularly in the presence of PMUs, which increase the linearity of the resulting models. Fully distributed schemes tend to be simpler, at the cost of poorer convergence to the optimal solution.

Acknowledgements

This work was performed in the context of the PEGASE project funded by European Community's 7th Framework Programme (grant agreement No. 211407). The Spanish authors also acknowledge the support of the Spanish DGI under grants ENE 2007-62997 and ENE 2010-18867.

References

[1] A.R. Khatib, L. Mili, A. Phadke, J. De La Ree, Y. Liu, Internet-based wide-area information sharing and its roles in power system state estimation, in: IEEE-PES Winter Power Meeting, Columbus, Ohio, January 28–February 1, 2001.
[2] www.coreso.eu.
[3] T. Van Cutsem, M. Ribbens-Pavella, Critical Survey of hierarchical methods for state estimation of electric power systems, IEEE Trans. Power Apparatus Syst. PAS-102 (October (10)) (1983) 3415–3424.

[4] M. Shahidehpour, Y. Wang, Communication and Control in Electric Power Systems: Applications of Parallel and Distributed Processing, Wiley–IEEE Press, July 2003.
[5] A. Abur, A. Gómez Expósito, Power System State Estimation: Theory and Implementation, Marcel Dekker, April 2004.
[6] A. Monticelli, A. Garcia, Reliable bad data processing for real-time state estimation, IEEE Trans. Power Apparatus Syst. PAS-102 (May (5)) (1983) 1126–1139.
[7] K.A. Clements, O.J. Denison, R.J. Ringlee, A multi-area approach to state estimation in power system networks, in: IEEE PES Summer Meeting, Paper C72 465-3, San Francisco, CA, July 1972.
[8] H. Kobayashi, S. Narita, M.S.A.A. Hamman, Model coordination method applied to power system control and estimation problems, in: Proc. of the IFAC/IFIP 4th Int. Conf. on Digital Computer Appl. to Process Control, 1974, pp. 114–128.
[9] M.R. Irving, M.J.M. Sterling, Multi-level state estimation and optimal control, in: Proc. of the IFAC Symp. on Comp. Applic. in Large-Scale Power Syst, New Delhi, August 1979, pp. 117–125.
[10] J.F. Marsh, R. Cristi, State estimation on electric power systems using partitioned network models, in: Proc. of the IFAC Symp. on Comp. Applic. in Large-Scale Power Syst, New Delhi, August 1979, pp. 97–103.
[11] T. Van Cutsem, J.-L. Horward, M. Ribbens-Pavella, A two-level static state estimator for electric power systems, IEEE Trans. Power Apparatus Syst. PAS-100 (August (8)) (1981) 3722–3732.
[12] Y. Wallach, E. Handschin, C. Bongers, An efficient parallel processing method for power system state estimation, IEEE Trans. Power Apparatus Syst. PAS-100 (November (11)) (1981) 4402–4406.
[13] C.W. Brice, R.K. Cavin, Multiprocessor static state estimation, IEEE Trans. Power Apparatus Syst. PAS-101 (February (2)) (1982) 302–308.
[14] K. Seidu, H. Mukai, Parallel multi-area state estimation, IEEE Trans. Power Apparatus Syst. 104 (May (5)) (1985) 1025–1034.
[15] M.S. Kurzin, Real-time state estimation for large-scale power systems, IEEE Trans. Power Apparatus Syst. PAS-102 (July (7)) (1983) 255–263.
[16] K.L. Lo, M.M. Salem, R.D. McColl, A.M. Moffatt, Two-level state estimation for large power system. I. Algorithms and II. Computational experience, IEE Proc.-Gener. Transm. Distrib. Part C 135 (July (4)) (1988) 299–318.
[17] S. Iwamoto, M. Kusano, V.H. Quintana, Hierarchical state estimation using a fast rectangular-coordinate method, IEEE Trans. Power Syst. 4 (August (3)) (1989) 870–879.
[18] A.A. El-Keib, C.C. Carroll, H. Singh, D.J. Maratukulam, Parallel state estimation in power systems, in: Proc. 22nd Southeastern Symposium on System Theory, CS Press, Los Alamitos, CA, Catalog No. 2038, March 1990, pp. 255–260.
[19] S.S. Ahmed, A. Brameller, New algorithm for diakoptical static state estimation, IEE Proc. Gener. Transm. Distrib., IEE Proc. C 138 (May (3)) (1991) 185–192.
[20] A.A. El-Keib, J. Nieplocha, H. Singh, D.J. Maratukulam, A decomposed state estimation technique suitable for parallel processor implementation, IEEE Trans. Power Syst. 7 (August (3)) (1992) 1088–1097.
[21] D.M. Falcao, F.F. Wu, L. Murphy, Parallel and distributed state estimation, IEEE Trans. Power Syst. 10 (May (2)) (1995) 724–730.
[22] G.N. Korres, G.C. Contaxis, Application of a reduced model to a distributed state estimator, IEEE Power Eng. Soc. Winter Meet. 2 (2000) 999–1004.
[23] R. Ebrahimian, R. Baldick, State estimation distributed processing, IEEE Trans. Power Syst. 15 (November (4)) (2000) 1240–1246.
[24] J.A. Aguado, C. Perez-Molina, V.H. Quintana, Decentralised power system state estimation: a decomposition–coordination approach, IEEE Porto Power Tech. Proc. 3 (September) (2001) 6.
[25] L. Zhao, A. Abur, Multiarea state estimation using synchronized phasor measurements, IEEE Trans. Power Syst. 20 (May (2)) (2005) 611–617.
[26] A. Abur, Distributed state estimation for mega grids, in: Proc. of the 15th PSCC Liege, August 2006, pp. 22–26.
[27] L. Yan, Z. Xiaoxin, Z. Jingyang, A new algorithm for distributed power system state estimation based on PMUs, in: International Conference on Power System Technology, PowerCon 2006, October 2006.
[28] R.F. Jeffers, Techniques for wide-area state estimation in power systems, Master Thesis, Virginia Polytechnic Institute, November 2006.
[29] W. Jiang, V. Vittal, G.T. Heydt, A distributed state estimator utilizing synchronized phasor measurements, IEEE Trans. Power Syst. 22 (May (2)) (2007) 563–571.
[30] M.Y. Patel, A.A. Girgis, Two-level state estimation for multi-area power system, in: IEEE Power Engineering Society General Meeting, June 2007.
[31] W. Jiang, V. Vittal, G.T. Heydt, Diakoptic state estimation using phasor measurement units, IEEE Trans. Power Syst. 23 (November (4)) (2008) 1580–1589.
[32] A. Gómez Expósito, A. de la Villa Jaén, Two-level state estimation with local measurement pre-processing, IEEE Trans. Power Syst. 24 (May (2)) (2009) 676–684.
[33] G.N. Korres, A distributed multiarea state estimation, IEEE Trans. Power Syst., 10.1109/TPWRS.2010.2047030.
[34] M.D. Mezarovic, D. Macko, Y. Takahara, Theory of Hierarchical Multilevel Systems, Academic Press, 1970.
[35] F.C. Schweppe, J. Wildes, D.B. Rom, Power system static state estimation, Part I, II, III, IEEE Trans. Power Apparatus Syst. PAS-89 (January (1)) (1970) 120–135.
[36] S.Y. Lin, A method to solve power system static-state estimation problem by using a network of processors, in: Proc. 29th CDC, Hawaii, Honolulu, December 5–7, 1990, pp. 3067–3072.
[37] S.Y. Lin, A distributed state estimator for electric power systems, IEEE Trans. Power Syst. 7 (May (2)) (1992) 551–557.
[38] S.-Y. Lin, C.-H. Lin, An implementable distributed state estimator and distributed bad data processing schemes for electric power systems, IEEE Trans. Power Syst. 9 (August (3)) (1994) 1277–1284.
[39] J.B. Carvalho, F.M. Barbosa, Parallel and distributed processing in state estimation of power system energy, in: Electrotechnical Conference, 1998 MELECON 98, 9th Mediterranean, vol. 2, May 18–20, 1998, pp. 969–973.
[40] G.M. Huang, J. Lei, A concurrent non-recursive textured algorithm for distributed multi-utility state estimation, in: Power Engineering Society Summer Meeting, 2002 IEEE, vol. 3, July 2002, pp. 1570–1575.
[41] A.J. Conejo, S. de la Torre, M. Canas, An optimization approach to multiarea state estimation, IEEE Trans. Power Syst. 22 (February (1)) (2007) 213–221.

CONTRIBUTED PAPER

A Multilevel State Estimation Paradigm for Smart Grids

The authors of this paper describe a multilevel state estimator architecture that can sustain growth in size, complexity of data, and information.

By Antonio Gómez-Expósito, *Fellow IEEE*, Ali Abur, *Fellow IEEE*, Antonio de la Villa Jaén, and Catalina Gómez-Quiles, *Student Member IEEE*

ABSTRACT | The main objective of this paper is to describe a multilevel framework that facilitates seamless integration of existing state estimators (SEs) that are designed to function at different levels of modeling hierarchy in order to accomplish very large-scale monitoring of interconnected power systems. This has been a major challenge for decades as power systems grew pretty much independently in different areas, which had to operate in an interconnected and synchronized fashion. The paper initially provides a brief historical perspective which also explains the existing state estimation paradigm. This is followed by a review of the recent technological and regulatory drivers that are responsible for the new developments in the energy management functions. The paper then shows that a common theoretical framework can be used to implement a hierarchical scheme by which even very large-scale power systems can be efficiently and accurately monitored. This is illustrated for substation level, transmission system level as well as for a level between different transmission system operators in a given power system. Finally, the paper describes the use and benefits of phasor measurements when incorporated at these different levels of the proposed infrastructure. Numerical examples are included to illustrate performance of the proposed multilevel schemes.

KEYWORDS | Distributed state estimation; hierarchical state estimation; phasor measurement units (PMUs); smart grids; substation estimation

I. INTRODUCTION

As the penetration of renewable and distributed energy sources along with the necessary means of centralized and distributed energy storage technologies increase to higher and higher levels, the existing electric energy network infrastructures are expected to evolve in two major directions. On the one hand, much longer and higher rated transmission lines carrying both alternating current (ac) and direct current (dc) power at the highest voltage levels will be needed to move huge amounts of renewable energy across very long distances. On the other hand, at the lowest voltage levels, the existing highly centralized power systems will be transformed into the so-called "smart grids" where the intelligence will be to a large extent distributed, through the use of distribution automation, power electronics, active load management, real-time metering, and other technical innovations in telecommunications and computer technologies [1].

State estimators (SEs) determine the most likely state of a power system from sets of remotely captured measurements that are collected periodically by SCADA systems via remote terminal units (RTUs). The role of the SE is crucial in modern energy management systems (EMSs), where a diversity of applications rely on accurate system snapshots [2]. The regulatory wave of the last decade has stressed the importance of the SE tool, in an open-access context in which many more transactions on much more congested networks have to be properly tracked in real time and also recorded for offline engineering studies. This paper serves two purposes. One is to provide a survey of SEs that have been developed and implemented in the past. The other is to present, in a tutorial manner, a novel

Manuscript received August 12, 2010; revised October 19, 2010; accepted January 12, 2011. Date of publication April 29, 2011; date of current version May 17, 2011. The work of A. Gómez-Expósito, A. de la Villa Jaén, and C. Gómez-Quiles was supported by the European Community's 7th Framework Programme (PEGASE Project) under Grant 211407 and the Ministry of Science & Innovation (DGI) under Grants ENE2007-62997 and ENE2010-18867. The work of A. Abur was supported by PSERC/NSF project S-22. The first author received the "VIII Javier Benjumea Puigcerver Research Award, granted by Fundación Focus-Abengoa, for this work.

A. Gómez-Expósito, **A. de la Villa Jaén**, and **C. Gómez-Quiles** are with the Department of Electrical Engineering, University of Seville, Seville 41092, Spain (e-mail: age@us.es; adelavilla@us.es; catalinagq@us.es).

A. Abur is with the Electrical and Computer Engineering Department, Northeastern University, Boston, MA 02115 USA (e-mail: abur@ece.neu.edu).

Digital Object Identifier: 10.1109/JPROC.2011.2107490

multilevel scheme that generalizes a number of previously developed ideas in one comprehensive proposal.

II. A BRIEF HISTORICAL PERSPECTIVE

Since its introduction by Schweppe in the late 1960s [3], the SE tool has benefited from a large number of theoretical developments and practical improvements, which were well documented in [4]–[6].

In addition to the pioneering work of Tinney and others on sparsity methods [7], there have been computational improvements such as the so-called fast decoupled state estimator (FDSE), proposed by Monticelli *et al.* [8], based on the successful fast decoupled load flow concept [9]. Another significant improvement was introduced in order to address the issues of numerical stability and ill-conditioning of the conventional weighted least squares (WLS) approach [10]. It was observed that the use of artificially high weights for very accurate measurements such as zero injections and rather low weights for much less accurate pseudomeasurements would lead to poor convergence of the so-called normal equations. It was shown that employing a computationally more expensive method of orthogonal or *QR* factorization would significantly improve the solution accuracy compared to the less stable Cholesky factorization scheme [11]. Alternatively, it was shown that these very accurate or exact measurements, such as zero injections, could be incorporated as explicit equality constraints into the optimization problem formulation, effectively eliminating the need to use artificially high weights [12]. One such formulation used the so-called Hachtel's augmented matrix approach, which could be implemented either directly [13] or using 2×2 pivots in combination with block arithmetic for improved computational efficiency [14], [15].

Network observability and bad data processing constitute two important functions related to the SE problem. Observability analysis is performed in advance in order to determine if the entire state vector is observable and, if not, to identify observable islands. Both numerical [16] and topological [17] algorithms have been proposed and implemented [18]. Strongly related to the network observability analysis was the optimal measurement design that would ensure full network observability under credible loss of RTUs or communication channels [19]. In many instances, the measurement set might be corrupted with gross errors (outliers) and thus, the assumption that all measurement errors were normally distributed would no longer be true. In such cases, if those gross errors (bad data) were not detected and removed by simple plausibility tests before the execution of the SE, the solution would be biased or even the algorithm could fail to converge. Hence, statistical tests such as the chi-squares test and largest normalized residual test, based on chi-squares and standard normal distributions, respectively, were developed in order to detect and identify bad data. Both tests relied on calculated measurement residuals once the SE algorithm converged [20]. More elaborate techniques, such as hypothesis testing identification (HTI), have also been proposed to handle cases involving multiple interacting bad data for which other methods were less effective [21].

There were numerous other developments that addressed a wide spectrum of issues ranging from statistical robustness of estimation [22], [23], hierarchical multiarea estimation [24], inclusion of Ampere measurements at the subtransmission level [25]–[27], incorporation of inequality constraints [28],[29], and detection and identification of network parameter and topology errors [30], [31].

In the last decade there has been an increased interest in the so-called generalized state estimator (GSE), aimed at developing circuit breaker (CB) models to improve the SE capability for topology error processing [5]. This involved a detailed physical level modeling of bus sections that should be used in combination with zooming techniques in order to cope with the huge size of the resulting model [32], [33]. An implicit GSE model has been recently developed. The model maintained the capability to identify topology errors, while using a slightly augmented state vector [34]. The same idea could be used to detect and identify network parameter errors [35].

More recently, the so-called phasor measurement units (PMUs), which provide global positioning system (GPS)-synchronized measurements, among which are voltage and current phasor magnitude and phase angles, are expected to introduce major improvements in SE performance and capabilities [36], [37].

III. EXISTING STATE ESTIMATION PARADIGM

This section briefly reviews the geographical scope, interactions, and solution methodology of existing SE tools.

A. Existing Configuration of SEs

Currently, the scope of SEs is mostly limited to the transmission level, where each transmission system operator (TSO) continuously tracks its own grid from a centralized EMS (see Fig. 1). However, in some power systems, trading and pricing may also take place at lower voltage levels, which are typically not closely monitored by a SE. This is mainly due to the unavailability of sufficiently redundant set of measurements at these voltage levels. Hence, there is a need to improve the monitoring capability for these parts of the systems and to facilitate reliable and effective operation of power markets at these levels as well.

At the uppermost voltage levels, where interconnections for energy trading exist, tie-line power flows have to be properly monitored, for which TSOs must get access to both the electrical model and real-time measurements of its neighbors, at least those in the adjacent buses. For this

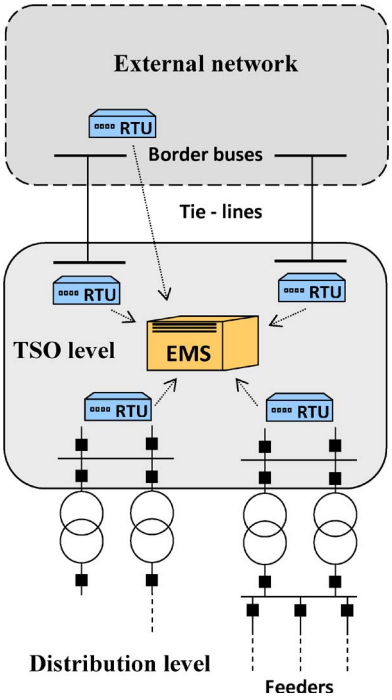


Fig. 1. *Conventional TSO-level state estimation paradigm. (RTU: remote terminal unit; EMS: energy management system; TSO: transmission system operator.)*

purpose, the external grid is usually represented by a reduced equivalent circuit [38]. Therefore, existing SEs, being tailored to the needs of a single TSO, with very neat borders, are not designed to significantly interact with its neighbors or subordinate networks. As explained later, this is not a satisfactory state of affairs, in view of both the needs and possibilities offered by smart grids, but commercial software evolves at a lower pace than theoretical developments.

B. Standard WLS-Based Solution

The SE relies on the following measurement equation [3]:

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

where x is the state vector to be estimated (size n), z is the known measurement vector (size $m > n$), h is the vector of functions, usually nonlinear, relating error free measurements to the state variables, and e is the vector of measurement errors, customarily assumed to have a normal distribution with zero mean and known covariance matrix R. When errors are independent, R is a diagonal matrix with values σ_i^2, where σ_i is the standard deviation of the error associated with measurement i.

In conventional bus-branch SE models the state vector is composed of voltage magnitudes and phase angles, whereas the measurement vector typically comprises power injections, branch power flows, and voltage magnitudes. Recently, the availability of synchro-phasors (PMUs) has made it possible the incorporation of phase angle measurements into the SE process.

The WLS estimator minimizes the weighted squares of residuals of the measurements given by

$$J = \sum_{i=1}^{m} W_i r_i^2$$

where $r_i = z_i - h_i(\hat{x})$ is the measurement residual, $\hat{x}$ is the estimated state vector, and W_i is the respective weighting coefficient.

The state estimate can be obtained by iteratively solving the normal equations

$$G_k \Delta x_k = H_k^T W [z - h(x_k)] \tag{2}$$

where $H_k = \partial h / \partial x$ is the Jacobian evaluated at $x = x_k$, $G_k = H_k^T W H_k$ is the gain matrix, $W = R^{-1} = \mathrm{diag}(W_i)$ is the weighting matrix, and $\Delta x_k = x_{k+1} - x_k$, k being the iteration counter.

Iterations are terminated when an appropriate tolerance is reached on Δx_k. The covariance of the estimate is

$$\mathrm{cov}(\hat{x}) = G_k^{-1}.$$

Then, the bad data processing function, aimed at detecting, identifying, and eliminating bad analog measurements, is activated. This is accomplished through the largest normalized residual test [20].

IV. TECHNOLOGICAL AND REGULATORY DRIVERS

The SE paradigm described above will have to drastically change with the advent of the smart grid. On the one hand, new generations of digital devices, such as PMUs or intelligent electronic devices (IEDs), intended for measurement, protection, and control, being less expensive and more flexible than existing analog equipment, will invade virtually every corner of future networks [37], [39]. This will provide a more accurate, complex, and highly redundant information system, allowing SEs to extend their scope well beyond presently observable areas and also to incorporate advanced functions that have not yet reached the industrial stage, in spite of being conceptually mature on the researcher blackboard. On the other hand, smart

transmission grids should further promote the development of regional energy markets, involving distant energy transactions. This also implies wide-area physical interactions, possibly of catastrophic consequences in case of cascaded failures [40], [41].

In the following sections, outstanding technological innovations associated with the smart grid concept will be succinctly analyzed from the point of view of their influence in the conceptual design of future SE architectures.

A. Intelligent Electronic Devices

The protection, metering, and control functions in substations are naturally distributed by the role and location of each device, being designed in general to provide primary protection or monitoring of an individual substation equipment. These functions may be performed by smart multifunctional and communicative units, so-called IEDs. They are broadly defined in [42] as "devices incorporating one or more processors with the capability to receive or send data/control signals from or to an external source (e.g., electronic multifunction meters, digital relays, controllers)."

The IEDs, employing efficient signal processing techniques, are becoming the source of much more information in real time than the one existing in old substations. Apart from implementing specific protection or control algorithms, they can provide externally electrical magnitudes measured by protection transformers as well as phase differences among them [39]. Those measurements can be synchronized, both at the substation and wide-area levels, by means of the GPS satellite clock time reference.

The quality of the SE process strongly relies on the redundancy of the measurement set. For this reason, the possibility of incorporating all the information provided by IEDs, including the ones whose primary function is not measurement but protection, is very attractive.

B. Communication and Architecture Standards: IEC 61850

New system architectures need to be devised for preprocessing the ever-increasing amount of information gathered by the IEDs. This goal is achieved by replacing the conventional centralized systems, based on RTUs and numerous protection and control devices, with local-area-network-based systems and advanced multifunctional protection and control IEDs [43].

IEC 61850, the global communication standard for substation automation system, defines the communication between IEDs and not only solves the interoperation problem but also specifies other system requirements, like message performance and information security in substation automation system network [43]. IEC 61850 allows

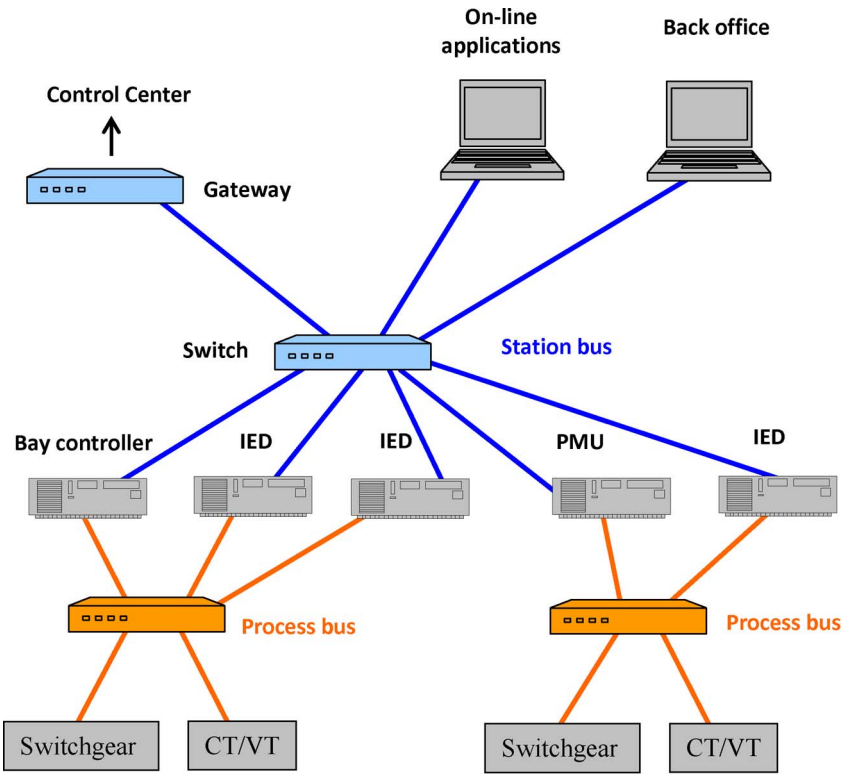


Fig. 2. *Substation-level hardware platform, according to IEC 61850. (IED: intelligent electronic device; PMU: phasor measurement unit; CT/VT: voltage/current transformer.)*

interoperability of IEDs from different manufacturers without the use of protocol converters.

The standard defines two communication buses between the different subsystems within a substation (see Fig. 2). The process bus is devoted to gathering information about electrical magnitudes, such as voltage or current, as well as switching status information, from the transformers and transducers connected to the primary power system process. The station bus is aimed at allowing primary communications between the logical nodes, which provide the various station protection, control, monitoring, and logging functions. The communication technologies involved in these buses include: Ethernet on fiber optic, TCP/IP, and MMS (ISO9506). This architecture supports remote network access for all types of data reads and writes.

C. Phasor Measurement Units

Most of the EMS applications assume that the system is in a pseudosteady state where ac circuit analysis can be carried out using phasors. Network model and the voltage phasors at all system buses are used to determine the state of the system. While the bus voltage phasors can be estimated based on a redundant set of measurements, direct measurement of phasors will be possible only if measurements are synchronized. PMUs are devices that take advantage of the GPS satellites in order to time synchronize the measurements. Voltage and current signals are collected at secondaries of instrument transformers (CT and PT) and are sampled via analog-to-digital converters at 48 samples/cycle. These samples are then processed and synchronized with universal time coordinated (UTC) time from a GPS receiver within 1-μs accuracy [36]. Time-synchronized samples are processed to obtain time-stamped voltage and current phasors, which are then transmitted over Ethernet to phasor data concentrators (PDCs), which will then send data to control center SCADA server. IEEE Standard C37.118 describes the requirements, format, and communication protocol for data provided by PMUs [44].

PMUs may have several channels, each of which will record one phase of a voltage or current signal. Two sets of three channels are typically used for three phase voltage measurements at the substation and several sets of three channels will be used for measurement of three phase currents along incident lines or transformers. Positive sequence components rather than individual phase signals are typically used by network applications, hence three phase signals are processed to compute the positive sequence components. The string of positive sequence phasors that are computed by the PMU will then be communicated at 30 samples/s to the PDC.

D. Distribution Automation

Distribution automation is a mature concept whose real potential never took off by a lack of reasonably priced infrastructures. In fact, unlike in transmission systems, most functions in this field (fault detection, service restoration, network reconfiguration, etc.) have been traditionally performed with the help of mobile service teams on call.

Regarding the information that can be found at the distribution level, virtually no measuring devices have been installed until recently to monitor the operating condition of medium voltage (MV) feeders. Typically, the current (or sometimes the power flow) at each feeder head, along with the voltage magnitude at the MV bus, are telemetered and gathered at dedicated distribution management systems (DMSs). But no real-time information is obtained of what happens downstream, unless a fault occurs. The substation bus voltage is kept almost constant by the use of automatic under-load tap changers, in the hope that customer voltages remain acceptable for nearly all operating points.

This situation is rapidly changing for several reasons.

- Distributed generators (DGs) connected at this level frequently reverse the sign of power flows, creating overvoltage problems that should be properly monitored and prevented. On the other hand, the energy they inject, frequently at premium prices, should also be carefully monitored and recorded.
- Smart meters, currently being deployed, provide hourly customer demands via PLC, regular cellular phone technology, or alternative means. There is a trend to concentrate all this information at the secondary transformer centers, from where it is then submitted upstream to the distribution substation or DMS.
- Cheap fault current detectors, along with automatic or remotely operated reclosers, are being installed at strategically selected points to speed up the service restoration process. These devices are capable of providing less accurate current values that could also be attached to the remaining information sent to the substation.

Consequently, the massive introduction of DG and a plethora of distribution automation devices, at network levels that are not currently supervised by TSOs, will also contribute to the development of ubiquitous monitoring systems [45].

E. Wide-Area Regional Energy Markets

Energy markets outcomes may be significantly affected by unique information regarding the present and likely states of the grid. But gathering such information is a real challenge when the energy transactions take place over networks that cross national or regional market borders. For this and other strategic reasons, international regulatory entities are promoting worldwide the creation of regional-level system operators, in an attempt to eliminate barriers and better coordinate multi-TSO transactions.

In the United States, for instance, realizing that competition was hindered because only a handful of utilities owned and controlled a large portion of the region's transmission, the Federal Energy Regulatory Commission (FERC) issued the Order 890 "Preventing Undue Discrimination and Preference in Transmission Service" and the 2010 Notice of Proposed Rulemaking on "Transmission Planning and Cost Allocation by Transmission Owning and Operating Public Utilities." These orders amended its regulations in order to remedy opportunities for this undue discrimination and address deficiencies in the proforma open-access transmission tariff. Enabling interconnection-wide operation via these rules that extend the local, regional, and inter-regional planning processes will be greatly facilitated by the proposed multilevel state estimation scheme.

Across the Atlantic, the European Regulators' Group for electricity and gas (ERGEG) launched in 2007 an initiative to create a series of Regional Energy Market projects within the European Union (EU), in order to remove barriers to cross-border trade between countries as a first step towards the completion of a single EU market for electricity [46]. Other long-term ambitious projects for long-distance bulk transmission of renewable energy, such as the Desertec initiative [47], will further stress the need for transnational cooperation in network monitoring.

V. MULTILEVEL STATE ESTIMATION ARCHITECTURE

The technological developments discussed above, along with more aggressive regulatory schemes intended to promote efficient trading of clean energy, allow to envision a future in which SEs will spread from MV distribution feeders to the bulk extra high voltage (EHV) transmission network, spanning several interconnected areas [45], [48].

But the crucial question is how the existing SE paradigm will have to be adapted in order to cope with such a diverse and extensive geographical scope as well as the formidable amount of information provided by the heterogeneous and distributed sources arising in the upcoming smart grid environment. When trying to answer that question, a dilemma arises about whether it will be feasible and convenient to keep on submitting all this information to a central EMS or it should be processed to a large extend in a local manner, as close as possible to the place in which it is generated. The first choice is discouraging for two main reasons:

1) the investment in new communication infrastructures would be prohibitive;
2) the required computing power for the central entity to be capable of processing the incoming data in real time should be one or two orders of magnitude larger than that of existing EMS.

The alternative and most natural choice to deal with the explosion of information arising in this multiagent

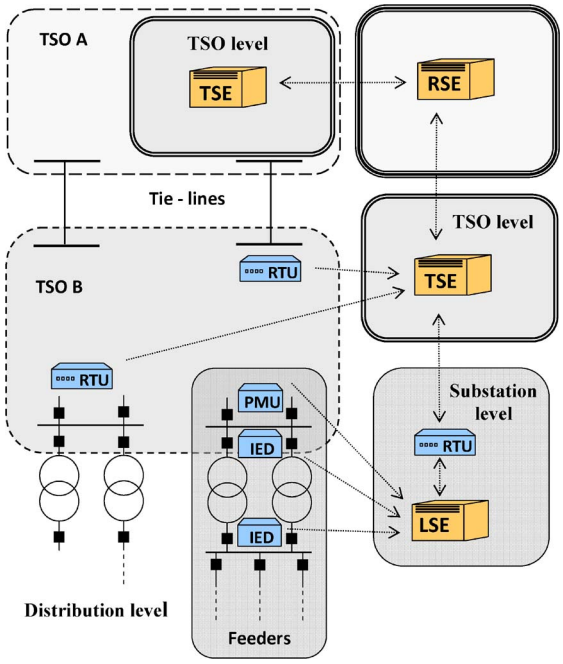


Fig. 3. *Smart-grid-oriented multilevel state estimation paradigm. (LSE: local SE; TSE: transmission-level SE; RSE: regional multi-TSO SE.)*

distributed environment calls necessarily for a multilevel, hierarchical SE paradigm. In this paper, at least three major levels are identified, as shown schematically in Fig. 3 (a fourth level associated with distribution feeders will also be distinguished).

At the lowest level, a local SE (LSE) can be implemented to preliminarily deal with the information collected within a substation or small set of adjacent substations. A great majority of raw measurements will be processed at this distributed level, where a modest but sufficient computing power already exists. Distribution substations, delivering power to a large number of secondary transformers through a set of radial feeders, constitute a particular relevant case. In those substations, it is advantageous and makes sense to process each radial feeder in a decoupled manner, leading to a fourth level of information processing.

The results provided by the LSE have to be transmitted through existing RTUs and communication channels to the TSO-level SE (TSE). At this intermediate level, commercially available software can be adopted with minor modifications, the major difference with respect to a conventional SE being that prefiltered rather than raw measurements are handled.

At the uppermost level, a regional SE (RSE) will be needed to synchronize and refine the results separately provided by each TSO affiliated with the interconnected system, particularly near the border nodes. The RSE will be a customized tool, designed in such a way that the amount of information exchanged with subordinate TSEs

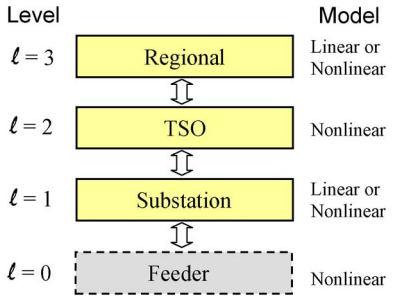


Fig. 4. *Hierarchical multilevel architecture.*

is kept to a minimum. This SE level will significantly benefit from wide-area measurements provided by PMUs.

Fig. 4 shows the resulting hierarchy, including the feeder level arising in distribution substations. The double arrow represents schematically the bidirectional flow of information between adjacent levels.

The interactions between adjacent levels can be better formulated and justified as particular customized cases of a common theoretical framework that will be presented in the next section. Then, a more detailed treatment of each level will be separately made in the remaining sections.

VI. THEORETICAL FRAMEWORK

The multilevel SE formulation advocated in this paper requires that the standard WLS solution approach be reconsidered. This stems from the following observations.

- The conventional TSE is based on the so-called bus-branch electrical model, which is not of direct application to the substation level. In this environment, the LSE must be capable of dealing with extra raw measurements and detailed topology information, not found ordinarily at the EMS level, for which augmented state vectors must be considered.
- The estimate of the multilevel SE should be optimal, i.e., as close as possible to that provided by an ideal solver simultaneously handling all raw measurements for the entire set of interconnected areas. Theoretically, this is possible, but only if the required statistical information, associated with each partial estimate, is duly exchanged between adjacent levels, usually in an iterative manner. Existing TSEs should be adapted to interact in this way with their neighbors.

The conventional WLS-based SE methodology, currently used by TSOs worldwide, was succinctly reviewed in Section III-B. In what follows, a recently introduced, two-stage WLS solution method, based on a factorization of the measurement model, will be presented [49]. Such a scheme, when combined with suitable geographical decomposition techniques and certain model transformations, determines the algorithmic steps and interactions involved in the multilevel SE paradigm, and provides the right mathematical framework supporting the overall estimate optimality.

First, the two-step approach introduced in [49] will be generalized to the fully nonlinear case. Then, two relevant particular cases, in which one of the resulting submodels is linear, will be considered. The final subsection provides important implementation guidelines allowing the factorized solution to be geographically distributed whenever possible, which is the main goal of the proposed multilevel architecture.

A. Factorized WLS Solution: General Nonlinear Case

The factorized approach to solve the WLS problem arises when the nonlinear measurement model (1) is "unfolded" into two sequential WLS problems, as follows:

$$z = f_1(y) + e \tag{3}$$
$$y = f_2(x) + e_y \tag{4}$$

where y is a vector of intermediate variables, selected in such a way that the solution of the pair (3)–(4) offers an advantage over that of (1). For the resulting factorized model to be equivalent to the original one, the following condition must be satisfied:

$$h(x) = f_1[f_2(x)] \quad \Rightarrow \quad H(x) = F_1(y)F_2(x) \tag{5}$$

where H, F_1, and F_2 represent the Jacobian matrices of h, f_1, and f_2, respectively.

As shown in the Appendix, the optimal estimate $\hat{x}$ provided by the conventional iterative process (2) can be alternatively obtained by successively solving the following pair of equations:

$$\left[F_1^T W F_1\right]\Delta y_k = F_1^T W[z - f_1(y_k)] \tag{6}$$
$$\left[F_2^T G_1 F_2\right]\Delta x_k = F_2^T G_1[\tilde{y} - f_2(x_k)] \tag{7}$$

where $\tilde{y}$ in (7) is the estimate of the intermediate vector provided by (6), and the weighting matrix G_1 satisfies

$$G_1 = [\mathrm{cov}(y)]^{-1} = F_1^T W F_1. \tag{8}$$

In other words, the weighting matrix of the second WLS problem is the gain matrix of the first one.

As explained in the Appendix, full equivalence between the original and the factorized models requires that the linearization of f_1 and f_2 be performed at the same point in

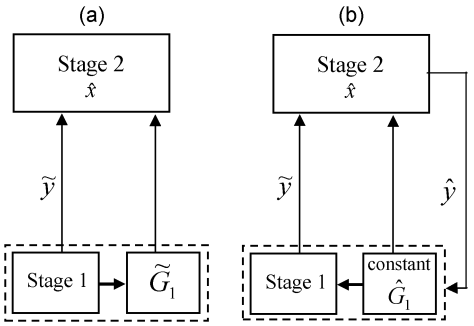


Fig. 5. *Schematic flowchart of the two-stage factorized procedure.*

the n-dimensional space represented by x. This leads in the general case to an outer iterative process in which two stages, one for each WLS subproblem, are successively solved.

Accordingly, the factorized procedure can be formally decomposed into the following sequence of steps, where the first run is separately considered for clarity of presentation.

First run:

- *Stage 1*: Find $\tilde{y}$ by repeatedly solving (6) until convergence. As a byproduct, $\tilde{G}_1$ is available.
- *Stage 2*: Find $\hat{x}$ by repeatedly solving (7) with $G_1 = \tilde{G}_1$.

Subsequent runs (if needed):

- *Stage 1*: Update the Jacobian, $\hat{F}_1$, and the gain matrix, $\hat{G}_1$, for $\hat{y} = f_2(\hat{x})$. Keeping these matrices constant, find $\tilde{y}$ by repeatedly solving (6).
- *Stage 2*: Find $\hat{x}$ by repeatedly solving (7) with $G_1 = \hat{G}_1$.

The process is repeated until two consecutive runs of Stage 2 provide close enough values of $\hat{x}$.

Fig. 5 represents schematically the two-stage procedure outlined above. The auxiliary vector y plays the role of a state vector when solving Stage 1 and that of a measurement vector when solving Stage 2. Also, as stated above, the gain matrix of Stage 1 becomes the weighting matrix of Stage 2. Note that this matrix is no longer diagonal, but its sparsity can and should be fully exploited. The fact that subsequent runs of Stage 1 involve only constant matrices can also be exploited.

In practice, unless the measurement vector is very noisy and/or contains key bad data, the first run of stages 1 and 2 will provide sufficiently accurate results.

For the sake of clarity, the description provided in this section assumes that all raw measurements z can be used in Stage 1. If this is not the case, then Stage 2 can be easily redesigned to handle simultaneously the pseudomeasurement $\tilde{y}$ and the components of z not yet used during Stage 1. This is explained in detail in [49].

The fully nonlinear factorized model presented above may reduce to any of the two particular cases discussed below.

B. Factorized WLS Solution: Linear–Nonlinear Case

This case, by far the most interesting in practice, arises when the intermediate vectory can be chosen in such a way that the measurement model of Stage 1 becomes linear [49]. Then, the nonlinear systems (3) and (4) reduce to

$$z = Ay + e \tag{9}$$

$$y = f_2(x) + e_y. \tag{10}$$

The general nonlinear model can be easily particularized to this case by systematically replacing $f_1(y)$ and F_1 in the expressions above by Ay and A, respectively. The most relevant implication is that the iterative system (6) reduces in this case to the following linear one:

$$[A^T W A]\tilde{y} = A^T W z. \tag{11}$$

Accordingly, the gain matrix

$$G_1 = A^T W A \tag{12}$$

remains constant, so long as the network topology and measurement set structure are unaltered. Therefore, changes in the measurement values and state variables originated by the daily load evolution do not alter the linear model of Stage 1, which is one of the sources of computational saving associated with the factorized approach.

The factorized procedure reduces in this case just to the first run of Stages 1 and 2. Stage 1 constitutes a linear prefilter of the raw measurement vector, which can be a valuable tool by itself (for instance, to perform preliminary bad data analysis when the redundancy if sufficiently high). Accuracy of the linear estimate delivered by Stage 1 could be subsequently checked by comparing $\tilde{y}$ with $\hat{y} = f_2(\hat{x})$.

C. Factorized WLS Solution: Nonlinear–Linear Case

In this case, the model of Stage 2 becomes linear. This may arise when all raw measurements are processed during Stage 1 and no lossy network components are involved in Stage 2, which reduces typically to a trivial model relating state variables with their estimates provided by Stage 1 [50]. Then, the nonlinear systems (3) and (4) reduce to

$$z = f_1(y) + e \tag{13}$$

$$y = Bx + e_y. \tag{14}$$

The general model can be easily particularized to this case by systematically replacing $f_2(x)$ and F_2 by Bx and B,

respectively. The most relevant implication is that the iterative system (7) reduces to the following linear one:

$$[B^T G_1 B]\hat{x} = B^T G_1 \tilde{y}. \tag{15}$$

Note that Stage 2 involves a constant, usually quite sparse, and trivial Jacobian, B, but a weighting matrix G_1, which compactly embeds most of the information associated with the original SE problem (raw measurement covariance, network topology, and parameters).

Like in the general nonlinear case, the factorized scheme may theoretically involve in this particular case several executions of Stages 1 and 2. In practice, however, it is seldom needed to repeat the two-stage process, unless the required accuracy is extremely high or the raw measurement set is abnormally noisy. A computationally less expensive alternative when subsequent solutions are needed consists in updating only the right-hand side of (15). This way, the *LU* or *QR* factors of the coefficient matrix do not have to be recomputed.

D. Distributed Implementation of the Two-Stage Procedure

The two-stage SE procedure, in the basic form outlined above, can be a better choice than the conventional scheme (particularly in the linear–nonlinear case) provided a suitable set of intermediate variables y can be found. The improved performance may take the form of reduced computational effort (simpler and/or constant Jacobian components, smaller size of the state vector during the iterative process), early bad data processing capability, simpler observability analysis, etc. [51].

However, for the multilevel paradigm envisioned in this paper, it is most important to consider a distributed environment, in which geographically scattered measurements can be naturally grouped in clusters, weakly coupled with neighboring sets. Fig. 6(a) illustrates this common situation for the case of four areas or clusters. Within each area, two types of variables can be distinguished: 1) internal variables, not involved in the measurement models of neighbor areas; 2) border variables, appearing in the measurement model of at least an adjacent area. Accordingly, the measurement model (1) can be decomposed in the following manner for the case of j clusters:

$$\begin{aligned} z_1 &= h_1(x_{i1}, x_b) + e_1 \\ z_2 &= h_2(x_{i2}, x_b) + e_2 \\ &\vdots \\ z_j &= h_j(x_{ij}, x_b) + e_j \end{aligned} \tag{16}$$

where x_{ik} represents the set of internal variables for area k and x_b comprises the union of all border variables. Ideally,

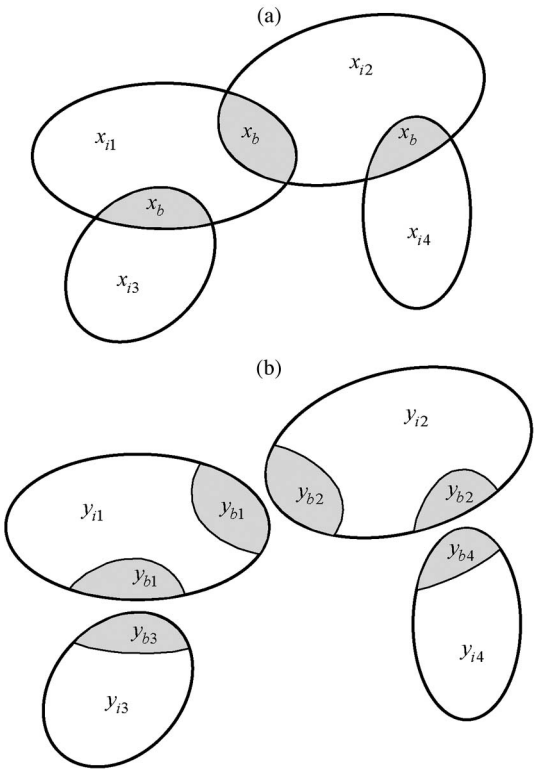


Fig. 6. *Measurement model composed of four natural clusters.*

the number of interior variables should be much larger than that of border variables, but in practice this may not be always the case. Notice that the set x_b constitutes the coupling term among all measurement submodels.

In this context, as justified in Section V, it is advantageous (sometimes even mandatory) to process the raw information as close as possible to the level (feeder, substation, etc.) in which it is captured, usually in a decoupled fashion. This calls for a distributed implementation of Stage 1, the aim of Stage 2 being essentially to coordinate or refine the preliminary solution provided by Stage 1 [49].

Keeping this goal in mind, the intermediate vectory should be selected in such a way that Stage 1 reduces to a set of fully decoupled SE subproblems, as shown in Fig. 6(b). Unlike x_b, whose components are shared by two or more areas, the vector of border variables y_b is composed of disjoint components, each involved in the measurement model of a single cluster. Depending on the measurement model and spatial structure of the particular SE problem being decomposed, the augmented vector y_b can be obtained by simply replicating border state variables of the original problem, adding certain measured magnitudes to the state vector, etc.

According to the geographical decomposition achieved, the measurement vector z is split into j components, each one exclusively related with the respective components of y. Therefore, the decoupled measurement

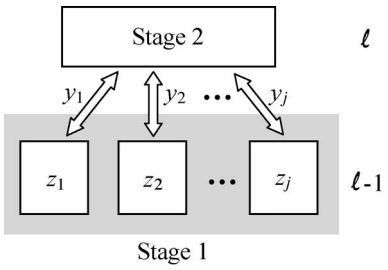


Fig. 7. ***Distributed implementation of Stage 1.***

models of Stage 1 can be mathematically formulated as follows:

$$\begin{aligned} z_1 &= f_{11}(y_{i1}, y_{b1}) + e_1 \\ z_2 &= f_{12}(y_{i2}, y_{b2}) + e_2 \\ &\vdots \\ z_j &= f_{1j}(y_{ij}, y_{bj}) + e_j. \end{aligned} \tag{17}$$

Fig. 7 schematically illustrates the interactions between both stages in case Stage 1 is performed in a distributed manner. The solution corresponding to each cluster can be iteratively obtained, in the nonlinear case, by solving the associated normal equation system (6).

Regarding Stage 2, consider for simplicity the nonlinear–linear case (extending the distributed formulation to the fully nonlinear context is straightforward). Let $\tilde{y}_b$ and $\tilde{y}_i$ be the estimates provided by Stage 1 of the border and interior variables, respectively, for all clusters $1, 2, \ldots, j$. Then, the mathematical model of Stage 2, for the linear case, can be written as follows:

$$\tilde{y}_b = Bx_b + e_b \tag{18}$$

$$\tilde{y}_i = x_i + e_i. \tag{19}$$

The weighting matrix arising in Stage 2 is a block diagonal matrix formed by simply juxtaposing the individual gain matrices corresponding to each cluster of Stage 1, yielding

$$G_1 = \begin{bmatrix} G_{11} & & & \\ & G_{12} & & \\ & & \ddots & \\ & & & G_{1j} \end{bmatrix} = \begin{bmatrix} \blacktriangle & \diamond & & & & & & \\ \diamond & \blacktriangledown & & & & & & \\ & & \blacktriangle & \diamond & & & & \\ & & \diamond & \blacktriangledown & & & & \\ & & & & \ddots & & & \\ & & & & & & \blacktriangle & \diamond \\ & & & & & & \diamond & \blacktriangledown \end{bmatrix} \tag{20}$$

where, as suggested by the rightmost matrix structure, each diagonal block is composed of four submatrices (self and mutual covariance terms between interior an border variables).

Symmetrically reordering the rows/columns so that border and interior variables are grouped together leads to the following blocked structure:

$$G_1 = \begin{bmatrix} G_{bb} & G_{ib} \\ G_{ib}^T & G_{ii} \end{bmatrix} = \left[\begin{array}{cccc|cccc} \blacktriangle & & & & \diamond & & & \\ & \blacktriangle & & & & \diamond & & \\ & & \ddots & & & & \ddots & \\ & & & \blacktriangle & & & & \diamond \\ \hline \diamond & & & & \blacktriangledown & & & \\ & \diamond & & & & \blacktriangledown & & \\ & & \ddots & & & & \ddots & \\ & & & \diamond & & & & \blacktriangledown \end{array}\right]. \tag{21}$$

Each and every major block above is composed of j decoupled blocks. Furthermore, the block sizes in G_{ii} (bottom right) are frequently much larger than those of G_{bb} (top left), which is an important feature from the computational point of view (e.g., partly distributed solution of Stage 2).

Based on the above notation, the normal equations to be solved at Stage 2 can be formulated

$$\begin{bmatrix} B^T & \\ & I \end{bmatrix} \begin{bmatrix} G_{bb} & G_{ib} \\ G_{ib}^T & G_{ii} \end{bmatrix} \begin{bmatrix} B & \\ & I \end{bmatrix} \begin{bmatrix} x_b \\ x_i \end{bmatrix} = \begin{bmatrix} B^T & \\ & I \end{bmatrix} \begin{bmatrix} G_{bb} & G_{ib} \\ G_{ib}^T & G_{ii} \end{bmatrix} \begin{bmatrix} \tilde{y}_b \\ \tilde{y}_i \end{bmatrix} \tag{22}$$

and, rearranging

$$\begin{bmatrix} B^T G_{bb} B & B^T G_{ib} \\ G_{ib}^T B & G_{ii} \end{bmatrix} \begin{bmatrix} x_b \\ x_i \end{bmatrix} = \begin{bmatrix} B^T G_{bb} & B^T G_{ib} \\ G_{ib}^T & G_{ii} \end{bmatrix} \begin{bmatrix} \tilde{y}_b \\ \tilde{y}_i \end{bmatrix}. \tag{23}$$

The above linear system could be directly solved by any of the generic techniques developed for symmetric systems (e.g., Cholesky or orthogonal factorization). However, this would be wasteful for two main reasons: 1) no advantage is taken of the distributed nature of the problem, reflected by the blocked structure of the coefficient matrix; 2) no benefits are obtained from the existing factorization of G_1, performed during Stage 1.

Many published procedures on multiarea SE ignore the mutual covariance between internal and border variables ($G_{ib} \approx 0$), as a means of quickly obtaining a suboptimal

solution. In fact, crude diagonal approximations of the covariance matrix components can also be found when formulating the coordination problem [52], [53]. As will be seen in the tutorial examples presented later, arbitrarily neglecting mutual covariance values may lead to poor solutions. Therefore, future SE implementations should be redesigned to account for nondiagonal weighting matrices, which is the price paid for the strengthened interactions of existing SEs with their neighborhood.

There exist at least two alternatives to more efficiently solve the system (23), taking into account the specific structure of the problem, that will be briefly discussed in the following.

1) Elimination of x_i: After trivial algebra, the system (23) can be reduced to

$$(B^T G_{\text{sch}} B) x_b = B^T G_{\text{sch}}\, \tilde{y}_b \tag{24}$$

where the resulting coefficient matrix

$$G_{\text{sch}} = G_{bb} - G_{ib} G_{ii}^{-1} G_{ib}^T$$

is known as the Schur's reduction of G_1.

The following remarks are in order regarding this solution approach.

- The vector component $\tilde{y}_i$, comprising the sets of interior variables for all clusters, is missing in (24), which is a nice feature. Note, however, that the influence in Stage 2 of the interior components of the solution provided by Stage 1 is properly exerted through the mutual covariance matrix G_{ib}.
- When computing G_{sch}, advantage should be taken of the block diagonal structure of G_1 components, as shown by (21).
- The inverse of G_{ii} is never computed, but the available sparse factors of its j block components should be instead resorted to. Therefore, computing G_{sch} reduces basically to repeated solutions of the sparse linear systems arising in Stage 1, which can be fully distributed.

2) Block Gauss–Seidel Iterations: The system (23) can be also rewritten as two coupled subsystems

$$(B^T G_{bb} B) x_b = B^T G_{bb}\, \tilde{y}_b - B^T G_{ib} (x_i - \tilde{y}_i) \tag{25}$$

$$G_{ii} (x_i - \tilde{y}_i) = G_{ib}^T (\tilde{y}_b - B x_b). \tag{26}$$

Then, a block iterative scheme can be implemented as follows.

1) Initialization: $x_i \leftarrow \tilde{y}_i$.
2) Obtain x_b by solving (25) with the most recent value of x_i.
3) Using x_b from the previous step, update x_i by solving (26).
4) Repeat steps 2) and 3) until convergence.

The computational effort and the amount of information exchanged is significantly reduced if step 3) above is fully distributed among the j processors in charge of Stage 1. In this regard, it is more convenient for the auxiliary vector

$$c_{ib} = G_{ib} (x_i - \tilde{y}_i) \tag{27}$$

to be locally computed and sent to the central processor, instead of x_i, in order to perform step 2).

Note that, when the iterative process is initialized in this way, the first value of x_b is the same as that obtained by neglecting G_{ib} in (23).

The former alternative, based on the reduced form, will be preferable in general when the number of border variables is comparatively small and/or the number of iterations for the latter scheme is high.

In the following sections, the two-stage model, information exchanged, and main features of the SE levels previously identified will be presented. A small example will be included illustrating each case. The basic idea is to consecutively apply the distributed version of the two-stage procedure (Fig. 7) to each pair of adjacent SE levels (Fig. 3).

VII. LOCAL STATE ESTIMATION

As discussed above, the advent of IED and PMU technologies at the substation level, along with the associated communication and computing infrastructures, is paving the way to the future "smart substation" [54]. Fig. 2 shows in schematic form the main actors arising in this distributed architecture, according to the IEC 61850 standard. This covers all aspects of the communications between a network of devices in the substation and the related systems.

In this environment, a huge number of measurement points, including those associated with protective devices, will provide information at a high sampling rate, which calls for the implementation of a LSE.

The LSE is useful not only to prefilter and reduce the size of the measurement set that should be sent to the TSE, but in many instances also for early detection of model and network inconsistencies at the substation level. For the redundancy ratios expected in next-generation substations, most topology errors and bad data could be handled at this level [55].

Several cases can be considered, depending on the substation voltage level and type.

A. Transmission Substation: Linear Model

This case arises by grouping together, in a single cluster or area, all busbar sections corresponding to the same rated voltage in a substation, along with the associated switching, protective and measurement devices [49]. A single substation then gives rise to as many areas as voltage levels. Fig. 8 shows a real substation with two voltage levels, each one leading to an area that is separately processed by the LSE.

An area so defined is therefore connected by nonzero impedance lossy branches (lines and power transformers, referred to hereafter as external branches) to its neighbors. On the other hand, internal connections take place exclusively through lossless components, leading to a straightforward linear model. This model, composed in turn of three decoupled submodels (active and reactive power and voltage magnitude), can be systematically built by resorting to well-known topological properties and concepts [56]. In the presence of PMU measurements, another submodel for phase angles could be considered.

Based on the so-called "proper tree," and the resulting links, a set of state variables can be chosen [34]. For convenience, the tree is selected in such a way that power flows through external branches are contained in the state vector. Those power flows, along with the voltage magnitudes of candidate electrical buses, constitute the border variables of each area or cluster (y_b). The power flows through the remaining links, necessary to complete the state vector, define the interior component (y_i).

A detailed model, relating existing measurements and topological constraints with the local state vector y, can be established for each area. Then, if sufficient redundancy locally exists, an estimate $\tilde{y}$ is provided by solving (11). The associate covariance matrix can also be obtained.

To illustrate the above ideas, consider for instance area 1 in Fig. 8, in which all CBs are closed. The associate

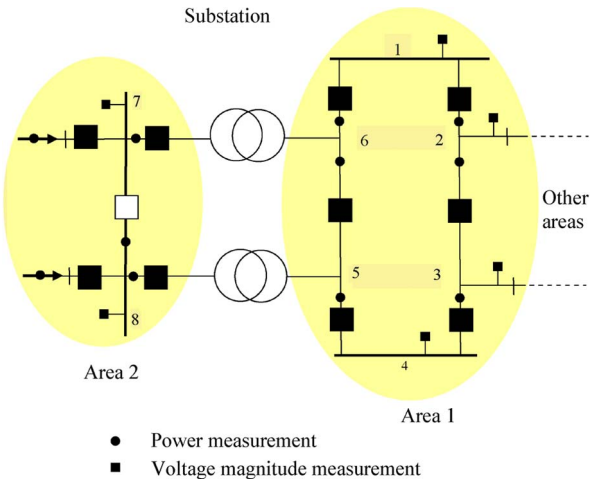


Fig. 8. *Sample substation and associate measurements.*

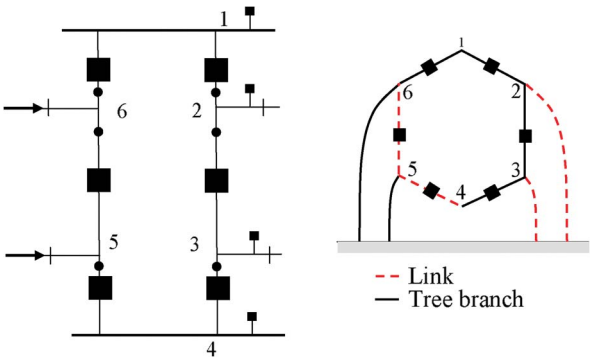


Fig. 9. *Sample linear substation and associate graph.*

graph and a suitable tree are shown in Fig. 9 (for simplicity, it is assumed that only buses 2 and 3 are connected to the rest of the TSO network, while buses 5 and 6 are representing external injections).

From the perspective of the two-stage procedure, in which the substation is considered a component (Stage 1) of a larger TSO system (Stage 2), the state vector can be decomposed as follows:

$$y_b^T = [P_{2e}, P_{3e}, Q_{2e}, Q_{3e}, V_1]$$
$$y_i^T = [P_{45}, P_{56}, Q_{45}, Q_{56}].$$

To estimate the above set of nine variables, 16 measurements are available, 12 power flows, and four voltage magnitudes (the measurement points correspond with those of a real substation).

The reader can easily verify that any measurement is a linear combination of the state variables. For instance, the expression for measurement P_{12} is

$$P_{12}^m = P_{2e} + P_{3e} + P_{45} + e_{p12}$$

and that of V_4

$$V_4^m = V_1 + e_{v4}.$$

Systematically gathering all the required relationships leads to a linear system of the form (9), which is composed in this case of three decoupled subsystems (P, Q, and V).

It is worth noting that the substation graph comprises just four loops, which means that the null-injection constraints P_1 and P_4 are implicitly considered when defining the graph [34], [56].

Table 1 Values of State Variables for the Linear Substation Case

		Exact	Estimated
y_b	P_{2e}	1.3000	1.3300
	P_{3e}	1.1000	1.0700
	Q_{2e}	0.1300	0.1265
	Q_{3e}	0.1100	0.1130
	V_1	1.0200	1.0180
y_i	P_{45}	-1.2000	-1.1900
	P_{56}	0.2000	0.1800
	Q_{45}	-0.1200	-0.1200
	Q_{56}	0.0200	0.0180

For a given measurement snapshot, generated by adding Gaussian noise to the exact values, the estimates shown in Table 1 are provided.

The resulting covariance matrix for the active subproblem is (same values result for the reactive one)

$$\operatorname{cov}\begin{bmatrix}\tilde{y}_b\\ \tilde{y}_i\end{bmatrix}_P = \operatorname{cov}\begin{bmatrix}P_{2e}\\ P_{3e}\\ P_{45}\\ P_{56}\end{bmatrix} = 10^{-5}\left[\begin{array}{cc|c} 15 & -10 & \\ -10 & 15 & -5 \\ \hline & -5 & 5 \\ & & 10 \end{array}\right].$$

As can be seen, the mutual terms coupling $\tilde{y}_b$ with $\tilde{y}_i$ are comparatively small or null in this case.

This example is useful to illustrate also a major advantage of the LSE over the conventional TSE, regarding the capability to individually handle all raw measurements. A standard SE, based on the bus-brunch model (B&B), relies on an auxiliary topology processor, aimed at identifying electrical buses and determining the net values of power measurements, as "seen" from the outer network. In this case, the topology processor could obtain the following active power "measurements" from actual raw measurements:

$$\begin{aligned} P_{2e} &= P_{12} - P_{23} \\ P_{3e} &= P_{23} - P_{34} \\ P_5 + P_6 &= -P_{16} - P_{45} \quad \text{(net power injection)} \end{aligned}$$

which means that six raw measurements are actually handled by the B&B model as three telemetered values (the same applies to the reactive power).

It should be noted that the local redundancy is relatively low in this example, which somewhat limits the possibility of properly identifying bad data at this level. For instance, if a 3.5% error is added to P_{12}^m the largest normalized residual is only 2.96 and the bad data flag is not triggered. On the other hand, when the added error is 4%, the largest normalized residual is 3.39 and bad data are detected. However, owing to the critical redundancy level of this example, both P_{12}^m and P_{61}^m share this abnormally high residual, which means that the bad data cannot be identified.

B. Transmission Substation: Nonlinear Integral Model

Extending the cluster notion to the entire substation (or even a few adjacent substations), and not only to the busbar sections of the same rated voltage, leads to a nonlinear model containing lossy elements. The advantage is that many more measurements can be locally processed, including phase angle differences provided by new generations of IEDs [55], increasing in this way the local redundancy. More detailed three-phase models could be even adopted if needed. The payoff is the added complexity of the nonlinear solution process, particularly when this constitutes the first stage of a two-stage TSE process.

The ideas will be illustrated with the help of the substation shown in Fig. 8, composed of two voltage levels ("linear substations" from the point of view of Section VII-A). For convenience, area 1 (right) will be again treated at the physical level, whereas the conventional B&B model will be adopted for area 2 (left). Fig. 10 represents the compact substation diagram, including one of the possible proper trees. The differences between the subtree corresponding to area 1 and the tree of Fig. 9 can be explained as follows (the reader can either skip these subtle details or see, for instance, [6, Ch. 8] for a deeper treatment).

- The external injections at buses 5 and 6 in the linear case become power transformers in this one. As nonzero impedance branches are links by definition, this implies that the CB branches 4–5 and 5–6 cannot be simultaneously links.
- A new branch at bus 4 (or at any other physical node within area 1) is added to take into account the null-injection constraints. These are needed to

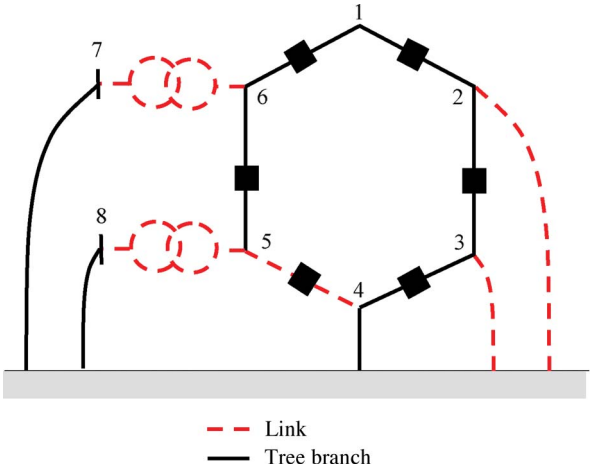


Fig. 10. *Graph associated with the substation of Fig. 8.*

assure that the physical subsystem composed of the CB ring is lossless (the sum of power flows through the four external branches is zero).

The state vector contains in this case a mixture of conventional and power flow variables (the latter required by the detailed physical model of area 1), as follows:

$$y = \begin{bmatrix} y_V \\ \hline y_P \\ y_Q \end{bmatrix} \quad \begin{cases} y_V^T = [V_1, V_7, V_8, \theta_7, \theta_8] \\ y_P^T = [P_{2e}, P_{3e}, P_{45}] \\ y_Q^T = [Q_{2e}, Q_{3e}, Q_{45}] \end{cases}$$

where the phase reference is taken at bus 1. This yields a total of 11 state variables, for which 26 measurements are available (six voltage magnitudes plus ten pairs of power flow measurements). In addition, nonlinear zero-injection constraints at bus 4 should be added.

The resulting measurement model comprises linear relationships, such as

$$P_{23}^m = P_{3e} + P_{45} + e_{p23}$$

as well as nonlinear ones

$$P_{76}^m = f_p(V_1, V_7, \theta_7) + e_{p76}$$

where $f_p(\cdot)$ is the well-known nonlinear function relating the active power flow through a transformer with the terminal bus voltages. The interested reader is encouraged to obtain the remaining expressions.

Notice that, unlike in the linear case, the variables in the subsets y_V, y_P, and y_Q become coupled in general by the presence of the nonlinear constraints.

From the perspective of the two-stage procedure, in which the substation is considered a component (Stage 1) of a larger TSO system (Stage 2), it is convenient to decompose the state vector as follows:

$$y = \begin{bmatrix} y_b \\ y_i \end{bmatrix}$$

where

$$y_b^T = [P_{2e}, P_{3e}, Q_{2e}, Q_{3e}, V_1]$$

and y_i contains the remaining variables (the internal ones).

For a measurement set which, regarding area 1, is identical to that of the linear substation case, the estimates corresponding to the state variables are shown in Table 2.

Table 2 Values of State Variables for the Nonlinear Substation Case

		Exact	Estimated
y_b	P_{2e}	1.3000	1.3296
	P_{3e}	1.1000	1.0755
	Q_{2e}	0.1300	0.1307
	Q_{3e}	0.1100	0.1232
	V_1	1.0200	1.0335
y_i	P_{45}	-1.2000	-1.1955
	Q_{45}	-0.1200	-0.1302
	V_7	1.0735	1.0851
	V_8	1.1099	1.1246
	θ_7	0.2322	0.2307
	θ_8	0.3169	0.3056

As stated earlier, the redundancy in the nonlinear case ($28/11 = 2.55$) is higher than in the linear one ($16/9 = 1.78$), which is an important aspect if topology errors or bad data have to be locally addressed. Referring again to the single bad datum analyzed in the linear substation case, affecting P_{12}^m, in this case even though the associated error is 3.5%, the largest normalized residual (3.45) exceeds the customary threshold 3, the next one in the ranking being 1.66. Therefore, such an error that could be barely detected in the linear case is safely identifiable when the entire substation is considered at once.

C. Distribution Substation and Associate Feeder System

Unlike transmission substations, in which very few lines and perhaps a lumped equivalent load are connected to each voltage level, distribution substations deliver power to a large amount of secondary transformers through a set of MV radial feeders (normally open switches allow backup service in case of a single failure).

As discussed in Section IV-D, the deployment of a new generation of sensors and meters, connected via relatively inexpensive communication channels with a DMS, in combination with well-established forecasting and data-mining techniques, is opening the way to the possibility of having a moderately redundant set of measurements and pseudomeasurements at the feeder level. In this upcoming context, the right tool capable of handling all these heterogeneous sources of information, according to their statistical quality, is the SE. In fact, SEs for distribution systems were considered long ago [57], when the required infrastructure was still far in the horizon.

In practice, up to 20 or more feeders may be connected to one or at most two MV busbar sections. As every feeder typically reaches dozens of secondary transformers (or even hundreds in rural dispersed areas), the size of the resulting SE problem may be discouraging. This section explains how the two-stage procedure can be used to take advantage of the weak electrical coupling among radial feeders.

A simplified distribution substation arrangement, with just two short feeders, is shown in Fig. 11 (the secondary

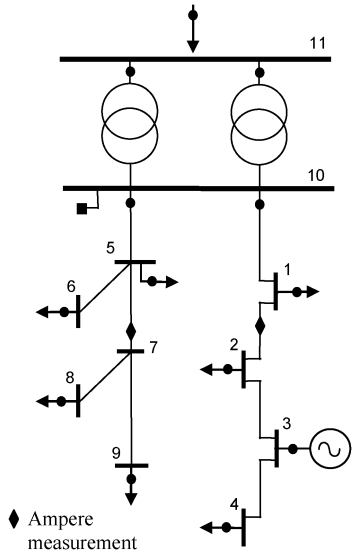


Fig. 11. *Sample distribution substation and associate feeders.*

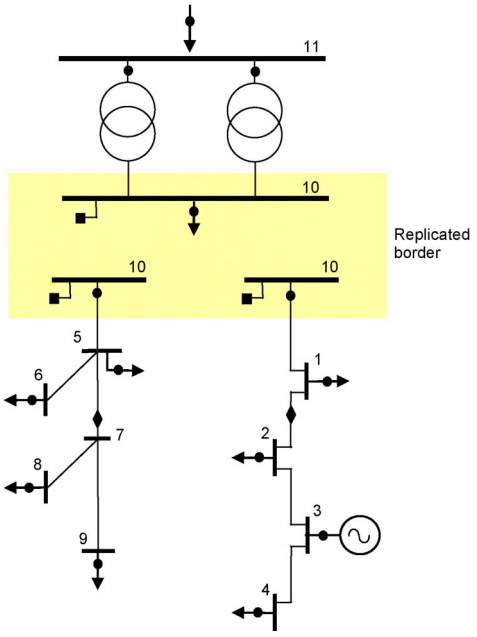


Fig. 12. *Resulting subsystems for the distribution substation.*

transformers at each feeder bus are not shown for simplicity). The feeder on the right, without laterals, is typically found in urban areas. It contains a DG, whose power injection exceeds in this example the load of the three remaining buses. While the power injected by the DG is accurately known, it is assumed that the load demand of buses 2 and 4, being provided by forecasting tools, is poorly defined. On the other hand, the feeder on the left corresponds with those of rural areas. In this case, a null-injection constraint (bus 7) coexists with power injections whose values are generally less accurate than those at transmission levels. Roughly in the middle of both feeders an Ampere measurement (provided for instance by a fault current detection device) is assumed. Typically, the measurement set within the substation is expected to be more accurate than the feeder measurements.

In this case, the border area allowing the overall SE problem to be decomposed into $f+1$ decoupled problems (f being the number of feeders) is simply the common bus from which the feeders hang, which is triplicated. Assuming the phase reference is taken for convenience at bus 10, the set of border variables reduces to V_{10} for all subproblems.

Fig. 12 shows the three decoupled systems that result in this example. The voltage magnitude measurement V_{10} can be used three times, once for each subsystem, but then its local weighting factor should be reduced accordingly (i.e., divided by three) so that its net effect, statistically speaking, is the same as if the global problem was solved at once. Similarly, the power flows at the head of both feeders can be used twice, once for each feeder and once more (in aggregated form, taking into account the null-injection constraint at bus 10) for the substation. Accordingly, their weighting coefficients should be divided by two.

As in any regular SE problem, the state vector for each subsystem comprises the voltage magnitudes and phase angles of all buses (the reference is in the border). Therefore, Stage 1 consists of solving three nonlinear systems of the form (17) by means of the iterative scheme (6). In turn, Stage 2 is composed of two linear systems, compactly represented by (18) and (19), which are coupled by the off-diagonal blocks of the gain matrices. In this simple case, the system (18) constraining the border variable reduces to

$$\begin{bmatrix} \tilde{V}_{10}^{(f1)} \\ \tilde{V}_{10}^{(f2)} \\ \tilde{V}_{10}^{(sub)} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} V_{10} + e_b.$$

Table 3 collects relevant results corresponding to the border magnitudes. The column under "Optimal" heading corresponds to the conventional solution of the entire problem, which is considered the optimal one. The minor differences that can be observed between "Optimal" and

Table 3 Values of Border Power Flows and Border Voltage for the Distribution Substation

	Exact	Optimal	Stage 2	$G_{ib}=0$
P_{10-1}	-0.1065	-0.1026	-0.1065	-0.1262
P_{10-5}	0.3819	0.3775	0.3763	0.3763
Q_{10-1}	-0.1204	-0.1213	-0.1266	-0.1398
Q_{10-5}	0.3910	0.3925	0.3897	0.3897
V_{10}	0.9913	0.9912	0.9912	0.9919

Stage 2" columns are due to the fact that Stage 2 has been run only once for this particular simulation. If the two-stage process was repeated, according to the complete methodology for the nonlinear–linear case (Section VI-D), the resulting differences between both columns would virtually vanish. Finally, the rightmost column shows the results that would be obtained by neglecting the off-diagonal blocks of G_1 (weighting matrix for Stage 2). As can be seen, the accuracy loss is not acceptable in this case, and presumably in all cases in which the coupling between border and interior variables is not negligible.

In this environment, both stages will run on the same computer, most likely located at the substation, which reduces the complexity associated with information exchanges between distant processors.

Upon completion of the two-stage local process, the border variables (in this case the power injections P_{11} and Q_{11} and the voltage magnitude V_{11}), along with the gain matrix arising during the last iteration, will be passed on to the TSO-level SE (the interior variables might also be involved if the block iterative scheme was adopted).

VIII. TSO-LEVEL STATE ESTIMATION

The remote data arriving from satellite LSEs (Stage 1), essentially composed of power flow and voltage magnitude estimates, as well as the associate statistical information, has to be integrated within the framework of the TSE (Stage 2). At this level, the resulting measurement model is mostly nonlinear, which involves the iterative solution of the system (7).

In this regard, the reader may recall that the design of the two schemes described in Section VI-D to solve Stage 2 in the linear case (reduced or iterative model) was based on the assumption that the number of clusters is moderate while the number of interior variables for each cluster is very high, which really calls for a distributed approach. However, the situation arising in the two-stage TSE is somewhat the opposite, namely the number of clusters j (i.e., substations) is very high, while the number of interior variables for each cluster is comparatively small (typically much lower than that of border variables, at least for transmission substations). Therefore, even though Stage 1 (substation level estimation) is performed in a distributed manner, the TSO-level coordination phase (Stage 2) should be better centralized, as the potential benefits of distributing the computations would be offset by the burden and complexity of exchanging information and coordinating the solution process with satellite processors.

Solving Stage 2 in a centralized manner provides the additional advantage that a standard SE can be resorted to, with minor adaptations. Major differences between the TSE that centrally performs Stage 2 and a conventional SE are [49] as follows.

- The two-stage TSE should be able to deal with nondiagonal yet sparse weighting matrices, characterizing estimates arriving from the substations. Indeed, conventional SEs customarily neglect coupling covariance terms between raw measurements provided by RTUs, which may not be acceptable in all cases [58]. Therefore, this should be a welcome addition. Anyway, as the weighting matrix is made up in this case of small diagonal blocks, one for each substation, this should not be a problem (in the linear substation case, each block is composed in turn of three decoupled blocks).
- A small set of additional state variables, corresponding to substation interior variables, must be accommodated into the state vector. Note however that this slightly augmented state vector would be anyway needed should the implicit GSE model be adopted [34].
- A reduced set of measurements, mostly composed of power flow and voltage magnitude measurements, arrives to Stage 2 (most injection measurements are locally processed). This may simplify to some extent certain auxiliary functions associated with the SE process.

It is worth noting that the computational effort of Stage 2 is virtually unaffected by the addition of more and more redundant measurements, as they are nearly always handled at the substation level (Stage 1). The relative computational efficiency of the two-stage SE increases therefore with the measurement redundancy. This will be particularly interesting in future "smart substations," with highly redundant measurement sets, where the proposed hierarchical approach will release the TSO-level SE from the burden of processing a huge number of raw measurements.

An added advantage of the two-stage scheme, arising from Stage 1 being implemented in a geographically distributed manner, is the reduction of the communication bandwidth requirements.

Other complementary functions, such as bad data processing, can also be partly performed in a distributed manner during Stage 1, provided the redundancy is high enough (this was illustrated in the previous section). Some bad data may have to be filtered out anyway at Stage 2.

In the following Section VIII-A and B, the TSE formulation will be separately addressed for the linear and nonlinear substation models considered above. A small example will be used to illustrate the ideas.

A. Linear Substation Model

As explained in Section VII-A, when a substation is decomposed into elemental subsystems, each associated with the physical equipment rated at the same voltage, the resulting local model is linear. The TSE model is then composed of as many linear clusters as electrical nodes, coupled in a nonlinear manner by external branches, including power transformers in this category. In this case, the two-stage linear–nonlinear methodology presented in

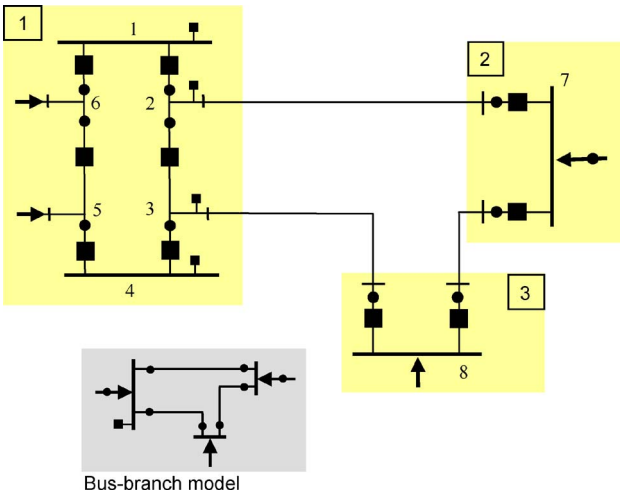


Fig. 13. *Sample transmission network with three linear substations.*

Section VI-B is directly of application. The role of the TSE is therefore to process in a centralized fashion, by solving (7), the prefiltered measurements provided by each cluster.

The sample system represented in Fig. 13 comprises three substations, one of them modeled in detail, as in Section VII-A. A conventional SE based on the B&B model would resort to the simplified diagram depicted at the bottom-left of the figure. In this case, the resulting measurements are generated during the topology processing stage by properly adding/subtracting raw measurements.

Table 4 shows, for a single snapshot, the simulation results corresponding to border power flows. From left to right, the table presents: 1) exact power flow values; 2) estimates provided by a generalized SE, capable of dealing with detailed substation models and, hence, with all raw measurements simultaneously (this is the optimal solution in the WLS sense); and 3) estimation errors corresponding to the conventional B&B model, Stage 1 and Stage 2, respectively, with respect to the optimal solution. In the last row, the mean of the absolute value errors (MAE) for the 12 power flow estimates is given. For the substation 1, the same measurements as in Section VII-A are used. In consequence, the results corresponding to Stage 1 are in agreement with those shown in Table 1 for y_b. Also, as Stage 1 is linear, the results provided by Stage 2 are optimal in this case after a single run (no need to recompute Stage 1 gain matrices). The results provided by the classical B&B model are suboptimal, but better than those after Stage 1. This is so because the B&B model, based on a subset of measurements generated by the topology processor, does not take full advantage of the redundancy that can be achieved when the detailed substation model is used. Moreover, as the raw measurements in substation 3 are locally critical, no possibility of local prefiltering exists in this case (Stage 1 is skipped for this substation).

B. Nonlinear Substation Model

In this case, each substation, including power transformers, is considered an indivisible entity, as in Section VII-B. The TSE model is then composed of as many nonlinear clusters as substations (several adjacent substations could be merged for convenience in a single cluster), coupled in a nonlinear manner by external branches (lines only). Note that transformerless substations reduce anyway to linear clusters. In this context, the fully nonlinear two-stage methodology presented in Section VI-A can be applied.

The system represented in Fig. 14 differs from that of Fig. 13 in substation 1, which now contains two power transformers (this affects also the B&B model, as shown at the bottom left).

Table 5 is the counterpart of Table 4. For substation 1, the same measurement data as in Section VII-B are used. Consequently, the results corresponding to Stage 1 are in agreement with those shown in Table 2 for y_b. In this case, it makes sense to repeat Stages 1 and 2 by updating the gain matrix of the only nonlinear cluster (substation 1). Interestingly, the results of this second run (estimation errors shown at the rightmost column) are hardly distinguishable

Table 4 Simulation Results for the Three-Bus System of Fig. 13

	Exact	Optimal	Estimation error (p.u.)		
			B & B	Stage 1	Stage 2
P_{12}	1.3000	1.2897	-4.9E-03	4.0E-02	4.0E-10
P_{13}	1.1000	1.0959	-1.0E-02	-2.6E-02	1.2E-10
Q_{12}	0.1300	0.1291	-5.8E-04	-2.6E-03	3.3E-11
Q_{13}	0.1100	0.1114	-8.1E-04	1.6E-03	4.2E-10
P_{21}	-1.2925	-1.2823	4.9E-03	1.7E-02	-3.6E-10
P_{23}	-0.5021	-0.4940	-3.3E-03	-1.1E-02	-3.2E-10
Q_{21}	-0.0923	-0.0918	2.9E-04	-1.1E-03	1.2E-10
Q_{23}	-0.0358	-0.0338	2.0E-05	2.3E-03	4.0E-10
P_{31}	-1.0958	-1.0917	9.9E-03		-1.0E-10
P_{32}	0.5031	0.4949	3.4E-03		3.3E-10
Q_{31}	-0.0889	-0.0903	4.3E-04		-3.4E-10
Q_{32}	0.0408	0.0386	7.0E-05		-3.8E-10
MAE			3.2E-03	1.3E-02	2.8E-10

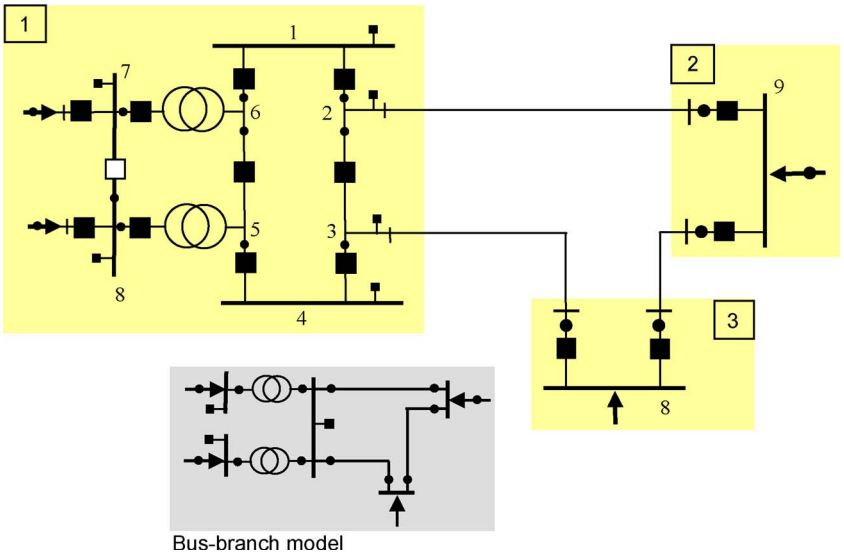


Fig. 14. ***Sample transmission network with a nonlinear substation.***

from those of the first one, and both can be deemed optimal for any practical purpose. Only when the measurement set contains unusually high noise levels is the second run justified. In the last row, the mean of the absolute value errors (MAE) for the 12 power flow estimates is given.

The results shown in Table 5 refer to a single snapshot. Repeating the experiment 1000 times provides a clearer picture of the resulting estimation accuracy for each SE model (random noise with $\sigma = 0.01$ is used to generate all measurements). Fig. 15 shows the probability density function (pdf) of the average estimation errors corresponding to the six pairs of border power flows. As expected, the conventional B&B model is more accurate than Stage 1 but less than Stage 2, according to the resulting redundancy levels (the pdf corresponding to the optimal solution is indistinguishable from that of Stage 2, first run).

In this context, it may be necessary to partly solve (7) in a distributed manner, for those more complex substations containing a large number of internal variables. This may apply, for instance, to distribution substations when feeders are involved in the SE process (Section VII-C).

IX. MULTI-TSO REGIONAL STATE ESTIMATION

The factorization-based approach described in Section VI-A provides also the theoretical basis for an upper SE hierarchy, in which the set of interconnected TSEs constitutes the first level and the RSE is the second level.

The geographical decomposition is achieved in a natural manner by replicating the state variables corresponding to tie-line terminal buses, which are shared by adjacent TSOs. In this context, both SE levels lead in general to nonlinear models, unless all raw measurements are processed at least once at the TSO level, which is generally the case. When this happens, the RSE solution reduces to a

Table 5 Simulation Results for the System of Fig. 14

	Exact	Optimal	Estimation error (p.u.)			
			B & B	Stage 1	Stage 2(I)	Stage 2(II)
P_{12}	1.3000	1.2937	-1.1E-02	3.6E-02	6.7E-07	-3.0E-08
P_{13}	1.1000	1.1025	-2.0E-02	-2.7E-02	9.4E-07	-4.2E-08
Q_{12}	0.1300	0.1317	-5.3E-03	-2.5E-03	-2.8E-06	-1.8E-07
Q_{13}	0.1100	0.1166	-8.3E-03	4.1E-03	-5.2E-06	-2.6E-07
P_{21}	-1.2925	-1.2864	1.1E-02	2.1E-02	-5.1E-07	2.8E-08
P_{23}	-0.5021	-0.4925	-5.3E-03	-1.3E-02	7.1E-08	-6.4E-09
Q_{21}	-0.0923	-0.0952	4.4E-03	3.8E-03	3.6E-06	1.7E-07
Q_{23}	-0.0358	-0.0326	-1.5E-03	5.6E-04	-1.2E-06	-3.1E-08
P_{31}	-1.0958	-1.0984	2.0E-02		-8.5E-07	4.0E-08
P_{32}	0.5031	0.4935	5.3E-03		-5.1E-08	6.2E-09
Q_{31}	-0.0889	-0.0958	7.5E-03		5.7E-06	2.5E-07
Q_{32}	0.0408	0.0372	1.6E-03		1.3E-06	3.0E-08
MAE			8.4E-03	1.3E-02	1.9E-06	8.9E-08

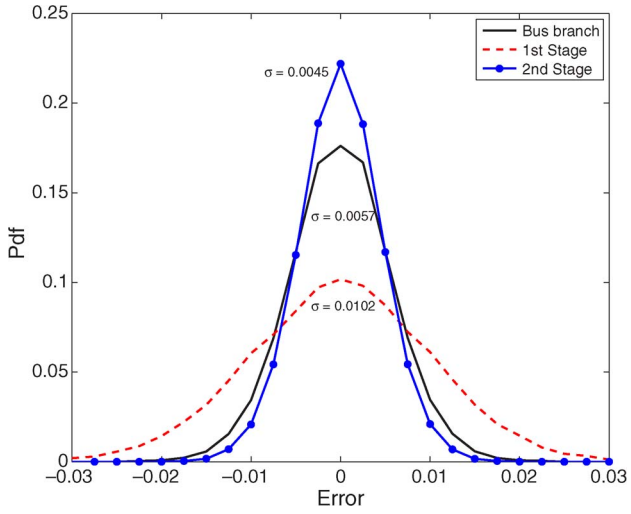


Fig. 15. ***Probability density function of average estimation errors.***

linear system in which the nonlinearity of the overall SE model is embedded into the weighting matrix, composed of the gain matrices provided by each TSE.

In the following, it will be assumed that the RSE model (Stage 2) is linear and that the blocked iterative solution scheme (25)–(26) described in Section VI-D is adopted. This increases the complexity of the interactions between each TSE and the RSE but has the advantage that all internal variables y_i are locally updated as a byproduct. Under these assumptions, main features of the RSE are as follows.

- Each TSO submits to the RSE its estimate of the border variables ($\tilde{y}_b$), along with the statistical information associated with the interior components, represented in compact form by the auxiliary vector (27). The overlapping degree determines the amount of border variables to be exchanged, but this is usually very small in relative terms. This is a welcome feature in nowadays deregulated systems, in which confidentiality of data is a major concern.
- The state vector at the RSE level comprises the border variables shared by the interconnected areas, as well as an auxiliary vector u composed of relative phase angle references, intended to synchronize all areas with respect to a global phase origin. Note that this was not an issue for the other applications of the two-stage procedure considered in this paper. In the distribution substation case, this is so because the common bus is chosen as the phase origin. At the TSO level phase angles are simply not included in the estimates provided by substations (in the presence of PMUs this may not be the case, as discussed in the next section).
- For each border component of the state vector two or more estimates are available in general, which increases the redundancy to estimate the auxiliary vector u.

A majority of multiarea SE procedures published up to date, combining heuristics and/or relaxation techniques with two-level WLS-based optimization, can be considered simplified particular cases of the general framework discussed herein, frequently providing suboptimal solutions [52], [53].

As an example, consider the system represented in Fig. 16, in which three IEEE 14-bus systems are interconnected by six tie-lines. The two scenarios reported in [50] will be summarized below.

A. Gaussian Errors

Two measurement sets with Gaussian errors are tested as follows.

- Set 1: containing the voltage magnitudes and the power injections at all the nodes, plus the power flows at both sides of all branches. This yields a total of 390 measurements and a redundancy of 4.64, for the entire 3-TSO system.
- Set 2: same as 1, except that branch power flows are measured at one terminal only. This yields a total of 258 measurements and a redundancy of 3.07.

The standard deviations associated with the measurements are: 0.01 for voltage magnitudes, 0.02 for power injections, and 0.015 for power flows. In the tests presented below, the voltage magnitude of border buses and tie-line power flow measurements are shared by neighboring TSOs, for which the respective weighting factors are suitably scaled.

Table 6 shows the 200-run average of the MAEs (taking exact values as references) associated with voltage magnitudes and phase angles for the complete set of state variables. The first row refers to the global SE solution, as if the entire system was a single TSO (optimal estimate). Two incomplete solutions of Stage 2 are added for comparison: $[1b-1i]$ stands for a single iteration of the block iterative scheme (25)–(26), in which both border and interior variables are updated only once after Stage 1; $[1b-0i]$ refers to the simplest possible arrangement, in which only border variables are updated.

As can be seen, even a single iteration of Stage 2 may provide sufficient estimation accuracy for practical purposes. Updating the interior variables at least once is however strongly recommended, particularly near the border. After four iterations, Stage 2 converges virtually to the reference solution, for a convergence threshold of 0.0001. As expected, more accuracy is obtained for measurement set 1.

B. Border Bad Data

The bad data detection and identification process is customarily implemented through the largest normalized residual ($|r_N|$) technique. For this purpose, several

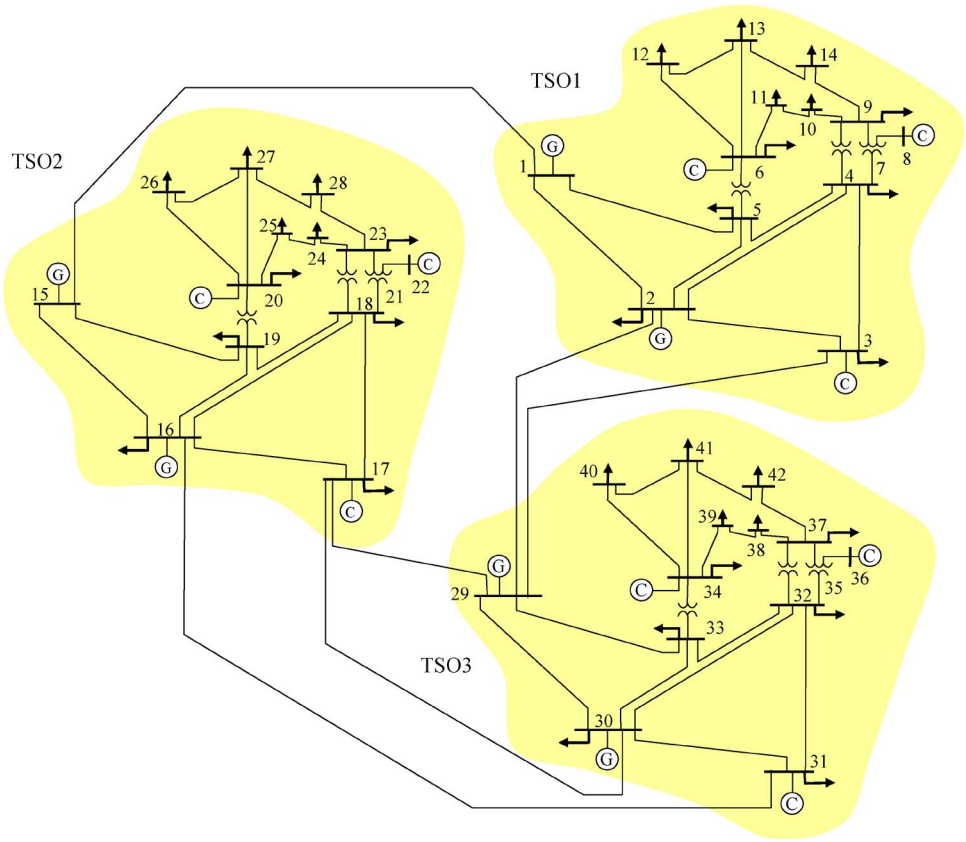


Fig. 16. ***Regional system composed of three 14-bus interconnected systems.***

scenarios are generated incorporating a single bad datum (10 σ) to measurements in the border area, the remaining ones (same two sets as before) being assumed as "exact." This way, the influence of random factors in the results is eliminated. Fig. 17 shows the five candidate bad data sequentially tested, all of them around the border bus 1.

Table 7 collects the two largest values of $|r_N|$ for each scenario, as well as the associate measurements, that result after the local TSO solution (Stage 1). Separate values are provided for the two involved partners (TSO 1 and TSO 2). It can be concluded that for the highly redundant

Table 6 Mean Absolute Value Errors for the Set of Border State Variables

Solution scheme	Meas. Set 1		Meas. Set 2	
	V	θ	V	θ
Optimal	0.00172	0.00136	0.00287	0.00227
[1*b*-1*i*]	0.00188	0.00137	0.00297	0.00229
[1*b*-0*i*]	0.00209	0.00138	0.00303	0.00228

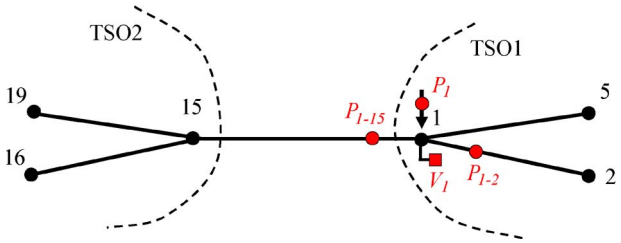


Fig. 17. ***Tested bad data around tie-line 1–15.***

Table 7 Two Largest $|r_N|$ for a Single Bad Datum After Stage 1

TSO 1				
Bad data	Meas. Set 1		Meas. Set 2	
	Max $\|r_N\|$	(meas.)	Max $\|r_N\|$	(meas.)
V_1	6.9213	V_1	6.8949	V_1
	0.4022	V_5	0.4991	$Q_{5,6}$
$P_{1,15}$	5.6626	$P_{1,15}$	4.4096	$P_{1,15}$
	3.1480	$P_{15,1}$	4.4059	P_1
P_1	7.1053	P_1	5.8763	P_1
	3.3575	$P_{1,15}$	5.8713	$P_{1,15}$
$P_{1,2}$	8.1889	$P_{1,2}$	6.5924	$P_{1,2}$
	3.8662	$P_{2,1}$	4.3346	P_1
TSO 2				
Bad data	Meas. Set 1		Meas. Set 2	
	Max $\|r_N\|$	(meas.)	Max $\|r_N\|$	(meas.)
V_1	6.9155	V_1	6.8732	V_1
	0.3717	V_19	0.7840	$Q_{1,15}$
$P_{1,15}$	5.6586	$P_{1,15}$	4.3957	$P_{1,15}$
	3.1599	$P_{15,1}$	4.3914	P_{15}

measurement set 1 the five bad data are correctly identified by both TSOs. However, with set 2 the bad data are detected but the identification is risky at least for $P_{1,15}$ and P_1.

Then, the first iteration of Stage 2 is partly performed, by updating only the border variables $(1b-0i)$, and the normalized residuals are computed again. In this case, all bad data are correctly identified by both TSOs.

Similar conclusions are obtained with not so extreme error levels. For instance, bad data on $P_{1,15}$ and P_1, when generated with 4σ, cannot be detected after Stage 1 and are barely identified after the first half iteration of Stage 2 $(1b-0i)$.

Therefore, bad data in tie-line and border bus measurements can be safely handled only at the RSE level (Stage 2), unless the redundancy is very high or the subset of measurements adjacent to the bad data is quite accurate. Indeed, this may be the case in the presence of PMUs, whose role in this multilevel paradigm will be discussed in Section X.

X. INCORPORATION OF PMUs

Phasor measurements provide phase angles of bus voltage and branch current phasors that have long been missing from measurement sets since the introduction of state estimation function in the control centers [36]. Availability of direct measurement of these phase angles necessitates some changes in the state estimation formulation and also opens up new opportunities of enhancement for existing SEs. While PMUs are being deployed in increasing numbers all over the globe, their placement is commonly carried out in limited number of units at a time due to physical, operational, and financial constraints. Moreover, depending upon the number of channels they have, their placement strategy will have to be modified.

One simple yet fundamental difference in the formulation of the state estimation problem is the disappearance of the user assigned reference angle [59]. When collecting phasor measurements from PMUs at a PDC, one of the PMUs is typically chosen as a reference and all other measurements are reported with respect to this reference angle. This practice may however lead to erroneous results if the chosen reference is incorrect or missing. This problem can be easily overcome by using a reference-free SE where conventional as well as phasor measurements are processed without externally assigning any reference angle.

Given a limited budget and communication infrastructure, PMUs can be placed at strategic locations in order to enhance energy management functions. The strategy to place PMUs will depend both on the type of PMUs being considered as well as the overall placement objective and priorities. PMUs can be assigned to buses or branches based on their available channels [60]. Assuming no channel limits on PMUs, it can be shown that full network observability can be achieved by strategically placing such

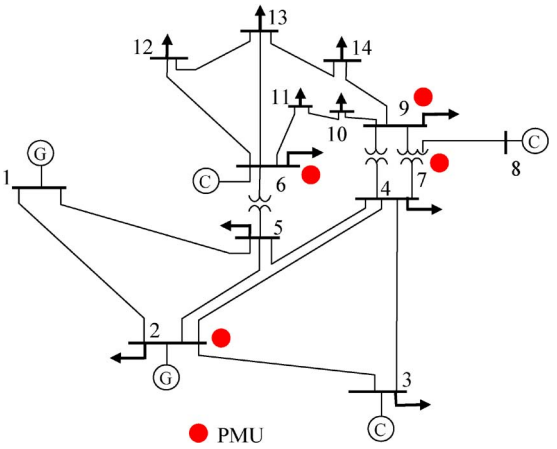


Fig. 18. *Optimal PMU placement for full network observability.*

PMUs at about one third of the system buses [61]. Considering the small 14-bus system of Fig. 18, the four required PMUs are indicated by circles next to their optimal bus locations. It is further shown that increasing the channel numbers will rapidly saturate the benefits of having extra channels when the objective is full network observability [62].

As in the case of conventional measurements, PMU measurements can be used to improve bad data processing capability in SEs [63]. It can be shown that using relatively few PMUs all existing critical measurements can be transformed into redundant ones. As a result, all bad data in any measurement can be ensured to be detected making the measurement configuration statistically more robust.

An important benefit of having access to synchronized phasor measurements at certain locations is shown to be related to network parameter error identification [64]. Errors creep into network data bases over the years due to poor maintenance or user mishandling. Other causes include environmental effects such as temperature that changes certain parameters by statistically significant amounts. Such errors can be detected and identified by techniques that can be built into the state estimation solution. However, there may be limits to the capability to identify such errors when using only conventional measurements. Examples of such cases can be found in [64]. Placement of PMUs at strategic locations in order to avoid such deficiencies of the measurement design belongs to the set of unique benefits provided by PMUs for state estimation.

A. Local State Estimation

In the case of transmission substations, as indicated in Section VII-A and B, availability of phasor measurements will introduce benefits. In the linear case, an additional submodel can be considered using the phase angles of measured voltage phasors. Depending upon the location,

number, and type of phasor measurements available at the substation (both areas 1 and 2 of Fig. 8), the model may remain linear despite the lossy elements. This is similar to the situation where exclusive use of phasor measurements leading to a linear SE problem formulation for the conventional single stage case. Furthermore, having synchronized measurements (at least one from each area) will allow independent handling of the two areas of the substation that can then be synchronized based on the phasor measurements, before communicating the results up to the TSO level. Caution has to be exercised by not relying on a set of critical phasor measurements in case any of them carry bad datum and bias the results with no possible way of bad data detection. Given that most PMUs provide redundant voltage measurements at the substation, bad data detection will practically not be a major issue.

The same benefits translate to the case of distribution substation as long as similar level of PMU deployment exists. Unfortunately, current emphasis in deploying PMUs is on transmission systems and most distribution feeders of today probably are not equipped with PMUs.

B. TSO-Level State Estimation

Any available PMU measurements will augment the set of data arriving from satellite LSEs (Stage 1) to the TSO. Again, availability of synchronized measurements will provide two benefits at the TSO level. One is the added redundancy, which will help strengthen bad data processing capability at the TSO level and will also provide a direct way to merge LSE results even when boundary measurements are not sufficient or bad. The second advantage is dependent on the number, type, and location of phasor measurements, which may allow the nonlinear case of Section VIII-B to be transformed into a linear problem.

As implied by theory of network observability, conventional measurements have to be incident to boundary buses and/or tie-lines in order to enable merging of observable islands, which can be considered as locally processed substations or subsystems. PMU measurements, however, have the advantage that they can be placed at any point within the observable islands and still will facilitate mergers of their islands with other islands, irrespective of them being physical neighbors or not. Hence, Stage 2 of the multilevel SE paradigm will be greatly enhanced as transmission and distribution substations continue to be populated with more PMUs.

C. Multi-TSO/Regional Level State Estimation

At the regional level, the main concerns of coordination between TSOs have been articulated in several earlier publications [24], [65]–[67]. As in the case of TSO-level coordination among various substations, proper and efficient handling of binding constraints are crucial. These appear in form of measurements that are incident at boundary buses as well as various types of synchronized phasor measurements, which may be located anywhere in the operating areas supervised by individual TSOs. Having PMUs spread around the different TSOs allows flexible coordination schemes without explicitly relying on boundary measurements, which may not have the necessary local redundancy. In principle, having one phasor measurement per transmission system should suffice for synchronization of individual SE solutions. However, in practice, this may lead to poor performance due to bad or missing data. This issue can be addressed by ensuring that none of the PMU measurements are critical. There are also extreme conditions such as having a single current phasor measurement along the tie-line connecting two transmission systems, in which case multiple solutions will be possible for the combined solution. However, such cases are highly unlikely, given the multiplicity of channels available on typical PMUs.

XI. CONCLUSION

SEs have been around for some 40 years. In spite of the tremendous progress made in computational efficiency, numerical stability, modeling issues, robustness against bad data, etc., the structural design of the existing SEs essentially remains identical to that of the 1970s, based on a centralized EMS in charge of a single TSO network with minor or no interactions with neighboring or subordinate systems.

This paper proposes an alternative vision based on the anticipated new role of SEs in the context of the future smart grids. The paper first provides an overview of the recent advances in sensors as well as signal processing and communication technologies (PMUs, IEDs, communication protocols, etc.). It then goes on to propose a new paradigm based on a multilevel computation and communication architecture, which can sustain growth and complexity of data and information flow in various parts of future energy systems as they are monitored more closely and synchronously.

The proposed multilevel scheme is a generalization of existing bi-level schemes, both at the local and regional geographical levels. A theoretical framework, supporting the adequacy and optimality of the resulting hierarchical scheme, is provided. ■

APPENDIX

Starting from the optimality conditions of the conventional WLS SE, the two-stage factorized approach will be inferred for the most general, fully nonlinear case. This provides the mathematical justification of the particular implementations and main steps described in Section VI.

The first-order optimality conditions (FOOCs) of the conventional WLS SE are compactly given by [6]

$$H(\hat{x})^T W[z - h(\hat{x})] = 0. \tag{28}$$

In geometrical terms, the above expression implies that the estimated residual vector, when properly weighted with matrix W, must be orthogonal to the columns of the Jacobian matrix H. The optimal estimate $\hat{x}$ satisfying the above expression is obtained by iteratively solving the normal equation (2).

The factorized approach to solve the WLS problem arises when a vectory of intermediate variables is introduced so that

$$\left.\begin{array}{l} z = f_1(y) \\ y = f_2(x) \end{array}\right\} \quad \Rightarrow \quad h(x) = f_1[f_2(x)] \tag{29}$$

which, by the chain rule, leads to the following relationship between the Jacobians:

$$H(x) = F_1(y)F_2(x). \tag{30}$$

Taking into account (29) and (30), the FOOCs (28) can be written as

$$F_2^T F_1^T W\{z - f_1[f_2(\hat{x})]\} = 0 \tag{31}$$

where the dependence of the Jacobians on x has been omitted for simplicity of notation. Then, adding and subtracting the term $f_1(\tilde{y})$, where the meaning of $\tilde{y}$ will be made clear later, the following FOOCs are obtained:

$$F_2^T\left\{F_1^T W[z - f_1(\tilde{y})]\right\} + F_2^T F_1^T W\{f_1(\tilde{y}) - f_1[f_2(\hat{x})]\} = 0. \tag{32}$$

Moreover, assuming $\tilde{y}$ is sufficiently close to $f_2(\hat{x})$, the following linearization of $f_1(\cdot)$ can be performed:

$$f_1(\tilde{y}) - f_1[f_2(x)] \cong F_1[\tilde{y} - f_2(\hat{x})]. \tag{33}$$

Finally, replacing the above expression into (32) leads to

$$F_2^T\left\{F_1^T W[z - f_1(\tilde{y})]\right\} + F_2^T\left(F_1^T W F_1\right)[\tilde{y} - f_2(\hat{x})] = 0. \tag{34}$$

Therefore, the optimal estimate $\hat{x}$ in (28) can be alternatively obtained by finding the pair $\langle \tilde{y}, \hat{x} \rangle$ satisfying simultaneously the set of equations

$$F_1^T W[z - f_1(\tilde{y})] = 0 \tag{35}$$

$$F_2^T\left(F_1^T W F_1\right)[\tilde{y} - f_2(\hat{x})] = 0 \tag{36}$$

where it is important to realize that the Jacobian matrices F_1 and F_2 should be computed at the solution point $\langle \hat{y}, \hat{x} \rangle$, with $\hat{y} = f_2(\hat{x})$.

The solution to the original problem (28) can be then decomposed into two successive stages, each involving the solution of a WLS problem, which can be summarized as follows.

1) Stage 1: Obtain the estimate $\tilde{y}$ satisfying (35) by iteratively solving the associated normal equations

$$\left[f_1^T W F_1\right]\Delta y_k = F_1^T W[z - f_1(y_k)]. \tag{37}$$

2) Stage 2: Using the value $\tilde{y}$ provided by Stage 1 as "measurement" vector in (36), obtain the estimate $\hat{x}$ by iteratively solving the resulting normal equations

$$\left[f_2^T G_1 F_2\right]\Delta x_k = F_2^T G_1[\tilde{y} - f_2(x_k)] \tag{38}$$

where the "weighting" matrix $G_1 = F_1^T W F_1$ is the gain matrix of Stage 1.

During the first solution of Stage 1 the Jacobian F_1 is updated at each iteration, according to the current value y_k. However, when a new value $\hat{x}$ is provided by Stage 2, the Jacobian F_1 gets obsolete. Therefore, after the first run of Stages 1 and 2, the Jacobian F_1 should be recomputed with $\hat{y} = f_2(\hat{x})$ and Stage 1 run again, but this time keeping the Jacobian and gain matrices constant. Then, Stage 2 should be repeated with updated values of $\tilde{y}$, and so on until Stage 2 provides close enough values of $\hat{x}$ in two consecutive runs. In summary, in order to reach the optimal solution, there should be a sequence of Stage 1 and Stage 2 runs until full convergence, with constant Jacobian in Stage 1 as required by new $\hat{x}$ values provided by Stage 2.

An interesting particular case arises when Stage 1 is linear, the factorized model being

$$\left.\begin{array}{l} z = Ay \\ y = f_2(x) \end{array}\right\} \quad \Rightarrow \quad h(x) = Af_2(x). \tag{39}$$

In this case, the Jacobian F_1 becomes the constant matrix A, and a single execution of Stages 1 and 2 provides the optimal solution. Even in the fully nonlinear case, a single run of Stages 1 and 2 provides acceptable results for practical purposes, provided the raw measurements are sufficiently accurate.

REFERENCES

[1] R. N. Anderson, The Pew Center on Global Climate Change and the National Commission on Energy Policy, "The distributed storage-generation smart electric grid of the future," in *Proc. Workshop, The 10–50 Solution: Technol. Policies for a Low-Carbon Future*, 2004.

[2] B. Stott, O. Alsac, and A. J. Monticelli, "Security analysis and optimization," *Proc. IEEE*, vol. 75, no. 12, pp. 1623–1644, Dec. 1987.

[3] F. C. Schweppe, J. Wildes, and D. B. Rom, "Power system static state estimation—Part I, II, III," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 1, pp. 120–135, Jan. 1970.

[4] F. C. Schweppe and E. J. Handschin, "Static state estimation in electric power systems," *Proc. IEEE*, vol. 62, no. 7, pp. 972–982, Jul. 1974.

[5] A. Monticelli, *State Estimation in Electric Power System. A Generalized Approach*. Norwell, MA: Kluwer, 1999.

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

[7] W. Tinney and R. Walker, "Direct solutions of sparse network equations by optimally ordered triangular factorization," *Proc. IEEE*, vol. 55, no. 11, pp. 1801–1809, Nov. 1967.

[8] A. Monticelli and A. Garcia, "Fast decoupled state estimators," *IEEE Trans. Power Syst.*, vol. 5, no. 2, pp. 556–564, May 1990.

[9] B. Stott, "Fast decoupled load flow," *IEEE Trans. Power App. Syst.*, vol. PAS-93, no. 3, pp. 859–869, May 1974.

[10] F. F. Wu, "Power system state estimation: A survey," *Int. J. Electr. Power Energy Syst.*, vol. 12, no. 2, pp. 80–87, Apr. 1990.

[11] M. Vempati, I. Slutsker, and W. Tinney, "Enhancements to givens rotations for power system state estimation," *IEEE Trans. Power Syst.*, vol. 6, no. 2, pp. 842–849, May 1991.

[12] A. Gjelsvik, "The significance of the lagrange multipliers in WLS state estimation with equality constraints," in *Proc. 11th Power Syst. Comput. Conf.*, Avignon, Aug. 1993, pp. 619–625.

[13] A. Gjelsvik, S. Aam, and L. Holten, "Hachtel's augmented matrix method—A rapid method improving numerical stability in power system static state estimation," *IEEE Trans. Power App. Syst.*, vol. PAS-104, no. 11, pp. 2987–2993, Nov. 1985.

[14] R. Nucera and M. Gilles, "A blocked sparse matrix formulation for the solution of equality-constrained state estimation," *IEEE Trans. Power Syst.*, vol. 6, no. 1, pp. 214–224, Feb. 1991.

[15] F. Alvarado and W. Tinney, "State estimation using augmented blocked matrices," *IEEE Trans. Power Syst.*, vol. 5, no. 3, pp. 911–921, Aug. 1990.

[16] A. Monticelli and F. F. Wu, "Network observability: Theory," *IEEE Trans. Power App. Syst.*, vol. PAS-104, no. 5, pp. 1042–1048, May 1985.

[17] G. R. Krumpholz, K. A. Clements, and P. W. Davis, "Power system observability: A practical algorithm using network topology," *IEEE Trans. Power App. Syst.*, vol. PAS-99, no. 4, pp. 1534–1542, Jul. 1980.

[18] K. A. Clements, "Observability methods and optimal meter placement," *Int. J. Electr. Power Energy Syst.*, vol. 12, no. 2, pp. 88–93, Apr. 1990.

[19] F. H. Magnago and A. Abur, "Unified approach to robust meter placement against bad data and branch outages," *IEEE Trans. Power Syst.*, vol. 15, no. 3, pp. 945–949, Aug. 2000.

[20] A. Monticelli and A. Garcia, "Reliable bad data processing for real-time state estimation," *IEEE Trans. Power App. Syst.*, vol. PAS-102, no. 5, pp. 1126–1139, May 1983.

[21] L. Mili, T. Van Cutsem, and M. Ribbens-Pavella, "Hypothesis testing identification: A new method for bad data analysis in power system state estimation," *IEEE Trans. Power App. Syst.*, vol. PAS-103, no. 11, pp. 3239–3252, Nov. 1984.

[22] M. K. Celik and A. Abur, "A robust WLAV state estimator using transformations," *IEEE Trans. Power Syst.*, vol. 7, no. 1, pp. 106–113, Feb. 1992.

[23] L. Mili, M. G. Cheniae, N. S. Vichare, and P. J. Rousseeuw, "Robust state estimation based on projection statistics of power systems," *IEEE Trans. Power Syst.*, vol. 11, no. 2, pp. 1118–1127, May 1996.

[24] T. Van Cutsem, J. L. Horward, and M. Ribbens-Pavella, "A two-level static state estimator for electric power systems," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 8, pp. 3722–3732, Aug. 1981.

[25] J. M. Ruiz Muñoz and A. Gómez-Expósito, "A line-current measurement based state estimator," *IEEE Trans. Power Syst.*, vol. 7, no. 2, pp. 513–519, May 1992.

[26] A. Abur and A. Gómez-Expósito, "Detecting multiple solutions in state estimation in the presence of current magnitude measurements," *IEEE Trans. Power Syst.*, vol. 12, no. 1, pp. 370–375, Feb. 1997.

[27] A. Gómez-Expósito and A. Abur, "Generalized observability analysis and measurement classification," *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 1090–1096, Aug. 1998.

[28] K. A. Clements, P. W. Davis, and K. D. Frey, "Treatment of inequality constraints in power system state estimation," *IEEE Trans. Power Syst.*, vol. 10, no. 2, pp. 567–574, May 1995.

[29] H. Singh, F. L. Alvarado, and W. Liu, "Constrained LAV state estimation using penalty functions," *IEEE Trans. Power Syst.*, vol. 12, no. 1, pp. 383–388, Feb. 1997.

[30] I. W. Slutsker, S. Mokhtari, and K. A. Clements, "Real time recursive parameter estimation in energy management systems," *IEEE Trans. Power Syst.*, vol. 11, no. 3, pp. 1393–1399, Aug. 1996.

[31] A. Monticelli, "Modeling circuit breakers in weighted least squares state estimation," *IEEE Trans. Power Syst.*, vol. 8, no. 3, pp. 1143–1149, Aug. 1993.

[32] O. Alsac, N. Vempati, B. Stott, and A. Monticelli, "Generalized state estimation," *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 1069–1075, Aug. 1998.

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

[34] A. de la Villa Jaén and A. Gómez-Expósito, "Implicitly constrained substation model for state estimation," *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 850–856, Aug. 2002.

[35] Z. Jun and A. Abur, "Identification of network parameter errors," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 586–592, May 2006.

[36] A. G. Phadke, "Synchronized phasor measurements in power systems," *IEEE Comput. Appl. Power*, vol. 6, no. 2, pp. 10–15, Apr. 1993.

[37] A. G. Phadke and J. S. Thorp, *Synchronized Phasor Measurements and Their Applications*. New York: Springer-Verlag, 2008.

[38] A. Monticelli and F. Wu, "A method that combines internal state estimation and external network modeling," *IEEE Trans. Power App. Syst.*, vol. PAS-104, no. 1, pp. 91–103, Jan. 1985.

[39] J. D. McDonald, *Electric Power Substation Engineering*. Boca Raton, FL: CRC Press, 2003.

[40] [Online]. Available: www.energie-schweiz.ch/imperia/md/content/energiemrkteetrgertechniken/elektrizitt/strompanne03/12.pdf

[41] [Online]. Available: www.energy-regulators.eu/portal/page/portal/eer_home/eer_publications/ceer_ergeg_papers/electricity/2007/e06-bag-01-06_blackout-finalreport_2007-02-06.pdf

[42] *IEEE Standard Definition, Specification and Analysis of Systems Used for Supervisory Control, Data Acquisition, and Automatic Control*, IEEE Std. C37.1.1994, 1994.

[43] K. Brand, V. Lohmann, and W. Wimmer, *Substation Automation Handbook*. Bremgarten, Switzerland: Utility Automation Consulting Lohmann, 2003.

[44] *IEEE Standard for Synchrophasors for Power Systems*, IEEE Std C37.118-2005.

[45] I. H. Valenzo, "Information architecture design for the electricity distribution network," M.S. thesis, Economics of Infrastructures Section of the Technology, Policy and Management, Delft Univ. Technol., Delft, The Netherlands, 2009.

[46] [Online]. Available: http://www.energy-regulators.eu/portal/page/portal/EER_HOME/EER_INITIATIVES

[47] [Online]. Available: http://www.desertec.org/

[48] A. Bose, "Smart transmission grid applications and their supporting infrastructure," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 11–19, Jun. 2010.

[49] A. Gómez-Expósito and A. de la Villa Jaén, "Two-level state estimation with local measurement pre-processing," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 676–684, May 2009.

[50] European Community's 7th Framework Programme, "Algorithms for State Estimation of ETN," Deliverable 2.1 part 1, PEGASE Project (Pan European Grid Advanced Simulation and state Estimation), 2010.

[51] C. Gómez-Quiles, A. de la Villa Jaén, and A. Gómez-Expósito, "A factorized approach to WLS state estimation," *IEEE Trans. Power Syst.*, DOI: 10.1109/TPWRS.2010.2096830.

[52] T. Van Cutsem and M. Ribbens-Pavella, "Critical survey of hierarchical methods for state estimation of electric power systems," *IEEE Trans. Power App. Syst.*, vol. PAS-102, no. 10, pp. 3415–3424, Oct. 1983.

[53] A. Gómez-Expósito, A. de la Villa Jaén, C. Gómez-Quiles, P. Rousseaux, and T. Van Cutsem, "A taxonomy of multi-area state estimation methods," *Electric Power Syst. Res.*, DOI: 10.1016/j.epsr.2010.11.012.

[54] H. M. Kim, J. J. Lee, and D. J. Kang, "A platform for smart substations," in *Proc. Future Generat. Commun. Netw.*, 2007, vol. 2, pp. 579–582.

[55] A. de la Villa Jaén, P. Cruz Romero, and A. Gómez-Expósito, "Substation data validation by local three-phase generalized state estimators," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 264–271, Feb. 2005.

[56] A. Gómez Expósito and A. de la Villa Jaén, "Reduced substation models for generalized state estimation," *IEEE Trans. Power Syst.*, vol. 16, no. 4, pp. 839–846, Nov. 2001.

[57] C. N. Lu, J. H. Teng, and W.-H. E. Liu, "Distribution state estimation," *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 229–240, Feb. 1995.

[58] E. Caro, A. Conejo, and R. Mínguez , "Power system state estimation considering measurement dependencies," *IEEE Trans. Power Syst.*, vol. 24, no. 4, pp. 1875–1885, Nov. 2009.

[59] J. Zhu and A. Abur, "Effect of phasor measurements on the choice of reference bus for state estimation," in *Proc. IEEE PES General Meeting*, Tampa, FL, Jun. 24–28, 2004, DOI: 10.1109/PES.2007.386175.

[60] B. Xu and A. Abur, "Observability analysis and measurement placement for system with PMUs," in *Proc. IEEE PES Power Syst. Conf. Expo.*, New York, Oct. 10–13, 2004, vol. 2, pp. 943–946.

[61] B. Xu, J. Y. Yeo, and A. Abur, "Optimal placement and utilization of phasor measurements for state estimation," in *Proc. Power Syst. Comput. Conf.*, Liege, Belgium, Aug. 22–26, 2005, pp. 1550–1555.

[62] M. Korkali and A. Abur, "Placement of PMUs with channel limits," in *Proc. IEEE Power Energy Soc. General Meeting*, Calgary, CA, Jul. 26–30, 2009, DOI: 10.1109/PES.2009.5275529.

[63] J. Chen and A. Abur, "Placement of PMUs to enable bad data detection in state estimation," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1608–1915, Nov. 2006.

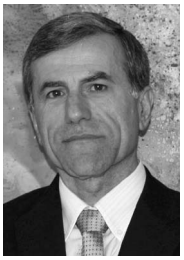
[64] J. Zhu and A. Abur, "Identification of network parameter errors using phasor measurements," in *Proc. IEEE Power Energy Soc. General Meeting*, Calgary, CA, Jul. 26–30, 2009, DOI: 10.1109/PES.2009.5275490.

[65] W. Jiang, V. Vittal, and G. T. Heydt, "Diakoptic state estimation using phasor measurement units," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1580–1589, Nov. 2008.

[66] L. Zhao and A. Abur, "Multiarea state estimation using synchronized phasor measurements," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 611–617, May 2005.

[67] J. Weiqing, V. Vittal, and G. T. Heydt, "A distributed state estimator utilizing synchronized phasor measurements," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 563–571, May 2007.

ABOUT THE AUTHORS



Antonio Gómez-Expósito (Fellow, IEEE) received the Industrial Engineering degree in electrical engineering and the Doctor Engineering degree, both with honors, from the University of Seville, Seville, Spain, in 1982 and 1985, respectively.

Currently, he is a Full Professor at the University of Seville, where he is chairing the Department of Electrical Engineering, the Endesa Red Professorship, and the Electrical Energy Systems Post-Graduate Program. In addition to nearly 250 technical publications, he has coauthored several textbooks and monographs about circuit theory and power system analysis, among which *Power System State Estimation: Theory and Implementation* (New York: Marcel Dekker, 2004) and *Electric Energy Systems: Analysis and Operation* (Boca Raton, FL: CRC Press, 2008), stand out.

Dr. Gómez-Expósito is a Fellow of the IEEE for his work on power systems analysis and operation, and an IEEE/PES Distinguished Lecturer. He serves on the Editorial Board of the IEEE TRANSACTIONS ON POWER SYSTEMS starting in 2011.

Ali Abur (Fellow, IEEE) received the B.S. degree from Orta Doğu Teknik Üniversitesi, Ankara, Turkey, in 1979 and the M.S. and Ph.D. degrees from the Ohio State University, Columbus, in 1981 and 1985, respectively.

He was a Professor at Texas A&M University until November 2005 when he joined the faculty of Northeastern University, Boston, MA, as a Professor and Chair of the Electrical and Computer Engineering Department. He coauthored the book *Power System State Estimation: Theory and Implementation* (New York: Marcel Dekker, 2004), contributed to several book chapters, and published widely in IEEE journals and conferences. His research and educational activities have been in the area of power systems.

Dr. Abur is a Fellow of the IEEE for his work on power system state estimation. He serves on the Editorial Board of the IEEE TRANSACTIONS ON POWER SYSTEMS and POWER ENGINEERING LETTERS. He is an IEEE/PES Distinguished Lecturer.

Antonio de la Villa Jaén was born in Riotinto, Spain, in 1960. He received the electrical and doctor engineering degrees from the University of Seville, Seville, Spain, in 1983 and 2001, respectively.

Currently, he is an Associate Professor at the Department of Electrical Engineering, University of Seville. His primary areas of interest are computer methods for power system state estimation problems.

Catalina Gómez-Quiles (Student Member, IEEE) received the Eng. degree from the University of Seville, Seville, Spain, in 2006 and the M.Eng. degree from McGill University, Montreal, QC, Canada, in 2008, both in electrical engineering. She is currently working towards the Ph.D. degree in power system state estimation at University of Seville.

Her research interests include risk assessment in competitive electricity markets and mathematical and computer models for power system state estimation.

A Factorized Approach to WLS State Estimation

Catalina Gómez-Quiles, *Student Member, IEEE*, Antonio de la Villa Jaén, and Antonio Gómez-Expósito, *Fellow, IEEE*

***Abstract*—This paper presents a factorized state estimation methodology based on the solution of two successive WLS problems. In the proposed scheme, a minimal set of intermediate variables is first introduced so that the resulting measurement model is linear. Estimates of those variables are then used as pseudo-measurements of a subsequent nonlinear estimator, along with the associated covariance matrix. The aim of the preliminary step is to reduce the size of the raw measurement vector to the maximum extent, without losing any relevant statistical information. Simulation results show that the proposed approach converges faster, is computationally more efficient, and provides accurate estimates after the first linear stage.**

***Index Terms*—Linear state estimation, measurement prefiltering, two-step state estimation.**

I. Introduction

STATE estimators (SE) determine the most likely state of a power system from sets of measurements remotely captured and periodically collected by SCADA systems through remote terminal units (RTU). The role of the SE is crucial in modern energy management systems (EMS), where a diversity of applications rely on accurate system snapshots. The new regulatory paradigm arisen in the last decade has stressed the importance of the SE tool, in an open-access context in which many more transactions on much more congested networks have to be properly tracked and recorded.

Since its introduction by Schweppe, in the late 1960s [1], the SE tool has benefited from a large number of theoretical developments and practical improvements, which can be tracked in [2] and [3]. From a computational point of view, in addition to the pioneering work of Tinney and others about sparsity exploitation [4], the most outstanding contribution is surely the so-called fast decoupled state estimator (FDSE), proposed by Monticelli *et al.* [5], after the introduction of the successful fast decoupled load flow (FDLF) concept [6]. The decoupled SE formulation requires less memory and saves a significant fraction of computations by solving smaller equation systems with constant coefficient matrices. For these reasons, the FDSE has been the subject of continued interest in an attempt to improve its performance [7], to develop an equivalent formulation in rectangular coordinates [8], or to explore its adequacy for distribution networks [9]. However, under certain operating conditions and/or combinations of network parameters, the overall convergence of the FDSE may significantly deteriorate. Moreover, when the decoupling principle is exploited to the limit, the Jacobian is not accurately represented on the right-hand side and the FDSE solution is no longer the maximum likelihood estimate.

During the last two decades or so, the amount of memory and speed of all categories of computers, including the smallest ones, have grown exponentially, the computational cost being frequently deemed a less relevant issue nowadays. Indeed, SEs in modern EMS can be run in theory several times a minute, provided the scan rate is high enough and the number of bad data is moderate. However, recent technological developments, such as the undergoing deployment of intelligent electronic devices (IED) at the substation level [10], will provide a huge number of measurement points, including those associated with protective devices. Furthermore, there is a clear trend to broaden the geographical scope of many SEs, in accordance to the needs of regional electricity markets, in which long-distance energy transactions have to be accurately and permanently monitored [11].

Therefore, SEs should be capable of coping efficiently with larger networks and huge measurements sets, which means that speed of response and convergence reliability will always be a concern, no matter how powerful the future computers are.

Recently, a two-level SE model has been proposed allowing a majority of raw measurements to be preprocessed at the substation level by a linear estimator [12]. The information provided by this stage, mostly composed of power flows and voltage magnitudes, is then optimally integrated within the framework of a conventional SE. The main advantage of this hierarchical procedure, particularly in environments with highly redundant measurement sets, is that the linear prefiltering phase can be run in a geographically distributed manner, significantly reducing the bandwidth requirements. As a byproduct, a modest but non-negligible reduction in the computational effort is also achieved.

In this paper, the possibility of previously filtering the raw measurements through a linear SE is further explored, for which an alternative and smaller set of intermediate variables is adopted. While the main goal of reference [12] was to distribute as much as possible the computations, reaching the substation granular level, in this work, the main outcome is the availability of a linear SE spanning the entire interconnected network, rather than isolated substations. For the noise levels typically found in real measurement sets, the solution of this linear SE, which always converges by definition, is a useful byproduct.

The structure of this paper is as follows: Section II reviews the conventional nonlinear WLS SE formulation. Section III describes the general framework of the two-level factorized SE formulation. Section IV presents the set of intermediate vari-

Manuscript received July 15, 2010; revised November 02, 2010 and November 26, 2010; accepted November 27, 2010. Date of publication January 10, 2011; date of current version July 22, 2011. This work was supported by the Spanish DGI, under grants ENE2007-62997 and ENE2010-18867. Paper no. TPWRS-00566-2010.

The authors are with the Department of Electrical Engineering, University of Seville, Seville, Spain (e-mail: catalinagq@us.es; adelavilla@us.es; age@us.es).

Digital Object Identifier 10.1109/TPWRS.2010.2096830

ables adopted in this work and explains in detail the resulting SE models. Section V briefly discusses the observability and bad data processing issues. Finally, test results, based on IEEE benchmark systems, are presented in Section VI, followed by the concluding remarks.

II. Conventional WLS State Estimation

Given the following measurement equation [1]:

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

where

- x state vector to be estimated (size $n = 2N - 1$);
- z known measurement vector (size $m > n$);
- h vector of functions, usually nonlinear, relating error-free measurements to the state variables;
- e vector of measurement errors, customarily assumed to have a Normal distribution with zero mean and known covariance matrix R. When errors are independent, R is a diagonal matrix with values σ_i^2, where σ_i is the standard deviation of the error associated with measurement i.

the weighted least squares (WLS) estimator will minimize the weighted squares of residuals of the measurements given by

$$J = \sum_{i=1}^{m} W_i r_i^2$$

where

- $r_i = z_i - h_i(\hat{x})$ measurement residual;
- $\hat{x}$ estimated state vector;
- W_i respective weighting coefficient.

The minimum of the scalar J can be obtained by iteratively solving the so-called Normal equations:

$$G_k \Delta x_k = H_k^T W [z - h(x_k)] \tag{2}$$

where:

- $H_k = \partial h / \partial x$ Jacobian evaluated at $x = x_k$;
- $G_k = H_k^T W H_k$ gain matrix;
- $W = R^{-1} = \mathrm{diag}(W_i)$ weighting matrix;
- $\Delta x_k = x_{k+1} - x_k$ k being the iteration counter.

Iterations finish when an appropriate tolerance is reached on Δx_k. The covariance of the estimate is

$$\mathrm{cov}(\hat{x}) = G_k^{-1}.$$

Upon convergence, the bad data processing function is activated to detect, identify, and eliminate bad analog measurements. Bad data detection is accomplished based on the largest normalized residual test [13]. If the detection test fails, then the measurement corresponding to the largest normalized residual will be declared bad and its value will be removed or corrected as given in [3, Ch. 5].

State estimation will be repeated as many times as needed after each identification and correction of a bad datum. Repetitive solutions will start from the most recent estimate instead of flat start, and hence will take fewer iterations to converge.

Note that multiple measurements of the same magnitude are fully compatible with the above formulation. For instance, bus voltage magnitudes may be available from two or more measuring instruments at a given substation bus. Instead of using a simple averaging, all those measurements can be included as separate entries in the measurement vector z, keeping track of their respective standard deviations of errors.

In conventional bus-branch SE models, the state vector is composed of voltage magnitudes and phase angles, whereas the measurement vector typically comprises measurements of power injections, branch power flows, and voltage magnitudes.

III. Factorized State Estimation Methodology

Following [12], the basic idea is to express the vector of nonlinear measurement functions, $h(x)$, arising in the conventional formulation (1), as the product of two terms, one of them linear. For this purpose, a set of intermediate variables y is introduced in such a way that the measurement model becomes

$$z = Ay + e \tag{3}$$

where A is a constant matrix of appropriate size and y is, in turn, a nonlinear function of x:

$$y = f(x). \tag{4}$$

Then, the factorized measurement model becomes

$$h(x) = Af(x). \tag{5}$$

Furthermore, the chain rule leads to

$$H(x) = AF(x) \tag{6}$$

where $F(x) = \partial f / \partial x$ is the Jacobian of $f(x)$.

The dimension l of the auxiliary vector y always satisfies the inequality $m > l > n$, but ideally it should be as small as possible. The extreme case $l = n$ arises only when the initial measurement model is fully linear. On the other hand, the unlikely case $l = m$ would mean that there is no way to factorize the original measurement model as suggested above.

The WLS solution $\tilde{y}$ to the linear problem (3) is obtained from

$$G_a \tilde{y} = A^T W z \tag{7}$$

where

$$G_a = A^T W A \tag{8}$$

is the respective gain matrix.

Taking into account the set of (5)–(8), the following relationships are obtained:

$$G = H^T W H = F^T G_a F \tag{9}$$

$$H^T W [z - h(x)] = F^T G_a [\tilde{y} - f(x)]. \tag{10}$$

Therefore, the original Normal (2) are fully equivalent to the system:

$$\left[F_k^T G_a F_k\right] \Delta x_k = F_k^T G_a [\tilde{y} - f(x_k)]. \tag{11}$$

Instead of directly performing iterations with (2), the above theoretical development provides the basis for an alternative but equivalent way of solving the original WLS problem in two stages: 1) compute $\tilde{y}$ by solving the linear system (7); 2) repeatedly solve (11) until convergence.

In this two-stage scheme, the auxiliary vector y plays the role of a state vector when solving (7) and that of a measurement vector when solving (11). The covariance matrix of $\tilde{y}$ is

$$\mathrm{cov}(\tilde{y}) = G_a^{-1}$$

the inverse of which (i.e., the sparse gain matrix G_a) appears as the weighting matrix in (11), like in a conventional WLS solution.

The resulting factorized scheme may or may not be computationally advantageous over the standard WLS solution, depending on the size of y and the sparsity of the involved matrices and vectors.

IV. Proposed Factorized Model

The factorized scheme proposed in an earlier work [12] was mainly devised to allow existing raw measurements to be pre-filtered at the substation level, in a fully distributed manner. This way, only a small subset of pseudomeasurements (mainly pre-estimated power flows and voltage magnitudes) is submitted for further processing and coordination of the solution at the central level. This approach makes sense particularly in the new generation of "intelligent" substations, where computing power locally exists and many more sources of information than in classical RTUs are deployed.

However, as every branch in the system contributes four power flow variables to the intermediate vector y (P and Q at each terminal substation), the resulting global redundancy is much lower during the linear preliminary phase (certain substations, equipped with conventional RTUs, may be locally unobservable, even if global observability is assured).

In this paper, a minimal set of two variables per branch are introduced in such a way that any power (flow or injection) measurement becomes strictly a linear function of those variables, along with squared voltage magnitudes.

Major differences between [12] and this work can be summarized as follows.

- In addition to bus voltage magnitudes, the set of auxiliary variables introduced in [12] comprises all power flows leaving each substation, which amounts to a total of $4b+N$ variables for a network with b branches and N buses. In this work, a set of intermediate variables, composed of just $2b$ branch variables plus N bus variables (squared voltage magnitudes), is introduced. This leads to a smaller model during the linear prefiltering stage and, what is more important, smaller Jacobian and residual vector during the subsequent nonlinear SE, which is the key to reduce the computational effort.
- The generous set of variables adopted in [12] allows the linear prefiltering phase to be fully distributed at the substation level, which implies an intrinsic parallelism when handling the raw measurements. The pay-off of the auxiliary variables selected in this paper is that the linear preprocessing stage cannot be so easily distributed, although the possibility of exploiting the existing parallelism to a certain extent should not be discarded.
- The model reduction achieved in this paper may allow larger network portions to be observable during the first linear procedure, even if certain power flows are not measured. When the enlarged model of [12] is adopted, missing power flow measurements lead to locally unobservable or critically observable substations.
- The substation level SE arising in [12], being a local processor, provides no information about phase angle differences between adjacent substations. As will be seen later, the linear prefilter proposed in this paper delivers a feasible set of initial values, both for voltage magnitudes and phase angles, which are sufficiently accurate in many instances.

A. Intermediate State Variables

For each network branch connecting buses i and j, the following pair of variables is defined:

$$K_{ij} = V_i V_j \cos\theta_{ij} \tag{12}$$
$$L_{ij} = V_i V_j \sin\theta_{ij} \tag{13}$$

where $V_i|\underline{\theta_i}$ is the complex voltage at bus i. In addition, squared voltage magnitudes

$$U_i = V_i^2 \tag{14}$$

replace voltage magnitudes in the intermediate state vector. Such a set of variables was first introduced in the context of radial load flows for distribution networks [14].

The intermediate state vector, composed of $2b+N$ variables, is then given in block partitioned form by

$$y^T = [\,K \quad L \quad U\,]. \tag{15}$$

B. First-Step State Estimator

Any measurement handled by conventional SEs can be linearly expressed in terms of the vector defined above as follows.

- *Power flow measurements*. For a branch connecting buses i and j, measurements taken at terminal bus i:

$$P_{ij} = (g_{sh.i} + g_{ij})U_i - g_{ij}K_{ij} - b_{ij}L_{ij} + \varepsilon_P \tag{16}$$
$$Q_{ij} = -(b_{sh.i} + b_{ij})U_i + b_{ij}K_{ij} - g_{ij}L_{ij} + \varepsilon_Q \tag{17}$$

  where g_{ij}, b_{ij} are the series conductance and susceptance, respectively, and g_{sh}, b_{sh} represent the shunt values.
- *Power injection measurements* at bus i:

$$P_i = \sum_{j\in i} P_{ij} + \varepsilon_P \tag{18}$$
$$Q_i = \sum_{j\in i} Q_{ij} + \varepsilon_Q. \tag{19}$$

TABLE I
JACOBIAN COMPONENTS FOR THE LINEAR STAGE

	$y = U_i$	$y = U_j$	$y = K_{ij}$	$y = L_{ij}$
$\partial U_i/\partial y$	1	0	0	0
$\partial P_{ij}/\partial y$	$g_{sh.i} + g_{ij}$	0	$-g_{ij}$	$-b_{ij}$
$\partial P_{ji}/\partial y$	0	$g_{sh.j} + g_{ij}$	$-g_{ij}$	b_{ij}
$\partial Q_{ij}/\partial y$	$-(b_{sh.i} + b_{ij})$	0	b_{ij}	$-g_{ij}$
$\partial Q_{ji}/\partial y$	0	$-(b_{sh.j} + b_{ij})$	b_{ij}	g_{ij}

- *Voltage magnitude measurement*. To retain linearity at this stage, squared voltage magnitude measurements should be used:

$$V_i^2 = U_i + \varepsilon_U. \tag{20}$$

The ε terms in the above expressions represent the measurement noise. The standard deviation of the error associated with V_i^2 is related to that of the original measurement V_i through

$$\sigma(V_i^2) = 2V_i\sigma(V_i) \tag{21}$$

which, for practical purposes ($V_i \approx 1$), can be approximated by

$$\sigma(V_i^2) \approx 2\sigma(V_i). \tag{22}$$

Table I provides the constant Jacobian terms corresponding to the above expressions.

Even though the network is fully observable in terms of the conventional state vector, it may happen that certain components of the enlarged intermediate vector adopted in the above formulation are not observable (the observability issue is dealt with in Section V-A). Therefore, the first-level linear state estimator should be restricted to the observable subnetwork.

Let the intermediate vector y be partitioned as follows:

$$y = \begin{bmatrix} y_o \\ y_u \end{bmatrix}$$

where the subscripts "o" and "u" refer to the "observable" and "unobservable" components, respectively. Accordingly, the measurement vector can be partitioned as

$$z = \begin{bmatrix} z_o \\ z_u \end{bmatrix}$$

where z_o contains the set of measurements linearly related to y_o and z_u is the subset that cannot be processed at this preliminary stage and must directly enter the second one (the reader should be aware that the partitions introduced in [12] have a slightly different meaning regarding y_u).

When restricted to the observable components, expressions (16)–(20) can be compactly written as follows:

$$z_o = A_o y_o + \varepsilon_o \tag{23}$$

where A_o is a constant rectangular matrix of appropriate size, the elements of which are given in Table I. In most cases of practical interest, the entire network will be observable at this preliminary phase, in which case $z_o = z$, $y_o = y$, and $A_o = A$.

The WLS estimate $\tilde{y_o}$ is obtained by solving

$$[A_o^T W_o A_o]\tilde{y}_o = A_o^T W_o z_o \tag{24}$$

where the weighting matrix W_o is the inverse of the covariance matrix associated with ε_o.

In turn, the inverse of the covariance matrix associated with $\tilde{y}_o$ is the gain matrix appearing in the above expression:

$$G_o = A_o^T W_o A_o = [\mathrm{cov}(\tilde{y}_o)]^{-1}. \tag{25}$$

Both $\tilde{y}_o$ and G_o are passed to the next stage, along with the raw measurements z_u still not used.

As shown in Table I, the nonzero pattern of the Jacobian, and hence the gain matrix G_o, depends on the network topology and composition of the measurement set. On the other hand, their elements are functions of the parameter values, but not the network state. Therefore, for a given measurement set, G_o should be updated almost exclusively when the topology gets modified (parameter values seldom change if resistance dependence on temperature is ignored, as done customarily, tap changers being a notable exception).

C. Second-Step State Estimator

Once $\tilde{y}_o$ and G_o are computed, the second nonlinear phase of the SE process, consisting of repeatedly solving (11), can be carried out.

In partitioned form, considering the possible existence of certain raw measurements, z_u, discarded during the previous phase, the resulting measurement model at this stage is given by

$$\begin{bmatrix} \tilde{y}_o \\ z_u \end{bmatrix} = \begin{bmatrix} f_o(x) \\ h_u(x) \end{bmatrix} + \begin{bmatrix} \varepsilon_{y_o} \\ \varepsilon_u \end{bmatrix} \tag{26}$$

where the error ε_{y_o} associated with $\tilde{y_o}$ is characterized by the covariance matrix (25).

Then, dropping for simplicity the iteration index and the dependence of the Jacobian components on x, the WLS solution is obtained by repeatedly solving

$$\left\{ \begin{bmatrix} F_o \\ H_u \end{bmatrix}^T \begin{bmatrix} G_o & 0 \\ 0 & W_u \end{bmatrix} \begin{bmatrix} F_o \\ H_u \end{bmatrix} \right\} \Delta x = \begin{bmatrix} F_o \\ H_u \end{bmatrix}^T \begin{bmatrix} G_o & 0 \\ 0 & W_u \end{bmatrix} \left\{ \begin{bmatrix} \tilde{y}_o \\ z_u \end{bmatrix} - \begin{bmatrix} f_o(x) \\ h_u(x) \end{bmatrix} \right\} \tag{27}$$

or, developing the block-wise matrix products:

$$\left[F_o^T G_o F_o + H_u^T W_u H_u\right] \Delta x = F_o^T G_o \left[\tilde{y}_o - f_o(x)\right] + H_u^T W_u \left[z_u - h_u(x)\right] \tag{28}$$

where the terms of the Jacobian F, according to (12)–(14), are given in Table II.

The above expression is to be compared with the conventional approach (2), in which all raw measurements are handled at once. Introducing the same partition, (2) can be rewritten as

$$\left[H_o^T W_o H_o + H_u^T W_u H_u\right] \Delta x = H_o^T W_o \left[z_o - h_o(x)\right] + H_u^T W_u \left[z_u - h_u(x)\right]. \tag{29}$$

TABLE II
JACOBIAN COMPONENTS FOR THE SECOND STAGE

$\partial U_i/\partial V_i=2V_i \quad ; \quad \partial U_i/\partial\theta_i=0$	$\partial U_i/\partial V_j=0 \quad ; \quad \partial U_i/\partial\theta_j=0$
$\partial K_{ij}/\partial V_i=V_j\cos\theta_{ij}$ $\partial K_{ij}/\partial\theta_i=-V_iV_j\sin\theta_{ij}$	$\partial K_{ij}/\partial V_j=V_i\cos\theta_{ij}$ $\partial K_{ij}/\partial\theta_j=V_iV_j\sin\theta_{ij}$
$\partial L_{ij}/\partial V_i=V_j\sin\theta_{ij}$ $\partial L_{ij}/\partial\theta_i=V_iV_j\cos\theta_{ij}$	$\partial L_{ij}/\partial V_j=V_i\sin\theta_{ij}$ $\partial L_{ij}/\partial\theta_j=-V_iV_j\cos\theta_{ij}$

The main difference between (28) and (29) lies in the dimension of both the Jacobian matrix F_o and measurement vector $\tilde{y}_o$ being much smaller than their counterparts H_o and z_o. This is achieved at the cost of previously computing the intermediate vector $\tilde{y}$. Note also that, unlike matrix W_o, the weighting matrix G_o is sparse but nondiagonal, which partially offsets the potential computational saving of the two-stage approach.

It is worth mentioning that the number of injection measurements and null injection constraints handled by (28) reduce to those in z_u, as the K and L components of $\tilde{y}_o$ preserve the structure of power flow measurements. In real life, the measurement redundancy is almost always high enough to guarantee that z_u is the empty set, in which case none of the injection measurements arise in (28).

D. Initialization of the Second-Step State Estimator

The state vector x must be assigned an initial value x_0 to start the iterative process (28). Although the well-known flat voltage profile is always a good choice, the solution $\tilde{y}_o$ of the first-level state estimator can easily provide an alternative enhanced value for x_0. In fact, if the raw measurement vector z lacks bad data, the power flow pattern and bus voltage magnitudes provided by $\tilde{y}_o$ may be sufficiently accurate in practice for several other purposes.

For simplicity of presentation, let us assume that the entire network is observable during the linear preliminary phase. Then, voltage magnitudes can be obviously initialized as

$$V_i^0 = \sqrt{\tilde{U}_i}. \tag{30}$$

Regarding phase angles, each pair of branch variables K and L provide an initial guess for the respective phase angle difference between the terminal buses:

$$\theta_{ij}^0 = \arctan\left[\frac{\tilde{L}_{ij}}{\tilde{K}_{ij}}\right]. \tag{31}$$

Then, by sweeping any spanning tree, bus phase angles θ_i^0 with respect to an arbitrary angle reference can be obtained. However, as the values of K and L above stem from noisy measurements, different trees will lead eventually to slightly different values of θ_i^0.

A more expensive but worth exploring approach consists of obtaining θ_i^0 by solving a linear LS problem. Let θ_b^0 be the vector of phase angle drops across the b network branches, θ_{ij}^0, and θ_n^0 be the vector of $n-1$ bus phase angles with respect to a reference bus. Then, the following equation holds:

$$A_I^T\theta_n^0 = \theta_b^0 \tag{32}$$

where A_I represents the so-called branch-to-node incidence matrix. The LS solution to the above overdetermined system is given by

$$\left[A_IW_bA_I^T\right]\theta_n^0 = A_IW_b\theta_b^0 \tag{33}$$

where the diagonal weighting matrix W_b should ideally take into account the accuracy of the elements in θ_b^0. For the sake of computational effort reduction, it will be assumed for this purpose that the variance of all components θ_{ij}^0 is roughly the same ($W_b \approx I$).

This approach to initialize bus phase angles is justified if at least an iteration is saved during the iterative process (28).

V. OTHER ISSUES

The factorized model poses new issues regarding observability analysis and bad data processing, which will be briefly addressed in this section (a thorough treatment of these topics is not possible within the scope of a single paper).

A. Observability Analysis

Like the SE itself, observability analysis should proceed in two phases.

1) Linear Prefiltering Phase: The objective is to assure that all components of vector y are observable in the linear model:

$$z = Ay.$$

If this is not the case, then the subset z_o should be determined in such a way that A_o in

$$z_o = A_oy_o$$

has full column rank while the size of y_o is as large as possible. For this purpose, a simple numerical procedure like that proposed in [15] can be employed. It is worth remarking that loop equations, considered as equality constraints in [15], do not appear in this formulation, which is the reason why the redundancy is not so high as with the true state vector x.

It should also be noted that, as y includes the branch variables K and L (directly related to phase angle differences across network branches), rather than phase angles with respect to a reference bus, the more involved logic associated with the introduction of a reference bus for every observable island is not needed at this stage. At the end of the process, each set of adjacent observable branches, along with the respective power flow and injection measurements, will constitute an observable island.

2) Second Nonlinear Phase: The objective is to determine whether x is observable for a given set of raw measurements z_u, not yet used, and a set of pseudo-measurements $\tilde{y}_o$. As the components K and L, on the one hand, and U on the other, can be biunivocally associated with pairs of power flows and voltage magnitudes, respectively, the pseudo-measurement vector $\tilde{y}_o$ can be replaced, just for observability analysis purposes, by a fictitious but fully equivalent measurement vector composed of ordinary measurements. This way, any existing observability method (numerical or topological) can be directly applied.

Note that if all elements of y turn out to be observable, then power injection measurements do not reach the second stage, which may simplify to a large extent the observability analysis process.

B. Bad Data and Topology Error Processing

Except for rounding errors, the second-stage SE provides the same solution as that of the conventional approach. Therefore, measurement residuals and associated covariances allow the largest normalized residual test to be implemented at the end of the two-step procedure, much in the same way as it is currently implemented within existing estimators [3].

However, assuming sufficiently high redundancy levels, the proposed factorized approach provides a chance for bad data to be preliminarily detected during the linear prefiltering phase, preventing in this manner the nonlinear phase to be contaminated by the influence of wrong measurements.

VI. Test Results

In this section, test results corresponding to the IEEE 118- and 298-bus systems are presented and discussed. Both the conventional and the proposed SE schemes are coded under the Matlab platform, taking advantage of available sparse matrix capabilities whenever possible. The main objective is to assess the potential benefits of the proposed factorized scheme, for different redundancy and measurement noise levels, in terms of convergence rate, computational effort, and accuracy of the linear prefilter.

A. Convergence Rate

From exact load flow solutions, measurement sets with varying noise and redundancy levels are simulated. For the results to be more representative, 100 measurement sets are randomly generated, for each combination of accuracy and redundancy level considered. Hence, unless otherwise noticed, each value provided henceforth is the average of 100 runs.

Figs. 1 and 2 present, for the 118- and 298-bus networks, respectively, the average number of iterations required by the conventional and the proposed schemes (in the latter case, the linear SE stage is counted as an iteration). Different noise levels, embracing a wide range of situations that could be found in practice, are simulated as follows: for voltage measurements, three values of σ_V are considered, namely, 0.001, 0.002, and 0.01. In turn, for each value of σ_V, three increasing values of s.d. for power measurements are tested ($\sigma_P = \rho\sigma_V$, with $\rho = 1, 2, 5$). The ideal case in which all measurements are exact has been also included for comparison purposes. In all cases, the convergence threshold is $1e-5$.

For each combination of noise values, two measurement sets are considered, as follows:

- low redundancy case (L): voltage magnitudes for all buses, power flows across all branches ("from" terminal only), power injections at one half of buses (one every other bus in the natural sequence);
- high redundancy level (H): voltage magnitudes measured twice for every bus, power flows at both "from" and "to" terminals of every branch, power injections at all buses.

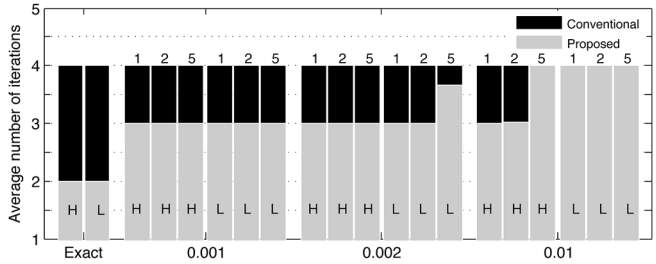


Fig. 1. Average number of iterations for the 118-bus system.

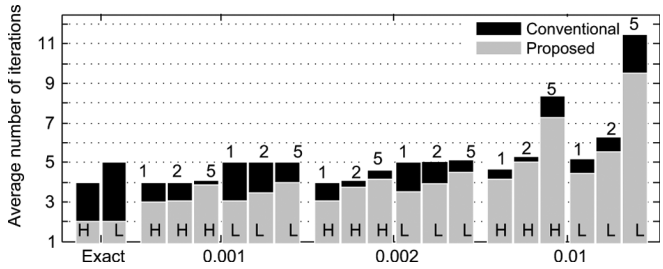


Fig. 2. Average number of iterations for the 298-bus system.

For the 118- and 298-bus systems, this leads to redundancy ratios ranging from 2.58 to 5.16 and 2.37 to 4.74, respectively.

The following conclusions are drawn from those figures.

- The conventional SE is not very much affected by the noise level, at least for the 118-bus system. However, for very noisy measurements, the number of iterations tends to increase, as seen in the rightmost columns of Fig. 2.
- The two-stage SE scheme is more sensitive to the noise level but, for accurate enough measurements, its speed of convergence clearly outperforms that of the conventional approach. This is a consequence of the linear stage providing a very good starting point for the second nonlinear phase. In fact, for the exact measurement (ideal) case, the solution of the linear prefilter is always optimal, and the second iteration accounted for in the leftmost columns is really useless (it is performed just as a means of checking that $\Delta x \approx 0$).
- Both the conventional and to a larger extent the proposed schemes improve their performance when the measurement redundancy is higher (columns with letter "H"). For very noisy and low-redundancy measurement sets, the convergence rate of the proposed scheme is only slightly better than that of the existing one. However, in nowadays transmission and distribution systems, the trend is just the opposite, which favors the adoption of the proposed methodology.

Table III provides the average saving in number of iterations achieved by the factorized scheme for the tested scenarios. Overall, considering only the 1200 cases with $\sigma_V = 0.001$ and $\sigma_V = 0.002$ (central columns in Figs. 1 and 2), the two-stage scheme takes an average of 0.94 and 0.98 iterations less than the conventional SE, for the 118- and 298-bus systems, respectively.

It is well known that the convergence rate of the load flow is significantly affected by the network load level, owing to the model nonlinearities and distance of the flat start to the solution

TABLE III
AVERAGE NUMBER OF ITERATIONS SAVED BY THE FACTORIZED SCHEME

		118		298		118-2P	
σ_V	ρ	H	L	H	L	H	L
0 (exact)	0	2.00	2.00	2.00	3.00	3.00	3.00
0.001	1	1.00	1.00	1.00	1.96	2.00	2.00
	2	1.00	1.00	0.96	1.55	2.00	2.00
	5	1.00	0.99	0.24	1.01	2.00	1.99
0.002	1	1.00	0.99	0.96	1.50	2.00	1.90
	2	1.00	0.99	0.33	1.06	2.00	1.83
	5	1.00	0.33	0.51	0.66	2.00	1.11
0.01	1	1.00	0.00	0.56	0.72	2.00	1.00
	2	0.98	0.00	0.27	0.74	1.71	1.00
	5	0.00	0.00	1.07	1.96	1.00	0.99

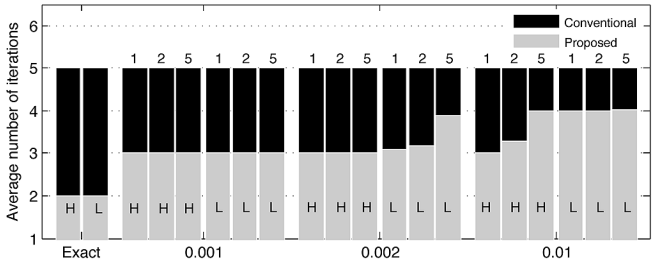


Fig. 3. Average number of iterations for the 118-2P system.

point. Although it has been found that this convergence deterioration affects the SE problem to a lesser extent, a new experiment is designed to assess how both SE methodologies behave under peak loading. For this purpose, the net power at PQ buses of the 118-bus system is doubled. Accordingly, the active power of PV buses is doubled and the reactive power limits are not enforced, allowing the generator voltages to remain at the scheduled value (the average bus voltage, however, reduces from 0.986 to 0.908). Fig. 3 is the counterpart of Fig. 1 for these new scenarios, referred to as 118-2P. A comparison of both figures reveals that the conventional scheme requires an extra iteration in all cases (2000 runs), whereas the two-stage scheme remains virtually unaffected by the network loading condition, which may be good in critical situations, close to voltage collapse. Considering again the 1200 cases with $\sigma_V = 0.001$ and $\sigma_V = 0.002$, the two-stage scheme takes now an average of 1.9 iterations less than the conventional SE (see Table III). This positive behavior can be explained by the linearity of the first stage. In fact, for accurate enough measurements, the estimate provided by the linear prefilter is quite good, irrespective of the load level (see Section VI-C).

B. Computational Efficiency

Table IV shows the average number of kflops[1] for the sets of scenarios with $\sigma_V = 0.002$ and $\sigma_P = 2\sigma_V$, including both redundancy levels. The first stage of the proposed method and the first iteration of the conventional one are accounted for separately. The cost of the initialization procedure for the second stage, as described in Section IV-D, is also shown. In parentheses, the average number of iterations needed by each method is provided, excluding the first stage and the first iteration, respectively.

[1]floating-point operations $\times 10^3$

TABLE IV
NUMBER OF KFLOPS ASSOCIATED WITH DIFFERENT STAGES OF THE SOLUTION PROCESSES FOR $\sigma_V = 0.002$ AND $\sigma_P = 2\sigma_V$

Test	Conventional		Proposed			%
Case	First	Rest (it.)	First	Init.	Sec. (it.)	Sav.
118-L	76	275 (3)	53	13	177 (2)	31
118-H	162	575 (3)	154	13	357 (2)	29
298-L	167	851 (4)	110	30	600 (2.9)	27
298-H	407	1414 (3.1)	311	30	1168 (2.7)	17
118-2P-L	76	366 (4)	53	13	191 (2.2)	42
118-2P-H	162	767 (4)	154	13	357 (2)	44

The following comments are in order regarding the results obtained.

- The average computational saving of the proposed method for the 600 cases collected in Table IV is about 32% (rightmost column of the table).
- The total saving stems from a combination of two factors: 1) the enhanced convergence rate (typically the two-stage scheme requires at least one iteration less); 2) the lower computational effort required by each iteration.
- Both the first stage of the proposed method and the first iteration of the conventional scheme are cheaper, as no trigonometric functions are involved.
- The computational effort reduction originates mainly in the computation of the Jacobian matrix and the residual vector (both entities are obtained simultaneously to avoid duplicated computations).
- The proposed method is more advantageous under peak loading conditions, owing to the extra iteration typically saved.

C. Accuracy and Robustness of the Linear Prefilter

The above results suggest that the linear estimator of the first stage, providing always a solution no matter how noisy measurements are or how loaded the network is, could be by itself a valuable byproduct of the proposed methodology.

In order to assess the accuracy of the solution provided by this step, the averages of the absolute errors associated with the estimates of state variables are computed, as follows:

$$S_V = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{V}_i - V_i^{\mathrm{ex}}\right|$$

$$S_\theta = \frac{1}{N-1}\sum_{i=1}^{N-1}\left|\hat{\theta}_i - \theta_i^{\mathrm{ex}}\right|$$

where V^{ex} and θ^{ex} represent the exact values, only known in simulated scenarios.

For the particular case with $\sigma_V = 0.002$ and $\sigma_P = 2\sigma_V$, these indices are compared in Table V with those arising after the first iteration of the conventional SE solution. As can be seen, the accuracy of the linear estimator is roughly an order of magnitude better than that obtained after the first conventional iteration. The accuracy gain is much higher when the doubly loaded case (118-2P) is considered.

Fig. 4 represents the probability density function (pdf) of the index S_V corresponding to 100 runs of the linear prefilter and the first iteration of the standard estimator, for both 118-bus loading scenarios (high redundancy case). The location of the

TABLE V
AVERAGE VALUES OF ESTIMATION ERRORS AFTER THE FIRST ITERATION (CONVENTIONAL) AND FIRST STAGE (PROPOSED) FOR $\sigma_V = 0.002$ AND $\sigma_P = 2\sigma_V$

Test Case	Conventional		Proposed	
	S_v	S_θ	S_v	S_θ
118-L	0.0040	0.0045	0.0011	0.0007
118-H	0.0058	0.0045	0.0002	0.0005
298-L	0.0108	0.0156	0.0011	0.0013
298-H	0.0105	0.0154	0.0002	0.0007
118-2P-L	0.0160	0.0636	0.0011	0.0010
118-2P-H	0.0214	0.0638	0.0002	0.0006

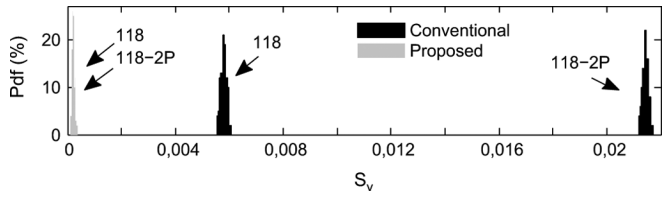


Fig. 4. Histograms of estimation errors for the 118-bus system.

histograms corresponding to the first conventional iteration (in black) clearly confirms the accuracy improvement provided by the linear stage of the factorized estimator (in grey), particularly in the 118-2P case. In this case, both histograms in grey fully overlap near the vertical zero-error axis, which means that the linear prefilter is quite insensitive to the network loading level.

The final test is aimed at checking the capability of the linear preliminary step to detect and identify bad data in noisy measurement environments. For this purpose, a single bad datum is introduced randomly in the high-redundancy measurement set corresponding to the 118-bus system, again for $\sigma_V = 0.002$ and $\sigma_P = 2\sigma_V$. The experiment is repeated 100 times both for voltage and power flow measurements. Two bad data sizes are tested, by adding gross errors with $\sigma_{\rm bd} = 8\sigma$ and $\sigma_{\rm bd} = 7\sigma$.

With $\sigma_{\rm bd} = 8\sigma$, the two hundred (voltage and power) bad data are correctly identified from the results of the first stage. With $\sigma_{\rm bd} = 7\sigma$, the 100 voltage bad data are correctly identified, but only 65 largest normalized residuals corresponding to power flow bad data exceed the customary threshold 3 (if this threshold is slightly reduced to 2.95, then all bad data are identified). This cannot be achieved after the first iteration of the conventional SE, simply because the resulting residuals are still meaningless. It is only after the second iteration (sometimes after the third) when the 200 bad data can be generally identified.

VII. CONCLUSIONS

In this paper, a two-stage state estimation methodology, arising from a factorization of the conventional WLS nonlinear model, is proposed. In the first stage, raw measurements are processed by a linear estimator, for which convergence is never an issue. Then, the resulting estimate is used to initialize the second phase, in which a reduced nonlinear model arises.

Experimental results on IEEE benchmark networks show that the proposed two-stage approach outperforms existing SEs, in terms of both convergence rate and computational cost. The benefits achieved are more noticeable in the presence of highly redundant and accurate measurement systems, particularly under peak loading conditions. For accurate enough measurement sets the results provided by the linear prefilter can be acceptable for many practical purposes, irrespective of the network loading level.

The only potential drawback of the proposed scheme is the risk for the auxiliary variables introduced in the first stage to be unobservable, which is quite unlikely considering the redundancy levels in nowadays bulk transmission systems. In any case, details are provided in the paper about how partially observable cases could be handled.

Future efforts will be directed to develop *ad hoc* observability methods for the linear prefilter. Robustness of this linear estimator against topology errors will be also assessed.

REFERENCES

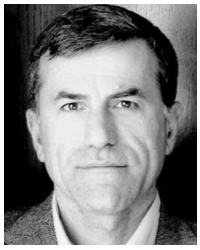
[1] F. C. Schweppe, J. Wildes, and D. B. Rom, "Power system static state estimation, part I, II, III," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 1, pp. 120–135, Jan. 1970.
[2] A. Monticelli, *State Estimation in Electric Power System. A Generalized Approach*. Norwell, MA: Kluwer, 1999.
[3] A. Abur and A. Gómez-Expósito, *Power System State Estimation: Theory and Implementation*. New York: Marcel Dekker, 2004.
[4] W. Tinney and R. Walker, "Direct solutions of sparse network equations by optimally ordered triangular factorization," *Proc. IEEE*, vol. 55, pp. 1801–1809, Nov. 1967.
[5] A. Garcia, A. Monticelli, and P. Abreu, "Fast decoupled state estimation and bad data processing," *IEEE Trans. Power App. Syst.*, vol. PAS-98, no. 5, pp. 1645–1652, 1979.
[6] B. Stott and O. Alsac, "Fast decoupled load flow," *IEEE Trans. Power App. Syst.*, vol. PAS-93, pp. 859–869, 1974.
[7] A. Monticelli and A. Garcia, "Fast decoupled state estimators," *IEEE Trans. Power Syst.*, vol. 5, no. 2, pp. 556–564, May 1990.
[8] I. O. Habiballa and V. H. Quintana, "Modified fast-decoupled state estimation in rectangular-coordinates," in *Proc. 22nd Annu. North Amer. Power Symp.*, Auburn, AL, 1990, pp. 102–111.
[9] W. Lin and J. Teng, "State estimation for distribution systems with zero-injection constraints," in *Proc. Power Industry Computer Application Conf.*, May 1995, pp. 523–529.
[10] A. de la Villa Jaén, P. C. Romero, and A. Gómez-Expósito, "Substation data validation by local three-phase generalized state estimators," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 264–271, Feb. 2005.
[11] W. Jiang, V. Vittal, and G. T. Heydt, "A distributed state estimator utilizing synchronized phasor measurements," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 563–571, May 2007.
[12] A. Gómez-Expósito and A. de la Villa, "Two-level state estimation with local measurement pre-processing," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 676–684, May 2009.
[13] A. Monticelli and A. Garcia, "Reliable bad data processing for real-time state estimation," *IEEE Trans. Power App. Syst.*, vol. PAS-102, pp. 1126–1139, May 1983.
[14] A. Gómez-Expósito and E. Romero, "Reliable load flow technique for radial distribution networks," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 1063–1069, Aug. 1999.
[15] A. Gómez-Expósito and A. Abur, "Generalized observability analysis and measurement classification," *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 1090–1096, Aug. 1998.

Catalina Gómez-Quiles (S'09) received the Eng. degree from the University of Seville, Seville, Spain, in 2006 and the M.Eng. degree from McGill University, Montreal, QC, Canada, in 2008, both in electrical engineering. She is currently pursuing the Ph.D. degree at the University of Seville.

Her research interests include state estimation and risk assessment in competitive electricity markets.

Antonio de la Villa Jaén was born in Riotinto, Spain, in 1960. He received the electrical and doctor engineering degrees from the University of Seville, Seville, Spain.

He is presently an Associate Professor in the Department of Electrical Engineering, University of Seville. His primary areas of interest are computer methods for power system state estimation problems and wave energy converter control.



Antonio Gómez-Expósito (F'05) received the electrical engineering and doctor degrees from the University of Seville, Seville, Spain.

Since 1982, he has been with the Department of Electrical Engineering, University of Seville, where he is currently a Professor and Chairman of the department. He also holds the Endesa Red Chair. His primary areas of interest are optimal power system operation, state estimation, digital signal processing, and control of flexible ac transmission system devices.

State Estimation for Smart Distribution Substations

Catalina Gómez-Quiles, Antonio Gómez-Expósito, *Fellow IEEE* and Antonio de la Villa Jaén

***Abstract*—In the upcoming smart grid environment many more measurements will be available, which can be locally processed by the so-called Substation State Estimator (SSE). Distribution substations serve energy to a large set of feeders, each one delivering power to a certain number of secondary transformers. In this context, the SSE may have to deal with a huge network model, comprising several hundred or even thousand buses.**

Taking advantage of the weak electrical coupling existing among the set of feeders connected to the same or adjacent substations, a two-stage procedure is proposed in this paper to efficiently solve the SSE. In the first stage the overall SE is decomposed into $f+s$ WLS subproblems (f and s being the total number of feeders and substations, respectively), which are then solved in a decoupled manner. The second stage, involving a linear WLS problem, consists of coordinating the solution provided by each subsystem (feeder or substation). The proposed solution scheme has a number of advantages, as shown by the case studies.

***Index Terms*—Substation state estimation, two-stage state estimation, intelligent electronic devices, smart grids.**

I. Introduction

CURRENTLY, the scope of State Estimators (SE) is mostly limited to the transmission level, where each Transmission System Operator (TSO) continuously tracks its own grid from a centralized EMS. Distribution and, to a certain extent, subtransmission networks have been traditionally excluded from this scrutiny, as the required infrastructure is not considered worth investing. In spite of that, the SE notion for distribution systems was considered long ago [1], [2], [3], when a complete measurement set warranting observability was still far in the horizon.

This paradigm will have to change drastically with the advent of the smart grid. On the one hand, new generations of digital devices, such as synchrophasors (PMU) or Intelligent Electronic Devices (IED), intended for measurement, protection and control, being more flexible than existing analog equipment and capable of exchanging data with the external world, are invading virtually every substation [4], [5]. On the other hand, distributed generation (DG), smart meters, and a plethora of distribution automation devices are increasingly providing a lot of additional information at the distribution feeder level.

This will eventually give rise to a more accurate, complex and highly redundant information system, allowing SEs to extend their scope well beyond presently observable areas and also to incorporate advanced functions that have not yet reached the industrial stage.

The authors are with the Department of Electrical Engineering, University of Sevilla, Sevilla, Spain.

This work has been supported by the Spanish Ministry of Science & Innovation (DGI grants ENE2007-62997 and EN2010-18867).

In this context the crucial question is how the existing SE paradigm will have to evolve in order to cope with such a formidable amount of information, originating in a huge number of heterogeneous and geographically distributed sources. Clearly, submitting all this information to a central EMS, or DMS for a distribution utility, would no longer be feasible. Therefore, a hierarchical scheme, processing data whenever possible in a local manner, as close as possible to the place where they are generated, should be implemented. As discussed in [6], [7], [8], the Substation State Estimator (SSE) concept becomes the cornerstone of this multi-level paradigm.

Distribution substations, delivering power to a large number of secondary transformers through a set of medium voltage (MV) feeders, constitute a relevant particular case, in which the size of the resulting SE problem may be discouraging. In fact, a single distribution feeder may comprise hundreds of electrical buses (i.e., MV/LV transformers). Even though the underlying topology of the distribution feeder system is meshed, for the sake of simplicity and cost reduction its operation is customarily radial (only in very dense urban areas, where highly demanding reliability indices are imposed, is the looped operation economically justified). For this reason, unlike in transmission or subtransmission systems, the electrical coupling among the feeder system usually reduces to the head bus, composed typically of one or two busbar sections (see subsection II-C for a more detailed treatment of this subject).

Although nowadays computers can handle at once the interconnected network model, encompassing in this case both the substation and the associated feeders [9], this particular topological arrangement suggests the use of a two-stage decomposition-coordination technique to solve the resulting SE problem. This paper explores such an approach, by which the weak electrical coupling among the involved subsystems is exploited. In the first stage the overall SE problem is decomposed into $f + s$ decoupled problems (f and s being the total number of feeders and substations, respectively) by simply replicating the coupling buses shared by two or more feeders. Then, the SE corresponding to each subsystem is solved in a decoupled manner. The second stage consists of coordinating the solution provided by each decoupled SE. This involves a linear WLS problem, which reduces typically to a trivial model relating state variables with their estimates provided by the first stage.

The two-stage SSE for distribution substations proposed in this work closely resembles the classical multi-area SE problem, oriented to the regional or multi-TSO case [10]. In this application, each feeder constitutes an area and the coordination problem (second stage) is aimed at refining the solution of the border area.

The structure of this paper is as follows: The next section describes succinctly the sources of information in smart dis-

tribution grids, as well as the need and role of SSEs in this context. Section III reviews the conventional WLS estimation model. Sections IV and V present the proposed two-stage methodology, which is illustrated with the help of a tutorial example in Section VI. Section VII shows the performance of the proposed method on a case study, while Section VIII summarizes the main conclusions.

II. The smart grid context

The distinguishing feature of smart grids is the deployment of a diversity of digital devices capable of communicating with each other and/or with a control center. At the transmission level this intelligence concentrates almost entirely in substations, whereas at the distribution level it is partly distributed along radial feeders (transformer centers and switching points). Both sources of information are separately discussed below.

A. Substation-level information

In order to cope with the ever-increasing amount of information gathered at the substation level, new system architectures need to be devised. This goal is achieved by replacing the conventional centralized systems, based on RTUs and numerous protection and control devices, with local area network-based systems and advanced multifunctional protection and control IEDs.

The global communication standard for Substation Automation System (IEC 61850) defines the communication rules between IEDs and specifies also other system requirements (Figure 1), like message performance and information security in substation automation [11], [12].

In this smart substation environment new flows of information among different sources can be established, such as:

- Collecting real-time data from protection relays. This information, which cannot obviously be sent to the DMS, can be locally used to increase the redundancy in order to monitor the operating state of a substation.
- Gathering information from all feeders connected to the substation.
- Submitting locally filtered information to the DMS and neighboring substations.
- Exchanging information with key agents connected to the substation (DG, microgrids, data concentrators and active demand).

B. Feeder-level information

Regarding the information that can be found at the distribution feeder level, very few measuring devices have been installed until recently to monitor the operating condition of MV feeders. Typically, the current at each feeder head, along with the voltage magnitude at the MV bus, are telemetered and gathered at dedicated DMS, but no real-time information is obtained of what happens downstream, unless a fault occurs. This situation is quickly changing for several reasons, including:

- Energy trading of DGs and new agents connected at this level (EV chargers, energy storage devices) must be properly monitored.

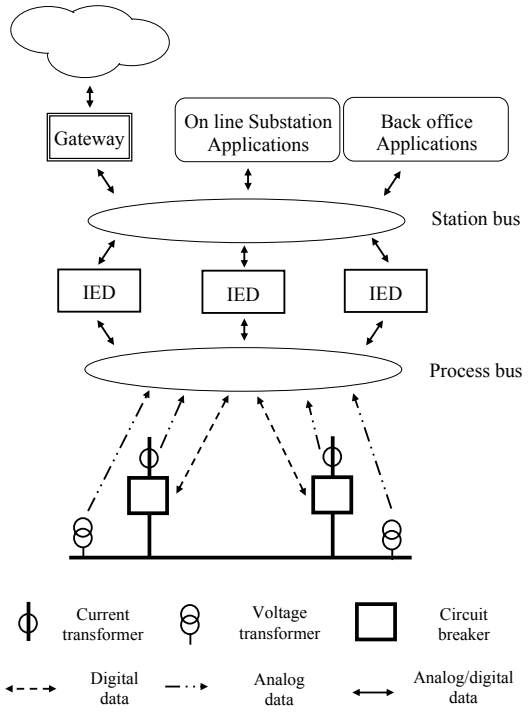


Fig. 1. IEC 61850 substation.

- Smart meters provide customer information which is concentrated at the secondary transformer centers before being submitted upstream to the distribution substation or DMS.
- Fault current detectors, along with proper switching devices (reclosers and sectionalizers), are being installed at strategically selected points to speed up the service restoration process. With a moderate extra investment those devices could provide as well Ampere measurements during normal operation.
- Higher levels of distribution automation allow more flexibility in the way distribution systems are operated, such as feeder reconfiguration and optimization of var resources for congestion or voltage violation relief, loss reduction and reliability enhancement.
- Aggregated consumers and DGs follow predictable patterns of power demand and injection, respectively, which can be helpful in combination with data mining techniques.

C. Substation and feeder topologies

In order to fulfill the N-1 reliability criteria, a distribution substation is customarily fed by at least two subtransmission lines. For the same reason, the substation contains several HV/MV transformers, which in turn feed a single or two busbar sections where the set of radial feeders are connected to. Normally open switches are located at strategically selected points so that neighboring feeders can help each other in case the main feeding route is interrupted for some reason (maintenance or isolation of faulty area). Reliability can be further enhanced if normally open switches are located between feeders originating at different substations.

In certain cases, where very exigent reliability standards are of application, the purely radial topology is not satisfactory and simple loop arrangements are resorted to [13], [14]. This can be achieved in a straightforward manner by closing the normally open tie switches between neighboring feeders, which calls for the deployment of more sophisticated automation and protection systems to cope with the resulting short-circuit levels and bidirectional power flows. In those cases, in addition to the head busbars, the pair of feeders forming a loop share also the remote bus where the circuit is closed. Although looped distribution feeders will become more common in the smart grid context, the total number of coupling buses shared by the set of feeders will remain very low in relative terms.

D. Substation state estimator

The technological developments discussed above allow to envision a future in which SEs will spread from MV distribution feeders to the TSO level [6], [15].

The SSE should deal with the information collected within a substation or small set of adjacent substations. A great majority of raw measurements will have to be processed at this distributed level, where a modest but sufficient computing power already exists. The SSE is useful not only to pre-filter and reduce the size of the measurement set that should be sent to the TSO estimator, but in many instances also for early detection of model and network inconsistencies at the substation level. For the redundancy ratios expected in next-generation substations, most topology errors and bad data could be handled at this level [7].

III. Conventional WLS state estimation

Consider the following measurement equation [16]:

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

where:

- x is the state vector composed of $n = 2N-1$ variables, N being the number of buses.
- z is the measurement vector of size $m > n$,
- h is the vector of usually nonlinear functions, relating error free measurements to the state variables,
- e is the vector of measurement errors, assumed to have a Normal distribution with zero mean and known covariance matrix R. When errors are independent R is a diagonal matrix with values σ_i^2, where σ_i is the standard deviation (s.d.) of the error associated with measurement i.

When the conventional bus-branch model is adopted the state variables are voltage magnitudes and phase angles, whereas the measurement vector comprises power injections, branch power flows and voltage magnitudes. At the subtransmission and distribution levels, Ampere measurements can also play an important role.

Under these assumptions the maximum likelihood estimate is provided by the WLS estimator, which minimizes the weighted squares of measurement residuals given by:

$$J = \sum_{i=1}^{m} W_i r_i^2$$

where:

$r_i = z_i - h_i(x)$ is the measurement residual,
W_i is the respective weighting coefficient.

The estimate $\hat{x}$ can be obtained by iteratively solving the Normal equations:

$$G^{(k)} \Delta x^{(k)} = \left[H^{(k)}\right]^T W \left[z - h\left(x^{(k)}\right)\right] \tag{2}$$

where:

$H^{(k)} = \partial h / \partial x$ is the Jacobian of h evaluated at $x^{(k)}$,
$G^{(k)} = H^T W H$ is the gain matrix,
$W = R^{-1} = \text{diag}(W_i)$ is the weighting matrix,
$\Delta x^{(k)} = x^{(k+1)} - x^{(k)}$, k being the iteration counter.

Iterations are terminated when all components of $|\Delta x_k|$ are within a specified tolerance. The covariance of the estimate is:

$$\text{cov}(\hat{x}) = \left[G^{(k)}\right]^{-1}$$

which is the basis for the computation of related covariance matrices, such as $\text{cov}(\hat{r})$. Then, the bad data processing function, aimed at detecting, identifying and eliminating bad analog measurements, is activated. This is accomplished through the so-called largest normalized residual test [17].

The above development assumes that a single global reference can be used for all phase angles, which is the case when the entire system constitutes a single and fully observable topological island. In the presence of several observable or disconnected islands a local reference should be chosen for each one, accordingly reducing the number of state variables.

IV. Two-stage estimation based on feeder-level decomposition

Let us consider a (normally large) set of distribution feeders along with the associated substation(s). As shown schematically in Figure 2, the electrical coupling between any pair of feeders, no matter how many buses they contain, reduces to the head busbar(s) and, if the tie switch is closed, another common bus downstream.

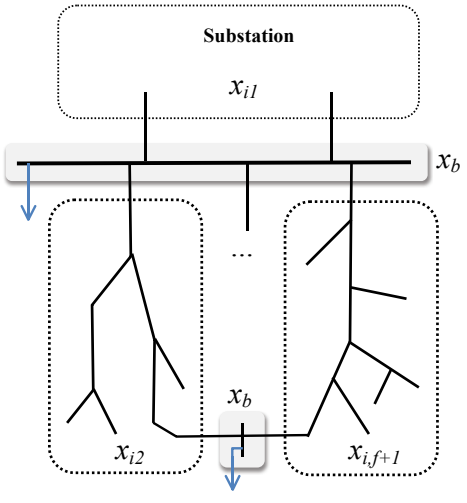


Fig. 2. Subsystems connected to a common bus in a distribution substation.

In this context, $f+s$ weakly coupled subsystems can be identified, where f is the number of feeders and s the number of substations interconnected by the feeders (in radially operated systems $s=1$, since each substation can be individually considered).

This weak coupling suggests the use of a two-stage decomposition-coordination technique. Let x_b denote the border state variables, associated with the set of buses which are shared by two or more subsystems. The remaining state variables, as illustrated in Figure 2, constitute the internal variables, x_{ij}, for each subsystem $j = 1, 2, \ldots, f+s$. Therefore, the state vector has the following structure:

$$x = \{x_i | x_b\} = \{x_{i1}, x_{i2}, \ldots, x_{i,f+s} | x_b\}$$

Accordingly, the measurement model (1) can be rearranged into the following $f+s$ subsystems, coupled only by the set of variables in x_b:

$$\begin{aligned} z_1 &= h_1(x_{i1}, x_b) + e_1 \\ z_2 &= h_2(x_{i2}, x_b) + e_2 \\ &\vdots \\ z_{f+s} &= h_{f+s}(x_{i,f+s}, x_b) + e_{f+s} \end{aligned} \tag{3}$$

Very frequently, the distribution network to be analyzed reduces to a radial system fed from a single substation busbar, in which case x_b contains just the voltage magnitude of that bus, provided its phase angle is taken as the global reference.

A. Model decomposition

The weakly coupled model (3) can be easily transformed into $f+s$ fully decoupled subsystems, simply by replicating each component of x_b as many times as subsystems to which it belongs. In addition, the phase angles within each decoupled area must be referred to a local phase reference, except for the area containing the global reference. These two manipulations can be algebraically represented by means of linear transforms, that will be separately discussed for the sake of clarity.

1) *State vector augmentation:* An augmented state vector, x^a, is defined by properly replicating the border subvector x_b:

$$x^a = \{x_i | x_b^a\} = \{x_{i1}, x_{i2}, \ldots, x_{i,f+s} | x_{b1}, x_{b2}, \ldots, x_{b,f+s}\}$$

In matrix form, the original and augmented vectors are then related as follows:

$$x^a = \begin{bmatrix} x_i \\ x_b^a \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & K \end{bmatrix} \begin{bmatrix} x_i \\ x_b \end{bmatrix} \tag{4}$$

where each column of matrix K has as many ones as number of times the respective element of x_b has been 'cloned'. Note that the augmentation applies to both components of the state vector (voltage magnitudes and phase angles).

1) *Phase angle rotations:* In the coupled model (3) all phase angles are referred to a single global reference, which is set arbitrarily to zero. However, when dealing with the augmented state vector introduced above, a local reference has to be chosen for each subsystem. In order to simplify the algebraic manipulations arising in Section V, local references are always chosen among border buses.

For the sake of clarity, let θ_j and δ_j denote the vectors of phase angles for any subsystem j in x^a, referred to the global and local references respectively. Then, the phase angle rotation can be expressed in vector form as follows:

$$\delta_j = \theta_j - e_1 \theta_{jr} \tag{5}$$

where e_1 is a vector of all ones and θ_{jr} is the phase angle of the border bus taken arbitrarily as local reference for subsystem j. Obviously, $\delta_{jr} = 0$ by definition, and hence it is removed from x^a (the rotation implies a reduction in the number of unknowns). In addition, for the subsystem containing the global reference (for instance one of the substations), labelled for convenience the first ($j=1$), the global and local references are the same and no rotation is needed ($\theta_{1r} = 0$ and $\delta_1 = \theta_1$).

In matrix form, considering that voltage magnitudes are not affected by the rotation, the following linear expression can be written for each area $j \neq 1$:

$$\underbrace{\begin{bmatrix} V_j \\ \delta_j \end{bmatrix}}_{y_j} = \begin{bmatrix} I & 0 \\ 0 & I - E_{1j} \end{bmatrix} \underbrace{\begin{bmatrix} V_j \\ \theta_j \end{bmatrix}}_{x_j^a} \tag{6}$$

where E_{1j} is a null matrix except for the column corresponding to the reference angle θ_{jr}, which is equal to e_1. Note that the null rows corresponding to local phase references ($\delta_{jr} = 0$) and the column corresponding to the global reference ($\theta_{1r} = 0$) should be removed from the above linear model.

By combining (4) and (6) a linear relationship arises between the augmented and rotated state vector, y, and the original state vector x:

$$y = Cx \tag{7}$$

which, in partitioned form, becomes,

$$\begin{bmatrix} y_i \\ y_b \end{bmatrix} = \begin{bmatrix} I & B_i \\ 0 & B_b \end{bmatrix} \begin{bmatrix} x_i \\ x_b \end{bmatrix} \tag{8}$$

It is important to realize that the particular simplified structure of the above system obeys the fact that local phase references are selected among border buses, which is always possible. Moreover, it is easy to show that, when there is at least a border bus which is shared by all areas, and this bus is both the global and local phase reference, then $B_i = 0$. This is the common case of a distribution substation feeding a set of feeders (radial or looped) from a single busbar.

The augmented and rotated vector y, composed of $f+s$ disjoint components, allows the weakly coupled system (3) to be fully decoupled, as follows:

$$\begin{aligned} z_1 &= f_1(y_{i1}, y_{b1}) + e_1 \\ z_2 &= f_2(y_{i2}, y_{b2}) + e_2 \\ &\vdots \\ z_{f+s} &= f_{f+s}(y_{i,f+s}, y_{b,f+s}) + e_{f+s} \end{aligned} \tag{9}$$

where the original functions $h_j(x)$ have been renamed as $f_j(y)$ to emphasize the qualitative difference arising from the change of variables (i.e. to distinguish a decoupled system from the original one). This is however a theoretical formalism, as both

functions are structurally identical. In compact form, (9) can be written as

$$z = f(y) + e \tag{10}$$

which is the counterpart of (1).

B. Decomposed solution stage

The estimates $\tilde{y}_j$ of the decoupled WLS models developed above are obtained by repeatedly solving the associated Normal equations, for $j = 1, 2, \ldots, f + s$:

$$G_{Fj}^{(k)} \Delta y_j^{(k)} = \left[F_j^{(k)}\right]^T W_j \left[z_j - f_j\left(y_j^{(k)}\right)\right] \tag{11}$$

where:

$F_j^{(k)} = \partial f_j / \partial y_j$ is the Jacobian of f_j,
$G_{Fj}^{(k)} = F_j^T W_j F_j$ is the gain matrix of subsystem j,
$\Delta y_j^{(k)} = y_j^{(k+1)} - y_j^{(k)}$, k being the iteration counter.

Accordingly, the covariance of each subvector $\tilde{y}_j$ is given by:

$$\mathrm{cov}(\tilde{y_j}) = \tilde{G}_{Fj}^{-1} \tag{12}$$

Measurements provided by PMUs, if any, can be directly used at this early stage, much like any other measurement (with their associated uncertainty modeled by the respective weighting coefficient). The only worth noting difference is the presence of phase angle measurements, which should be taken into account when dealing with phase angle references. This issue is fully addressed in [18].

C. Coordinating stage

The estimates of border variables obtained by independently solving the $f + s$ subsystems (11), will be slightly different due to the measurement noise. On the other hand, the resulting phase angles refer to a local rather than global reference. Therefore, a coordination phase is needed aimed at: 1) reconciling the decoupled estimates of border variables; 2) estimating the phase reference of each subsystem (θ_{jr}, $j > 1$); and 3) eventually refining the estimates of the internal variables.

Taking into account the uncertainty of the estimates $\tilde{y}$, the overdetermined linear system (8) can be written as follows:

$$\begin{bmatrix} \tilde{y}_i \\ \tilde{y}_b \end{bmatrix} = \begin{bmatrix} I & B_i \\ & B_b \end{bmatrix} \begin{bmatrix} x_i \\ x_b \end{bmatrix} + e_y \Rightarrow \tilde{y} = Cx + e_y \tag{13}$$

where the error terms of the pseudomeasurements $\tilde{y}$ are characterized by the covariance matrix (12).

The WLS solution $\hat{x}$ satisfying (13) is obtained from

$$\left[C^T W_y C\right] \hat{x} = C^T W_y \tilde{y} \tag{14}$$

where the weighting matrix W_y is given by

$$W_y = \mathrm{cov}^{-1}(\tilde{y}) = \tilde{G}_F \tag{15}$$

It is worth noting that the weighting matrix W_y is in this case a non-diagonal but sparse matrix composed of $f + s$ major blocks, $\tilde{G}_{Fj}$, each one representing the gain matrix of each decoupled subsystem (11).

In summary, the original WLS nonlinear coupled problem is transformed into two simpler problems, namely:

1) Obtain $\tilde{y}_j$, for $j = 1, 2, \ldots, f + s$, by repeatedly solving (11). These are the WLS estimates for the set of nonlinear decoupled subsystems (9). As a byproduct, the associated gain matrices $\tilde{G}_{Fj}$ are also obtained. It is worth stressing that all raw measurements should be used at this stage, except for border injections (these are dealt with in Section IV-E).
2) Obtain $\hat{x}$ by solving the WLS linear problem (14), using as pseudomeasurements and weighting matrices the estimates and gain matrices of the previous step, respectively.

At the end of the process the covariance of $\hat{x}$ can be computed from the respective gain matrix,

$$\mathrm{cov}(\hat{x}) = \left[C^T W_y C\right]^{-1}, \tag{16}$$

allowing the bad data identification process to be implemented through the largest normalized residual test, in the same way as in the conventional WLS solution. However, depending on the existing redundancy, bad data can also be detected in a majority of cases at the end of the first stage.

D. Theoretical justification

This two-stage approach can be considered a particular case of the general factorized scheme recently presented in [15]. From this point of view, the original nonlinear problem (1) is unfolded into the two simpler problems (10)-(13), repeated here for convenience:

$$z = f(y) + e \tag{17}$$

$$y = Cx + e_y \tag{18}$$

where the intermediate state vector y is chosen in such a way that (17) is geographically decomposed and (18) becomes linear. By the chain rule, after direct substitution, it is easy to conclude that:

$$h(x) = f(Cx) \quad \Rightarrow \quad H = FC \tag{19}$$

which proves the full equivalence between the resulting gain matrices:

$$H^T W H = C^T (F^T W F) C = C^T G_F C = C^T W_y C$$

As discussed in [15], the estimate $\hat{x}$ is not, rigorously speaking, the optimal one. This is due to the fact that the gain matrix $\tilde{G}_F$, playing also the role of weighting matrix W_y in the final linear phase, does not refer exactly to the same linearization point as that of $G^{(k)}$, k being the last iteration of the standard solution process (2). Nevertheless, as shown in Section VII, the resulting accuracy is satisfactory for practical purposes.

E. Solution refinement

As shown in Figure 2, there may be power injections at border buses, representing non observable feeder sections, which have been fully ignored until now. If they are not measured, which is very likely, then the solution provided by the two-stage process described so far is the correct one, and no extra steps are needed. However, if those injections are

measured, or they are exactly null, they should be somehow taken into account for the estimate to be the same as that provided by the conventional system-wide SE.

It is clear, for the restricted overlapping level considered in this work (no tie-lines are shared between any two subsystems), that border injections cannot be handled during the first decoupled stage, simply because the state variables of neighboring areas are missing [10]. On the other hand, they could be easily included in the coordination phase, but then the resulting model would become partly nonlinear, which is an undesirable feature.

In order to retain the simplicity and advantages of the proposed two-stage scheme (see the next section), the solution adopted in this work consists of refining or compensating the initial estimate, $\hat{x}$, to take into account the information provided by border injections. As shown in the Appendix, the state vector correction can be computed from

$$\left[C^T W_y C\right] \Delta x_c = -P^T \lambda \tag{20}$$

where P is the Jacobian of the border injections and λ is an auxiliary vector obtained by solving the small system (27) or (30), depending on whether null injections or regular measurements are being considered. Note that, as the table of factors of the coefficient matrix above is available from the previous solution of (14), computing Δx_c reduces to performing just a low-cost forward/backward elimination process. Furthermore, as the right-hand side vector is very sparse in this case, the so-called fast forward elimination can be applied to speed up the computations [19].

Then, the compensated solution is given by:

$$\hat{x}_c = \Delta x_c + \hat{x}$$

V. Practical implementation

This section discusses relevant implementation issues regarding the second linear stage. The first stage, involving the decoupled solution of $f+s$ conventional WLS problems, each one related to a single feeder or substation, requires no additional explanations (standard solution techniques, based on the sparse Cholesky or QR factorization should be applied to the Normal equations).

The idea is to exploit the special blocked structure of the equations arising during the solution of the second linear stage. While the Jacobian matrix C in (14) is composed of trivial submatrices, the gain matrix G_F (weighting matrix W_y) is a non-diagonal but sparse matrix composed of $f+s$ major blocks, each one representing the gain matrix of the decoupled subsystems (11). Note that any block G_{Fj} is composed of four submatrices (self and mutual terms between interior an border variables). Reordering the rows/columns so that the border and interior variables are gathered together, leads to the following blocked structure,

$$G_F = \begin{bmatrix} G_{ii} & G_{ib} \\ G_{ib}^T & G_{bb} \end{bmatrix} = \left[\begin{array}{ccc|ccc} \blacktriangle & & & \diamond & & \\ & \blacktriangle & & & \diamond & \\ & & \ddots & & & \ddots \\ \hline & & & \blacktriangle & & \diamond \\ \diamond & & & & \blacktriangledown & \\ & \diamond & & & & \blacktriangledown \\ & & \ddots & & & \ddots \\ & & \diamond & & & \blacktriangledown \end{array}\right] \tag{21}$$

Every major block above is composed of $f+s$ decoupled blocks. For the application considered in this work the block sizes in G_{ii} (top left) are significantly much larger than those of G_{bb} (bottom right), which is an important feature to be computationally exploited.

Based on the above partitioning, the Normal equations (14) can be rewritten as,

$$\begin{bmatrix} I & \\ B_i^T & B_b^T \end{bmatrix} \begin{bmatrix} G_{ii} & G_{ib} \\ G_{ib}^T & G_{bb} \end{bmatrix} \begin{bmatrix} I & B_i \\ & B_b \end{bmatrix} \begin{bmatrix} \hat{x}_i \\ \hat{x}_b \end{bmatrix} = \begin{bmatrix} I & \\ B_i^T & B_b^T \end{bmatrix} \begin{bmatrix} G_{ii} & G_{ib} \\ G_{ib}^T & G_{bb} \end{bmatrix} \begin{bmatrix} \tilde{y}_i \\ \tilde{y}_b \end{bmatrix} \tag{22}$$

Solving the above linear system from scratch would be wasteful for two main reasons: 1) no advantage is taken of the blocked structure of the coefficient matrix; 2) no benefits are obtained from the existing factorization of each block G_{Fj}, performed during the first nonlinear stage.

As discussed in [15], there are two alternatives to efficiently solve the system (22), taking into account the specific structure of the problem. In this paper, considering the very small size of the subvector $\hat{x}_b$, the direct method based on the blocked elimination of the variables $\hat{x}_i$ is applied. This elimination leads to the following reduced system,

$$(B_b^T G_{\text{sch}} B_b)\hat{x}_b = B_b^T G_{\text{sch}} \tilde{y}_b \tag{23}$$

where the resulting coefficient matrix

$$G_{\text{sch}} = G_{bb} - G_{ib}^T G_{ii}^{-1} G_{ib}$$

is known as the Schur's reduction of G_F.

When computing G_{sch}, advantage should be taken of the block diagonal structure of the G_F components, as shown by (21). If the proper computing environment exists, this block-based computation could be fully distributed among $f+s$ processors, much in the same way as the decoupled solution of the $f+s$ subsystems (11). Note that the inverse of G_{ii} is never computed, but the available sparse factors of its $f+s$ components should be rather used. In fact, computing G_{sch} reduces basically to repeated solutions of the sparse linear systems arising during the last iteration of (11).

Once $\hat{x}_b$ is computed, the remaining state variables $\hat{x}_i$ can be obtained by resorting to the eliminated subsystem of (22), as follows:

$$G_{ii}\Delta x_i = G_{ib}(\tilde{y}_b - B_b \hat{x}_b) \tag{24}$$

$$\hat{x}_i = \Delta x_i + \tilde{y}_i - B_i \hat{x}_b \tag{25}$$

Note that the above equation is actually composed of $f+s$ decoupled subsystems.

Advantages of the proposed solution approach are:

- Potential reduction of the computational effort (simpler or constant Jacobian components, smaller size of the state vector during the iterative process, etc.).
- Early bad data processing capability within each subsystem.
- Possibility of tailoring the solution technique in accordance with the peculiarities of each subsystem (ill-conditioning typically arises when solving distribution feeders).
- No need to adopt the same modeling details for each subsystem (e.g. three-phase feeder model versus single-phase substation model).
- Less risk of convergence problems, as a problematic feeder will be identified during the first stage and replaced by equivalent injections.

VI. Tutorial example

The proposed two-stage decomposition-based methodology to solve the SSE problem will be illustrated with the help of the tutorial example shown in Figure 3.

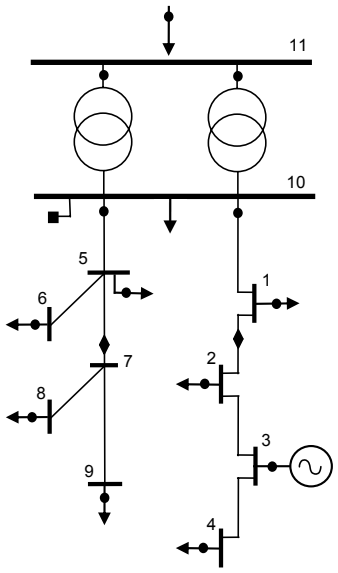


Fig. 3. Sample distribution substation and associate feeders.

This small example is representative of a diversity of situations that can be found at this level:

- Long rural feeders, with many single- or three-phase laterals and relatively smaller loads (left), may coexist with urban, densely loaded feeders, usually lacking laterals (right).
- Very short feeder sections have to be modeled sometimes.
- A majority of measurements or pseudomeasurements associated with the feeders are injections. Flow measurements, if any, take usually the form of Ampere measurements. In any case, the redundancy is rather small.
- Power injected by DG is more accurately known than load delivered by secondary distribution transformers, usually provided by load forecasting tools. Some nodes are zero-injection buses (e.g. bus 7).
- Typically, the measurement set within the substation is expected to be more accurate than feeder measurements.

As discussed in [16], Chapter 3, all those conditions (heterogeneous weights, short branches, many injection measurements, etc.) significantly contribute to the ill-conditioning of the resulting equation system. Indeed, it is a well-known fact that distribution networks tend to be more troublesome regarding iterative solution schemes, such as load flow [20] and state estimation [21].

In this case, the border area allowing the overall SE problem to be decomposed into 3 decoupled problems (two feeders plus the substation) is simply the common bus, which is therefore triplicated (see Figure 4). If bus 10 is chosen as phase reference, the set of border variables reduces in this case to V_{10} for all subproblems.

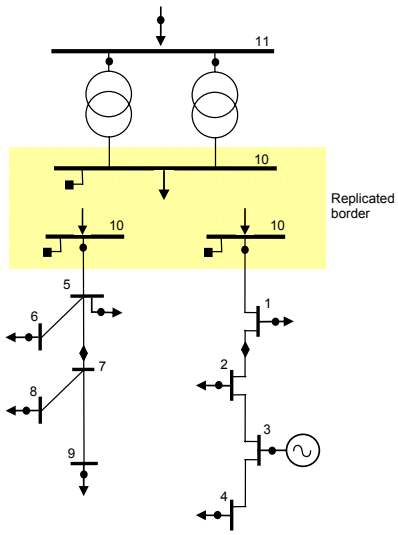


Fig. 4. Resulting subsystems for the distribution substation.

The following remarks are worth noting for the sake of generalization of this tutorial example:

- All measurements of the initial problem are used at least once in the decomposed problem. This guarantees that the coordination phase will involve a linear model.
- The voltage magnitude measurement of the common bus (V_{10}^m) is triplicated, which is good to retain as much redundancy as possible within each subsystem. Consequently, the weight of each and every replicated measurement should be divided by three so that the decomposed problem is statistically equivalent to the original one.

The first SE stage consists of solving three nonlinear decoupled systems by means of the iterative scheme (11). A major advantage of the resulting decoupling is that different solution techniques could be applied if needed to each subsystem. For instance, while the cheaper Cholesky-based solution of the Normal equations can be adopted for the substation subsystem, some ill-conditioned feeders might require the use of more expensive QR factorization, in order to overcome potential numerical problems.

On the other hand, the coordination stage reduces to a linear system, in which internal and border variables become coupled by the off-diagonal blocks of the gain matrices. In this simple example, the subsystem constraining the border variable reduces to,

$$\begin{bmatrix} \tilde{V}_{10}^{(f1)} \\ \tilde{V}_{10}^{(f2)} \\ \tilde{V}_{10}^{(sub)} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} V_{10} + e_b$$

VII. Case study

A realistic case study is used to assess the two-stage SSE procedure. The test system is composed of a substation with two parallel transformers, like in the tutorial example. The MV bus of this substation feeds a set of 4 radial feeders, comprising two replicas of the 69- and 85-bus benchmark systems, which are well-known as ill-conditioned cases in the distribution load flow literature [20]. The measurement set is as follows:

- Substation: The two bus voltage magnitudes plus the transformers power flows on the high voltage side are measured, yielding a total of 6 measurements.
- Feeders: P & Q injections at all buses. No voltage magnitude measurements are assumed. A modest redundancy level is achieved by means of 5 Ampere measurements in two feeders and 10 Ampere measurements in the remaining two, labelled as 69L-85L and 69H-85H (low and high redundancy, respectively). For each feeder, one of the Ampere measurements is assigned to the first section (incident to the substation busbar), while the remaining 26 are randomly located downstream. Note that, as recommended in [16], Chapter 9, squared current magnitudes I^2 are adopted to prevent numerical problems for very short or unloaded branch sections.

It is assumed that the border (common substation busbar) injections are not measured. The overall system redundancy is therefore 1.054, ranging from 1.024 for the 85L feeder to 1.5 for the substation. These are much lower values than those typically found in transmission systems, but they are sufficient to implement a workable SE function.

A. Gaussian errors

Simulated measurements are obtained by adding Gaussian noise to exact values provided by the load flow solution. For the results to be statistically significant one thousand measurement sets are randomly generated for each scenario. Two error levels are separately considered for the whole set of measurements, namely $\sigma = 0.01$ and $\sigma = 0.025$.

Both the standard WLS methodology, repeatedly solving the Normal equations (2), and the proposed two-stage procedure are applied to each scenario, with a convergence threshold of $1E-4$.

In order to compare the resulting accuracies, the averages of the absolute errors associated with the estimates of state variables are computed, as follows:

$$S_V = \frac{1}{N}\sum_{i=1}^{N} |\hat{V}_i - V_i^{\text{ex}}| \quad ; \quad S_\theta = \frac{1}{N-1}\sum_{i=1}^{N-1} |\hat{\theta}_i - \theta_i^{\text{ex}}|$$

where V^{ex} and θ^{ex} represent the exact values.

Figures 5 and 6 present the probability density functions (pdf) of the one thousand S_V and S_θ values, respectively, corresponding to the conventional method (gray continuous line) and the proposed two-stage estimator (dashed black line). As can be seen, both curves almost perfectly match, indicating that in practice the performance of the proposed method is comparable to that of the conventional SE, in spite of the low redundancy assumed in this case study.

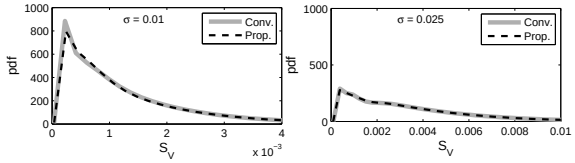


Fig. 5. Pdf of the estimation errors (absolute value) associated with voltage magnitudes.

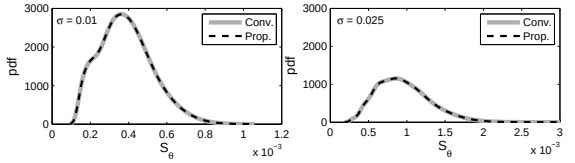


Fig. 6. Pdf of the estimation errors (absolute value) associated with phase angles.

Table I collects the average number of iterations required by the conventional method and the nonlinear decoupled stage of the proposed scheme. In the latter case, each subsystem (substation and feeders) is separately shown. In average, the convergence of the decoupled subsystems is better than that of the entire system. Note that the substation, not so ill-conditioned as the feeders, involves typically one iteration less to converge. The last row refers to a modified case in which the active power of all buses in the 69L feeder is multiplied by 3.2 (for the scenarios with $\sigma = 0.01$). This significantly increases the number of iterations required by the standard method and, to a lesser extent, the affected feeder. However, the substation and the remaining feeders are obviously unaffected, which is a clear advantage of the proposed methodology.

TABLE I
Number of iterations

	Conv.	Subs.	69H	69L	85H	85L
Exact	5	4	5	5	5	5
$\sigma = 0.01$	5.64	3.90	5.00	5.00	4.99	5.00
$\sigma = 0.025$	5.59	4.05	5.00	5.00	4.98	5.01
$3.2P(69L)$	10.02	3.90	5.00	7.00	4.99	5.00

B. Bad data

Single bad data identification in the substation area can be safely implemented based on the largest normalized residual ($|r_N|$) test [16]. In this subsystem, both the conventional and the decoupled solution of the two-stage approach, provide

essentially the same residual values. Therefore, bad data could be removed before running the coordination phase.

However, for the low redundancy levels arising in the feeders, the normalized residual test is not always reliable, no matter which method is employed for SE. This is illustrated by means of 30 different scenarios, each one containing a single Ampere measurement contaminated with a large enough error to be detected (the remaining measurements are exact). The following conclusions are reached, applying to both the conventional solution and the first decoupled stage of the proposed decomposed solution:

- For the 26 Ampere measurements which are not incident to the substation busbar, the largest $|r_N|$ clearly exceeds the second one in the ranking, which means that if bad data are detected they can be safely identified.
- The remaining four measurements, each located at the first section of the respective feeder, constitute critical k-tuples, with k varying from 15 to 71 (this size depends on the location of the remaining Ampere measurements downstream). Therefore, k barely distinguishable normalized residuals are candidates for bad data, which can be detected but not identified. In such cases, if the largest $|r_N|$ is automatically sought, one can reach the wrong conclusion, since in floating-point arithmetic strict equalities never hold owing to numerical round-off errors (the difference between any two $|r_N|$ values in the critical k-tuple is related to the convergence threshold). For instance, the conventional SE misidentifies two out of the four bad data, whereas the decoupled solution stage fails to correctly target any of them, but this is really meaningless in this case. Note that adding border injection measurements and/or forming loops by closing tie switches will alleviate this problem to a certain extent when the conventional SE is applied (in the two-stage procedure this will be noticed only after the second coordination phase).

Figure 7 represents the 50 largest $|r_N|$ for two cases, corresponding to the situations just discussed. In the identifiable case, a single $|r_N|$ value stands out, properly flagging the actual bad datum. In the other case, however, there are 15 candidates for bad data, and there is no way to ascertain which one is the real one.

In the presence of noisy measurements and/or multiple bad data, the identification process will be even more tricky. Therefore, higher redundancy levels are strongly recommended before the bad data detection function is activated. Such levels are not currently available in distribution feeders, though.

VIII. Conclusion

This paper focuses on the SSE tool, for the particular configuration arising in distribution substations directly feeding MV radial feeders. The driving idea is to exploit the weak electrical coupling among the involved subsystems, for which a decomposition-coordination technique is adopted based on a straightforward two-stage procedure. In the first stage $f+s$ nonlinear decoupled problems are solved while the coordination phase reduces to a WLS linear problem (an additional solution refinement may be needed in the presence of border injections).

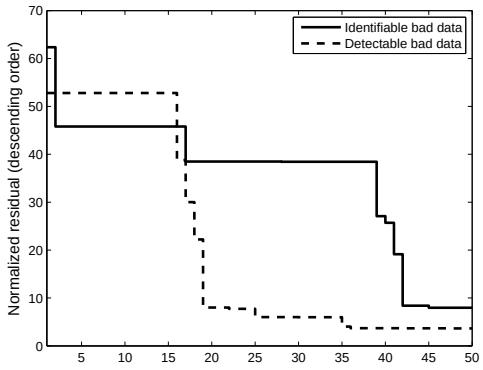


Fig. 7. Fifty largest normalized residuals in identifiable and non-identifiable bad data cases.

Main advantages of the proposed solution scheme, which have been shown by means of a realistic case study, are: early bad data processing capability for each subsystem during the decoupled solution phase, provided the redundancy is large enough, and possibility of tailoring both the network model and solution technique according to the peculiarities of each subsystem (e.g. unbalanced or ill-conditioned feeders).

Appendix

The two-stage procedure proposed in this paper provides a solution $\hat{x}$ which is the correct one if the net power injections at border buses are unknown. In those cases in which they are measured or should be enforced as null-injection equality constraints, the initial estimate $\hat{x}$ given by (14) can be easily corrected as explained below (each case will be separately addressed).

Border bus with null-injection equality constraints

In this case, the linear model (13), relating the estimate $\tilde{y}$ with the state vector x, must be completed with the null-injection border constraints, compactly written as $p(x)=0$. The resulting Lagrangian function to be minimized is therefore:

$$\mathcal{L}=\frac{1}{2}(\tilde{y}-Cx)^T W_y(\tilde{y}-Cx)+\lambda^T p(x)$$

Taking into account that $\hat{x}$ satisfies (14), the first-order optimality conditions lead to the following system:

$$\begin{bmatrix} C^T W_y C & P^T \\ P & 0 \end{bmatrix}\begin{bmatrix} \Delta x_c \\ \lambda \end{bmatrix}=\begin{bmatrix} 0 \\ -p(\hat{x}) \end{bmatrix} \qquad (26)$$

where P is the Jacobian of $p(x)$ and Δx_c the required state vector correction. As the table of factors of matrix $C^T W_y C$ is already available and the size of λ is very small, the best way to solve the above system is as follows:

1) Eliminate Δx_c and obtain λ by solving the resulting system

$$\left[P(C^T W_y C)^{-1}P^T\right]\lambda=p(\hat{x}) \qquad (27)$$

The size of this reduced system is proportional to the number of border buses with equality constraints, which

is comparatively very small in a radial distribution system.

2) Obtain Δx_c from,

$$(C^T W_y C)\Delta x_c = -P^T \lambda \tag{28}$$

1) *Border bus with regular injection measurements*

In this case, the linear model of the coordination phase must be completed with the border injection measurements, $z_p = p(x) + e_p$. The Lagrangian function to be minimized is therefore:

$$\mathcal{L} = \frac{1}{2}(\tilde{y} - Cx)^T W_y (\tilde{y} - Cx) + \frac{1}{2}(z_p - p(x))^T W_p (z_p - p(x))$$

where W_p is the weighting matrix of the border measurements.

When linearized around $\hat{x}$, the first-order optimality conditions lead to:

$$\begin{bmatrix} C^T W_y C & P^T \\ P & -W_p^{-1} \end{bmatrix} \begin{bmatrix} \Delta x_c \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ z_p - p(\hat{x}) \end{bmatrix} \tag{29}$$

The above linear system can be easily solved by resorting to the table of factors of $C^T W_y C$ as follows:

1) Eliminate Δx_c and obtain λ by solving the resulting system

$$\left[P(C^T W_y C)^{-1} P^T + W_p^{-1}\right] \lambda = p(\hat{x}) - z_p \tag{30}$$

2) Obtain Δx_c from,

$$(C^T W_y C)\Delta x_c = -P^T \lambda \tag{31}$$

Note that this case reduces to the previous one for null injections ($z_p = 0$) with arbitrarily large weights ($W_p \to \infty$). Hence, the augmented formulation of the border constraints provides a common framework to deal indistinctly with regular or virtual (exact) measurements.

REFERENCES

[1] I. Roytelman, S. M. Shahidehpour, "State estimation for electric power distribution systems in quasi real-time conditions," *IEEE Trans. on Power Delivery*, vol. 8, no. 4, pp. 2009-2015, Oct 1993.

[2] C.N. Lu, J.H. Teng, W.-H.E. Liu , "Distribution state estimation", *IEEE Transactions on Power Systems*, vol.10, no. 1, pp. 229-240, February 1995.

[3] R. Hoffman, S. Lefebvre, J. Prevost, "Distribution State Estimation : A Fundamental Requirement for the Smart Grid", DistribuTECH (2010).

[4] A. G. Phadke, J. S. Thorp, *Synchronized Phasor Measurements and Their Applications*, Springer, 2008.

[5] J. D. McDonald, *Electric Power Substation Engineering*, CRC Press, Boca Raton, FL, 2003.

[6] A. Bose, "Smart Transmission Grid Applications and Their Supporting Infrastructure," *IEEE Transactions on Smart Grid*, vol.1, no.1, pp. 11-19, June 2010.

[7] A. de la Villa Jaén, P. Cruz Romero, A. Gómez-Expósito, "Substation data validation by local three-phase generalized state estimators," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 264-271, February 2005.

[8] A. Gómez-Expósito, A. de la Villa Jaén, "Two-Level State Estimation With Local Measurement Pre-Processing," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 676-684, May 2009.

[9] I. Dzafic, S. Henselmeyer, H.T. Neisius, "Real-time Distribution System State Estimation," IPEC, 2010 Conference Proceedings, pp. 22-27, Oct. 2010.

[10] A. Gómez-Expósito, A. de la Villa Jaén, C. Gómez-Quiles, T. van Cutsem, P. Rousseaux, "A Taxonomy of Multi-Area State Estimation Methods", *Electric Power Systems Research*, vol. 81, pp. 1060-1069, 2011.

[11] K. Brand, V. Lohmann, W. Wimmer, *Substation Automation Handbook*, Utility Automation Consulting Lohmann, Bremgarten, Switzerland, 2003.

[12] H. M. Kim, J. J. Lee, D. J. Kang, "A Platform for Smart Substations," fgcn, vol. 1, pp.579-582, Future Generation Communication and Networking (FGCN 2007), vol. 2, 2007.

[13] M. Behnke et al, "Secondary Network Distribution Systems: Background and Issues Related to the Interconnection of Distributed Resources", Technical Report NREL/TP-560-38079, July 2005, available at http://www.nrel.gov/docs/fy05osti/38079.pdf

[14] T.H. Chen et al, "Feasibility Study of Upgrading Primary Feeders From Radial and Open-Loop to Normally Closed-Loop Arrangement," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1308-1316, May 2004.

[15] A. Gómez-Expósito, A. Abur, A. de la Villa Jaén, C. Gómez-Quiles "A Multilevel State Estimation Paradigm for Smart Grids, *Proceeedings of the IEEE*, vol. 99, No. 6, pp. 952 - 976, Jun 2011.

[16] A. Abur, A. Gómez-Expósito, *Power System State Estimation: Theory and Implementation*, Marcel Dekker, April 2004.

[17] A. Monticelli, A. Garcia, "Reliable Bad Data Processing for Real-Time State Estimation", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-102, No. 5, pp. 1126-1139, May 1983.

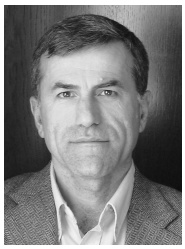
[18] J. Zhu and A. Abur, "Effect of Phasor Measurements on the Choice of Reference Bus for State Estimation", *Proceedings of the IEEE PES General Meeting*, June 24-28, 2004, Tampa, FL.

[19] W.F. Tinney, V. Brandwajn, S.M. Chan, "Sparse Vector Methods," *IEEE Trans. on Power Apparatus and Systems*, vol. PAS-104(2), pp. 295-301, Feb. 1985.

[20] A. Gómez-Expósito, E. Romero-Ramos, "Reliable Load Flow Technique for Radial Distribution Networks," *IEEE Trans. on Power Systems*, vol. 14 (3), pp. 1063-1069, August 1999.

[21] A. Gómez-Expósito, A. de la Villa Jaén, J. Ramírez-Izaga, "An Alternative State Estimation Formulation for Radial Distribution Networks", Proceedings of the IEEE PowerTech'07, Paper 563, Lausanne, July 2007.

Catalina Gómez-Quiles (SM'09) received the Eng. degree from the University of Seville, Spain, in 2006 and the M.Eng. degree from McGill University, Montreal, Canada in 2008, both in Electrical Engineering. She is currently pursuing her PhD at the University of Seville. Her research interests include state estimation and risk assessment in competitive electricity markets.



Antonio Gómez-Expósito (F'05) received the electrical engineering and doctor degrees from the University of Seville, Spain, where he is currently the Endesa Red Industrial Chair Professor. His primary areas of interest are optimal power system operation, state estimation, digital signal processing and control of flexible ac transmission system devices.

Antonio de la Villa Jaén was born in Riotinto, Spain, in 1960. He received his electrical and doctor engineering degrees from the University of Seville. He is presently an Associate Professor at the Department of Electrical Engineering, University of Seville. His primary areas of interest are computer methods for power system state estimation problems and wave energy converter control.

On the Use of PMUs in Power System State Estimation

Antonio Gómez-Expósito
University of Seville
Seville, Spain
age@us.es

Ali Abur
Northeastern University
Boston, USA
abur@ece.neu.edu

Patricia Rousseaux
University of Liège
Liège, Belgium
p.rousseaux@ulg.ac.be

Antonio de la Villa Jaén
University of Seville
Seville, Spain
adelavilla@us.es

Catalina Gómez-Quiles
University of Seville
Seville, Spain
catalinagq@us.es

Abstract - Synchronized phasor measurement units (PMUs) are becoming a reality in more and more power systems, mainly at the transmission level. This paper presents, in a tutorial manner, the benefits that existing and future State Estimators (SE) can achieve by incorporating these devices in the monitoring process. After a review of the relevant PMU technological aspects and the associated deployment issues (observability, optimal location, etc.), the alternative SE formulations in the presence of PMUs are revisited. Then, several application environments are separately addressed, regarding the enhancements potentially brought about by the use of PMUs.

Keywords - PMUs, state estimation, observability, measurement placement, dynamic estimation, multi-area estimation

1 Introduction

STATE estimation (SE) has been for decades one of the essential applications in Energy Management Systems (EMS), allowing secure operation of transmission grids. Measurements received and processed by the state estimators typically include power flows, net power injections and voltage and current magnitudes [1]. A basic assumption behind the SE is that the measurement set constitutes a single snapshot of the system being monitored, which is not fulfilled in practice because it takes a while to remotely capture and centrally gather all the information to be processed by the SE. In fact, it would be very difficult, if not impossible to assure that all measurements are synchronized (i.e., refer to the same instant). However, as long as the time elapsed between the first and the last component in the measurement set is small enough, compared to the time constant of the system load, this assumption will be acceptable in practice.

Nowadays, there is a clear trend to broaden the geographical scope of many SEs, in accordance to the needs of regional electricity markets, in which long-distance energy transactions have to be accurately and permanently monitored. In this context, the task of collecting wide area measurements and synchronizing the solutions provided by each control area will be a challenging one [2].

Recently, the introduction of more sophisticated protection and measurement components, such as the Intelligent Electronic Device (IED), has provided the capability of certain phase angle differences between adjacent (voltage and current) phasors to be added at the local area or substation level [3, 4]. While this may benefit to a certain extent the accuracy of the SE at the TSO level, it can hardly be helpful in the multi-TSO case.

Occasionally, a seemingly simple technical innovation can become a major player in changing the entire industry. There are several historical examples of such innovations one can identify, including the light bulb, transistor, laser, etc. Global positioning satellite (GPS) system can be included in this vein, perhaps not unlike the internet (or the original DARPA-net), which was also created initially for a small group of users, but then rapidly became a universal tool employed by almost anyone worldwide. The GPS system provides two important benefits which were previously not readily or easily acquired. One is the ability to determine geographical coordinates and the other is to have global access to a very accurate clock allowing to time stamp measured quantities irrespective of the physical coordinates at which these measurements are taken. Development and installation of the GPS system was soon followed by numerous engineering applications, which mainly consisted of a receiver and a processor. The receiver's function is to capture the signals transmitted from a redundant set of satellites and then process them in various different ways based on the objective of the implemented application. One example of such an application is the phasor measurement unit (PMU). These devices are installed at substations in electric power systems, and their objective is to accurately determine the frequency of the alternating current and voltage, and also produce the phasor representation of these signals defined with respect to the global clock.

Deployment of PMUs started at a slow pace about a decade ago and accelerated after a number of successive blackouts experienced in power systems all around the globe.

Phasor measurements are synchronized with respect to the time reference provided by the GPS satellites. Thanks to this accurate global time reference, synchrophasors with identical time-stamp received from various substations allow to create a coherent picture of the system state at a given instant, eliminating in this way the need to artificially set a phase angle, arbitrarily taken as the reference

angle in conventional state estimators.

Devices that can measure synchronized phasors were developed and potential benefits of PMU measurements were recognized over twenty years ago by Phadke et al. [5, 6]. The first implementation of GPS-synchronized phase angle measurements in an industrial power system SE was presented in [7].

Utilization of PMU measurements will impact state estimation in different ways. On the one hand, since the number of PMUs installed in existing power systems is not yet sufficient to carry out SE exclusively based on PMU measurements, SE formulation and solution remains nonlinear and iterative respectively. More imaginative solutions, sequentially handling conventional and PMU measurements in a two-step procedure have been also proposed [8].

On the other hand, SE related issues such as network observability and measurement placement [9, 14], solution accuracy and reliability (convergence rate), processing of bad data and other (parameter and topological) types of errors [15, 16] will have to be reconsidered. Moreover, in view of the higher sampling rates at which PMUs can work, the possibility of estimating the dynamic evolution of certain critical variables is being explored [17].

This paper covers in a succinct manner all of the above issues. Then, with the help of small tutorial examples, the potential benefits provided by the incorporation of PMUs in several SE environments are shown.

2 Phasor Measurement Units

A PHASOR Measurement Unit is a digital device providing synchronized voltage and current phasor measurements, referred to as synchrophasors [18].

PMU features were first implemented in stand alone units whose most relevant function is its capability to provide synchrophasors. Nowadays, many IEDs (RTUs, protective relays, ...) have been upgraded to produce synchrophasor measurements in addition to their own function.

2.1 General PMU architecture

Most generally, PMUs provide multi-channel input so that in addition to the voltage at the installation bus, currents in more than one line and possibly in all incident lines can be processed by a single unit. The three-phase voltages and currents are converted to appropriate analog inputs by instrument transformers and anti-aliasing filtering. Each analog signal is digitized by the A/D converter with sampling rate usually varying from 12 to 128 samples per cycle of the nominal power frequency. The sampling clock is phase-locked with the GPS clock pulse which provides the Universal Time Coordinated (UTC) time reference used to time-tag the outputs.

Phasors of phase voltages and currents are computed from sampled data by the PMU microprocessor using a signal processing technique as described in section 2.2 below. The calculated phasors are combined to form the positive sequence phasor measurements (as well as negative and zero sequences in case of unbalanced conditions). Other estimates of interest are frequency and rate of change of frequency.

Computed phasor measurements are transmitted through a digital communication network to higher level applications at a rate of 10 up to 60 frames per second. Many PMUs offer storage capacity enabling local exploitation of synchrophasors. However, in many real-time applications, and in particular state estimation, the phasor measurements are not used locally but rather at remote locations. Phasor data from a number of PMUs is then collected by a special-purpose computer, called Phasor Data Concentrator (PDC), which correlates phasor data by time-stamp to create a system-wide measurement set. PDCs can provide a number of specialized outputs such as a direct interface to a SCADA, EMS system or an upper level PDC.

2.2 Syncrophasors and measurement techniques

The basic definition of the phasor representation of a sinusoidal waveform at nominal frequency f_0, both in its polar and rectangular form, is illustrated in Fig. 1. The phasor angle is given by the angular difference between the peak of the sinusoid and the reference time $t = 0$. When considering synchrophasor, this reference time corresponds to the time-tag. If the waveform is not a pure sine signal, the computed phasor represents its fundamental frequency component.

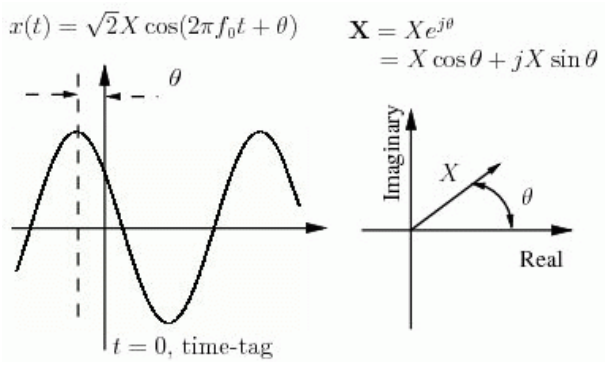


Figure 1: Sinusoidal waveform and its phasor representation

There have been several digital algorithms proposed to estimating the phasor data. The most commonly used technique relies on the Discrete Fourier Transform (DFT) [19]. The signal phasor is computed in a continuous process from successive samples in a moving data window of one or several fundamental cycles. Adding the contribution of the new sample while removing that of the oldest one provides the more computationally efficient recursive DFT algorithm. The sampling clocks are usually kept constant at a multiple of f_0. When frequency varies by a small amount around its nominal value, the leakage error introduced in phasor estimates can be compensated with high accuracy by a post-processing filtering. It can also be shown that the computed phasor rotates in the complex plane with an angular velocity equal to the difference between f_0 and the actual frequency so that frequency and rate of change of frequency estimates are given by the first

and second derivatives of the phasor angle.

An alternative consists in replacing the DFT by a wavelet transform [20]. A number of algorithms based on nonlinear estimation techniques (nonlinear weighted least squares estimation, Kalman filtering, neural networks, etc.) have also been proposed in the literature [21, 22, 23]. According to these approaches, the phase signal is modeled as a nonlinear function of the phasor data (amplitude, phase angle, frequency, rate of change of frequency) considered as parameters to be estimated from the waveform samples.

2.3 PMU performances and standards

Performances of a PMU device, in terms of accuracy or processing time, are dictated by its components, mainly, the instrumentation channel, the A/D converter and the parameters of the phasor estimation algorithm.

Regarding state estimation, the accuracy of PMU data is a very important issue. It is recognized that synchrophasor measurements are usually more precise than conventional SCADA ones. Conceptually, PMU data are time tagged with precision better than 1 microsecond and magnitude accuracy that is better than 0.1%. However, this potential performance is not achieved due mainly to errors from instrumentation channels and system imbalances. Presently, evaluation of PMU data accuracy is still a challenging problem discussed in the scientific literature [24].

PMU are manufactured by a variety of companies defining specifications for each particular PMU device. The specifications concern: the window length, sampling rate and type of phasor estimation algorithm, the phasor estimate reporting rate, the communication protocol and the measurement accuracy.

In order to achieve interoperability among PMUs, it is essential that their behavior complies with a common standard. The most recent IEEE C37.118-2005 standard [25] defines the synchrophasor convention and the time-tagging process, provides the definition of an accuracy measure as well as requirements for measurement performances under steady-state conditions. It also defines data communication formats. Requirements for response to power system transients are not considered.

3 Background on State Estimation

Given the following measurement equation [26]:

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

where:

x is the state vector (size $n = 2N - 1$),

z is the measurement vector (size $m > n$),

h is the vector of functions, usually nonlinear, relating error free measurements to the state variables,

e is the vector of measurement errors, customarily assumed to have a Normal distribution with zero mean and known covariance matrix R. When errors are independent R is a diagonal matrix with values σ^2, where σ is the standard deviation of the measurement errors.

the maximum likelihood estimate $\hat{x}$ is obtained by minimizing the Weighted Least Squares (WLS) function:

$$J = \sum_{i=1}^{m} [z_i - h_i(\hat{x})]^2 / \sigma_i^2$$

The minimum of the scalar J is reached by iteratively solving the so-called Normal equations:

$$G_k \Delta x_k = H_k^T W [z - h(x_k)] \tag{2}$$

where:

$H_k = \partial h / \partial x$ is the Jacobian evaluated at $x = x_k$,

$G_k = H_k^T W H_k$ is the gain matrix,

$W = R^{-1}$ is the weighting matrix,

$\Delta x_k = x_{k+1} - x_k$, k being the iteration counter.

Iterations finish when Δx_k is within an appropriate tolerance. It can be shown that the covariance of the estimate is:

$$\mathrm{cov}(\hat{x}) = \hat{G}^{-1} \tag{3}$$

where $\hat{G}$ is the gain matrix computed in the last iteration.

Upon convergence, the bad data processing function is activated to detect, identify and eliminate bad analog measurements. Bad data detection is accomplished based on the largest normalized residual test [27]. If the detection test fails, then the measurement corresponding to the largest normalized residual will be declared bad and its value will be removed or corrected [1].

In conventional bus-branch SE models the state vector is composed of voltage magnitudes and phase angles, whereas the measurement vector typically comprises measurements of power injections, branch power flows and voltage magnitudes. At lower voltage levels, though, line current magnitudes and bus current injection magnitudes can play a key role to obtain a sufficiently redundant system. The inclusion of PMU measurements is the subject of this paper.

In the so-called generalized SE model the state vector is augmented with power flows through circuit breakers (CB) at certain substations where a topology error is suspected, and the measurement vector may likewise include existing current or power flow measurements through any CB.

Transformer taps and suspected network parameters can also be handled, if sufficient redundancy exists, both by conventional and generalized SE.

4 Network observability and measurement placement

WHEN using conventional measurements network observability tests can be carried out based on the properties of the measurement Jacobian H which is the gradient of the nonlinear measurement equation evaluated at an arbitrary operating point such as the flat start. If the Jacobian has full column rank, then the system will be declared fully observable. In principle, the same approach will work when the Jacobian is modified in the presence of synchronized phasor measurements, which are either of voltage or current type. On the other hand, considering the long term outlook where sufficient number of PMUs are available to carry out state estimation exclusively using PMUs and disregarding all the conventional measurements, network observability analysis can be formulated as a simple graph covering problem.

Phasor measurement units may have different number of channels, i.e. they may support different number of inputs for voltage and current signals. It should also be noted that each input will be connected to one phase of a three phase voltage or current. While all three phases of voltage and current signals are monitored, most PMUs will output only the positive sequence values or the corresponding three phase quantity. Given the fact that power systems are sparsely connected, i.e. each bus has only a limited number of neighbors irrespective of the system size, channel limits on PMUs may be assumed to be sufficient to monitor as many signals as needed at a given bus. This assumption will be relaxed later and its impact will be investigated.

Assuming that a PMU is placed at bus k, the following quantities can then be assumed to be available:

- Voltage phasor at bus k,
- Current phasors along all lines/branches incident to bus k.

Network model and parameters being known, the above information will allow computation of phasor voltages at all the neighboring buses as well. Hence, placing a PMU at a given bus implies observability of all branches incident to that bus. This simple observation will lead to the following integer programming formulation of the network observability problem using only phasor measurements:

$$Minimize \quad C^T x_i \tag{4}$$

$$Subjectto \quad AX \quad \geq 1 \tag{5}$$

where
C is the cost vector for installation of PMUs,
X is a binary vector indicating the presence (1) or absence (0) of PMUs at buses,
A is a binary matrix mapping nonzero entries of the bus admittance matrix to ones,
1 is a vector of ones.

Buses corresponding to the nonzero values in the solution of (5) will yield the locations to place PMUs for full network observability [10]. While the conventional measurements are excluded in this formulation, equality constraints can be incorporated in order to reduce the number of required PMUs. The most common equality constraints in power systems are those provided by the net zero injections at passive buses with no generation or load. These can be readily incorporated into the problem of (5) as done in [11].

4.1 Methods of placing PMUs for different objectives

While the ultimate goal is to populate power systems with enough PMUs to facilitate full observability based only on PMU measurements, this will still be a few years away. In the meantime, as investment decisions are to be made where and how many PMUs to place in a given system, different objectives may be considered. Furthermore, since PMUs may be considered for specific applications such as special protection schemes, secondary voltage control, voltage or angle stability monitoring, etc. there may be already a constrained set of buses where PMUs may have to be placed. In such cases, secondary considerations in order to make the best of these investments will be important. This section will briefly review two such cases.

4.2 PMU placement to detect topology errors

Topology errors are caused by incomplete or wrong information about one or more circuit breakers at the substations. These errors can be very difficult to detect and identify due to the specific measurement configuration around the affected substation. A certain type of topology error that is referred to as *branch topology error* is defined as the error in the status of a given branch, i.e. whether or not the branch is in or out of service [12]. These types of errors are relatively easier to detect and identify compared to the more complex ones involving several breakers leading to bus splits or mergers.

Detectability of a branch topology error is closely linked to the measurement redundancy and configuration. Hence, it is possible to make a previously undetectable branch topology error detectable by strategic meter placement. If a limited number of PMUs are being considered to be placed in a given system, one consideration may be to improve branch topology detectability. The optimal case would be to have all branches topology error detectable, but even making a subset of branches topology error detectable would be a welcome improvement.

This can be accomplished in three steps:

- Identify all branches that are topology error undetectable,
- For each identified branch, determine all candidate PMU locations so that if a PMU is placed, this will make the branch topology error detectable,
- Set up an optimal selection problem so that a minimum number of the candidates identified in the pre-

vious step can be selected to make topology errors associated with all identified branches in step 1, detectable.

Details of the problem formulation along with illustrative examples can be found in [13]. It should be noted that, as shown in [13], phasor measurements have some unique advantages over the conventional measurements when it comes to topology error detection and identification.

4.3 PMU placement for measurement error detection

Every measurement system may have vulnerability pockets where errors in one or more measurements can not be detected. Such measurements are referred to as *critical* and their measurement residuals will be null irrespective of their measured values [15]. A robust measurement design will address these vulnerabilities and strategically place meters in order to transform these critical measurements into redundant ones whose errors will be detectable. Techniques for identifying all critical measurements in a given power system exist and can be used to determine all such measurements. Once they are identified, a numerical factorization based approach (whose details are given in [15]) can be used to determine all possible locations where PMUs can be placed so that a given critical measurement will be transformed. These candidate PMU locations are then considered simultaneously and a minimum number that will transform all critical measurements, can be determined again using an integer programming formulation.

Simulation results shown in [15] clearly imply dependence of the optimal number of required PMUs on the network topology and existing measurement configuration. However, there are cases where with only a handful of PMUs, a very large number of critical measurements can be transformed, thus drastically improving the robustness of state estimation against bad data.

4.4 PMU placement for parameter error detection and identification

Every power system data base requires constant maintenance due to changes in network parameters either due to environmental conditions such as temperature, humidity, wind, etc. or due to human error in entering data corresponding to equipment parameters such as transformer taps, shunt capacitor banks, etc. A typical power system model will have a huge number of parameters associated with its line, transformer, shunt capacitor/reactor models. Hence, use of state estimator as a tool to detect and identify parameter errors has been a topic of numerous investigations. These investigations mainly focused on parameter estimation based on the assumption that a suspect set of parameters have already been identified. However, selecting a suspect set which is guaranteed to contain erroneous parameters simply based on measurement residuals is not always possible. A method that overcomes this limitation is recently developed and then applied to the case of strategic placement of PMUs for parameter error detection and identification [16].

It is noted that, if strategically placed, PMUs will enable error identification of certain parameters which are not possible to identify using conventional measurements, no matter how high of a measurement redundancy is introduced. This result is validated with some simple examples in [16]. Along with the branch topology error detection problem, the parameter error identification problem constitutes one of the examples where PMUs will have a unique edge over conventional measurements.

5 State estimation formulation in the presence of PMUs

UP to the advent of the PMU technology the measurement vector z contained only power (flow and injection), voltage magnitude and, in certain particular cases, current magnitude measurements, all of them taken in a non-synchronized manner.

At the local level [4], phase angle differences among adjacent voltage and/or current waveforms can also be provided by the new generation of intelligent electronic devices (IEDs).

This section discusses several modeling issues arising in the formulation of the SE problem when PMUs are to be incorporated.

5.1 Measurement models

As stated previously, PMUs can provide both voltage and current phasors, collectively denoted as z_V and z_I respectively.

Depending on whether the state vector is represented in polar or rectangular coordinates, different expressions will result for the measurement model, as follows:

- State vector in polar coordinates:

$$x = \begin{bmatrix} V \\ \theta \end{bmatrix}$$

In this case, the error-free models corresponding to each type of measurements take the form:

$$z = h(x) \quad (6)$$

$$z_V = Kx \quad (7)$$

$$z_I = h_I(x) \quad (8)$$

where it is assumed that z_V is represented in polar coordinates so that K is a trivial matrix with a single 1 in each row. If phase angle differences, $\theta_i - \theta_j$, are also considered, then the respective rows in K will contain a 1 and a -1. Note that the measurement model associated with z_I is nonlinear, irrespective of z_I being expressed in polar or rectangular coordinates. The rectangular version is preferable, however, owing to the numerical problems (undefined Jacobian terms) that may arise for very small currents.

The expressions for $h(.)$ can be found elsewhere [1] while those corresponding to $h_I(.)$ are given in [28].

- State vector in rectangular coordinates:

$$x_r = \begin{bmatrix} V_{Re} \\ V_{Im} \end{bmatrix}$$

The error-free models corresponding to each type of measurements take the form:

$$z = h_r(x_r) \quad (9)$$
$$z_{Vr} = Kx_r \quad (10)$$
$$z_{Ir} = Y_r x_r \quad (11)$$

where the measurement vectors z_{Vr} and z_{Ir} are assumed in rectangular coordinates,

$$z_{Vr} = \begin{bmatrix} V_{Re} \\ V_{Im} \end{bmatrix} \quad ; \quad z_{Ir} = \begin{bmatrix} I_{Re} \\ I_{Im} \end{bmatrix}$$

because this way the measurement models (10) and (11) become linear, which is one of the key advantages associated with PMUs. In this case, $h_r(.)$ reduces to a set of quadratic functions, provided V^2 rather than V measurements are included in z.

Note that, unlike in the polar state vector case, null injection constraints can be linearly formulated if null currents rather than null powers are enforced.

The expressions for $h_r(.)$ can be found elsewhere [1] while those corresponding to Y_r are given in [29].

5.2 Simultaneous SE formulation

The information provided by PMUs can and should be handled in theory at the same time the conventional measurements are processed by the SE. This requires modifications to the existing software in order to accommodate the new Jacobian terms and components of the residual vector [30].

Depending on the proportion of conventional versus PMU measurements, and the number of PMU channels, the polar or rectangular model will be preferable, as follows:

- If z_I is empty, that is, PMUs provide only voltage phasors, then the conventional polar model will be preferable. This requires minor adaptations of existing SEs to accommodate phase angle measurements (see the subsection 5.4 below for a discussion on the reference angle issue).
- In future environments, however, where PMU measurements will eventually replace conventional RTU measurements, the rectangular model will be preferable. Eventually, in the absence of any power or isolated voltage magnitude measurement, the resulting SE model will become fully linear in rectangular coordinates, which is a nice feature to exploit.

In certain cases, conventional raw measurements could be previously manipulated so that they are converted to PMU-like pseudomeasurements. For instance, a pair of power (flow or injection) measurements associated with a bus whose voltage phasor is provided by a local PMU, can be transformed into an equivalent current phasor, increasing in this way the linearity of the resulting model. This requires of course that the covariance of the pseudomeasurements be computed from that of the raw measurements.

Note also that, in this approach, the PMU measurements are taken into account from the very beginning during the observability and bad data analyses.

5.3 Sequential SE formulation

In this scheme, conventional measurements are first processed in the usual manner, and then a new SE is designed aimed at improving the initial estimates by incorporating the information provided by PMUs. A nonlinear transformation is required in between to switch between polar and rectangular coordinates [31].

The three stages involved are as follows:

1) Disregarding PMU measurements, obtain a preliminary estimate $\tilde{x}$ by solving the conventional nonlinear SE problem, given by (1)-(2). This requires that the entire network be observable in the presence of just RTU measurements. As a byproduct, the inverse of the gain matrix arising in the last iteration provides the estimate's covariance, according to (3):

$$\text{cov}(\tilde{x}) = \tilde{G}^{-1}$$

2) The estimate $\tilde{x}$ is transformed to rectangular coordinates:

$$\tilde{x}_r = f(\tilde{x}) \quad (12)$$

where the nonlinear functions $f(.)$ represent the well-known relationships,

$$\tilde{V}_{Re} = \tilde{V}\cos\tilde{\theta}$$
$$\tilde{V}_{Im} = \tilde{V}\sin\tilde{\theta}$$

In addition to $\tilde{x}_r$ its covariance is required. This is obtained from:

$$\text{cov}(\tilde{x}_r) = \tilde{F} \cdot \text{cov}(\tilde{x}) \cdot \tilde{F}^T \quad (13)$$

where $\tilde{F}$ is the Jacobian of $f(.)$ computed for $\tilde{x}$. Note that this Jacobian is a 2×2-block diagonal square matrix.

3) The phasor measurements provided by PMUs, in rectangular coordinates, along with the estimate $\tilde{x}_r$, lead to the following linear measurement model:

$$\tilde{x}_r = x_r + \varepsilon_x \quad (14)$$
$$z_{Vr} = Kx_r + \varepsilon_V \quad (15)$$
$$z_{Ir} = Y_r x_r + \varepsilon_I \quad (16)$$

where the covariance of ε_V and ε_I is a known diagonal matrix and that of ε_x is given by (13). Accordingly, the final estimate $\hat{x}$ is the solution to the Normal equations arising at this linear stage:

$$\begin{bmatrix} I \\ K_x \\ Y_r \end{bmatrix}^T \begin{bmatrix} W_x & & \\ & W_V & \\ & & W_I \end{bmatrix} \begin{bmatrix} I \\ K_x \\ Y_r \end{bmatrix} \hat{x} = \begin{bmatrix} I \\ K_x \\ Y_r \end{bmatrix}^T \begin{bmatrix} W_x & & \\ & W_V & \\ & & W_I \end{bmatrix} \begin{bmatrix} \tilde{x}_r \\ z_{Vr} \\ z_{Ir} \end{bmatrix} \quad (17)$$

where the weighting matrices W_x, W_V and W_I are the inverse of the respective covariance matrices. Note that W_V and W_I are customarily considered diagonal matrices, which ignores the fact that the measurement noises for the set of "raw" measurements gathered by a single PMU are correlated. On the other hand, the matrix W_x is given by:

$$W_x = \text{cov}^{-1}(\tilde{x}_r) = \tilde{F}^{-T} \cdot \tilde{G} \cdot \tilde{F}^{-1} \qquad (18)$$

where the inverse of $\tilde{F}$ is trivially obtained by computing the 2×2 inverse of its constitutive diagonal blocks. This matrix is no longer diagonal but its sparsity should be exploited.

The main advantage of this alternative is that a conventional SE is resorted to at the beginning, allowing existing software to be adopted. Moreover, the third stage consists of a linear SE, which implies that the solution is obtained in a single iteration, preventing the risk of divergence in the presence of bad data. Although the solution reached through this three-stage approach is not the optimal one, it is sufficiently accurate for practical purposes. Its main drawback is that PMU measurements cannot be used during the first step to potentially enlarge the observable network and to enhance the bad data detection and identification process.

5.4 Reference bus issues

State estimation problem is commonly formulated by choosing a reference bus (typically but not necessarily the same as the slack bus used for the power flow analysis) and setting its voltage phase angle equal to zero. This also implies that the reference phase angle will be excluded from the state vector and the corresponding column of H will be removed when building the measurement Jacobian. Alternatively, the reference phase angle can be retained in the state vector but then a phase angle pseudo-measurement of arbitrary value (zero for convenience) must be added for each observable island.

In the absence of any phase angle measurement, this practice presents no problems and provides a suitable framework to define the system state where the actual value of the reference bus voltage phase angle is irrelevant.

However, as the phasor measurements start populating the systems, the choice of a reference bus will no longer be an arbitrary decision. There are two possibilities:

1) Choose a bus where no PMU exists: This will create inconsistencies between the arbitrarily assigned reference angle at the chosen bus and actual phase angle measurements provided by PMUs at other buses.

2) Choose one of the buses with PMUs as the reference bus: This will work as long as the PMU at the chosen bus functions perfectly. If the measurements provided by this PMU contain errors, then these errors will not be detectable and will bias the estimated state.

This issue has been recognized early on and alternative approaches were considered. Among them is a document [32] which is produced by the Eastern Interconnection Phasor Project (EIPP) group. In this document a virtual bus angle reference, which is computed as the average of several phase angle measurements by PMUs located in the vicinity of a chosen bus is introduced. This approach still remains vulnerable to errors in individual PMU measurements despite the use of averaging.

In the presence of numerous PMUs it will be logical to use the absolute phase angle information provided by those devices. Hence, the measurement Jacobian will have to include columns corresponding to all bus voltage magnitudes as well as phase angles, the dimension of the system state vector being twice the number of buses [33].

In this case, the system will be declared observable if no zero pivots are encountered while factorizing G. When there is more than one observable island in the system excluding the phasor measurements, then there has to be at least one phase angle measurement in every observable island to make the overall system observable.

When there is only one phase angle measurement in the system, then this case can be reduced to the conventional formulation with an assigned reference bus. Since the value of its phase angle is irrelevant, errors in this measurement will not affect the estimation results (critical information).

A more realistic case is when there are two or more phasor measurements in the system. In this case, detection of phasor measurement errors requires higher redundancy as discussed below. Disregarding the phasor measurements, conventional network observability analysis [1] will yield the number of observable islands in a given system. Having at least one phasor measurements in every observable island will ensure observability for the entire network. In order to be able to detect and identify errors in the phasor measurements, their redundancy should be further increased in their respective observable islands. Definition of critical k-tuples can be found in [1]. Following this definition, it can be shown that two phasor measurements will ensure detectability and three will be necessary for identification of bad data associated with any phasor measurement in a given observable island.

6 Application environments

IN this section, the improvement in SE performance due to the incorporation of PMU measurements is illustrated. First, the effect of including PMU measurements in the accuracy of the TSO-level estimator is assessed. Then, the enhanced synchronization capability provided by PMUs in the multi-TSO SE case is addressed. Owing to space limitations, only small tutorial examples are considered, but the main conclusions remain valid for realistic networks.

Other issues, such as the improved monitoring of smart distribution systems by combining SE and PMUs, or the use of the high volume of historical data collected by PMUs to develop reliable load forecasting, still in their

infancy, are not addressed for the sake of brevity.

For the simulations below the state variables are expressed in polar form. The complex voltage phasors provided by PMUs are assumed in polar form while the current phasors are represented in rectangular form in order to avoid ill-conditioning problems [4]. Under these assumptions, the PMU measurements can be expressed as follows:

$$\begin{aligned} V_i^m &= V_i + \varepsilon_{V_i} \\ \theta_i^m &= \theta_i + \varepsilon_{\theta_i} \\ I_{Re,ij}^m &= V_i(A\sin\theta_i + B\cos\theta_i) + \\ &\quad V_j(C\sin\theta_j + D\cos\theta_j) + \varepsilon_{I_{Re,ij}} \\ I_{Im,ij}^m &= V_i(E\sin\theta_i + F\cos\theta_i) + \\ &\quad V_j(G\sin\theta_j + H\cos\theta_j) + \varepsilon_{I_{Im,ij}} \end{aligned}$$

where the constants A to H depend on the parameters of the π model associated with branch i-j [28]. In all simulated scenarios the PMU and conventional measurements have been handled simultaneously during the estimation process.

6.1 Utilization of PMUs at the TSO level

The accuracy improvement arising by incorporating PMU measurements in a conventional SE at the TSO level has been assessed with the help of the 5-bus network sketched in Fig. 2. Different scenarios with increased redundancies have been considered:

1) Conventional measurements only: the set of measurements used for these simulations is shown in Fig. 2. The standard deviation for this type of measurements has been set to 0.01.

2) Inclusion of 1 PMU: in addition to the conventional set of measurements a single PMU ('PMU1' in Fig. 3) has been located at node 3. This device has several channels measuring the complex voltage at bus 3 as well as the current phasors through lines 3-4 and 3-5.

3) Inclusion of 3 PMUs: two more PMUs are added at nodes 4 and 5 ('PMU2' and 'PMU3' in Fig. 3). These PMUs measure only the voltage phasor at the corresponding node.

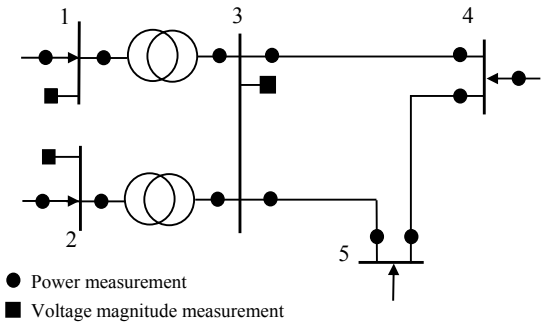


Figure 2: 5-bus illustrative network with conventional measurements

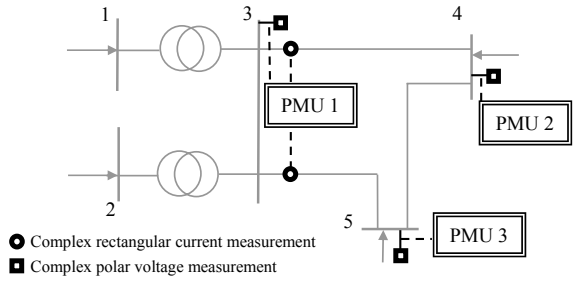


Figure 3: 5-bus illustrative network with PMU measurements

Different values for the standard deviation associated with PMU measurements have been tested, ranging from 0.0005 to 0.01 (same quality as conventional measurements). A parameter K, relating the standard deviation of conventional (σ_c) and PMU (σ_{pmu}) measurements, has been defined:

$$K = \sigma_c/\sigma_{pmu}$$

Hence, $K = 10$, for instance, means that the PMU measurements are 10 times more accurate in average than conventional measurements.

The measurements for the different scenarios have been created by adding gaussian noise to the 'exact' measurements corresponding to a given network state. The randomly generated noise has been scaled according to the standard deviation of the corresponding measurement, as follows:

$$z_i^m = z_i^{ex} + k_i\sigma_i \tag{19}$$

where z_i^m is the i-th measurement, z_i^{ex} the exact calculated value, k_i a randomly generated gaussian number $N(0,1)$ and σ_i the standard deviation assumed.

For each value of K one hundred Monte Carlo simulations have been performed. In order to evaluate the improvement brought about by PMUs, the accuracy of voltage magnitude and power flow estimates are separately analyzed (these are the most interesting magnitudes for EMS operators). For this purpose, the following indices have been defined:

$$\begin{aligned} S_V &= \sum_{i=1}^{n} |\tilde{V}_i - V_i^{\text{ex}}|/n \\ S_{PQ} &= \sum_{i=1}^{n_{PQ}} |\tilde{PQ}_{ij} - PQ_{ij}^{\text{ex}}|/n_{PQ} \end{aligned}$$

where n and n_{PQ} are the number of nodes and power flow (active and reactive) measurements, respectively and PQ_{ij} represents any power flow measurement.

Fig. 4 shows the resulting S_V index for the different scenarios. It can be observed how, as more PMUs are incorporated, lower values are obtained, which means more accurate estimates. Moreover, the estimates improve as the parameter K increases, which happens when the standard deviation associated with the PMU measurements decreases. Fig. 5 shows similar results for the index S_{PQ}.

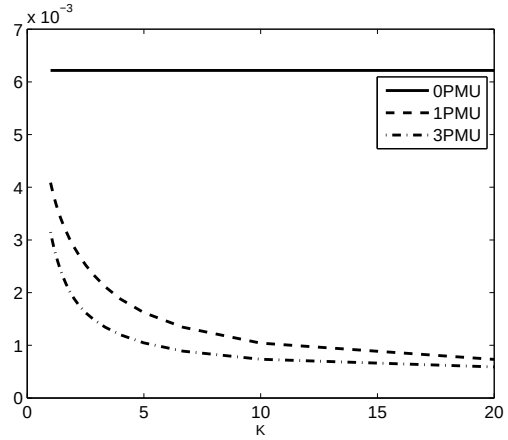


Figure 4: Simulation results for index S_V

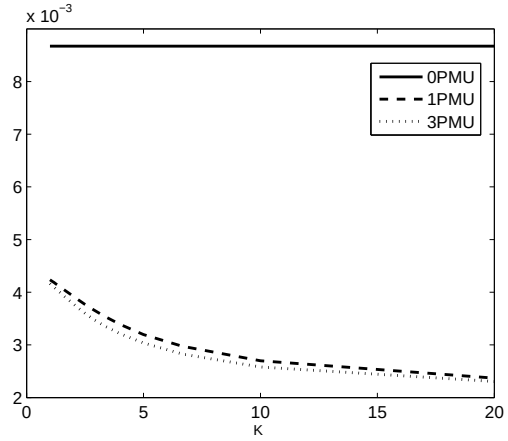


Figure 5: Simulation results for index S_{PQ}

The following remarks are in order:

- The incremental benefit of adding voltage phasors ('PMU2' and 'PMU3') is less noticeable than that of current phasors ('PMU1'), particularly when power flows are considered (Fig. 5).
- For values of K larger than 10 the accuracy improvement brought about by PMUs somewhat saturates. In general, however, this will also depend on the redundancy of the conventional measurement set.
- From the point of view of estimate accuracy, it is probably better to invest in further improving the quality of PMU measurements rather than increasing the number of PMUs. However, robustness against failure or loss of PMU channels is higher when a larger number of PMUs are installed.

6.2 Utilization of PMUs at the regional multi-TSO level

In regional power systems, where real-time measurements are gathered within neighboring areas by the various control centers distributed over the grid, the Multi-Area State Estimation (MASE) has got renewed interest. In this environment, the need of properly monitoring energy transactions across TSO borders via large interconnections, while at the same time processing the real-time data at the most appropriate place, make the MASE a good alternative. In general, MASE relies on some kind of decomposition-coordination scheme, taking advantage of the usually weaker geographical or measurement coupling among areas.

The MASE consists of a sequence of hierarchical SE processes comprising two main stages [2]: 1) each TSO independently solves the state estimation of its own area, including the tie-lines, border nodes and border measurements of adjacent areas; 2) the results of these decoupled estimators are then used by an *ad hoc* procedure which coordinates the estimate for the entire system. Under this scheme, when the different areas are not synchronized in time with PMU measurements, each of the areas sets a local phase angle reference for the TSO-level estimation process (first step). This requires the introduction of new state variables u, one for each area, relating the different phase angle references of the system. These variables, whose values are estimated at the second step, coordinate the results of the areas by referring the estimates to a global phase angle reference.

When synchronized PMU measurements are available at all areas, the local phase angle references are no longer needed. The PMU measurements will implicitly coordinate the independent estimates to the Universal Time Coordinated reference (UTC). In case some areas do not have PMU measurements available, only those areas will need to set a local phase angle reference, and for each one a u variable will have to be defined and estimated in order to coordinate the local estimates to the UTC.

Since the u variables coordinate the estimates of different areas, their role is crucial in the computation of power flows through tie-lines. The quality of the u estimates will affect the accuracy of the estimated flows through the tie-lines and, as a consequence, the estimates of the energy transactions among TSOs. If PMUs are available, the lack of u variables along with the enhanced accuracy usually provided by the PMU measurements, imply better estimates of power flows at tie-lines.

Some simulations have been carried out in order to evaluate the improvement of the multi-area state estimation in the presence of PMUs. The network used in these tests is made up of three IEEE 14-bus test networks, connected to each other as shown in Fig. 6, where only the tie-lines and border buses are represented. Table 1 shows the tie-line parameters adopted for these experiments.

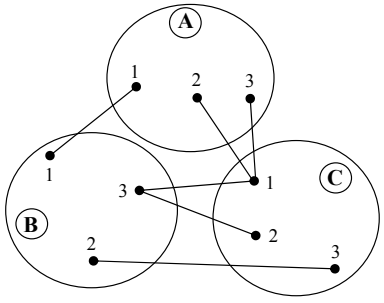


Figure 6: Multi-area system composed of 3 IEEE 14-bus networks

Table 1: Tie-line parameters

From	To	R (p.u.)	X (p.u.)	B_{sh} (p.u.)
1A	1B	0.0194	0.0592	0.0528
2A	1C	0.0540	0.2230	0.0528
3A	1C	0.0470	0.1980	0.0528
2B	3C	0.0581	0.1763	0.0528
3B	1C	0.0570	0.1739	0.0528
3B	2C	0.0670	0.1710	0.0346

Different sets of 'realistic' measurements have been generated from a given state, following the procedure described in the previous section, in accordance to (19). The base case contains a complete set of conventional measurements comprising: voltage magnitudes at all nodes ($\sigma = 0.01$), active and reactive power injections at all nodes ($\sigma = 0.02$) and active and reactive power flows at both ends of all branches ($\sigma = 0.015$). Four scenarios with additional PMUs have been considered:

1) A single PMU located at node 5 of TSO A. In this case, the local phase angle reference of TSO A can be taken as the global phase angle reference. The u variables for TSOs B and C will have to be computed in order to coordinate the estimates and calculate the power flows through the tie-lines.

2) PMU at node 5 of TSOs A and C. TSO B is the only area without PMU measurements and, therefore, it needs to set a local phase angle reference in order to perform its own SE. Since the estimate at TSO B is not synchronized to the rest of the system, its corresponding u variable will have to be estimated.

3) All TSOs with PMU at node 5. In this case, u variables are not necessary since the independent TSO estimates will be synchronized. Note that node 5 is an internal node, which is not directly connected to tie-lines.

4) A single PMU at node 1 of TSO C. Unlike in the previous scenarios, in this case the PMU is located at a border node, directly connected to TSOs A and B through tie-lines. Since the TSO areas of influence used for the first step include the tie-lines, the three TSOs can 'see' the PMU measurements in this case and, therefore, the estimates of the first step will be directly synchronized without the need of u variables.

Each PMU incorporates the voltage phasor measurement at the corresponding node and the current flow phasors measurements at the adjacent lines. The standard deviation of the PMU measurements has been set to 0.001.

Tables 2 and 3 show the exact and estimated values of the u variables for scenarios 1 and 2 respectively. For scenarios 3 and 4, no u variables are involved. In scenario 2 the estimate of u_B is more accurate (i.e., closer to the exact value) than in the first scenario, due to the higher redundancy level.

Table 2: Only one PMU at TSO A

Variable	Exact	Estimate
u_B	-0.0972	-0.0955
u_C	-0.1345	-0.1324

Table 3: PMU at TSOs A and C

Variable	Exact	Estimate
u_B	-0.0972	-0.0964

For each scenario, the tie-line power flows have been estimated and the index S_{PQ} defined in the previous section has been computed. In this case, an S_θ index, similar to the S_V index employed before, has been defined:

$$S_\theta = \sum_{i=1}^{n} |\tilde{\theta}_i - \theta_i^{ex}|/n$$

Table 4 shows the indices obtained for the different scenarios. For scenarios 1 to 3 the quality of the estimates increases (lower errors) with the number of PMUs incorporated. Note that scenario 4, with a single PMU located at a strategic border node so that no u variables are needed, gives better estimates of power flows through tie-lines than scenario 3, with 3 PMUs installed and no u variables involved either. However, in terms of phase angles estimates, scenarios 2 and 3 provide better results.

Table 4: Average estimation errors for the scenarios considered in MASE

Indices	Scen. 1	Scen. 2	Scen. 3	Scen. 4
S_{PQ}	0.00373	0.00338	0.00324	0.00118
S_θ	0.00234	0.00134	0.00119	0.00196

From these results, it can be concluded that PMU measurements improve the estimates of power flows through the tie-lines in a MASE process, due to the removal of u variables and the increased accuracy of the added measurements. Not only is the number of PMUs important, but also the location. A lower number of PMU measurements in carefully selected locations may give better estimates of the tie-line power flows. However, as expected, the number of phase angle measurements provided by PMUs is directly related to the accuracy of the phase angles estimates.

7 Use of PMUs in dynamic state estimation

IN the electric power engineering terminology, SE refers to static SE; this computes the state vector at one time instant from measurements captured at the same time instant. The process is repeated at successive times k but does not include any physical modeling of the time behavior of the system. The Dynamic State Estimation (DSE) method on the contrary relies on the following general dynamic model, written here in its discrete state transition form [34] :

$$x_{k+1} = f(x_k, w_k, k) \tag{20}$$

with noise w_k accounting for modeling errors.

Most DSE algorithms rely on the extended Kalman filter consisting of alternate sequences of filtering and prediction steps: at time k

$$\hat{x}_k = \bar{x}_{k-1} + K_k\left(z_k - h(\bar{x}_{k-1})\right) \tag{21}$$

$$K_k = \left(H_k^T W H_k + M_k^{-1}\right)^{-1} H_k^T W \tag{22}$$

$$\bar{x}_{k+1} = F_k \hat{x}_k + u_k \tag{23}$$

$$M_{k+1} = F_k \left(H_k^T W H_k + M_k^{-1}\right)^{-1} F_k^T + Q \tag{24}$$

These equations are obtained after proper linearization of (20) and (1) where:

$\hat{x}_k$ $(\bar{x}_k)$ is the estimated (predicted) state at time k,

F_k is the Jacobian of f evaluated at time k,

u_k acts as a command term coming from the linearization,

M_k is the covariance matrix of predicted state $\bar{x}_k$,

Q is the covariance matrix of noise w, assumed to have a Normal distribution with zero mean.

Benefits which could be encountered from DSE are linked to its predictive ability which provides the necessary information to perform preventive analysis and control and can also help observability analysis, identification of bad data and detection of topology errors.

In practice, however, DSE is faced with the problem of the availability of a reliable modeling of the system state evolution. According to the power system state estimation paradigm, the system dynamics are modeled as a succession of steady states, the transitions between states being caused by the variations of the loads and by the corresponding adaptations of the generations. Up to the advent of the PMU technology, the available measurements were limited to slow varying quantities captured at relatively low rates thus limiting the application of DSE. The introduction of phasor measurements, precisely synchronized and available at higher rates, makes it possible to derive dynamic estimators capable of following faster system variations.

PMU measurements z_V and z_I (7), (8) are easily embedded in the EKF filtering step. The simultaneous formulation described in section 5.2 is used by handling at the same time the PMU and the conventional measurements in (21).

Regarding the dynamics modeling two main approaches can be distinguished. The first one [35] relies on the generic linear model :

$$x_{k+1} = F_k x_k + d_k + w_k \tag{25}$$

Here F_k is a diagonal matrix accounting for the state transition and d_k is associated with the trend component. These parameters are identified on-line from archived data of the system state using the Holt's linear exponential smoothing method.

The second approach [36] recognizes that, rather than the voltage state vector components, the variables which actually drive the system dynamics considered are the nodal power injections. The prediction is as follows : (1) the load flow data are predicted at the next time instant using a short-term nodal load forecasting technique; (2) the predicted state vector $\bar{x}_{k+1}$ is obtained through a standard load flow calculation.

The concept of DSE can be extended to short-term dynamics such as the generator speed or acceleration using the additional information of frequency and rate of change of frequency provided by PMU devices [37]. The state of the system is composed of the voltage phasor at each bus and also of the frequency of the voltage phasor at each bus of the system. Additional internal dynamical or algebraic states are also introduced for each device. The model of the system is described by a set of differential and algebraic equations as follows :

$$\frac{dx(t)}{dt} = f(x(t), y(t), t)\,, \qquad 0 = g(x(t), y(t), t) \tag{26}$$

where x and y are the dynamic and algebraic state vector respectively.

The estimation is distributed at the substation level using a three-phase breaker oriented, instrumentation channel inclusive model. The set of physical measurements z comprises all available data coming from PMUs, relays and other IEDs. These values are compared to the computed values $h(\bar{x})$ deduced from the model forming a measurement error vector e. To ease the computations, the model is quadratized and digitized through a numerical integration technique. A standard WLS estimation is performed on the sum of the squared errors e.

This algorithm can be run at rates comparable to those recommended in the synchrophasors standard IEEE-C37.118, thus enabling to track the system real-time dynamic evolution. This estimator can be used in several applications dealing with power system wide area monitoring and control, such as transient stability monitoring [38].

8 Conclusions

THIS paper aims to provide a candid evaluation of the way large, multi-area power systems will be monitored as their operations become more interdependent. Multi-areas can be defined either geographically or based on voltage levels. This paper's contributions build on the numerous innovative works done so far by various researchers and the hierarchical perspective to monitor very large scale power systems which is arising in the upcoming smart grid context. An important technological driver in this development is the synchronized phasor measurements, which provide benefits in identification of topological and parameter errors, maintaining network observability, improving statistical as well as numerical robustness of the estimators. They also pave the way of developing estimators with very high scan rates, making it possible to capture system dynamics which are currently ignored by existing state estimators. Another important driver is the set of computational and communication technologies that are rapidly becoming available at all substations, facilitating the implementation of hierarchical solutions like those discussed in this paper. Such hierarchical decoupling appears inevitable in order to efficiently address the growing complexity of the system due to the penetration of renewable distributed generation and storage, primarily at the lower voltage levels. In a near future, as the operation becomes more heavily dependent on these technologies and their automation, issues of cyber and physical security will need to be addressed.

9 Acknowledgements

This work was performed in the context of the PEGASE project funded by European Community's 7th Framework Programme (grant agreement No. 211407). The Spanish authors also thank the support of the DGI, under grants ENE2007-62997 and ENE2010-18867.

REFERENCES

[1] A. Abur, A. Gómez-Expósito, *Power System State Estimation: Theory and Implementation*, Marcel Dekker, April 2004.

[2] A. Gómez-Expósito, A. de la Villa Jaén, C. Gómez-Quiles, P. Rousseaux, T. Van Cutsem, "A Taxonomy of Multi-Area State Estimation Methods,", *Electric Power Systems Research*, in press.

[3] A. Gómez-Expósito and A. Abur A, "Use of Locally Synchronized Voltage and Current Measurements for State Estimation", Probabilistic Methods Applied to Power Systems Conference (PMAPS), September 22-26, 2002, Naples, Italy.

[4] A. de la Villa Jaén, P. Cruz Romero, A. Gómez-Expósito, "Substation data validation by local three-phase generalized state estimators", *IEEE Transactions on Power Systems*, Vol. 20, No. 1, pp. 264-271, February 2005.

[5] A. G. Phadke, J. S. Thorp and K. J. Karimi, "State Estimation with Phasor Measurements", *IEEE Transactions on Power Systems*, Vol. 1, No. 1, pp. 233- 241, February 1986.

[6] A. G. Phadke, "Synchronized phasor measurements in power systems", *IEEE Computer Applications in Power*, Vol. 6, Issue 2, pp. 10-15, April 1993.

[7] I. W. Slutsker, S. Mokhtari, L. A. Jaques, J. M. G. Provost, M. B. Perez, J. B. Sierra, F. G. Gonzalez, J. M. M. Figueroa "Implementation of Phasor Measurements in State Estimator at Sevillana de Electricidad", Proc. of the Power Industry Computer Application Conference, pp. 392-398, May 1995.

[8] R. F. Nuqui and A. G. Phadke "Hybrid Linear State Estimation Utilizing Synchronized Phasor Measurements", Proc. of PowerTech 2007, Lausane, Switzerland, July 2007.

[9] Bei Xu and A. Abur, "Observability Analysis and Measurement Placement for System with PMUs", *IEEE PES Power Systems Conference & Exposition*, New York, Oct.10-13, 2004.

[10] Bei Xu, Y. J. Yoon, and A. Abur, "Optimal Placement and Utilization of Phasor Measurements for State Estimation", *Power System Computation Conference*, Liege, Belgium, Aug. 2005.

[11] S. Chakrabarti, E. Kyriakides, and D. G. Eliades, "Placement of synchronized measurements for power system observability", *IEEE Transactions on Power Delivery*, vol. 24, no. 1, pp. 1219, Jan. 2009.

[12] K.A. Clements, P.W. Davis, "Detection and identification of topology errors in electric power systems", *IEEE Transactions on Power Systems*, Vol. 3, No. 4, pp. 1748-1753, Nov. 1988.

[13] J. Chen and A. Abur, "Enhanced Topology Processing via Optimal Measurement Design", *IEEE Transactions on Power Systems*, Vol. 23, No. 3, pp. 845-852, Aug 2008.

[14] M. Korkali and A. Abur, "Placement of PMUs with Channel Limits", *Proceedings of the IEEE Power and Energy Society General Meeting*, Calgary, CA, July 26-30, 2009.

[15] J. Chen and A. Abur, "Placement of PMUs to Enable Bad Data Detection in State Estimation", *IEEE Transactions on Power Systems*, Vol. 21, No. 4, pp. 1608-1915, Nov 2006.

[16] Jun Zhu and A. Abur, "Improvements in Network Parameter Error Identification via Synchronized Phasors", *IEEE Transactions on Power Systems*, Vol. 25 , No. 1, 2010 , pp. 44 - 50.

[17] A. P. S. Meliopoulos, G. J. Cokkinides, F. Galvan, B. Fardeanesh, P. Myrda, "Delivering accurate and timely data to all", *IEEE Power and Energy Magazine*, Vol. 5, No. 3, pp 74 - 86, May-June 2007.

[18] A. G. Phadke, J. S. Thorpe, *Synchronized Phasor Measurements and Their Applications*, Springer, 2008.

[19] A. G. Phadke, J. S. Thorpe, M. G. Adamiak, "A New Measurement Technique for Tracking Voltage Phasors, Local System Frequency, and Rate of Change of Frequency", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-102, No. 5, pp. 1025-1038, 1983.

[20] W. Chi-Kong, L. Ieng-Tak, W. Jing-Tao, H. Ying-Duo, "A novel algorithm for phasor calculation based on wavelet analysis", IEEE Power Engineering Society Summer Meeting, Vol. 3, pp. 1500-1503, July 2001.

[21] V. V. Terzija, M. B. Djuric, B. D. Kovacevic, "Voltage phasor and local system frequency estimation using Newton type algorithm", *IEEE Transactions on Power Delivery*, Vol. 9, No. 3, pp. 1000-1007, 1991.

[22] A. A. Girgis, R. G. Brown, "Application of Kalman filtering in computer relaying", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-100, No. 7, pp. 3387-3397, 1981.

[23] P. K. Dash, S. K. Panda, B. Mishra, D. P. Swain, "Fast estimation of voltage and current phasors in power networks using an adaptive neural network", *IEEE Transactions on Power Systems*, Vol. 12, No. 4, pp. 1494-1499, 1997.

[24] Z. Huang, B. Kasztenny, V. Madani, K. Martin, S. Meliopoulos, D. Novosel, "Performance Evaluation of Phasor Measurement Systems", IEEE Power Engineering Society General Meeting, Pittsburgh, PA, 2008.

[25] "IEEE Standard for Synchrophasors for Power Systems", C37.118-2005.

[26] F. C. Schweppe, J. Wildes, D. B. Rom, "Power system static state estimation, Part I, II, III", *IEEE Trans. Power Apparat. Syst.*, Vol. PAS-89, No. 1, pp. 120-135, January 1970.

[27] A. Monticelli, A. Garcia, "Reliable Bad Data Processing for Real-Time State Estimation", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-102, No. 5, pp. 1126-1139, May 1983.

[28] J. Chen, "Measurement enhancement for state estimation", Ph. D. dissertation. Texas A&M University, May 2008.

[29] R. F. Nuqui, "State Estimation and Voltage Security Monitoring Using Synchronized Phasor Measurements", Ph. D. dissertation. Virginia Polytechnic Institute, 2001.

[30] T.S. Bi, X.H. Qin, Q.X. Yang, "A novel hybrid state estimator for including synchronized phasor measurements", *Electric Power Systems Research*, Vol. 78, pp. 1343?1352, 2008.

[31] M. Zhou, V. A. Centeno, J. S. Thorpe, A. G. Phadke, "An Alternative for Including Phasor Measurements in State Estimators", *IEEE Transactions on Power Systems*, Vol. 21, No. 4, pp. 1930-1937, Nov. 2006.

[32] D. Novosel, H. Huang, K. Martin, "Implementation of Virtual Bus Angle Reference", Eastern Interconnection Phasor Project, November, 2005.

[33] Jun Zhu and A. Abur, "Effect of Phasor Measurements on the Choice of Reference Bus for State Estimation", *Proceedings of the IEEE PES General Meeting*, June 24-28, 2004, Tampa, FL.

[34] P. Rousseaux, T. Van Cutsem, T. E. Dy Liacco, "Whither dynamic state estimation?", *Electrical Power & Energy Systems*, Vol. 12, No. 2, pp. 104-116, April 1990.

[35] A. Jain, N. R. Shivakumar, "Impact of PMU in Dynamic State Estimation of Power Systems", 40th North American Power Symposium (NAPS), 2008.

[36] H. Xue, Q. Jia, N. Wang, Z. Bo, H. Wang, H. Ma, "A Dynamic State Estimation Method with PMU and SCADA Measurement for Power Systems", International Power Engineering Conference IPEC2007, pp. 848-853, 2007.

[37] E. Farantatos, G. K. Stefopoulos, G. J. Cokkinides, A. P. Meliopoulos, "PMU-Based Dynamic State Estimation for Electric Power Systems", IEEE Power & Energy Society General Meeting, 2009.

[38] S. Meliopoulos, G. Cokkinides, R. Huang, E. Farantatos, S. Choi, Y. Lee, "Wide Area Dynamic Monitoring and Stability Controls", Bulk Power System Dynamics and Control Symposium (IREP) VIII, Brazil, August 2010.

Bilinear Power System State Estimation

Antonio Gómez-Expósito, *Fellow, IEEE*, Catalina Gómez-Quiles, *Student Member, IEEE*, and Antonio de la Villa Jaén

***Abstract*—This paper presents a three-stage state estimation methodology based on the sequential solution of two weighted least squares (WLS) linear problems with a nonlinear explicit transformation in between. This requires that appropriate sets of auxiliary variables be introduced so that the resulting measurement models become linear. Simulation results show that the proposed approach yields, in a non-iterative manner, virtually the same solution as that provided by the Gauss-Newton iterative scheme arising in the conventional WLS method.**

***Index Terms*—Factorized state estimation, linear weighted least squares (WLS), measurement prefiltering.**

I. Introduction

SOON after its introduction by Schweppe, late in the 1960s [1], the state estimator (SE) became an essential tool in energy management systems (EMS), where a diversity of applications rely on accurate information about the system state.

Given a set of measurements arriving from remote terminal units (RTU), the SE filters out the associated noise and determines the maximum likelihood state. When measurement errors are Gaussian, uncorrelated, and with zero mean, obtaining such a state involves a weighted least squares (WLS) problem. As most measurements are nonlinearly related to the state variables, this leads to an iterative process in which the so-called Normal equations are repeatedly solved [2], [3].

Several improvements and alternatives have been considered throughout these decades regarding the solution of the resulting Normal equations, which can be roughly grouped in two major categories.

- Computational saving. Apart from the very basic issues related to sparsity exploitation [4], the main attempt to reduce the computational cost of the SE was the introduction of the decoupling principle by Monticelli *et al.* [5]. The fast decoupled SE formulation (FDSE) requires less memory and saves a significant fraction of computations by solving smaller equation systems with constant coefficient matrices. However, under certain operating conditions and/or combinations of network parameters, the overall convergence of the FDSE may deteriorate.
- Numerical issues. In this category, the orthogonal or QR factorization, numerically more stable yet more expensive than the Cholesky factorization, can be cited [6], [7]. Explicitly using equality constraints to model virtual measurements, rather than adopting disproportionately high weights, was also a major development [8], [9].

In addition to the WLS scheme, nonquadratic SEs have also been considered when statistical robustness against outliers is a concern [10]–[12].

Even though the computing power in nowadays EMS is orders of magnitude higher than that of the 1970s, the need for computationally more efficient and reliable techniques to solve the WLS problem still remains. On the one hand, the undergoing deployment of new digital devices, such as the intelligent electronic devices (IED) at the substation level [13], will provide a huge number of measurement points, including those associated with protective relays. On the other hand, there is a clear trend to broaden the geographical scope of SEs, in accordance to the needs of regional electricity markets, in which long-distance energy transactions have to be accurately and permanently monitored [14]. Therefore, SEs should be designed in the future to cope efficiently with very large networks and highly redundant measurement sets, for which solution speed and convergence reliability will always be relevant issues.

Recently, a factorized SE model has been proposed allowing the raw measurements to be preprocessed by a linear WLS filter, either at the substation level [15] or in a centralized fashion [16]. The information provided by the linear stage is then optimally processed by a nonlinear SE. The convergence pattern of the resulting two-stage procedure is usually better than that of the conventional iterative approach based on the Normal equations, particularly when relatively accurate measurement sets are available.

In this paper, a three-stage SE method is proposed. The first linear step is the same as that of [16]. Then an explicit nonlinear transformation is applied to the estimate provided by the first step, yielding a set of variables which are in turn linearly related to the conventional state vector. The third step simply solves the linear WLS problem arising after the nonlinear change of variables. An important feature of the proposed method is that, throughout the process, the associated weighting matrices are sparse.

The structure of this paper is as follows: Section II reviews both the conventional and the two-stage factorized WLS formulations. Section III presents in general form the three phases of the proposed methodology. Section IV introduces the set of intermediate variables adopted in this work and explains in detail the main steps of the resulting bilinear model. Finally, test results, based on IEEE benchmark systems, are presented in Section V, followed by the concluding remarks.

Manuscript received February 17, 2011; revised June 08, 2011; accepted July 09, 2011. This work was supported by the Spanish DGI, under grants ENE2007-62997 and ENE2010-18867. Paper no. TPWRS-00138-2011.

The authors are with the Department of Electrical Engineering, University of Seville, Seville, Spain (e-mail: age@us.es; catalinagq@us.es; adelavilla@us.es).

Digital Object Identifier 10.1109/TPWRS.2011.2162256

II. Background on State Estimation

A. Conventional WLS State Estimation

Given the following measurement equation [2]:

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

where

- x state vector to be estimated (size $n = 2N - 1$, N being the number of buses);
- z known measurement vector (size $m > n$);
- h vector of functions, usually nonlinear, relating error-free measurements to the state variables;
- e vector of measurement errors, customarily assumed to have a Normal distribution with zero mean and known covariance matrix R;

the Weighted Least Squares (WLS) estimator provides the maximum likelihood estimation by minimizing the following scalar function:

$$J = r^T W r = \sum_{i=1}^{m} W_i r_i^2 \tag{2}$$

where

- $r = z - h(\hat{x})$ measurement residual vector;
- $\hat{x}$ estimated state vector;
- $W = R^{-1}$ weighting matrix.

When errors are independent, R is a diagonal matrix with values σ_i^2, where σ_i is the standard deviation of the error associated with measurement i.

The minimum of the scalar J can be obtained by iteratively solving the so-called Normal equations:

$$G_k \Delta x_k = H_k^T W [z - h(x_k)] \tag{3}$$

where

- $H_k = \partial h / \partial x$ Jacobian evaluated at $x = x_k$;
- $G_k = H_k^T W H_k$ gain matrix;
- $\Delta x_k = x_{k+1} - x_k$, k iteration counter.

Iterations finish when an appropriate tolerance is reached on Δx_k. The covariance of the estimate is given by

$$\mathrm{cov}(\hat{x}) = G_k^{-1}.$$

After convergence, the bad data processing function, based on the largest normalized residual test [17], is run to detect, identify, and eliminate bad analog measurements.

B. Factorized WLS State Estimation

Recently, a factorized scheme [15], [16] has been proposed providing a computationally efficient two-stage procedure usually converging in one or two less iterations than the conventional approach described above. In addition, the first stage constitutes itself a valuable byproduct as a reliable linear prefilter, capable of providing accurate enough results for the measurement noise and redundancy levels typically found nowadays.

The factorized approach to solve the WLS problem arises when a vector of intermediate variables, y, is introduced in such a way that the nonlinear measurement model (1) is "unfolded" into a sequence of two WLS problems, the first of them linear, as follows:

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

$$y = f(x) + e_y. \tag{5}$$

For the resulting factorized model to be equivalent to the original one, the following condition must be satisfied:

$$h(x) = Bf(x) \quad \Rightarrow \quad H(x) = BF(x) \tag{6}$$

where F represents the Jacobian of $f(\cdot)$.

The WLS solution $\tilde{y}$ to the linear problem (4) is directly obtained by solving

$$G_b \tilde{y} = B^T W z \tag{7}$$

where

$$G_b = B^T W B \tag{8}$$

is the associated gain matrix.

Then, the optimal estimate $\hat{x}$ is computed by iteratively solving the system

$$[F_k^T G_b F_k] \Delta x_k = F_k^T G_b [\tilde{y} - f(x_k)] \tag{9}$$

where the dependence of F on x is removed for simplicity of notation.

In summary, instead of directly performing iterations with (3), the factorized approach provides the optimal solution in two stages:

1) compute $\tilde{y}$ by solving the linear system (7);
2) repeatedly solve (9) until convergence.

In this two-stage scheme, the auxiliary vector y plays the role of a state vector when solving (7) and that of a measurement vector when solving (9). The covariance matrix of $\tilde{y}$ is

$$\mathrm{cov}(\tilde{y}) = G_b^{-1} \tag{10}$$

the inverse of which (i.e., the sparse gain matrix G_b) appears as the weighting matrix in (9), like in a conventional WLS solution.

While the above factorized scheme was used in [16] mainly as a means of alleviating the computational cost by reducing the size of the auxiliary vector y as much as possible, in [15], the driving idea was to choose an enlarged vector y so that the first linear stage could be distributed at the substation level.

III. Three-Stage WLS Solution

In this work, the factorized SE scheme is further pursued in order to decompose the original nonlinear WLS problem into the sequential solution of two linear WLS stages with a nonlinear transformation interleaved.

The first linear stage is the same as that described above, relating the raw measurements z with the auxiliary vector y. Next, an explicit nonlinear transformation is performed to obtain a new set of variables u as follows:

$$u = f_u(y). \tag{11}$$

Note that the size of u is the same as that of y. Given an estimate $\tilde{y}$, along with its covariance $\mathrm{cov}(\tilde{y})$, the corresponding value $\tilde{u}$ can be directly computed from (11). Moreover, its associated covariance, when needed, is provided by

$$\mathrm{cov}(\tilde{u}) = \tilde{F}_u \mathrm{cov}(\tilde{y}) \tilde{F}_u^T \tag{12}$$

where $\tilde{F}$ is the Jacobian of $f_u(\cdot)$ computed at $\tilde{u}$. Actually, the inverse of the covariance will be needed in subsequent steps, which taking into account (10) can be obtained from

$$W_u = \mathrm{cov}^{-1}(\tilde{u}) = \tilde{F}_u^{-T} G_b \tilde{F}_u^{-1}. \tag{13}$$

In the implementation described in the next section, each variable in u will be related to one or at most two variables in y. Therefore, $\tilde{F}_u$ will be composed of scalar or 2×2 diagonal blocks with trivial inverses. Furthermore, the sparsity pattern of W_u will be the same as that of G_b.

Nothing has been said yet on the criteria to select the vector u. The main idea proposed in this paper is that u should be chosen in such a way that it is strictly a linear function of the conventional state vector x, as follows:

$$u = Cx + e_u \tag{14}$$

where C is an appropriate rectangular matrix and e_u represents the uncertainty of u, characterized by (13).

Given $\tilde{u}$ and W_u, the WLS estimate $\hat{x}$ is finally computed by solving

$$G_c \hat{x} = C^T W_u \tilde{u} \tag{15}$$

where

$$G_c = C^T W_u C \tag{16}$$

is the resulting gain matrix.

In summary, the conventional nonlinear measurement model (1) is replaced by the three following models:

$$z = By + e \tag{17}$$
$$u = f_u(y) \tag{18}$$
$$u = Cx + e_u \tag{19}$$

yielding a three-stage solution approach:

1) compute $\tilde{y}$ by solving the linear system (7);
2) obtain $\tilde{u}$ and W_u from (11) and (13);
3) compute $\hat{x}$ by solving the linear system (15).

The covariance of $\hat{x}$ and that of the resulting residuals can be computed, much as in the conventional WLS approach, by resorting to the table of factors of the gain matrix G_c. As a matter of fact, by direct substitution, keeping in mind the chain rule, G_c can be shown to be mathematically equivalent to the conventional gain matrix

$$G_c = C^T W_u C = C^T [\tilde{F}_u^{-T} G_b \tilde{F}_u^{-1}] C$$
$$= [C^T \tilde{F}_u^{-T} B^T] W [B \tilde{F}_u^{-1} C] = H^T W H = G$$

provided the Jacobian $\tilde{F}_u$ is obtained exactly at the same linearization point as G. In practice, this is not strictly the case as $\tilde{F}_u$ is computed around the point $\tilde{u}$, which is not the best possible estimate for u, whereas $G = G_k$ refers to the last iteration k of the Gauss-Newton scheme (3).

For the same reason, the estimate $\hat{x}$ provided by the three-stage procedure is not, from a mathematically rigorous point of view, the same as the one obtained by iteratively solving (3). If deemed necessary, a better estimate for u can be obtained from $\hat{x}$:

$$\hat{u} = C\hat{x} \tag{20}$$

which can be used to recompute F_u. Then, repeating the second linear filter (stage 3) yields an improved estimate $\hat{\hat{x}}$.

Nevertheless, as shown by the experiments below, even for unrealistically high measurement errors, a single run of the three-stage process provides results which are indistinguishable for all practical purposes from the optimal ones given by (3).

Note that, even though the network is fully observable in terms of the conventional state vector x, it may happen that certain components of the enlarged vector y are not observable (the observability issue is addressed in [16]). In this paper, it will be assumed that the measurement redundancy is appropriate to make y fully observable, which is the case in nowadays transmission networks.

IV. Proposed State Estimation Model

This section particularizes the three-stage bilinear WLS model to the power system SE problem, providing the required details about the state variables involved at each phase.

A. First Linear Stage

The measurement model at this stage is exactly the same as that proposed in [16]. The state vector y is composed of two variables per branch along with the squared voltage magnitudes. In order to make the paper self-contained, the resulting model will be summarized below.

For each branch connecting buses i and j, the following pair of variables is defined:

$$K_{ij} = V_i V_j \cos\theta_{ij} \tag{21}$$
$$L_{ij} = V_i V_j \sin\theta_{ij} \tag{22}$$

where V_i and θ_i represent the voltage magnitude and phase angle, respectively, of bus i. In addition, squared voltage magnitudes

$$U_i = V_i^2 \tag{23}$$

rather than plain voltage magnitudes are retained in the intermediate state vector y, which is then composed of $2b+N$ variables (b being the number of branches):

$$y = \{U_i, K_{ij}, L_{ij}\}. \tag{24}$$

TABLE I
JACOBIAN COMPONENTS FOR THE FIRST LINEAR STAGE

	$y=U_i$	$y=U_j$	$y=K_{ij}$	$y=L_{ij}$
$\partial U_i/\partial y$	1	0	0	0
$\partial P_{ij}/\partial y$	$g_{sh.i}+g_{ij}$	0	$-g_{ij}$	$-b_{ij}$
$\partial P_{ji}/\partial y$	0	$g_{sh.j}+g_{ij}$	$-g_{ij}$	b_{ij}
$\partial Q_{ij}/\partial y$	$-(b_{sh.i}+b_{ij})$	0	b_{ij}	$-g_{ij}$
$\partial Q_{ji}/\partial y$	0	$-(b_{sh.j}+b_{ij})$	b_{ij}	g_{ij}

Such a set of variables was first introduced in the context of radial load flows for distribution networks [18].

Any measurement handled by conventional SEs can be linearly expressed in terms of the vector defined above as follows:

- *Power flow measurements*. For a branch connecting buses i and j, measurements taken at terminal bus i:

$$P_{ij}^m = (g_{sh.i}+g_{ij})U_i - g_{ij}K_{ij} - b_{ij}L_{ij} + \varepsilon_P \quad (25)$$
$$Q_{ij}^m = -(b_{sh.i}+b_{ij})U_i + b_{ij}K_{ij} - g_{ij}L_{ij} + \varepsilon_Q \quad (26)$$

  where g_{ij}, b_{ij} are the series conductance and susceptance, respectively, and g_{sh}, b_{sh} represent the shunt values.
- Power injection measurements at bus i:

$$P_i^m = \sum_{j\in i} P_{ij} + \varepsilon_P \quad (27)$$
$$Q_i^m = \sum_{j\in i} Q_{ij} + \varepsilon_Q. \quad (28)$$

- *Voltage magnitude measurement*. To retain linearity at this stage, squared voltage magnitude measurements should be used:

$$\left(V^2\right)_i^m = U_i + \varepsilon_U. \quad (29)$$

  The ε terms in the above expressions represent the measurement noise. The standard deviation of the error associated with V_i^2 is related to that of the original measurement V_i through

$$\sigma(V_i^2) = 2E\{V_i\}\sigma(V_i) \quad (30)$$

  where $E\{V_i\}$ is the expected value of V_i which, in practice, can be approximated by the measured value.

The constant Jacobian terms corresponding to the above expressions can be found in [16], but they are collected in Table I for the sake of self-sufficiency. This defines the components of matrix B in the proposed implementation.

Even though Ampere measurements are seldom gathered in bulk transmission systems (see [2, Ch. 9], for a detailed discussion about the difficulties in using Ampere measurements), they could be useful to achieve observability at the subtransmission and distribution levels. Should the need arise, squared current magnitudes can be also linearly expressed in terms of the augmented set of variables y, as follows:

$$\left(I^2\right)_{ij}^m = \frac{P_{ij}^2+Q_{ij}^2}{U_i} + \varepsilon_{I2}$$
$$= \left[(g_{sh.i}+g_{ij})^2 + (b_{sh.i}+b_{ij})^2\right]U_i + (g_{ij}^2+b_{ij}^2)U_j - 2\left[g_{ij}(g_{sh.i}+g_{ij}) + b_{ij}(b_{sh.i}+b_{ij})\right]K_{ij} + 2(g_{ij}b_{sh.i} - b_{ij}g_{sh.i})L_{ij} + \varepsilon_{I2}. \quad (31)$$

B. Intermediate Nonlinear Transformation

The intermediate vector u, composed like y of $2b+N$ variables, is defined as follows:

- *Bus-related variable*. For each bus i, the following variable is adopted:

$$\alpha_i = \ln U_i = 2\ln V_i \quad \Rightarrow \quad V_i^2 = \exp\alpha_i. \quad (32)$$

  Note that, unlike V_i, which is positive by definition and close to 1, α_i can take a (usually small) positive or negative value, much like the conventional phase angle θ_i. Such a logarithmic version of the voltage magnitude was used just for observability purposes in [19].
- *Branch-related variables*. For each branch between buses i and j, a pair of variables is introduced:

$$\theta_{ij} = \theta_i - \theta_j \quad (33)$$
$$\alpha_{ij} = \alpha_i + \alpha_j. \quad (34)$$

  A key point in the way α_{ij} is defined lies in the use of the "+" sign, unlike the "−" sign adopted in [19].

Then, the components of vector u

$$u = \{\alpha_i, \alpha_{ij}, \theta_{ij}\} \quad (35)$$

can be explicitly related to those of vector y as follows:

$$\alpha_i = \ln U_i \quad (36)$$
$$\alpha_{ij} = \ln(K_{ij}^2 + L_{ij}^2) \quad (37)$$
$$\theta_{ij} = \arctan\left(\frac{L_{ij}}{K_{ij}}\right). \quad (38)$$

The above relationships (36)–(38) constitute the nonlinear transformation (18) for the particular set of variables adopted in this paper. From those expressions, it is straightforward to derive the $(2b+N)\times(2b+N)$ Jacobian F_u, composed of N diagonal scalars plus $2b$ diagonal blocks (size 2×2), as follows:

$$\frac{\partial\alpha_i}{\partial U_i} = \frac{1}{U_i} \quad (39)$$
$$\begin{bmatrix} \frac{\partial\alpha_{ij}}{\partial K_{ij}} & \frac{\partial\alpha_{ij}}{\partial L_{ij}} \\ \frac{\partial\theta_{ij}}{\partial K_{ij}} & \frac{\partial\theta_{ij}}{\partial L_{ij}} \end{bmatrix} = \frac{1}{K_{ij}^2+L_{ij}^2}\begin{bmatrix} 2K_{ij} & 2L_{ij} \\ -L_{ij} & K_{ij} \end{bmatrix} \quad (40)$$

the inverse of which are needed to compute W_u, in accordance to (13)

$$\left[\frac{\partial\alpha_i}{\partial U_i}\right]^{-1} = U_i \quad (41)$$
$$\begin{bmatrix} \frac{\partial\alpha_{ij}}{\partial K_{ij}} & \frac{\partial\alpha_{ij}}{\partial L_{ij}} \\ \frac{\partial\theta_{ij}}{\partial K_{ij}} & \frac{\partial\theta_{ij}}{\partial L_{ij}} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{K_{ij}}{2} & -L_{ij} \\ \frac{L_{ij}}{2} & K_{ij} \end{bmatrix}. \quad (42)$$

Therefore, computing F_u^{-1} involves virtually no computations, which is a nice added feature of the proposed nonlinear transformation.

C. Second Linear Stage

The final stage consists of linearly relating vector u to a slightly modified state vector x. For this purpose, let us express u and x in blocked form as follows:

$$u = \begin{bmatrix} \alpha \\ \alpha_b \\ \theta_b \end{bmatrix} \quad ; \quad x = \begin{bmatrix} \alpha \\ \theta \end{bmatrix}$$

where the subindex "b" refers to the set of branch variables. Note that, for convenience, the logarithmic version α replaces the ordinary voltage magnitude V in x. Once α is available V can be readily computed if needed.

Then, the branch components of u can be related to the nodal components of x by means of

$$\theta_b = A_r^T \theta \tag{43}$$

$$\alpha_b = |A^T| \alpha \tag{44}$$

where A represents the well-known branch-to-node incidence matrix and A_r is the reduced matrix obtained by omitting the reference phase angle in A.

In blocked form, the linear measurement model (19) can be expressed as follows:

$$\begin{bmatrix} \alpha \\ \alpha_b \\ \theta_b \end{bmatrix} = \underbrace{\left[\begin{array}{c|c} I & 0 \\ |A^T| & 0 \\ \hline 0 & A_r^T \end{array}\right]}_{C} \begin{bmatrix} \alpha \\ \theta \end{bmatrix} + e_u. \tag{45}$$

It is worth noting that a decoupled linear model arises at this stage, characterized by a trivial matrix C composed of 0's and 1's. Furthermore, this matrix is just a function of the network topology, which seldom changes.

D. Simplifications for Repeated Solutions

The bilinear SE requires that the two linear systems (7) and (15) be solved. In the general case, solving a linear system involves the following main steps: 1) creating the coefficient matrix; 2) node ordering and symbolic factorization to determine the resulting sparsity structure (fill-ins) of the table of factors; 3) actual numerical factorization; 4) building the independent vector; 5) forward/backward solutions on the independent vector.

However, the elements of the gain matrix in (7) must be updated only when any of the following events takes place: a) there are topology changes, b) the composition of the measurement set is modified, or c) any parameter value (e.g., tap changer) is altered, which happens only occasionally. When solving (7) during repeated executions of the SE, neither B nor the table of factors of the gain matrix G_b have to be recomputed unless any of those modifications occurs. Therefore, the successive solutions of (7) involve in those cases only step 4, which reduces to a sparse matrix-vector product if $B^T W$ was saved, and step 5, significantly reducing the computational cost of the first linear filter.

This saving is not possible when solving (15), since the gain matrix G_c involves, in addition to the trivial constant matrix C, the weighting matrix W_u, which is a function of the operating point.

V. Test Results

In this section, test results corresponding to the IEEE 118- and 298-bus benchmark systems are presented and discussed. Both the conventional and the proposed SE schemes are coded in Matlab, taking full advantage of available sparse matrix capabilities. The main objective is to assess the potential benefits of the proposed bilinear scheme, for different redundancy and measurement noise levels, in terms of overall solution accuracy and computational effort. The reader is referred to [16], where additional results are provided regarding the accuracy of the first linear stage as well as the possibility of detecting and identifying bad data after this preliminary step.

A. Accuracy Analysis

From exact load flow solutions, several measurement sets with varying noise and redundancy levels are simulated. For the results to be more representative, 1000 measurement sets are randomly generated for each scenario.

Different error levels are simulated as follows: for voltage measurements, two values of σ_V are considered (0.002 and 0.01). In turn, for each value of σ_V, two increasing values of s.d. for power measurements are tested ($\sigma_P = \rho\sigma_V$, with $\rho = 2$ and $\rho = 5$).

For each simulated error, two measurement sets are considered, as follows:

- low redundancy case (L): voltage magnitudes at all buses, power flows across all branches ("from" terminal only), power injections at one half of buses (one every other bus in the natural sequence). In this case, $m = 2(b + n)$.
- high redundancy level (H): voltage magnitudes measured twice for every bus, power flows at both terminal nodes of every branch, power injections at all buses $(m = 4(b+n))$.

For the 118- and 298-bus systems, this leads to redundancy ratios ranging from 2.58 to 5.16 and 2.37 to 4.74, respectively. In practice, the redundancy found in bulk transmission systems lies in between (for instance, the redundancy level for the Spanish case is currently 3.6).

The standard WLS methodology, based on the Gauss-Newton iterative method, is applied to each scenario, yielding a number of iterations ranging typically from 4 to 6 for a convergence threshold of $1E{-}5$ (applied to the largest absolute value change in any state vector component). Then, a single run of the proposed three-stage procedure is also performed for each test case.

In order to compare the accuracy of the solution provided by the proposed procedure (single run) with that of the conventional solution approach, the averages of the absolute errors associated with the estimates of state variables are computed, as follows:

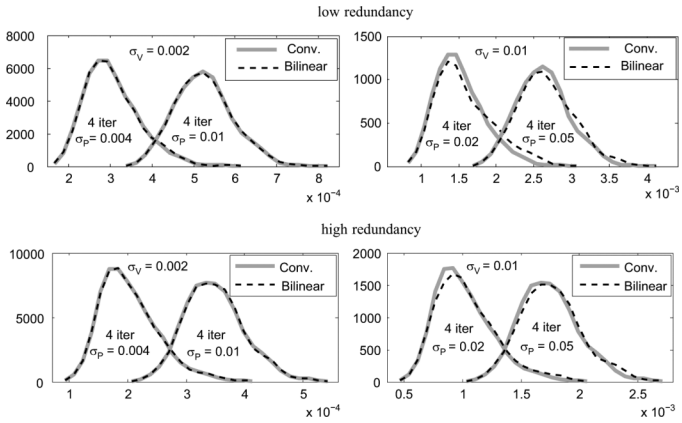


Fig. 1. Pdf of S_V values for the 118-bus system.

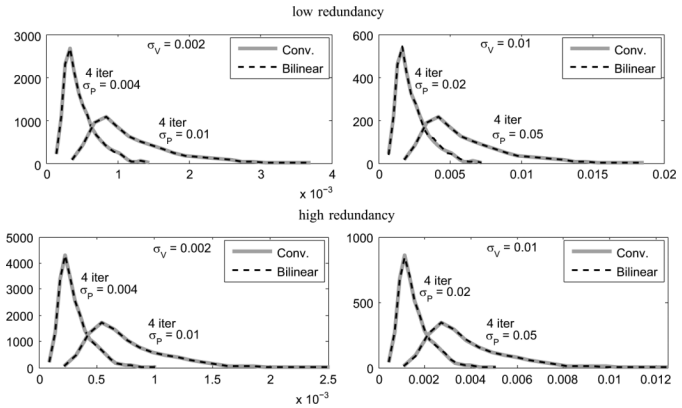


Fig. 2. Pdf of S_θ values for the 118-bus system.

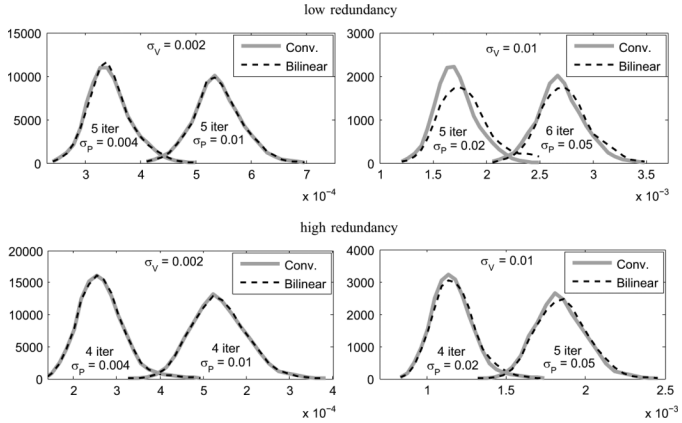


Fig. 3. Pdf of S_V values for the 298-bus system.

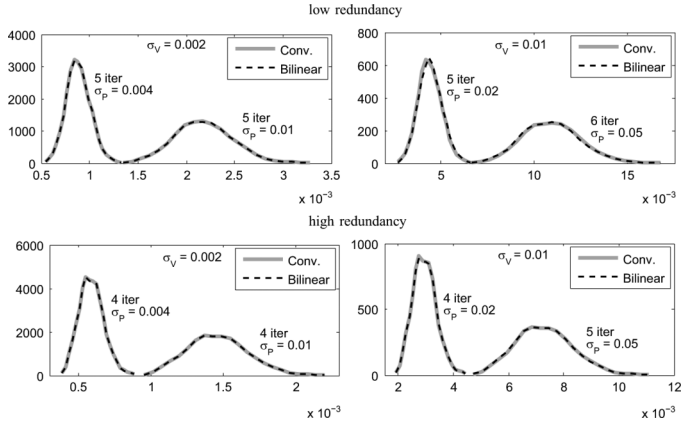


Fig. 4. Pdf of S_θ values for the 298-bus system.

$$S_V = \frac{1}{N}\sum_{i=1}^{N} |\hat{V}_i - V_i^{\text{ex}}|;$$

$$S_\theta = \frac{1}{N-1}\sum_{i=1}^{N-1} |\hat{\theta}_i - \theta_i^{\text{ex}}|$$

where V^{ex} and θ^{ex} represent the exact values.

Figs. 1 and 2 present, for the 118-bus network, the probability density functions (pdf) corresponding to the 1000 S_V and S_θ values, respectively, after full convergence of the conventional method (gray continuous line) and after a single execution of the bilinear estimator (dashed black line). Close to each pdf, representing a unique combination of measurement noise and redundancy level, the minimum number of iterations required by the standard Gauss-Newton approach to converge is shown (the average number of iterations for the entire set of 1000 simulations is slightly higher, as there is always a small percentage of anomalous cases requiring more iterations). Figs. 3 and 4 provide the same information for the 298-bus system.

By definition, the area under a pdf along the entire x axis is 1. Therefore, narrower pdfs give rise to proportionally higher values on the y axis. The following conclusions are drawn from those figures:

- Wider pdfs with larger expected values arise for S_θ, probably due to the fact that phase angle measurements are not considered.
- As expected, the average values of S_V and S_θ are higher for lower redundancy levels and larger measurement errors. Sometimes (see for instance the 298-bus case) an extra iteration is required by the conventional method for very noisy measurements.
- For highly redundant measurement systems, the resulting pdf for the proposed method virtually matches that of the Gauss-Newton method, particularly regarding S_θ. Very small differences can only be noticed for the pdfs of S_V in scenarios with low redundancy levels and very large measurement errors.

It has been found that the convergence speed of the Gauss-Newton method is somewhat affected by the network loading level, as a consequence of the model nonlinearities and distance of the flat start to the solution point. In order to assess this issue, new scenarios are created by doubling the net power at all PQ buses of the 118-bus system. Accordingly, the active power of PV buses is doubled and the reactive power limits are not enforced, allowing the generator voltages to remain at the scheduled value (the average bus voltage, however, reduces from 0.986 to 0.908).

Fig. 5 represents the pdfs corresponding to S_V and S_θ for these new scenarios, referred to in the sequel as 118–2P (due to

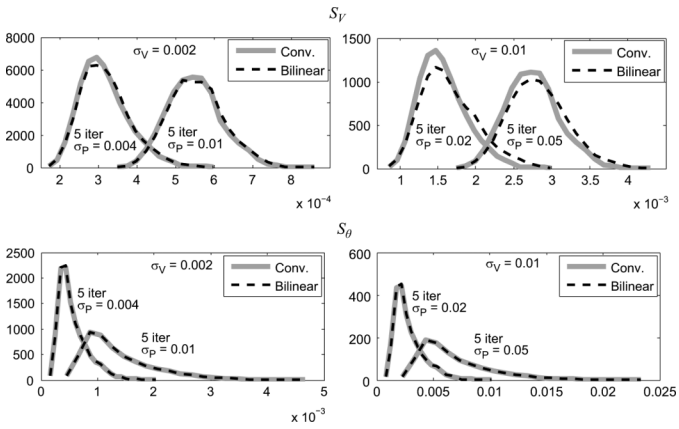


Fig. 5. Pdf of S_V and S_θ values for the 118–2P system (low redundancy cases).

space limitations only the worst cases with low measurement redundancy are shown). From a comparison between Figs. 1 and 2 (118) and Fig. 5 (118–2P), the following conclusions are reached:

- The conventional scheme requires at least an extra iteration in all cases compared to the base case. Note that, owing perhaps to the presence of the power flow measurements, and the fact that the generator voltages are kept artificially high, the convergence deterioration is anyway not so significant in the SE as in the load flow case.
- On the other hand, the bilinear scheme remains virtually unaffected by the network loading condition, which may be good in critical situations, close to voltage collapse. This positive behavior can be explained by the linearity of the first stage, providing quite a good estimate for accurate enough measurements irrespective of the load (i.e., nonlinearity) level.

The values of the objective function J in (2) are also computed, both for the fully converged solution of the conventional approach and after a single execution of the bilinear scheme. The average values of J/m, for the 1000 trials corresponding to each scenario, are given in Table II. The table also shows (rightmost column) the maximum difference between the diagonal elements of $\mathrm{cov}(\hat{x})$ provided by the conventional method (G_k^{-1}) and the proposed one (G_c^{-1}). As can be seen, the differences between both approaches are very small.

It is also interesting to check the condition number of the involved gain matrices, as this may affect the numerical performance of any solution method. Table III shows the resulting condition numbers for the conventional method (matrix G_k) as well as those associated with the two linear systems arising in the proposed method (matrices G_b and G_c). Note that these are always approximate values provided by the Matlab function "cond". While $\mathrm{cond}(G_b)$ tends to be consistently the smallest one, the relatively small discrepancies between $\mathrm{cond}(G_k)$ and $\mathrm{cond}(G_c)$ are due to round-off errors and the different linearization points to which both matrices refer.

In summary, the results presented so far confirm that the three-stage procedure proposed in this paper provides solutions which, for practical purposes, are barely distinguishable from the optimal ones. What is more important, this is achieved in a non-iterative fashion.

Should much more exigent accuracy levels be needed and/or unrealistically poor measurements be used, the second linear stage could be repeated by previously updating the Jacobian of the nonlinear transformation, as explained in Section III. Fig. 6 compares, for the low redundancy scenarios with $\sigma_V = 0.01$ and $\sigma_P = 0.02$, the pdf of the S_V values provided by a single run of the three-stage procedure (already shown in the previous figures) with that resulting after an extra execution of the second linear estimator. As can be seen, the pdfs of the

TABLE II
AVERAGE VALUES OF J/m AND MAXIMUM DIFFERENCES OF COVARIANCE MATRIX ELEMENTS

Case	σ_V	ρ	J/m Conv.	J/m Bilin.	J/m Diff.	Cov. matrix Max. Diff.
118(L)	0.002	2	0.6189	0.6191	1.27E-04	6.67E-10
	0.002	5	0.6187	0.6188	9.72E-05	1.49E-09
	0.01	2	0.6189	0.6216	2.69E-03	8.88E-08
	0.01	5	0.6187	0.6200	1.23E-03	1.96E-07
118(H)	0.002	2	0.8085	0.8085	1.46E-05	2.81E-10
	0.002	5	0.8084	0.8085	2.87E-05	1.38E-09
	0.01	2	0.8085	0.8086	1.52E-04	3.68E-08
	0.01	5	0.8084	0.8086	1.80E-04	1.74E-07
298(L)	0.002	2	0.5804	0.5807	3.11E-04	1.03E-07
	0.002	5	0.5804	0.5806	1.61E-04	6.22E-07
	0.01	2	0.5804	0.5869	6.53E-03	2.57E-06
	0.01	5	0.5804	0.5826	2.23E-03	1.11E-04
298(H)	0.002	2	0.7909	0.7909	2.58E-05	3.79E-08
	0.002	5	0.7907	0.7907	4.41E-05	2.29E-07
	0.01	2	0.7909	0.7910	1.04E-04	9.47E-07
	0.01	5	0.7907	0.7909	1.61E-04	8.25E-06
118-2P(L)	0.002	2	0.6189	0.6195	6.03E-04	7.22E-10
	0.002	5	0.6187	0.6200	1.31E-03	1.56E-09
	0.01	2	0.6189	0.6225	3.63E-03	9.64E-08
	0.01	5	0.6187	0.6215	2.83E-03	2.27E-07
118-2P(H)	0.002	2	0.8084	0.8086	2.25E-04	2.62E-10
	0.002	5	0.8084	0.8090	6.02E-04	1.38E-09
	0.01	2	0.8084	0.8088	3.94E-04	3.41E-08
	0.01	5	0.8084	0.8092	8.44E-04	1.74E-07

TABLE III
CONDITION NUMBERS

Case	σ_V	ρ	Condition Number $\mathrm{cond}(G_k)$	$\mathrm{cond}(G_b)$	$\mathrm{cond}(G_c)$
118(L)	0.002	2	8.13E+05	1.49E+06	8.28E+05
	0.002	5	1.71E+06	2.76E+05	7.85E+05
	0.01	2	7.02E+06	1.50E+06	8.04E+05
	0.01	5	4.26E+07	2.77E+05	7.59E+05
118(H)	0.002	2	4.46E+05	5.25E+05	4.50E+05
	0.002	5	7.20E+05	8.73E+04	4.26E+05
	0.01	2	3.05E+06	5.24E+05	4.50E+05
	0.01	5	1.80E+07	8.71E+04	4.27E+05
298(L)	0.002	2	1.02E+09	4.91E+08	9.86E+08
	0.002	5	1.02E+09	4.81E+08	9.49E+08
	0.01	2	1.01E+09	4.92E+08	9.47E+08
	0.01	5	1.00E+09	4.81E+08	1.19E+09
298(H)	0.002	2	4.58E+08	2.69E+08	4.78E+08
	0.002	5	4.57E+08	2.64E+08	5.09E+08
	0.01	2	4.56E+08	2.69E+08	5.71E+08
	0.01	5	4.55E+08	2.64E+08	8.35E+08
118-2P(L)	0.002	2	1.10E+06	1.41E+06	1.14E+06
	0.002	5	2.20E+06	2.58E+05	1.02E+06
	0.01	2	9.44E+06	1.41E+06	1.11E+06
	0.01	5	5.50E+07	2.59E+05	9.88E+05
118-2P(H)	0.002	2	6.43E+05	4.43E+05	6.00E+05
	0.002	5	9.34E+05	7.74E+04	5.55E+05
	0.01	2	4.55E+06	4.42E+05	6.00E+05
	0.01	5	2.33E+07	7.72E+04	5.56E+05

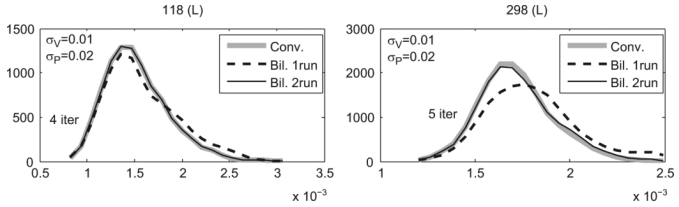


Fig. 6. Pdf of S_V arising after a single run of the proposed method and after a second execution of the second linear estimator.

proposed method perfectly match those of the conventional solution after the second run.

B. Computational Cost

The conventional WLS iterative scheme, the two-stage factorized solution of [16], and the three-stage solution approach proposed in this paper have been coded in Matlab (version R2008 a) and run under Windows 7 on a 64-bit i5 Intel Core laptop (2.27 GHz, 4 GB of RAM). Every effort has been made to fully optimize all codes. Specifically, in addition to the built-in Matlab capability to handle sparse matrices, the following key aspects have been born in mind:

- In the conventional formulation, the sine and cosine functions are computed only once for each branch and conveniently stored. Then, they are systematically used to compute both the related Jacobian elements and power flow measurements. Power injections are obtained by adding the necessary power flows, already available. The first iteration is separately coded, as the flat start involves neither trigonometric functions nor multiplications by voltage magnitudes. Furthermore, when computing the gain matrix, the partial result $H^T W$ is saved for the computation of the right-hand side vector.
- In the proposed three-stage procedure, similar considerations apply, the main difference being that expensive functions arise only in the nonlinear transform given by (36)–(38). As explained in Section IV-D, the first linear stage is computationally less expensive for repeated solutions, in which only the measurement values get modified. Therefore, both a "cold" and a "hot" start are separately considered.
- The above implementation details apply in part to the existing factorized scheme [16], taking into account that the first linear stage is identical to that of this paper and that the second nonlinear phase also involves trigonometric functions but a reduced "measurement" model.

As a result, very low solution times have been achieved for the software and hardware platform adopted. Table IV shows the average solution times for the sets of scenarios with $\sigma_V = 0.002$ and $\sigma_P = 2\sigma_V$, including both redundancy levels. The following data are collected:

- Individual solution times for each of the three direct phases involved in the proposed approach. For the first linear phase, the "cold" and "hot" starts are separately considered.
- Additional solution time eventually required by running once more the last two phases (2nd run).

TABLE IV
AVERAGE SOLUTION TIMES CORRESPONDING TO DIFFERENT STAGES OF THE SOLUTION PROCESSES AND RESULTING SPEEDUPS

TIMES (msec)		Low redundancy			High redundancy		
		118	298	118-2P	118	298	118-2P
1st Lin. Phase	cold	10.75	25.77	10.75	17.28	42.53	17.28
	hot	0.14	0.25	0.15	0.22	0.41	0.22
Nonlin. Trans.		1.04	2.21	1.05	1.18	2.42	1.18
2nd Lin. Phase		1.86	6.61	1.85	2.38	7.79	2.37
2nd run		2.90	8.81	2.90	3.56	10.20	3.55
Stage 2 [16] + init.		20.92	56.61	20.93	22.44	60.41	22.42
Conventional		40.69	126.39	50.23	70.72	175.23	87.61

SPEEDUPS	Low redundancy			High redundancy		
	118	298	118-2P	118	298	118-2P
Single run (cold)	2.98	3.65	3.68	3.39	3.32	4.21
Single run (hot)	13.37	13.93	16.50	18.69	16.49	23.24
Two runs (cold)	2.46	2.91	3.04	2.90	2.78	3.60
Two runs (hot)	6.85	7.07	8.45	9.63	8.41	11.98
Factor. [16] (cold)	1.29	1.53	1.59	1.78	1.70	2.21
Factor. [16] (hot)	1.93	2.22	2.39	3.13	2.88	3.87

- Solution time involved in the iterative solution of the second stage proposed in [16], including the nontrivial initialization of the state vector.
- Solution time associated with the fully converged conventional solution.

The speedup, defined as T_c/T_a, where T_c and T_a represent the solution times of the conventional and the alternative factorized methodologies, respectively, has been computed for each case and included also in Table IV. The following comments are in order regarding the results obtained:

- For a single cold-start execution of the proposed bilinear scheme, the average speedup ranges from 3 to 4.2, higher values being obtained in general for larger networks, peak loading conditions, and more redundant measurement sets.
- Amazingly high speedup values, ranging from 13 to 23, are obtained for the hot-start execution. This can be explained as follows: the first linear step is by far the most expensive one in the proposed three-stage methodology, owing to the fact that all raw measurements and an augmented state vector are handled during this step. Indeed, both the nonlinear transformation and the inverse of the associated Jacobian are trivially carried out, whereas computing the gain matrix of the final linear step involves only sums, since C is composed of 1's and -1's. Furthermore, as discussed above, for repeated solutions (hot start) it is possible to skip a majority of computations associated with the first linear step, provided they were saved during the cold-start solution.
- For the same reason, even if the last two phases (Jacobian of the nonlinear transformation and the second linear phase), are repeated in order to obtain much more accurate solutions, the resulting speedups are still very large (from 2.5 to 3.6 and from 6.9 to 12 for the cold- and hot-start cases, respectively).
- As expected, the proposed three-stage procedure is computationally much more efficient than the former factorized scheme [16].

VI. CONCLUSIONS

In this paper, a three-stage state estimation methodology, arising from a factorization of the conventional WLS nonlinear

model, is proposed. In the first stage, raw measurements are processed by an augmented-state linear estimator, for which convergence is never an issue. Then, a nonlinear but explicit transform is applied to this initial estimate, the result of which is finally used by another reduced-order linear estimator. It turns out that the Jacobian of the nonlinear transform, which is needed to compute the weighting matrix of the final linear stage, has a trivial inverse. Moreover, this latter filter involves only the well-known incidence matrix, topologically relating branch variables to the nodal ones.

Experimental results on IEEE benchmark networks show that the proposed three-stage approach yields, in a single run, virtually the same solution as that provided by the Gauss-Newton iterative scheme. Only in the presence of extremely high measurement errors and very low redundancy levels might a second run of the last filter be justified, for which the Jacobian of the nonlinear transform should be previously updated. Accordingly, the resulting solution times are between three to four times smaller than those of the conventional iterative method. In tracking mode, over an order of magnitude savings have been recorded.

The only potential limitation of the proposed scheme is the risk for the augmented state vector introduced in the first stage to be unobservable, which is very unlikely considering the redundancy levels nowadays. But even if this happened, the proposed methodology could be applied to those areas for which the intermediate state vector is observable.

Future efforts will be directed to extend the proposed methodology so that equality constraints can be explicitly handled.

REFERENCES

[1] F. C. Schweppe, J. Wildes, and D. B. Rom, "Power system static state estimation, part I, II, III," *IEEE Trans. Power App. Syst*, vol. PAS-89, no. 1, pp. 120–135, Jan. 1970.

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

[3] A. Monticelli, *State Estimation in Electric Power System. A Generalized Approach*. Norwell, MA: Kluwer, 1999.

[4] W. Tinney and R. Walker, "Direct solutions of sparse network equations by optimally ordered triangular factorization," *Proc. IEEE*, vol. 55, no. 11, pp. 1801–1809, Nov. 1967.

[5] A. Garcia, A. Monticelli, and P. Abreu, "Fast decoupled state estimation and bad data processing," *IEEE Trans. Power App. Syst*, vol. PAS-98, no. 5, pp. 1645–1652, 1979.

[6] A. Simoes-Costa and V. Quintana, "A robust numerical technique for power system state estimation," *IEEE Trans. Power App. Syst*, vol. PAS-100, pp. 691–698, Feb. 1981.

[7] M. Vempati, I. Slutsker, and W. Tinney, "Enhancements to givens rotations for power system state estimation," *IEEE Trans. Power Syst.*, vol. 6, no. 2, pp. 842–849, May 1991.

[8] F. Aschmoneit, N. Peterson, and E. Adrian, "State estimation with equality constraints," in *Proc. 10th PICA Conf.*, Toronto, ON, Canada, May 1977, pp. 427–430.

[9] A. Gjelsvik, S. Aam, and L. Holten, "Hachtel's augmented matrix method—a rapid method improving numerical stability in power system static state estimation," *IEEE Trans. Power App. Syst*, vol. PAS-104, pp. 2987–2993, Nov. 1985.

[10] M. R. Irving, R. C. Owen, and M. J. H. Sterling, "Power system state estimation using linear programming," *Proc. Inst. Elect. Eng., Part C*, vol. 125, no. 9, pp. 879–885, Sep. 1978.

[11] M. K. Celik and A. Abur, "Robust state estimation of electric power systems," *IEEE Trans. Power Syst.*, vol. 7, no. 1, pp. 106–113, Feb. 1992.

[12] L. Mili, M. G. Cheniac, and P. J. Rousseeuw, "A robust WLAV state estimator using transformations," *IEEE Trans. Circuits Syst.*, vol. 41, no. 5, pp. 349–358, May 1994.

[13] A. de la Villa Jaén, P. C. Romero, and A. Gómez-Expósito, "Substation data validation by local three-phase generalized state estimators," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 264–271, Feb. 2005.

[14] W. Jiang, V. Vittal, and G. T. Heydt, "A distributed state estimator utilizing synchronized phasor measurements," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 563–571, May 2007.

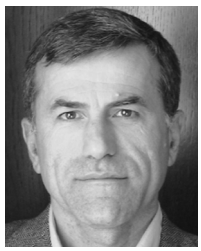
[15] A. Gómez-Expósito and A. de la Villa, "Two-level state estimation with local measurement pre-processing," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 676–684, May 2009.

[16] C. Gómez-Quiles, A. de la Villa, and A. Gómez-Expósito, "A factorized approach to WLS state estimation," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1724–1732, Aug. 2011.

[17] A. Monticelli and A. Garcia, "Reliable bad data processing for real-time state estimation," *IEEE Trans. Power App. Syst.*, vol. PAS-102, no. 5, pp. 1126–1139, May 1983.

[18] A. Gómez-Expósito and E. Romero, "Reliable load flow technique for radial distribution networks," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 1063–1069, Aug. 1999.

[19] A. Gómez-Expósito and A. Abur, "Generalized observability analysis and measurement classification," *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 1090–1096, Aug. 1998.



Antonio Gómez-Expósito (F'05) received the electrical engineering and Ph.D. degrees from the University of Seville, Seville, Spain.

Since 1982, he has been with the Department of Electrical Engineering, University of Seville, where he is currently a Professor and leader of the Power Engineering group. He also holds the Endesa Red Chair. His primary areas of interest are optimal power system operation, state estimation, digital signal processing, and control of flexible ac transmission system devices.

Catalina Gómez-Quiles (S'09) received the Eng. degree from the University of Seville, Seville, Spain, in 2006 and the M.Eng. degree from McGill University, Montreal, QC, Canada, in 2008, both in electrical engineering. She is currently pursuing the Ph.D. degree at the University of Seville.

Her research interests include state estimation and risk assessment in competitive electricity markets.

Antonio de la Villa Jaén was born in Riotinto, Spain, in 1960. He received the electrical engineering and Ph.D. degrees from the University of Seville, Seville, Spain.

He is presently an Associate Professor with the Department of Electrical Engineering, University of Seville. His primary areas of interest are computer methods for power system state estimation problems and wave energy converter control.

Equality-Constrained Bilinear State Estimation

Catalina Gomez-Quiles, Hugo A. Gil, Antonio de la Villa Jaén and Antonio Gómez-Expósito, *Fellow, IEEE*

***Abstract*—This paper generalizes the recently introduced bilinear formulation of the WLS state estimation problem to those cases in which equality constraints must be explicitly considered. This leads to augmented equation systems in which Lagrange multipliers and the resulting covariance matrices get updated throughout the three stages arising in the bilinear approach (two linear filters with a nonlinear transformation in between). The proposed formulation prevents the ill-conditioning typically arising when exact-injection constraints are handled as virtual measurements with huge weights, while the excellent convergence speed of the bilinear scheme, which for practical purposes reaches the optimal solution in a single iteration, is fully preserved.**

***Index Terms*—Factorized state estimation, linear WLS, equality constraints.**

I. Introduction

THE goal of a State Estimator (SE) is to process the redundant set of noisy measurements which are remotely captured via Remote Terminal Units (RTU) and collected periodically in Energy Management Systems, where several crucial applications rely on accurate system snapshots [1].

Given a set of independent measurements with normally (Gaussian) distributed errors, the maximum likelihood state is obtained by solving a nonlinear Weighted Least Squares (WLS) problem. In turn, this leads to an iterative process in which the so-called Normal equations are repeatedly solved [2].

In addition to handling regular measurements, the capability to explicitly incorporate equality constraints adds substantial value to a SE for the following main reasons:

- Null injections, which are known with certainty, frequently arise in transmission grids and have to be accounted for in the SE. The traditional strategy of using huge weights, as if they were very accurate measurements, significantly deteriorates the condition number of the Normal equations, thus creating a mathematical hurdle in their solution. As described in [3]–[6], considering exact injection values as equality constraints improves considerably the stability and accuracy of the solution.
- The presence of Circuit Breakers (CB) and the need to handle the associated power flow measurements led to their inclusion into the SE as branches with very small impedance. Plugging in these nearly zero impedance elements into the Jacobian of the WLS optimization contributed to the surge of numerical problems. An effective way-around was proposed in [7], in which CBs are modeled via additional state variables with an accompanying set of equality constraints representing the discrete on-off status of a CB. However, this strategy leads to a substantially large WLS model when topology errors are to be systematically checked at all substations. As a response to this challenge, [8] shows how trivial constraints like those associated to CBs may be implicitly incorporated into the WLS optimization.
- The same applies when abnormally large parameter errors are to be identified. An equality-constrained implicit model has been also proposed in [9] for this purpose.

A bilinear SE (BSE) formulation has been recently introduced in [11] as a valuable alternative to the conventional SE (CSE) based on the well-known Gauss-Newton (GN) iterative scheme. Instead of performing successive linearizations of the first-order optimality conditions, the original WLS nonlinear model is decomposed into the following three stages with the help of auxiliary variables: 1) an augmented-state linear estimator to reliably filter raw measurements; 2) a trivial and explicit nonlinear transform of the initial estimate (change of variables); 3) a reduced-order linear estimator providing the estimate of the conventional state vector. For measurement errors well beyond those typically encountered nowadays, a single run of the above procedure provides virtually optimal results, leading to terrific computational savings when compared to the GN iterative solution.

In this paper, the unconstrained BSE proposed in [11] is extended to the case in which equality constraints must be explicitly considered. It is shown that the three steps arising in the bilinear approach can be easily generalized by augmenting the resulting equation systems with the required Lagrange multipliers, much like in the conventional GN-based equality-constrained solution. A relevant feature of the resulting scheme is that the associated weighting matrices remain sparse throughout the solution process.

The structure of this paper is as follows: Section II reviews the conventional GN-based and the bilinear WLS SE formulations. Section III presents the proposed equality-constrained BSE. Finally, test results, based on IEEE benchmark systems, are presented in Section V, followed by the concluding remarks and an Appendix.

II. Power System State Estimation

For the sake of self-sufficiency, both the unconstrained and the equality-constrained WLS SE formulations are reviewed below. Also, the unconstrained BSE is summarized. The reader is referred to [11] for further details about this new solution approach.

A. Unconstrained WLS State Estimation

In absence of equality constraints, the measurement model is [2]:

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

Authors are with the Department of Electrical Engineering, University of Seville, Seville, Spain. e-mail: age@us.es. This work was supported by the Spanish Ministry of Science and Innovation, under grant ENE2010-18867 and the Ramon y Cajal research incentive program.

where:

- x is the state vector to be estimated, composed of bus voltage magnitudes, V_i, and phase angles, θ_i (size $n = 2N - 1$, N being the number of buses),
- z is the measurement vector (size $m > n$),
- h is the vector of functions relating error free measurements to the state variables,
- e is the vector of measurement errors, customarily assumed to be $N(0, \sigma)$ and uncorrelated, with covariance matrix $\mathrm{cov}(e) = R = \mathrm{diag}(\sigma_i^2)$.

Under this assumption, the Weighted Least Squares (WLS) estimator provides the maximum likelihood estimate $\hat{x}$ by repeatedly solving the Normal equations until Δx_k is below an appropriate tolerance:

$$G_k \Delta x_k = H_k^T W \Delta z_k \tag{2}$$

where:

$W = R^{-1}$ is the weighting matrix
H_k is the Jacobian of $h(x)$ evaluated at $x = x_k$,
$G_k = H_k^T W H_k$ is the gain matrix,
$\Delta z_k = z - h(x_k)$
$\Delta x_k = x_{k+1} - x_k$, k being the iteration counter.

The covariance of the estimate can be approximately obtained from:

$$\mathrm{cov}(\hat{x}) = G_l^{-1}$$

where l denotes the last iteration.

B. Equality-Constrained WLS State Estimation

In practice, exactly known magnitudes, such as zero injections, must be accommodated by WLS estimators, which is sometimes done by considering them as very accurate measurements with arbitrarily large weights. However, the coexistence of extremely uneven weights may lead to convergence problems due to ill-conditioning of the gain matrix [2]. The equality-constrained formulation avoids the use of high weights by explicitly adding the necessary constraints to the WLS estimation model [3], [5], which becomes:

$$\begin{aligned} z &= h(x) + e \quad (3)\\ b &= h_e(x) \quad (4) \end{aligned}$$

where $b = 0$ for null injections.

Then, the associated optimization problem can be formulated as:

$$\begin{aligned} \text{minimize} \quad & J(x) = \frac{1}{2}[z - h(x)]^T W [z - h(x)] \quad (5)\\ \text{subject to} \quad & h_e(x) = b \end{aligned}$$

By applying the GN method to the first-order optimality conditions of the resulting Lagrangian function, the solution of the constrained optimization problem is obtained by repeatedly solving the following system:

$$\begin{bmatrix} \beta H^T W H & H_e^T \\ H_e & 0 \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \lambda \end{bmatrix} = \begin{bmatrix} \beta H^T W \Delta z_k \\ b - h_e(x_k) \end{bmatrix} \tag{6}$$

where λ is the vector of Lagrange multipliers and β is a scaling factor which can be introduced to improve the condition number of the coefficient matrix. According to [2], Chapter 3, a reasonable value for β is,

$$\beta^{-1} = \max W_{ii} \tag{7}$$

The main shortcoming of the above system is that the coefficient matrix is indefinite. However, efficient solvers exist fully exploiting the resulting matrix sparsity [13]–[15].

It is worth stressing, as explained in [2], page 48, that the equality-constrained formulation can be alternatively developed by considering a WLS problem with two sets of measurements, one of them overweighted by a factor ρ arbitrarily large with respect to the other, and then letting $\rho \to \infty$. Considering the equality-constrained formulation as the limiting case of a regular WLS problem is theoretically appealing, since it allows well-known statistical properties to be extrapolated to this particular condition (e.g., the interpretation and use of normalized Lagrange multipliers for bad data analysis).

C. Bilinear WLS State Estimation

In the BSE the nonlinear measurement model (1) is replaced by two linear models, which become coupled through a nonlinear change of variables, as follows [11]:

$$\begin{aligned} z &= By + e \quad (8)\\ u &= f(y) \quad (9)\\ u &= Cx + e_u \quad (10) \end{aligned}$$

In the above factorized model the following auxiliary vectors are introduced:

- Intermediate state vector y,

$$y = \{U_i, K_{ij}, L_{ij}\} \tag{11}$$

  where, for each branch connecting buses i and j the following pair of variables is adopted:

$$\begin{aligned} K_{ij} &= V_i V_j \cos\theta_{ij} \quad (12)\\ L_{ij} &= V_i V_j \sin\theta_{ij} \quad (13) \end{aligned}$$

  and, for each bus i, squared voltage magnitudes rather than plain voltage magnitudes are retained,

$$U_i = V_i^2 \tag{14}$$

  Vector y then comprises $2b + N$ variables (b being the number of branches).
- Pseudo-measurement vector u, composed also of $2b + N$ variables,

$$u = \{\alpha_i, \alpha_{ij}, \theta_{ij}\} \tag{15}$$

  where,
  $\alpha_i = \ln U_i = 2 \ln V_i$
  $\alpha_{ij} = \alpha_i + \alpha_j$
  $\theta_{ij} = \theta_i - \theta_j$.

The nonlinear functions $f(\cdot)$, as well as the constant matrices B and C, are easily obtained from the above definitions (see [11] for the details).

The solution process comprises the following three steps:

1) Compute $\tilde{y}$ by solving the linear system,

$$G_B \tilde{y} = B^T W z \tag{16}$$

where

$$G_B = B^T W B \tag{17}$$

is the associated gain matrix.

2) Obtain $\tilde{u}$ from

$$\tilde{u} = f(\tilde{y}) \tag{18}$$

and the weighting matrix,

$$\tilde{W}_u = \mathrm{cov}^{-1}(\tilde{u}) = \tilde{F}^{-T} G_B \tilde{F}^{-1} \tag{19}$$

where $\tilde{F}$ is the Jacobian of $f(\cdot)$, composed of scalar or 2×2 diagonal blocks with trivial inverses. In fact, $\tilde{F}^{-1}$ is directly obtained at virtually no cost [11].

3) Compute $\hat{x}$ by solving the linear system,

$$G_C \hat{x} = C^T \tilde{W}_u \tilde{u} \tag{20}$$

where:

$$G_C = C^T \tilde{W}_u C \tag{21}$$

is the associated gain matrix.

The covariance of $\hat{x}$ and that of the resulting residuals can be computed, much like in the conventional WLS approach, by resorting to the table of factors of the gain matrix G_C.

III. Equality-Constrained Bilinear State Estimation

Like the unconstrained BSE, the constrained version proposed in this work is composed of three main steps, which are discussed in the logical sequence below.

A. First linear stage

It is assumed that the equality constraints become linear in terms of the auxiliary vector y, which is indeed the case when power injections are considered. Therefore, the original SE model (3)-(4), which is nonlinear in terms of x, can be written as follows:

$$z = By + e \tag{22}$$
$$b = Ey \tag{23}$$

where matrix E is simply the counterpart of B for exact injection measurements. By direct comparison of (3)-(4) with (22)-(23), keeping in mind (9) and (10), the following relationships result:

$$h(x) = Bf^{-1}(Cx) \quad ; \quad h_e(x) = Ef^{-1}(Cx) \tag{24}$$

The WLS estimate $\tilde{y}$, given the set of measurements (22) and the constraints (23), is obtained by solving the augmented system,

$$\begin{bmatrix} B^T W B & E^T \\ E & 0 \end{bmatrix} \begin{bmatrix} \tilde{y} \\ \tilde{\lambda} \end{bmatrix} = \begin{bmatrix} B^T W z \\ b \end{bmatrix} \tag{25}$$

which also provides an estimate of the vector of Lagrange multipliers, $\tilde{\lambda}$. As in the original nonlinear formulation (6), a scaling factor β can be introduced in the above system, which is omitted for simplicity of notation.

The gain matrix in (25) is the inverse of the associated covariance matrix [12],

$$G_{Ba} = \begin{bmatrix} G_B & E^T \\ E & 0 \end{bmatrix} = \mathrm{cov}^{-1} \begin{bmatrix} \tilde{y} \\ \tilde{\lambda} \end{bmatrix} \tag{26}$$

where $G_B = B^T W B$ and the subscript a denotes 'augmented'. This matrix will appear in subsequent steps as a weighting matrix.

B. Intermediate nonlinear transformation

In order to get equation systems with homogeneous sizes, the nonlinear transform (9) is trivially augmented as follows:

$$\begin{bmatrix} \tilde{u} \\ \tilde{\lambda} \end{bmatrix} = \begin{bmatrix} f(\tilde{y}) \\ \tilde{\lambda} \end{bmatrix} \tag{27}$$

which, in incremental form, yields,

$$\begin{bmatrix} \Delta u \\ \Delta \lambda \end{bmatrix} = \underbrace{\begin{bmatrix} \tilde{F} & 0 \\ 0 & I \end{bmatrix}}_{\tilde{F}_a} \begin{bmatrix} \Delta y \\ \Delta \lambda \end{bmatrix} \tag{28}$$

This allows the covariance matrix of vector $[\tilde{u}, \tilde{\lambda}]$ to be related with that of $[\tilde{y}, \tilde{\lambda}]$, given by (26). In terms of inverses,

$$\mathrm{cov}^{-1} \begin{bmatrix} \tilde{u} \\ \tilde{\lambda} \end{bmatrix} = \tilde{F}_a^{-T} G_{Ba} \tilde{F}_a^{-1} = \tilde{W}_a \tag{29}$$

where the inverse of $\tilde{F}_a$ is trivially obtained.

C. Second linear stage

The last stage reduces to the following WLS linear model:

$$\tilde{u} = Cx + e_u \tag{30}$$
$$\tilde{\lambda} = \lambda + e_\lambda \tag{31}$$

where the two subsystems become coupled by the error covariance (29). The corresponding estimates are provided by solving the system,

$$\underbrace{\left(C_a^T \tilde{W}_a C_a\right)}_{\tilde{G}_{Ca}} \begin{bmatrix} \hat{x} \\ \hat{\lambda} \end{bmatrix} = C_a^T \tilde{W}_a \begin{bmatrix} \tilde{u} \\ \tilde{\lambda} \end{bmatrix} \tag{32}$$

where the following augmented matrix is introduced,

$$C_a = \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix}$$

In summary, the equality-constrained bilinear formulation, involving (25), (27) and (32), is directly obtained by properly augmenting the unconstrained version, (16), (18) and (20), with the vector of Lagrange multipliers.

As shown in Table II of [11], except for round-off errors, the covariance matrix values provided by the BSE are the same as those of the CSE. Hence, virtually the same normalized residuals result, and the bad data processing can proceed as usual (provided the redundancy is enough). The same applies to the equality-constrained version of the BSE (the required covariance matrix elements can be computed as explained in [12]).

D. Simplifications for successive snapshots

A great deal of computational effort can be saved when consecutive snapshots arrive in which only the measurement values change (i.e., same measurement set, network parameters and topology), which happens almost all the time.

In those cases, when processing the first snapshot (cold start) it is worth saving the matrix $B^T W$ and the LU factors of G_{Ba}, so that they can be recovered for subsequent snapshots. This way, solving (25) reduces to computing the independent vector (sparse matrix-vector product) and performing low-cost forward/backward solution steps (hot start).

E. Solution refinement

One of the main advantages of the bilinear approach, compared to the conventional GN iterative scheme, is that there is no need to choose initial values to iterate, as they are simply and reliably provided by solving the first linear filter (25), which is also the reason why the second linear filter (32) yields virtually the optimal solution after a single run of the three-stage procedure.

However, in the presence of very large measurement errors, the first estimate $\{\hat{x}, \hat{\lambda}\}$ may not be accurate enough, owing to the fact that the Jacobian $\tilde{F}$ is computed around the point $\tilde{u}$, which is not the best possible value for u. The accuracy of $\tilde{u}$ can be checked by recomputing u as follows,

$$\hat{u} = C\hat{x} \tag{33}$$

and then comparing $\hat{u}$ with $\tilde{u}$. If both values of u are close enough, then the first estimate is acceptable. Otherwise, as explained in the Appendix, improved estimates can be obtained by solving,

$$\tilde{G}_{Ca} \begin{bmatrix} x - \hat{x} \\ \lambda - \hat{\lambda} \end{bmatrix} = C_a^T \left[\tilde{F}_a \hat{F}_a^{-1}\right]^T \tilde{W}_a \begin{bmatrix} \tilde{u} - \hat{u} \\ \tilde{\lambda} - \hat{\lambda} \end{bmatrix} \tag{34}$$

A careful analysis of the above system reveals that the second run may be needed by a combination of two factors:

1) Non-linearity of the transformation $f(\cdot)$. Should this function be linear (constant Jacobian), then the term $\tilde{F}_a \hat{F}_a^{-1}$ would be the identity matrix and the right side of the above equation would vanish according to (32), no matter how large the measurement errors are.
2) Measurement errors. Should the measurements be exact, then the first estimate would be also exact ($\hat{u} = \tilde{u}$), no matter how nonlinear $f(\cdot)$ is, and the right side would also vanish.

The nonlinearity of $f(\cdot)$ is somewhat related to the network loading level, but the most critical factor affecting the accuracy of the first estimate is the measurement errors. For error levels typically found in power system measurements, and beyond (σ values up to 2%), two runs are seldom justified, as will be shown in the experiments below. On the other hand, with flat-start initialization, the conventional GN iterative scheme takes typically between 3 to 5 iterations to reach an estimate of acceptable accuracy.

IV. Test Results

In order to show the advantage and improvements brought about by the proposed equality-constrained bilinear procedure over previously-developed methods, different tests have been carried out. These tests require a series of choices that may be summarized as:

- Two networks will be tested, namely the standard IEEE 118- and 298-bus systems.
- Assumed voltage and power measurement error levels (represented by σ_V and σ_P) of 0.5% and 1.0%, respectively.
- A pre-conditioning parameter β in (6) of 0.01.
- Measurement redundancy of 3.4 for both methods, according to redundancy levels typically found in transmission networks.
- Different sets of equality constraints have been assumed. More specifically, all nodal injections have been sorted in ascending order according to apparent power. Needless to say, power injections are zero at all transfer buses and thus top the sorted list. The issue is that the number of transfer buses at a given network cannot be controlled for the sake of performance comparison. Therefore, from the sorted list of injections, starting from the top, a given percentage of them are fixed to their exact value from the reference state as equality constraints, even though they may not be actually null. Percent values are selected at 5, 10 and 20 with respect to N, the total number of buses.

In addition, the following acronyms will be used for practical reasons and ease of reference hereinafter. The reader is invited to refer to this list for the remainder of the text when needed (the corresponding sections that covered the listed methods are also indicated):

- UE: Unconstrained Estimator (II-A)
- ECE: Equality-Constrained Estimator (II-B)
- UBE: Unconstrained BSE (II-C)
- ECBE: Equality-Constrained BSE, the method proposed in this work (III)

With the previous assumptions and definitions in mind, the first experiment consists in validating the benefit of setting out to implement an equality-constrained state-estimation over the previously developed methods that do not explicitly account for this feature. More specifically, the traditional UE method and the recently proposed UBE hinge on the factorization of matrix G_k in (2) and matrices G_B and G_C in (16)-(20), respectively. As it is now conventional knowledge, the weights associated to the exactly-known measurements must be high enough so that the WLS output fixes their estimated values as close as possible to the known or desired values. The 'closeness' of the estimates depend on the relative weight ascribed to these 'fixed' measurements with respect to the remaining regular measurements.

The problem here is that the relatively large weight difference deteriorates the condition of the involved gain matrices (it reduces the inverse of the eigenvalue spectrum width). A poor condition number may be reflected in numerical instability at matrix factorization stage [2]. To illustrate this issue, let ρ be a parameter representing the factor by which the weights

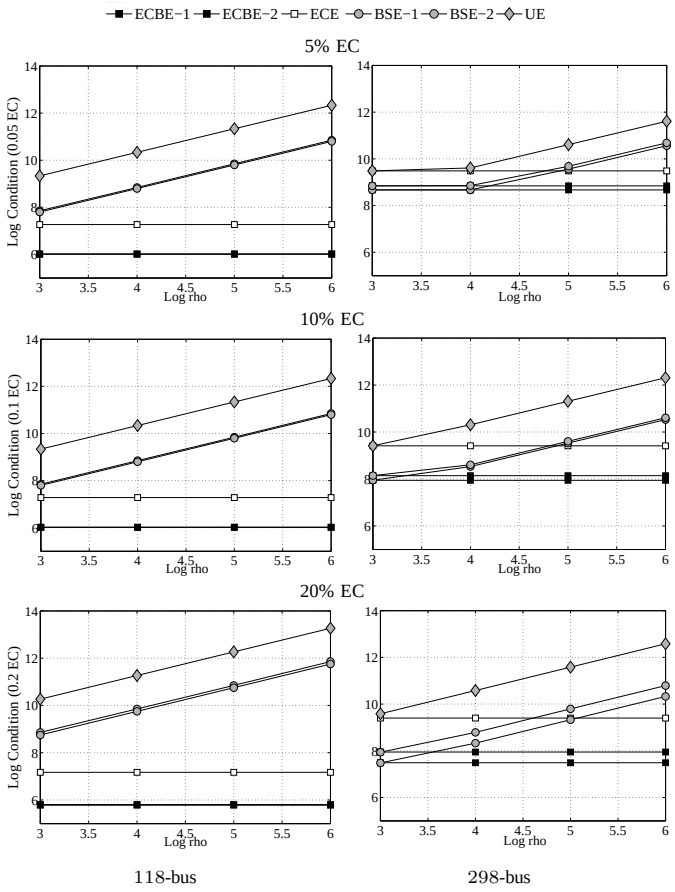


Fig. 1. Log of condition numbers of gain matrices for all tested methods and different percentages of equality-constrained buses

applied to the exact measurements are multiplied with respect to the average weight associated to regular measurements. In other words, if $\overline{w}_i$ is the average weight applied to all regular measurements i, the weight, w_e, applied to all exact measurements is given by $w_e = \rho \times \overline{w}_i$. Therefore, the aim of this experiment is to look at how different values of ρ affect the conditioning of all matrices involved in the different estimation methods described in the previous section. The results of this analysis are summarized in Fig. 1.

The left column of plots in Fig. 1 refers to the 118-bus network while the right column refers to the 298-bus network. The x-axis on all plots shows the $\log_{10}$ of ρ. For instance, a ρ equal to five indicates that $w_e = 10^5\,\overline{w}_i$. The y-axis label on all plots shows the $\log_{10}$ of the condition number for the corresponding gain matrices in each method (the y-axis ranges between five and 14 in all plots in order to bring up the relative differences across the examined cases). From top to bottom, the plots differ in the percentage of buses (five, 10 and 20) for which the injection measurements have been either: a) set as equality constraints in the standard ECE and the proposed ECBE method, or b) being applied the ρ factor in the UE and UBE methods.

The legend on top of Fig. 1 refers to different matrices involved in all the estimation methods examined in this work. Note that both the recently proposed UBE method [11] and the ECBE method developed here rely on two linear systems, each one requiring the factorization of a gain matrix as described in sections II-C and II-B, respectively. The legend indicates the results for each step in these procedures (ECBE-1, ECBE-2 and UBE-1, UBE-2).

Different key results stem from Fig. 1. The horizontal lines in the different subplots are the natural consequence of the condition numbers related to those methods that explicitly account for the equality constraints: parameter ρ plays no role at all in these methods. What is important for the purpose of this work is that the condition numbers for the matrices involved in the proposed ECBE method are, in all examined cases, better than those of the traditional ECE. Refer for instance to the bottom-right plot (the 298-bus network with 20% of injections set as equality constraints). The condition number for the gain matrix in the traditional ECE is almost two orders of magnitude larger than those for the ECBE method (about 10 versus eight). A similar pattern is observed for the 118-bus network.

More striking differences are observed, however, when the proposed ECBE method is compared to those that make use of the weight scaling factor ρ (the traditional UE and the recently-developed UBE). As expected, the condition numbers increase with ρ. The relative impact remains fairly similar across all cases, even matching closely the increase in ρ. In other words, as ρ grows by three orders of magnitude (from 10^3 to 10^6) the condition numbers for the gain matrices of the traditional UE and UBE methods grow also approximately by three orders of magnitude (refer, for instance, to the top-left plot in which the $\log_{10}$ of the condition numbers for the UBE and the UE methods grow from about eight and nine to 11 and 12, respectively). This result provides a significant evidence that the proposed ECBE method may prove to be substantially more stable from the numerical standpoint in the presence of equality constraints, even when compared to the recently-developed UBE. There is no definite conclusion about the impact of the network size at least for the ones adopted in this work (condition numbers increase by similar amounts in all plots).

The impact of the condition number on the accuracy of the solution of linear equation systems has been thoroughly studied elsewhere, and several upper bounds to the solution errors have been theoretically found. For LS problems, the following relationship applies to the solution based on the Cholesky factorization (see [16], section 2.2.3):

$$||\tilde{x} - x_{ex}||_2 \leq 2.5 n^{3/2} u \kappa(G) ||x_{ex}||_2 \tag{35}$$

where

$\tilde{x}$ is the computed solution,
x_{ex} is the exact value,
n is the size of the equation system,
u is the machine precision (10^{-16} in double-precision arithmetic),
$\kappa(G)$ is the condition number of the gain matrix.

For instance, for the 298-bus system ($n \simeq 600$), with $u = 10^{-16}$, the following modest bound on the solution accuracy,

$$||\tilde{x} - x_{ex}||_2 \leq 0.01 ||x_{ex}||_2 \tag{36}$$

requires that

$$\kappa(G) \leq 2.72 \cdot 10^9$$

Fortunately, particularly for very large sparse systems, such bounds are rarely reached in practice and, in any case, unacceptable solution errors usually arise only in state variables of areas in which the ill-conditioning originates, the remaining errors being much smaller. For this reason, the infinite norm (largest absolute value) is frequently a more appropriate indicator than the Euclidean norm (sum of squares).

In order to more specifically assess the actual solution inaccuracy brought about by the ill-conditioning of the gain matrix, let us consider an exact measurement vector z_{ex}. The computed solution $\tilde{x}$ of the corresponding WLS linear problem is given by:

$$\tilde{x} = \tilde{G}^{-1} H^T W z_{ex} \tag{37}$$

where the inverse $\tilde{G}^{-1}$ is explicitly used for compactness of notation (in practice, the Normal equations are solved through sparse triangular factorization). Being exact, z_{ex} satisfies

$$z_{ex} = H x_{ex} \tag{38}$$

which means that there exists a reference state vector, x_{ex}, yielding null residuals. Therefore, by direct substitution, the following expression immediately comes about:

$$x_{ex} - \tilde{x} = \underbrace{\left(I - \tilde{G}^{-1} G\right)}_{R_{IC}} x_{ex} \tag{39}$$

For a given $\kappa(G)$, the residual matrix R_{IC} would be null should sufficiently precise arithmetic be adopted during the solution process. Otherwise, certain components may be large enough to yield inaccurate solutions. Therefore, $||R_{IC}||_\infty$ can be used in practice to assess the impact of $\kappa(G)$ on the solution accuracy.

Let us illustrate the above discussion with the help of the 118-bus system, conveniently modified to increase the intrinsic ill-conditioning. It is well known that the presence of very short branches significantly deteriorates the conditioning of the Normal equations (see [2], Chapter 3). Accordingly, the following modifications have been carried out:

- The smallest series impedance, Z_{min}, corresponding to branch 68-116 has been divided by 100. This way, the Z_{max}/Z_{min} network's ratio is about 10^4, which is still within the maximum range frequently found in practice. For instance, in the European interconnected system (UCTE), this ratio approaches 10^5 (one order of magnitude larger) for several subsystems.
- Power injections at both terminal buses, 68 and 116, are considered exact (i.e handled via either high weights or equality constraints, depending on the solution procedure). This way, both sources of ill-conditioning are geographically close.

Firstly, Table I shows the condition numbers of gain matrices corresponding to methods ECBE and UBE for parameter ρ ranging between 10^3 and 10^5. These numbers are shown for the original and modified 118-bus networks and 20% of buses with injections set as equality constraints. Notice once again that condition numbers under method ECBE are not affected by the choice of ρ. As a matter of fact, the horizontal lines in the bottom-left plot of Fig. 1 for method ECBE reflect the values shown in first row of Table I for the (original) 118-bus network. The same holds for condition numbers in both steps of method UBE for growing values of ρ. The point behind this table is, however, the effect of the impedance ratio increase on the condition numbers in both methods, leading to a 100-fold increase from the original to the modified network.

TABLE I
LOG OF CONDITION NUMBERS OF DIFFERENT GAIN MATRICES FOR THE ORIGINAL AND MODIFIED 118-BUS SYSTEMS

Method	$Log(\rho)$	118-bus		118-bus modified	
		Step 1	Step 2	Step 1	Step 2
ECBE	-	5.81	5.78	7.83	7.81
UBE	3	8.86	8.75	10.5	10.5
	4	9.85	9.75	11.5	11.5
	5	10.9	10.8	12.5	12.5

TABLE II
$||R_{IC}||_\infty$ VALUES CORRESPONDING TO DIFFERENT GAIN MATRICES FOR THE ORIGINAL AND MODIFIED 118-BUS SYSTEMS

Method	$Log(\rho)$	118-bus		118-bus modified	
		Step 1	Step 2	Step 1	Step 2
ECBE	-	1.34E-09	3.48E-10	5.43E-07	1.69E-05
UBE	3	1.80E-03	3.10E-03	1.43E+01	4.02E+00
	4	4.91E-02	4.42E-02	1.18E+02	1.52E+01
	5	2.44E+01	2.87E+01	4.14E+04	1.26E+03

The key issue is then, to elucidate how the previous results influence the numerical performance of the examined methods. To this end, Table II shows the $||R_{IC}||_\infty$ values corresponding to the gain matrices of the two WLS problems arising in the BSE procedures, namely G_B and G_C (UBE) and G_{Ba} and G_{Ca} (ECBE) for both, the original and modified 118-bus networks.

The largest $||R_{IC}||_\infty$ values, and consequently the largest solution errors, always originate in those buses where the sources of ill-conditioning are located. It is worth noting that the remaining components of matrix R_{IC} in (39) are several orders of magnitude smaller. As discussed above this issue calls for the need of using the infinite rather than the Euclidean or other norms as accuracy measure. The key points behind these results are that:

- The infinite norm of R_{IC}, ($||R_{IC}||_\infty$, the largest error) is highly affected by the choice of ρ. This can be noticed by working the way down on the original 118-bus system in Table II. Increasing ρ by two orders of magnitude (from three to five) leads to an equivalent or even larger increase in $||R_{IC}||_\infty$.
- Looking across the table along the horizontal direction, the network's maximum to minimum impedance ratio has a large influence in $||R_{IC}||_\infty$ as well. For $\rho = 10^3$, for instance, increasing this ratio by 100 (modified 118-bus system) amplifies $||R_{IC}||_\infty$ by a factor of almost 10^4. As mentioned above, this reactance ratio is still smaller than those observed in practice in very large-scale networks, such as the UCTE system, where the effect of

TABLE III
ACCURACY COMPARISON BETWEEN ECE AND ECBE METHODS

Buses	% EC	ECE			ECBE (single run)	
		Mean No. Iter.	Mean S_v $\times 10^{-3}$	CV	Mean S_v $\times 10^{-3}$	CV
118	5	4.00	1.2735	0.35	1.2738	0.35
	10	4.00	1.2666	0.36	1.2669	0.36
	20	4.00	1.2330	0.37	1.2332	0.37
298	5	5.02	1.6644	0.13	1.6642	0.13
	10	5.00	1.1423	0.17	1.1426	0.17
	20	5.00	0.7906	0.15	0.7935	0.15

TABLE IV
ACCURACY COMPARISON BETWEEN ECE AND ECBE METHODS FOR HIGHER NOISE SCENARIO

Buses	% EC	ECE			ECBE (single run)	
		Mean No. Iter.	Mean S_v $\times 10^{-3}$	CV	Mean S_v $\times 10^{-3}$	CV
118	5	4.00	2.5470	0.35	2.5489	0.35
	10	4.00	2.5331	0.36	2.5351	0.36
	20	4.00	2.4659	0.37	2.4676	0.37
298	5	5.49	3.3287	0.13	3.3269	0.13
	10	5.00	2.2844	0.17	2.2869	0.17
	20	5.00	1.5811	0.15	1.6036	0.15

ill-conditioning will certainly be more dramatic.

Having looked into the details of how the conditioning of the gain matrices is improved by the proposed ECBE method and the implications of ill-conditioning on accuracy, the second test consists in measuring the accuracy of the estimations that the ECBE method provides with respect to its analogous counterpart, that is, the conventional ECE. Estimating the accuracy of either method hinges on the availability of an 'exact' or reference state, which is estimated from an AC load flow based on a set of injections. The reference state is named here $\overline{\mathbf{V}}_r$. The over-lined, bold symbol indicates that the reference state is a vector of voltage phasors, one entry for each bus (naturally, these phasors are calculated from the load-flow voltage magnitudes and angles).

Thus, for any state estimated via the ECE or the ECBE methods, say $\overline{\mathbf{V}}_e$, its relative accuracy with respect to the reference state is estimated as:

$$S_v = \frac{1}{N}\,\mathbf{e}^\top \left|\overline{\mathbf{V}}_e - \overline{\mathbf{V}}_r\right| \tag{40}$$

where N is the number of buses of the network under study and $\mathbf{e}$ a vector of ones of size N. The previous expression basically says that the accuracy of a method is calculated by taking the norms of the voltage phasor differences (which is a vector itself) and estimating the vector's mean entry.

Since each estimate is the result of chance depending on the noise randomly picked in generating the measurement set, the index S_v has been estimated for 500 different measurement sets based on the σ_V, and σ_P stated above.

Table III shows the results. On the first column, the examined network is indicated. Column two refers to the different scenarios for the percent number of buses with respect to N (total number of buses) for which the active and reactive injections have been fixed as equality constraints. Column three shows the average number of iterations over the 500 cases for which the ECE declared convergence. The remaining columns show the mean S_v and coefficient of variation (CV) over the 500 runs. It is of key importance to point out that the S_v shown for the ECBE model (penultimate column) corresponds to the accuracy reached after a *single* run of the model in all 500 cases.

Two meaningful conclusions stem from Table III: a) it can be seen how after just a single run, the accuracy of the proposed ECBE method matches almost exactly the accuracy that the traditional ECE reaches after four or five iterations, depending on the network. An analogous result was previously shown in [11] regarding the UBE; b) as expected, and as the relative number of equality constraints increase, the overall accuracy of both methods increase. This can be observed by looking at the decreasing values for the mean S_v from top to bottom for each network. This remark is more noticeable for the 298-bus case.

The third experiment of interest corresponds to looking at the effect of larger measurement errors on the performance of the proposed method with respect to the conventional ECE. To that end, the standard deviations for voltage and injection measurements have been doubled with respect to the previous case, that is, values of 1.0% for σ_V and 2.0% for σ_P have been assumed. Arguably, these values are substantially higher than those observed in typical power systems. The analysis is carried out, nonetheless, as a further performance test for the proposed model under extreme conditions. Analogously to the previous case, 500 artificial set of measurements have been created based on the assumed σ_V and σ_P and the conventional ECE and the proposed ECBE have been run for each artificial set. The key results are summarized in Table IV. Similarly to the previous case, it can be observed how under such high noise scenario, the ECBE method, after just a single run, still matches quite closely the required accuracy for which the conventional ECE takes four or five iterations to reach.

The observation that the proposed ECBE method matches after just one run the accuracy reached by the ECE method deserves a closer look. To this end, Fig 2 shows non-parametric kernel probability densities (PDF) of parameter S_v calculated for each of the 500 runs of the ECE after five iterations (298-bus network, 20% proportion of equality constraints, the bottom case in Table III). Both errors levels are included in the same figure for direct visual comparison. Similar PDFs are shown for the ECBE after the first and single run (in a black dotted line). Once again, it can be clearly observed that the PDF corresponding to the ECBE method, after just a single run, lies right over the final ECE's PDF.

The need for a second run of the ECBE method would be justified in case of substantially high accuracy requirements or the presence of measurements subject to atypical errors. The way the second run is performed is explained in Section III-E. In order to illustrate the solution refinement achieved by this process, Fig. 3 compares the PDFs of parameter S_v based on the 500 runs after, a) full convergence of the ECE

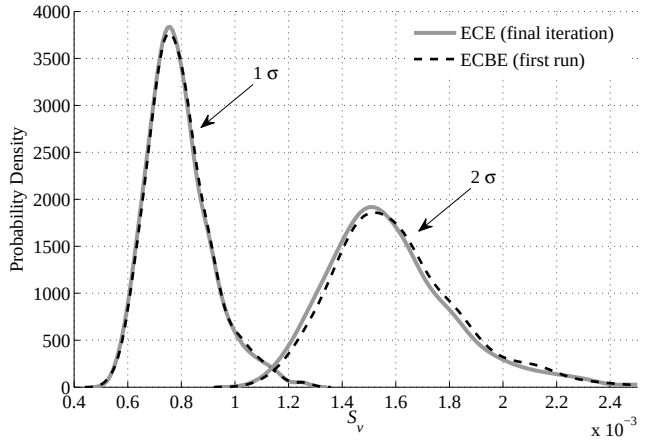


Fig. 2. PDFs of parameter S_v after convergence of ECE method and a single run of ECBE

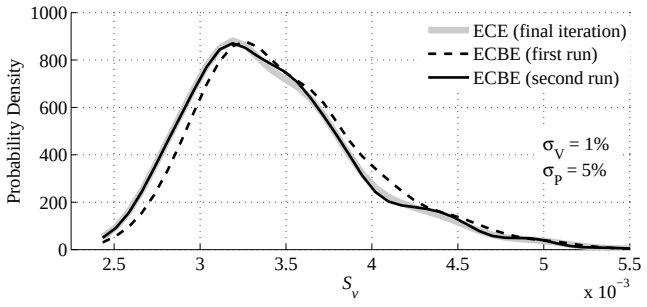


Fig. 3. PDFs of parameter S_v after convergence of ECE method and first and second runs of ECBE

method, and, b) after first and second runs of ECBE method. Standard deviations of 1.0% for σ_V and 5.0% for σ_P have been assumed. As it can be seen, the PDF for the ECBE method after the second run matches closely that of the ECE method even under this high-noise scenario.

Finally, computation times for the ECE and ECBE methods have been compared. The algorithms have been coded in Matlab (version R2008a), Windows 7 on a 64-bit Intel i5, 2.27 GHz processor. The code has been optimized under different options and strategies. More precisely, in addition to the built-in sparse matrix Matlab handling capability and vectorized operations, other code optimization aspects described in [11] regarding vector computation and storage have been also applied in this work. Solution times have been substantially lowered as a result. Table V shows the average solution times for the noise scenario given by a σ_V of 0.5% and σ_P of 1.0% and 20% proportion of equality constraints. Based on the simulations, the following data are collected:

- Individual solution times for each of the three direct phases involved in the proposed ECBE approach. For the first linear phase, 'cold' and 'hot' starts have been evaluated.
- Additional solution time for the second run of the ECBE method.
- Solution time for the converged conventional solution (ECE).

TABLE V
AVERAGE SOLUTION TIMES FOR DIFFERENT STAGES AND RESULTING SPEEDUPS. TEST NETWORKS INDICATED BY BOLD CHARACTERS.

TIMES (ms)			**118**	**298**
ECBE	1st Linear Phase	cold	15.6	39.2
		hot	0.16	0.28
	Nonlinear Transfomation		1.49	2.52
	2nd Linear Phase		4.15	14.4
	2nd Run		1.15	2.17
ECE	Complete		94.2	427

SPEEDUPS		**118**	**298**
T_c/T_a	Single Run (cold)	4.43	7.60
	Single Run (hot)	16.2	24.8
	Two Runs (cold)	4.20	7.32
	Two Runs (hot)	13.6	22.0

The speedup, defined as T_c/T_a, where T_c and T_a represent the solution times of the ECE and the ECBE methodologies, respectively, has been computed for each case and included also in Table V. The results may be summarized as:

- For a single 'cold-start' run of the proposed ECBE method, the obtained speedups are 4.43 and 7.60 (naturally, the speedup increases with the problem size).
- For the 'hot-start' run (where a majority of computations associated with the first linear step, the most computational consuming of the three steps, are avoided), considerably larger speedups are observed (16.2 and 24.8).
- In the case of the two runs of the ECBE method, speedups of 4.20 to 7.32 ('cold start) and 13.6 to 22.0 ('hot-start') are observed, slightly lower than the single-run cases.

V. CONCLUSION

The recently introduced Bilinear State Estimation scheme has been enhanced in this work to incorporate equality constraints explicitly by augmenting, with a vector of Lagrange multipliers, the matrices involved in each of the three steps (two linear filters with an intermediate nonlinear transformation). The main and foremost advantage of this approach is to do away with the use of high weights for exact measurements, a strategy that deteriorates the condition of gain matrices in state estimation methods that do not account for equality constraints explicitly. Experiments carried out for two standard network benchmarks as well as different set of equality constraints and noise scenarios show important improvements in matrix conditioning with respect to previous methods (with all the numerical advantages behind it) yet the solutions reached after just a single run of the equality-constrained bilinear formulation, is substantially close to those provided by traditional methods after four or five iterations. Speed-ups of the same order of magnitude as in the unconstrained case have been obtained.

VI. APPENDIX

It will be shown that the BSE and the CSE are fully equivalent, provided the required number of iterations are performed in both cases. The unconstrained version will be first addressed (as a complement to [11] where this proof is missing), and then the equality-constrained formulation will be

developed by simply augmenting the involved matrices. The reader is referred to the body of the paper for the notation.

A. Unconstrained state estimation

The CSE obtains the unconstrained WLS estimate by minimizing the scalar,

$$J(x) = \frac{1}{2}[z - h(x)]^T W [z - h(x)].$$

When the factorized model (8)-(10) is adopted, the same WLS solution is obtained by minimizing the corresponding Lagrangian function, which can be expressed in terms of vectors y and x as follows (for the sake of compactness, the auxiliary vector $u = f(y)$ is not explicitly used),

$$\mathcal{L}(y, x, \mu) = \frac{1}{2}[z - By]^T W [z - By] + \mu^T [f(y) - Cx] \quad (41)$$

The First Order Optimality Conditions (FOOC) for this case are,

$$\begin{array}{rcl} G_B y + F^T \mu &=& B^T W z \\ -C^t \mu &=& 0 \\ f(y) - Cx &=& 0 \end{array} \quad (42)$$

where the only nonlinear term is $f(y)$. It is worth noting that both $f(\cdot)$ and its Jacobian F can be trivially inverted, which involves only scalars or 2x2 blocks. This function is next linearized around the point $\tilde{y}$, obtained by solving the WLS linear problem (16), yielding:

$$\begin{bmatrix} G_B & 0 & \tilde{F}^T \\ 0 & 0 & -C^T \\ \tilde{F} & -C & 0 \end{bmatrix} \begin{bmatrix} y - \tilde{y} \\ x \\ \mu \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\tilde{u} \end{bmatrix} \quad (43)$$

Eliminating $y - \tilde{y}$ leads to the following reduced system,

$$\begin{bmatrix} 0 & C^T \\ C & \tilde{F} G_B^{-1} \tilde{F}^T \end{bmatrix} \begin{bmatrix} x \\ \mu \end{bmatrix} = \begin{bmatrix} 0 \\ \tilde{u} \end{bmatrix} \quad (44)$$

and eliminating μ,

$$(C^T \tilde{W}_u C)\hat{x} = C^T \tilde{W}_u \tilde{u} \quad (45)$$

The resulting system is the same as that of (20) and provides the BSE solution, $\hat{x}$, after the first run. Substituting into the eliminated equations yields:

$$\begin{array}{rcl} \hat{\mu} &=& \tilde{W}_u(\tilde{u} - \hat{u}) \\ G_B(\hat{y} - \tilde{y}) &=& \tilde{F}^T \tilde{W}_u(\hat{u} - \tilde{u}) \end{array} \quad (46)$$

with $\hat{u} = C\hat{x}$. Note that the triplet $\{\hat{x}, \hat{\mu}, \hat{y}\}$ satisfies the first two equations of the FOOC (42), and also the third one in the WLS sense (with weights provided by $cov(\tilde{y})$), which explains the surprisingly good results provided by the BSE after the first run.

Next, the FOOC are linearized around $\hat{y}$, leading to the incremental model:

$$\begin{bmatrix} G_B & 0 & \hat{F}^T \\ 0 & 0 & -C^T \\ \hat{F} & -C & 0 \end{bmatrix} \begin{bmatrix} y - \hat{y} \\ x \\ \mu \end{bmatrix} = \begin{bmatrix} G_B(\tilde{y} - \hat{y}) \\ 0 \\ -\hat{u} \end{bmatrix} \quad (47)$$

Eliminating again $y - \hat{y}$ and μ yields:

$$\underbrace{(C^T \hat{W}_u C)}_{\hat{G}_C}(x - \hat{x}) = C^T \hat{F}^{-T} G_B(\tilde{y} - \hat{y}) \quad (48)$$

the solution of which provides the estimate of the second run. Replacing $\hat{G}_C$ in the above system by $\tilde{G}_C$ (already built and factorized from the first run) significantly reduces the computational burden of this step, while the computed values are virtually the same (solution times shown in Table V have been obtained this way).

Finally, taking into account (46), an equivalent system is obtained in terms of Δu:

$$(C^T \hat{W}_u C)(x - \hat{x}) = C^T \left[\tilde{F}\hat{F}^{-1}\right]^T \tilde{W}_u(\tilde{u} - \hat{u}) \quad (49)$$

B. Equality-constrained state estimation

In this case, the Lagrangian function is augmented so that the equality constraints are enforced, as follows,

$$\mathcal{L} = \frac{1}{2}[z - By]^T W [z - By] + \mu^T [f(y) - Cx] - \lambda^T (b - Ey) \quad (50)$$

The First Order Optimality Conditions (FOOC) for this case become,

$$\begin{array}{rcl} G_B y + E^T \lambda + F^T \mu &=& B^T W z \\ Ey &=& b \\ -C^t \mu &=& 0 \\ f(y) - Cx &=& 0 \end{array} \quad (51)$$

The nonlinear function $f(\cdot)$ is linearized around the point $\tilde{y}$, which is obtained by solving the augmented system (25), yielding:

$$\left[\begin{array}{cc|cc} G_B & E^T & 0 & \tilde{F}^T \\ E & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -C^T \\ \tilde{F} & 0 & -C & 0 \end{array}\right] \left[\begin{array}{c} y - \tilde{y} \\ \lambda - \tilde{\lambda} \\ \hline x \\ \mu \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ \hline 0 \\ -\tilde{u} \end{array}\right] \quad (52)$$

In order to obtain augmented matrices of proper dimensions, an additional dummy variable ($\gamma = 0$) is introduced, as well as the trivial identity,

$$(\lambda - \tilde{\lambda}) - \lambda = -\tilde{\lambda}$$

which leads to,

$$\left[\begin{array}{cc|cc|cc} G_B & E^T & 0 & 0 & \tilde{F}^T & 0 \\ E & 0 & 0 & 0 & 0 & I \\ \hline 0 & 0 & 0 & 0 & -C^T & 0 \\ 0 & 0 & 0 & 0 & 0 & -I \\ \hline \tilde{F} & 0 & -C & 0 & 0 & 0 \\ 0 & I & 0 & -I & 0 & 0 \end{array}\right] \left[\begin{array}{c} y - \tilde{y} \\ \lambda - \tilde{\lambda} \\ \hline x \\ \lambda \\ \hline \mu \\ \gamma \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ \hline 0 \\ 0 \\ \hline -\tilde{u} \\ -\tilde{\lambda} \end{array}\right] \quad (53)$$

Next, in addition to the augmented matrices G_{Ba}, F_a and C_a defined in Section III, the following augmented vectors are introduced,

$$y_a = \begin{bmatrix} y \\ \lambda \end{bmatrix}; \; x_a = \begin{bmatrix} x \\ \lambda \end{bmatrix}; \; \mu_a = \begin{bmatrix} \mu \\ \gamma \end{bmatrix}; \; u_a = \begin{bmatrix} u \\ \lambda \end{bmatrix}$$

With this compact notation, the system (53) becomes,

$$\begin{bmatrix} G_{Ba} & 0 & \tilde{F}_a^T \\ 0 & 0 & -C_a^T \\ \tilde{F}_a & -C_a & 0 \end{bmatrix} \begin{bmatrix} y_a - \tilde{y}_a \\ x_a \\ \mu \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\tilde{u}_a \end{bmatrix} \quad (54)$$

which, except for the subindex 'a', is a replica of (43). Therefore, proceeding exactly as in the unconstrained case, the first estimate is obtained by solving (32), repeated here for convenience,

$$\tilde{G}_{Ca}\hat{x}_a = C_a^T \tilde{W}_a \tilde{u}_a \quad (55)$$

while the second run involves the solution of the following incremental system

$$\tilde{G}_{Ca}(x_a - \hat{x}_a) = C_a^T \hat{F}_a^{-T} G_{Ba}(\tilde{y}_a - \hat{y}_a) \quad (56)$$

which can be rewritten in terms of $(\tilde{u}_a - \hat{u}_a)$,

$$\tilde{G}_{Ca}(x_a - \hat{x}_a) = C_a^T \left[\tilde{F}_a \hat{F}_a^{-1}\right]^T \tilde{W}_a(\tilde{u}_a - \hat{u}_a) \quad (57)$$

REFERENCES

[1] F. C. Schweppe, J. Wildes, D. B. Rom, "Power system static state estimation, Part I, II, III", *IEEE Trans. Power Apparat. Syst.*, Vol. PAS-89, No. 1, pp. 120-135, Jan. 1970.

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

[3] Aschmoneit F., Peterson N., Adrian E., "State Estimation with Equality Constraints", 10th PICA Conference Proceedings, Toronto, pp. 427-430, May 1977.

[4] A. Gjelsvik, "The significance of the Lagrange Multipliers in WLS State Estimation with Equality Constraints," in Proceedings of the 11th Power Systems Computation Conference, Avignon, pp. 619-625, August 1993.

[5] Gjelsvik A., Aam S. and Holten L., "Hachtel's Augmented Matrix Method - A Rapid Method Improving Numerical Stability in Power System Static State Estimation", *IEEE Trans. Power Apparat. Syst.*, Vol. PAS-104, pp. 2987-2993, November 1985.

[6] R. A. M. van Amerongen, "On the exact incorporation of virtual measurements in orthogonal-transformation based state-estimation procedures," *Int. J. Electrical Power & Energy Systems*, vol. 13, no. 3, pp. 167-174, June 1991.

[7] Monticelli, A., "Electric power system state estimation," *Proceedings of the IEEE*, vol. 88, no.2, pp.262-282, Feb 2000.

[8] A. de la Villa Jaén, A. Gómez-Expósito, "Implicitly Constrained Substation Model for State Estimation," *IEEE Transactions on Power Systems*, vol. 17, no. 3, pp. 850-856, August 2002.

[9] Z. Jun, A. Abur;, "Identification of network parameter errors," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 586-592, May 2006.

[10] C. Gómez-Quiles, A. de la Villa Jaén and A. Gómez-Expósito, "A Factorized Approach to WLS State Estimation", *IEEE Transactions on Power Systems*, Vol. 26, No. 3, pp. 1724-1732, 2011.

[11] A. Gómez-Expósito, C. Gomez-Quiles, A. de la Villa Jaén, "Bilinear Power System State Estimation", *IEEE Transactions on Power Systems* Vol. 27, No. 1, pp. 493-501, 2012.

[12] Wu, F.F.; Liu, W.-H.E.; Lun, S.-M.; "Observability analysis and bad data processing for state estimation with equality constraints", *IEEE Transactions on Power Systems*, Vol. 3, No. 2, pp. 541-548, May 1988.

[13] P. Chen, H. Runesha, D. T. Nguyen, P. Tong and T. Y. P. Chang, "Sparse algorithms for indefinite system of linear equations", *Computational Mechanics*, Vol. 25, No. 1, pp. 33-42, 2000.

[14] J.D. Hogg and J.A. Scott, "An indefinite sparse direct solver for multicore machines", Technical Report TR-RAL-2010-011. Available from http://www.numerical.rl.ac.uk/reports/reports.shtml

[15] E. Anderson et. al., "Lapack User's Guide", http://www.netlib.org/lapack/lug/lapack_lug.html

[16] A. Bjorck, *Numerical Methods for Least Squares Problems*, SIAM 1996.

Catalina Gómez-Quiles (SM'09) received the Eng. degree (Hon) from University of Seville, Spain, in 2006 and the M.Eng Degree from McGill University, Montreal, Canada in 2008, both in Electrical Engineering. In 2012 she received the PhD from University of Seville. Her research interests include risk assessment in competitive electricity markets and mathematical and computer models for power system analysis.

Hugo A. Gil (M'00) received the Eng. degree from the Universidad Nacional de Colombia, Medelln, Colombia in 1995 and the Dr. Eng. degree from Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil in 2001, both in Electrical Engineering. He spent several years as postdoctoral fellow, associate researcher and lecturer at McGill University, Montreal, Canada. He has also been a renewable energy integration analyst in the public and private sectors in Canada and the United States. In 2009, he joined University of Seville, Spain. His work is focused on the economic and regulatory issues of renewable energy integration and the power transmission business.

Antonio de la Villa Jaén was born in Riotinto, Spain, in 1960. He received his electrical and doctor engineering degrees from the University of Seville. He is presently an Associate Professor at the Department of Electrical Engineering, University of Seville. His primary areas of interest are computer methods for power system state estimation problems and wave energy converter control.

Antonio Gómez-Expósito (F'05) received the electrical engineering and doctor degrees from the University of Seville, Spain, where he is currently the Endesa Red Industrial Chair Professor. His primary areas of interest are optimal power system operation, state estimation, digital signal processing and control of flexible ac transmission system devices.