Electrical and Electronics Engineering
2024-2025
Master Semester 2

Course
Smart grids technologies

Power Systems state estimation

Prof. Mario Paolone
Distributed Electrical Systems Laboratory
École Polytechnique Fédérale de Lausanne (Switzerland)

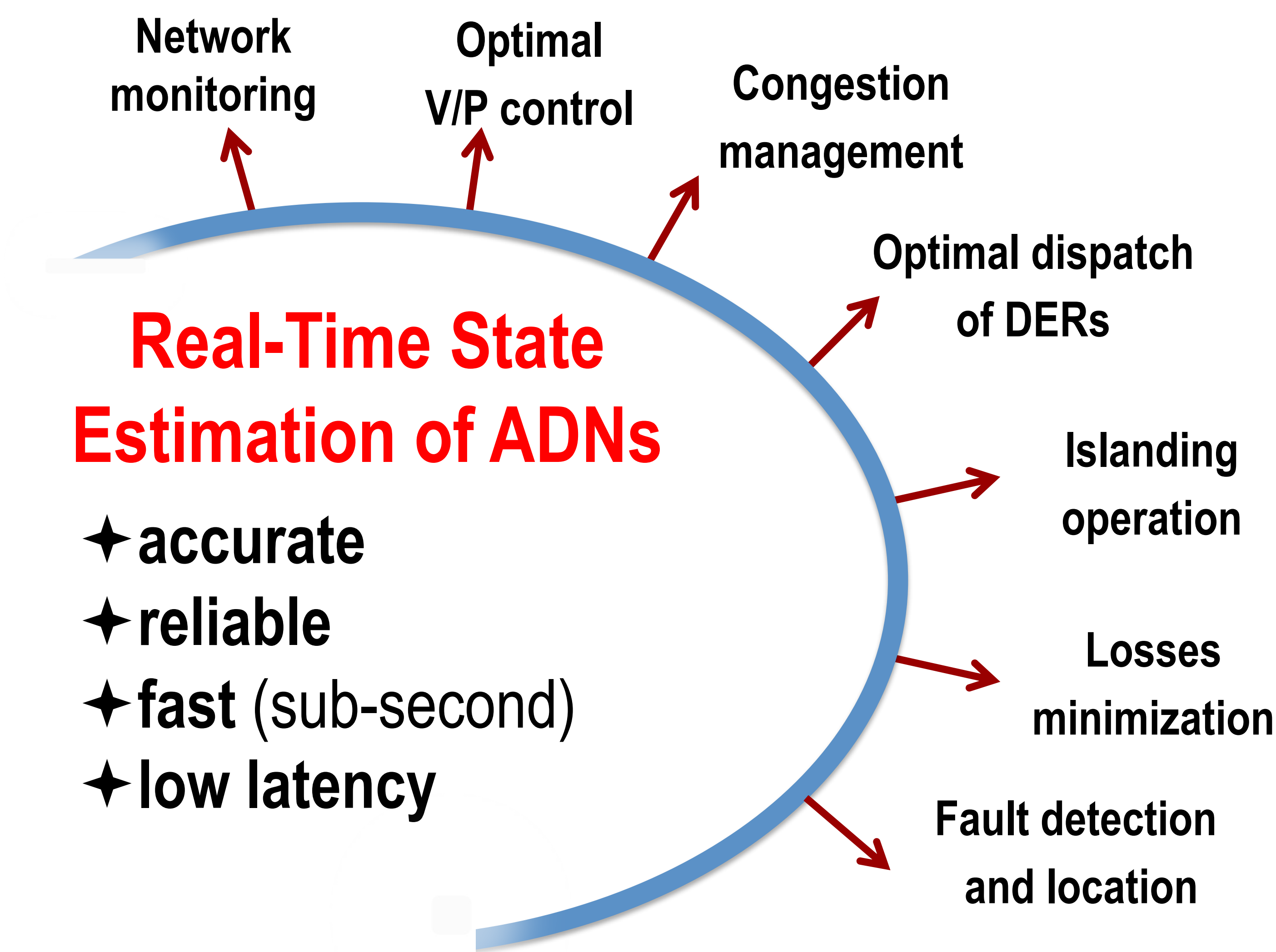
Definition

To fix the ideas, in what follows with the term

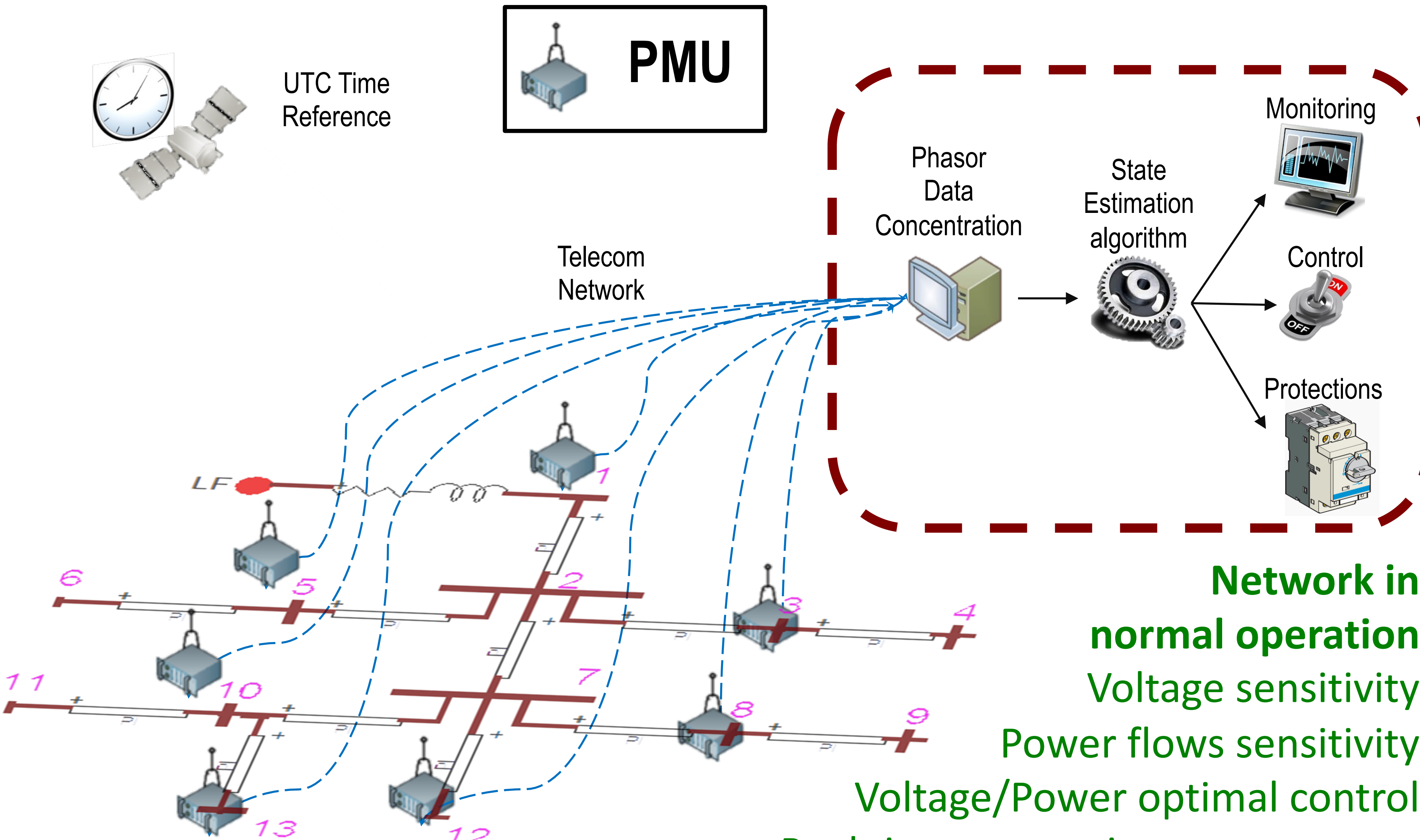
Real-Time State Estimation – RTSE

we make reference to the process of **estimating the network state** (i.e., **phase-to-ground node voltages**) with an **extremely high refreshing rate** (typically of **several tens of frames per second**) enabled by the use of **synchrophasor measurements**.

Potential applications of RTSE in ADNs



Potential applications of RTSE in ADNs



Network in normal operation

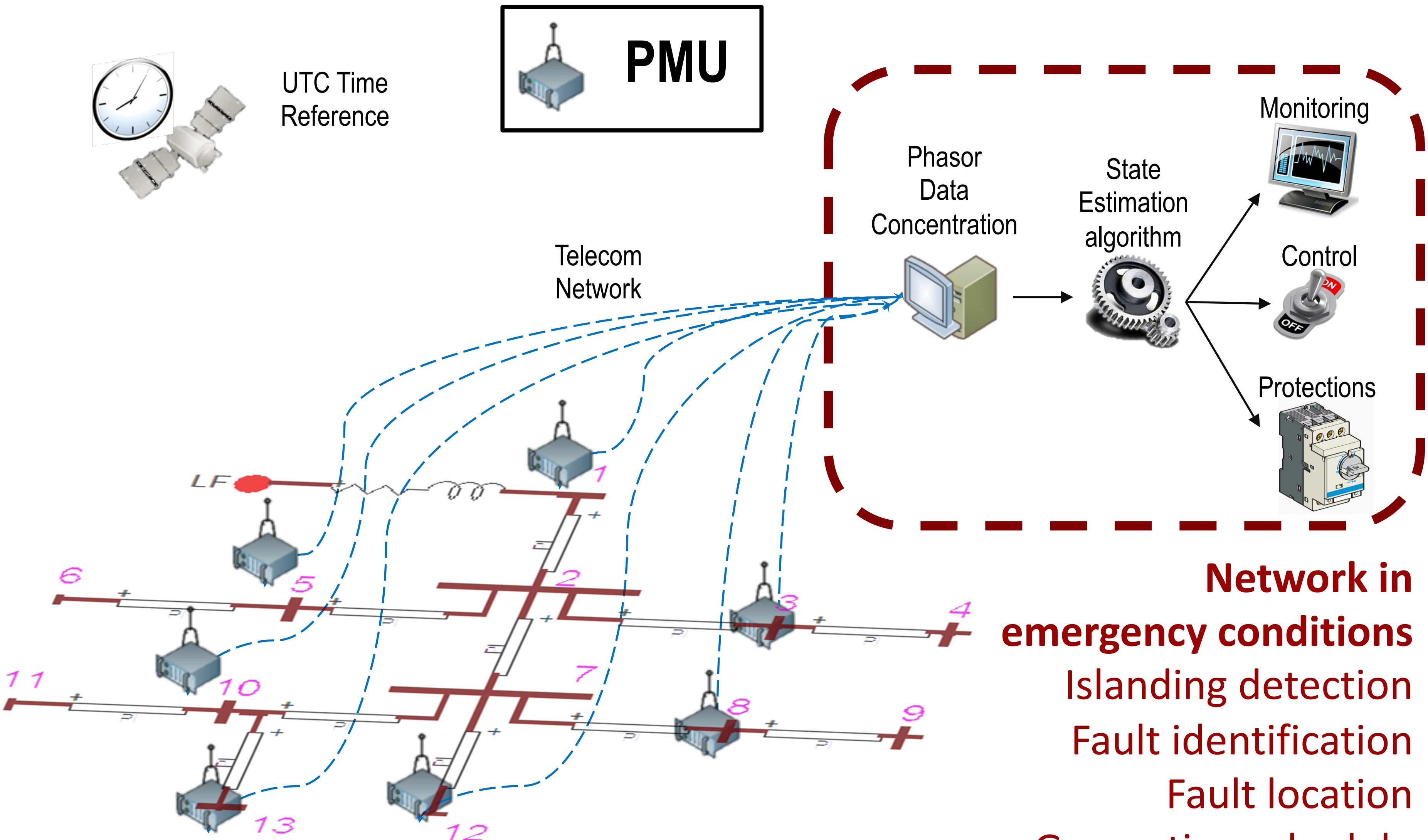
Voltage sensitivity

Power flows sensitivity

Voltage/Power optimal control

Real time congestion management

Potential applications of RTSE in ADNs



Network in emergency conditions

Islanding detection

Fault identification

Fault location

Generation schedule

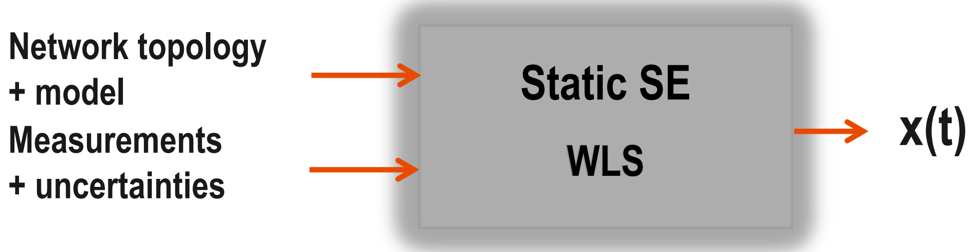
Introduction to the SE algorithms

One of the challenging tasks related to the real-time control of **modern Power Systems** → development of **fast** (i.e. sub-second) State Estimation (SE) processes → **major advantages associated to the use of Phasor Measurements Units**.

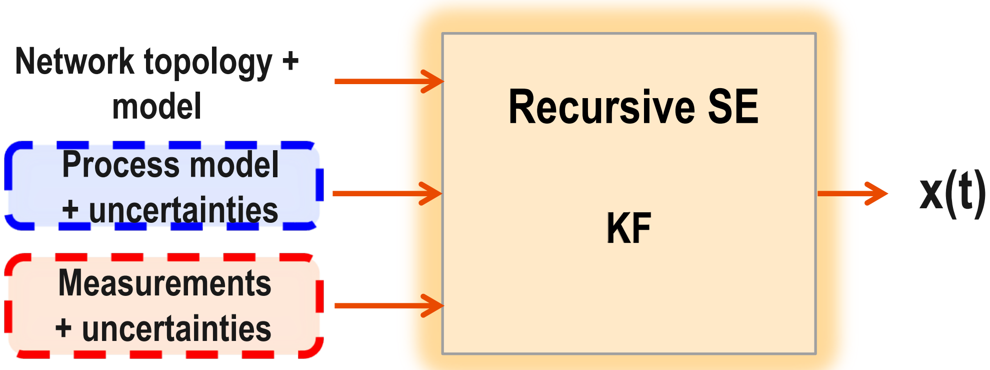
- **Delays**:
  - Synchrophasor estimation (finite window length to infer a measurement)
  - Telecom
  - Measurement concentration and data retrieve from a Data Base
  - State Estimation algorithm itself.
- The calculation of the system states is accomplished by **solving a minimization problem** by using, for instance:
- **Static algorithms (i.e. based on Weighted Least Squares (WLS), or**
- **Recursive algorithms (i.e. based on Kalman Filter (KF) methods).**
- In a first step, we consider the case of **balanced** networks. Therefore, we make reference to the **direct sequence only**. In the second half of the lecture, we will consider **unbalanced systems**.

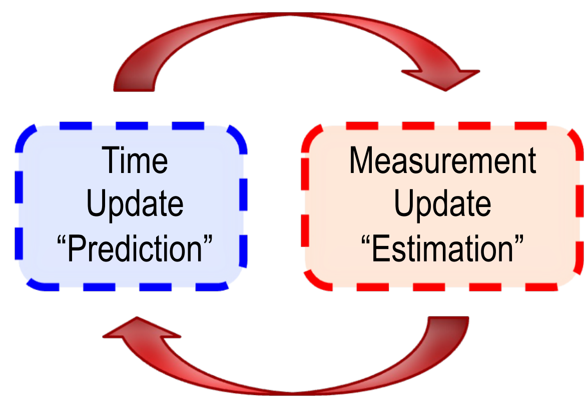
Introduction to the SE algorithms

- **Static SE:** infers the system state by using only current time information (e.g., **Weighted Least Squares – WLS – method**).



- **Recursive SE:** takes into account information available from previous time steps and predict the state vector in time (e.g., **Kalman Filter – KF – method**).





The Weighted Least Squares (WLS) method

Static SE refers to the procedure of inferring the **system state**, given by the **phase-to-ground voltage phasors at all system buses** at a **given point in time** and is expressed by the **Weighted Least Squares (WLS) method**.

In this first development, we will approach the SE problem from its basics avoiding a too-compact formalism.

For a system that has s buses, the system state vector has $(2s\text{-}1)$ elements, namely s bus voltage magnitudes V and $(s\text{-}1)$ voltage phase angles δ. **The angle of the slack bus is chosen to be the reference angle** and is set to a fixed value, in general, equal to 0. Therefore, the network state vector $\mathbf{x} \in \mathbb{R}^{2s-1}$ is a follows:

$$\mathbf{x} = \left[\delta_2, \ldots, \delta_s, V_1, \ldots, V_s\right]^T \qquad (1)$$

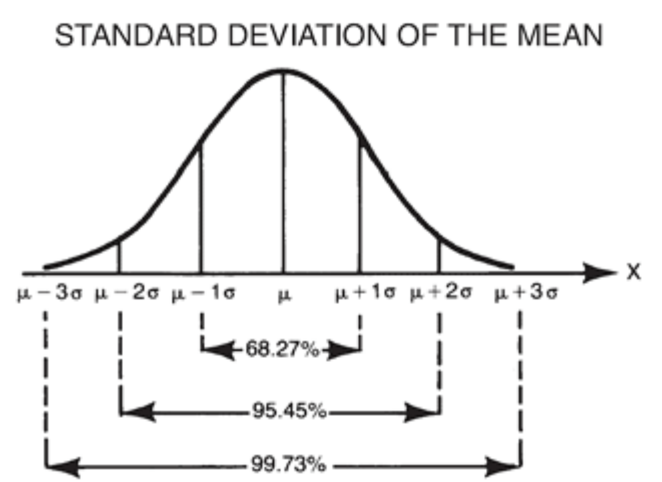
The Weighted Least Squares (WLS) method

Theoretical background:

The **main goal** of SE is to **compute the most likely system state**, based on **some measured quantities**. A way to do this is by using the **maximum likelihood estimation (MLE)** method, where the measurement errors are assumed to have **a known probability distribution**.

When the system state is chosen so that it is closest to the real one, the likelihood function attains its peak value. Therefore, an optimization problem must be solved, and the solution provides the maximum likelihood estimates for the system state.

In WLS-type state estimators, the **measurement errors are assumed to have a Gaussian (normal) distribution**. The parameters that are used are the mean, μ and the variance σ^2.



The Weighted Least Squares (WLS) method

We assume that the **normal probability density function** (p.d.f.) of a generic measurement z_i is defined as:

$$f(z_i) = \frac{1}{\sqrt{2\pi\sigma_i}} e^{-\frac{1}{2}\left(\frac{z_i-\mu_i}{\sigma_i}\right)^2} \triangleq N(\mu_i, \sigma_i^2) \qquad (2).$$

The **joint p.d.f.** $f_m(\mathbf{z})$ **is expressed as the product of the individual probability density functions**, given that **each measurement is assumed to be independent of the others**. All the measurements are assumed to have a Gaussian-type p.d.f.

$$f_m(\mathbf{z}) = f(z_1) f(z_2) \ldots f(z_m) \qquad (3)$$

Where z_i is the i^{th} measurement, m is the total number of measurements and:

$$\mathbf{z}^T = \left[z_{1,} z_2, \ldots, z_m \right] \qquad (4).$$

The Weighted Least Squares (WLS) method

The function $f_m(\mathbf{z})$ **expresses the probability of observing the specific set of measurements** in the measurements array $\mathbf{z}$. After **linking** $f_m(\mathbf{z})$ **with the system state**, the objective of MLE is to **maximize** $f_m(\mathbf{z})$ **by varying the unknown system state**. The p.d.f. can be replaced by its logarithm, as in this way the optimization procedure has a convex objective function. The so-called **Log-Likelihood Function** L is given by:

$$L = \log f_m(\mathbf{z}) = \sum_{i=1}^{m} \log f(z_i) = -\frac{1}{2}\sum_{i=1}^{m}\left(\frac{z_i-\mu_i}{\sigma_i}\right)^2 - \frac{m}{2}\log 2\pi - \sum_{i=1}^{m}\log\sigma_i \qquad (5).$$

MLE will maximize function L **for a given set of measurements** $z_1, z_2, \ldots, z_m$:

$$\underset{\mathbf{x}}{\text{maximize}} \quad \log f_m(\mathbf{z}) \qquad \text{OR} \qquad \underset{\mathbf{x}}{\text{minimize}} \quad \sum_{i=1}^{m}\left(\frac{z_i-\mu_i}{\sigma_i}\right)^2 \qquad (6).$$

The Weighted Least Squares (WLS) method

Equation (6) **does not contain explicitly the system state**. In order to express this link, **it can be formulated as a function of the residual r_i of measurement** i, which is defined as:

$$r_i = z_i - \mu_i \qquad (7).$$

The mean μ_i can be expressed as $h_i(\mathbf{x})$: a **non-linear function relating the system state vector $\mathbf{x}$** to the i^{th} measurement (in what follow $h_i(\mathbf{x})$ will be called **measurement function**). The square of each residual r_i^2 is weighted by $W_{ii} = \sigma_i^{-2}$ and, as a consequence, equation (6) can be re-written as follows:

$$\begin{aligned} &\underset{\mathbf{x}}{\text{minimize}} \quad \sum_{i=1}^{m} W_{ii} r_i^2 \\ &\text{subject to } \; z_i = h_i(\mathbf{x}) + r_i, \quad i = 1, \ldots, m \end{aligned} \qquad (8).$$

The Weighted Least Squares (WLS) method

By solving the above problem, the WLS estimator for $\mathbf{x}$ can be obtained. The WLS estimator will minimize the following objective function:

$$J(\mathbf{x}) = \sum_{i=1}^{m} \frac{\left(z_i - h_i(\mathbf{x})\right)^2}{R_{ii}} \qquad (9)$$

where $\mathbf{R} = diag\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2\right)$ is the so-called **measurement noise covariance matrix**.

In order to clarify the meaning and the role of the **measurement function** $h(\mathbf{x})$ in SE, in the following slides the formulation of $h(\mathbf{x})$ is given for the case where the measurements consist of **power injections, line power flows, line current flow magnitudes and bus voltage phasors**.

The next task is to write the measurements as a function of the system state.

The Weighted Least Squares (WLS) method – The non-linear case

The formal approach and the algorithm

Let's assume that at a given point in time defined by the time-step index t, the set of measurements $\mathbf{z}$ is linked to the system state $\mathbf{x}$ by means of a nonlinear function h:

$$\mathbf{z}_t = h(\mathbf{x}_t) + \mathbf{v}_t$$

where $\mathbf{v}$ is the measurement noise, assumed to be white and with a normal probability distribution of covariance $\mathbf{R}$.

The aim of the WLS estimator is the minimization of the following objective function:

$$J_t = \left[\mathbf{z}_t - h(\mathbf{x}_t)\right]^T \mathbf{R}_t^{-1} \left[\mathbf{z}_t - h(\mathbf{x}_t)\right]$$

If $\mathbf{R} = diag\left(\sigma_1^2, \sigma_2^2, ..., \sigma_m^2\right)$:

$$J_t = \sum_{i=1}^{m} \frac{\left(z_{t,i} - h_{t,i}(\mathbf{x})\right)^2}{R_{t,ii}}$$

The Weighted Least Squares (WLS) method – The non-linear case

At the minimum, **the first-order optimality conditions will have to be satisfied**. These can be expressed in compact form as follows:

$$\mathbf{g}\left(\hat{\mathbf{x}}_t\right) = \left.\frac{\partial J\left(\mathbf{x}_t\right)}{\partial \mathbf{x}_t}\right|_{\hat{\mathbf{x}}} = \mathbf{H}^T\left(\hat{\mathbf{x}}_t\right)\mathbf{R}_t^{-1}\left[\mathbf{z}_t - h\left(\hat{\mathbf{x}}_t\right)\right] = 0$$

where: $\mathbf{H}\left(\hat{\mathbf{x}}_t\right) = \left.\frac{\partial h\left(\mathbf{x}_t\right)}{\partial \mathbf{x}_t}\right|_{\hat{\mathbf{x}}_t}$

We may expand the non-linear function $\mathbf{g}(\boldsymbol{x})$ (that must be null in view of the above) into its Taylor series around the state vector:

$$\mathbf{g}(\boldsymbol{x}) = \mathbf{g}(\hat{\boldsymbol{x}}_t) + \mathbf{G}(\hat{\boldsymbol{x}}_t)(\boldsymbol{x} - \hat{\boldsymbol{x}}_t) + \cdots = \mathbf{0}$$

where:

$$\mathbf{G}(\hat{\boldsymbol{x}}_t) = \left[\frac{\partial \boldsymbol{g}(\hat{\boldsymbol{x}}_t)}{\partial \hat{\boldsymbol{x}}_t}\right] = \mathbf{H}^T(\hat{\boldsymbol{x}}_t)\mathbf{R}_t^{-1}\mathbf{H}(\hat{\boldsymbol{x}}_t)$$

$\mathbf{G}$ is called the so-called gain matrix.

By combining the previous expressions we get the following iterative process (where k indicates the generic iteration of the process):

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

$$\hat{\mathbf{x}}_t^{k+1} = \hat{\mathbf{x}}_t^k + \left[\mathbf{G}(\hat{\mathbf{x}}_t^k)\right]^{-1}\mathbf{H}^T\left(\hat{\mathbf{x}}_t^k\right)\mathbf{R}_t^{-1}\left[\mathbf{z}_t - h\left(\hat{\mathbf{x}}_t^k\right)\right]$$

The Weighted Least Squares (WLS) method – The non-linear case

Let consider a **balanced** electrical network, let define $(\mathbf{Z}_{i\ell})^{-1}$ as the generic element of the network admittance martix:

$$(\mathbf{Z}_{i\ell})^{-1} = G_{i\ell} + jB_{i\ell} \qquad (10).$$

We remind that the active and reactive power injection, P_{inj} and Q_{inj} equations of $h(\mathbf{x})$ at bus i can be inferred from the load flow problem:

$$\begin{aligned} P_i &= V_i^2 G_{ii} + V_i \sum_{\substack{\ell=1 \\ \ell \neq i}}^{s} V_\ell (G_{i\ell} \cos\delta_{i\ell} + B_{i\ell} \sin\delta_{i\ell}) \\ Q_i &= -V_i^2 B_{ii} + V_i \sum_{\substack{\ell=1 \\ \ell \neq i}}^{s} V_\ell (G_{i\ell} \sin\delta_{i\ell} - B_{i\ell} \cos\delta_{i\ell}) \end{aligned} \qquad (11)$$

where $\delta_{i\ell} = \delta_i - \delta_\ell$ is the angle difference between voltage phasors of buses i and ℓ.

The Weighted Least Squares (WLS) method – The non-linear case

We also remind the expressions of active and reactive power flows P_{flow}, Q_{flow} from bus i to bus ℓ are:

$$
\begin{aligned}
P_{i\ell} &= V_i^2(g_{si} + g_{i\ell}) - V_i V_\ell (g_{i\ell}\cos\delta_{i\ell} + b_{i\ell}\sin\delta_{i\ell}) \\
Q_{i\ell} &= -V_i^2(b_{si} + b_{i\ell}) - V_i V_\ell (g_{i\ell}\sin\delta_{i\ell} - b_{i\ell}\cos\delta_{i\ell})
\end{aligned}
\qquad (12)
$$

where $g_{i\ell}+jb_{i\ell}$ is the admittance of the series branch composing the π-equivalent line connecting buses i and ℓ and $g_{si}+jb_{si}$ is the admittance of the shunt branch connected to bus i.

The line current flow magnitude I_{magn} from bus i to bus ℓ is simply:

$$
I_{i\ell} = \frac{\sqrt{P_{i\ell}^2 + Q_{i\ell}^2}}{V_i} \qquad (13).
$$

The Weighted Least Squares (WLS) method – The non-linear case

Let us assume that the measurements are:

$$\mathbf{z} = [P_{inj}, P_{flow}, Q_{inj}, Q_{flow}, , I_{mah}, V_{mag}, , \delta]$$

It is worth observing the inherent non-linear nature of $h(\mathbf{x})$.

Therefore, In order to re-formulate the optimal problem stated by (6) or (8) as an **iterative convex optimisation problem**, there is the need of linearizing $h(\mathbf{x})$. The linearized version of $h(\mathbf{x})$ is indicated as $\mathbf{H}$.

$$\mathbf{H} = \begin{bmatrix} \frac{\partial P_{inj}}{\partial \delta} & \frac{\partial P_{inj}}{\partial V} \\ \frac{\partial P_{flow}}{\partial \delta} & \frac{\partial P_{flow}}{\partial V} \\ \frac{\partial Q_{inj}}{\partial \delta} & \frac{\partial Q_{inj}}{\partial V} \\ \frac{\partial Q_{flow}}{\partial \delta} & \frac{\partial Q_{flow}}{\partial V} \\ \frac{\partial I_{magn}}{\partial \delta} & \frac{\partial I_{magn}}{\partial V} \\ \frac{\partial V_{magn}}{\partial \delta} & \frac{\partial V_{magn}}{\partial V} \\ \frac{\partial \delta}{\partial \delta} & \frac{\partial \delta}{\partial V} \end{bmatrix} \quad (14).$$

The Weighted Least Squares (WLS) method – The non-linear case

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

$$\frac{\partial P_i}{\partial \delta_i} = \sum_{\ell=1}^{s} V_i V_\ell (-G_{i\ell} \sin\delta_{i\ell} + B_{i\ell} \cos\delta_{i\ell}) - V_i^2 B_{ii}$$

$$\frac{\partial P_i}{\partial \delta_\ell} = V_i V_\ell (G_{i\ell} \sin\delta_{i\ell} - B_{i\ell} \cos\delta_{i\ell})$$

$$\frac{\partial P_i}{\partial V_i} = \sum_{\ell=1}^{s} V_\ell (G_{i\ell} \cos\delta_{i\ell} + B_{i\ell} \sin\delta_{i\ell}) + V_i G_{ii}$$

$$\frac{\partial P_i}{\partial V_\ell} = V_i (G_{i\ell} \cos\delta_{i\ell} + B_{i\ell} \sin\delta_{i\ell})$$

(15).

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

$$\frac{\partial Q_i}{\partial \delta_i} = \sum_{\ell=1}^{s} V_i V_\ell (G_{i\ell} \cos\delta_{i\ell} + B_{i\ell} \sin\delta_{i\ell}) - V_i^2 G_{ii}$$

$$\frac{\partial Q_i}{\partial \delta_\ell} = V_i V_\ell (-G_{i\ell} \cos\delta_{i\ell} - B_{i\ell} \sin\delta_{i\ell})$$

$$\frac{\partial Q_i}{\partial V_i} = \sum_{\ell=1}^{s} V_\ell (G_{i\ell} \sin\delta_{i\ell} - B_{i\ell} \cos\delta_{i\ell}) - V_i B_{ii}$$

$$\frac{\partial Q_i}{\partial V_\ell} = V_i (G_{i\ell} \sin\delta_{i\ell} - B_{i\ell} \cos\delta_{i\ell})$$

(16).

The Weighted Least Squares (WLS) method – The non-linear case

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

$$\frac{\partial P_{i\ell}}{\partial \delta_i} = V_i V_\ell (g_{i\ell} \sin\delta_{i\ell} - b_{i\ell} \cos\delta_{i\ell})$$

$$\frac{\partial P_{i\ell}}{\partial \delta_j} = -V_i V_\ell (g_{i\ell} \sin\delta_{i\ell} - b_{i\ell} \cos\delta_{i\ell}) \qquad (17).$$

$$\frac{\partial P_{i\ell}}{\partial V_i} = -V_\ell (g_{i\ell} \cos\delta_{i\ell} + b_{i\ell} \sin\delta_{i\ell}) + 2(g_{i\ell} + g_{si}) V_i$$

$$\frac{\partial P_{i\ell}}{\partial V_\ell} = -V_i (g_{i\ell} \cos\delta_{i\ell} + b_{i\ell} \sin\delta_{i\ell})$$

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

$$\frac{\partial Q_{i\ell}}{\partial \delta_i} = -V_i V_\ell (g_{i\ell} \cos\delta_{i\ell} + b_{i\ell} \sin\delta_{i\ell})$$

$$\frac{\partial Q_{i\ell}}{\partial \delta_\ell} = V_i V_\ell (g_{i\ell} \cos\delta_{i\ell} + b_{i\ell} \sin\delta_{i\ell}) \qquad (18).$$

$$\frac{\partial Q_{i\ell}}{\partial V_i} = -V_\ell (g_{i\ell} \sin\delta_{i\ell} - b_{i\ell} \cos\delta_{i\ell}) - 2(b_{i\ell} + b_{si}) V_i$$

$$\frac{\partial Q_{i\ell}}{\partial V_\ell} = -V_i (g_{i\ell} \sin\delta_{i\ell} - b_{i\ell} \cos\delta_{i\ell})$$

The Weighted Least Squares (WLS) method – The non-linear case

The partial derivatives that correspond to the current magnitudes (if the shunt admittance of the branch is ignored) are:

$$\frac{\partial I_{i\ell}}{\partial \delta_i} = \frac{g_{i\ell}^2 + b_{i\ell}^2}{I_{i\ell}} V_i V_\ell \sin \delta_{i\ell}$$

$$\frac{\partial I_{i\ell}}{\partial \delta_\ell} = -\frac{g_{i\ell}^2 + b_{i\ell}^2}{I_{i\ell}} V_i V_\ell \sin \delta_{i\ell} \qquad (19).$$

$$\frac{\partial I_{i\ell}}{\partial V_i} = \frac{g_{i\ell}^2 + b_{i\ell}^2}{I_{i\ell}} (V_i - V_\ell \cos \delta_{i\ell})$$

$$\frac{\partial I_{i\ell}}{\partial V_\ell} = \frac{g_{i\ell}^2 + b_{i\ell}^2}{I_{i\ell}} (V_\ell - V_i \cos \delta_{i\ell})$$

Finally, the partial derivatives that correspond to the voltage magnitudes and the voltage phases provided by PMUs are:

$$\frac{\partial V_i}{\partial \delta_i} = 0 \quad \frac{\partial V_i}{\partial \delta_\ell} = 0 \quad \frac{\partial V_i}{\partial V_i} = 1 \quad \frac{\partial V_i}{\partial V_\ell} = 0$$

$$\frac{\partial \delta_i}{\partial \delta_i} = 1 \quad \frac{\partial \delta_i}{\partial \delta_\ell} = 0 \quad \frac{\partial \delta_i}{\partial V_i} = 0 \quad \frac{\partial \delta_i}{\partial V_\ell} = 0 \qquad (20).$$

The Weighted Least Squares (WLS) method – The non-linear case

The iterative algorithm for the non-linear case

1. Initialize the state vector $\mathbf{x}^{\mathbf{0}}$, typically as a **'flat-start'** (all bus voltages are assumed to be 1 per unit (pu) and in phase with each other);

Iteration loop (index k)

2. Calculate the nonlinear function $h(\mathbf{x}^k)$ and the matrix $\mathbf{H}\left(\hat{\mathbf{x}}_t^k\right)$;
3. Calculate the so-called "Gain matrix" $\mathbf{G}(\hat{\mathbf{x}}_t^k)$ and the function $\mathbf{g}\left(\hat{\mathbf{x}}_t^k\right)$;
4. Calculate $\hat{\mathbf{x}}_t^{k+1} = \hat{\mathbf{x}}_t^k + \left[\mathbf{G}(\hat{\mathbf{x}}_t^k)\right]^{-1} \mathbf{H}^T\left(\hat{\mathbf{x}}_t^k\right)\mathbf{R}_t^{-1}\left[\mathbf{z}_t - h\left(\hat{\mathbf{x}}_t^k\right)\right]$;
5. Calculate $J\left(\hat{\mathbf{x}}_t^k\right)$ and **stop** if the following conditions are satisfied:
   - Condition 1: $\max\left|\hat{\mathbf{x}}_t^{k+1} - \hat{\mathbf{x}}_t^k\right| \le \varepsilon_1$
   - Condition 2: $\left|J\left(\hat{\mathbf{x}}_t^{k+1}\right) - J\left(\hat{\mathbf{x}}_t^k\right)\right| < \varepsilon_2$
   - Condition 3: $J\left(\hat{\mathbf{x}}_t^{k+1}\right) < \varepsilon_3$

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

The Weighted Least Squares (WLS) method – The linear case

Let us now suppose that the measurements consist of **phasors of bus phase-to-ground voltages** and **phasors of nodal current injections**.

Of course, these measurements are provided by PMUs and are, also, **synchronous (i.e., they are time-tagged using the UTC time reference)**.

The system state is always given by the equation (1) but, in order to simplify the problem, we rewrite the system state in **rectangular coordinates**:

$$\mathbf{x} = \left[V_{1,re},\dots,V_{s,re},,V_{1,im},\dots,V_{s,im}\right]^T \quad (21).$$

Note that, as for the non-linear case, **if the slack bus is assumed to be the reference, it corresponds to have**

$$V_{1,im} = 0$$

The Weighted Least Squares (WLS) method – The linear case

We assume that the measurements come only from PMUs. Therefore, the measurement set is composed of:

- d_1 phase-to-ground voltage phasors
- d_2 nodal-injected current phasors.

We also assume that the $d_1+d_2 \geq s$ do that the network is **observable.**

Note that the concept of observability has not been defined.

In this case, the set of measurements is a $m = 2d_1+2d_2$ array:

$$\mathbf{z}^T = \left[\mathbf{z}_V, \mathbf{z}_I\right] \tag{22}$$

where:

$$\begin{aligned}\mathbf{z}_V &= \left[V_{1,re},\ldots,V_{d_1,re},V_{1,im},\ldots,V_{d_1,im}\right] \\ \mathbf{z}_I &= \left[I_{1,re},\ldots,I_{d_2,re},I_{1,im},\ldots,I_{d_2,im}\right]\end{aligned} \quad (23).$$

The Weighted Least Squares (WLS) method – The linear case

Let us now see a more compact form of the SE problem. The equation linking the measurements with the system state can be also written as:

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{v} \tag{24}$$

where:

- $\mathbf{H}$ is a m x $2s$ matrix representing the measurement Jacobian (*) which connects the state with the measurements **for the case of null measurement noise**;
- $\mathbf{v}$ is the measurement noise.

(*) Observation: in this case, it might be improper to call this matrix a 'Jacobian' since it is a **constant and state-independent matrix**.

The Weighted Least Squares (WLS) method – The linear case

Also in this case, we assume that the measurement noise is **white and Gaussian** and the **measurement errors are independent**. So, we get:

$$p(\mathbf{v}) \sim N(0, \mathbf{R}) \qquad (25)$$

$$\mathbf{R} = diag\left(\sigma_1^2, \ldots, \sigma_m^2\right) \qquad (26)$$

where $\mathbf{R}$ is the so-called **measurement noise covariance matrix** and σ_i $(i = 1, \ldots, m)$ is the standard deviation of the i^{th} measurement. Therefore, $\mathbf{R}$ represents the accuracies of the measurement devices.

The Weighted Least Squares (WLS) method – The linear case

<u>IMPORTANT OBSERVATION</u>: the real and imaginary part of the nodal current injections are linked to the system state directly via the admittance matrix:

$$I_{i,re} = \sum_{\ell=1}^{s} \left(G_{i\ell} V_{\ell,re} - B_{i\ell} V_{\ell,im} \right)$$

$$I_{i,im} = \sum_{\ell=1}^{s} \left(G_{i\ell} V_{\ell,im} + B_{i\ell} V_{\ell,re} \right) \qquad (27)$$

where:

- i is the bus index;
- $G_{i\ell}$ and $B_{i\ell}$ are the real and imaginary parts of the $i\ell$ admittance matrix elements, respectively.

The Weighted Least Squares (WLS) method – The linear case

In this case, the measurements Jacobian $\mathbf{H}$ is equal to:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H_V} \\ \mathbf{H_I} \end{bmatrix} \tag{28}$$

where:

- $\mathbf{H_V}$ is is the part of the Jacobian (*) that is related to the **partial derivatives of the real and imaginary part of the voltages as a function of the state**;
- $\mathbf{H_I}$ is the part of the Jacobian (*) that is related to the **partial derivatives of the real and imaginary part of the injected currents as a function of the state**.

(*) Observation: as said before, it might be improper to call this matrix a 'Jacobian' since it is a **constant and state-independent matrix**.

The Weighted Least Squares (WLS) method – The linear case

Therefore, we have for $\mathbf{H_V}$:

$$\mathbf{H_V} = \begin{bmatrix} [\beta] & [\upsilon] \\ [\zeta] & [\eta] \end{bmatrix} \qquad (29)$$

$$where \begin{cases} \beta = \begin{cases} 1 \ if \ \ i = \ell \\ 0 \ if \ \ i \neq \ell \end{cases} \\ \upsilon = \zeta = 0 \\ \eta = \begin{cases} 1 \ if \ \ i = \ell \\ 0 \ if \ \ i \neq \ell \end{cases} \end{cases} \qquad (30).$$

The Weighted Least Squares (WLS) method – The linear case

And for $\mathbf{H_I}$:

$$\mathbf{H_I} = \begin{bmatrix} [G_{i\ell}] & [-B_{i\ell}] \\ [B_{i\ell}] & [G_{i\ell}] \end{bmatrix} \qquad (31).$$

The Weighted Least Squares (WLS) method – The linear case

As it is evident, the link between the system state and the measurements is **linear**.

As a consequence, the SE process refers to the **minimization of the following quadratic objective function**:

$$J(\mathbf{x}) = \sum_{i=1}^{m} \frac{\left(z_i - \sum_{j=1}^{2s} H_{ij} x_j \right)^2}{R_{ii}} \quad (32).$$

The minimum of this objective, as a function of the state $\mathbf{x}$, can be analytically computed as:

$$\hat{\mathbf{x}} = \mathbf{G}^{-1}\mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{z}$$
$$\textit{with} \quad (33).$$
$$\mathbf{G} = \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H}$$

The Weighted Least Squares (WLS) method – Bad Data Processing

Bad data issue

Measurements can contain errors due to instrumentation malfunction, incorrect sensor compensation, telecommunication system failures, user misinterpretations…

The presence of **erroneous measurements** can be **detected** by analyzing the normalized measurement estimation residual vector.

The residuals $\mathbf{r}$ are expected to be Gaussian distributed: $r_{\mathrm{i}} \sim \mathcal{N}(0, \Omega_{\mathrm{ii}})$, with:

$$\mathbf{r} = \mathbf{z} - \mathbf{H}\hat{\mathbf{x}}$$
$$\boldsymbol{\Omega} = \mathrm{cov}(\mathbf{r}) = \mathbf{R} - \mathbf{H}\mathbf{G}^{-1}\mathbf{H}^{\mathrm{T}}$$

The Weighted Least Squares (WLS) method – Bad Data Processing

Largest normalized residual (LNR) test

1. Solve the WLS problem and compute the estimation residuals:

$$r_{\mathrm{i}} = z_{\mathrm{i}} - \mathbf{H}\hat{x}_{\mathrm{i}} \qquad \mathrm{i} = 1,..,\mathrm{m}$$

2. Compute the normalized residuals:

$$r_{\mathrm{i}}^{\mathrm{N}} = \frac{|\mathrm{r_i}|}{\sqrt{\mathbf{\Omega}_{\mathrm{ii}}}} \qquad \mathrm{i} = 1,..,\mathrm{m}$$

3. Find the largest normalized residual:

$$r_{\mathrm{k}}^{\mathrm{N}} = \max_{\mathrm{i}} r_{\mathrm{i}}^{\mathrm{N}} \qquad \mathrm{i} = 1,..,\mathrm{m}$$

4. If $r_{\mathrm{k}}^{\mathrm{N}} > c$ the kth measurement is suspected to be bad data. c is the user defined identification threshold, e.g. 3.0

5. Eliminate the kth measurement and go to step 1

The Weighted Least Squares (WLS) method – Bad Data Processing

Strengths and limitations of the LNR test

The LNR test will perform differently depending upon the type of bad data and their configuration

- **Single bad data**: only one measurement with a large error.

  The LNR will correspond with the bad data provided that it is not critical or its removal does not create a critical measurement.

  N.B. a critical measurement is a measurement that makes the network unobservable when removed.

- **Multiple bad data**: more than one measurement with large error
  - If the residuals are **weakly correlated:** $\boldsymbol{\Omega}_{ik} \approx 0$, the bad data is non-interacting and easier to detect
  - If the residuals are **strongly correlated:** $\boldsymbol{\Omega}_{ik}$ is significantly large, we have consistent bad data that are more difficult to detect

The Kalman Filter (KF) method

Recursive SE refers to procedures aiming at obtaining the system state at a given time by **taking into account information available from both measurements and a process model**. One of the typical techniques adopted in this field is represented by the **KF method.**

There are more than one versions of the KF method. The **Discrete Kalman Filter** (DKF) is used for linear systems, whereas the **Extended Kalman Filter** (EKF) and the **Iterated Kalman Filter** (IKF) are used when the **process to be estimated and/or the measurement relationship to the process is non-linear**. Since the equations for each version are similar, let give the ones that describe the DKF.

The KF consists of a set of equations that implement a **"predictor-measurement update"** process that **minimizes the estimated error covariance - provided that some specific conditions are met**.

The Kalman Filter (KF) method

The objective is to estimate the state $\mathbf{x} \in \mathbb{R}^{2n-1}$ of a discrete-time controlled process, governed by the linear stochastic difference equation that represents the **process model**:

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_{t-1} + \mathbf{w}_{t-1} \tag{34}$$

where:

- $\mathbf{x}_t$ and $\mathbf{x}_{t-1}$ represent the state of the system in correspondence of discrete time steps t and t-1, respectively;
- $\mathbf{u}_{t-1}$ represents a set of u_c control variables (independent from the system state) of the system at time step t -1;
- $\mathbf{w}_{t-1}$ represents the system process noise assumed white and with a normal probability distribution;
- $\mathbf{A}$ is a $(2s\text{-}1)\text{x}(2s\text{-}1)$ matrix that links that state of the system at time step t-1 with the one of the current time step t for the case of null active injections and process noise;
- $\mathbf{B}$ is a $(2s\text{-}1)$ x u_c matrix that links the time evolution of the state of the system with the u_c controls at time step t-1 for the case of null process noise.

The Kalman Filter (KF) method

The measurement array $\mathbf{z} \in \mathbb{R}^m$ is given by the equation we have already seen:

$$\mathbf{z}_t = \mathbf{H}\mathbf{x}_t + \mathbf{v}_t \tag{35}$$

where:

- the $m\times(2s\text{-}1)$ matrix $\mathbf{H}$ represents the measurements Jacobian, as defined before;
- $\mathbf{v}_t$ represents the measurement noise at the same time step t; it is assumed white and with a normal probability distribution. $\mathbf{v}_t$ is also assumed independent from $\mathbf{w}_t$ (clearly, $\mathbf{v}_t \in \mathbb{R}^m$ as $\mathbf{z}_t$ does).

Similarly to the WLS, the variables $\mathbf{v}_t$ and $\mathbf{w}_t$ are assumed to have the following normal probability distributions (the mean value μ is equal to zero):

$$p(w) \sim N(0,Q)$$
$$p(v) \sim N(0,R) \tag{36}$$

The Kalman Filter (KF) method

The **process covariance** $\mathbf{Q}$ and **measurement noise covariance** $\mathbf{R}$ matrices derived from (36) are assumed to be constant. In practice, $\mathbf{A}$ might change with each time step, but here it is assumed as constant.

Hence, we can define $\tilde{\mathbf{x}}_t \in \mathbb{R}^{2s-1}$ as the **"a priori" state estimate** at step t given knowledge of the process prior to step t and $\hat{\mathbf{x}}_t \in \mathbb{R}^{2s-1}$ as the **"a posteriori" state estimate** at step t given the measurement array $\mathbf{z}_t$. The **"a priori"** and **"a posteriori"** estimate errors can be defined as

$$\tilde{\mathbf{e}}_t \equiv \mathbf{x}_t - \tilde{\mathbf{x}}_t$$
$$\hat{\mathbf{e}}_t \equiv \mathbf{x}_t - \hat{\mathbf{x}}_t \quad (37).$$

The **"a priori"** estimate error covariance is

$$\tilde{\mathbf{P}}_t \equiv \mathbb{E}[\tilde{\mathbf{e}}_t \tilde{\mathbf{e}}_t^T] \quad (38).$$

The Kalman Filter (KF) method

The "a posteriori" estimate error covariance is

$$\hat{\mathbf{P}}_t \equiv \mathbb{E}[\hat{\mathbf{e}}_t \hat{\mathbf{e}}_t^T] \qquad (39).$$

Note that in (38) and (39) $\mathbb{E}$ indicates the **expected or mean value operator**.

The next objective is to find an equation that calculates the "a posteriori" state estimate $\hat{\mathbf{x}}_t$ as a linear combination of an "a priori" estimate $\tilde{\mathbf{x}}_t$ and a weighted difference between the actual measurement array $\mathbf{z}_t$ and the measurement prediction $\mathbf{H}\tilde{\mathbf{x}}_t$.

$$\hat{\mathbf{x}}_t = \tilde{\mathbf{x}}_t + \mathbf{K}(\mathbf{z}_t - \mathbf{H}\tilde{\mathbf{x}}_t) \qquad (40).$$

The difference $(\mathbf{z}_t - \mathbf{H}\tilde{\mathbf{x}}_t)$ is called **measurement innovation**, or "measurement residual" and expresses the discrepancy between the predicted measurement $\mathbf{H}\tilde{\mathbf{x}}_t$ and the actual measurement array $\mathbf{z}_t$.

The Kalman Filter (KF) method

The $(2s\text{-}1) \times m$ matrix $\mathbf{K}$ in (40) is called **"Kalman Gain"** or **"blending factor"** and it **minimizes the "a posteriori" error covariance** $\hat{\mathbf{P}}_t$. It is calculated as:

$$\mathbf{K}_t = \tilde{\mathbf{P}}_t \mathbf{H}^T (\mathbf{H} \tilde{\mathbf{P}}_t \mathbf{H}^T + \mathbf{R})^{-1} \qquad (41).$$

The KF method estimates a process by using a **kind of feedback**: firstly, it estimates the process state at some time and then obtains a feedback in the form of (noisy) measurements. The KF equations are therefore divided in two groups:

- time-update equations and
- measurement-update equations.

The **time-update equations are responsible for projecting forward (in time) the current state and error covariance estimates** in order to obtain the **"a priori"** estimates for the next time step, whereas the **measurement-update equations are responsible for the feedback**, namely for the incorporation of new measurements into the "a priori" estimate so as to obtain an improved **"a posteriori" estimate**.

The Kalman Filter (KF) method

- DKF time update equations (**"prediction"**):

$$\tilde{\mathbf{x}}_t = \mathbf{A}\hat{\mathbf{x}}_{t-1} + \mathbf{B}\mathbf{u}_{t-1} \quad (42)$$

$$\tilde{\mathbf{P}}_t \equiv \mathbf{A}\hat{\mathbf{P}}_{t-1}\mathbf{A}^T + \mathbf{Q}_{t-1} \quad (43).$$

- DKF measurement update equations (kind of **"correction of the estimation"**):

$$\mathbf{K}_t = \tilde{\mathbf{P}}_t\mathbf{H}^T(\mathbf{H}\tilde{\mathbf{P}}_t\mathbf{H}^T + \mathbf{R})^{-1} \quad (44)$$

$$\hat{\mathbf{x}}_t = \tilde{\mathbf{x}}_t + \mathbf{K}_t(\mathbf{z}_t - \mathbf{H}\tilde{\mathbf{x}}_t) \quad (45)$$

$$\hat{\mathbf{P}}_t \equiv (\mathbf{I} - \mathbf{K}_t\mathbf{H})\tilde{\mathbf{P}}_t \quad (46).$$

where $\mathbf{I}$ is the identity matrix.

Example of DKF-based SE (linear)

Since we are targeting ADNs, it is worth reminding that the peculiar characteristics of these networks (e.g., high level of imbalance of lines, loads, and Distributed Generators) require the adoption of **3-phase unbalanced SE process**. Moreover, the Discrete Kalman Filter (DKF)-SE here described relies only on measurements provided by PMUs that, as above-mentioned, enable to obtain a measurement matrix $\mathbf{H}$ consisting **of constant elements**, namely: zeros, ones, and elements of the 3-ph compound admittance matrix of the network.

Example of DKF-based SE (linear)

PMUs can measure both **nodal voltage and injected current synchrophasors**. If this is the case, it is possible to take advantage of the **linear dependence** between the **network state (nodal voltages)** and the **measured injected currents**, when the equations are written in **rectangular coordinates**.

The **system state** for a network with *s* buses can be expressed in **rectangular coordinates** as:

$$\mathbf{x} = \left[\mathbf{V}_{1,re}^{a,b,c},\ldots,\mathbf{V}_{s,re}^{a,b,c},\mathbf{V}_{1,im}^{a,b,c},\ldots,\mathbf{V}_{s,im}^{a,b,c}\right]^{T} \quad (47)$$

where

$$\begin{aligned}\mathbf{V}_{i,re}^{a,b,c} &= \left[V_{i,re}^{a},V_{i,re}^{b},V_{i,re}^{c}\right]\\ \mathbf{V}_{i,im}^{a,b,c} &= \left[V_{i,im}^{a},V_{i,im}^{b},V_{i,im}^{c}\right]\end{aligned} \quad (48).$$

Example of DKF-based SE (linear)

HP: measurements coming only from PMUs → the measurement set is :

$3d_1$ phase-to-ground voltage phasors;

$3d_2$ injected current phasors;

Remark: the observability constraints are not discussed here (in any case, $2d_1+2d_2 \geq 2s$).

The measurement set $\mathbf{z}$ is:

$$\mathbf{z} = \left[\mathbf{z}_V, \mathbf{z}_I\right]^T \quad (49)$$

where

$$\mathbf{z}_V = \left[\mathbf{V}_{1,re}^{a,b,c}, \ldots, \mathbf{V}_{d_1,re}^{a,b,c}, \mathbf{V}_{1,im}^{a,b,c}, \ldots, \mathbf{V}_{d_1,im}^{a,b,c}\right]$$

$$\mathbf{z}_I = \left[\mathbf{I}_{1,re}^{a,b,c}, \ldots, \mathbf{I}_{d_2,re}^{a,b,c}, \mathbf{I}_{1,im}^{a,b,c}, \ldots, \mathbf{I}_{d_2,im}^{a,b,c}\right] \quad (50).$$

Example of DKF-based SE (linear)

As we have seen, the real and imaginary part of the nodal current injections are linked to the system state directly via the admittance matrix:

$$I_{i,re} = \sum_{\ell=1}^{s} \left(G_{i\ell} V_{\ell,re} - B_{i\ell} V_{\ell,im} \right)$$

$$I_{i,im} = \sum_{\ell=1}^{s} \left(G_{i\ell} V_{\ell,im} + B_{i\ell} V_{\ell,re} \right)$$

Past (27)

where:

- i is the bus index;
- $G_{i\ell}$ and $B_{i\ell}$ are the real and imaginary parts of the $i\ell$ admittance matrix elements, respectively.

Example of DKF-based SE (linear)

For the case the measurements are composed of nodal voltages and nodal-injected/absorbed currents, the measurements function $\mathbf{H}$ is **constant and exact**:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{\mathbf{V}} \\ \mathbf{H}_{\mathbf{I}} \end{bmatrix} \qquad \text{Past (28)}$$

where:

- $\mathbf{H}_{\mathbf{V}}$ is the part of the measurement function $\mathbf{H}$ that is related to the **link between the real and imaginary parts of the voltages as a function of the state**;
- $\mathbf{H}_{\mathbf{I}}$ is the part of the measurement function $\mathbf{H}$ that is related to the partial derivatives of the real and imaginary part of the injected currents as a function of the state.

Example of DKF-based SE (linear)

$$\mathbf{H_V} = \begin{bmatrix} [\beta] & [\upsilon] \\ [\zeta] & [\eta] \end{bmatrix}$$

$$where \begin{cases} \beta = \begin{cases} 1 \ if \ \ i = \ell \\ 0 \ if \ \ i \neq \ell \end{cases} \\ \upsilon = \zeta = 0 \\ \eta = \begin{cases} 1 \ if \ \ i = \ell \\ 0 \ if \ \ i \neq \ell \end{cases} \end{cases}$$

$$\mathbf{H_I} = \begin{bmatrix} [G_{i\ell}] & [-B_{i\ell}] \\ [B_{i\ell}] & [G_{i\ell}] \end{bmatrix}$$

Past (29)-(31)

Example of DKF-based SE (linear)

Time-step t

Prediction Equations
Prediction of the state: $\tilde{\mathbf{x}}_t = \mathbf{A}\hat{\mathbf{x}}_{t-1} + \mathbf{B}\mathbf{u}_{t-1}$; $\tilde{\mathbf{P}}_t \equiv \mathbf{A}\hat{\mathbf{P}}_{t-1}\mathbf{A}^T + \mathbf{Q}_{t-1}$

Estimation Equations
(1) Computation of the *Kalman Gain*: $\mathbf{K}_t = \tilde{\mathbf{P}}_t\mathbf{H}^T(\mathbf{H}\tilde{\mathbf{P}}_t\mathbf{H}^T + \mathbf{R})^{-1}$
(2) Estimation of the state: $\hat{\mathbf{x}}_t = \tilde{\mathbf{x}}_t + \mathbf{K}_t(\mathbf{z}_t - \mathbf{H}\tilde{\mathbf{x}}_t)$; $\hat{\mathbf{P}}_t \equiv (\mathbf{I} - \mathbf{K}_t\mathbf{H})\tilde{\mathbf{P}}_t$

where:

- $\tilde{\mathbf{P}}_t$: prediction error covariance matrix ;
- **K** : "Kalman gain" ;
- $\hat{\mathbf{P}}_t$: estimation error covariance matrix ;

Example of IKF-based SE (non-linear)

In order to provide an example, let us now suppose to have two types of measurements:

- type-d_1 nodes where **phase-to-ground voltage phasors are measured directly by means of PMUs**;

- type-u_1 nodes where active and reactive power injections are measured, so that $d_1 + u_1 \geq s$.

Example of IKF-based SE (non-linear)

The resulting measurement array $\mathbf{z}_t$ is therefore:

$$\mathbf{z} = \left[\overbrace{\delta_2, \ldots, \delta_{d_1}, V_1, \ldots, V_{d_1}}^{z^{d_1}}, \overbrace{P_1, \ldots, P_{u_1}, Q_1, \ldots, Q_{d_1}}^{z^{u_1}} \right]^T \quad (51).$$

It is worth noting that is has deliberately been disregarded the problem of network observability since it is out of the scope of this section. For sake of simplicity the measurement Jacobian $\mathbf{H}$ is formulated as:

$$\mathbf{H} = \begin{bmatrix} & \mathrm{T} & \\ \frac{\partial P_{inj}}{\partial \delta} & & \frac{\partial P_{inj}}{\partial V} \\ \frac{\partial P_{flow}}{\partial \delta} & & \frac{\partial P_{flow}}{\partial V} \\ \frac{\partial Q_{inj}}{\partial \delta} & & \frac{\partial Q_{inj}}{\partial V} \\ \frac{\partial Q_{flow}}{\partial \delta} & & \frac{\partial Q_{flow}}{\partial V} \\ \frac{\partial I_{magn}}{\partial \delta} & & \frac{\partial I_{magn}}{\partial V} \end{bmatrix} = \begin{bmatrix} \mathbf{T} \\ \mathbf{H}_s \end{bmatrix} \quad (52).$$

Example of IKF-based SE (non-linear)

The $(2d_1\text{-}1)\text{x}(2s\text{-}1)$ sub-matrix $\mathbf{T}$, that is composed of the rows of the $(2s\text{-}1)$ identity matrix corresponding to the type-d_1 nodes, allows linking the first part of the measurement array $\mathbf{z}_t^{d_1}$ to the system state variables:

$$\mathbf{z}_t^{d_1} = \mathbf{T}\mathbf{x}_t + \mathbf{v}_t^{d_1} \qquad (53).$$

where $\mathbf{v}_t^{d_1}$ represents the array of uncertainties of the state variables, measured by the PMUs. Therefore:

$$\mathbf{z}_t = \overbrace{\begin{bmatrix} \mathbf{T} \\ \mathbf{H}^{u_1} \end{bmatrix}}^{\mathbf{H}} \mathbf{x}_t + \mathbf{v}_t \qquad (54)$$

and

$$\mathbf{v}_t = \begin{bmatrix} \mathbf{v}_t^{d_1} \\ \hline \mathbf{v}_t^{u_1} \end{bmatrix} \qquad (55).$$

Due to the **combination of conventional power measurements and PMU measurements** the SE provides, in general, more accurate results. However, it is important to point out that **the number and the location of the measurement devices has a strong impact on the SE results.**