Chapter 6

Static and recursive PMU-based state estimation processes for transmission and distribution power grids

Mario Paolone, Jean-Yves Le Boudec, Styliani Sarri and Lorenzo Zanni

In the operation of power systems, the knowledge of the system state is required by several fundamental functions, such as security assessment, voltage control and stability analysis. By making reference to the *static* state of the system represented by the voltage phasors at all the network buses, it is possible to infer the system operating conditions. Until the late 1970s, conventional load flow calculations provided the system state by directly using the raw measurements of voltage magnitudes and power injections. The loss of one measurement made the calculation impossible and the presence of measurement errors affected dramatically the computed state. To overcome these limitations, load flow theory has been combined with statistical estimation constituting the so-called *state estimation* (SE). The latter consists in the solution of an optimization problem that processes the measurements together with the network model to determine the optimal estimate of the system state. The outputs of load flow and SE are composed of the same quantities, typically the voltage magnitude and phase at all the network buses, but SE uses all the types of measurements (e.g., voltage and current magnitudes, nodal power injections and flows, synchrophasors) and evaluates their consistency using the network model. The measurement redundancy is key to tolerate measurement losses, identify measurement and network parameter errors, and filter out the measurement noise. The foregoing properties of SE allow the system operator to obtain an accurate and reliable estimate of the system state that consequently improves the performance of the functions relying on it.

Traditionally, SE has been performed at a relatively low refresh rate of a few minutes, dictated by the time requirements of the related functions together with the low measurement acquisition rate of remote terminal units (RTUs).[1] Nowadays, the emerging availability of phasor measurement units (PMUs) allows to acquire

[1]RTUs are devices installed at the network substations that regularly send measurements, usually unsynchronized, to the network operator control centre. Typically, these measurements are composed of voltage and current magnitudes as well as active and reactive powers.

accurate and time-aligned phasors, called *synchrophasors*, with typical streaming rates in the order of some tens of measurements per second [1, 2]. This technology is experiencing a fast evolution, which is triggered by an increasing number of power system applications that can benefit from the use of synchrophasors. SE processes can exploit the availability of synchrophasor measurements to achieve better accuracy performance and higher refresh rate (sub-second).

PMUs already compose the backbone of wide area monitoring systems in the context of transmission networks to which several real-time functionalities are connected, such as inter-area oscillations, relaying, fault location and real-time SE [3, 4]. However, PMUs might represent fundamental monitoring tools even in the context of distribution networks for applications such as: SE [5, 6], loss of main [7], fault event monitoring [8], synchronous islanded operation [9] and power quality monitoring [10]. The recent literature has discussed the use of PMUs for SE in distribution networks both from the methodological point of view [11] and also via dedicated real-scale experimental setups [12, 13].

Since the pioneering works of Schweppe on power system SE in 1970 [14–16], most of the research on the subject has investigated *static* SE methods based on weighted least squares (WLS) [17–19]. Static SE computes the system state performing a "best fit" of the measurements belonging only to the current time-step. Another category of state estimators are the *recursive* methods, such as the Kalman filter (KF). In addition to the use of the measurements and their statistical properties, they also predict the system state by modelling its time evolution. In general, recursive estimators are characterized by higher complexity and the prediction introduces an additional source of uncertainty that, if not properly quantified, might worsen the accuracy of the estimated state. Besides, their ability to filter out measurement noise could not be exploited due to the low SE refresh rate: even in quasi-steady state conditions, the measurement noise was smaller than the state variations between two consecutive time-steps. However, the effectiveness of power system SE based on KF has been recently reconsidered thanks to the possibility to largely increase the SE refresh rate by using synchrophasor measurements.

The chapter starts by providing the measurement and process model of WLS and KF SE algorithms and continues with the analytical formulation of the two families of state estimators, including their linear and non-linear versions as a function of the type of available measurements. Finally, two case studies targeting IEEE transmission and distribution reference networks are given.

6.1 State estimation measurement and process model

In general, any SE algorithm relies on a *measurement model*, which expresses how the system state variables are related to the measurements and the measurement noise. The link between the state variables and the measurements can be either linear or non-linear, depending on the kind of measurements used. For the case of power systems, the link is represented by the network model that is composed of the network topology and the electrical parameters of the various components, such as transmission lines

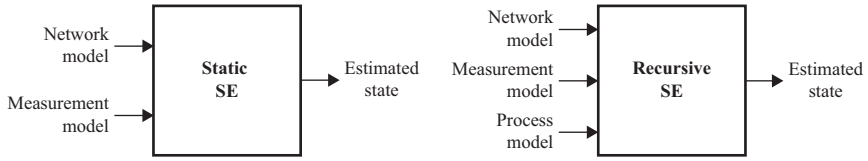

Figure 6.1 Inputs and outputs of static and recursive state estimators

and transformers. Additionally to the measurement model, recursive state estimators, such as KF, use a *process model* to represent the time evolution of the system state as a function of the previous system states, the controllable inputs and the process noise. Figure 6.1 shows the inputs and outputs of static and recursive state estimators.

6.1.1 Measurement model

The measurement model of the SE can be formulated in a common way regardless of the estimator type (static or recursive). In particular, if the measurements come only from PMUs, the measurement model is linear, whereas if the measurements come from RTUs or a hybrid set of RTUs and PMUs, the measurement model is non-linear. In what follows, examples of formulations of these measurement models are given.

6.1.1.1 Linear measurement model

In case the SE uses only measurements coming from PMUs composed of voltage and/or current phasors, the SE can be formulated in a linear way. The state variables are represented by the phase-to-ground nodal voltages, or the branch voltages, or nodal current injections or current flows (or even a mix of them, provided that they are independent state variables).

By defining the set of network buses $\mathcal{S}$, the number of network buses is equal to $s = |\mathcal{S}|$, where the operator $|\ \ |$ denotes the cardinality of a set. Then, the state of a three-phase (3-ph) network with s buses is denoted by $\mathbf{x} \in \mathbb{R}^n$ (where $n = 3 \cdot 2s$ is the number of state variables that compose the set of state variables $\mathcal{N}$) that, in most of the literature on the subject, is represented by the phase-to-ground nodal voltages. The set of three phases a, b, c is denoted by $\mathcal{P}$. Also, the set of network branches is denoted by $\mathcal{B}$. To obtain an exact linear measurement model, measurements and state variables are expressed in rectangular coordinates. Hence, the state is composed of the real and imaginary parts of the voltage phasors at every bus:

$$\mathbf{x} = [\mathbf{V}_{1,re}^{a,b,c}, \ldots, \mathbf{V}_{i,re}^{a,b,c}, \ldots, \mathbf{V}_{s,re}^{a,b,c}, \mathbf{V}_{1,im}^{a,b,c}, \ldots, \mathbf{V}_{i,im}^{a,b,c}, \ldots, \mathbf{V}_{s,im}^{a,b,c}]^T \tag{6.1}$$

where

$$\begin{aligned} \mathbf{V}_{i,re}^{a,b,c} &= [\mathbf{V}_{i,re}^{a}, \mathbf{V}_{i,re}^{b}, \mathbf{V}_{i,re}^{c}] \\ \mathbf{V}_{i,im}^{a,b,c} &= [\mathbf{V}_{i,im}^{a}, \mathbf{V}_{i,im}^{b}, \mathbf{V}_{i,im}^{c}] \end{aligned} \tag{6.2}$$

are, respectively, the real and imaginary parts of the voltage phasor at bus i in the three phases a, b and c. It is worth mentioning that, as explained in Reference 20, the existence of PMU measurements can eliminate the need to choose a reference bus, i.e., a bus where the phase or the imaginary part of the voltage is not included in the state and is assumed to have a certain value (usually equal to zero). Therefore, unlike the conventional SE formulation, in this case the phase or the imaginary part of the voltage is estimated at every bus.

To provide an example of linear measurement model, it is assumed that measurements come only from PMUs that measure nodal voltage phasors, current injection phasors and current flow phasors. The set of buses where PMUs measure nodal voltage phasors is $\mathcal{D}_1$ ($d_1 = |\mathcal{D}_1|$). Similarly, the set of buses where current injection phasors are measured is $\mathcal{D}_2$ ($d_2 = |\mathcal{D}_2|$) and the set of branches where current flow phasors are measured is $\mathcal{D}_3$ ($d_3 = |\mathcal{D}_3|$). Hence, the set of measurements $\mathcal{M}$ is composed of:

- $3d_1$ phase-to-ground voltage phasors;
- $3d_2$ current injection phasors;
- $3d_3$ current flow phasors;

and its cardinality is equal to $m = 2 \cdot (3d_1 + 3d_2 + 3d_3)$. Therefore, the measurement array $\mathbf{z} \in \mathbb{R}^m$ is equal to:

$$\mathbf{z} = [\mathbf{z}_V, \mathbf{z}_{I_{\text{inj}}}, \mathbf{z}_{I_{\text{flow}}}]^T \tag{6.3}$$

where

$$\begin{aligned}
\mathbf{z}_V &= [\mathbf{V}^{a,b,c}_{1,re}, \ldots, \mathbf{V}^{a,b,c}_{d_1,re}, \mathbf{V}^{a,b,c}_{1,im}, \ldots, \mathbf{V}^{a,b,c}_{d_1,im}] \\
\mathbf{z}_{I_{\text{inj}}} &= [\mathbf{I}^{a,b,c}_{\text{inj},1,re}, \ldots, \mathbf{I}^{a,b,c}_{\text{inj},d_2,re}, \mathbf{I}^{a,b,c}_{\text{inj},1,im}, \ldots, \mathbf{I}^{a,b,c}_{\text{inj},d_2,im}] \\
\mathbf{z}_{I_{\text{flow}}} &= [\mathbf{I}^{a,b,c}_{\text{flow},1,re}, \ldots, \mathbf{I}^{a,b,c}_{\text{flow},d_3,re}, \mathbf{I}^{a,b,c}_{\text{flow},1,im}, \ldots, \mathbf{I}^{a,b,c}_{\text{flow},d_3,im}].
\end{aligned} \tag{6.4}$$

The convention employed in this chapter for the currents injections, and also the most common, is to consider with positive sign the current flowing from the loads/generators to the respective bus.

The equation that relates the measurements with the system state variables is:

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \boldsymbol{\varepsilon} \tag{6.5}$$

where $\mathbf{H}$ is an $m \times n$ matrix that represents the link between measurements and state variables and $\boldsymbol{\varepsilon}$ is the measurement noise. It is important to point out that, in the case of linear SE, $\mathbf{H}$ does not represent a linear approximation of the measurement model, since it corresponds to the exact link between measurements and state variables. The measurement noise is assumed to be white and Gaussian:

$$p(\boldsymbol{\varepsilon}) \sim \mathcal{N}(0, \mathbf{R}) \tag{6.6}$$

where $\mathbf{R}$ is the measurement noise covariance matrix. Note that the normality of PMU errors is based on experimental evidences of error distributions of actual PMUs (e.g., Reference 21). Therefore, the diagonal entries of $\mathbf{R}$ represent the variances

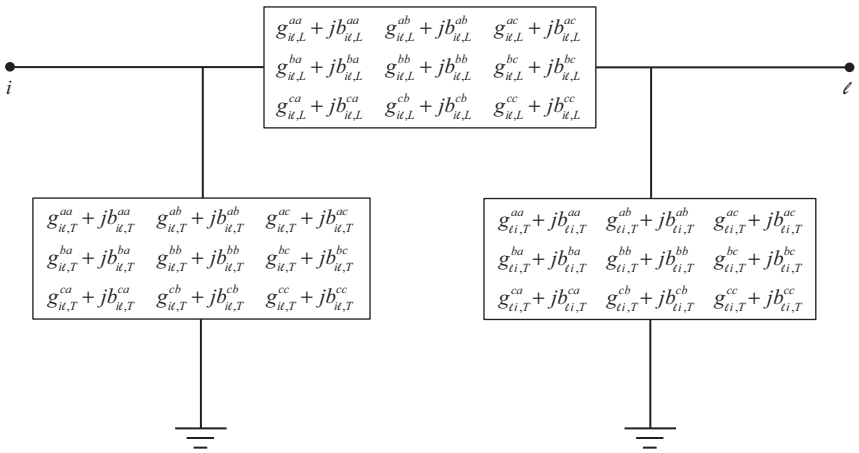

Figure 6.2 3-ph two-port π-model of a generic network branch

of the measurements, which correspond to the cumulative uncertainty of sensors[2] and meters. The off-diagonal entries account for eventual correlation between the measurements that occurs if mutual influence among meters is present.

For the case of power networks, matrix **H** of (6.5) is derived from the network topology and the electrical parameters of the various network components. To this end, the so-called *nodal analysis* linearly links the nodal current injections with the considered system state, i.e., the phase-to-ground nodal voltages. As it is described in what follows, such link consists in the so-called *compound network admittance matrix* **Y**, whose derivation is given below.

A generic 3-ph network branch between buses i and ℓ can be represented by the two-port π-model shown in Figure 6.2. Its parameters are 3×3 matrices of complex numbers, i.e., the longitudinal admittance $\mathbf{y}_{i\ell,L}$ and the two transverse admittances $\mathbf{y}_{i\ell,T}$ and $\mathbf{y}_{\ell i,T}$:

$$\mathbf{y}_{i\ell,L} = \mathbf{g}_{i\ell,L} + j\mathbf{b}_{i\ell,L} \tag{6.7}$$

$$\mathbf{y}_{i\ell,T} = \mathbf{g}_{i\ell,T} + j\mathbf{b}_{i\ell,T} \tag{6.8}$$

$$\mathbf{y}_{\ell i,T} = \mathbf{g}_{\ell i,T} + j\mathbf{b}_{\ell i,T} \tag{6.9}$$

where

- j is the imaginary unit;
- $\mathbf{g}_{i\ell,L}$ and $\mathbf{b}_{i\ell,L}$ are, respectively, the longitudinal conductance and susceptance;

[2]In this chapter, the sensors refer to the transducers (such as voltage or current instrument transformers) that scale down the input voltage and current signals in order to interface the meters (represented, in this case, by PMUs and/or RTUs) with the electrical network.

- $\mathbf{g}_{i\ell,T}$ and $\mathbf{b}_{i\ell,T}$ are, respectively, the transverse conductance and susceptance from the side of bus i.
- $\mathbf{g}_{\ell i,T}$ and $\mathbf{b}_{\ell i,T}$ are, respectively, the transverse conductance and susceptance from the side of bus ℓ;

The 3×3 matrix representing the longitudinal admittance is given by (the same matrix form applies also to the transverse admittances as shown in Figure 6.2):

$$\mathbf{y}_{i\ell,L} = \begin{bmatrix} g^{aa}_{i\ell,L} + jb^{aa}_{i\ell,L} & g^{ab}_{i\ell,L} + jb^{ab}_{i\ell,L} & g^{ac}_{i\ell,L} + jb^{ac}_{i\ell,L} \\ g^{ba}_{i\ell,L} + jb^{ba}_{i\ell,L} & g^{bb}_{i\ell,L} + jb^{bb}_{i\ell,L} & g^{bc}_{i\ell,L} + jb^{bc}_{i\ell,L} \\ g^{ca}_{i\ell,L} + jb^{ca}_{i\ell,L} & g^{cb}_{i\ell,L} + jb^{cb}_{i\ell,L} & g^{cc}_{i\ell,L} + jb^{cc}_{i\ell,L} \end{bmatrix} \tag{6.10}$$

The expression of these parameters depends on the type of network branch. For instance, a 3-ph transmission line characterized by a longitudinal impedance $\mathbf{z}_{\text{line}}$ and a transverse admittance $\mathbf{y}_{\text{line}}$ can be represented by the following two-port π-model parameters:

$$\mathbf{y}_{i\ell,L} = \mathbf{z}^{-1}_{\text{line}} \tag{6.11}$$

$$\mathbf{y}_{i\ell,T} = \mathbf{y}_{\ell i,T} = \mathbf{y}_{\text{line},T}/2 \tag{6.12}$$

Note that $\mathbf{y}_{i\ell,T}$ can be different from $\mathbf{y}_{\ell i,T}$ for some network components, such as tap changing or phase-shifting transformers. For the particular case of phase-shifting transformers, the two-port network is not reciprocal.

A detailed description of the procedure to construct the network admittance matrix is given in Reference 17. Considering a 3-ph network of s buses, $\mathbf{Y}$ is a $3s \times 3s$ matrix with the following form:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{11} & \mathbf{Y}_{12} & \cdots & \mathbf{Y}_{1s} \\ \mathbf{Y}_{21} & \mathbf{Y}_{22} & \cdots & \mathbf{Y}_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{Y}_{s1} & \mathbf{Y}_{s2} & \cdots & \mathbf{Y}_{ss} \end{bmatrix} \tag{6.13}$$

where

- the off-diagonal element $\mathbf{Y}_{i\ell}$ ($i \neq \ell$) is a 3×3 matrix equal to the opposite of the longitudinal admittance of the branch between buses i and ℓ:

  $$\mathbf{Y}_{i\ell} = -\mathbf{y}_{i\ell,L} \tag{6.14}$$

  and $\mathbf{Y}_{i\ell} = \mathbf{Y}_{\ell i}$, so that $\mathbf{Y}$ is symmetrical;
- the diagonal element $\mathbf{Y}_{ii}$ is a 3×3 matrix equal to the sum of the longitudinal admittances of the branches connected to bus i and the transverse admittances of the branches connected between bus i and the neutral:

  $$\mathbf{Y}_{ii} = \sum_{\ell=1}^{s} (\mathbf{y}_{i\ell,L} + \mathbf{y}_{i\ell,T}) \tag{6.15}$$

  Note that $\mathbf{y}_{i\ell,L}$ and $\mathbf{y}_{i\ell,T}$ are null matrices if bus i is not connected to bus ℓ.

The expression of the generic element $\mathbf{Y}_{i\ell}$ of the compound admittance matrix is:

$$\mathbf{Y}_{i\ell} = \begin{bmatrix} G_{i\ell}^{aa} + jB_{i\ell}^{aa} & G_{i\ell}^{ab} + jB_{i\ell}^{ab} & G_{i\ell}^{ac} + jB_{i\ell}^{ac} \\ G_{i\ell}^{ba} + jB_{i\ell}^{ba} & G_{i\ell}^{bb} + jB_{i\ell}^{bb} & G_{i\ell}^{bc} + jB_{i\ell}^{bc} \\ G_{i\ell}^{ca} + jB_{i\ell}^{ca} & G_{i\ell}^{cb} + jB_{i\ell}^{cb} & G_{i\ell}^{cc} + jB_{i\ell}^{cc} \end{bmatrix} \tag{6.16}$$

where **G** and **B** are the real and imaginary parts of the admittance matrix **Y**, respectively.

The expressions of the real and imaginary parts of the 3-ph nodal current injection phasors at bus $i \in \mathcal{S}$ and phase $p \in \mathcal{P}$ are:

$$I_{i,re}^{p} = \sum_{\ell=1}^{s} \sum_{l \in \mathcal{P}} \left[G_{i\ell}^{pl} V_{\ell,re}^{l} - B_{i\ell}^{pl} V_{\ell,im}^{l} \right] \tag{6.17}$$

$$I_{i,im}^{p} = \sum_{\ell=1}^{s} \sum_{l \in \mathcal{P}} \left[G_{i\ell}^{pl} V_{\ell,im}^{l} + B_{i\ell}^{pl} V_{\ell,re}^{l} \right] \tag{6.18}$$

where the subscripts i and ℓ refer to the bus indices whereas the superscripts p and l refer to the phase indices.

The current flow phasor $I_{\text{flow},u}^{p}$ at branch $u \in \mathcal{B}$ and phase $p \in \mathcal{P}$ can be also indicated as $I_{i\ell}^{p}$ with respect to the two terminal buses i and ℓ of this branch. Note that $I_{i\ell}^{p}$ is measured from the side of bus i, while $I_{\ell i}^{p}$ is measured from the side of bus ℓ. The expressions of the real and imaginary parts of the 3-ph current flow phasors at the branch between buses i and ℓ and phase $p \in \mathcal{P}$ are:

$$I_{i\ell,re}^{p} = \sum_{l \in \mathcal{P}} [g_{i\ell,L}^{pl}(V_{i,re}^{l} - V_{\ell,re}^{l}) - b_{i\ell,L}^{pl}(V_{i,im}^{l} - V_{\ell,im}^{l}) + g_{i\ell,T}^{pl} V_{i,re}^{l} - b_{i\ell,T}^{pl} V_{i,im}^{l}] \tag{6.19}$$

$$I_{i\ell,im}^{p} = \sum_{l \in \mathcal{P}} [g_{i\ell,L}^{pl}(V_{i,im}^{l} - V_{\ell,im}^{l}) + b_{i\ell,L}^{pl}(V_{i,re}^{l} - V_{\ell,re}^{l}) + g_{i\ell,T}^{pl} V_{i,im}^{l} + b_{i\ell,T}^{pl} V_{i,re}^{l}] \tag{6.20}$$

The structure of the matrix **H** of (6.5) is:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{V} \\ \mathbf{H}_{I_{\text{inj}}} \\ \mathbf{H}_{I_{\text{flow}}} \end{bmatrix} \tag{6.21}$$

The part related to the nodal voltage phasor measurements $\mathbf{H}_{V}$ is:

$$\mathbf{H}_{V} = \begin{bmatrix} \boldsymbol{\beta} & \boldsymbol{\upsilon} \\ \boldsymbol{\zeta} & \boldsymbol{\eta} \end{bmatrix} \tag{6.22}$$

where

$$\beta_{\ell l,re}^{ip,re} = \begin{cases} 1, \text{ if } i = \ell \text{ and } p = l \\ 0, \text{ if } i \neq \ell \text{ or } p \neq l \end{cases} \tag{6.23}$$

$$\upsilon_{\ell l,im}^{ip,re} = 0 \tag{6.24}$$

$$\zeta_{\ell l,re}^{ip,im} = 0 \tag{6.25}$$

$$\eta_{\ell l,im}^{ip,im} = \begin{cases} 1, \text{ if } i = \ell \text{ and } p = l \\ 0, \text{ if } i \neq \ell \text{ or } p \neq l \end{cases} \tag{6.26}$$

In (6.23)–(6.26), the superscripts refer, respectively, to the bus, the phase and the real or imaginary part of the measurements, while the subscripts refer to the state variables. For instance, $\beta_{\ell l,re}^{ip,re}$ is the scalar that links the measurement $V_{i,re}^{p}$ with the state variable $V_{\ell,re}^{l}$ (in this specific case, the scalar is simply zero or one).

The part related to the current injection phasor measurements $\mathbf{H}_{I_{\text{inj}}}$ can be derived in a straightforward way from (6.17) and (6.18):

$$\mathbf{H}_{I_{\text{inj}}} = \begin{bmatrix} \boldsymbol{\Theta} & \boldsymbol{\Lambda} \\ \boldsymbol{\Xi} & \boldsymbol{\Upsilon} \end{bmatrix} \tag{6.27}$$

where

$$\Theta_{\ell l,re}^{ip,re} = G_{i\ell}^{pl} \tag{6.28}$$

$$\Lambda_{\ell l,im}^{ip,re} = -B_{i\ell}^{pl} \tag{6.29}$$

$$\Xi_{\ell l,re}^{ip,im} = B_{i\ell}^{pl} \tag{6.30}$$

$$\Upsilon_{\ell l,im}^{ip,im} = G_{i\ell}^{pl} \tag{6.31}$$

In (6.28)–(6.31), the superscripts and subscripts have the same meaning as in (6.23)–(6.26).

The part related to the current flow phasor measurements $\mathbf{H}_{I_{\text{flow}}}$ can be derived in a straightforward way from (6.19) and (6.20):

$$\mathbf{H}_{I_{\text{flow}}} = \begin{bmatrix} \boldsymbol{\theta} & \boldsymbol{\vartheta} & \boldsymbol{\iota} & \boldsymbol{\kappa} \\ \boldsymbol{\lambda} & \boldsymbol{\nu} & \boldsymbol{\xi} & \boldsymbol{\varpi} \end{bmatrix} \tag{6.32}$$

where

$$\theta_{il,re}^{i\ell p,re} = g_{i\ell,L}^{pl} + g_{i\ell,T}^{pl} \tag{6.33}$$

$$\vartheta_{\ell l,re}^{i\ell p,re} = -g_{i\ell,L}^{pl} \tag{6.34}$$

$$\iota_{il,im}^{i\ell p,re} = -(b_{i\ell,L}^{pl} + b_{i\ell,T}^{pl}) \tag{6.35}$$

$$\kappa_{\ell l,im}^{i\ell p,re} = b_{i\ell,L}^{pl} \tag{6.36}$$

$$\lambda_{il,re}^{i\ell p,im} = b_{i\ell,L}^{pl} + b_{i\ell,T}^{pl} \tag{6.37}$$

$$\nu_{\ell l,re}^{i\ell p,im} = -b_{i\ell,L}^{pl} \tag{6.38}$$

$$\xi_{il,im}^{i\ell p,im} = g_{i\ell,L}^{pl} + g_{i\ell,T}^{pl} \tag{6.39}$$

$$\varpi_{\ell l,im}^{i\ell p,im} = -g_{i\ell,L}^{pl} \tag{6.40}$$

In (6.33)–(6.40), the superscripts refer, respectively, to the two terminal buses of the branch, the phase and the real or imaginary part of the measurements, while the subscripts refer to the state variables. For instance, $\theta_{il,re}^{i\ell p,re}$ is the scalar that links the measurement $I_{i\ell,re}^{p}$ with the state variable $V_{i,re}^{l}$.

6.1.1.2 Non-linear measurement model

In the case of a mixed set of measurements that includes PMUs and conventional power and magnitude[3] measurements, the SE becomes non-linear due to the non-linear equations that link the non-phasor measurements with the system state. In this case, the system state for a network with s buses is usually expressed in polar coordinates:

$$\mathbf{x} = [\boldsymbol{\delta}_1^{a,b,c}, \ldots, \boldsymbol{\delta}_i^{a,b,c}, \ldots, \boldsymbol{\delta}_s^{a,b,c}, \mathbf{V}_1^{a,b,c}, \ldots, \mathbf{V}_i^{a,b,c}, \ldots, \mathbf{V}_s^{a,b,c}]^T \tag{6.41}$$

where

$$\begin{aligned} \boldsymbol{\delta}_i^{a,b,c} &= [\delta_i^a, \delta_i^b, \delta_i^c] \\ \mathbf{V}_i^{a,b,c} &= [V_i^a, V_i^b, V_i^c] \end{aligned} \tag{6.42}$$

are, respectively, the phase and magnitude of the voltage phasor at bus i in the three phases a, b and c.

The measurements are assumed to come from PMUs that measure voltage phasors and from conventional power meters.[4] The set of network buses where active and reactive power injections are measured is $\mathcal{U}_1$ ($u_1 = |\mathcal{U}_1|$). Then, the set of network branches where active and reactive power flows are measured is $\mathcal{U}_2$ ($u_2 = |\mathcal{U}_2|$). Hence, it is assumed that the set of measurements $\mathcal{M}$ is composed of:

- $3d_1$ phase-to-ground voltage phasors;
- $3 \cdot (2u_1)$ active and reactive power injections;
- $3 \cdot (2u_2)$ active and reactive power flows.

[3]The *magnitude* indicates the amplitude of voltage and current signals.

[4]The magnitude and phase of current injections and flows could be added to the measurement set. The equations related to current magnitude measurements can be found in Reference 19.

and its cardinality is equal to $m = 2 \cdot (3d_1 + 3u_1 + 3u_2)$. Therefore, the measurement array $\mathbf{z} \in \mathbb{R}^m$ is equal to:[5]

$$\mathbf{z} = [\mathbf{z}_V, \mathbf{z}_{PQ_{\text{inj}}}, \mathbf{z}_{PQ_{\text{flow}}}]^T \tag{6.43}$$

where

$$\begin{aligned}
\mathbf{z}_V^T &= [\boldsymbol{\delta}_1^{a,b,c}, \ldots, \boldsymbol{\delta}_{d_1}^{a,b,c}, \mathbf{V}_1^{a,b,c}, \ldots, \mathbf{V}_{d_1}^{a,b,c}] \\
\mathbf{z}_{PQ_{\text{inj}}}^T &= [\mathbf{P}_{\text{inj},1}^{a,b,c}, \ldots, \mathbf{P}_{\text{inj},u_1}^{a,b,c}, \mathbf{Q}_{\text{inj},1}^{a,b,c}, \ldots, \mathbf{Q}_{\text{inj},u_1}^{a,b,c}] \\
\mathbf{z}_{PQ_{\text{flow}}}^T &= [\mathbf{P}_{\text{flow},1}^{a,b,c}, \ldots, \mathbf{P}_{\text{flow},u_2}^{a,b,c}, \mathbf{Q}_{\text{flow},1}^{a,b,c}, \ldots, \mathbf{Q}_{\text{flow},u_2}^{a,b,c}]
\end{aligned} \tag{6.44}$$

As stated before, in this case, the equation that links the measurements with the system state variables is:

$$\mathbf{z} = h(\mathbf{x}) + \boldsymbol{\varepsilon} \tag{6.45}$$

where the vector $h(\mathbf{x}) \in \mathbb{R}^m$ represents the non-linear function relating the system state variables to the measurements, i.e., the so-called *measurement function*. As a consequence, matrix $\mathbf{H}$ used in the SE process is a linearization of $h(\mathbf{x})$ and does not represent the exact link between measurements and state variables.

By using the same notation of (6.17) and (6.18), the expressions of the active and reactive power injections with respect to the state variables are given by:

$$P_i^p = V_i^p \sum_{\ell=1}^{s} \sum_{l \in \mathcal{P}} V_\ell^l (G_{i\ell}^{pl} \cos \delta_{i\ell}^{pl} + B_{i\ell}^{pl} \sin \delta_{i\ell}^{pl}) \tag{6.46}$$

$$Q_i^p = V_i^p \sum_{\ell=1}^{s} \sum_{l \in \mathcal{P}} V_\ell^l (G_{i\ell}^{pl} \sin \delta_{i\ell}^{pl} - B_{i\ell}^{pl} \cos \delta_{i\ell}^{pl}) \tag{6.47}$$

where $\delta_{i\ell}^{pl} = \delta_i^p - \delta_\ell^l$.

By using the same notation of (6.19) and (6.20), the expressions of the active and reactive power flows with respect to the state variables are given by:

$$\begin{aligned}
P_{i\ell}^p = {} & V_i^p \sum_{l \in \mathcal{P}} \{V_i^l [(g_{i\ell,L}^{pl} + g_{i\ell,T}^{pl}) \cos \delta_{ii}^{pl} + (b_{i\ell,L}^{pl} + b_{i\ell,T}^{pl}) \sin \delta_{ii}^{pl}]\} \\
& - V_i^p \sum_{l \in \mathcal{P}} [V_\ell^l (g_{i\ell,L}^{pl} \cos \delta_{i\ell}^{pl} + b_{i\ell,L}^{pl} \sin \delta_{i\ell}^{pl})]
\end{aligned} \tag{6.48}$$

[5]In general, the voltage phasor measurements from PMUs and the power measurements from RTUs are not obtained at the same time. In particular, the RTU measurements are characterized by lack of global positioning system (GPS) synchronization, lower refresh rates and lower accuracy compared to the PMU measurements. However, here it is assumed that at time t a full set of measurements composed of synchronized voltage phasor and power measurements is available, as shown in (6.43).

$$Q_{i\ell}^{p} = V_i^p \sum_{l\in\mathcal{P}} \{V_i^l [(g_{i\ell,L}^{pl} + g_{i\ell,T}^{pl}) \sin\delta_{ii}^{pl} - (b_{i\ell,L}^{pl} + b_{i\ell,T}^{pl}) \cos\delta_{ii}^{pl}]\}$$
$$- V_i^p \sum_{l\in\mathcal{P}} [V_\ell^l (g_{i\ell,L}^{pl} \sin\delta_{i\ell}^{pl} - b_{i\ell,L}^{pl} \cos\delta_{i\ell}^{pl})] \tag{6.49}$$

As it will be clarified later, both static and recursive state estimators require the measurement model to be linear. Therefore, in the case of the non-linear measurement model, $\mathbf{H}$ is the Jacobian matrix of the measurement function $h(\mathbf{x})$:

$$\mathbf{H}(\mathbf{x}) = \frac{\partial h(\mathbf{x})}{\partial \mathbf{x}} \tag{6.50}$$

It is called *measurement Jacobian* and its sub-matrices for the considered available measurements are:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_V \\ \mathbf{H}_{PQ_{\text{inj}}} \\ \mathbf{H}_{PQ_{\text{flow}}} \end{bmatrix} \tag{6.51}$$

where $\mathbf{H}_V$ is the part of the Jacobian that is related to the partial derivatives of the phase and magnitude of the voltages as a function of the state, $\mathbf{H}_{PQ_{\text{inj}}}$ is related to the partial derivatives of the active and the reactive power injections as a function of the state, and $\mathbf{H}_{PQ_{\text{flow}}}$ is related to the partial derivatives of the active and the reactive power flows as a function of the state.

The first part $\mathbf{H}_V$ is given by:

$$\mathbf{H}_V = \begin{bmatrix} \boldsymbol{\Psi} & \boldsymbol{\Gamma} \\ \boldsymbol{\Lambda} & \boldsymbol{\Phi} \end{bmatrix} \tag{6.52}$$

where

$$\Psi_{\ell l,re}^{ip,re} = \begin{cases} 1, \text{ if } i = \ell \text{ and } p = l \\ 0, \text{ if } i \neq \ell \text{ or } p \neq l \end{cases} \tag{6.53}$$

$$\Gamma_{\ell l,im}^{ip,re} = 0 \tag{6.54}$$

$$\Lambda_{\ell l,re}^{ip,im} = 0 \tag{6.55}$$

$$\Phi_{\ell l,im}^{ip,im} = \begin{cases} 1, \text{ if } i = \ell \text{ and } p = l \\ 0, \text{ if } i \neq \ell \text{ or } p \neq l \end{cases} \tag{6.56}$$

In (6.53)–(6.56), the superscripts and subscripts have the same meaning as in (6.23)–(6.26).

The second part $\mathbf{H}_{PQ_{\text{inj}}}$ is equal to:

$$\mathbf{H}_{PQ_{\text{inj}}} = \begin{bmatrix} \frac{\partial P_i^p}{\partial \delta_\ell^l} & \frac{\partial P_i^p}{\partial V_\ell^l} \\ \frac{\partial Q_i^p}{\partial \delta_\ell^l} & \frac{\partial Q_i^p}{\partial V_\ell^l} \end{bmatrix} \tag{6.57}$$

The partial derivatives that correspond to the active power injections are:

$$\frac{\partial P_i^p}{\partial \delta_i^p} = -(V_i^p)^2 B_{ii}^{pp} + V_i^p \sum_{\ell=1}^{s} \sum_{l \in \mathcal{P}} V_\ell^l (-G_{i\ell}^{pl} \sin \delta_{i\ell}^{pl} + B_{i\ell}^{pl} \cos \delta_{i\ell}^{pl}) \tag{6.58}$$

$$\frac{\partial P_i^p}{\partial \delta_\ell^l} = V_i^p V_\ell^l (G_{i\ell}^{pl} \sin \delta_{i\ell}^{pl} - B_{i\ell}^{pl} \cos \delta_{i\ell}^{pl}) \tag{6.59}$$

$$\frac{\partial P_i^p}{\partial V_i^p} = V_i^p G_{ii}^{pp} + \sum_{\ell=1}^{s} \sum_{l \in \mathcal{P}} V_\ell^l (G_{i\ell}^{pl} \cos \delta_{i\ell}^{pl} + B_{i\ell}^{pl} \sin \delta_{i\ell}^{pl}) \tag{6.60}$$

$$\frac{\partial P_i^p}{\partial V_\ell^l} = V_i^p (G_{i\ell}^{pl} \cos \delta_{i\ell}^{pl} + B_{i\ell}^{pl} \sin \delta_{i\ell}^{pl}) \tag{6.61}$$

The partial derivatives that correspond to the reactive power injections are:

$$\frac{\partial Q_i^p}{\partial \delta_i^p} = -(V_i^p)^2 G_{ii}^{pp} + V_i^p \sum_{\ell=1}^{s} \sum_{l \in \mathcal{P}} V_\ell^l (G_{i\ell}^{pl} \cos \delta_{i\ell}^{pl} + B_{i\ell}^{pl} \sin \delta_{i\ell}^{pl}) \tag{6.62}$$

$$\frac{\partial Q_i^p}{\partial \delta_\ell^l} = V_i^p V_\ell^l (-G_{i\ell}^{pl} \cos \delta_{i\ell}^{pl} - B_{i\ell}^{pl} \sin \delta_{i\ell}^{pl}) \tag{6.63}$$

$$\frac{\partial Q_i^p}{\partial V_i^p} = -V_i^p B_{ii}^{pp} + \sum_{\ell=1}^{s} \sum_{l \in \mathcal{P}} V_\ell^l (G_{i\ell}^{pl} \sin \delta_{i\ell}^{pl} - B_{i\ell}^{pl} \cos \delta_{i\ell}^{pl}) \tag{6.64}$$

$$\frac{\partial Q_i^p}{\partial V_\ell^l} = V_i^p (G_{i\ell}^{pl} \sin \delta_{i\ell}^{pl} - B_{i\ell}^{pl} \cos \delta_{i\ell}^{pl}) \tag{6.65}$$

The third part of the Jacobian $\mathbf{H}_{PQ_{\text{flow}}}$ is equal to:

$$\mathbf{H}_{PQ_{\text{flow}}} = \begin{bmatrix} \dfrac{\partial P_{i\ell}^p}{\partial \delta_\ell^l} & \dfrac{\partial P_{i\ell}^p}{\partial V_\ell^l} \\ \dfrac{\partial Q_{i\ell}^p}{\partial \delta_\ell^l} & \dfrac{\partial Q_{i\ell}^p}{\partial V_\ell^l} \end{bmatrix} \tag{6.66}$$

The partial derivatives that correspond to the active power flows are:

$$\begin{aligned} \frac{\partial P_{i\ell}^p}{\partial \delta_i^p} &= -(V_i^p)^2 (b_{i\ell,L}^{pl} + b_{i\ell,T}^{pl}) \\ &\quad + V_i^p \sum_{l \in \mathcal{P}} \{V_i^l [-(g_{i\ell,L}^{pl} + g_{i\ell,T}^{pl}) \sin \delta_{ii}^{pl} + (b_{i\ell,L}^{pl} + b_{i\ell,T}^{pl}) \cos \delta_{ii}^{pl}]\} \\ &\quad + V_i^p \sum_{l \in \mathcal{P}} [V_\ell^l (g_{i\ell,L}^{pl} \sin \delta_{i\ell}^{pl} - b_{i\ell,L}^{pl} \cos \delta_{i\ell}^{pl})] \end{aligned} \tag{6.67}$$

$$\frac{\partial P_{i\ell}^p}{\partial \delta_\ell^l} = -V_i^p V_\ell^l (g_{i\ell,L}^{pl} \sin \delta_{i\ell}^{pl} - b_{i\ell,L}^{pl} \cos \delta_{i\ell}^{pl}) \tag{6.68}$$

$$\frac{\partial P_{i\ell}^{p}}{\partial V_{i}^{p}} = V_{i}^{p}(g_{i\ell,L}^{pl} + g_{i\ell,T}^{pl}) + \sum_{l\in\mathcal{P}}\{V_{i}^{l}[(g_{i\ell,L}^{pl} + g_{i\ell,T}^{pl})\cos\delta_{ii}^{pl} + (b_{i\ell,L}^{pl} + b_{i\ell,T}^{pl})\sin\delta_{ii}^{pl}]\}$$

$$- \sum_{l\in\mathcal{P}}[V_{\ell}^{l}(g_{i\ell,L}^{pl}\cos\delta_{i\ell}^{pl} + b_{i\ell,L}^{pl}\sin\delta_{i\ell}^{pl})] \tag{6.69}$$

$$\frac{\partial P_{i\ell}^{p}}{\partial V_{\ell}^{l}} = -V_{i}^{p}(g_{i\ell,L}^{pl}\cos\delta_{i\ell}^{pl} + b_{i\ell,L}^{pl}\sin\delta_{i\ell}^{pl}) \tag{6.70}$$

The partial derivatives that correspond to the reactive power flows are:

$$\frac{\partial Q_{i\ell}^{p}}{\partial \delta_{i}^{p}} = -(V_{i}^{p})^{2}(g_{i\ell,L}^{pl} + g_{i\ell,T}^{pl})$$

$$+ V_{i}^{p}\sum_{l\in\mathcal{P}}\{V_{i}^{l}[(g_{i\ell,L}^{pl} + g_{i\ell,T}^{pl})\cos\delta_{ii}^{pl} + (b_{i\ell,L}^{pl} + b_{i\ell,T}^{pl})\sin\delta_{ii}^{pl}]\}$$

$$- V_{i}^{p}\sum_{l\in\mathcal{P}}[V_{\ell}^{l}(g_{i\ell,L}^{pl}\cos\delta_{i\ell}^{pl} + b_{i\ell,L}^{pl}\sin\delta_{i\ell}^{pl})] \tag{6.71}$$

$$\frac{\partial Q_{i\ell}^{p}}{\partial \delta_{\ell}^{l}} = V_{i}^{p}V_{\ell}^{l}(g_{i\ell,L}^{pl}\cos\delta_{i\ell}^{pl} + b_{i\ell,L}^{pl}\sin\delta_{i\ell}^{pl}) \tag{6.72}$$

$$\frac{\partial Q_{i\ell}^{p}}{\partial V_{i}^{p}} = -V_{i}^{p}(b_{i\ell,L}^{pl} + b_{i\ell,T}^{pl})$$

$$+ \sum_{l\in\mathcal{P}}\{V_{i}^{l}[(g_{i\ell,L}^{pl} + g_{i\ell,T}^{pl})\sin\delta_{ii}^{pl} - (b_{i\ell,L}^{pl} + b_{i\ell,T}^{pl})\cos\delta_{ii}^{pl}]\}$$

$$- \sum_{l\in\mathcal{P}}[V_{\ell}^{l}(g_{i\ell,L}^{pl}\sin\delta_{i\ell}^{pl} - b_{i\ell,L}^{pl}\cos\delta_{i\ell}^{pl})] \tag{6.73}$$

$$\frac{\partial Q_{i\ell}^{p}}{\partial V_{\ell}^{l}} = -V_{i}^{p}(g_{i\ell,L}^{pl}\sin\delta_{i\ell}^{pl} - b_{i\ell,L}^{pl}\cos\delta_{i\ell}^{pl}) \tag{6.74}$$

6.1.2 Network observability

A power grid is fully observable if it is possible to calculate all the system state variables using a given set of measurements. For a given network model, the network observability is influenced by the type of available measurements (e.g., nodal voltages, current/power injections and flows) and their locations.

A necessary (but not sufficient) condition for the network observability is that the total number of measurements should be equal or larger than the total number of state variables, i.e., $m \geq n$. However, the criterion that needs to be always satisfied for the network to be fully observable is that matrix **H** must be of full rank.

At the design stage of the network measurement infrastructure, the type and location of the measurements are chosen in order to satisfy the network observability criteria. However, in case of data losses or topology changes, a new observability study must be conducted in real-time before the SE computation. If the network becomes unobservable, it can be split in multiple observable sub-networks called *observable islands*, and a separate SE is performed for every island. Additionally, both at the design stage and in real-time, the observability criteria can be met by adding pseudo-measurements.

There are several methods to perform the observability analysis, using graph theory or mathematical techniques, e.g., References 19, 22, 23. The minimum set of n measurements that can guarantee the full network observability is called *critical set* and if one measurement is removed from this set, the system state cannot be calculated. Note that the critical set of measurements is not unique since it can be composed of measurements of different type and location. Adding measurements to the critical set is very beneficial for the SE process since it results in higher estimation accuracy, improved robustness against data loss, and enhanced bad data identification capability.

6.1.3 Process model

As shown in Figure 6.1, recursive SE exploits the statistical properties of the system state by modelling its time evolution via a process model. In particular, the considered linear discrete-time process model can be formulated as [24]:

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_{t-1} + \mathbf{w}_{t-1} \tag{6.75}$$

where

- t is the time-step index;
- $\mathbf{x} \in \mathbb{R}^n$ represents the system state;
- $\mathbf{u} \in \mathbb{R}^{u_c}$ represents a set $\mathcal{U}_c$ ($u_c = |\mathcal{U}_c|$) of known controllable variables;
- $\mathbf{w} \in \mathbb{R}^n$ represents the process noise;
- $\mathbf{A}$ is an $n \times n$ matrix that links the system state $\mathbf{x}$ at time-step $t-1$ with the one at the time-step t, for the case of null controllable variables and null process noise;
- $\mathbf{B}$ is an $n \times u_c$ matrix that links the system state $\mathbf{x}$ at time-step t with the controllable variables $\mathbf{u}$ at time-step $t-1$, for the case of null process noise.

In general, also the matrices $\mathbf{A}$ and $\mathbf{B}$ might change at each time-step. The process noise $\mathbf{w}_{t-1}$ is assumed to be a Gaussian white sequence:

$$p(\mathbf{w}_{t-1}) \sim \mathcal{N}(0, \mathbf{Q}_{t-1}) \tag{6.76}$$

where $\mathbf{Q}_{t-1}$ is the *process noise covariance matrix*.

As it has been already clarified, power system SE is facilitated by the use of synchrophasor measurements streamed by PMUs at a high frame rate, e.g., 50 frames-per-second – fps. The advantage of this working condition is that the state exhibits small variations between two consecutive time-steps, so that a good approximation of matrix $\mathbf{A}$ can be the identity matrix $\mathbf{I}$. In addition, the power system inputs are typically not controllable, i.e. $\mathbf{B}$ is set equal to the null matrix $\mathbf{0}$. Therefore, a suitable

process model for the case of power system SE is the autoregressive integrated moving average – ARIMA (0,1,0) model given by:

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_{t-1} \tag{6.77}$$

The model of (6.77) can be derived from (6.75) by imposing $\mathbf{A} = \mathbf{I}$ and $\mathbf{B} = \mathbf{0}$ for the aforementioned reasons. The application of this process model to power system SE using KF was firstly proposed in 1970 by Debs and Larson [25] and then it was adopted by other authors, e.g., References 26, 27. An advantage of using this process model is that only $\mathbf{Q}$ has to be assessed. An heuristic method for the assessment of $\mathbf{Q}$ in the context of power system SE is proposed in Reference 28 and recalled in Section 6.3.4. The numerical validation of the correctness of this process model is provided in Section 6.7 by using the procedure presented in Reference 5.

Observation. In most of the cases, including power system SE, it is reasonable to assume that the process and measurement noises are uncorrelated:

$$\mathbb{E}[\mathbf{w}\boldsymbol{\varepsilon}^T] = \mathbf{0} \tag{6.78}$$

6.2 Static state estimation: the weighted least squares

The WLS problem is generally formulated as an unconstrained optimization problem. While the least squares requires the measurement noises to have the same variances, the WLS is used when measurements are characterized by different accuracies (*heteroscedasticity*). Indeed, the WLS method is able to weight the measurements according to their accuracies. The WLS relies on the following assumptions:

1. The measurement noises are Gaussian-distributed with null mean value.
2. The measurement noises are uncorrelated; therefore, the measurement noise covariance matrix $\mathbf{R}$ is diagonal.
3. The measurement matrix $\mathbf{H}$ that links the measurements with the state variables is of full rank, so that the network is observable.

Observation. Assumption 2 could be relaxed if significant mutual correlation among measurement errors is present. For such purpose, the problem becomes the so-called *generalized least squares*, where $\mathbf{R}$ is a full matrix [29, 30]. However, this case is unlikely to occur in power systems due to the following reasons [5]:

- Measurements provided by different meters can be reasonably considered independent [30], and it is assumed to use no 3-ph multi-function meters [29].
- The sensors are typically installed separately in each of the three phases and the cross-talk interferences are assumed to be negligible.
- The voltage and current magnitudes measured by the same PMU can be usually considered uncorrelated [30].
- As demonstrated in Reference 30, neglecting PMU correlations (both in magnitude and phase) in the estimator model does not lead to a significant decrease of the SE accuracy.

The WLS method can be derived using the maximum likelihood estimation concept as in Reference 31. Indeed, the goal is to compute the *most likely* system state given a set of measured quantities. The measurement noises are assumed to have the same and known Gaussian probability distribution with zero mean. The variances of the measurement noises compose the diagonal elements of $\mathbf{R}$. The *measurement residual* vector $\mathbf{r}$ is defined as:

$$\mathbf{r} = \mathbf{z} - h(\mathbf{x}) \tag{6.79}$$

where the measurement function $h(\mathbf{x})$ can be either linear or non-linear, as in (6.5) and (6.45), respectively. Then, the objective of the WLS optimization problem is to minimize the weighted sum of the squares of the measurement residuals:[6]

$$J(\mathbf{x}) = \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} \tag{6.80}$$

Equation (6.80), since $\mathbf{R}$ is diagonal (assumption 2 of the WLS), becomes:

$$J(\mathbf{x}) = \sum_{i=1}^{m} \frac{r_i^2}{R_{ii}} \tag{6.81}$$

It can be seen that the reciprocal of the measurement noise variances represents the weights assigned to each measurement, so that the higher the accuracy, the higher the weight.

At this point, the WLS algorithm has two different formulations depending on the types of measurements available. In particular, as it is explained in Sections 6.1.1.1 and 6.1.1.2, the choice of the types of measurements makes the measurement function linear or non-linear leading to the formulation of linear or non-linear SE, respectively.

6.2.1 Linear weighted least squares state estimator

In the case of linear weighted least squares (LWLS), the state is defined by using (6.1) and the exact linear measurement model is (6.5). The measurement residual vector is given by:

$$\mathbf{r}_t = \mathbf{z}_t - \mathbf{H}\mathbf{x}_t \tag{6.82}$$

where $\mathbf{H}$ is the exact matrix linking measurements and state variables, which contains no approximations and it is constant in time for a given network model. In this case, J is a quadratic function of the unknown state $\mathbf{x}$ with a minimum that can be computed analytically by imposing the derivative of J equal to zero evaluated in correspondence of the *estimated state* $\widehat{\mathbf{x}}$, as follows:

$$\left.\frac{\partial J(\mathbf{x}_t)}{\partial \mathbf{x}_t}\right|_{\widehat{\mathbf{x}}} = \left.\frac{\partial (\mathbf{r}_t^T \mathbf{R}_t^{-1} \mathbf{r}_t)}{\partial \mathbf{x}_t}\right|_{\widehat{\mathbf{x}}} = 0 \tag{6.83}$$

that yields:

$$\mathbf{H}^T \mathbf{R}_t^{-1} (\mathbf{z}_t - \mathbf{H}\widehat{\mathbf{x}}_t) = 0 \tag{6.84}$$

[6] See Reference 19 for the formal derivation of (6.80) and (6.81).

Solving for $\widehat{\mathbf{x}}_t$ yields:

$$\widehat{\mathbf{x}}_t = \mathbf{G}_t^{-1}\mathbf{H}^T\mathbf{R}_t^{-1}\mathbf{z}_t \tag{6.85}$$

where $\mathbf{G}$ is the so-called *Gain matrix* that is defined as:

$$\mathbf{G}_t = \mathbf{H}^T\mathbf{R}_t^{-1}\mathbf{H} \tag{6.86}$$

The covariance matrix of $\widehat{\mathbf{x}}_t$ is:

$$\text{cov}(\widehat{\mathbf{x}}_t) = \mathbf{G}_t^{-1} \tag{6.87}$$

Note that while $\mathbf{H}$ is constant in time for a given network model, $\mathbf{R}$ may change at each time-step for the reasons explained in Section 6.4.

As discussed in detail later in Section 6.5, the presence of erroneous measurements can be detected by analysing the vector of the *normalized measurement estimation residual* vector $\widehat{\mathbf{r}}_t^N$. The measurement estimation residual vector and its covariance matrix are:

$$\widehat{\mathbf{r}}_t = \mathbf{z}_t - \mathbf{H}\widehat{\mathbf{x}}_t \tag{6.88}$$

$$\text{cov}(\widehat{\mathbf{r}}_t) = \mathbf{C}_t = \mathbf{R}_t - \mathbf{H}\mathbf{G}_t^{-1}\mathbf{H}^T \tag{6.89}$$

then, the i^{th} element of $\widehat{\mathbf{r}}_t^N$ is computed as:

$$\widehat{r}_{t,i}^N = \frac{|\widehat{r}_{t,i}|}{\sqrt{C_{t,ii}}} \tag{6.90}$$

6.2.2 Non-linear weighted least squares

In the case of non-linear weighted least squares (NLWLS), the state is defined by (6.41) and the measurement model is non-linear and given by (6.45). The measurement residual vector is given by:

$$\mathbf{r}_t = \mathbf{z}_t - h(\mathbf{x}_t) \tag{6.91}$$

where $h(\mathbf{x})$ is the non-linear measurement function linking measurements and state variables. Thus, unlike the linear case, substituting (6.91) into (6.83) yields:

$$\mathbf{H}^T(\widehat{\mathbf{x}}_t)\mathbf{R}_t^{-1}[\mathbf{z}_t - h(\widehat{\mathbf{x}}_t)] = 0 \tag{6.92}$$

where $\mathbf{H}(\widehat{\mathbf{x}}_t)$ is the *measurement Jacobian* evaluated at $\widehat{\mathbf{x}}_t$:

$$\mathbf{H}(\widehat{\mathbf{x}}_t) = \left.\frac{\partial h(\mathbf{x}_t)}{\partial \mathbf{x}_t}\right|_{\widehat{\mathbf{x}}_t} \tag{6.93}$$

To solve (6.92), the non-linear function $h(\mathbf{x}_t)$ can be linearized around an initial vector $\mathbf{x}_{t,k}$ as:

$$h(\mathbf{x}_t) = h(\mathbf{x}_{t,k}) + \mathbf{H}(\mathbf{x}_{t,k})(\mathbf{x}_t - \mathbf{x}_{t,k}) \tag{6.94}$$

that substituted in (6.92) yields:

$$\mathbf{H}^T(\widehat{\mathbf{x}}_{t,k})\mathbf{R}_t^{-1}[\mathbf{z}_t - h(\widehat{\mathbf{x}}_{t,k})] - \mathbf{G}(\widehat{\mathbf{x}}_{t,k})(\widehat{\mathbf{x}}_{t,k+1} - \widehat{\mathbf{x}}_{t,k}) = 0 \tag{6.95}$$

where **G** is the Gain matrix already defined in (6.86). It can be seen that the linearization leads to an iterative algorithm where k is the iteration index. The matrices $\mathbf{H}(\widehat{\mathbf{x}}_{t,k})$ and $\mathbf{G}(\widehat{\mathbf{x}}_{t,k})$ can be abbreviated as $\mathbf{H}_{t,k}$ and $\mathbf{G}_{t,k}$, respectively. Rearranging (6.95) leads to the so-called *normal equation*:

$$\mathbf{G}_{t,k}(\widehat{\mathbf{x}}_{t,k+1} - \widehat{\mathbf{x}}_{t,k}) = \mathbf{H}_{t,k}^T \mathbf{R}_t^{-1}[\mathbf{z}_t - h(\widehat{\mathbf{x}}_{t,k})] \tag{6.96}$$

Note that for the case of NLWLS, **H** is a function of the network state; consequently, it is not constant in time and needs to be recomputed at every time-step. Equation (6.96) can be solved with various techniques (e.g., forward/backward substitution as proposed in Reference 19) or by directly inverting the Gain matrix as:

$$\widehat{\mathbf{x}}_{t,k+1} = \widehat{\mathbf{x}}_{t,k} + \mathbf{G}_{t,k}^{-1} \mathbf{H}_{t,k}^T \mathbf{R}_t^{-1}[\mathbf{z}_t - h(\widehat{\mathbf{x}}_{t,k})] \tag{6.97}$$

Equations (6.96) and (6.97) are iterated until convergence. Some possible stopping criteria are:

1. $\max|\widehat{\mathbf{x}}_{t,k+1} - \widehat{\mathbf{x}}_{t,k}| \leq \epsilon_1$
2. $|J(\widehat{\mathbf{x}}_{t,k+1}) - J(\widehat{\mathbf{x}}_{t,k})| < \epsilon_2$
3. $J(\widehat{\mathbf{x}}_{t,k+1}) < \epsilon_3$

where ϵ_1, ϵ_2 and ϵ_3 are a priori selected thresholds.

6.3 Recursive state estimation: the Kalman filter

In 1960, R.E. Kalman published his most famous paper, i.e., Reference 32, describing a recursive solution to the discrete data linear filtering problem. Since then, the KF has been used in a large number of different fields including power systems SE.

The KF relies on the following assumptions:

1. The process and measurement noises are Gaussian white sequences.
2. The process and measurement noises are uncorrelated, as indicated by (6.78).
3. The measurement matrix **H** that links the measurements with the state variables is of full rank, so that the network is observable.

There are several versions of KF. The Discrete Kalman Filter (DKF) is used for linear formulations of the SE problem, whereas the Extended Kalman Filter (EKF) is used when the process and/or the measurement model are non-linear. These two versions of KF are described, respectively, in Sections 6.3.1 and 6.3.2 with reference to the ARIMA (0,1,0) process model of (6.77).

6.3.1 Discrete Kalman filter

As known (e.g., Reference 24), DKF consists of two different parts, the so-called *time-update* (*prediction*) and the *measurement-update* (*estimation*). In what follows, the derivation of the KF equations is presented. The goal is to find an unbiased and minimum variance estimator.

Recursivity: The first assumption is that at time $t-1$ there is already an estimate $\widehat{\mathbf{x}}_{t-1}$ which includes all information up to – and including – time $t-1$. At the next time-step t, this estimate is used to compute the predicted state $\widetilde{\mathbf{x}}_t$. The goal is to find an estimate $\widehat{\mathbf{x}}_t$ for the state at time t that incorporates the new set of measurements $\mathbf{z}_t$ and the predicted state $\widetilde{\mathbf{x}}_t$, e.g., Reference 33:

$$\widehat{\mathbf{x}}_t = \mathbf{K}\prime_t\widetilde{\mathbf{x}}_t + \mathbf{K}_t\mathbf{z}_t \tag{6.98}$$

where $\mathbf{K}\prime_t$ and $\mathbf{K}_t$ are two weighting matrices whose values are determined as reported below.

The prediction and estimation errors are, respectively, defined as:

$$\widetilde{\mathbf{e}}_t \triangleq \mathbf{x}_t - \widetilde{\mathbf{x}}_t \tag{6.99}$$

$$\widehat{\mathbf{e}}_t \triangleq \mathbf{x}_t - \widehat{\mathbf{x}}_t \tag{6.100}$$

and the related prediction and estimation error covariance matrices are, respectively:

$$\widetilde{\mathbf{P}}_t = \mathbb{E}[\widetilde{\mathbf{e}}_t\widetilde{\mathbf{e}}_t^T] \tag{6.101}$$

$$\widehat{\mathbf{P}}_t = \mathbb{E}[\widehat{\mathbf{e}}_t\widehat{\mathbf{e}}_t^T] \tag{6.102}$$

where $\mathbb{E}$ is the expected value operator.

By subtracting $\mathbf{x}_t$ in both parts, (6.98) becomes:

$$\widehat{\mathbf{x}}_t - \mathbf{x}_t = \mathbf{K}\prime_t\widetilde{\mathbf{x}}_t + \mathbf{K}_t\mathbf{z}_t - \mathbf{x}_t \tag{6.103}$$

Then, by substituting (6.5) in (6.103) for the current time-step t, (6.103) becomes:

$$\widehat{\mathbf{x}}_t - \mathbf{x}_t = \mathbf{K}\prime_t\widetilde{\mathbf{x}}_t + \mathbf{K}_t(\mathbf{H}\mathbf{x}_t + \boldsymbol{\varepsilon}_t) - \mathbf{x}_t = \mathbf{K}\prime_t\widetilde{\mathbf{x}}_t + (\mathbf{K}_t\mathbf{H} - \mathbf{I})\mathbf{x}_t + \mathbf{K}_t\boldsymbol{\varepsilon}_t \tag{6.104}$$

By adding and subtracting $\mathbf{K}\prime_t\mathbf{x}_t$ in the right-hand side of (6.104), the latter becomes:

$$\begin{aligned}\widehat{\mathbf{x}}_t - \mathbf{x}_t &= \mathbf{K}\prime_t\widetilde{\mathbf{x}}_t + (\mathbf{K}_t\mathbf{H} - \mathbf{I})\mathbf{x}_t + \mathbf{K}_t\boldsymbol{\varepsilon}_t + \mathbf{K}\prime_t\mathbf{x}_t - \mathbf{K}\prime_t\mathbf{x}_t \\ &= \mathbf{K}\prime_t(\widetilde{\mathbf{x}}_t - \mathbf{x}_t) + (\mathbf{K}\prime_t + \mathbf{K}_t\mathbf{H} - \mathbf{I})\mathbf{x}_t + \mathbf{K}_t\boldsymbol{\varepsilon}_t\end{aligned} \tag{6.105}$$

In order for the estimation to be *unbiased* the expected value of the estimation error must be zero:

$$\mathbb{E}[\widehat{\mathbf{x}}_t - \mathbf{x}_t] = \mathbf{K}\prime_t\mathbb{E}[\widetilde{\mathbf{x}}_t - \mathbf{x}_t] - (\mathbf{I} - \mathbf{K}_t\mathbf{H} - \mathbf{K}\prime_t)\mathbb{E}[\mathbf{x}_t] + \mathbf{K}_t\mathbb{E}[\boldsymbol{\varepsilon}_t] = \mathbf{0} \tag{6.106}$$

or

$$\mathbb{E}[\widehat{\mathbf{e}}_t] = \mathbf{K}\prime_t\mathbb{E}[\widetilde{\mathbf{e}}_t] + (\mathbf{I} - \mathbf{K}_t\mathbf{H} - \mathbf{K}\prime_t)\mathbb{E}[\mathbf{x}_t] - \mathbf{K}_t\mathbb{E}[\boldsymbol{\varepsilon}_t] = \mathbf{0} \tag{6.107}$$

Since $\mathbb{E}[\boldsymbol{\varepsilon}_t] = \mathbf{0}$ (assumption 1 of the KF) and $\mathbb{E}[\widetilde{\mathbf{e}}_t] = \mathbf{0}$, the estimator is unbiased only if:

$$\mathbf{K}\prime_t = \mathbf{I} - \mathbf{K}_t\mathbf{H} \tag{6.108}$$

By substituting (6.108) in (6.98), the KF estimation equation is obtained:

$$\widehat{\mathbf{x}}_t = \widetilde{\mathbf{x}}_t + \mathbf{K}_t(\mathbf{z}_t - \mathbf{H}\widetilde{\mathbf{x}}_t) \tag{6.109}$$

where $\mathbf{K}_t$ is the so-called *Kalman Gain*.[7] The term $\boldsymbol{\gamma}_t = \mathbf{z}_t - \mathbf{H}\widetilde{\mathbf{x}}_t$ in (6.109) is known as *innovation*, which is a Gaussian white sequence with covariance matrix:

$$\mathbf{S}_t = \mathbf{R}_t + \mathbf{H}\widetilde{\mathbf{P}}_t\mathbf{H}^T \tag{6.110}$$

Some methods for erroneous measurement detection are based on the analysis of the normalized innovation, as explained in Section 6.5. The i^{th} element of the normalized innovation vector is computed as:

$$\gamma_{t,i}^N = \frac{|\gamma_{t,i}|}{\sqrt{S_{t,ii}}} \tag{6.111}$$

Note that $\mathbf{H}\widetilde{\mathbf{x}}_t$ is essentially the vector of the predicted measurements, so that (6.109) can be interpreted as the sum of the prediction, plus the weighted difference between actual and predicted measurements.

Calculation of the estimation error covariance matrix: by definition, (6.102) can be also written as:

$$\widehat{\mathbf{P}}_t = \text{cov}(\widehat{\mathbf{e}}_t) = \text{cov}(\mathbf{x}_t - \widehat{\mathbf{x}}_t) \tag{6.112}$$

By using (6.109), (6.112) becomes:

$$\begin{aligned}\widehat{\mathbf{P}}_t &= \text{cov}(\mathbf{x}_t - (\widetilde{\mathbf{x}}_t + \mathbf{K}_t(\mathbf{z}_t - \mathbf{H}\widetilde{\mathbf{x}}_t))) \\ &= \text{cov}(\mathbf{x}_t - (\widetilde{\mathbf{x}}_t + \mathbf{K}_t(\mathbf{H}\mathbf{x}_t + \boldsymbol{\varepsilon}_t - \mathbf{H}\widetilde{\mathbf{x}}_t))) \\ &= \text{cov}((\mathbf{I} - \mathbf{K}_t\mathbf{H})(\mathbf{x}_t - \widetilde{\mathbf{x}}_t) - \mathbf{K}_t\boldsymbol{\varepsilon}_t)\end{aligned} \tag{6.113}$$

Since $\boldsymbol{\varepsilon}_t$ is not correlated with the other terms, (6.113) becomes:

$$\widehat{\mathbf{P}}_t = \text{cov}((\mathbf{I} - \mathbf{K}_t\mathbf{H})(\mathbf{x}_t - \widetilde{\mathbf{x}}_t)) + \text{cov}(\mathbf{K}_t\boldsymbol{\varepsilon}_t) \tag{6.114}$$

and thus:

$$\widehat{\mathbf{P}}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{H})\text{cov}(\mathbf{x}_t - \widetilde{\mathbf{x}}_t)(\mathbf{I} - \mathbf{K}_t\mathbf{H})^T + \mathbf{K}_t\text{cov}(\boldsymbol{\varepsilon}_t)\mathbf{K}_t^T \tag{6.115}$$

By definition, $\text{cov}(\boldsymbol{\varepsilon}_t) = \mathbf{R}_t$ and $\text{cov}(\mathbf{x}_t - \widetilde{\mathbf{x}}_t) = \widetilde{\mathbf{P}}_t$, therefore:

$$\widehat{\mathbf{P}}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{H})\widetilde{\mathbf{P}}_t(\mathbf{I} - \mathbf{K}_t\mathbf{H})^T + \mathbf{K}_t\mathbf{R}_t\mathbf{K}_t^T \tag{6.116}$$

The formula in (6.116) is also known as the *Joseph form* of the covariance update equation and is valid for any value of the Kalman Gain $\mathbf{K}_t$. However, if $\mathbf{K}_t$ is the optimal gain, (6.116) can be further simplified.

Optimal Kalman Gain derivation: KF is a *minimum variance* estimator. The goal is to minimize the expected value of the square of the magnitude of the estimation

[7]Note that the Kalman Gain $\mathbf{K}_t$ needs to be determined since, at this stage, it is still unknown.

error $\mathbb{E}[\|\mathbf{x}_t - \widehat{\mathbf{x}}_t\|^2]$. This is equivalent to minimizing the *trace* of $\widehat{\mathbf{P}}_t$. By expanding out the terms in (6.116):

$$\widehat{\mathbf{P}}_t = \widetilde{\mathbf{P}}_t - \mathbf{K}_t\mathbf{H}\widetilde{\mathbf{P}}_t - \widetilde{\mathbf{P}}_t\mathbf{H}^T\mathbf{K}_t^T + \mathbf{K}_t\mathbf{S}_t\mathbf{K}_t^T \tag{6.117}$$

The trace of $\widehat{\mathbf{P}}_t$ is minimized when its matrix derivative with respect to the Kalman Gain matrix is zero:

$$\frac{\partial \mathrm{tr}(\widehat{\mathbf{P}}_t)}{\partial \mathbf{K}_t} = 0 \tag{6.118}$$

Using the following identities:

$$\frac{\partial \mathrm{tr}(\mathbf{AB})}{\partial \mathbf{A}} = \mathbf{B}^T \tag{6.119}$$

$$\frac{\partial \mathrm{tr}(\mathbf{BA}^T)}{\partial \mathbf{A}} = \mathbf{B} \tag{6.120}$$

$$\frac{\partial \mathrm{tr}(\mathbf{ABA}^T)}{\partial \mathbf{A}} = 2\mathbf{AB}, \qquad \text{if } \mathbf{B} \text{ is a symmetric matrix} \tag{6.121}$$

and the fact that $\widetilde{\mathbf{P}}_t$ is a symmetric matrix, (6.118) becomes:

$$\frac{\partial \mathrm{tr}(\widehat{\mathbf{P}}_t)}{\partial \mathbf{K}_t} = -2\widetilde{\mathbf{P}}_t\mathbf{H}^T + 2\mathbf{K}_t\mathbf{S}_t = 0 \tag{6.122}$$

Solving (6.122) for $\mathbf{K}_t$ yields the optimal Kalman Gain:

$$\mathbf{K}_t = \widetilde{\mathbf{P}}_t\mathbf{H}^T\mathbf{S}_t^{-1} = \widetilde{\mathbf{P}}_t\mathbf{H}^T(\mathbf{H}\widetilde{\mathbf{P}}_t\mathbf{H}^T + \mathbf{R}_t)^{-1} \tag{6.123}$$

Hence, by multiplying both sides of (6.123) on the right by $\mathbf{S}_t\mathbf{K}_t^T$:

$$\mathbf{K}_t\mathbf{S}_t\mathbf{K}_t^T = \widetilde{\mathbf{P}}_t\mathbf{H}^T\mathbf{K}_t^T \tag{6.124}$$

Going back to (6.117), the last two terms are cancelled out, therefore:

$$\widehat{\mathbf{P}}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{H})\widetilde{\mathbf{P}}_t \tag{6.125}$$

Calculation of the prediction error covariance matrix: the ARIMA (0,1,0) process model of (6.77) is recalled here below:

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_{t-1} \tag{6.126}$$

Thus, the KF state prediction equation is:

$$\widetilde{\mathbf{x}}_t = \widetilde{\mathbf{x}}_{t-1} \tag{6.127}$$

Adding the true state $\mathbf{x}_t$ at both sides of (6.127) and using (6.126) yields:

$$\mathbf{x}_t - \widetilde{\mathbf{x}}_t = \mathbf{x}_{t-1} - \widehat{\mathbf{x}}_{t-1} + \mathbf{w}_{t-1} \tag{6.128}$$

and it follows that:

$$\widetilde{\mathbf{P}}_t = \text{cov}(\widehat{\mathbf{e}}_{t-1} + \mathbf{w}_{t-1}) \tag{6.129}$$

Since $\mathbf{w}_{t-1}$ is the process noise between time $t-1$ and t, whereas $\widehat{\mathbf{e}}_{t-1}$ is the estimation error up to time $t-1$, they are uncorrelated; therefore:

$$\widetilde{\mathbf{P}}_t = \widehat{\mathbf{P}}_t + \mathbf{Q}_{t-1} \tag{6.130}$$

To sum up, the formulation of the DKF-SE for the optimal Kalman Gain is the following:

1. Time-update/prediction:

$$\widetilde{\mathbf{x}}_t = \widehat{\mathbf{x}}_{t-1} \tag{6.131}$$

$$\widetilde{\mathbf{P}}_t = \widehat{\mathbf{P}}_{t-1} + \mathbf{Q}_{t-1} \tag{6.132}$$

2. Measurement-update/estimation:

$$\mathbf{K}_t = \widetilde{\mathbf{P}}_t\mathbf{H}^T(\mathbf{H}\widetilde{\mathbf{P}}_t\mathbf{H}^T + \mathbf{R}_t)^{-1} \tag{6.133}$$

$$\widehat{\mathbf{x}}_t = \widetilde{\mathbf{x}}_t + \mathbf{K}_t(\mathbf{z}_t - \mathbf{H}\widetilde{\mathbf{x}}_t) \tag{6.134}$$

$$\widehat{\mathbf{P}}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{H})\widetilde{\mathbf{P}}_t \tag{6.135}$$

6.3.2 Extended Kalman filter

The EKF is used when the process to be estimated and/or the relationship between measurements and state variables are non-linear, e.g., References 34, 35. In this case, the equations that describe the EKF are similar to the ones of the DKF except for the following differences:

- Similarly to (6.45), the equation that relates measurements and state variables is:

$$\mathbf{z}_t = h(\mathbf{x}_t) + \boldsymbol{\varepsilon}_t \tag{6.136}$$

- Matrix $\mathbf{H}$ is the measurement Jacobian defined by (6.50). It is not constant in time for a given network model, but it changes at every time-step as a function of $\mathbf{x}$.

The equations that describe the EKF-SE are:

1. Time-update/prediction:

$$\widetilde{\mathbf{x}}_t = \widehat{\mathbf{x}}_{t-1} \tag{6.137}$$

$$\widetilde{\mathbf{P}}_t = \mathbf{P}_{t-1} + \mathbf{Q}_{t-1} \tag{6.138}$$

2. Measurement-update/estimation:

$$\mathbf{K}_t = \widetilde{\mathbf{P}}_t \mathbf{H}_t^T (\mathbf{H}_t \widetilde{\mathbf{P}}_t \mathbf{H}_t^T + \mathbf{R}_t)^{-1} \tag{6.139}$$

$$\widehat{\mathbf{x}}_t = \widetilde{\mathbf{x}}_t + \mathbf{K}_t (\mathbf{z}_t - h(\widetilde{\mathbf{x}}_t)) \tag{6.140}$$

$$\widehat{\mathbf{P}}_t = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \widetilde{\mathbf{P}}_t \tag{6.141}$$

6.3.3 Kalman Filter sensitivity with respect to the measurement and process noise covariance matrices

As it has been already explained above, the KF equations involve two covariance matrices, the measurement noise covariance matrix $\mathbf{R}$ and the process noise covariance matrix $\mathbf{Q}$. The values of these two matrices influence significantly the performance of the KF. If they are not properly assessed, the quality of the estimated state given by the KF-SE is not guaranteed. Therefore, it is important to perform a sensitivity analysis of the KF with respect to $\mathbf{R}$ and $\mathbf{Q}$ (e.g., Reference 36).

The measurement noise covariance matrix represents the accuracies of the measurement devices and it weights how much the KF trusts the measurements. Its value can be easily inferred provided the knowledge of the accuracy of the measurement infrastructure (sensors plus PMUs and/or conventional metering devices) and pseudo-measurements (i.e., historical data and/or zero-injection buses). The process noise covariance matrix represents the uncertainties introduced by the process model to predict the system state. It is worth pointing out that in the literature dealing with power systems SE using KF, the values of $\mathbf{Q}$ are, usually, arbitrarily selected although they should be assessed in order to improve the estimation accuracy (e.g., References 25, 37, 38).

The influence of $\mathbf{R}_t$ and $\mathbf{Q}_t$ on the Kalman gain $\mathbf{K}_t$ can be explained by looking at (6.139). As $\mathbf{R}_t$ decreases, i.e., increasing the confidence on the measurement model, $\mathbf{K}_t$ increases. As a consequence, this leads to an increase of the contribution of the innovation $\boldsymbol{\gamma}_t = \mathbf{z}_t - \mathbf{H}\widetilde{\mathbf{x}}_t$ in the estimation equation (6.134), so that the KF estimates approach the WLS ones. In particular, as $\mathbf{R}_t \to 0$, also $\widehat{\mathbf{P}}_t \to 0$, so that $\widetilde{\mathbf{P}}_t \to \mathbf{Q}_{t-1}$. Therefore, the limit of $\mathbf{K}_t$ as $\mathbf{R}_t \to 0$ is:

$$\lim_{\mathbf{R}_t \to 0} \mathbf{K}_t = \mathbf{Q}_{t-1} \mathbf{H}^T (\mathbf{H} \mathbf{Q}_{t-1} \mathbf{H}^T)^{-1} \tag{6.142}$$

On the other hand, as $\mathbf{Q}_{t-1}$ decreases, i.e., the process model is trusted more, $\mathbf{K}_t$ decreases leading to an increasing weight of the predicted state with respect to the innovation. As $\mathbf{Q}_{t-1} \to 0$, both $\widehat{\mathbf{P}}_t$ and $\widetilde{\mathbf{P}}_t$ approach zero, so that also $\mathbf{K}_t$ tends to zero:

$$\lim_{\mathbf{Q}_{t-1} \to 0} \mathbf{K}_t = \mathbf{0} \tag{6.143}$$

Based on the above analysis, it becomes evident that a proper assessment of the matrix **Q** is of fundamental importance to maximize the performance of the KF-SE. The following section discusses this aspect.

6.3.4 Assessment of the process noise covariance matrix

The process noise covariance matrix can be assessed by using the heuristic method firstly proposed in Reference 28 and then further investigated in Reference 5. This method is effective in terms of estimation accuracy and suitable for real-time applications, although it needs a pre-tuning stage due to its heuristic nature. It assesses the diagonal elements of **Q** by using a moving window composed of the previous N estimated states. The method requires high-resolution state estimates to be effective. In this respect, reference is made to SEs based on PMU measurements (exclusively or through a mixed measurement set), characterized by high refresh rates in the order of tens of fps. The application of this method to the case of power system SE is given in Section 6.7.

In what follows, the formulation of the aforementioned method is briefly recalled. At time-step t, the estimation of $\mathbf{Q}_{t-1}$ is performed by using the last N estimated states. The procedure is the following:

1. Compute the c^{th} element of the vector $\mathbf{g} \in \mathbb{R}^n$ as the sample variance of the c^{th} element of the last N estimated states:

$$g_c = \text{var}[\widehat{x}_{t-1,c}, \ldots, \widehat{x}_{t-N,c}] \tag{6.144}$$

2. Then, the diagonal of $\mathbf{Q}_{t-1}$ is composed of the elements of $\mathbf{g}$ calculated in step 1:

$$\mathbf{Q}_{t-1} = \text{diag}(\mathbf{g}) \tag{6.145}$$

In general, the diagonal elements of **Q** are not all equal to each other.

6.4 Assessment of the measurement noise covariance matrix

The measurement noise covariance matrix **R** is a matrix that may change from one time-step to another. It can be diagonal or full, but it has to be *positive definite*. The elements of **R** include the accuracy of both the measurement electrical sensors and the PMUs/RTUs. The reasons making the measurement noise covariance matrix **R** changing at every time-step are:

- The measurement errors are calculated with respect to the measured values.
- When the state is expressed in rectangular coordinates, e.g., linear SE, whereas the measurement errors are given in polar coordinates, a transformation of coordinates (projection from polar to rectangular coordinates) is performed at each time-step. Another factor that generates time-dependent accuracies is that the

true system frequency is different from the rated one, since phasors rotate.[8] It is important to mention that, in general, the transformed measurement errors in rectangular coordinates are not Gaussian-distributed, unless the standard deviations of the original magnitude and phase errors are small enough.

Here below are described in detail the steps that lead to the derivation of **R**. In Section 6.7, it is given a numerical example of how matrix **R** is derived.

1. The first step is the calculation of the *maximum errors* that consist in the cumulative errors of the sensors and the PMUs/RTUs. The maximum errors are expressed in percentage for the magnitude and in radians for the phase.
2. The second step consists in the expression of the maximum magnitude errors with respect to the measured value.[9]
3. The third step includes the calculation of the cumulative standard deviations, which are equal to one-third of the maximum errors, assuming that the 99.73% of the values lie within three standard deviations from the mean.
4. This step, in case the SE is linear, includes the transformation of the measurement errors and the associated standard deviations from polar coordinates (magnitudes and phases) to rectangular coordinates (real and imaginary parts).

In what follows, the polar-to-rectangular projection of the measurement uncertainties of phasors is described. Indeed, phasor uncertainties are usually expressed in polar coordinates and their transformation to rectangular coordinates requires the knowledge of the true value of the measured quantity that is, obviously, a hidden value. For this reason, it is interesting to derive the analytical relationship that allows to express the phasors uncertainties in rectangular coordinates as a function of the measurements. The process here illustrated is based on the work presented in Reference 39. In particular, compared to Reference 39, it is illustrated a more detailed derivation of the variances of the projected errors as well as a more detailed description of the mathematical justification of the individual steps of the process.

The measured voltage magnitude and phase are denoted with V_z and δ_z.[10] Then:

$$V_z = V_x + \widetilde{V} \tag{6.146}$$

$$\delta_z = \delta_x + \widetilde{\delta} \tag{6.147}$$

where V_x and δ_x are the true voltage magnitude and phase, respectively, whereas $\widetilde{V}$ and $\widetilde{\delta}$ are the voltage magnitude and phase measurement errors. The latter are assumed to be independent white Gaussian sequences with standard deviations equal to σ_V and σ_δ, respectively.

[8]Note that even if all the phasor phases are referred to the phase of one of the phasors, i.e., they do not rotate, their phases change with time, making the projection from polar to rectangular coordinates time dependent.

[9]For instance, if the maximum magnitude error is equal to 5% and the measurement is 1.1 pu, then the maximum error is equal to $0.05 \times 1.1 = 0.055$ pu.

[10]The same process applies to the case of the current measurements.

The measurements in polar coordinates given in (6.146) and (6.147) are transformed to measurements in rectangular coordinates:

$$V_{re,z} = V_z \cos \delta_z \tag{6.148}$$

$$V_{im,z} = V_z \sin \delta_z \tag{6.149}$$

where $V_{re,z}$ and $V_{im,z}$ are the measured voltage real and imaginary parts, respectively.

The measured real and imaginary voltages ($V_{re,z}$ and $V_{im,z}$) can be expressed as a function of the corresponding true quantities ($V_{re,x}$ and $V_{im,x}$) and errors ($\widetilde{V}_{re}$ and $\widetilde{V}_{im}$) in the following way:

$$V_{re,z} = V_{re,x} + \widetilde{V}_{re} = (V_x + \widetilde{V}) \cos(\delta_x + \widetilde{\delta}) \tag{6.150}$$

$$V_{im,z} = V_{im,x} + \widetilde{V}_{im} = (V_x + \widetilde{V}) \sin(\delta_x + \widetilde{\delta}) \tag{6.151}$$

By making use of the trigonometric identities:

$$\cos(a + b) = \cos a \cos b - \sin a \sin b \tag{6.152}$$

$$\sin(a + b) = \sin a \cos b + \cos a \sin b \tag{6.153}$$

equations (6.150) and (6.151) become:

$$\begin{aligned} V_{re,z} &= (V_x + \widetilde{V})(\cos \delta_x \cos \widetilde{\delta} - \sin \delta_x \sin \widetilde{\delta}) \\ &= V_x \cos \delta_x \cos \widetilde{\delta} + \widetilde{V} \cos \delta_x \cos \widetilde{\delta} - V_x \sin \delta_x \sin \widetilde{\delta} - \widetilde{V} \sin \delta_x \sin \widetilde{\delta} \end{aligned} \tag{6.154}$$

and

$$\begin{aligned} V_{im,z} &= (V_x + \widetilde{V})(\sin \delta_x \cos \widetilde{\delta} + \cos \delta_x \sin \widetilde{\delta}) \\ &= V_x \sin \delta_x \cos \widetilde{\delta} + \widetilde{V} \sin \delta_x \cos \widetilde{\delta} + V_x \cos \delta_x \sin \widetilde{\delta} + \widetilde{V} \cos \delta_x \sin \widetilde{\delta} \end{aligned} \tag{6.155}$$

The measurement error of the real part of the voltage, using (6.150) and (6.154) is equal to:

$$\begin{aligned} \widetilde{V}_{re} &= V_{re,z} - V_{re,x} \\ &= V_x \cos \delta_x \cos \widetilde{\delta} + \widetilde{V} \cos \delta_x \cos \widetilde{\delta} - V_x \sin \delta_x \sin \widetilde{\delta} - \widetilde{V} \sin \delta_x \sin \widetilde{\delta} - V_x \cos \delta_x \\ &= V_x \cos \delta_x (\cos \widetilde{\delta} - 1) + \widetilde{V} \cos \delta_x \cos \widetilde{\delta} - V_x \sin \delta_x \sin \widetilde{\delta} - \widetilde{V} \sin \delta_x \sin \widetilde{\delta} \end{aligned} \tag{6.156}$$

and the measurement error of the imaginary part of the voltage, using (6.151) and (6.155) is equal to:

$$\begin{aligned}\widetilde{V}_{im} &= V_{im,z} - V_{im,x} \\ &= V_x \sin\delta_x \cos\widetilde{\delta} + \widetilde{V} \sin\delta_x \cos\widetilde{\delta} + V_x \cos\delta_x \sin\widetilde{\delta} + \widetilde{V} \cos\delta_x \sin\widetilde{\delta} - V_x \sin\delta_x \\ &= V_x \sin\delta_x(\cos\widetilde{\delta} - 1) + \widetilde{V} \sin\delta_x \cos\widetilde{\delta} + V_x \cos\delta_x \sin\widetilde{\delta} + \widetilde{V} \cos\delta_x \sin\widetilde{\delta}\end{aligned} \tag{6.157}$$

As it can be observed from (6.156) and (6.157):

- The errors in rectangular coordinates depend on both the true quantities (magnitude and phase) as well as the errors in polar coordinates.
- The link between the errors in rectangular coordinates and the above-mentioned quantities is non-linear through trigonometric functions.

The mean value of the error of the real part of the voltage, starting from (6.156), is calculated as:

$$\begin{aligned}\mu_{\widetilde{V}re,x} &= \mathbb{E}[\widetilde{V}_{re}] = \mathbb{E}[V_x \cos\delta_x \cos\widetilde{\delta}] - \mathbb{E}[V_x \cos\delta_x] \\ &\quad + \mathbb{E}[\widetilde{V} \cos\delta_x \cos\widetilde{\delta}] - \mathbb{E}[V_x \sin\delta_x \sin\widetilde{\delta}] - \mathbb{E}[\widetilde{V} \sin\delta_x \sin\widetilde{\delta}]\end{aligned} \tag{6.158}$$

Since the errors in polar coordinates are assumed to be *zero-mean Gaussian and independent*:

$$\begin{aligned}&\mathbb{E}[\widetilde{V} \cos\delta_x \cos\widetilde{\delta}] = \cos\delta_x \mathbb{E}[\widetilde{V}]\mathbb{E}[\cos\widetilde{\delta}] = 0, \text{ since } \mathbb{E}[\widetilde{V}] = 0 \text{ and similarly} \\ &\mathbb{E}[\widetilde{V} \sin\delta_x \sin\widetilde{\delta}] = 0\end{aligned} \tag{6.159}$$

Then (6.158) becomes:

$$\mu_{\widetilde{V}re,x} = V_x \cos\delta_x \mathbb{E}[\cos\widetilde{\delta}] - V_x \cos\delta_x - V_x \sin\delta_x \mathbb{E}[\sin\widetilde{\delta}] \tag{6.160}$$

To calculate the value of $\mathbb{E}[\cos\widetilde{\delta}]$ and $\mathbb{E}[\sin\widetilde{\delta}]$, where $\widetilde{\delta} \sim \mathcal{N}(0, \sigma_\delta^2)$, Euler's formula is used:

$$e^{j\delta} = \cos\delta + j\sin\delta \tag{6.161}$$

Then, the *characteristic function* $\Phi_X(t)$ for a Gaussian distributed variable $X \sim \mathcal{N}(\mu, \sigma^2)$ can be defined as:

$$\Phi_X(t) = \mathbb{E}[e^{jtX}] = e^{jt\mu - \frac{1}{2}\sigma^2 t^2} \tag{6.162}$$

where $t \in \mathbb{R}$ is the argument of the characteristic function. For $t = 1$ and since $\mu_\delta = 0$, (6.162) becomes:

$$\mathbb{E}[e^{j\widetilde{\delta}}] = e^{0-\frac{1}{2}\sigma_\delta^2} = e^{-\frac{1}{2}\sigma_\delta^2} \tag{6.163}$$

Going back to (6.161), by taking the expectations in both parts of the equation and by using (6.163), the imaginary part $\mathbb{E}[\sin(\widetilde{\delta})]$ is equal to zero and:

$$\mathbb{E}[\cos\widetilde{\delta}] = \mathbb{E}[e^{j\widetilde{\delta}}] = e^{-\frac{1}{2}\sigma_\delta^2} \tag{6.164}$$

Hence, (6.160) becomes:

$$\mu_{\widetilde{V}_{re,x}} = V_x \cos\delta_x\left(e^{-\frac{1}{2}\sigma_\delta^2} - 1\right) \tag{6.165}$$

Similarly, the mean error of the imaginary part of the voltage is equal to:

$$\mu_{\widetilde{V}_{im,x}} = V_x \sin\delta_x\left(e^{-\frac{1}{2}\sigma_\delta^2} - 1\right) \tag{6.166}$$

The bias becomes significant only for large values of magnitude and/or large phase errors.

The measurement variance of the real part of the voltage, without doing any approximations, is calculated as:

$$\mathrm{Var}_{\widetilde{V}_{re,x}} = \mathrm{Var}[\widetilde{V}_{re}] \tag{6.167}$$

According to the properties of the variance, if X, Y and Z are independent random variables, then the covariance terms are zero and:

$$\mathrm{Var}[aX + bY - cZ] = a^2\mathrm{Var}[X] + b^2\mathrm{Var}[Y] + c^2\mathrm{Var}[Z] \tag{6.168}$$

In this respect, (6.167) becomes:

$$\begin{aligned}\mathrm{Var}_{\widetilde{V}_{re,x}} = {} & V_x^2\cos^2\delta_x\mathrm{Var}[\cos\widetilde{\delta}] \\ & + \cos^2\delta_x\mathrm{Var}[\widetilde{V}\cos\widetilde{\delta}] + V_x^2\sin^2\delta_x\mathrm{Var}[\sin\widetilde{\delta}] + \sin^2\delta_x\mathrm{Var}[\widetilde{V}\sin\widetilde{\delta}]\end{aligned} \tag{6.169}$$

By using the definition of the variance

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \tag{6.170}$$

and its following property for the product of two independent variables X, Y:

$$\mathrm{Var}[XY] = (\mathbb{E}[X])^2\mathrm{Var}[Y] + (\mathbb{E}[Y])^2\mathrm{Var}[X] + \mathrm{Var}[X]\mathrm{Var}[Y] \tag{6.171}$$

Equation (6.169) becomes:

$$\begin{aligned}\mathrm{Var}_{\widetilde{V}re,x} &= V_x^2\cos^2\delta_x(\mathbb{E}[\cos^2\widetilde{\delta}]-(\mathbb{E}[\cos\widetilde{\delta}])^2)\\ &+\cos^2\delta_x((\mathbb{E}[\widetilde{V}])^2\mathrm{Var}[\cos\widetilde{\delta}]+(\mathbb{E}[\cos\widetilde{\delta}])^2\mathrm{Var}[\widetilde{V}]+\mathrm{Var}[\widetilde{V}]\mathrm{Var}[\cos\widetilde{\delta}])\\ &+V_x^2\sin^2\delta_x(\mathbb{E}[\sin^2\widetilde{\delta}]-(\mathbb{E}[\sin\widetilde{\delta}])^2)\\ &+\sin^2\delta_x((\mathbb{E}[\widetilde{V}])^2\mathrm{Var}[\sin\widetilde{\delta}]+(\mathbb{E}[\sin\widetilde{\delta}])^2\mathrm{Var}[\widetilde{V}]+\mathrm{Var}[\widetilde{V}]\mathrm{Var}[\sin\widetilde{\delta}])\end{aligned} \tag{6.172}$$

The expectation $\mathbb{E}[\cos^2\widetilde{\delta}]$ is obtained in the following way:

$$\mathbb{E}[\cos^2\widetilde{\delta}]=\mathbb{E}\left[\frac{1}{2}+\frac{\cos(2\widetilde{\delta})}{2}\right]=\frac{1}{2}(1+e^{-2\sigma_\delta^2}) \tag{6.173}$$

The expectation $\mathbb{E}[\sin^2\widetilde{\delta}]$ is computed in a similar way and is equal to:

$$\mathbb{E}[\sin^2\widetilde{\delta}]=\frac{1}{2}(1-e^{-2\sigma_\delta^2}) \tag{6.174}$$

By using the fact that $\mathbb{E}[\sin(\widetilde{\delta})]=0$, $\mathrm{Var}[\widetilde{V}]=\sigma_V^2$, $\mathbb{E}[\widetilde{V}]=0$, as well as (6.164), (6.173) and (6.174), equation (6.172) becomes:

$$\begin{aligned}\mathrm{Var}_{\widetilde{V}re,x} &= V_x^2\cos^2\delta_x\frac{1}{2}(1+e^{-2\sigma_\delta^2})-V_x^2\cos^2\delta_x e^{-\sigma_\delta^2}+V_x^2\sin^2\delta_x\frac{1}{2}(1-e^{-2\sigma_\delta^2})\\ &+\sigma_V^2\sin^2\delta_x\frac{1}{2}(1-e^{-2\sigma_\delta^2})+\sigma_V^2\cos^2\delta_x\frac{1}{2}(1+e^{-2\sigma_\delta^2})\end{aligned} \tag{6.175}$$

Equation (6.175) can be re-written as:

$$\begin{aligned}\mathrm{Var}_{\widetilde{V}re,x} &= \frac{V_x^2\cos^2\delta_x}{2}+\frac{V_x^2\cos^2\delta_x e^{-2\sigma_\delta^2}}{2}-V_x^2\cos^2\delta_x e^{-\sigma_\delta^2}\\ &+\frac{V_x^2\sin^2\delta_x}{2}-\frac{V_x^2\sin^2\delta_x e^{-2\sigma_\delta^2}}{2}+\frac{\sigma_V^2\sin^2\delta_x}{2}\\ &-\frac{\sigma_V^2\sin^2\delta_x e^{-2\sigma_\delta^2}}{2}+\frac{\sigma_V^2\cos^2\delta_x}{2}+\frac{\sigma_V^2\cos^2\delta_x e^{-2\sigma_\delta^2}}{2}\end{aligned} \tag{6.176}$$

and

$$\begin{aligned}\mathrm{Var}_{\widetilde{V}re,x} &= V_x^2\cos^2\delta_x e^{-\sigma_\delta^2}\left(\frac{e^{\sigma_\delta^2}+e^{-\sigma_\delta^2}}{2}\right) &(6.177)\\ &-V_x^2\cos^2\delta_x e^{-\sigma_\delta^2}+V_x^2\sin^2\delta_x e^{-\sigma_\delta^2}\left(\frac{e^{\sigma_\delta^2}-e^{-\sigma_\delta^2}}{2}\right)\\ &=\sigma_V^2\cos^2\delta_x e^{-\sigma_\delta^2}\left(\frac{e^{\sigma_\delta^2}+e^{-\sigma_\delta^2}}{2}\right)+\sigma_V^2\sin^2\delta_x e^{-\sigma_\delta^2}\left(\frac{e^{\sigma_\delta^2}-e^{-\sigma_\delta^2}}{2}\right) &(6.178)\end{aligned}$$

By using the trigonometric identities for the hyperbolic sine and cosine:

$$\sinh x = \frac{e^x - e^{-x}}{2} \tag{6.179}$$

$$\cosh x = \frac{e^x + e^{-x}}{2} \tag{6.180}$$

the measurement variance of the real part of the voltage can be finally written as:

$$\begin{aligned}\mathrm{Var}_{\widetilde{V}_{re,x}} = {} & V_x^2 e^{-\sigma_\delta^2}\left[\cos^2\delta_x(\cosh(\sigma_\delta^2)-1)+\sin^2\delta_x\sinh(\sigma_\delta^2)\right] \\ & + \sigma_V^2 e^{-\sigma_\delta^2}\left[\cos^2\delta_x\cosh(\sigma_\delta^2)+\sin^2\delta_x\sinh(\sigma_\delta^2)\right]\end{aligned} \tag{6.181}$$

The measurement variance of the imaginary part of the voltage is calculated in the same way and is equal to:

$$\begin{aligned}\mathrm{Var}_{\widetilde{V}_{im,x}} = {} & V_x^2 e^{-\sigma_\delta^2}\left[\sin^2\delta_x(\cosh(\sigma_\delta^2)-1)+\cos^2\delta_x\sinh(\sigma_\delta^2)\right] \\ & + \sigma_V^2 e^{-\sigma_\delta^2}\left[\sin^2\delta_x\cosh(\sigma_\delta^2)+\cos^2\delta_x\sinh(\sigma_\delta^2)\right]\end{aligned} \tag{6.182}$$

So far, the mean values and the variances of the errors in rectangular coordinates have been derived as a function of both the true quantities and the measurement errors expressed in polar coordinates. The mean values are given in (6.165) and (6.166), whereas the variances are given in (6.181) and (6.182). However, the true system state is *not* known. Therefore, the error statistics have to be re-expressed as a function of the measurements by introducing *secondary errors* [39].

Starting from (6.165) and using (6.146) and (6.147), the mean value of the error of the real part of the voltage can be expressed as:

$$\mu_{\widetilde{V}_{re,x}} = (V_z - \widetilde{V})\cos(\delta_z - \widetilde{\delta})\left(e^{-\frac{1}{2}\sigma_\delta^2} - 1\right) \tag{6.183}$$

Then, by using the trigonometric identities (6.152) and (6.153) and after some algebraic manipulations, (6.183) becomes:

$$\begin{aligned}\mu_{\widetilde{V}_{re,x}} = {} & V_z\cos\delta_z\cos\widetilde{\delta}\left(e^{-\frac{1}{2}\sigma_\delta^2}-1\right) - \widetilde{V}\cos\delta_z\cos\widetilde{\delta}\left(e^{-\frac{1}{2}\sigma_\delta^2}-1\right) \\ & + V_z\sin\delta_z\sin\widetilde{\delta}\left(e^{-\frac{1}{2}\sigma_\delta^2}-1\right) - \widetilde{V}\sin\delta_z\cos\widetilde{\delta}\left(e^{-\frac{1}{2}\sigma_\delta^2}-1\right)\end{aligned} \tag{6.184}$$

As mentioned in Reference 39, the new mean value of the error of the real part of the voltage, conditioned on the measured state, is:

$$\mu_{\widetilde{V}_{re,z}} = \mathbb{E}[\mu_{\widetilde{V}_{re,x}} | V_z, \delta_z] \tag{6.185}$$

Since $\widetilde{V}$ and $\widetilde{\delta}$ are assumed to be independent and $\mathbb{E}[\sin(\widetilde{\delta})] = 0$, $\mathbb{E}[\widetilde{V}] = 0$, (6.185) becomes:

$$\mu_{\widetilde{V}_{re,z}} = V_z\cos\delta_z\left(e^{-\sigma_\delta^2} - e^{-\frac{1}{2}\sigma_\delta^2}\right) \tag{6.186}$$

The mean value of the error of the imaginary part of the voltage, conditioned on the measured state, is equal to:

$$\mu_{\widetilde{V}_{im,z}} = V_z\sin\delta_z\left(e^{-\sigma_\delta^2} - e^{-\frac{1}{2}\sigma_\delta^2}\right) \tag{6.187}$$

The new variances, associated to the new mean values, e.g., Reference 39, are given by:

$$\begin{aligned}\mathrm{Var}_{\tilde{V}re,z} = {} & V_z^2 e^{-2\sigma_\delta^2}[\cos^2\delta_z(\cosh(2\sigma_\delta^2)-\cosh(\sigma_\delta^2)) \\ & +\sin^2\delta_z(\sinh(2\sigma_\delta^2)-\sinh(\sigma_\delta^2))] \\ & +\sigma_V^2 e^{-2\sigma_\delta^2}[\cos^2\delta_z(2\cosh(2\sigma_\delta^2)-\cosh(\sigma_\delta^2)) \\ & +\sin^2\delta_z(2\sinh(2\sigma_\delta^2)-\sinh(\sigma_\delta^2))]\end{aligned} \tag{6.188}$$

and

$$\begin{aligned}\mathrm{Var}_{\tilde{V}im,z} = {} & V_z^2 e^{-2\sigma_\delta^2}[\sin^2\delta_z(\cosh(2\sigma_\delta^2)-\cosh(\sigma_\delta^2)) \\ & +\cos^2\delta_z(\sinh(2\sigma_\delta^2)-\sinh(\sigma_\delta^2))] \\ & +\sigma_V^2 e^{-2\sigma_\delta^2}[\sin^2\delta_z(2\cosh(2\sigma_\delta^2)-\cosh(\sigma_\delta^2)) \\ & +\cos^2\delta_z(2\sinh(2\sigma_\delta^2)-\sinh(\sigma_\delta^2))]\end{aligned} \tag{6.189}$$

Note that the variances conditioned on the measured values are larger than the ones conditioned on the true state. This is normal, since they account for the additional errors due to the evaluation at the measured position.

6.5 Data conditioning and bad data processing in PMU-based state estimators

One of the main features of SE is the ability to detect and identify errors in the measurements, the so-called *bad data*. Detection means recognizing the presence of bad data in the measurement set; identification means determining which measurements are bad data. In general, the inclusion of bad data in the estimation process significantly deteriorates the accuracy of the estimate, although the same bad data can have different influence on the estimate depending on the employed SE method and measurement redundancy. Besides, the presence of bad data is common in power systems, since the amount of collected measurements is usually very large and the error sources are several. Therefore, the coupling of a bad data processor to state estimators that are vulnerable to bad data, such as the WLS, is of fundamental importance to maintain the reliability of the estimates and, consequently, the confidence of the system operator on the state estimator. Instead, some state estimators belonging to the category of the robust state estimators, are designed to have an intrinsic bad data rejection capability, without the need of being equipped with a separate bad data processor [19]. One of the most common robust estimators is the least absolute value (LAV) estimator [40]. The main drawback of LAV is the high computational time that can be significantly decreased if the measurement model is linear. Thus, the

use of synchrophasor-only measurements allows LAV to become computationally competitive with the LWLS.

Bad data consisting of gross errors are measurements very distant from their expected value, such as negative magnitude values, and can be identified with simple plausibility checks. Examples of possible causes of gross errors are telecommunication network or meter failure, erroneous wire connections, software bugs and loss of GPS signal (if GPS-synchronized meters are used). Thus, an algorithm that analyses the incoming measurements seeking for gross errors and missing data is suggested. This process can also include data conditioning algorithms that refine the raw measurements, such as procedures for the replacement of missing data or for the filtering of undesired disturbances. Data conditioning is particularly important for PMU measurements since they are vulnerable to additional sources of errors due to their high accuracy and time resolution [41, 42].

Some other bad data consist of measurement errors with magnitude a few times larger than their expected standard deviation that still may significantly affect the estimate accuracy, but require more advanced methods to be detected. These errors can be due, for instance, to the degradation of the meter accuracy or to electromagnetic interferences. Static state estimators dealing with this kind of bad data are coupled with bad data processors placed after the estimation, called *post-estimation* methods. The availability of the estimate allows post-estimation methods to exploit the statistical properties of SE. Typically, the analysis is carried out on the objective function and the measurement estimation residuals, since they quantify how well the set of measurements fits the network model. The common assumption of post-estimation methods is the exact knowledge of the network model. The consequence of such hypothesis is that any inconsistency between measurements and network model is attributed to the former as a bad data. However, network parameter errors can have a similar effect of bad data on the state estimate, so that a bad data processor may identify persistent bad data in presence of network parameter errors.[11] In most of the cases, post-estimation methods are reliable and precise, but this usually comes at the expense of computational time, because after the identification of bad data the state has to be re-estimated iteratively until no more bad data are diagnosed.

A well-known bad data detection method is the χ^2-test, but it usually fails to detect bad data if the error is lower than 20 standard deviations [19, 44]. It is based on the assumption that each measurement residual r_i is a randomly distributed Gaussian variable with zero mean and variance R_{ii}. Thus, it can be easily demonstrated that the objective function $J(\widehat{\mathbf{x}})$ has a χ^2-distribution with $(m-n)$ degrees of freedom. The test reveals the presence of bad data if $J(\widehat{\mathbf{x}}) > \chi^2_{(m-n),\zeta}$, where ζ is the chosen detection confidence probability, e.g., 95% or 99%. The values of $\chi^2_{(m-n),\zeta}$ can be found in dedicated tables or using the Matlab function $chi2inv(\zeta, m-n)$.

Another widely used post-estimation method for bad data detection and identification is the largest normalized residual (LNR) test, thanks to its accuracy and

[11]An effective method to distinguish between bad data and network parameter errors is presented in Reference 43.

straightforward implementation [19, 44]. It requires the calculation of the normalized measurement estimation residual vector $\widehat{\mathbf{r}}^N$ as in (6.90). Each normalized residual should be distributed as $\sim \mathcal{N}(0, 1)$; therefore, the LNR test detects bad data if at least one element of $\widehat{\mathbf{r}}^N$ exceeds a certain threshold, e.g., 3 or 4. Assuming that $\widehat{\mathbf{r}}_i^N$ is the largest residual exceeding the threshold, the i^{th} measurement is flagged as bad data and the SE is re-computed without using this measurement. This procedure is iterated until no bad data are detected. The LNR test limits its identification capability to a single bad data or multiple non-interacting bad data, namely multiple bad data appearing simultaneously whose residuals are not correlated. In case of multiple interacting bad data (whose residuals are correlated), the hypothesis testing identification method has shown better performance [19, 45]. Moreover, critical measurements or measurements belonging to a critical pair cannot be identified as bad data by post-estimation algorithms, since their residuals are always equal to zero. A critical measurement is a measurement that makes the network unobservable if removed, while a critical pair is composed of two measurements that makes the network unobservable if removed simultaneously. It is evident that an abundant measurement redundancy helps substantially the detection and identification of bad data.

Concerning estimators that include a process model, such as the recursive ones, bad data processing can be performed before and/or after the estimation using, respectively, pre- and/or post-estimation methods. Pre-estimation methods consist in statistical procedures or simply logical checks that typically involve the examination of the normalized measurement innovation vector defined in (6.111) [46, 47, 42]. Therefore, they cannot be optimal and their reliability is enhanced if combined with a post-estimation method accounting also for the measurement estimation residuals [47]. High values of some elements of the innovation vector can be due to bad data or to sudden variation of the system state caused, for instance, by inrushes, faults and disconnection of loads or generators. Before performing bad data identification, pre-estimation methods need an algorithm that distinguishes between the two.

Historically, static state estimators, and therefore post-estimation methods, have been the main subject of the literature research and have been employed in real applications. The reasons are multiple and the main ones are listed hereafter. Static state estimators are characterized by lower computational complexity with respect to recursive ones. Furthermore, they are, in general, more reliable since they use only one measurement set without the need to consider a process model that may lead to estimation errors in case of sudden and unexpected changes of the power system state. The meters providing measurements for SE are usually voltage, current and power meters that send the measurement packet at intervals of several seconds or a few minutes. Hence, there is sufficient time to perform post-estimation analysis. However, the recent inclusion of PMU data in SE have made possible to dramatically increase the SE refresh rate to a few tens of milliseconds and to use effectively the process model of (6.77). Although the bad data processing theory proposed by the literature remains valid, PMUs have raised new interest on recursive state estimators, and consequently on pre-estimation methods, as well as on data conditioning of raw measurements.

6.6 Kalman filter vs. weighted least squares

This section compares theoretically the accuracy of KF vs. WLS. To formally quantify this difference, it is useful to recall that the KF process makes use of all the available measurements, past and present, whereas the WLS algorithm uses only measurements of the current time-step. The former should intuitively perform better, provided that the process model hypotheses that underlie the KF are correct. The following theorem formalizes this aspect. It states that the estimation error with the KF algorithm is always less than the estimation error with the WLS algorithm, the difference being given equal to the mean square difference between the two methods:

Theorem 6.1. *Assume that the true (unobserved) state* $\mathbf{x}_t$ *satisfies the process model in* (6.77). *Assume that the system parameters are known. Let* $\widehat{\mathbf{x}}_{t,WLS}$ *and* $\widehat{\mathbf{x}}_{t,KF}$ *be the state estimates obtained at time-step t with the WLS and KF algorithms, respectively. Then*

$$\mathbb{E}[\|\mathbf{x}_t - \widehat{\mathbf{x}}_{t,WLS}\|^2] = \mathbb{E}[\|\mathbf{x}_t - \widehat{\mathbf{x}}_{t,KF}\|^2] + \mathbb{E}[\|\widehat{\mathbf{x}}_{t,WLS} - \widehat{\mathbf{x}}_{t,KF}\|^2] \tag{6.190}$$

Proof. First of all, by standard KF theory, e.g., Reference 48, the estimation of the non-observable state is equal to its conditional expectation, given the sequence of measurements, i.e.,

$$\widehat{\mathbf{x}}_{t,KF} = \mathbb{E}[\mathbf{x}_t|\mathcal{F}_t] \tag{6.191}$$

where $\mathcal{F}_t$ the σ-field generated by all measurements up to and including time-step t.

Second, consider the Hilbert space of random vectors (with values in $\mathbb{R}^n$) equipped with the inner product $\langle \mathbf{X}, \mathbf{Y}\rangle_H \triangleq \mathbb{E}[\sum_{c=1}^n X_c Y_c]$. It is requested to show that the random vector $\mathbf{x}_t - \widehat{\mathbf{x}}_{t,KF}$ is orthogonal,[12] in the sense of this Hilbert space, to all random vectors $\mathbf{Y}$ that are $\mathcal{F}_t$-measurable (i.e., that are a function of the measurements up to time t; in this context, the initial conditions of the estimation algorithms are assumed to be known and non-random). Note that what needs to be shown is that:

$$\mathbb{E}[\langle \mathbf{Y}, \mathbf{x}_t\rangle] = \mathbb{E}[\langle \mathbf{Y}, \widehat{\mathbf{x}}_{t,KF}\rangle] \tag{6.192}$$

with $\langle \mathbf{Y}, \mathbf{x}_t\rangle = \sum_{c=1}^n Y_c x_{t,c}$.

To prove (6.192), observe that, for any (real-valued) random variable U that is measurable with respect to $\mathcal{F}_t$, it can be written that $\mathbb{E}[U|\mathcal{F}_t] = U$ and further [48]:

$$\mathbb{E}[Ux_{t,c}|\mathcal{F}_t] = U\mathbb{E}[x_{t,c}|\mathcal{F}_t] = U\widehat{x}_{t,c,KF}$$

Note that such a U can be any non-linear real-valued function of $(\mathbf{z}_1, \ldots, \mathbf{z}_t)$. Take expectations on both sides and use the fact that the expectation of the conditional expectation is the same as the original expectation (*law of total expectation*, e.g., Reference 48) and obtain

$$\mathbb{E}[Ux_{t,c}] = \mathbb{E}[U\widehat{x}_{t,c,KF}] \tag{6.193}$$

[12]We remind here that when two vectors are orthogonal, their inner product is equal to zero.

Consider now any $\mathcal{F}_t$-measurable random vector $\mathbf{Y}$, apply (6.193) to $U = Y_c$ for all coordinates c and sum over c; it comes:

$$\mathbb{E}\left[\sum_{c=1}^{n} Y_c x_{t,c}\right] = \mathbb{E}\left[\sum_{c=1}^{n} Y_c \widehat{x}_{t,c,KF}\right] \tag{6.194}$$

which shows (6.192) as required and it means that:

$$\mathbb{E}\left[\sum_{c=1}^{n} Y_c (\widehat{x}_{t,c} - \widehat{x}_{t,c,KF})\right] = \mathbb{E}[\langle \mathbf{Y}, \mathbf{x}_t - \widehat{\mathbf{x}}_{t,KF}\rangle] = 0 \tag{6.195}$$

namely $\mathbf{Y}$ and $\mathbf{x}_t - \widehat{\mathbf{x}}_{t,KF}$ are *orthogonal*.

Now observe that both $\widehat{\mathbf{x}}_{t,WLS}$ and $\widehat{\mathbf{x}}_{t,KF}$ are $\mathcal{F}_t$-measurable because they are derived from the measurements. Therefore, the previous result can be applied to $\mathbf{Y} = \widehat{\mathbf{x}}_{t,WLS} - \widehat{\mathbf{x}}_{t,KF}$. Equation (6.190) then follows from Pythagoras's equality. □

Final remark: The theorem applies as long as the process model in (6.77) holds. This explains why it is important to verify the adequacy of the process model.

6.7 Numerical validation and performance assessment of the state estimation

6.7.1 Linear state estimation case studies

6.7.1.1 Distribution network case study: IEEE 13-bus distribution test feeder

The adopted IEEE 13-bus distribution test feeder is shown in Figure 6.3 and is based on Reference 49. It is a 3-ph feeder of 13 buses where Bus 1 represents the connection to the sub-transmission network with a short-circuit power $S_{sc} = 300$ MVA and a ratio between the real and imaginary parts of the short-circuit impedance $R_{sc}/X_{sc} = 0.1$. The network has a rated voltage equal to 15 kV line-to-line root mean square (RMS) and the voltage base $V_b = 15$ kV (line-to-line RMS). The power base $S_b = 10$ MVA. The lines are unbalanced and the used line configuration is the #602 of Reference 49. The values of the resistance, reactance and susceptance and the line lengths are given in Appendix B. The loads/distributed energy resources (DERs) are also characterized by unbalanced power absorptions/injections, respectively. Figure 6.4 shows the aggregated active and reactive power consumption of the loads in three phases, whereas Figure 6.5 shows the power injected by DERs. The data comes from an experimental campaign in real distribution grids in the South-West region of Switzerland, in the EPFL campus and in a feeder in the Netherlands (BML 2.10 distribution feeder operated by Alliander). The DERs (one Mini-Hydro power plant at Bus 4 and one photovoltaic unit at Bus 12) inject only active power.

Table 6.1 shows the location of PMUs for the IEEE 13-bus distribution test feeder. In this case, the number of PMUs is equal to 7. It is assumed that each PMU measures the bus phase-to-ground voltage phasors and the nodal current injection phasors; therefore in this case, d_1 and d_2 of (6.4) are equal to 7 and d_3 is equal to 0.

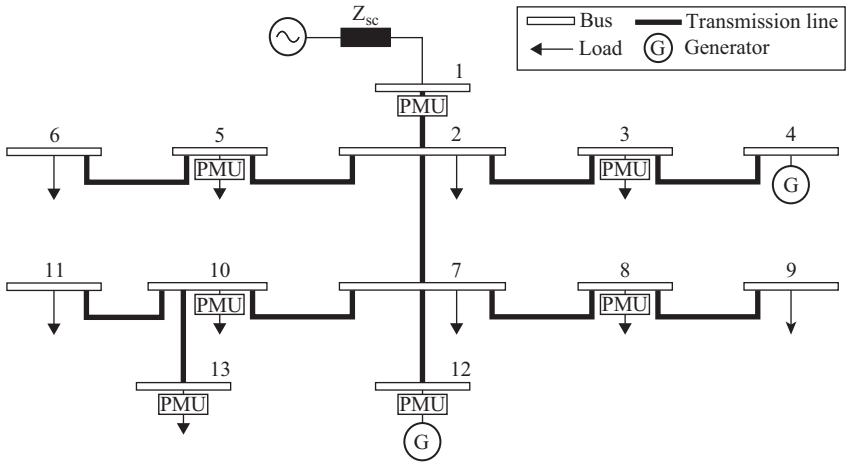

Figure 6.3 Network topology of the IEEE 13-bus distribution test feeder [49], together with the adopted PMU placement. We assume that Bus 1 is the connection point of the system to an external network that is represented by a voltage source in series with the short-circuit impedance Z_{sc}

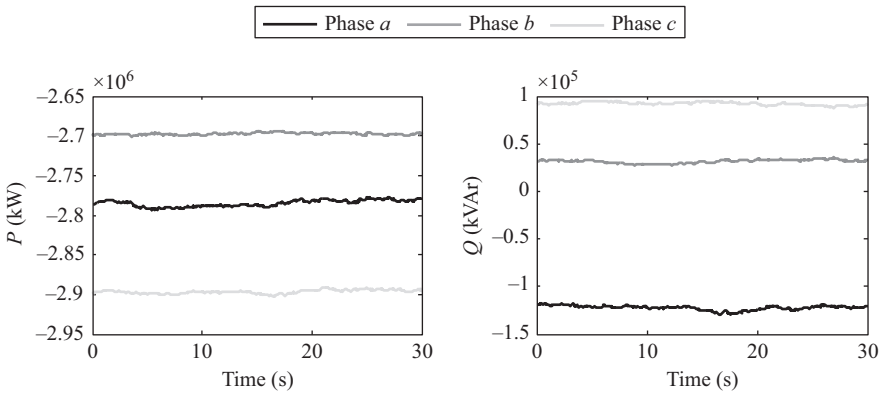

Figure 6.4 Time evolution of the active and reactive powers of the loads, per phase, used in the IEEE 13-bus distribution test feeder

The SE accuracy assessment of the IEEE 13-bus distribution test feeder is performed by using the LWLS-SE and the DKF-SE. This is justified by the fact that the measurements come only from PMUs; therefore, the SE is linear. The value of $\mathbf{Q}$ used by the DKF-SE is assessed using the method presented in Section 6.3.4, where the parameter $N = 20$. Current and voltage sensors used to perform the tests are assumed to be of 0.1-class. Limits of ratio error and phase displacement imposed by References 50, 51 are shown in Tables 6.2 and 6.3, respectively.

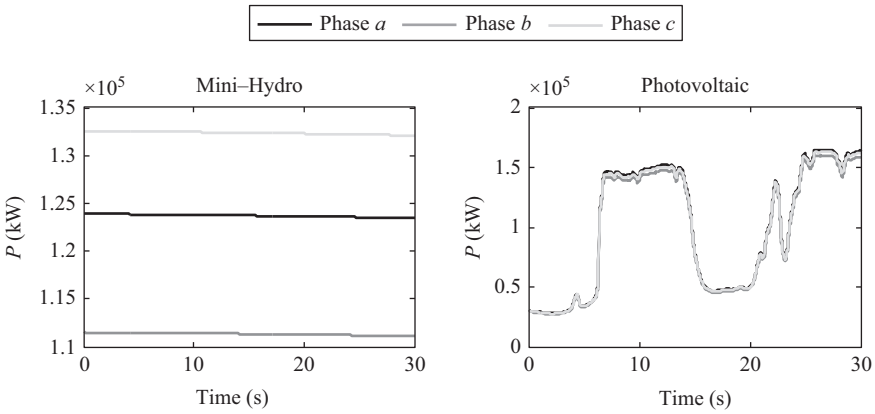

Figure 6.5 Time evolution of the active powers of the two DERs (one Mini-Hydro power plant at Bus 4 and one photovoltaic unit at Bus 12), per phase, used in the IEEE 13-bus distribution test feeder

Table 6.1 PMUs location in the IEEE 13-bus distribution test feeder

	Number	**Buses**
PMUs	7	1 3 5 8 10 12 13

Table 6.2 Limits of ratio error and phase displacement of the used current sensors in the IEEE 13-bus distribution test feeder, according to Reference 50

Class	**Ratio error (%)**	**Phase displacement (rad)**
0.1	0.1	1.5×10^{-3}

Table 6.3 Limits of ratio error and phase displacement of the used voltage sensors in the IEEE 13-bus distribution test feeder, according to Reference 51

Class	**Ratio error (%)**	**Phase displacement (rad)**
0.1	0.1	1.5×10^{-3}

Table 6.4 Limits of magnitude and phase errors for the used PMUs in the IEEE 13-bus distribution test feeder

TVE (in %)	Magnitude error (%)	Phase error (rad)
0.14	0.1	10^{-3}

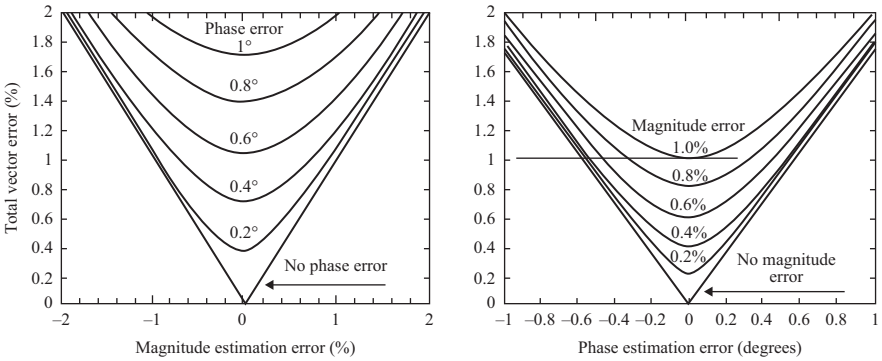

Figure 6.6 Equivalence of magnitude and phase error with PMUs TVE values (adapted from Reference 1)

The limits of magnitude and phase errors for the adopted PMUs are shown in Table 6.4 and correspond to a total vector error (TVE) equal to 0.14%. Such a TVE value results from assuming PMU class-P devices characterized by typical maximum errors in magnitude and phase of 0.1% and 10^{-3} rad, respectively.

Here below are given the steps for the numerical derivation of the measurement noise covariance matrix **R**, using the theory in Section 6.4.

1. The values of the maximum errors for the 0.1-class electrical sensors, in terms of magnitude error and phase displacement, are given in Tables 6.2 and 6.3. Hence, the limit for the magnitude error is 0.1%, whereas the maximum phase displacement is 1.5×10^{-3} rad.
   The maximum magnitude error of PMUs is obtained from Figure 6.6 by assuming a TVE equal to 0.1% and a null phase error. Then, the maximum phase error is obtained from Figure 6.6 by assuming a TVE equal to 0.1% and a null magnitude error. The maximum magnitude and phase errors of PMUs are, respectively, 0.1% and 1×10^{-3} rad, which corresponds to a TVE of 0.14%.
   Then, the cumulative maximum errors of sensors and PMUs are $(0.1 + 0.1)$% for the voltage magnitude and $(1.5 \times 10^{-3} + 1 \times 10^{-3})$ rad for the voltage phase.

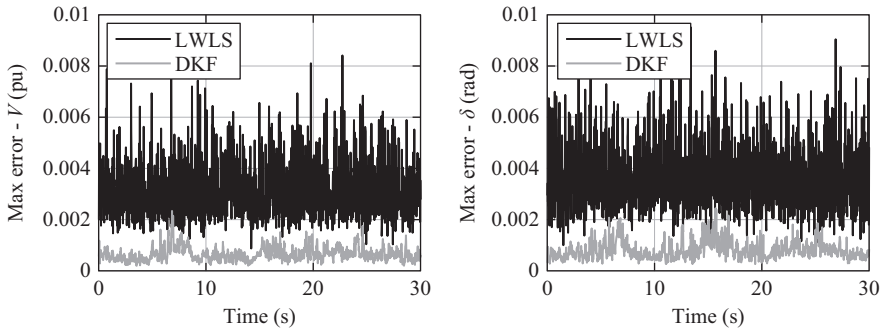

Figure 6.7 Time evolution of the maximum estimation errors calculated by using (6.196)*, for the LWLS-SE and DKF-SE, case of the IEEE 13-bus test feeder*

2. The second step consists in the expression of the maximum magnitude errors with respect to the measured values.
3. The third step includes the calculation of the cumulative standard deviations, which are equal to one-third of the maximum errors, assuming that the 99.73% of the values lie within three standard deviations from the mean.
4. By using the cumulative standard deviations from the previous step and the measurements, the variances in rectangular coordinates are calculated by using (6.188) and (6.189).
5. The diagonal of the measurement noise covariance matrix is composed of the aforementioned variances.

Figure 6.7 shows the time evolution of the maximum errors of the estimated state vs. the true one for the LWLS-SE and DKF-SE, for a time window of 30 s and a resolution of 20 ms.[13] Errors refer to both voltage magnitude V and phase δ. The magnitude error is expressed in pu and the phase error is expressed in radians. At time-step t, the maximum estimation error (MEE) is calculated considering the estimation errors in all the buses and the three phases as:

$$\begin{aligned} \text{MEE}(V_t) &= \max[\text{err}V^a_{1,t}, \text{err}V^b_{1,t}, \text{err}V^c_{1,t}, \ldots, \text{err}V^a_{n,t}, \text{err}V^b_{n,t}, \text{err}V^c_{n,t}] \\ \text{MEE}(\delta_t) &= \max[\text{err}\delta^a_{1,t}, \text{err}\delta^b_{1,t}, \text{err}\delta^c_{1,t}, \ldots, \text{err}\delta^a_{n,t}, \text{err}\delta^b_{n,t}, \text{err}\delta^c_{n,t}] \end{aligned} \tag{6.196}$$

Both the magnitude and the phase maximum errors of the LWLS-SE are larger than the ones of the DKF-SE along the overall simulation time.

To clarify further this aspect, Figure 6.8 shows the root mean squared errors (RMSEs) for the LWLS-SE and the DKF-SE and considering the same time window of Figure 6.7. Only the maximum RMSE among the three phases for each bus is

[13]The KF estimates have always an initial phase in which they converge from an arbitrary initial value (e.g., a flat-start initialization) towards the true state. For brevity, in all the figures that show the KF accuracy performance, this initial phase is not shown.

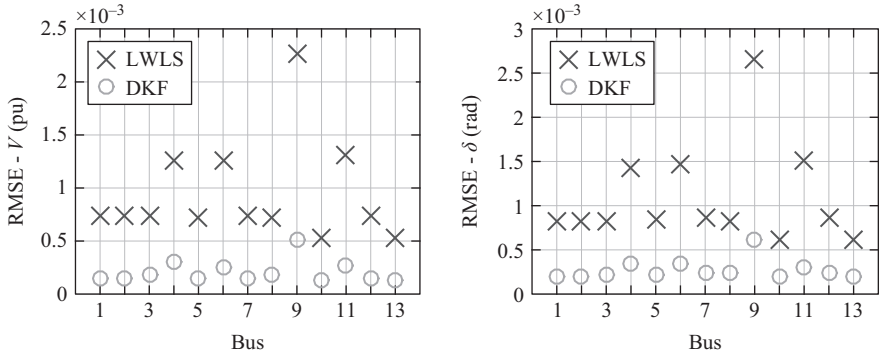

Figure 6.8 RMSEs calculated by using (6.197)*, for the LWLS-SE and DKF-SE, case of the IEEE 13-bus test feeder. Only the maximum RMSE among the three phases for each bus is shown*

shown. The errors refer to both voltage magnitude V and phase δ. The magnitude error is expressed in pu and the phase error is expressed in radians. As known, the RMSE for the voltage magnitude and phase at bus i and phase p is calculated as:

$$\begin{aligned} \text{RMSE}(V_i^p) &= \sqrt{\frac{1}{N_s}\sum_{n_s=1}^{N_s}(\widehat{V}_{i,n_s}^p - V_{i,n_s}^p)^2} \\ \text{RMSE}(\delta_i^p) &= \sqrt{\frac{1}{N_s}\sum_{n_s=1}^{N_s}(\widehat{\delta}_{i,n_s}^p - \delta_{i,n_s}^p)^2} \end{aligned} \tag{6.197}$$

As it can be observed, the RMSEs of LWLS-SE are larger than the ones of the DKF-SE for all the 13 network buses.

6.7.1.2 Transmission system case study: IEEE 39-bus transmission test system

The IEEE 39-bus test system [52], shown in Figure 6.9, is balanced; therefore, only the direct sequence has been considered. Bus 31 in Figure 6.9 is the connection point of the system to an external grid characterized by a short-circuit power $S_{sc} = 50$ GVA and a ratio between real and imaginary parts of the short-circuit impedance $R_{sc}/X_{sc} = 0$, which is a standard assumption for transmission power systems. The IEEE 39-bus system is assumed to have four different voltage levels, i.e., 380 kV, 230 kV, 125 kV and 15 kV. The transformer ratios are given in Appendix B. The chosen value for the base power is $S_b = 100$ MVA. The values of the resistance, reactance and susceptance are given in Appendix B.

Table 6.5 gives the values of the active and reactive power injections at the respective buses, which are in accordance with Reference 52. The convention is that the absorbed powers are marked with a minus, whereas the generated powers have a positive sign.

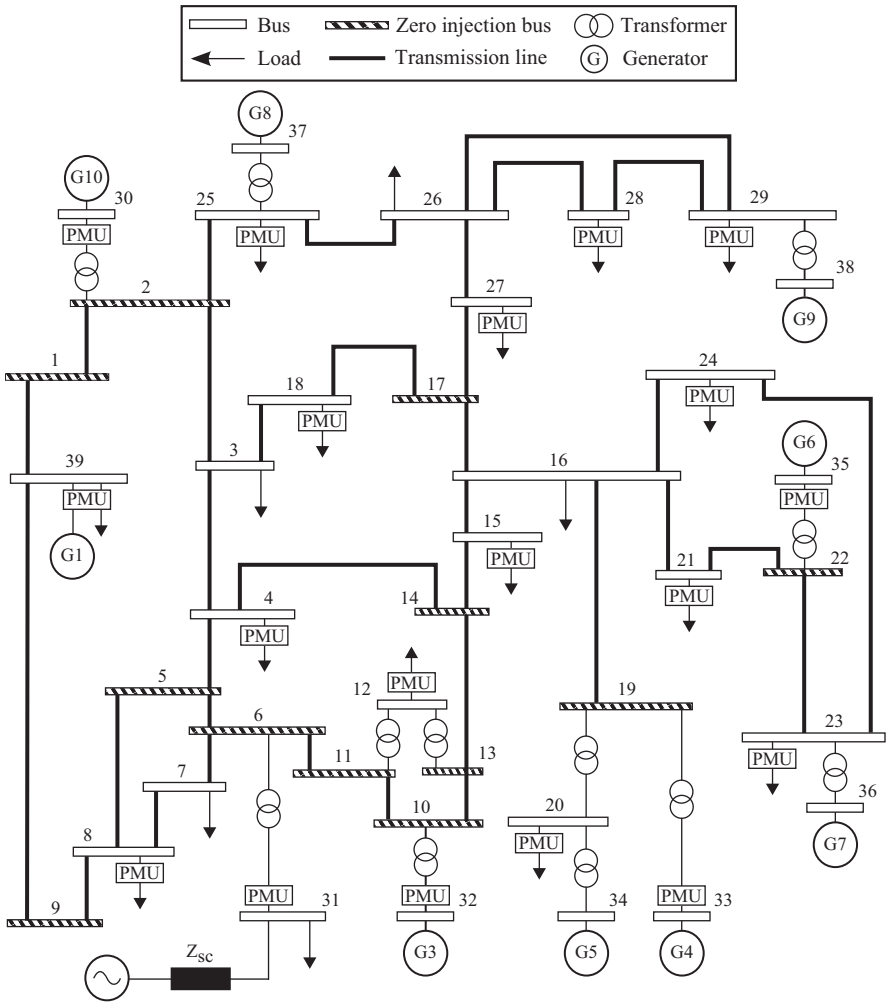

Figure 6.9 Network topology of the IEEE 39-bus test system [52], together with the adopted PMU placement. We assume that Bus 31 is the connection point of the system to an external network that is represented by a voltage source in series with the short-circuit impedance Z_{sc}

The SE accuracy assessment of the IEEE 39-bus transmission test system is performed by using the LWLS-SE and the DKF-SE. The current and voltage sensors used to perform the tests are assumed to be of 0.5-class. The PMU and sensor accuracies are reported in Tables 6.6–6.8. The assumed PMU locations and the zero-injection buses are given in Table 6.9. The variance assigned to zero-injection buses is lower than the ones of the other measurements since they are not affected by error. It is assumed that each PMU measures the bus phase-to-ground voltage phasors and the

Table 6.5 Active and reactive power injections for the IEEE 39-bus transmission test system

Bus	Type	*P* (MW)	*Q* (MW)
3	Load	−322	−2.4
4	Load	−500	−184
7	Load	−233.8	−84
8	Load	−522	−176
12	Load	−7.5	−88
15	Load	−320	−153
16	Load	−329	−32.3
18	Load	−158	−30
20	Load	−628	−103
21	Load	−274	−115
23	Load	−247.5	−84.6
24	Load	−308.6	92
25	Load	−224	−47.2
26	Load	−139	−17
27	Load	−281	−75.5
28	Load	−206	−27.6
29	Load	−283.5	−26.9
30	Generator	250	189.9
31	Load	−9.2	−4.6
32	Generator	650	204.8
33	Generator	632	72.3
34	Generator	508	149.8
35	Generator	650	258.7
36	Generator	560	212.8
37	Generator	540	30.8
38	Generator	830	63.0
39	Load	−104	−250.0

Table 6.6 Limits of magnitude and phase errors for the used PMUs of the IEEE 39-bus transmission test system

TVE (in %)	Magnitude error (%)	Phase error (rad)
0.14	0.1	10^{-3}

nodal current injection phasors; therefore in this case, d_1 and d_2 of (6.4) are equal to 19 and d_3 is equal to 0.

Figure 6.10 shows the time evolution of the maximum errors of the estimated state vs. the true one for the LWLS-SE and DKF-SE. The MEEs are calculated by using (6.196). As it can be observed, both the magnitude and the phase maximum errors of the LWLS-SE are significantly larger than the ones of the DKF-SE along the overall simulation time.

Table 6.7 Limits of ratio error and phase displacements of the used current sensors of the IEEE 39-bus transmission test system, according to Reference 50

Class	Ratio error (%)	Phase displacement (rad)
0.5	0.5	9×10^{-3}

Table 6.8 Limits of ratio error and phase displacement of the used voltage sensors of the IEEE 39-bus transmission test system, according to Reference 51

Class	Ratio error (%)	Phase displacement (rad)
0.5	0.5	6×10^{-3}

Table 6.9 PMU locations and zero-injection buses in the IEEE 39-bus test system

Case	Number	Buses
Zero-injection buses	12	1 2 5 6 9 10 11 13 14 17 19 22
PMUs	19	4 8 12 15 18 20 21 23 24 25 27 28 29 30 31 32 33 35 39

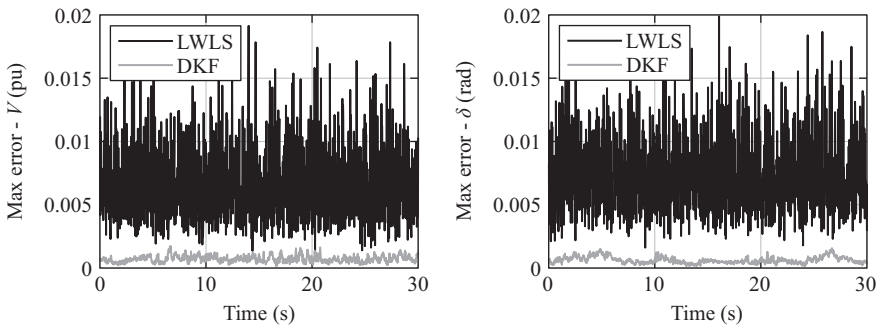

Figure 6.10 Time evolution of the maximum estimation errors calculated by using (6.196)*, for the LWLS-SE and DKF-SE, case of the IEEE 39-bus test system*

Table 6.10 PMUs and RTUs location in the IEEE 13-bus distribution test feeder

	Number	Buses
PMUs	7	1 3 5 8 10 12 13
RTUs	7	1 3 5 8 10 12 13

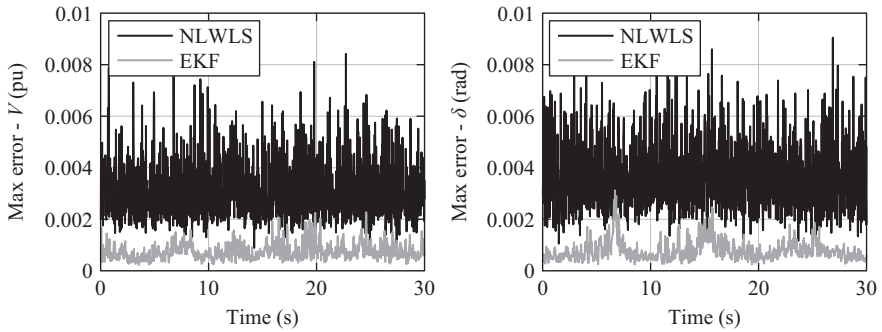

Figure 6.11 Time evolution of the maximum estimation errors calculated by using (6.196)*, for the NLWLS-SE and EKF-SE, case of the IEEE 13-bus test feeder*

6.7.2 Non-linear SE case studies

The network used as a case study to show the performance of the non-linear SE is the 13-bus distribution test case. The network data is the same as the one reported in Section 6.7.1.1. Table 6.10 shows the location of PMUs and RTUs for the IEEE 13-bus distribution test feeder. It is assumed that each PMU measures the bus phase-to-ground voltage phasors and each RTU measures the nodal active and reactive power injections; therefore in this case, d_1 and u_1 of (6.44) are equal to 7 and u_2 is equal to 0.

The accuracy of the PMUs is the one reported in Table 6.4. The cumulative error of sensors and RTUs is assumed to be equal to 3%. The SE accuracy assessment of the IEEE 13-bus distribution test feeder is in this case performed by using the NLWLS-SE and the EKF-SE.

Figure 6.11 shows the time evolution of the maximum errors of the estimated state vs. the true one for the NLWLS-SE and EKF-SE, for a time window of 30 s and a resolution of 20 ms. The errors refer to both voltage magnitude V and phase δ. The magnitude error is expressed in pu and the phase error is expressed in radians. At time-step t, the MEE is calculated by using (6.196). Both the magnitude and the

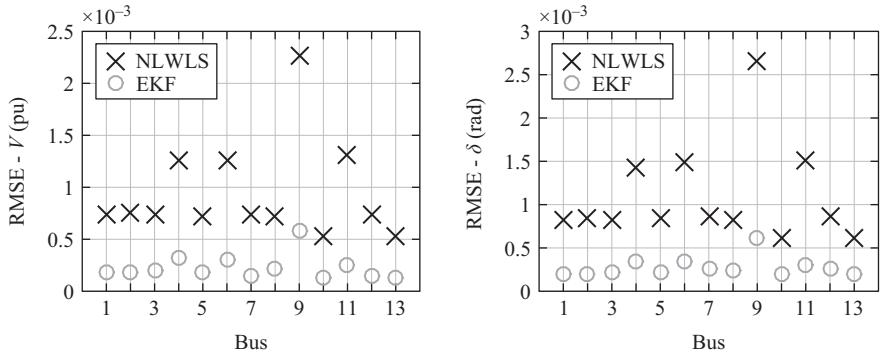

Figure 6.12 RMSEs calculated by using (6.197)*, for the NLWLS-SE and EKF-SE, case of the IEEE 13-bus test feeder. Only the maximum RMSE among the three phases for each bus is shown*

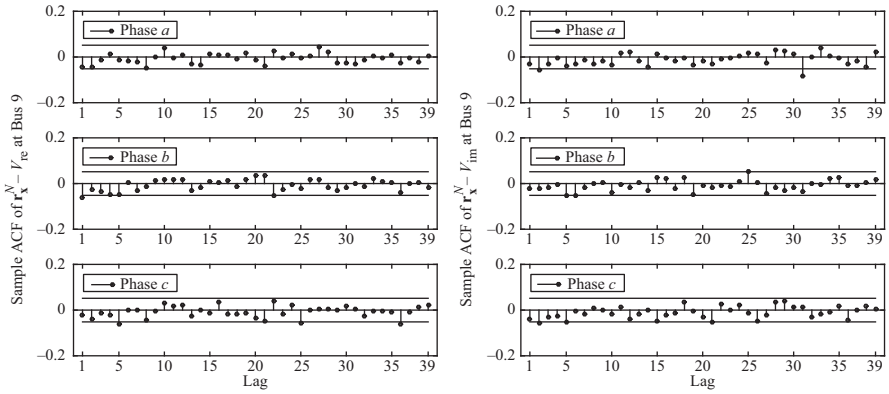

Figure 6.13 Process model validation for the case of linear SE (DKF) of the IEEE 13-bus distribution test feeder: sample ACFs of the normalized residuals of the state estimates referred to the real and imaginary parts of the voltage at Bus 9

phase maximum errors of the NLWLS-SE are larger than the ones of the EKF-SE along the overall simulation time.

Figure 6.12 shows the RMSEs calculated by using (6.197), for the NLWLS-SE and the EKF-SE and considering the same time window of Figure 6.11. Only the maximum RMSE among the three phases for each bus is shown. The errors refer to both voltage magnitude V and phase δ. The magnitude error is expressed in pu and the phase error is expressed in radians. As it can be observed, the RMSEs of NLWLS-SE are larger than the ones of the EKF-SE for all the 13 network buses.

6.8 Kalman filter process model validation

The validity of the KF algorithm depends on how accurate the underlying process model is. More precisely, since the KF equations use only second-order properties, it is sufficient to verify that the covariance properties of the model do hold. It is thus needed to verify whether the normalized residuals of the state estimates $\mathbf{r}_{\mathbf{x}}^{N}$ are uncorrelated by analyzing the sample auto correlation functions (ACFs). The i^{th} element of $\mathbf{r}_{\mathbf{x}}^{N}$ is calculated as:

$$r_{\mathbf{x},i}^{N} = \frac{\widehat{x}_{t,i} - \widehat{x}_{t-1,i}}{\sqrt{Q_{t-1,ii}}} \tag{6.198}$$

The sample ACFs are computed by considering the first $\sim\sqrt{N_S}$ lags, where $N_S = 1\,500$ is the number of simulation time-steps. If the residuals are uncorrelated, the ACFs should be within the noise margins $\pm 1.96/\sqrt{N_S}$ with 95% of probability [53].

The above-mentioned condition is fulfilled for the case of the linear SE of the IEEE 13-bus distribution test feeder, as shown in Figure 6.13. The results refer to the real and imaginary parts of the voltage at Bus 9, which is the bus where the largest estimation error is observed. Since the normalized residuals of the state estimates are uncorrelated, the KF process model is accurate. Note that the sample ACF is, as a matter of fact, a statistically distributed quantity. The fact that, in few cases, the ACFs are slightly beyond the noise margins do not violate the validity of the result and the numerical proof of the statistical correctness of the process model. The same analysis for the case of the non-linear SE of the IEEE 13-bus distribution test feeder is shown in Figure 6.14 with respect to the magnitude and phase of the voltage at Bus 9.

It is worth pointing out that the validation of the KF process model together with the assessment of the process noise covariance matrix $\mathbf{Q}$ require further investigation. Future research shall focus on the analysis of the innovation and measurement residual vectors.

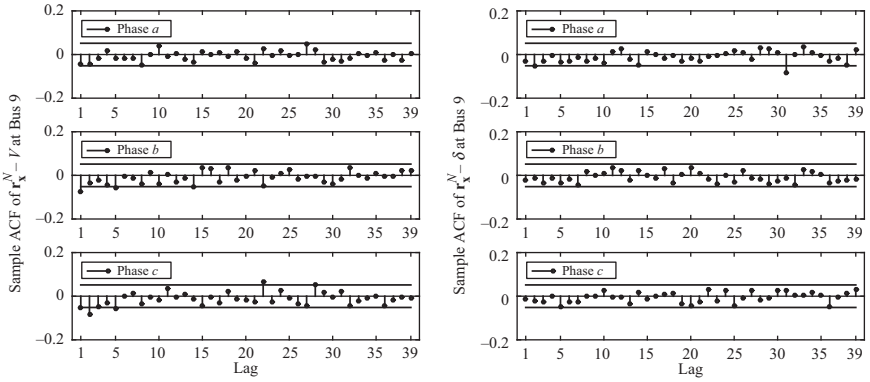

Figure 6.14 Process model validation for the case of non-linear SE (EKF) of the IEEE 13-bus distribution test feeder: sample ACFs of the normalized residuals of the state estimates referred to the magnitude and phase of the voltage at Bus 9

6.9 Numerical validation of Theorem 6.1

In this section, it is verified whether the quantitative conclusion of Theorem 6.1 of Section 6.6, namely (6.190), numerically holds. Figure 6.15 shows the left- and right-hand sides (LHS and RHS, respectively) of (6.190) for the case of the linear SE of the IEEE 13-bus distribution test feeder. The LHS and RHS of (6.190) are close to each other, which means that the equality in Theorem 6.1 is in expectation. In Figure 6.15, expectations are estimated by empirical averages; therefore, a small discrepancy is expected. Figure 6.15 also shows the contributions of the two terms of the RHS of (6.190). It can be observed that the contribution of the second one is predominant with respect to the first one, proving that the DKF is applied correctly.

Figure 6.16 shows the verification of (6.190) for the case of the non-linear SE of the IEEE 13-bus distribution test feeder. As it can be observed, the quantitative conclusion of Theorem 6.1 holds also in the case of the non-linear SE, proving that EKF is applied correctly.

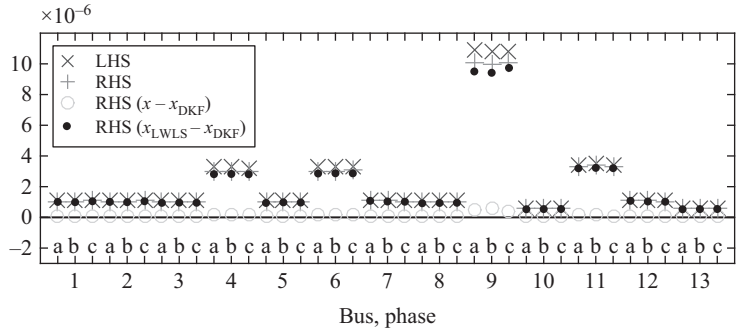

Figure 6.15 Numerical validation of (6.190) for the case of linear SE (DKF) of the IEEE 13-bus distribution test feeder: LHS vs. RHS of the equation. The separate contribution of the two terms of the RHS is also shown

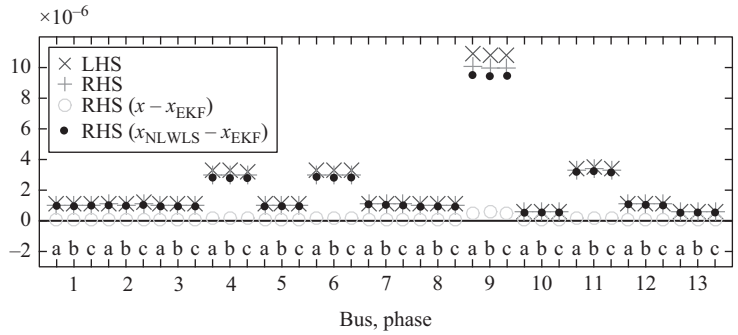

Figure 6.16 Numerical validation of (6.190) for the case of non-linear SE (EKF) of the IEEE 13-bus distribution test feeder: LHS vs. RHS of the equation. The separate contribution of the two terms of the RHS is also shown

Bibliography

[1] *IEEE Standard for Synchrophasor Measurements for Power Systems*. Standard IEEE C37.118.1; 2011.

[2] *IEEE Standard for Synchrophasor Data Transfer for Power Systems*. Standard IEEE C37.118.2; 2011.

[3] Jones KD, Thorp JS, Gardner RM. "Three-phase linear state estimation using phasor measurements". In: *IEEE PES Gen. Meet.* Vancouver, Canada; 2013.

[4] Zhang L, Bose A, Jampala A, Madani V, Giri J. "Design, testing, and implementation of a linear state estimator in a real power system". *IEEE Transactions on Smart Grid* (in press). 2016.

[5] Sarri S, Zanni L, Popovic M, Boudec JYL, Paolone M. "Performance assessment of linear state estimators using synchrophasor measurements". *IEEE Transactions on Instrumentation and Measurement*. 2016 Mar;65(3): 535–548.

[6] Powalko M, Rudion K, Komarnicki P, Blumschein J. "Observability of the distribution system". In: *Proc. 20th Int. Conf. and Exhib. on Electr. Distr., CIRED*. Prague, Czech Republic; 2009.

[7] Laverty DM, Morrow DJ, Best RJ, Crossley PA. "Differential ROCOF relay for loss-of-mains protection of renewable generation using phasor measurement over Internet protocol". In: *CIGRE/IEEE Power Energy Soc. Joint Symp. Integr. of Wide-Scale Ren. Res. into the Power Del. Syst*. Calgary, Canada; 2009.

[8] Samuelsson O, Hemmingsson M, Nielsen AH, Pedersen KOH, Rasmussen J. "Monitoring of power system events at transmission and distribution level". *IEEE Transactions on Power Systems*. 2006;21(2):1007–1008.

[9] Best RJ, Morrow DJ, Laverty DM, Crossley PA. "Synchrophasor broadcast over internet protocol for distributed generator synchronization". *IEEE Transactions on Power Delivery*. 2010 Oct;25(4):2835–2841.

[10] Carta A, Locci N, Muscas C. "GPS-based system for the measurement of synchronized harmonic phasors". *IEEE Transactions on Instrumentation and Measurement*. 2009 Mar;58(3):586–593.

[11] Liu J, Tang J, Ponci F, Monti A, Muscas C, Pegoraro PA. "Trade-offs in PMU deployment for state estimation in active distribution grids". *IEEE Transactions on Smart Grid*. 2012 Jun;3(2):915–924.

[12] Pignati M, Popovic M, Andrade SB, *et al.* "Real-time state estimation of the EPFL campus medium-voltage grid by using PMUs". In: *Sixth Conf. on Innov. Smart Grid Techn. (ISGT 2015)*. Washington, DC; 2015.

[13] Meier AV, Culler D, McEachern A, Arghandeh R. "Microsynchrophasors for distribution systems". In: *IEEE PES Innov. Smart Grid Techn. (ISGT) Conf.* Washington, DC; 2014.

[14] Schweppe FC, Wildes J. "Power system static-state estimation, part I: exact model". *IEEE Transactions on Power Apparatus and Systems*. 1970 Jan;PAS-89(1):120–125.

[15] Schweppe FC, Rom DB. "Power system static-state estimation, part II: approximate model". *IEEE Transactions on Power Apparatus and Systems*. 1970 Jan;PAS-89(1):125–130.
[16] Schweppe FC. "Power system static-state estimation, part III: implementation". *IEEE Transactions on Power Apparatus and Systems*. 1970 Jan;PAS-89(1):130–135.
[17] Grainger JJ, Stevenson WD. *Power System Analysis*. New York, NY: McGraw-Hill International Editions; 1994.
[18] Monticelli A. *State Estimation in Electric Power Systems: A Generalized Approach*. Berlin: Springer; 1999.
[19] Abur A, Expósito AG. *Power System State Estimation: Theory and Implementation*. New York, NY: CRC Press, Marcel Dekker; 2004.
[20] Zhu J, Abur A. "Effect of phasor measurements on the choice of reference bus for state estimation". In: *IEEE PES Gen. Meet.* Tampa, FL; 2007.
[21] Paolone M, Borghetti A, Nucci CA. "A synchrophasor estimation algorithm for the monitoring of active distribution networks in steady state and transient conditions". In: *Power Syst. Comp. Conf. (PSCC)*. Wrocław, Poland; 2014.
[22] Manousakis NM, Korres GN, Georgilakis PS. "Taxonomy of PMU placement methodologies". *IEEE Transactions on Power Systems*. 2012 May;27(2): 1070–1077.
[23] Xu B, Abur A. "Observability analysis and measurement placement for systems with PMUs". In: *Proc. of the IEEE PES Power System Conf. and Exp. 2004*, vol. 2. New York, NY; 2004. p. 943–946.
[24] Welch G, Bishop G. "An introduction to the Kalman filter". Department of Computer Science, University of North Carolina, Chapel Hill, NC; 2006. TR 95-041.
[25] Debs AS, Larson RE. "A dynamic estimator for tracking the state of a power system". *IEEE Transactions on Power Apparatus and Systems*. 1970 Sep;PAS-89(7):1670–1678.
[26] Nishiya K, Hasegawa J, Koike T. "Dynamic state estimation including anomaly detection and identification for power systems". In: *IEE Proc. on Gen., Transm. and Distr.*, vol. 129, no. 5. IET; 1982. p. 192–198.
[27] Zhang J, Welch F, Bishop G, Huang Z. "A two-stage Kalman filter approach for robust and real-time power system state estimation". *IEEE Transactions on Sustainable Energy*. 2014 Apr;5(2):629–636.
[28] Zanni L, Sarri S, Pignati M, Cherkaoui R, Paolone M. "Probabilistic assessment of the process-noise covariance matrix of discrete Kalman filter state estimation of active distribution networks". In: *Intern. Conf. Probab. Meth. Appl. to Power Syst. (PMAPS)*. Durham, UK; 2014.
[29] Caro E, Conejo AJ, Mínguez R. "Power system state estimation considering measurement dependencies". *IEEE Transactions on Power Systems*. 2009;24(4):1875–1885.
[30] Muscas C, Pau M, Pegoraro PA, Sulis S. "Effects of measurements and pseudo-measurements correlation in distribution system state estimation". *IEEE Transactions on Instrumentation and Measurement*. 2014;63(12):2813–2823.

[31] Boyd S, Vandenberghe L. *Convex Optimization*. Cambridge: Cambridge University Press; 2004.
[32] Kalman RE. "A new approach to linear filtering and prediction problems". *Transactions of the ASME-Journal of Basic Engineering*. 1960;82(1):33–45.
[33] Gelb A, Kasper JF, Nash RA, Price CF, Sutherland AA. *Applied Optimal Estimation*. Gelbe A, editor. Cambridge, MA: The MIT Press; 2001.
[34] Ghahremani E, Kamwa I. "Dynamic state estimation in power system by applying the extended Kalman filter with unknown inputs to phasor measurements". *IEEE Transactions on Power Systems*. 2011 Nov;26(4):2556–2566.
[35] Huang Z, Schneider K, Nieplocha J. "Feasibility studies of applying Kalman filter techniques to power system dynamic state estimation". In: *Inter. Power Eng. Conf. 2007 (IPEC 2007)*, Singapore; 2007. p. 376–382.
[36] Sarri S, Paolone M, Cherkaoui R, Borghetti A, Napolitano F, Nucci CA. "State estimation of active distribution networks: comparison between WLS and iterated Kalman-filter algorithm integrating PMUs". In: *Third Conf. on Innov. Smart Grid Techn. (ISGT Europe 2012)*. Berlin, Germany. IEEE; 2012.
[37] Jain A, Shivakumar NR. "Power system tracking and dynamic state estimation". In: *Power Syst. Conf. and Expos. (PSCE), 2009*. Seattle, WA. IEEE; 2009.
[38] Zhang J, Welch G, Bishop G, Huang Z. "Reduced measurement–space dynamic state estimation (ReMeDySE) for power systems". In: *2011 IEEE PowerTech*. Trondheim, Norway; 2011.
[39] Lerro D, Bar-Shalom Y. "Tracking with debiased consistent converted measurements versus EKF". *IEEE Transactions on Aerospace and Electronics Systems*. 1993 Jul;29(3):1015–1022.
[40] Göl M, Abur A. "LAV based robust state estimation for systems measured by PMUs". *IEEE Transactions on Smart Grid*. 2014 Jul;5(4):1808–1814.
[41] Jones KD, Pal A, Thorp JS. "Methodology for performing synchrophasor data conditioning and validation". *IEEE Transactions on Power Systems*. 2015 May;30(3):1121–1130.
[42] Pignati M, Zanni L, Sarri S, Cherkaoui R, Boudec JYL, Paolone M. "A pre-estimation filtering process of bad data for linear power systems state estimators using PMUs". In: *Power Syst. Comp. Conf. (PSCC)*. Wrocław, Poland; 2014.
[43] Zhu J, Abur A. "Identification of network parameter errors". *IEEE Transactions on Power Systems*. 2006 May;21(2):586–592.
[44] Monticelli A, Garcia A. "Reliable bad data processing for real-time state estimation". *IEEE Transactions on Power Apparatus and Systems*. 1983 May;PAS-102(5):1126–1139.
[45] Van Cutsem T, Ribbens-Pavella M, Mili L. "Hypothesis testing identification: a new method for bad data analysis in power system state estimation". *IEEE Transactions on Power Apparatus and Systems*. 1984 Nov;PAS-103(11): 3239–3252.
[46] Falcao DM, Cooke PA, Brameller A. "Power system tracking state estimation and bad data processing". *IEEE Transactions on Power Apparatus and Systems*. 1982 Feb;PAS-101(2):325–333.

[47] Leite da Silva AM, Filho MBDC, Cantera JMC. "An efficient dynamic state estimation algorithm including bad data processing". *IEEE Power Engineering Review*. 1987 Nov;PER-7(11):49–49.

[48] Krishnan V. *Nonlinear Filtering and Smoothing: An Introduction to Martingales, Stochastic Integrals, and Estimation*. Mineola, NY: Courier Dover Publications; 2005.

[49] Group IDPW. "Radial distribution test feeders". *IEEE Transactions on Power Systems*. 1991;6:975–985.

[50] *Instrument Transformers: Additional Requirements for Electronic Current Transformers*. Standard IEC 61869-7; 2014.

[51] *Instrument Transformers: Additional Requirements for Electronic Voltage Transformers*. Standard IEC 61869-8; 2011.

[52] Athay T, Podmore R, Virmani S. "A practical method for the direct analysis of transient stability". *IEEE Transactions on Power Apparatus and Systems*. 1979 Mar;PAS-98(2):573–584.

[53] Le Boudec JY. *Performance Evaluation of Computer and Communication Systems*. Lausanne: EPFL Press; 2010.