Power System Static-State Estimation With Uncertain Network Parameters as Input Data

Gabriele D'Antona

Abstract—This paper introduces a novel Least Squares (LS) observer for the state estimation in power system under static conditions. The proposed state estimator takes into consideration not only the measurement uncertainty but also the uncertainty associated with all the network parameters used in the load flow model of the power grid. A comparison with the classical weighted LS method is reported.

Index Terms—Least squares (LS) methods, power system monitoring, power system state estimation (SE), SE, uncertain systems.

I. INTRODUCTION

ELECTRICAL power system observers for state estimation (SE) play an important role for operating the energy management system (EMS) in control centers [1]–[3]. The SE is based on the availability of a redundant set of data together with power system models that link the unknown state variables with the available data [4]. Both data and models are characterized by a certain degree of uncertainty that the state observer must be capable to address properly, because it will influence the subsequent EMS actions [5].

Usually, the data provided to state observers are gathered from direct telemetered measurements and pseudomeasurements (prior knowledge, e.g., information stemmed from system forecasts); these two sources of information are affected by uncertainty. For what concerns the power system models typically just one is available, derived from a quasi-static mathematical representation of the electrical power grid. The power system model is uncertain as well because of uncertainties affecting the network parameters, such as transmission line and transformer longitudinal and transversal impedances, transformer tap values, and circuit breaker positions; moreover, further inaccuracies could result from unmodeled unbalanced operating conditions and system asymmetries, network transients, nonlinear loads, and so on [6]–[10]. The models available in the literature for the network parameters do not specify the model uncertainty, also because the geographical distribution of the electrical networks makes the statistical multivariate correlation analysis quite complex.

The classical nonlinear Least Squares (LS) approach is the most widespread technique employed for the SE in electrical power systems, first proposed in [11]–[13] in the early 1970's, and known in the electrical power system literature as

weighted LS (WLS) observer. The WLS framework assumes a probabilistic description for the measurement uncertainty and the availability of an exact (i.e., perfectly known and not uncertain) static power system model.

The aforementioned assumption, which considers the power system model not uncertain, is far to be true. As a matter of fact, often network parameters are generally less accurate than the measurement data [9]. The lack of the model uncertainty contribution in the WLS algorithm has the advantage of a simpler problem formulation but with the serious risk of a significant bias upon the state estimates, together with a noteworthy undervaluation of the uncertainty affecting the state estimates [14], [15]. Besides, it is important to point out that this risk increases with the improvement of the measurement quality, being in this case more important the network parameter uncertainty contribution.

This concern has led some investigators to propose effective network parameter estimation (PE) methods [16]–[19]. Some of these PE techniques are based on residual sensitivity analysis [20], [21], while others require augmenting the state vector either using the normal equation [22]–[24] or adopting a Kalman filter [25]–[27].

Although the PE improves the SE performances, it does not really address the problem of including the model uncertain parameters as input data to the SE process. As a matter of fact, the PE based on residual sensitivity analysis does not have any effect on the state estimates, but it just detects and identifies bad parameters by unexpected large residuals resulting from the WLS algorithm. For what concerns the PE based on augmenting the state vector instead, which estimates together the state variable with some parameters, the assigned number of measurement data limits the applicability of these methods, because increasing the number of estimated parameters reduces the redundancy of the information provided by the data, eventually leading to local or global observability limits.

This paper proposes a new perspective of the SE problem that considers the power grid model as an uncertain system because of the limited accuracy of the network parameters. The proposed SE technique does not require any PE procedure, because the network parameters and their uncertainty are treated as input data to the estimator, which are known *a priori*, exactly as the pseudomeasurements. If necessary those parameters estimated in real time (*a posteriori*), using a PE algorithm, may be handled as input data to the estimator as well, exactly as the data resulting from real-time meters (as power flows, voltage magnitudes, and phases).

In the following discussion, a probability-based approach to uncertainty will be adopted, according to the recommendation

provided by the ISO/IEC guides [28]–[30]. The mean and variance–covariance matrix of the uncertain model parameters and measurements can derive either from prior knowledge (e.g., system design data and pseudomeasurements) or *a posteriori* from PE algorithms and real-time meters. Other authors have analyzed the SE problem with uncertain parameters before, but uncertainty was not described by a covariance matrix. For example, in [31], the measurement and parameter uncertainties were modeled as bounded intervals.

This paper will be set out as follows. Section II introduces the notation and conventions adopted; Section III reviews the WLS static SE approach, and Section IV generalizes this approach to cope with parameter uncertainty; Section V presents the new and more effective LS formulation for the SE, which considers an uncertain power system model; finally, in Section IV, the performance of the new SE algorithm is discussed and exemplified by few test cases showing the effects of parameter uncertainties.

II. Definitions and Notation

The following notation conventions will be used. An uppercase bold face letter denotes a matrix, while a lowercase bold face letter denotes a vector. All the vectors are column vectors. The notation of [32] is used here for matrix algebra and differential calculus. For easy reference, some rules and definitions of matrix differential calculus are listed in Appendix A.

The measurement and pseudomeasurement data are organized in the M-vector $\mathbf{y}$. It is related to the system static-state N-vector $\mathbf{x}$ by the following relationship expressing the power system model:

$$\mathbf{y}_T = \mathbf{f}(\mathbf{x}_T; \boldsymbol{\pi}_T, \boldsymbol{\kappa}_T) \in \mathbb{R}^{M \times 1} \qquad (1)$$

where the subscript T stands for *true* value.

In (1), the P-vector $\boldsymbol{\pi}$ gathers together the values of the network parameters (transmission line resistances, reactances and charging capacitances, transformer reactances and tap values, and shunt capacitor/reactor values are the examples of network parameters), while the K-vector $\boldsymbol{\kappa}$ represents the network topology (breaker positions and so on.), with

$$\boldsymbol{\pi}_T \in \mathbb{R}^{P \times 1}, \quad \boldsymbol{\kappa}_T \in \mathbb{Z}^{K \times 1}. \qquad (2)$$

Since measurements, pseudomeasurements, and network parameters and topology are affected by uncertainties due to random errors $\Delta \mathbf{y}$ in the measurements and $\Delta \boldsymbol{\pi}$ and $\Delta \boldsymbol{\kappa}$ in the network parameters, if $M > N$, the system of equations (1) does not have any solution $\mathbf{x}$ in general, i.e., it is inconsistent

$$\mathbf{y} \neq \mathbf{f}(\mathbf{x}_T; \boldsymbol{\pi}, \boldsymbol{\kappa}) \qquad (3)$$

where

$$\mathbf{y} = \mathbf{y}_T + \Delta \mathbf{y} \qquad (4a)$$

$$\boldsymbol{\pi} = \boldsymbol{\pi}_T + \Delta \boldsymbol{\pi}, \quad \boldsymbol{\kappa} = \boldsymbol{\kappa}_T + \Delta \boldsymbol{\kappa} \qquad (4b)$$

are the available measurements and parameters while $\Delta \mathbf{y}$, $\Delta \boldsymbol{\pi}$, and $\Delta \boldsymbol{\kappa}$ are the measurement and parameter errors.

The combination of the consistent model formulation (1) with (4a) and (4b) yields the following equation relating the

available measurements and parameters to the unknown values of the true state and measurement-parameter errors:

$$\mathbf{y} - \Delta \mathbf{y} = \mathbf{f}(\mathbf{x}_T; \boldsymbol{\pi} - \Delta \boldsymbol{\pi}, \boldsymbol{\kappa} - \Delta \boldsymbol{\kappa}). \qquad (5)$$

In the following developments, only the effects of errors $\Delta \boldsymbol{\pi}$ affecting the network parameters will be considered, while the case of errors $\Delta \boldsymbol{\kappa}$ on the network topology vector will not be considered in this paper. Therefore, the perfectly known vector $\boldsymbol{\kappa}$ will be dropped from the arguments of the measurement model $\mathbf{f}(\cdot)$ for the sake of simplicity.

Linearization with a Taylor series expansion of (5) about the available value of the network parameter vector $\boldsymbol{\pi}$ yields

$$\mathbf{y} - \Delta \mathbf{y} \cong \mathbf{f}(\mathbf{x}_T; \boldsymbol{\pi}) - \left. \frac{\partial \mathbf{f}}{\partial \boldsymbol{\pi}^T} \right|_{\mathbf{x}_T, \boldsymbol{\pi}} \cdot \Delta \boldsymbol{\pi}. \qquad (6)$$

Introducing the data vector $\mathbf{d}$, which includes all the measurement and parameter data necessary for carry on the SE

$$\mathbf{d} = \begin{bmatrix} \mathbf{y} \\ \boldsymbol{\pi} \end{bmatrix} \in \mathbb{R}^{(M+P) \times 1} \qquad (7)$$

equation (6) can be recasted in the more compact matrix form

$$\mathbf{r}(\mathbf{x}_T; \mathbf{d}) \cong \mathbf{A}(\mathbf{x}_T; \boldsymbol{\pi}) \cdot \Delta \mathbf{d} \qquad (8)$$

with $\mathbf{r}$ being the model-measurement misfit (residuals)

$$\mathbf{r}(\mathbf{x}_T; \mathbf{d}) = \mathbf{y} - \mathbf{f}(\mathbf{x}_T; \boldsymbol{\pi}) \qquad (9)$$

$\Delta \mathbf{d}$ being the data error

$$\Delta \mathbf{d} = \begin{bmatrix} \Delta \mathbf{y} \\ \Delta \boldsymbol{\pi} \end{bmatrix} \in \mathbb{R}^{(M+P) \times 1} \qquad (10)$$

and $\mathbf{A}$ is the matrix

$$\mathbf{A}(\mathbf{x}; \boldsymbol{\pi}) = \begin{bmatrix} \mathbf{I}_M & -\dfrac{\partial \mathbf{f}}{\partial \boldsymbol{\pi}^T} \end{bmatrix} \in \mathbb{R}^{M \times (M+P)}. \qquad (11)$$

III. Classical WLS State Estimation

The classical WLS approach assumes that only measurements $\mathbf{y}$ are affected by errors, while the network parameter $\boldsymbol{\pi}$ values are considered known exactly, i.e., the data error (10) is

$$\Delta \mathbf{d} = \begin{bmatrix} \Delta \mathbf{y} \\ \mathbf{0}_{P \times 1} \end{bmatrix} \in \mathbb{R}^{(M+P) \times 1} \qquad (12)$$

and consequently just the measurement errors concur to the misfit (8)

$$\mathbf{r}(\mathbf{x}_T; \mathbf{d}) = \mathbf{y} - \mathbf{f}(\mathbf{x}_T; \boldsymbol{\pi}) = \Delta \mathbf{y}. \qquad (13)$$

In this case, an estimate $\widetilde{\mathbf{x}}$ of the state vector and of the measurement errors $\Delta \widetilde{\mathbf{y}}$ is defined as those values, which minimize the weighted squared norm of the residuals $\mathbf{r}$ constrained to satisfy (13) [11]

$$[\text{WLS}] : \underset{\{\mathbf{x}, \Delta \mathbf{y}\}}{\arg \min} \| \Delta \mathbf{y} \|_{\boldsymbol{\Sigma}_\mathbf{y}}^2$$
$$\text{s.t.} \ \Delta \mathbf{y} = \mathbf{y} - \mathbf{f}(\mathbf{x}; \boldsymbol{\pi}) \qquad (14)$$

where s.t. stands for subject to.

The norm in (14) is defined as the Mahalanobis norm [33] weighted by the variance–covariance matrix $\mathbf{\Sigma_y}$ of the random measurement errors, representing the measurement (and pseudomeasurement) uncertainty, that is

$$g(\mathbf{x}, \mathbf{d}) = \|\Delta \mathbf{y}\|^2_{\mathbf{\Sigma_y}} = \mathbf{r}^T \cdot \mathbf{\Sigma_y}^{-1} \cdot \mathbf{r}. \qquad (15)$$

The unconstrained minimization of (15) with respect to $\mathbf{x}$ leads to the solution of the constrained nonlinear LS problem (14), providing the estimation of the following quantities [1], [11].

1) The static-state vector $\widetilde{\mathbf{x}}$, by the following iterative equation:

$$\widetilde{\mathbf{x}}_{k+1} = \widetilde{\mathbf{x}}_k - \left[\left(\frac{\partial^2 g}{\partial \mathbf{x}^2} \right)^{-1} \left(\frac{\partial g}{\partial \mathbf{x}^T} \right)^T \right]_{\widetilde{\mathbf{x}}_k, \mathbf{y}}. \qquad (16)$$

2) The variance–covariance matrix $\mathbf{\Sigma}_{\widetilde{\mathbf{x}}}$ of the state estimate as

$$\mathbf{\Sigma}_{\widetilde{\mathbf{x}}} = 2 \left(\frac{\partial^2 g}{\partial \mathbf{x}^2} \right)^{-1}_{\widetilde{\mathbf{x}}, \mathbf{y}}. \qquad (17)$$

3) The measurement errors $\Delta \widetilde{\mathbf{y}}$ as

$$\Delta \widetilde{\mathbf{y}} = \mathbf{y} - \mathbf{f}(\widetilde{\mathbf{x}}; \boldsymbol{\pi}). \qquad (18)$$

4) The variance–covariance matrix $\mathbf{\Sigma}_{\Delta \widetilde{\mathbf{y}}}$ of the measurement errors as

$$\mathbf{\Sigma}_{\Delta \widetilde{\mathbf{y}}} = \mathbf{\Sigma_y} + \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}^T} \mathbf{\Sigma}_{\widetilde{\mathbf{x}}} \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}^T} \right)^T \right]_{\widetilde{\mathbf{x}}_k, \mathbf{y}} +$$
$$- \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}^T} \mathbf{\Sigma}_{\widetilde{\mathbf{x}}, \mathbf{y}} \right]^T_{\widetilde{\mathbf{x}}_k, \mathbf{y}} - \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}^T} \mathbf{\Sigma}_{\widetilde{\mathbf{x}}, \mathbf{y}} \right]_{\widetilde{\mathbf{x}}_k, \mathbf{y}}. \qquad (19)$$

When the network parameters uncertainty is far to be negligible, the estimates (16) are biased and the uncertainty (17) is underestimated. In other words, the assumption that there is no error in the system model does not hold anymore because of the uncertainty of the network parameters.

IV. New Generalization of WLS

In this section, a new generalization of the WLS approach will be introduced. The uncertainty of the network parameters can be handled using the linearized expression (6) yielding to the following expression of the measurement misfit that can be thought totally due to an artificially increased measurement error $\Delta \mathbf{y}'$ [14], [16], [21]:

$$\Delta \mathbf{y}' = \mathbf{y} - \mathbf{f}(\mathbf{x}_T, \boldsymbol{\pi}) \cong \Delta \mathbf{y} - \frac{\partial \mathbf{f}}{\partial \boldsymbol{\pi}^T} \bigg|_{\mathbf{x}_T, \boldsymbol{\pi}} \Delta \boldsymbol{\pi}. \qquad (20)$$

It is important to underline that (20), which is an input information to the LS estimator, cannot be used in practice, because it relies on the unknown true value $\mathbf{x}_T$ of the state. If $\mathbf{x}_0$ represents an initial guess of the state vector $\mathbf{x}$, (20) can be rewritten as

$$\Delta \mathbf{y}' \cong \Delta \mathbf{y} - \frac{\partial \mathbf{f}}{\partial \boldsymbol{\pi}^T} \bigg|_{\mathbf{x}_0, \boldsymbol{\pi}} \Delta \boldsymbol{\pi} - \frac{\partial}{\partial \mathbf{x}^T} \left(\frac{\partial \mathbf{f}}{\partial \boldsymbol{\pi}^T} \Delta \boldsymbol{\pi} \right)_{\mathbf{x}_0, \boldsymbol{\pi}} \Delta \mathbf{x}_0. \qquad (21)$$

The random augmented measurements errors (21) leads to the following expression of the variance–covariance matrix $\mathbf{\Sigma}_{\mathbf{y}'}$ (for analytical details see Appendix C):

$$\mathbf{\Sigma}_{\mathbf{y}'} = \mathbf{\Sigma_y} + \left(\frac{\partial \mathbf{f}}{\partial \boldsymbol{\pi}^T} \right)_{\mathbf{x}_0, \boldsymbol{\pi}} \mathbf{\Sigma}_{\boldsymbol{\pi}} \left(\frac{\partial \mathbf{f}}{\partial \boldsymbol{\pi}^T} \right)^T_{\mathbf{x}_0, \boldsymbol{\pi}}$$
$$+ \left(\frac{\partial^2 \mathbf{f}}{\partial \boldsymbol{\pi}^T \partial \mathbf{x}^T} \right)_{\mathbf{x}_0, \boldsymbol{\pi}} \left(\mathbf{\Sigma}_{\mathbf{x}_0} \otimes \mathbf{\Sigma}_{\boldsymbol{\pi}} \right) \left(\frac{\partial^2 \mathbf{f}}{\partial \boldsymbol{\pi}^T \partial \mathbf{x}^T} \right)^T_{\mathbf{x}_0, \boldsymbol{\pi}} \qquad (22)$$

evaluated considering uncorrelated the errors related to the measurements, the parameters, and the state initial guess. If some network parameters are estimated online from the measurement data the parameters/measurements uncorrelation hypothesis in (22) does not hold anymore. This case can be easily handled considering the generalized law for the propagation of uncertainty presented in [34] for the evaluation of the variance matrix (22).

The WLS estimation problem (14) can be redefined correspondingly as

$$[\text{GWLS}] \; : \; \underset{\{\mathbf{x}, \Delta \mathbf{y}\}}{\arg \min} \|\Delta \mathbf{y}\|^2_{\mathbf{\Sigma}_{\mathbf{y}'}}$$
$$\text{s.t. } \Delta \mathbf{y} = \mathbf{y} - \mathbf{f}(\mathbf{x}; \boldsymbol{\pi}) \qquad (23)$$

whose solution can be attained minimizing the function

$$g(\mathbf{x}, \mathbf{d}) = \|\Delta \mathbf{y}\|^2_{\mathbf{\Sigma}_{\mathbf{y}'}} = \mathbf{r}^T \cdot \mathbf{\Sigma}_{\mathbf{y}'}^{-1} \cdot \mathbf{r}. \qquad (24)$$

The solution of the *generalized* WLS (GWLS) problem is expressed by (16) and (17) with $\mathbf{\Sigma}_{\mathbf{y}'}$ in place of $\mathbf{\Sigma_y}$.

If the state vector initial guess $\mathbf{x}_0$ is accurate enough, the GWLS approach (23) has the advantage of providing better estimates for the static-state vector.

A major limitation of the GWLS estimator is the inability to identify single and multiple measurement and parameters outliers, since the uncertainties of the measured data, the parameters, and the initial state guess are mixed together by (21), and the error estimate cannot separate the three contributions.

Due to the fact that: 1) parameter uncertainties can badly affect both the WLS and GWLS state estimates and 2) the identification of bad parameter and measurement data is unfeasible, a new method to the SE will be introduced next.

V. Proposed Extension of the WLS Algorithm

An extension of the SE algorithm is based on the solution of a new optimization problem, named *extended* WLS (EWLS), which minimizes the whole data error vector $\Delta \mathbf{d}$ rather than the misfit $\mathbf{r}$ only, like in the WLS and GWLS algorithms. In the EWLS problem, the estimates of the state vector $\widetilde{\mathbf{x}}$ and of the data errors $\Delta \widetilde{\mathbf{d}}$ are defined as the values, which minimize the weighted squared norm of the data error $\Delta \mathbf{d}$, constrained to satisfy (8)

$$[\text{EWLS}] \; : \; \underset{\{\mathbf{x}, \Delta \mathbf{d}\}}{\arg \min} \|\Delta \mathbf{d}\|^2_{\mathbf{\Sigma}_{\mathbf{d}}}$$
$$\text{s.t. } \mathbf{r} = \mathbf{A}(\mathbf{x}; \boldsymbol{\pi}) \cdot \Delta \mathbf{d}. \qquad (25)$$

Although not strictly necessary, in the following, we will assume that the measurement and parameter errors are

uncorrelated, as for the GWLS algorithm, i.e., the variance–covariance matrix of the data vector is the following block-diagonal matrix:

$$\boldsymbol{\Sigma}_{\mathbf{d}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{y}} & \mathbf{0}_{M \times P} \\ \mathbf{0}_{P \times M} & \boldsymbol{\Sigma}_{\boldsymbol{\pi}} \end{bmatrix} \in \mathbb{R}^{(M+P) \times (M+P)}. \quad (26)$$

The constrained problem (25), in the unknown $\Delta \mathbf{d}$, is a classical overdetermined linear least norm problem, whose solution is [35]

$$\Delta \mathbf{d} = \boldsymbol{\Sigma}_{\mathbf{d}} \mathbf{A}^T [\mathbf{A} \boldsymbol{\Sigma}_{\mathbf{d}} \mathbf{A}^T]^{-1} \mathbf{r}. \quad (27)$$

It is important to highlight that (27) expresses the data error $\Delta \mathbf{d}$ as a function of $\mathbf{x}$; therefore, the EWLS estimate of $\mathbf{x}$ results from the minimization of the nonlinear function $g(\mathbf{x})$, which expresses the weighted squared norm of (27)

$$g(\mathbf{x}, \mathbf{d}) = \|\Delta \mathbf{d}\|_{\boldsymbol{\Sigma}_{\mathbf{d}}}^2 = \mathbf{r}^T [\mathbf{A} \boldsymbol{\Sigma}_{\mathbf{d}} \mathbf{A}^T]^{-1} \mathbf{r}. \quad (28)$$

Considering now the definition (11) of the matrix $\mathbf{A}$ and the expression (26) of the data variance–covariance matrix, it is simple to verify that, if the matrix $\mathbf{A}(\mathbf{x}; \boldsymbol{\pi})$ is evaluated at the initial guess $\mathbf{x}_0$ of the state vector, the function (28) is the same as (24), which defines the GWLS problem.

The EWLS algorithm instead recursively evaluates the matrix $\mathbf{A}$ at each iteration using the kth state estimate $\widetilde{\mathbf{x}}_k$. The state estimate provided by this algorithm will be more accurate than the GWLS, because the approximation (21) of the augmented measurement error is replaced by the iterative expression

$$\Delta \mathbf{y}' \cong \Delta \mathbf{y} - \left. \frac{\partial \mathbf{f}}{\partial \boldsymbol{\pi}^T} \right|_{\widetilde{\mathbf{x}}_k, \boldsymbol{\pi}} \Delta \boldsymbol{\pi} - \frac{\partial}{\partial \mathbf{x}^T} \left(\frac{\partial \mathbf{f}}{\partial \boldsymbol{\pi}^T} \Delta \boldsymbol{\pi} \right)_{\widetilde{\mathbf{x}}_k, \boldsymbol{\pi}} \Delta \mathbf{x}_k \quad (29)$$

which is more accurate since $\Delta \mathbf{x}_k \to 0$ as $k \to +\infty$.

A. State Estimate

The solution of the EWLS problem is attained iteratively by the unconstrained minimization of (28) by the Newton method [36] leading to the same WLS formula (16), but with g previously defined by (28). The expressions of the Jacobian and Hessian of g to be used in (16) for the EWLS estimate are reported in Appendix B.

B. State-Estimate Variance–Covariance Matrix

The covariance matrix of the estimated state is obtained, with similar considerations made in [37], [38], from the observation that the state-estimate perturbation $\Delta \mathbf{x}$, due to a data error $\Delta \mathbf{d}$, is given by

$$\Delta \mathbf{x} = \frac{d\mathbf{x}}{d\mathbf{d}^T} \Delta \mathbf{d}. \quad (30)$$

In order to obtain an expression for the derivative in (30), let us consider the following implicit system of N functions:

$$\boldsymbol{\phi}(\mathbf{x}, \mathbf{d}) = \frac{\partial g}{\partial \mathbf{x}} = \mathbf{0}_{N \times 1} \quad (31)$$

obtained taking into account that when the function g is at minimum its first derivative (B.6) vanishes. If the system of functions (31) is differentiable at $(\mathbf{x}, \mathbf{d})$, it can be solved for $\mathbf{x}$

as a function of $\mathbf{d}$: $\mathbf{x} = \mathbf{x}(\mathbf{d})$, and its derivative can be found by differentiating (31) [39]

$$\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{d}^T} + \frac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}^T} \frac{d\mathbf{x}}{d\mathbf{d}^T} = \mathbf{0} \Rightarrow \frac{d\mathbf{x}}{d\mathbf{d}^T} = -\left(\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}^T} \right)^{-1} \frac{\partial \boldsymbol{\phi}}{\partial \mathbf{d}^T}. \quad (32)$$

The partial derivatives in (32) can be expressed considering the definitions of $\boldsymbol{\phi}$ just given in (31)

$$\frac{d\mathbf{x}}{d\mathbf{d}^T} = -\left(\frac{\partial^2 g}{\partial \mathbf{x}^2} \right)^{-1} \frac{\partial^2 g}{\partial \mathbf{x} \partial \mathbf{d}^T} \quad (33)$$

with (B.10) as the Hessian of g and (B.15) as the mixed second derivative of g.

From (30), the expression of the state-estimate covariance matrix can be easily derived from the data covariance matrix by the relation

$$\boldsymbol{\Sigma}_{\mathbf{x}} = \frac{d\mathbf{x}}{d\mathbf{d}^T} \boldsymbol{\Sigma}_{\mathbf{d}} \left(\frac{d\mathbf{x}}{d\mathbf{d}^T} \right)^T$$
$$= \left(\frac{\partial^2 g}{\partial \mathbf{x}^2} \right)^{-1} \frac{\partial^2 g}{\partial \mathbf{x} \partial \mathbf{d}^T} \boldsymbol{\Sigma}_{\mathbf{d}} \frac{\partial^2 g}{\partial \mathbf{d} \partial \mathbf{x}^T} \left(\frac{\partial^2 g}{\partial \mathbf{x}^2} \right)^{-1}. \quad (34)$$

Since according to (B.15) and (B.10)

$$\frac{\partial^2 g}{\partial \mathbf{x} \partial \mathbf{d}^T} \boldsymbol{\Sigma}_{\mathbf{d}} \frac{\partial^2 g}{\partial \mathbf{d} \partial \mathbf{x}^T} = 2 \frac{\partial^2 g}{\partial \mathbf{x}^2} \quad (35)$$

equation (34) simplifies to the same expression of the WLS state-estimate covariance matrix expression (17), but considering the EWLS expression for g previously defined in (28) and the Hessian expression derived in Appendix B.

C. Data Error Estimate

A significant advantage of the EWLS algorithm compared with the WLS and GWLS state estimators is that it is allows to estimate both the measurement and the parameter errors, while the classical WLS approach deals with measurement error only. This characteristic of the EWLS opens the way to identify and detect gross errors in both measured data and network parameters. For the reasons of brevity, it is not possible to develop in this paper the identifiability and detectability process of the EWLS algorithm, but an expression for the data errors estimate and its covariance matrix will be given.

The expression of the data error estimate $\Delta \widetilde{\mathbf{d}}$ is attained replacing in (27) the estimate of the state vector. The substitution of the derivative (B.13) into (27) provides the following expression for the data error estimate:

$$\Delta \widetilde{\mathbf{d}} = \frac{1}{2} \boldsymbol{\Sigma}_{\mathbf{d}} \left(\frac{\partial g}{\partial \mathbf{d}^T} \right)_{\widetilde{\mathbf{x}}}^T. \quad (36)$$

The definition of the above derivative $\partial g / \partial \mathbf{d}^T$ is reported in Appendix B.

D. Data Error Variance–Covariance Matrix

The variance–covariance matrix of the estimated data errors can be evaluated deftly considering that the data error (36) expresses the data misfit as a function of the state estimate $\widetilde{\mathbf{x}}$ and the data $\mathbf{d}$

$$\Delta \widetilde{\mathbf{d}} = \Delta \mathbf{d}(\widetilde{\mathbf{x}}, \mathbf{d}).$$

According to the generalized law of propagation of uncertainty [34], the data error estimate variance–covariance matrix is the following function of $\tilde{\mathbf{x}}$ and $\mathbf{d}$, and the respective data-state covariance matrices:

$$\boldsymbol{\Sigma}_{\Delta\tilde{\mathbf{d}}} = \left[\frac{\partial \Delta\tilde{\mathbf{d}}}{\partial \mathbf{d}^T} \boldsymbol{\Sigma}_{\mathbf{d}} \left(\frac{\partial \Delta\tilde{\mathbf{d}}}{\partial \mathbf{d}^T} \right)^T \right]_{\tilde{\mathbf{x}}}$$
$$+ \left[\frac{\partial \Delta\tilde{\mathbf{d}}}{\partial \mathbf{x}^T} \boldsymbol{\Sigma}_{\tilde{\mathbf{x}}} \left(\frac{\partial \Delta\tilde{\mathbf{d}}}{\partial \mathbf{x}^T} \right)^T \right]_{\tilde{\mathbf{x}}}$$
$$+ \left[\frac{\partial \Delta\tilde{\mathbf{d}}}{\partial \mathbf{x}^T} \boldsymbol{\Sigma}_{\tilde{\mathbf{x}},\mathbf{d}} \left(\frac{\partial \Delta\tilde{\mathbf{d}}}{\partial \mathbf{d}^T} \right)^T + \frac{\partial \Delta\tilde{\mathbf{d}}}{\partial \mathbf{d}^T} \boldsymbol{\Sigma}_{\mathbf{d},\tilde{\mathbf{x}}} \left(\frac{\partial \Delta\tilde{\mathbf{d}}}{\partial \mathbf{x}^T} \right)^T \right]_{\tilde{\mathbf{x}}} . \quad (37)$$

The derivatives of (36) in the previous relationship have the following expressions:

$$\frac{\partial \Delta\tilde{\mathbf{d}}}{\partial \mathbf{d}^T} = \frac{1}{2} \boldsymbol{\Sigma}_{\mathbf{d}} \frac{\partial^2 g}{\partial \mathbf{d}^2}, \quad \frac{\partial \Delta\tilde{\mathbf{d}}}{\partial \mathbf{x}^T} = \frac{1}{2} \boldsymbol{\Sigma}_{\mathbf{d}} \frac{\partial^2 g}{\partial \mathbf{d}\partial \mathbf{x}^T}. \quad (38)$$

The data-state covariance matrix can be derived directly from the data covariance matrix as [34]

$$\boldsymbol{\Sigma}_{\tilde{\mathbf{x}},\mathbf{d}} = \left(\frac{d\mathbf{x}}{d\mathbf{d}^T} \right)_{\tilde{\mathbf{x}}} \boldsymbol{\Sigma}_{\mathbf{d}} \quad (39)$$

with the first derivative of the state with respect to the data given by (33).

Replacing (17), (38), (39) into (37) the following expression of the data errors covariance matrix is achived (see the derivative definitions listed in Appendix B):

$$\boldsymbol{\Sigma}_{\Delta\tilde{\mathbf{d}}} = \boldsymbol{\Sigma}_{\mathbf{d}}(\mathbf{V} - \mathbf{W})\boldsymbol{\Sigma}_{\mathbf{d}}$$

with

$$\mathbf{V} = \frac{1}{4} \left[\frac{\partial^2 g}{\partial \mathbf{d}^2} \boldsymbol{\Sigma}_{\mathbf{d}} \frac{\partial^2 g}{\partial \mathbf{d}^2} \right]_{\tilde{\mathbf{x}}} \quad (40a)$$

$$\mathbf{W} = \frac{1}{2} \left[\frac{\partial^2 g}{\partial \mathbf{d}\partial \mathbf{x}^T} \left(\frac{\partial^2 g}{\partial \mathbf{x}^2} \right)^{-1} \frac{\partial^2 g}{\partial \mathbf{x}\partial \mathbf{d}^T} \right]_{\tilde{\mathbf{x}}} . \quad (40b)$$

VI. CASE STUDIES

The simple ac circuit shown in Fig. 1 is used to exemplify the behavior of the proposed SE and the data detection/identification approaches in the case of parameter uncertainty and outliers.

Let $S_{\text{base}} = 100\,\text{MVA}$ and $V_{\text{base}} = 100\,\text{kV}$ be the base power and the voltage of the power system, while let y_1 denote the measurement of bus 1 voltage magnitude, and y_2, y_3, y_4, and y_5 denote the measurements of the active and reactive power injections at buses 1 and 2, respectively.

The state variables x_1 and x_2 are the voltage magnitudes at buses 1 and 2, while the state variable x_3 is the phase of the voltage at bus 2 (bus 1 is considered as the slack bus). The state is assumed to have the exact (unknown) value

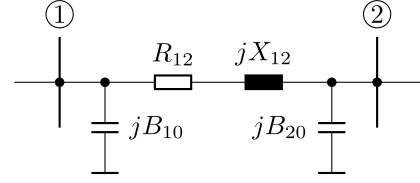$$\mathbf{x}_T = [x_1 \ x_2 \ x_3]^T = [4 \ 3.97 \ -0.2]^T \text{ p.u.} \quad (41)$$

Fig. 1. AC circuit used as a simple test case.

The measurements error are assumed uncorrelated and normally distributed, with the following standard uncertainty (1 % of the base quantities):

$$u(y_1) = u(y_2) = u(y_3) = u(y_4) = u(y_5) = 0.01 \text{ p.u.} \quad (42)$$

For what concerns the network parameters, they have the following nominal values:

$$\pi_1 = R_{12} = 3\ \Omega\ (0.03\,\text{p.u.})$$
$$\pi_2 = X_{12} = 40\ \Omega\ (0.40\,\text{p.u.})$$
$$\pi_3 = B_{10} = 150\ \mu\text{S}\ (0.015\,\text{p.u.})$$
$$\pi_3 = B_{20} = 200\ \mu\text{S}\ (0.020\,\text{p.u.}). \quad (43)$$

When the parameters (43) are considered uncertain, they will be affected by independent Gaussian errors with 10 % relative standard uncertainty $u(\pi_i)$, that is

$$u(\pi_1) = 0.3\ \Omega\ (0.003\,\text{p.u.}), \quad u(\pi_2) = 4\ \Omega\ (0.04\,\text{p.u.})$$
$$u(\pi_3) = 15\ \mu\text{S}\ (0.0015\,\text{p.u.}), \quad u(\pi_4) = 20\ \mu\text{S}\ (0.0020\,\text{p.u.}). \quad (44)$$

This level of uncertainty was simply assumed on the bases of the transmission system operator confidence on the network parameters.

In order to characterize statistically the algorithm behavior, the test case circuit was repeatedly analyzed $n = 10\,000$ times with the following state vector initial guess:

$$\mathbf{x}_0 = [4 \ 4 \ 0]^T \text{ p.u.} \quad (45)$$

For all the test cases analyzed, but the classical WLS without parameter uncertainty, the statistical distributions of the state estimate failed to pass the normality tests [40]. For this reason, in order to test, if the state estimates are unbiased, the multivariate generalization of the sign test proposed in [41] was executed (see Appendix D). According to (D.8), the critical value for all the upcoming test cases, at significance level $\alpha = 5\,\%$, was

$$n\tilde{R}_\alpha^2 = Q_N^{-1}(1 - \alpha) = 7.8. \quad (46)$$

For what concerns the state estimates, the sampling mean $\overline{\Delta\tilde{\mathbf{x}}}$ and the sampling standard deviation $\mathbf{s}_{\Delta\tilde{\mathbf{x}}}$ of the state error together with the standard deviation $\boldsymbol{\sigma}_{\tilde{\mathbf{x}}} = (\text{diag}(\boldsymbol{\Sigma}_{\tilde{\mathbf{x}}}))^{1/2}$ of the estimates, computed analytically using (17) for the 10 000th sample data, are reported together with the $n\tilde{R}_{0.05}^2$ statistic.

A. WLS Test Cases

The WLS algorithm has been tested for two cases:

1) with perfectly known parameters (certain parameters);
2) with parameters affected by uncertainty.

TABLE I
STATISTICS OF THE WLS OBSERVER TEST CASES

	sampling mean $\overline{\Delta \widetilde{\mathbf{x}}}$ (p.u.)	sampling std. dev. $\mathbf{s}_{\Delta \widetilde{\mathbf{x}}}$ (p.u.)	computed std. dev. $\boldsymbol{\sigma}_{\widetilde{\mathbf{x}}}$ (p.u.)	$n\widetilde{R}^2_{0.05}$
certain param.	$\begin{bmatrix} 0.0000 \\ 0.0000 \\ 0.00000 \end{bmatrix}$	$\begin{bmatrix} 0.0079 \\ 0.0079 \\ 0.00080 \end{bmatrix}$	$\begin{bmatrix} 0.0080 \\ 0.0081 \\ 0.00081 \end{bmatrix}$	0.7
uncertain param.	$\begin{bmatrix} 0.013 \\ 0.014 \\ 0.000 \end{bmatrix}$	$\begin{bmatrix} 0.063 \\ 0.067 \\ 0.014 \end{bmatrix}$	$\begin{bmatrix} 0.0080 \\ 0.0081 \\ 0.00082 \end{bmatrix}$	15.5

TABLE II
STATISTICS OF THE GWLS AND EWLS OBSERVER TEST CASES

	sampling mean $\overline{\Delta \widetilde{\mathbf{x}}}$ (p.u.)	sampling std. dev. $\mathbf{s}_{\Delta \widetilde{\mathbf{x}}}$ (p.u.)	calculated std. dev. $\boldsymbol{\sigma}_{\widetilde{\mathbf{x}}}$ (p.u.)	$n\widetilde{R}^2_{0.05}$
GWLS	$\begin{bmatrix} 0.003 \\ 0.005 \\ -0.002 \end{bmatrix}$	$\begin{bmatrix} 0.016 \\ 0.025 \\ 0.019 \end{bmatrix}$	$\begin{bmatrix} 0.0098 \\ 0.013 \\ 0.0010 \end{bmatrix}$	102.4
EWLS	$\begin{bmatrix} 0.000 \\ 0.000 \\ 0.0000 \end{bmatrix}$	$\begin{bmatrix} 0.010 \\ 0.011 \\ 0.0052 \end{bmatrix}$	$\begin{bmatrix} 0.010 \\ 0.011 \\ 0.0052 \end{bmatrix}$	2.5

The results of the analysis are shown briefly in Table I.

For the first test case, the WLS problem (14) is linear, and the solution could be attained analytically but, for consistency with the analysis of the other test cases, it was obtained statistically. The results confirm that the estimates are unbiased and the consistency between the sample statistics and the analytical formula (17).

In the second test case, with uncertain parameters, according to the multivariate sign test (D.6), the state estimate is biased (at 5 % significance level). The standard deviations $\boldsymbol{\sigma}_{\widetilde{\mathbf{x}}}$ of the state estimate evaluated analytically are about $10 \div 20$ times smaller than the actual statistical determination. As mentioned before, the WLS, in the case of uncertain model parameters, largely underestimates the estimate uncertainty.

B. GWLS Test Cases

The GWLS algorithm has been tested with parameters affected by uncertainty, and assuming the state $\mathbf{x}_0$ in (45) affected by uncertainty, whose magnitude is defined by the following covariance matrix:

$$\boldsymbol{\Sigma}_{\mathbf{x}_0} = \begin{bmatrix} 1 \times 10^{-10} & 0 & 0 \\ 0 & 0.02^2 & 0 \\ 0 & 0 & 0.2^2 \end{bmatrix} \text{ p.u.} \qquad (47)$$

The results of the analysis are shown briefly in the first line of Table II, with the data organized as in Table I.

According to the condition (D.6), the state estimate is biased (at 5 % significance level). The calculated variance–covariance matrix of the state estimate, computed analytically using (17) (with the 10 000th estimate), is not consistent with the actual statistical computation as shown by the discrepancy between the standard deviations reported in Table II. Anyhow, the approximation in this case is better than the WLS results with uncertain parameters. The values of the actual standard deviations are less than twice the values evaluated with equation (17), for what concerns the magnitude of the bus voltages. The standard deviation of the bus voltage phase is instead much higher than the theoretical value.

C. EWLS Test Cases

The EWLS algorithm has been tested under the same conditions of the GWLS test case. The results of the analysis are shown briefly in the second line of Table II.

According to the condition (D.4), the state estimate is unbiased (at 5 % significance level). Unlike the GWLS results,
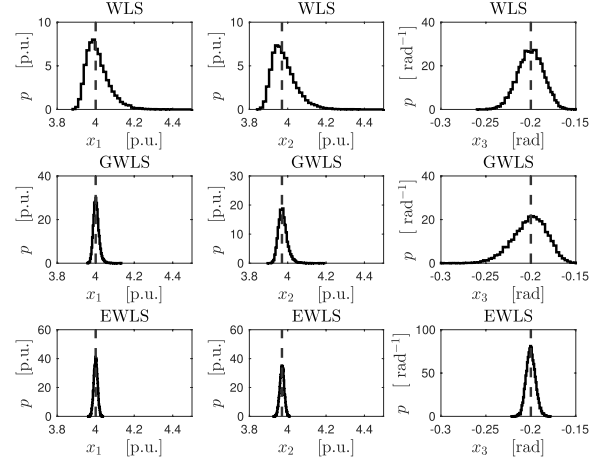
Fig. 2. Comparison of the marginal pdf for each state estimate (the dashed line represents the *true* value of each state variable).

the calculated variance–covariance matrix of the state estimate, computed analytically using (17) with the 10 000th sample data, is consistent with the actual statistical computation.

D. Analysis of the Results

Fig. 2 shows the comparison of the estimator performances when the parameters are affected by uncertainties in terms of the empirical probability density functions (pdf's) evaluated form the statistical estimates ensemble. The plots obviously confirm the results of the test cases summarized in Tables I and II, but with the advantage to display clearly the relative dispersion of the estimation errors for each estimator. Fig. 2 reveals the shape of the estimate distributions; in particular, the WLS and the GWLS results are far to be normally distributed. It is also evident the greater relative efficiency of the EWLS estimator.

Figs. 3 and 4 show the scatter plots of the WLS and GWLS results for x_1 and x_2 (bus voltage magnitudes) and the empirical and theoretical pdf of x_3 (bus voltages relative phase). The scatter plot of the voltage magnitudes shows also the ellipsoids determined from the analytical expression (17) of the state estimate covariance matrix $\boldsymbol{\Sigma}_{\widetilde{\mathbf{x}}}$ applied to the 10 000th sample together with the actual ellipsoid computed from the sample covariance matrix. Both the ellipsoids are assigned defining a 95 % confident region. The plot showing the pdf of the estimated phase compares the theoretical normal pdf with variance taken from $\boldsymbol{\Sigma}_{\widetilde{\mathbf{x}}}$ with the empirical pdf derived from the state estimate samples. Figs. 3 and 4 show distinctly
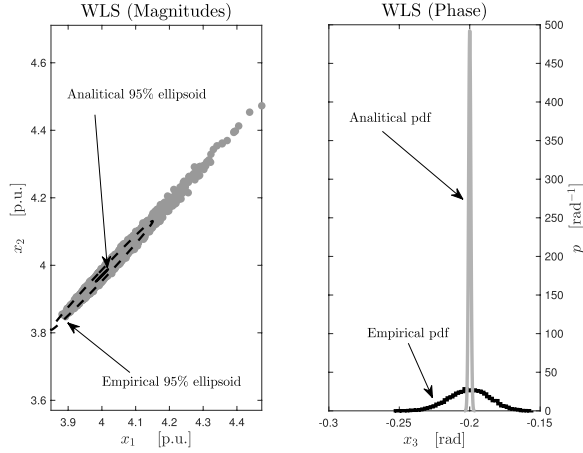
Fig. 3. Scatter plot of the WLS estimate of x_1 and x_2 (left) and the pdf of the estimate x_3 (right).
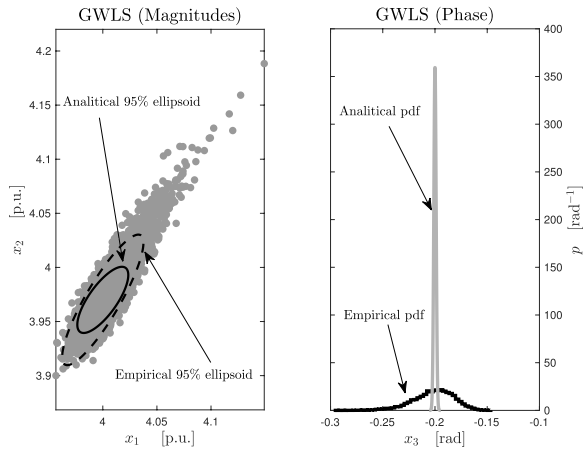

Fig. 4. Scatter plot of the GWLS estimate of x_1 and x_2 (left) and the pdf of the estimate x_3 (right).

how the WLS and GWLS algorithms may underestimate the uncertainty of their results.

Finally, Fig. 5 shows for the EWLS estimator the scatter plot of x_1 and x_2, and the pdf of x_3. In this case, the analytical and the statistical determination of the 95 % confident region ellipsoids and the phase pdf are practically coincident.

Furthermore Fig. 6 shows the pdf of the data error estimates, obtained statistically using 10 000 samples. The distributions are well approximated by Gaussian pdf with mean (36) and covariance matrix (37) (applied to the 10 000th sample in Fig. 6).

Although the WLS, the GWLS, and the EWLS estimation algorithms were not implemented with the aim to investigate the computational burdens, it may be interesting to report the mean time required by each estimator: $T_{\text{WLS}} = 587\ \mu$s, $T_{\text{GWLS}} = 716\ \mu$s, and $T_{\text{EWLS}} = 990\ \mu$s (simulations done using MATLAB R2016a Prerelease, on a MacBook Pro computer 2,9 GHz Intel Core i7, OS X 10.11.4).

VII. CONCLUSION

This paper, after the introduction of a generalization of the WLS static-state estimator capable to cope with network parameter and the initial state guess uncertainties, presents a new approach to the LS SE problem, named the EWLS method.
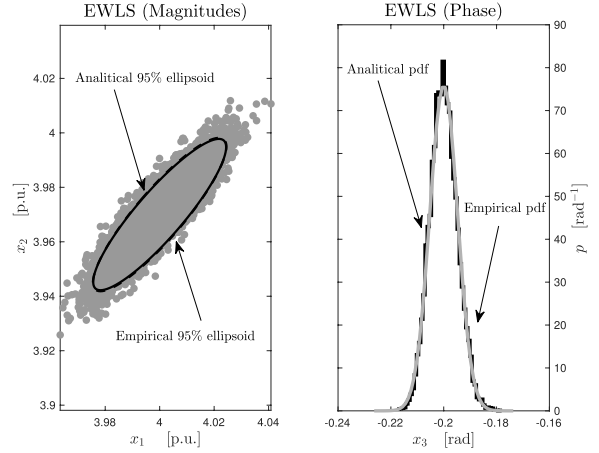

Fig. 5. Scatter plot of the EWLS estimate of x_1 and x_2 (left) and the pdf of the estimate x_3 (right). The theoretical and the statistical ellipsoids and pdf's overlap nearly perfectly in the pictures.
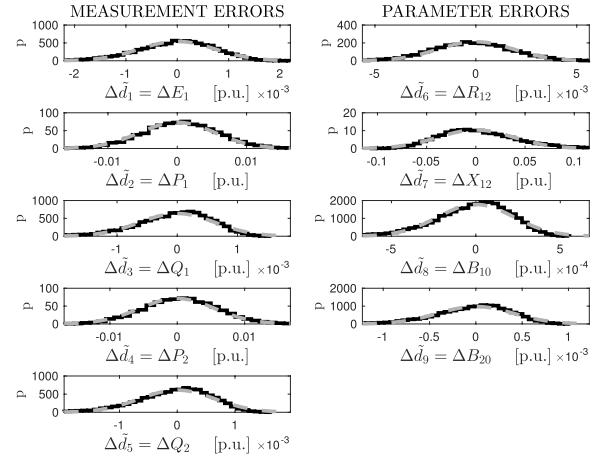

Fig. 6. Marginal pdf of the data error estimates. Black lines: statistical computation using 10 000 samples. Gray dashed lines: analytical evaluation from the 10 000th sample.

The novelty of the proposed estimator is twofold. First, it allows to deal with the uncertainties of the network parameters with higher accuracy than the GWLS. Second, the EWLS formulation provides the estimation of the measurement and parameter errors and the associated covariance matrix, which are not feasible with the WLS and its generalization and are of paramount importance for the successful detection and identification of single and multiple gross errors in measurements and parameters.

For the reasons of brevity, the subject of error identification, computational complexity, and algorithm implementation is not treated in this paper, but will be dealt with in the future works.

APPENDIX A
SOME MATRIX DIFFERENTIAL CALCULUS DEFINITIONS AND PROPERTIES

If $\mathbf{A}$ is an $M \times N$ matrix and $\mathbf{B}$ is a $T \times U$ matrix, the Kronecker product $\mathbf{A} \otimes \mathbf{B}$ is a $MT \times NU$ matrix defined

TABLE III
USEFUL PROPERTIES OF MATRIX KRONECKER PRODUCTS

$(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T$	(A.4)
$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$	(A.5)
$\mathrm{vec}\,(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\,\mathrm{vec}\,(\mathbf{B})$	(A.6)
$\mathrm{vec}\,(\mathbf{ab}^T) = (\mathbf{b} \otimes \mathbf{a})$	(A.7)

as [32], [42]

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \dots & a_{1s}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{r1}\mathbf{B} & \dots & a_{rs}\mathbf{B} \end{bmatrix}. \tag{A.1}$$

The vec($\cdot$) operator transforms a matrix into a vector by stacking the columns of the matrix one underneath the other [32]. Suppose $\mathbf{A}$ is an $M \times N$ matrix specified in terms of its column as

$$\mathbf{A} = [\mathbf{a}_1 \quad \dots \quad \mathbf{a}_N] \in \mathbb{R}^{M \times N} \tag{A.2}$$

with $\mathbf{a}_i \in \mathbb{R}^{M \times 1}$. Then, vec($\mathbf{A}$) is the $MN \times 1$ vector

$$\mathrm{vec}\,(\mathbf{A}) = \begin{bmatrix} \mathbf{a}_1^T & \dots & \mathbf{a}_N^T \end{bmatrix}^T \in \mathbb{R}^{MN \times 1}. \tag{A.3}$$

For convenience, in Table III, are reported some properties of the Kronecker product and the vec($\cdot$) operator [32].

The derivative of a vector or matrix function with respect to a vector or matrix can be defined in several ways [32], [42]–[47]. The notation of [32] will be adopted, and the matrix derivative is defined as

$$\frac{d\mathbf{A}}{d\mathbf{B}} = \frac{d}{d\mathbf{B}} \otimes \mathbf{A}. \tag{A.8}$$

The derivative of vectors and scalars with respect to a vector can be derived from definition (A.8). Table IV displays the derivatives of general expressions involving vectors and matrices. The derivative of any order n can be defined accordingly as the derivative of the matrix obtained as the derivative of order $n - 1$.

APPENDIX B
SOME USEFUL EXPRESSIONS OF THE FUNCTION $g(\mathbf{x})$ DERIVATIVES

The EWLS algorithm proposed in this paper requires to evaluate the expression of the first and second derivatives of the function

$$g(\mathbf{x}, \mathbf{d}) = \|\Delta\mathbf{d}\|_{\Sigma_\mathbf{d}}^2 = \mathbf{r}^T \cdot [\mathbf{A}\Sigma_\mathbf{d}\mathbf{A}^T]^{-1} \cdot \mathbf{r} \tag{B.1}$$

with

$$\mathbf{r}(\mathbf{x}, \mathbf{d}) = \mathbf{y} - \mathbf{f}(\mathbf{x}; \boldsymbol{\pi}). \tag{B.2}$$

This operation is far to be simple and requires the application of the matrix derivative rules provided in Appendix A. The necessary operations and the relevant results are collected in the following.

TABLE IV
MATRIX DERIVATIVES OF GENERAL EXPRESSIONS

$\mathbf{f}(\mathbf{x}) - f(\mathbf{x})$	$d\mathbf{f}/d\mathbf{x}^T - df/d\mathbf{x}^T$	
$\mathbf{f}(\mathbf{x}) = \mathbf{x}$	$\mathbf{I}_m$	(A.9)
$\mathbf{f}(\mathbf{x}) = \mathbf{Ay}$	$\dfrac{d\mathbf{A}}{d\mathbf{x}^T}(\mathbf{I}_m \otimes \mathbf{y}) + \mathbf{A}\dfrac{d\mathbf{y}}{d\mathbf{x}^T}$	(A.10)
$f(\mathbf{x}) = \mathbf{y}^T\mathbf{Qy}$	$2\mathbf{y}^T\mathbf{Q}\dfrac{d\mathbf{y}}{d\mathbf{x}^T} + (\mathbf{y}^T \otimes \mathbf{y}^T)\dfrac{d\mathrm{vec}(\mathbf{Q})}{d\mathbf{x}^T}$	(A.11)

$\mathbf{A} \in \mathbb{R}^{r \times s}$, $\mathbf{B} \in \mathbb{R}^{s \times r}$, $\mathbf{Q} = \mathbf{Q}^T \in \mathbb{R}^{s \times s}$, $\mathbf{x} \in \mathbb{R}^{m \times 1}$, $\mathbf{y} \in \mathbb{R}^{s \times 1}$

The expression of the first and second derivatives with respect to $\mathbf{x}$ are given by

$$\frac{\partial g}{\partial \mathbf{x}^T} = 2\mathbf{r}^T\left[\mathbf{A}\Sigma_\mathbf{d}\mathbf{A}^T\right]^{-1}\frac{\partial \mathbf{r}}{\partial \mathbf{x}^T} \tag{B.6}$$
$$+\left(\mathbf{r}^T \otimes \mathbf{r}^T\right)\frac{\partial}{\partial \mathbf{x}^T}\mathrm{vec}\left(\left[\mathbf{A}\Sigma_\mathbf{d}\mathbf{A}^T\right]^{-1}\right)$$

$$\frac{\partial^2 g}{\partial \mathbf{x}^2} = \frac{\partial^2 g}{\partial \mathbf{x}\partial \mathbf{x}^T} = \frac{\partial}{\partial \mathbf{x}^T}\left(\frac{\partial g}{\partial \mathbf{x}^T}\right)^T$$
$$= \frac{\partial}{\partial \mathbf{x}^T}\left(2\frac{\partial \mathbf{r}^T}{\partial \mathbf{x}}\left[\mathbf{A}\Sigma_\mathbf{d}\mathbf{A}^T\right]^{-1}\mathbf{r}\right)$$
$$= \frac{\partial}{\partial \mathbf{x}^T}\left(2\frac{\partial \mathbf{r}^T}{\partial \mathbf{x}}\left[\mathbf{A}\Sigma_\mathbf{d}\mathbf{A}^T\right]^{-1}\right)(\mathbf{I}_N \otimes \mathbf{r})$$
$$+ 2\frac{\partial \mathbf{r}^T}{\partial \mathbf{x}}\left[\mathbf{A}\Sigma_\mathbf{d}\mathbf{A}^T\right]^{-1}\frac{\partial \mathbf{r}}{\partial \mathbf{x}^T} \tag{B.7}$$

with

$$\frac{\partial \mathbf{r}}{\partial \mathbf{x}^T} = -\frac{\partial \mathbf{f}}{\partial \mathbf{x}^T}. \tag{B.8}$$

Neglecting the second-order term $(\mathbf{r}^T \otimes \mathbf{r}^T)$ in (B.6) and the first-order term $(\mathbf{I}_N \otimes \mathbf{r})$ in (B.7), since $\mathbf{r} \approx \mathbf{0}_{M \times 1}$, the approximate expressions of the first and the second-order derivatives with respect to $\mathbf{x}$ of the function g are

$$\frac{\partial g}{\partial \mathbf{x}^T} \approx -2\mathbf{r}^T[\mathbf{A}\Sigma_\mathbf{d}\mathbf{A}^T]^{-1}\frac{\partial \mathbf{f}}{\partial \mathbf{x}^T} \tag{B.9}$$

$$\frac{\partial^2 g}{\partial \mathbf{x}^2} \approx 2\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}}[\mathbf{A}\Sigma_\mathbf{d}\mathbf{A}^T]^{-1}\frac{\partial \mathbf{f}}{\partial \mathbf{x}^T}. \tag{B.10}$$

The expression of the first derivative with respect to $\mathbf{d}$ is given by

$$\frac{\partial g}{\partial \mathbf{d}^T} = 2\mathbf{r}^T[\mathbf{A}\Sigma_\mathbf{d}\mathbf{A}^T]^{-1}\frac{\partial \mathbf{r}}{\partial \mathbf{d}^T} \tag{B.11}$$
$$+(\mathbf{r}^T \otimes \mathbf{r}^T)\frac{\partial}{\partial \mathbf{d}^T}\mathrm{vec}([\mathbf{A}\Sigma_\mathbf{d}\mathbf{A}^T]^{-1})$$

with [see definitions (9) and (11)]

$$\frac{\partial \mathbf{r}}{\partial \mathbf{d}^T} = \begin{bmatrix} \dfrac{\partial \mathbf{r}}{\partial \mathbf{y}^T} & \dfrac{\partial \mathbf{r}}{\partial \boldsymbol{\pi}^T} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_M & -\dfrac{\partial \mathbf{f}}{\partial \boldsymbol{\pi}^T} \end{bmatrix} = \mathbf{A}. \tag{B.12}$$

The resulting approximate expressions of the first and the second-order derivatives with respect to $\mathbf{d}$ of the function g

are

$$\frac{\partial g}{\partial \mathbf{d}^T} \approx 2\mathbf{r}^T [\mathbf{A}\boldsymbol{\Sigma_d}\mathbf{A}^T]^{-1}\mathbf{A} \tag{B.13}$$

$$\frac{\partial^2 g}{\partial \mathbf{d}^2} \approx 2\mathbf{A}^T [\mathbf{A}\boldsymbol{\Sigma_d}\mathbf{A}^T]^{-1}\mathbf{A}. \tag{B.14}$$

For what concerns the mixed second derivative of g, it results by partially differentiating (B.9) with respect to $\mathbf{d}$ and considering that the vector $(\mathbf{r}^T \otimes \mathbf{r}^T)$ is negligible because $\mathbf{r} \approx \mathbf{0}$ together with (B.12)

$$\frac{\partial^2 g}{\partial \mathbf{x}^T \partial \mathbf{d}} \approx -2\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}}[\mathbf{A}\boldsymbol{\Sigma_d}\mathbf{A}^T]^{-1}\mathbf{A}. \tag{B.15}$$

APPENDIX C
ANALYSIS OF THE INITIAL STATE ERROR CONTRIBUTION TO THE MEASUREMENT UNCERTAINTY

In Section IV, the expression (21) of the augmented measurement error was introduced. In particular, the term representing the initial state error contribution using (A.10) can be written as

$$\frac{\partial}{\partial \mathbf{x}^T}\left(\frac{\partial \mathbf{f}}{\partial \boldsymbol{\pi}^T}\Delta\boldsymbol{\pi}\right)_{\mathbf{x}_0,\boldsymbol{\pi}} \Delta\mathbf{x}_0$$
$$= \left(\frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^T \partial \boldsymbol{\pi}^T}\right)_{\mathbf{x}_0,\boldsymbol{\pi}} (\mathbf{I}_N \otimes \Delta\boldsymbol{\pi})\Delta\mathbf{x}_0 = \left(\frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^T \partial \boldsymbol{\pi}^T}\right)_{\mathbf{x}_0,\boldsymbol{\pi}} \Delta\boldsymbol{\epsilon} \tag{C.1}$$

with [see (A.7)]

$$\Delta\boldsymbol{\epsilon} = (\mathbf{I}_N \otimes \Delta\boldsymbol{\pi})\Delta\mathbf{x}_0 = (\mathbf{I}_N \otimes \Delta\boldsymbol{\pi})\text{vec}(\Delta\mathbf{x}_0)$$
$$= \text{vec}(\Delta\boldsymbol{\pi} \cdot \Delta\mathbf{x}_0^T) = \Delta\mathbf{x}_0 \otimes \Delta\boldsymbol{\pi}. \tag{C.2}$$

The covariance matrix representing the contribution of the initial state error is defined by the following expectation [34]:

$$E\left\{\left(\frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^T \partial \boldsymbol{\pi}^T}\right)_{\mathbf{x}_0,\boldsymbol{\pi}} \Delta\boldsymbol{\epsilon} \cdot \Delta\boldsymbol{\epsilon}^T \left(\frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^T \partial \boldsymbol{\pi}^T}\right)_{\mathbf{x}_0,\boldsymbol{\pi}}^T\right\}$$
$$= \left(\frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^T \partial \boldsymbol{\pi}^T}\right)_{\mathbf{x}_0,\boldsymbol{\pi}} E\{\Delta\boldsymbol{\epsilon}\Delta\boldsymbol{\epsilon}^T\} \left(\frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^T \partial \boldsymbol{\pi}^T}\right)_{\mathbf{x}_0,\boldsymbol{\pi}}^T. \tag{C.3}$$

Since according to (A.4)

$$\Delta\boldsymbol{\epsilon} \cdot \Delta\boldsymbol{\epsilon}^T = (\Delta\mathbf{x}_0 \otimes \Delta\boldsymbol{\pi}) \cdot (\Delta\mathbf{x}_0^T \otimes \Delta\boldsymbol{\pi}^T)$$
$$= (\Delta\mathbf{x}_0 \Delta\mathbf{x}_0^T \otimes \Delta\boldsymbol{\pi} \Delta\boldsymbol{\pi}^T) \tag{C.4}$$

if the parameters errors and the initial state errors are uncorrelated from (C.3), the covariance contribution of the initial state errors to the augmented measurement error is:

$$\boldsymbol{\Sigma_\epsilon} = \left(\frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^T \partial \boldsymbol{\pi}^T}\right)_{\mathbf{x}_0,\boldsymbol{\pi}} \boldsymbol{\Sigma_{\mathbf{x}_0}} \otimes \boldsymbol{\Sigma_\pi} \left(\frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^T \partial \boldsymbol{\pi}^T}\right)_{\mathbf{x}_0,\boldsymbol{\pi}}^T. \tag{C.5}$$

APPENDIX D
MULTIVARIATE HOTELLING'S T^2 TEST STATISTIC

The Hotelling's T-squared distribution [48] is the multivariate generalization of Student's t-distribution. It is utilized to decide about the value of a population mean vector $\boldsymbol{\mu}$. Given a plausible value $\boldsymbol{\mu}_0$ for the population mean, the decision is the result of a proper statistical hypothesis testing problem [49] with the following two competing hypothesis.

1) $\mathscr{H}_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ (null hypothesis).
2) $\mathscr{H}_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$.

Harold Hotelling [48] proved that, if $\bar{\mathbf{x}}$ is the sampling mean vector and $\mathbf{S}$ is the sampling variance–covariance matrix obtained from a random sample of size n drawn from a p-variate Gaussian population with mean $\boldsymbol{\mu}_0$, the random quadratic residual

$$T^2 = n(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)^T \mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0) \tag{D.1}$$

has the following cumulative distribution function:

$$\frac{n - p}{p(n - 1)}T^2 \sim F_{p,n-p} \tag{D.2}$$

with $F_{p,n-p}$ representing the Fisher's F-distribution with p degrees of freedom for the numerator and $n - p$ for the denominator. The random variable T^2 is called the Hotelling's T-squared statistic.

Thus, if $\boldsymbol{\mu}$ values were specified to be $\boldsymbol{\mu}_0$, this could be tested by taking a single p-variate sample of size n, then computing T^2, and comparing it with the critical value

$$\overline{T}_\alpha^2 = \frac{p(n - 1)}{n - p}F_{p,n-p}^{-1}(1 - \alpha) \tag{D.3}$$

for a suitably chosen level of significance α. At significance level α, we reject the null hypothesis if the Hotelling's T^2 statistic exceed the critical value $\overline{T}_\alpha^2$

$$T^2 > \overline{T}_\alpha^2. \tag{D.4}$$

If the random sample is drawn from a nonnormal p-variate distribution, the Hotelling's T^2 statistics fails to satisfy the Fisher's distribution (D.2). In [41], a generalization of this test is presented valid under much weaker condition of the random sample distribution. The $\mu_0 = 0$ hypothesis is tested substituting the original p-variate random sample $\mathbf{x}$ of size n by the following random sample:

$$\mathbf{z} = \text{sign}(\mathbf{x}) \tag{D.5}$$

and considering the statistic

$$n\widetilde{R}^2 = n\bar{\mathbf{z}}\left(\frac{\mathbf{S}_z}{n}\right)^{-1}\bar{\mathbf{z}}^T \tag{D.6}$$

where $\bar{\mathbf{z}}$ is the sampling mean vector, and $\mathbf{S}_z$ is the sampling variance–covariance matrix of the random variable $\mathbf{z}$. The statistic $n\widetilde{R}^2$ is a multivariate generalization of the sign test [50], which is distributed according to the chi-squared distribution with p degrees of freedom [41]

$$n\widetilde{R}^2 \sim \chi_p. \tag{D.7}$$

For the statistic $n\widetilde{R}^2$, at significant level α, the upper-tail one-sided test critical value is

$$\overline{n\widetilde{R}_\alpha^2} = Q_p^{-1}(1 - \alpha) \tag{D.8}$$

where Q_p is the chi-square cumulative distribution function with p degrees of freedom.

References

[1] 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.

[2] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, Feb. 2000.

[3] A. Abur and A. G. Expósito, *Power System State Estimation: Theory and Implementation*, vol. 24. Boca Raton, FL, USA: CRC Press, 2004.

[4] A. Bose and K. A. Clements, "Real-time modeling of power networks," *Proc. IEEE*, vol. 75, no. 12, pp. 1607–1622, Dec. 1987.

[5] Z. Bubnicki, *Analysis and Decision Making in Uncertain Systems*. London, U.K.: Springer, 2004.

[6] M. Assadian, R. J. Goddard, H. W. Hong, and D. French, "Field operational experiences with on line state estimator," *IEEE Trans. Power Syst.*, vol. 9, no. 1, pp. 50–58, Feb. 1994.

[7] B. Porretta and R. Dhillon, "Performance evaluation of state estimation from line flow measurements on Ontario Hydro power system," *IEEE Trans. Power App. Syst.*, vol. PAS-92, no. 5, pp. 1702–1712, Sep. 1973.

[8] A. P. S. Meliopoulos, B. Fardanesh, and S. Zelingher, "Power system state estimation: Modeling error effects and impact on system operation," in *Proc. 34th Annu. Hawaii Int. Conf. Syst. Sci.*, Jan. 2001, pp. 682–690.

[9] T. A. Stuart and C. J. Herczet, "A sensitivity analysis of weighted least squares state estimation for power systems," *IEEE Trans. Power App. Syst.*, vol. PAS-92, no. 5, pp. 1696–1701, Sep. 1973.

[10] G. Sivanagaraju, S. Chakrabarti, and S. C. Srivastava, "Uncertainty in transmission line parameters: Estimation and impact on line current differential protection," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 6, pp. 1496–1504, Jun. 2014.

[11] F. Schweppe and J. Wildes, "Power system static-state estimation, part I: Exact model," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 1, pp. 120–125, Jan. 1970.

[12] F. C. Schweppe and D. B. Rom, "Power system static-state estimation, part II: Approximate model," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 1, pp. 125–130, Jan. 1970.

[13] F. C. Schweppe, "Power system static-state estimation, part III: Implementation," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 1, pp. 130–135, Jan. 1970.

[14] G. D'Antona, "Uncertainty of power system state estimates due to measurements and network parameter uncertainty," in *Proc. IEEE Int. Workshop Appl. Meas. Power Syst. (AMPS)*, Sep. 2010, pp. 37–40.

[15] G. D'Antona and M. Davoudi, "Effects of parameter and measurement uncertainties on the power system WLS state estimation," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf. (I2MTC)*, May 2012, pp. 1015–1020.

[16] P. Zarco and A. G. Expósito, "Power system parameter estimation: A survey," *IEEE Trans. Power Syst.*, vol. 15, no. 1, pp. 216–222, Feb. 2000.

[17] A. Debs, "Parameter estimation for power systems in the steady state," *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 882–886, Dec. 1974.

[18] J. Zhu and A. Abur, "Identification of network parameter errors," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 586–592, May 2006.

[19] J. Zhu and A. Abur, "Improvements in network parameter error identification via synchronized phasors," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 44–50, Feb. 2010.

[20] W.-H. E. Liu, F. F. Wu, and S.-S. Lun, "Estimation of parameter errors from measurement residuals in state estimation [power systems]," *IEEE Trans. Power Syst.*, vol. 7, no. 1, pp. 81–89, Feb. 1992.

[21] T. Van Cutsem and V. H. Quintana, "Network parameter estimation using online data with application to transformer tap position estimation," *IEE Proc. C Generat., Transmiss. Distrib.*, vol. 135, no. 1, pp. 31–40, Jan. 1988.

[22] 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.

[23] W.-H. E. Liu and S.-L. Lim, "Parameter error identification and estimation in power system state estimation," *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 200–209, Feb. 1995.

[24] P. A. Teixeira, S. R. Brammer, W. L. Rutz, W. C. Merritt, and J. L. Salmonsen, "State estimation of voltage and phase-shift transformer tap settings," in *Proc. Power Ind. Comput. Appl. Conf.*, May 1991, pp. 416–422.

[25] S. A. Arafeh and R. Schinzinger, "Estimation algorithms for large-scale power systems," *IEEE Trans. Power App. Syst.*, vol. PAS-98, no. 6, pp. 1968–1977, Nov. 1979.

[26] A. S. Debs, "Estimation of steady-state power system model parameters," *IEEE Trans. Power App. Syst.*, vol. PAS-93, no. 5, pp. 1260–1268, Sep. 1974.

[27] 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.

[28] JCGM, BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement*, 1st ed., document JCGM 100:2008, Sep. 2008.

[29] JCGM, BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, *Evaluation of Measurement Data—An Introduction to the 'Guide to the Expression of Uncertainty in Measurement' and Related Documents*, 1st ed., document JCGM 104:2009, Jul. 2009.

[30] JCGM, BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, *Evaluation of Measurement Data—Supplement 2 to the 'Guide to the Expression of Uncertainty in Measurement'—Extension to Any Number of Output Quantities*, 1st ed., document JCGM 102:2011, Oct. 2011.

[31] A. K. Al-Othman and M. R. Irving, "Uncertainty modelling in power system state estimation," *IET Generat., Transmiss. Distrib.*, vol. 152, no. 2, pp. 233–239, Mar. 2005.

[32] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus With Applications in Statistics and Econometrics*. New York, NY, USA: Wiley, 1999.

[33] P. C. Mahalanobis, "On the generalized distance in statistics," in *Proc. Nat. Inst. Sci. India*, vol. 2, no. 1, pp. 49–55, 1936.

[34] G. D'Antona, "Measurement data processing using random matrices: A generalized formula for the propagation of uncertainty," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 2, pp. 537–545, Apr. 2004.

[35] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*. New York, NY, USA: Wiley, 1971, vol. 7.

[36] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[37] T. J. Abatzoglou, J. M. Mendel, and G. A. Harada, "The constrained total least squares technique and its applications to harmonic superresolution," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1070–1087, May 1991.

[38] G. D'Antona, "The full least-squares method," *IEEE Trans. Instrum. Meas.*, vol. 52, no. 1, pp. 189–196, Feb. 2003.

[39] C. H. Edwards, *Advanced Calculus of Several Variables*. New York, NY, USA: Dover, 1994.

[40] C. J. Mecklin and D. J. Mundfrom, "An appraisal and bibliography of tests for multivariate normality," *Int. Statist. Rev.*, vol. 72, no. 1, pp. 123–138, 2004.

[41] M. L. Eaton and B. Efron, "Hotelling's T^2 test under non-normal conditions," Dept Statist., Stanford Univ., Stanford, CA, USA, Tech. Rep. 11, Aug. 1968.

[42] W. H. Marlow, *Mathematics for Operations Research*. New York, NY, USA: Dover, 1993.

[43] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740–2746, Jun. 2007.

[44] W. J. Vetter, "Matrix calculus operations and Taylor expansions," *SIAM Rev.*, vol. 15, no. 2, pp. 352–369, 1973.

[45] P. M. Bentler and S.-Y. Lee, "Matrix derivatives with chain rule and rules for simple, Hadamard, and Kronecker products," *J. Math. Psychol.*, vol. 17, no. 3, pp. 255–262, 1978.

[46] A. Weinmann, *Uncertain Models and Robust Control*. Vienna, Austria: Springer-Verlag, 1991.

[47] A. Hjorungnes, D. Gesbert, and D. P. Palomar, "Unified theory of complex-valued matrix differentiation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, vol. 3. Apr. 2007, pp. III-345–III-348.

[48] H. Hotelling, "The generalization of student's ratio," *Ann. Math. Statist.*, vol. 2, no. 3, pp. 360–378, 1931.

[49] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis* (Probability and Mathematical Statistics), 1st ed. San Diego, CA, USA: Academic, 1980.

[50] W. J. Conover, *Practical Nonparametric Statistics*. New York, NY, USA: Wiley, 1998.

Gabriele D'Antona received the M.Sc. degree in electrical engineering and the Ph.D. degree in electrical science from the Politecnico di Milano, Milan, Italy, in 1989 and 1994, respectively.

He was with the Joint European Torus Laboratory, Abingdon, U.K., from 1990 to 1996. From 1996 to 2007, he was with the Department of Electrical Engineering, Politecnico di Milano, where he became an Associate Professor of Electrical and Electronic Measurements in 2002. Since 2008, he has been with the Electrical Engineering Division, Department of Energy, Politecnico di Milano. His current research interests include the development of measurement systems based on probabilistic and variational signal processing techniques for the diagnosis and control of spatially distributed systems.