Parameter Estimation of Three Phase Untransposed Short Transmission Lines from Synchrophasor Measurements

Antoine Wehenkel

***Abstract*—This thesis presents estimation methods for the parameters of a three-phase untransposed short transmission line from voltage/current synchro-phasor measurements taken at the two ends of the line. Due to high sensitivity to measurement noise, conventional approaches, which use the admittance matrix to model the line, are not applicable for parameter estimation for short lines. We propose to use the transmittance matrix to model the line. We show that the parameters of the transmittance matrix are less sensitive to measurement noise than that of the admittance matrix or the impedance matrix. Therefore, using this model we can obtain more accurate and robust estimates of the line parameters for short lines. The different methods proposed are based on different assumptions made on the projection of the complex measurements in rectangular coordinates. These methods are ordinary/weighted least squares (OLS/WLS) and ordinary/weighted total least squares (OTLS/WTLS). Thanks to an accurate model of the noise in polar coordinates we are able to compute the covariance matrices of the noise in rectangular coordinates needed for some of these methods. We evaluate the accuracy of the proposed methods for different classes of instrument transformers, and on a variety of lines with different parameters. We show that OTLS has a poor accuracy, WTLS has a good accuracy but has prohibitive computational requirements, OLS has good accuracy and least computational requirements, and finally WLS has the best accuracy with fairly low computational requirements. Furthermore we show the superiority of our approach by comparing it with the state of the art techniques on a realistic dataset.**

I. Introduction

Large-scale integration of renewable energy sources has made necessary the development of control methods for distribution networks because of their stochasticity. Most modern control methods for electrical grids require the knowledge of the parameter of the lines to do state estimation, to solve optimal power flow problems (e.g [1], [2], [3], [4], [5], [6]) or to improve accuracy in relay settings ([7]). However this knowledge is often absent or very crude due to incomplete records of installation and updates.

Historically, parameters of electrical lines are obtained by using the physical properties of the line such as conductor dimensions, type of wires, tower geometries, etc. (e.g. [8]). Although this method has been applied successfully for long transmission lines (TL) present in transport network, it cannot be easily applied to distribution networks where there is often no good knowledge of the characteristics of the lines. Another approach, which requires to disconnect the line during the measurements, is based on generation of known signals at one end of the line and precise measurements on the other side ([9]). Both methods, in addition to being time consuming and prone to human errors, have the drawback to provide only static knowledge of the line parameters whereas the parameters change over time. State of the art control methods for distribution networks already use phasor measurement units (PMUs). These devices can also be exploited to do line parameters estimation while circumventing the drawbacks of the methods mentioned earlier.

Modern day PMUs have become accurate enough so that many recent works propose their utilization to estimate line parameters. More specifically [10], [11] use a single-phase model of the line and PMU measurements taken from both ends of a line to estimate its parameters. These methods are already used in practice for transmission networks ([12]). However for distribution networks the single phase model of the line is often not accurate due to the presence of untransposed lines. Although the problem of three-phase line parameter estimation is very similar to single-phase estimation, it requires the estimation of many more parameters which makes the problem harder. In [13] a method for three-phase line parameters has been proposed which is based on ordinary least squares. Although the method is accurate for long lines it fails in the case of short lines because the model used for estimation is highly sensitive to measurement noise. Moreover as reported in [14] the method seems to require prior knowledge of the ranges of the parameters to have acceptable quality of estimation. Finally all these works lack a realistic noise model. It makes the results of their simulations of limited use. Several other methods based on only 1 or 2 measurements (e.g [15]) have been proposed but [13] has proven these methods to be very inaccurate in the presence of noise.

Contributions: The main contribution of this thesis is to propose and evaluate methods for accurate and robust estimation of untransposed line parameters for short lines. All the existing methods for line parameters estimation fails for short lines due to the high noise sensitivity of the model being used. To overcome this difficulty, we propose a model based on the transmittance matrix of the line and show that this model has less noise sensitivity for short lines compared to other estimation models. We compare several estimation methods based on least squares and total least squares techniques. We show that, for short lines, our method outperforms the state-of-the-art technique proposed in [13].

We use an accurate model for measurement noise obtained from transforming them from polar to rectangular coordinates. The proposed methods are numerically evaluated for a range of different instruments transformers and on a variety of

lines with different characteristics. From these evaluations we conclude that the proposed methods based on the transmittance matrix formulation is more robust to measurement noises.

The rest of the thesis is organized as follows. In Section II, we formulate the problem of line parameter estimation using the transmittance matrix model of the line, we introduce the noise model used in the thesis, we also present 2 other line models and we give the definition of the general statistical model used to describe the estimation methods. In Section III we describe the proposed estimation methods. Section IV evaluates the proposed algorithms numerically and compares the different line models for different lengths of line. The thesis is finally concluded in Section V.

II. Problem Formulation

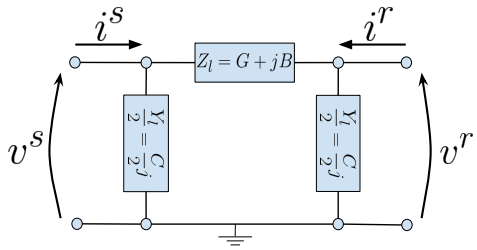

Fig. 1. Simplified π-model of three phase TL.

A. Transmittance matrix model

We are interested in identifying the parameters of an untransposed 3-phase line modelled via its π equivalent circuit (see Figure 1). The three-phase series impedance and shunt admittance matrices of the line between the two nodes s and r are denoted by $\mathbf{Z_l} \in \mathbb{C}^{3\times3}$ and $\mathbf{Y_l} \in \mathbb{C}^{3\times3}$, respectively. Let $i^x, v^x \in \mathbb{C}^3$ denote the complex three-phase current and voltage phasors at node $x \in \{s, r\}$, respectively. The relation between the phasors at the two terminals can be written as follows (e.g. [16]):

$$\begin{bmatrix} v^s \\ i^s \end{bmatrix} = \begin{bmatrix} \mathbf{I} + \frac{\mathbf{Z_l}\mathbf{Y_l}}{2} & -\mathbf{Z}_l \\ \mathbf{Y}_l\left(\mathbf{I} + \frac{\mathbf{Z_l}\mathbf{Y_l}}{4}\right) & -\left(\mathbf{I} + \frac{\mathbf{Y_l}\mathbf{Z_l}}{2}\right) \end{bmatrix} \begin{bmatrix} v^r \\ i^r \end{bmatrix}, \quad (1)$$

$$\approx \begin{bmatrix} \mathbf{I} & -\mathbf{Z}_l \\ \mathbf{Y}_l & -\mathbf{I} \end{bmatrix} \begin{bmatrix} v^r \\ i^r \end{bmatrix}, \quad (2)$$

where $\mathbf{I} \in \mathbb{C}^{3\times3}$ denotes the identity matrix. The approximation in (2) is obtained by using the fact that the product matrix $\mathbf{Z_l}\mathbf{Y_l}$ is usually close to the null matrix $\mathbf{0} \in \mathbb{C}^{3\times3}$ for lines that are electrically short ($\ll$ 100km) (see [17]), i.e., $\mathbf{Z_l}\mathbf{Y_l} \approx \mathbf{0}$. In what follows, we refer to the matrix on the right hand side of (2) as the transmittance matrix $\mathbf{T}$.

The objective of the thesis is to estimate the elements of $\mathbf{T}$ or equivalently the elements of $\mathbf{Y}_l$ and $\mathbf{Z}_l$ using noisy observations of v^x and i^x, $x \in \{s, r\}$. Exploiting the symmetric structure of $\mathbf{Y}_l$, $\mathbf{Z}_l$, and the fact that $\mathbf{Y}_l$ is assumed to be purely imaginary, we express the matrices with the following parametrization

$$\mathbf{Y}_l(\theta^{\mathbf{T}}) = j \begin{bmatrix} \theta_1^{\mathbf{T}} & \theta_{16}^{\mathbf{T}} & \theta_{17}^{\mathbf{T}} \\ \theta_{16}^{\mathbf{T}} & \theta_2^{\mathbf{T}} & \theta_{18}^{\mathbf{T}} \\ \theta_{17}^{\mathbf{T}} & \theta_{18}^{\mathbf{T}} & \theta_3^{\mathbf{T}} \end{bmatrix}, \quad (3)$$

$$\mathbf{Z}_l(\theta^{\mathbf{T}}) = \begin{bmatrix} \theta_4^{\mathbf{T}} + \theta_5^{\mathbf{T}} j & \theta_6^{\mathbf{T}} + \theta_7^{\mathbf{T}} j & \theta_8^{\mathbf{T}} + \theta_9^{\mathbf{T}} j \\ \theta_6^{\mathbf{T}} + \theta_7^{\mathbf{T}} j & \theta_{10}^{\mathbf{T}} + \theta_{11}^{\mathbf{T}} j & \theta_{12}^{\mathbf{T}} + \theta_{13}^{\mathbf{T}} j \\ \theta_8^{\mathbf{T}} + \theta_9^{\mathbf{T}} j & \theta_{12}^{\mathbf{T}} + \theta_{13}^{\mathbf{T}} j & \theta_{14}^{\mathbf{T}} + \theta_{15}^{\mathbf{T}} j \end{bmatrix}, \quad (4)$$

where $\theta^{\mathbf{T}}_{=}[\theta_1^{\mathbf{T}}, ..., \theta_{18}^{\mathbf{T}}]^T \in \mathbb{R}^{18}$ is the unknown parameter vector. The transmittance matrix as a function of the parameter vector θ is therefore expressed as

$$\mathbf{T}(\theta^{\mathbf{T}}) = \begin{bmatrix} \mathbf{I} & -\mathbf{Z_l}(\theta^{\mathbf{T}}) \\ \mathbf{Y}_l(\theta^{\mathbf{T}}) & -\mathbf{I} \end{bmatrix}.$$

To simplify the estimation problem, we first express (2) in rectangular coordinates as follows:

$$l^s = \bar{\mathbf{T}}(\theta^{\mathbf{T}}) l^r, \quad (5)$$

where

$$l^x = \begin{bmatrix} \mathcal{R}(v^x) \\ \mathcal{R}(i^x) \\ \mathcal{I}(v^x) \\ \mathcal{I}(i^x) \end{bmatrix} \in \mathbb{R}^{12}, \quad x \in \{r, s\}, \quad (6)$$

$$\bar{\mathbf{T}}(\theta^{\mathbf{T}}) = \begin{bmatrix} \mathbf{I} & -\mathcal{R}(\mathbf{Z}_l(\theta^{\mathbf{T}})) & \mathbf{0} & \mathcal{I}(\mathbf{Z}_l(\theta^{\mathbf{T}})) \\ \mathbf{0} & -\mathbf{I} & -\mathcal{I}(\mathbf{Y}_l(\theta^{\mathbf{T}})) & \mathbf{0} \\ \mathbf{0} & -\mathcal{I}(\mathbf{Z}_l(\theta^{\mathbf{T}})) & \mathbf{I} & -\mathcal{R}(\mathbf{Z}_l(\theta^{\mathbf{T}})) \\ \mathcal{I}(\mathbf{Y}_l(\theta^{\mathbf{T}})) & \mathbf{0} & \mathbf{0} & -\mathbf{I} \end{bmatrix}, \quad (7)$$

and $\mathcal{R}(\cdot), \mathcal{I}(\cdot)$ respectively denote the real and imaginary parts of a complex variable.

To further simplify the estimation, we convert (5) to the following standard form

$$l^s = \mathbf{H_T}(l^r)\theta^{\mathbf{T}} + \lambda(l^r), \quad (8)$$

where

$$\lambda(l) = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{I} \end{bmatrix} l, \quad \forall l \in \mathbb{R}^{12}, \quad (9)$$

and $\mathbf{H_T} : \mathbb{R}^{12} \to \mathbb{R}^{12\times18}$ is a linear map such that $\mathbf{H_T}(l)\theta^{\mathbf{T}} = \bar{\mathbf{T}}(\theta^{\mathbf{T}})l - \lambda(l)$ for all $\theta^{\mathbf{T}} \in \mathbb{R}^{18}$ and all $l \in \mathbb{R}^{12}$. Given a vector $l \in \mathbb{R}^{12}$, the matrix $\mathbf{H_T}(l)$ can be very easily computed. The key idea is that since $\mathbf{H_T} : \mathbb{R}^{12} \to \mathbb{R}^{12\times18}$ is a linear map, each column i of $\mathbf{H_T}(l)$ is of the form $\mathbf{\Omega}_i l$ for some coefficient matrix $\mathbf{\Omega}_i \in \mathbb{R}^{12\times12}, i = 1:18$, i.e.,

$$\mathbf{H_T}(l) = [\mathbf{\Omega}_1 l \quad \mathbf{\Omega}_2 l \ldots \mathbf{\Omega}_{18} l].$$

It is straightforward to see that the j^{th} column of $\mathbf{\Omega}_i$ can be recovered as

$$\mathbf{\Omega}_i(:,j) = \left[\bar{\mathbf{T}}(\theta^{\mathbf{T}})l - \gamma(l)\right]_{\theta^{\mathbf{T}}=e_i^{(18)}, l=e_j^{(12)}},$$

where $e_k^{(r)}$ denotes the k^{th} unit vector in $\mathbb{R}^r$. Pseudo code for this computation is given in appendix V-F.

Remark. *Although the problem formulation presented assumes a model with 18 parameters, the formulation can be generalized to any other parameter space of size $\mathcal{S}_\theta$. Indeed until now we only considered the most generic type of line, however in the simulations section we present numerical results for lines which can be expressed with smaller parameter space.*

B. Other line models

There are other line models that are classically used to link the voltage and current phasors at two ends of a three phase line. They are described briefly below:

First, the currents and the voltages of the two ends of the line are related through the admittance matrix ($\mathbf{Y}$) in the following way:

$$i = \mathbf{Y}(\theta^{\mathbf{T}})v, \tag{10}$$

where

$$i = \begin{bmatrix} i^r \\ i^s \end{bmatrix} \quad v = \begin{bmatrix} v^r \\ v^s \end{bmatrix},$$

and

$$\mathbf{Y}(\theta^{\mathbf{T}}) = \begin{bmatrix} \left(\mathbf{Z_l}(\theta^{\mathbf{T}})\right)^{-1} + \frac{\mathbf{Y_l}(\theta^{\mathbf{T}})}{2} & -\left(\mathbf{Z_l}(\theta^{\mathbf{T}})\right)^{-1} \\ -\left(\mathbf{Z_l}(\theta^{\mathbf{T}})\right)^{-1} & \left(\mathbf{Z_l}(\theta^{\mathbf{T}})\right)^{-1} + \frac{\mathbf{Y_l}(\theta^{\mathbf{T}})}{2} \end{bmatrix}.$$

Similarly, they can also be related through the impedance matrix ($\mathbf{Z}$) in the following way: (the derivation is very easy using the T model of the line, see Figure 2)

$$v = \mathbf{Z}(\theta^{\mathbf{T}})i, \tag{11}$$

where

$$\mathbf{Z}(\theta^{\mathbf{T}}) = \begin{bmatrix} \frac{\mathbf{Z_l}(\theta^{\mathbf{T}})}{2} + \mathbf{Y_l}(\theta^{\mathbf{T}})^{-1} & \mathbf{Y_l}(\theta^{\mathbf{T}})^{-1} \\ \mathbf{Y_l}(\theta^{\mathbf{T}})^{-1} & \frac{\mathbf{Z_l}(\theta^{\mathbf{T}})}{2} + \mathbf{Y_l}(\theta^{\mathbf{T}})^{-1} \end{bmatrix}.$$

In order to express the relationship between the measured phasors and the unknown parameters in a way similar to what has been done for the $\mathbf{T}$ line model, we need to express the $\mathbf{Y}$ and $\mathbf{Z}$ matrices linearly in terms of the unknown parameters of the line. In terms of the unknowns in $\theta^{\mathbf{T}}$, the matrices $\mathbf{Y}$ and $\mathbf{Z}$ are not linear because of the presence of $\mathbf{Z}_l^{-1}$ and $\mathbf{Y}_l^{-1}$. To have a linear estimation model, therefore, we need to define two new parameter vectors $\theta^{\mathbf{Y}}$ and $\theta^{\mathbf{Z}}$ which respectively parametrizes $\mathbf{Y}_l^{-1}$ and $\mathbf{Z}_l^{-1}$ and with respect to which the $\mathbf{Y}$ and $\mathbf{Z}$ models are truly linear. The mapping between $\theta^{\mathbf{Y}}$ and $\theta^{\mathbf{T}}$ is defined as

$$\mathbf{Y_l}(\theta^{\mathbf{Y}}) = \mathbf{Y_l}(\theta^{\mathbf{T}})$$

and

$$\mathbf{Z_l}(\theta^{\mathbf{T}})^{-1} = \begin{bmatrix} \theta_4^{\mathbf{Y}} + \theta_5^{\mathbf{Y}} j & \theta_6^{\mathbf{Y}} + \theta_7^{\mathbf{Y}} j & \theta_8^{\mathbf{Y}} + \theta_9^{\mathbf{Y}} j \\ \theta_6^{\mathbf{Y}} + \theta_7^{\mathbf{Y}} j & \theta_{10}^{\mathbf{Y}} + \theta_{11}^{\mathbf{Y}} j & \theta_{12}^{\mathbf{Y}} + \theta_{13}^{\mathbf{Y}} j \\ \theta_8^{\mathbf{Y}} + \theta_9^{\mathbf{Y}} j & \theta_{12}^{\mathbf{Y}} + \theta_{13}^{\mathbf{Y}} j & \theta_{14}^{\mathbf{Y}} + \theta_{15}^{\mathbf{Y}} j \end{bmatrix} \tag{12}$$

(13)

The parameter vector $\theta^{\mathbf{Z}}$ can also be defined similarly.

With the new parameter vectors defined above, we can now express the complex relationships (10), (11) in rectangular coordinates (separating the real and imaginary parts) in terms of the real matrices $\bar{\mathbf{Z}}(\theta^{\mathbf{Z}})$ and $\bar{\mathbf{Y}}(\theta^{\mathbf{Y}})$, whose definitions are similar to that of $\bar{\mathbf{T}}(\theta^{\mathbf{T}})$ in the $\mathbf{T}$ line model (see (5)-(7)).

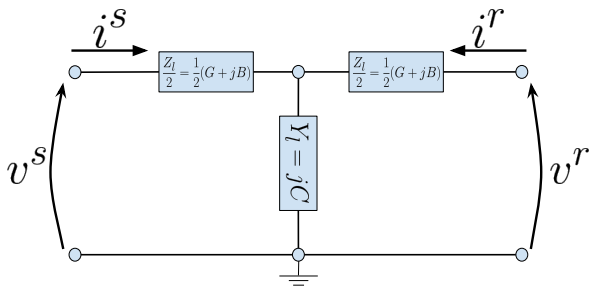

Fig. 2. Simplified T-model of three phase TL.

C. Noise model

In this subsection, we describe the noise model used in this thesis. The noisy measurements of l^x is denoted as $\tilde{l}^x$ for $x \in \{s, r\}$ and are defined as

$$\tilde{l}^x = l^x + \Delta_{l^x}, \quad x \in \{s, r\},$$

where Δ_{l^x} denotes the noise on l^x. To statistically characterize the noise, we make use of standard specifications of the measurement devices.

The noise in the measurements are introduced from two sources: 1) from the instrument transformers (IT's), i.e., the voltage transformers for voltage measurements and the current transformers for current measurements, 2) from the PMU's. In this thesis we consider class .1 PMU which are characterised by a supplementary maximum relative noise of 0.1% on the magnitude and a maximum absolute noise on the phase equal to 10^{-4} (see [18]). For this reason we compute the total maximum error by summing the errors of the IT's and PMU's. The manufacturers of instrument transformers specify - as percentages in case of magnitudes and as absolute values in case of phases - the maximum errors introduced by such devices. The maximum errors introduced by different classes of instrument transformers are given in Table I [19]. Note that the maximum errors are the same for voltage and current transformers.

We assume that the noise on any variable is Gaussian, unbiased, and lies within the specified maximum bound with probability 0.9973. Hence to find the standard deviation, we divide the value of the maximum error by 3. More specifically, let ρ and ϕ respectively denote the magnitude and phase of a complex quantity w, i.e., $w = \rho e^{j\phi}$. Then the noise Δ_ρ and Δ_ϕ

on ρ and ϕ are assumed to be distributed as $\Delta_\rho \sim \mathcal{N}(0, \frac{\alpha\rho}{3})$, and $\Delta_\phi \sim \mathcal{N}(0, \beta/3)$, where α and β denote the maximum error in percentage for magnitude and maximum absolute error for phase, respectively (see Table I for values).

TABLE I
MAXIMUM ERRORS FOR DIFFERENT CLASSES OF INSTRUMENT TRANSFORMERS (IT).

Transformer Class	Max. magnitude error (α) (%)	Max. phase error (β) (rad)
0.1	0.1	15e-04
0.2	0.2	30e-04
0.5	0.5	90e-04
1	1	180e-04

While the noise is Gaussian in polar coordinates, we use rectangular coordinates in our estimation procedures. In general, the noise projection from polar to rectangular coordinates does not preserve normality. However, for the parameters of Table I we observe numerically that the distribution of the noise transformed to the rectangular coordinates is very close to Gaussian distribution. This is shown by quantile-quantile (QQ) plots in Figures 3 and 4 for Class 1 instrument transformers. We observe that the quantiles of the transformed noise on both voltage and current match very closely with those of the standard normal distribution. For this reason we treat the noise in the rectangular coordinates as Gaussian random variables. We are conscious that QQ plots only verify the Gaussianity of marginal density and that other tests would be required to ensure that the joint density is also Gaussian.

To find the first and second moments of the transformed noise, we first denote the noisy measurements of magnitude and phase of the variable $w = \rho e^{j\phi}$ as $\tilde{\rho}$ and $\tilde{\phi}$, respectively. If Δ_r and Δ_i respectively denote the noise in the real and imaginary parts, then we have

$$\Delta_r = \tilde{\rho}\cos(\tilde{\phi}) - \rho\cos(\phi) \tag{14}$$

$$E[\Delta_r] = \left(e^{-\frac{1}{2}\sigma_\phi^2} - 1\right)\cos(\phi)\rho \tag{15}$$

$$E[\Delta_r^2] = \frac{1}{2}(1 + \frac{\alpha^2}{9})\rho^2(1 + e^{-2\sigma_\phi^2}\cos(2\phi)) + \rho^2\cos^2(\phi)(1 - 2e^{(-\frac{1}{2}\sigma_\phi^2)}) \tag{16}$$

$$\Delta_i = \tilde{\rho}\sin(\tilde{\phi}) - \rho\sin(\phi) \tag{17}$$

$$E[\Delta_i] = \left(e^{-\frac{1}{2}\sigma_\phi^2} - 1\right)\sin(\phi)\rho \tag{18}$$

$$E[\Delta_i^2] = \frac{1}{2}\left(1 + \frac{\alpha^2}{9}\right)\rho^2\left(1 + e^{-2\sigma_\phi^2}\cos(2\phi)\right) + \rho^2\sin^2(\phi)\left(1 - 2e^{-\frac{1}{2}\sigma_\phi^2}\right) \tag{19}$$

$$E[\Delta_r\Delta_i] = \frac{1}{2}\sin(2\phi)\left[\left(1 + \frac{\alpha^2}{9}\right)\rho^2 e^{-2\sigma_\phi^2} - 2\rho^2 e^{\left(-\frac{1}{2}\sigma_\phi^2\right)} + \rho^2\right], \tag{20}$$

where $E[\cdot]$ is the expectation operator and $\sigma_\phi = \beta/3$. The derivations of (14) to (20) are given in V-A We note from (15) and (18) that the noise in the rectangular coordinates is biased in general. However, for the parameters of Table I we observe that the bias is negligible. This is also confirmed by the QQ plots in Figure 4 and 3. Furthermore, we observe that for the parameters of interest the real and imaginary part of the noise have non-negligible correlation. This is usually neglected in all existing models for line parameter estimation.

The entries of the covariance matrix for Δ_x, $x \in \{l^s, l^r, i, v\}$ can therefore be obtained from the parameters of Table I using (15)-(20). We note that these entries in general depend on the true value of x which is unknown in practice. So in our estimation procedure the entries of the covariance matrices are computed using the observed values $\tilde{x}$. We have numerically observed that replacing the true values with the observed values have a negligible impact on the estimation procedures proposed in this thesis.

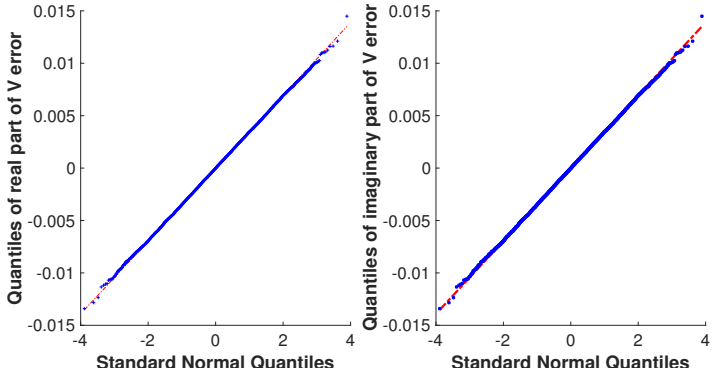

Fig. 3. QQ plots of errors on voltages measurement taken by class 1 IT.

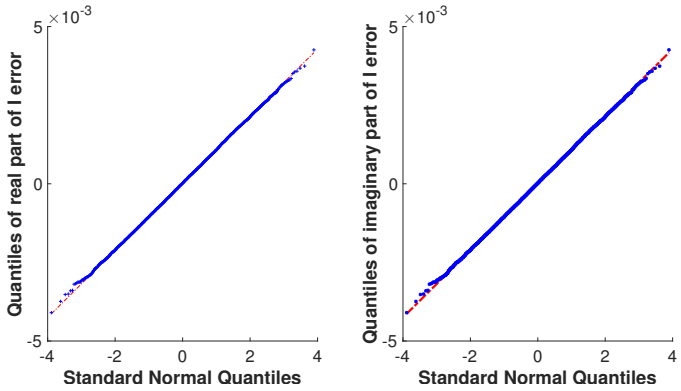

Fig. 4. QQ plots of errors on currents measurement taken by class 1 IT.

D. Statistical model

In this section we define a general statistical model which incorporates any of the three previously discussed line models ($\mathbf{Z}$, $\mathbf{Y}$, $\mathbf{T}$) and the noise model discussed in Section II-C.

In our cases, the statistical model corresponds to an "error-in-variables (EIV) linear regression model" [20] and can be

written as follows if we assume the noise to be zero-mean Gaussian:

$$y_i = \mathbf{B}(\theta)x_i, \tag{21}$$
$$\tilde{y}_i = y_i + \Delta_{y_i} \quad \Delta_{y_i} \sim \mathcal{N}(0, \mathbf{Q}_{y_i}), \tag{22}$$
$$\tilde{x}_i = x_i + \Delta_{x_i} \quad \Delta_{x_i} \sim \mathcal{N}(0, \mathbf{Q}_{x_i}), \tag{23}$$

where $\tilde{y}_i, \tilde{x}_i \in \mathbb{R}^{12}$ denote the ith noisy phasor measurements of the true phasors $y_i, x_i \in \mathbb{R}^{12}$; $\Delta_{y_i}, \Delta_{x_i} \in \mathbb{R}^{12}$ denote the noise on y_i and x_i respectively; and $\mathbf{Q}_{y_i}$, $\mathbf{Q}_{x_i} \in \mathbb{R}^{12\times 12}$ respectively denote the noise covariance matrices of $\Delta_{y_i}, \Delta_{x_i}$ (easily computable thanks to (16), (19) and (20)). The link between y_i and x_i is given by a matrix $\mathbf{B}(\theta) \in \mathbb{R}^{12\times 12}$ where $\theta \in \mathbb{R}^{\mathcal{S}_\theta}$, the unknown parameter, is a vector parametrizing linearly $\mathbf{B}$. Because the parametrization is linear we can also rewrite (21) as

$$y_i = \mathbf{B}(\theta)x_i = \mathbf{H}(x_i)\theta + \gamma(x_i),$$

with $\mathbf{H} \in \mathbb{R}^{12\times \mathcal{S}_\theta}$ can be computed as detailed for $\mathbf{T}$ model in II-A, and $\gamma(.) : \mathbb{R}^{12} \to \mathbb{R}^{12}$ is a known linear function of x_i.

We can write the statistical model incorporating all the snapshots of measurements in a more compact way as given below:

$$y = \bar{\mathbf{B}}(\theta)x \tag{24}$$
$$\tilde{y} = y + \Delta_y \quad \Delta_y \sim \mathcal{N}(0, \mathbf{Q}_y) \tag{25}$$
$$\tilde{x} = x + \Delta_x \quad \Delta_x \sim \mathcal{N}(0, \mathbf{Q}_x) \tag{26}$$

where $x, y \in \mathbb{R}^{12N}$ are defined as:

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \tag{27}$$

and similarly for their noisy version $\tilde{x}, \tilde{y} \in \mathbb{R}^{12N}$, $\bar{\mathbf{B}}(\theta) \in \mathbb{R}^{12N\times 12N}$ is a block diagonal matrix composed of $\mathbf{B}(\theta)$ matrices. The covariance matrices $\mathbf{Q}_x$ and $\mathbf{Q}_y$ are also block diagonal matrices composed of $\mathbf{Q}_{x_i}$ and $\mathbf{Q}_{y_i}$ blocks respectively. Reformulating the right term of (24) in term of vector θ we have:

$$y = \mathbf{H}(x)\theta + \gamma(x)$$

with $\mathbf{H}(x) \in \mathbb{R}^{12N\times \mathcal{S}_\theta}$ defined as

$$\mathbf{H}(x) = \begin{bmatrix} \mathbf{H}(x_1) \\ \vdots \\ \mathbf{H}(x_N) \end{bmatrix}$$

and

$$\gamma(x) = \begin{bmatrix} \gamma(x_1) \\ \vdots \\ \gamma(x_N) \end{bmatrix}.$$

It is straightforward to define the log-likelihood function of this model:

$$L(\tilde{x}, \tilde{y}|\theta, x) = (\tilde{x} - x)^T \mathbf{Q}_x^{-1} (\tilde{x} - x) + \left(\tilde{y} - \bar{\mathbf{B}}(\theta)x\right)^T \mathbf{Q}_y^{-1} \left(\tilde{y} - \bar{\mathbf{B}}(\theta)x\right) + const. \tag{28}$$

A mapping between statistical model and the three line models are given in Table II.

TABLE II
MAPPING BETWEEN STATISTICAL MODEL AND LINE MODELS.

Stat. Model	$\mathbf{T}$	$\mathbf{Z}$	$\mathbf{Y}$
y_i	$l_i^s + \gamma(l_i^r)$	$\left[\mathcal{R}(v_i^T)\, \mathcal{I}(v_i^T)\right]^T$	$\left[\mathcal{R}(i_i^T)\, \mathcal{I}(i_i^T)\right]^T$
x_i	l_i^r	$\left[\mathcal{R}(i_i^T)\, \mathcal{I}(i_i^T)\right]^T$	$\left[\mathcal{R}(v_i^T)\, \mathcal{I}(v_i^T)\right]^T$
$\mathbf{B}$	$\bar{\mathbf{T}}$	$\bar{\mathbf{Z}}$	$\bar{\mathbf{Y}}$
$\gamma(.)$	$\lambda(.)$	0	0
θ	$\theta^{\mathbf{T}}$	$\theta^{\mathbf{Z}}$	$\theta^{\mathbf{Y}}$

Where the subscript i denotes the ith observation.

III. Proposed Estimation Methods

We now describe different estimation methods in terms of the statistical model defined in the previous section. The objective is to estimate the parameter vector θ from the noisy observations $\tilde{x}$ and $\tilde{y}$. To evaluate different estimation methods, in addition to the true statistical model, we use different approximations of the true model as described below.

A. Maximum Likelihood estimator - Weighted Total Least Squares

The maximum likelihood estimator, also named Weighted Total Least Squares (WTLS) uses the true statistical model defined before to compute an estimator of θ. Clearly, from (28) the estimator can be found by solving the following minimization problem:

$$\min_\theta \min_x (\tilde{x} - x)^T \mathbf{Q}_x^{-1} (\tilde{x} - x) + \left(\tilde{y} - \bar{\mathbf{B}}(\theta)x\right)^T \mathbf{Q}_y^{-1} \left(\tilde{y} - \bar{\mathbf{B}}(\theta)x\right) \tag{29}$$

We note that the inner minimization problem on x is convex and therefore can be solved in closed form as a function of θ. The solution $\hat{x}(\theta)$ is given as

$$\hat{x}(\theta) = \left(\mathbf{Q}_x^{-1} + \bar{\mathbf{B}}^T(\theta)\mathbf{Q}_x^{-1}\bar{\mathbf{B}}(\theta)\right)^{-1} \times \left(\mathbf{Q}_x^{-1}\tilde{x} + \bar{\mathbf{B}}^T(\theta)\mathbf{Q}_y^{-1}\tilde{y}\right), \tag{30}$$

Finally, to solve for the optimal θ we solve the following problem:

$$\hat{\theta}_{ml} = \arg\min_\theta (\tilde{x} - \hat{x}(\theta))^T \mathbf{Q}_x^{-1} (\tilde{x} - \hat{x}(\theta)) + (\tilde{y} - \mathbf{H}(\hat{x}(\theta))\theta - \gamma(\hat{x}(\theta)))^T \mathbf{Q}_y^{-1} \times (\tilde{y} - \mathbf{H}(\hat{x}(\theta))\theta - \gamma(\hat{x}(\theta))).$$

We note that the above optimization problem is non-convex and we use MATLAB fminunc to solve it.

B. Ordinary Total Least Squares

The ordinary total least squares (OTLS) method is classically used in error in variables regression [21], [20]. Here, the objective is to estimate a parameter matrix $\mathbf{U} \in \mathbb{C}^{p\times q}$ which satisfies the linear relationship

$$\mathbf{C}\mathbf{U} = \mathbf{D} \tag{31}$$

with $\mathbf{C} \in \mathbb{C}^{n\times p}$ and $\mathbf{D} \in \mathbb{C}^{n\times q}$ from noisy observations $\tilde{\mathbf{C}}$ of $\mathbf{C}$ and $\tilde{\mathbf{D}}$ of $\mathbf{D}$. Note that in ordinary least squares regression, errors are assumed to be present only in the matrix $\mathbf{D}$. In the special case where $\mathbf{D}$ and $\mathbf{U}$ are vectors ($q = 1$) we denote them by lower cases d and u.

In general (31) can be rewritten in the following form

$$\begin{bmatrix}\mathbf{C}\,\mathbf{D}\end{bmatrix}\begin{bmatrix}\mathbf{U}\\ -\mathbf{I}_{q\times q}\end{bmatrix} = \mathbf{0}_{n\times q}. \tag{32}$$

The above can be used to find the denoised version $\hat{\mathbf{C}}$ and $\hat{\mathbf{D}}$ of $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{D}}$ as follows:

$$\hat{\mathbf{C}}, \hat{\mathbf{D}} = \underset{\mathbf{C},\mathbf{D}}{\arg\min} \left\| \begin{bmatrix}\mathbf{C}\,\mathbf{D}\end{bmatrix} - \begin{bmatrix}\tilde{\mathbf{C}}\,\tilde{\mathbf{D}}\end{bmatrix} \right\|_F \tag{33}$$

$$\text{subject to} \quad \mathsf{rank}\left(\begin{bmatrix}\mathbf{C}\,\mathbf{D}\end{bmatrix}\right) \le q. \tag{34}$$

Note that the rank constraint in the above optimization is equivalent to the existence of a parameter matrix $\mathbf{U} \in \mathbb{R}^{p\times q}$ satisfying (32). The above optimization is an instance of the low rank approximation problem for the observed data matrix $\begin{bmatrix}\mathbf{C}\,\mathbf{D}\end{bmatrix}$ and can be solved using the singular value decomposition (SVD) of $\begin{bmatrix}\tilde{\mathbf{C}}\,\tilde{\mathbf{D}}\end{bmatrix}$. Let the SVD of $\begin{bmatrix}\tilde{\mathbf{C}}\,\tilde{\mathbf{D}}\end{bmatrix}$ be

$$\begin{bmatrix}\tilde{\mathbf{C}}\,\tilde{\mathbf{D}}\end{bmatrix} = \begin{bmatrix}\mathbf{U}_C\,\mathbf{U}_D\end{bmatrix}\begin{bmatrix}\mathbf{\Sigma}_C & \mathbf{0}_{p\times q}\\ \mathbf{0}_{q\times p} & \mathbf{\Sigma}_D\end{bmatrix}\begin{bmatrix}\mathbf{V}_{C,C} & \mathbf{V}_{C,D}\\ \mathbf{V}_{D,C} & \mathbf{V}_{D,D}\end{bmatrix}^*$$

where $\mathbf{A}^*$ denotes the matrix conjugate of $\mathbf{A} \in \mathbb{C}^{n\times n}$.

Then the parameter matrix $\mathbf{U}$ is estimated as (see [21]):

$$\hat{\mathbf{U}}_{\text{tls}} = -\mathbf{V}_{C,D}\mathbf{V}_{D,D}^{-1}. \tag{35}$$

We note that $\hat{\mathbf{C}}, \hat{\mathbf{D}}$ are the maximum likelihood estimates of $\mathbf{C}, \mathbf{D}$ only when the noise $\begin{bmatrix}\Delta_{\mathbf{C}}\,\Delta_{\mathbf{D}}\end{bmatrix}$ on the data matrix $\begin{bmatrix}\mathbf{C}\,\mathbf{D}\end{bmatrix}$ has independent and identically distributed (zero-mean Gaussian) rows which is an approximation of the true statistical model.

1) Structured OTLS: The structured OTLS (SOTLS) problem in the context of the estimation problem of θ is solved by putting

$$\mathbf{C} = \mathbf{H}(x), \quad u = \theta, \quad d = y - \gamma(x). \tag{36}$$

The estimate $\hat{\theta}_{tls}$ of θ can therefore be obtained using (35).

2) Unstructured OTLS: The unstructured OTLS (UOTLS) problem in the context of the estimation problem of $\mathbf{B}$ is solved by putting

$$\mathbf{C} = \begin{bmatrix}x_1^T\\ \vdots\\ x_N^T\end{bmatrix}, \quad \mathbf{U} = \mathbf{B}^T, \quad \mathbf{D} = \begin{bmatrix}y_1^T\\ \vdots\\ y_N^T\end{bmatrix}. \tag{37}$$

Remark. *In the particular case of* $\mathbf{T}$ *model we can estimate independently* $\mathbf{Z}_l$ *and* $\mathbf{Y}_l$ *by splitting* (2) *into two independent equations such as:*

$$v^s - v^r = -\mathbf{Z}_l i^r, \tag{38}$$

$$i^s + i^r = \mathbf{Y}_l v^r. \tag{39}$$

We can equivalently write:

$$v^s - v^r = \mathbf{Z}_l i^s, \tag{40}$$

$$i^s + i^r = \mathbf{Y}_l v^s. \tag{41}$$

We can solve these 4 equations by unstructured TLS. We then obtain 4 estimates $\hat{\mathbf{Z}}_l^1, \hat{\mathbf{Z}}_l^2, \hat{\mathbf{Y}}_l^1, \hat{\mathbf{Y}}_l^2$. *In order to also incorporate symmetry in the estimated matrices we get the final estimates by:*

$$\hat{\mathbf{Z}}_l = \frac{(\hat{\mathbf{Z}}_l^1 + \hat{\mathbf{Z}}_l^2) + (\hat{\mathbf{Z}}_l^1 + \hat{\mathbf{Z}}_l^2)^T}{4} \tag{42}$$

$$\hat{\mathbf{Y}}_l = \frac{(\hat{\mathbf{Y}}_l^1 + \hat{\mathbf{Y}}_l^2) + (\hat{\mathbf{Y}}_l^1 + \hat{\mathbf{Y}}_l^2)^T}{4} \tag{43}$$

C. Weighted and Ordinary Least Squares

In the method of ordinary or weighted least squares, the noise on the right hand side of (24) is ignored, i.e., it is assumed that $\tilde{x} = x$. In this case, the classical least squares estimation formula for weighted least squares (WLS) yields the following estimate of θ:

$$\hat{\theta}_{wls} = \left(\mathbf{H}(\tilde{x})^T\mathbf{Q}_y^{-1}\mathbf{H}(\tilde{x})\right)^{-1}\mathbf{H}(\tilde{x})^T\mathbf{Q}_y^{-1}(\tilde{y} - \gamma(x)), \tag{44}$$

The ordinary least squares (OLS) estimate of θ is obtained by replacing $\mathbf{Q}_\mathbf{y}$ in the above by the identity matrix.

If the assumption $\tilde{x} = x$ holds, the estimations methods presented are optimal and are equivalent to maximum likelihood solution. However this assumption does not hold in practice and it is a-priori difficult to know what is its impact on the quality of the estimation. In the appendix we show that under EIV model WLS and OLS are asymptotically biased. We provide the derivation and expression of the asymptotic bias in V-B.

Remark. *For the* $\mathbf{T}$ *model, in addition to WLS methods we can derive another method which re-establishes the symmetry of the problem. Indeed WLS makes the assumption that the noise is only present on one side of the line which is not true in practice. In particular the problem has a symmetry which should be exploited, we have*

$$l^s = \bar{\mathbf{T}}(\theta^{\mathbf{T}})l^r \quad \textit{and} \quad l^r = \bar{\mathbf{T}}(\theta^{\mathbf{T}})l^s.$$

An iterative scheme where WLS is used to denoise the observation from one side and then use the denoised version of the

observation to estimate θ and also a denoised version of the observation of the other side and so on could help to make the assumption $\tilde{l}_i^r = l_i^r$ true such that the estimation could be improved. The pseudocode of this iterative scheme is given in Algorithm 1. In the rest of the thesis we name this method iterative WLS (IWLS).

Algorithm 1 IWLS

procedure $\theta_{IWLS}(\tilde{L}^r, \tilde{L}^s, \mathbf{Q_{L^s}}, \mathbf{Q_{L^r}})$
  $\hat{L}^r \leftarrow \tilde{L}^r$
  $\theta^1 \leftarrow [0 \dots 0]$
  $\theta^2 \leftarrow [0 \dots 0]$
  $d^s \leftarrow \text{DMAPPING}(\hat{L}^s, \hat{L}^r)$
  $d^r \leftarrow \text{DMAPPING}(\hat{L}^r, \hat{L}^s)$
  do
    $\theta_{sum} \leftarrow \theta^1 + \theta^2$
    $\mathbf{C}^r \leftarrow \text{CMAPPING}(\hat{L}^r)$
    $\theta^1 \leftarrow \left(\mathbf{C}^{r^T}\mathbf{Q}_{L^s}^{-1}\mathbf{C}^r\right)^{-1}\mathbf{C}^{r^T}\mathbf{Q}_{L^s}^{-1}d^s$
    $\hat{L}^s \leftarrow \mathbf{C}^r\theta$
    $\mathbf{C}^s \leftarrow \text{CMAPPING}(\hat{L}^s)$
    $\theta^2 \leftarrow \left(\mathbf{C}^{s^T}\mathbf{Q}_{L^r}^{-1}\mathbf{C}^s\right)^{-1}\mathbf{C}^{s^T}\mathbf{Q}_{L^r}^{-1}d^r$
    $\hat{L}^r \leftarrow \mathbf{C}^s\theta$
  while $\left|(\theta^1 + \theta^2) - \theta_{sum}\right| \geq crit$
  return $\frac{\theta^1+\theta^2}{2}$
end procedure
With $\text{CMAPPING}(L)$ $\mathbb{R}^{12N} \rightarrow \mathbb{R}^{12N\times\mathcal{S}_\theta}$ and $\text{DMAPPING}(L^s, L^r)$ $\mathbb{R}^{12N} \rightarrow \mathbb{R}^{12N}$ 2 procedures which respectively create $\mathbf{C}$ and d as in (36).

IV. Simulations

Realistic simulations based on data from real PMU measurements of a 125kV sub-transmission grid (see Figure 5) has been used to assess the performances of the proposed methods on different line models. As it can be observed from Figure 5, the grid contains 7 buses connected by 10 lines. The lines can be grouped into two classes. First, the underground coaxial lines (e.g. line 2), characterised by null terms for the non-diagonals of the shunt admittance matrix. Consequently these lines can be expressed with $\mathcal{S}_\theta = 15$. The second class contains overhead lines (e.g. line 10), resulting in the generic TL model with $\mathcal{S}_\theta = 18$. Another important property of the network's lines is that they are transposed which permits their modelling with $\mathcal{S}_\theta = 5$ or $\mathcal{S}_\theta = 6$ depending on if the line is underground or overhead. Hereunder the methodology used to create the dataset and the results of the different estimation methods applied to some lines of the grid are presented.

A. Methodology

The dataset used to compare the estimation methods contains one day of PMU measurements. The dataset contains current and voltage phasor measurements of both ends of all the lines. The PMUs measure currents and voltages at a frequency of 50 frames per second. This results in a maximum of $24 \times 3600 \times 50 \approx 4.3e6$ samples. The dataset also provides lines parameter values.

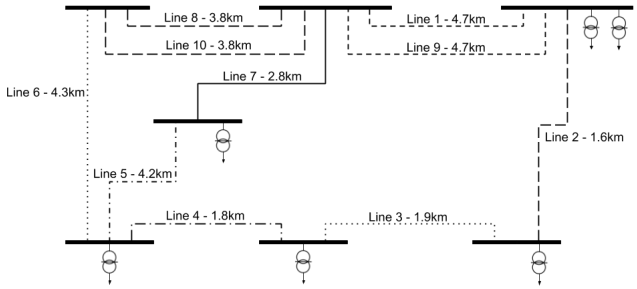

Fig. 5. Grid topology of the 125kV sub-transmission grid at Lausanne, Switzerland.

Although the value of the line parameter provided by the grid constructor may not be accurate, we assume them to be the ground truth to asses the accuracy of the estimation methods. For each line, using the line parameter and the measurements at one end of the line (either L^s or L^r) we generate artificial exact measurements of the other end of the line. We then add artificial noise as described in section II-C to generate realistic measurements. This procedure is described in Algorithm 2. The noise is generated using values from Table I.

Remark. *It is important to notice that depending on the line model used, the generation method could give rise to an unrealistic dataset. Indeed a very similar generation method could make use of the impedance ($\mathbf{Z}$) matrix or the admittance ($\mathbf{Y}$) matrix of the line instead of the transmittance ($\mathbf{T}$) matrix used. However the $\mathbf{Z}$ and $\mathbf{Y}$ matrix formulations are very sensitive to measurement noise. Table III presents a comparison of the average relative deviation of the samples artificially generated by the line models ($\mathbf{Z}$, $\mathbf{Y}$ or $\mathbf{T}$ matrix) from the samples given in the dataset. From the table we observe that the deviations are least when the transmittance matrix is used to model the line. The small deviations we observe for this case are due to measurement noise.*

TABLE III
Average relative deviation between samples generated and measurements.

Model	T	Z	Y
Average relative error	0.00773	0.0505	6.35

The methodology used to assess the performance of the proposed methods consists in 1) the samples generation, 2) the application of the estimation method(s) to these samples, and 3) then to assess the performance of the estimation method(s). The metrics used to evaluate the performance of an estimation method are the average relative error or the average relative error per component of the parameter vector. The first is mathematically defined as follows:

$$e_\theta = \frac{1}{\mathcal{S}_\theta}\sum_{i=1}^{\mathcal{S}_\theta}\left|\frac{\theta_i - \hat{\theta}_i}{\theta_i}\right| \times 100$$

and the latter is defined as follows:

$$e_{\theta_i} = \left| \frac{\theta_i - \hat{\theta}_i}{\theta_i} \right| \times 100 \quad \forall i \in [1, \mathcal{S}_\theta].$$

In the following, each experiment consists in repeating this methodology 10 times and averaging the results. To simplify our notations, in the following we use θ without a superscript to denote the parameter vector $\theta^{\mathbf{T}}$ of the $\mathbf{T}$ model unless specified otherwise.

Algorithm 2 Samples generation

procedure GENSAMPLES
  for *line_id* $= 1:10$ **do**
    $\mathbf{Z_l} \leftarrow$ GETZLINE(*line_id*)
    $\mathbf{Y_l} \leftarrow$ GETYLINE(*line_id*)
    for $n = 1:N$ **do**
      $\begin{bmatrix} v_n^r \\ i_n^r \end{bmatrix} \leftarrow$ GETMEASUREMENTS(*line_id*, N)
      $\begin{bmatrix} v_s^s \\ i_n^s \end{bmatrix} \leftarrow \begin{bmatrix} \mathbf{I} + \frac{\mathbf{Z_l Y_l}}{2} & -\mathbf{Z} \\ \mathbf{Y}_l \left(\mathbf{I} + \frac{\mathbf{Z_l Y_l}}{4}\right) - \left(\mathbf{I} + \frac{\mathbf{Z_l Y_l}}{2}\right) \end{bmatrix} \begin{bmatrix} v_n^r \\ i_n^r \end{bmatrix}$
      for $x = [v_n^r, v_n^s, i_n^r, i_n^s]$ **do**
        $\Delta_\rho \leftarrow \mathcal{N}(0, \frac{\alpha}{3}|x|)$
        $\tilde{\rho} \leftarrow |x| + \Delta_\rho$
        $\Delta_\phi \leftarrow \mathcal{N}(0, \frac{\beta}{3})$
        $\tilde{\phi} \leftarrow \angle_x + \Delta_\phi$
        $w = \tilde{\rho} e^{j\tilde{\phi}}$
      end for
      $\tilde{l}_n^r \leftarrow [\mathcal{R}\left((v_n^r)^T, (i_n^r)^T\right), \mathcal{I}\left((v_n^r)^T, (i_n^r)^T\right)]$
      $\tilde{l}_n^s \leftarrow [\mathcal{R}\left((v_n^r)^T, (i_n^r)^T\right), \mathcal{I}\left((v_n^r)^T, (i_n^r)^T\right)]$
    end for
  end for
  $\tilde{L}^r \leftarrow \left[\tilde{l}_1^r \ldots, \tilde{l}_N^r\right]^T$
  $\tilde{L}^s \leftarrow \left[\tilde{l}_1^s \ldots, \tilde{l}_N^s\right]^T$
  return $\tilde{L}^s$, $\tilde{L}^r$
end procedure

GETMEASUREMENTS(*line_id*, N) being a procedure which draws N measurements of a the line *line_id* uniformly at random without replacement from the dataset ($N < 4.2e6$).

B. Results

In this sub-section we first compare the different estimation methods in terms of computing time and accuracy. Based on these results we also compare different models discussed in this thesis and previous works.

We observed during our experiments that the error of OTLS methods was decreasing with the number of samples. However, the rate of decrease is very small, in particular even for 3 millions of samples we obtain the result of Table IV. This table presents the mean squared error (MSE) obtained with both unstructured and structured OTLS (SOTLS) and OLS. The methods are compared for the 2 types of lines with 3 millions of samples measured by class 1 IT. As it can be observed form the table, the OLS significantly outperforms the OTLS methods. The structured and unstructured OTLS completely fails at recovering the right parameters in the presence of noise. To find an accurate explanation why whether or not OTLS should work is difficult in this case because the assumption of homoscedasticity is not respected. Indeed all the theoretical results about the convergence of OTLS require homoscedasticity.

TABLE IV
MSE OF ESTIMATED θ FOR UOTLS, SOTLS AND OLS WITH IT OF CLASS 1 AND $N = 3 \times 10^6$.

Line	Coaxial	Non coaxial
$\mathcal{S}_\theta$	15	18
OLS	3.7014e-08	6.6537e-07
UOTLS	5.7140e-04	3.8823e+03
SOTLS	1.2709e+06	1.9371e+05

Table V presents the MSE of WLS and IWLS estimation methods in comparison with OLS. The estimation methods have been applied for sampleset of size $N = 3 \times 10^5$ and for IT of class 1. It can be observed from this table that WLS methods slightly outperforms OLS. Moreover IWLS does not improve the quality of the results for this number of samples.

TABLE V
MSE OF ESTIMATED θ FOR OLS, WLS AND IWLS WITH IT OF CLASS 1 AND $N = 3 \times 10^5$.

Line	Coaxial	Non coaxial
$\mathcal{S}_\theta$	15	18
OLS	2.1251e-05	1.7893e-05
WLS	7.5926e-07	8.6151e-07
IWLS	1.1170e-06	1.1213e-06

In Figure 6 we show the evolution of the average relative error as a function of the number of samples, N for OLS, WLS and WTLS. It can be observed from this figure that 1) all the estimation methods improve the quality of the estimation with the number of samples and 2) all the estimation methods tend to have similar accuracy with a large number of samples. To evaluate the robustness of these estimation methods we also analysed the variance of the estimation methods as a function of the number of samples. The variance of each component of the parameter vector is different. Therefore, we plot the component with the maximum value which is representative of the worst case performance of the estimation method. In Figure 7 we plot the maximum standard deviation as a function of the number of samples. It can be observed from this figure that 1) all the estimation methods have a decreasing worst case standard deviation with the increase in the number of samples and 2) all the estimation methods have similar behaviours in terms of worst case standard deviation. Same type of graphics are shown in the appendix V-D for other lines, the discussion is also valid. It is worth to notice that WTLS does not always

outperform the other methods. This can be explained by 1) the fact that there is no guarantee that the solution returned by fminunc is a global optimum and 2) for a finite number of samples there is no guarantee of the superiority of the ML estimate over other estimates. Moreover in the EIV model the effective parameter space increases at the same rate as the number of samples which is contrary to the standard least square models.

We also compared the estimation methods in terms of their computation time. In Table VI we present the computation time of each estimation method for different number, N, of samples. It can be observed that OLS outperforms the other estimation methods from this perspective. Indeed, on one side WLS requires the inverse of the covariance matrices to be computed and to be kept in the memory. And on the other side brute force approach requires gradient and function evaluation which has a complexity of $O(N)$. Although all these methods have linear complexity, in the subsequent experiments we keep only WLS method because it requires an acceptable amount of computation while being the most accurate for almost all the cases.

TABLE VI
AVERAGE COMPUTING TIME (IN SECONDS) OF THE MAIN ESTIMATION METHODS.

N° samples	WLS	OLS	WTLS
10^2	0.02	0.008	1.38
10^3	0.21	0.013	15.9
10^4	1.9	0.021	165
10^5	19	0.15	x
10^6	267	2.7	x

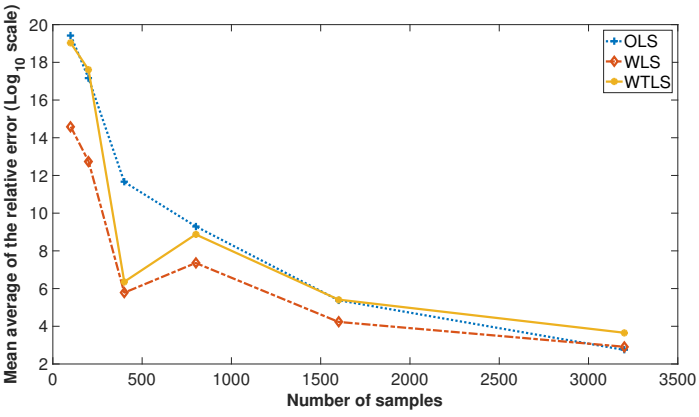

Fig. 6. Average relative error evolution for $\mathbf{T}$ line model.

A more accurate overview of the performance of the WLS method with $N = 3 \times 10^6$ is presented in Table VII for an overhead line ($\mathcal{S}_\theta = 18$) and an underground coaxial line ($\mathcal{S}_\theta = 15$). For this table and all the results presented in this section, the results of each experiment have been averaged over 10 experiments as said above. Although overhead lines are characterized by 18 unknown parameters ($\mathcal{S}_\theta = 18$), as described in the beginning of the section, the parameter vector provided by the grid constructor only contains 6 distinct values. We exploit this property to ease the visualization of

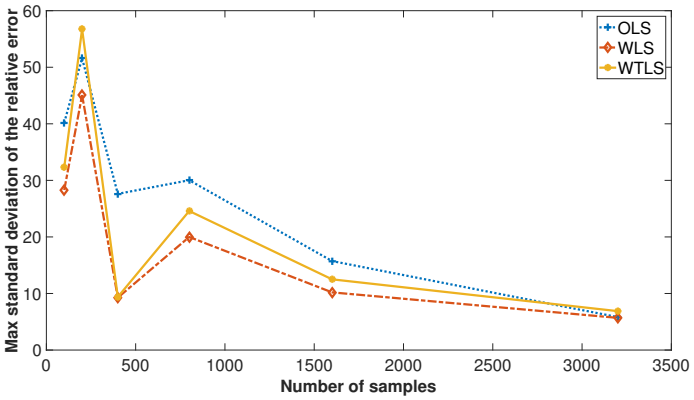

Fig. 7. Evolution of the maximun standard deviation of the relative error over θ parameters for $\mathbf{T}$ line model.

the results. In particular $\mathbf{Y_1}(\theta)$ and $\mathbf{Z_1}(\theta)$ can be expressed as in (45) and (46) with $\mathcal{S}_\theta = 6$. Similarly, the underground coaxial line has only 5 distinct values, where $\theta_6 = 0$ in (45). Although the estimation method is made with $\mathcal{S}_\theta = 18$ or $\mathcal{S}_\theta = 15$, we report the average relative error made by the parameters estimating the same value. For example, the fifth row of the Table VII is computed as the average value of $\frac{1}{3}(e_{\theta_1} + e_{\theta_2} + e_{\theta_3})$ over the 10 repetitions of the estimation process. For both lines we observe 1) that with weaker noise the accuracy of the estimation improves and 2) the estimation accuracy depends on the parameter component considered.

TABLE VII
ABSOLUTE RELATIVE ERROR (IN %) PER θ COMPONENT OF $\mathbf{T}$ LINE MODEL FOR 2 UNTRANSPOSED LINES AND 2 EXTREME CLASS OF IT - $N = 3 \times 10^6$.

Line	Coaxial ($\mathcal{S}_\theta = 15$)		Non coaxial ($\mathcal{S}_\theta = 18$)	
IT class	0.1	1	0.1	1
e_{θ_1}	28	59	2.8	12
e_{θ_2}	17	38	1.1	7.9
e_{θ_3}	58	120	8.1	25
e_{θ_4}	64	160	2.8	19
e_{θ_5}	0.004	0.01	24	25
e_{θ_6}	/	/	180	190

$$\mathbf{Y_1}(\theta) = \begin{bmatrix} \theta_5\, \theta_6\, \theta_6 \\ \theta_6\, \theta_5\, \theta_6 \\ \theta_6\, \theta_6\, \theta_5 \end{bmatrix} j \tag{45}$$

$$\mathbf{Z_1}(\theta) = \begin{bmatrix} \theta_1 + \theta_2 j\, \theta_3 + \theta_4 j\, \theta_3 + \theta_4 j \\ \theta_3 + \theta_4 j\, \theta_1 + \theta_2 j\, \theta_3 + \theta_4 j \\ \theta_3 + \theta_4 j\, \theta_3 + \theta_4 j\, \theta_1 + \theta_2 j \end{bmatrix} \tag{46}$$

We can further take advantage of this property of the lines considered to use a parameter space with a smaller size $\mathcal{S}_\theta$. Using a smaller parameter space we evaluate the sensitivity of the WLS method on different type of IT classes in Table VIII. It is interesting to compare this table with Table VII. Indeed the 2 observations made for Table VII are also valid for the Table VIII. However, unsurprisingly, it can be observed that

the quality of the estimation is much better when we use a smaller parameter space.

TABLE VIII
ABSOLUTE RELATIVE ERROR (IN %) PER θ COMPONENT OF $\mathbf{T}$ LINE MODEL FOR 2 TRANSPOSED LINES AND 2 EXTREME CLASS OF IT - $N = 3 \times 10^6$.

Line	Coaxial ($\mathcal{S}_\theta = 5$)		Non coaxial ($\mathcal{S}_\theta = 6$)	
IT class	0.1	1	0.1	1
e_{θ_1}	1.4	16	0.19	2.4
e_{θ_2}	2.2	6.3	0.17	0.87
e_{θ_3}	13	48	1.9	7.7
e_{θ_4}	18	54	0.73	3.1
e_{θ_5}	0.0033	0.01	17	25
e_{θ_6}	/	/	130	190

C. Comparison of line models: Fisher Information Matrix

In this section we use 2 metrics to compare the different line models. The first is theoretical and does not rely on the estimation method but only on the statistical model, it is called Cramer-Rao bound (CRB) and is defined through the Fisher Information Matrix (FIM). The FIM is a matrix which measures the quantity of information that some observations carry out about a parameter vector. The second metric is the per component relative error, although it relies on the estimation methods, we show that both CRB and per component relative error lead to the same conclusion.

Using the log-likelihood function it is possible to derive the FIM of the statistical model (see V-E for detailed derivation). It is given by:

$$\mathcal{I}(x,\theta) = \begin{bmatrix} \mathbf{Q}_x^{-1} + \bar{\mathbf{B}}(\theta)\mathbf{Q}_y^{-1}\bar{\mathbf{B}}(\theta)^T & \bar{\mathbf{B}}(\theta)\mathbf{Q}_y^{-1}\mathbf{H}(x)] \\ \mathbf{H}(x)^T\mathbf{Q}_y^{-1}\bar{\mathbf{B}}(\theta)^T & \mathbf{H}(x)^T\mathbf{Q}_y^{-1}\mathbf{H}(x). \end{bmatrix},$$

for which each term can be computed from the statistical model. The CRB can be derived from the FIM as follows:

$$Cov_{x,\theta}(T(\tilde{x},\tilde{y})) \succeq [\mathcal{I}(x,\theta)]^{-1} \tag{47}$$

where $T(\tilde{x},\tilde{y})$ is an unbiased estimator of $[x,\theta]$ and $A \succeq B$ denote that $A - B$ is positive semi definite. In particular it also gives a bound on the mean squared error of any unbiased estimator:

$$MSE(T(\tilde{x},\tilde{y})) = \operatorname{trace}\left(Cov_{x,\theta}(T(\tilde{x},\tilde{y}))\right) \tag{48}$$

$$\geq \operatorname{trace}\left([\mathcal{I}(x,\theta)]^{-1}\right) \tag{49}$$

As previously mentioned the transmittance ($\mathbf{T}$) matrix formulation is preferred for its greater robustness to noise. Hereunder we show the superiority of T matrix for short TL parameter estimation in comparison to impedance ($\mathbf{Z}$) and admittance ($\mathbf{Y}$) matrix formulation.

CRB cannot be applied directly to compare the performance of the different estimation methods defined previously because they are not unbiased. However it can be used to get insights on the quality of the different line models ($\mathbf{Z}$, $\mathbf{Y}$ and $\mathbf{T}$) for parameter estimation purpose. Table IX presents a comparison between the different line models in term of their CRB. Although the CRBs are respectively related to the line model parameters, it can still be concluded from it that theoretically the $\mathbf{T}$ formulation should be an easier estimation problem than the 2 other formulations. This can be explained because on one side the $\mathbf{T}$ formulation do the estimation of line impedance ($\mathbf{Z}_l$) and shunt admittance ($\mathbf{Y}_l$) independently. Whereas on the other sides the 2 other formulations shuffle the estimation of $\mathbf{Z}_l$ (or its inverse) and of $\mathbf{Y}_l$ (or its inverse). Moreover, as shown in Tables XII and XIII, the magnitudes of the parameters are very diverse and so these two models makes the parameter estimation problem more difficult. It is clear that for Y model there will be a length of line for which the order of magnitude of the parameters will be very close (because the small parameters increase whereas the others decrease) and so increasing the line length should improve the quality of the estimates. Indeed if we keep on lengthen the line the estimation based on $\mathbf{Y}$ reach an error minimum and then start increasing again. There is a similar effect on parameter of the $\mathbf{Z}$ model when the line is lengthen. However, probably because of the sensitivity of the formulation we see that the minimum error of $\mathbf{Z}$ is reached before the one of $\mathbf{Y}$ and is not very accurate in term of $\mathbf{Y}$ matrix reconstruction. Figure 8 shows the quality of estimation depending on the line length. It can be observed that $\mathbf{T}$ models become worse whereas the effect is opposite for $\mathbf{Z}$ and $\mathbf{Y}$.

TABLE IX
COMPARISON BETWEEN T, Z AND Y IN TERM OF CRB - $\mathcal{S}_\theta = 18$ AND $N = 1000$.

Model	T		Z		Y	
IT	.1	1	.1	1	.1	1
$\operatorname{tr}(\Sigma_{CR})$	2.0e-04	2.2e-02	8.2e+05	9.2e+07	8.5e+03	9.6e+05

Table X presents the empirical sensitivity analysis for an underground line with $\mathcal{S}_\theta = 18$. In order to be able to compare the 3 different formulations ($\mathbf{Z}$, $\mathbf{Y}$ and $\mathbf{T}$ matrix formulations) we mapped the estimated values to a common parameter space. The space of $\mathbf{Y}$ parameters ($\theta^{\mathbf{Y}}$) is chosen because it is the most commonly used for specifying line parameters.

Table X presents results for the 3 line models presented ($\mathbf{T}$, $\mathbf{Z}$ and $\mathbf{Y}$) with the best and the worst IT classes. These results confirm what was discussed earlier about the superiority of $\mathbf{T}$ formulation in comparison to $\mathbf{Z}$ and $\mathbf{Y}$ formulations for short TL.

Moreover we also compare our formulation with the one used in [13], [14], in [13] the authors proposed to use OLS with equations very similar to the ones given by the $\mathbf{Y}$ model of the line. In [14] the authors used the same objective function as in [13] but added constraint on the values of the parameters. Table XI presents the relative errors obtained by the method proposed in [13], it can be observed that for the type of lines considered the method proposed in [13] fails to find the right parameters. Even for the best class of IT the results approach 100% of relative error. Indeed the results of XI are very similar to the one obtained by the $\mathbf{Y}$ model. Di shi et al. in their work already noticed the poor quality of their methods for short TL. We also show in Figure 8 the impact of line length on the

quality of the reconstructed $\mathbf{Y}$ matrix, the metric used is the normalized MSE. This metric has 2 advantages, 1) it does not rely too much on the true value of the parameters (and so of the length) and 2) it gives a good insight if the reconstruction can be used for practical use or not.

TABLE X
ABSOLUTE RELATIVE ERROR (IN %) OBTAINED WITH $\mathbf{T}$, $\mathbf{Z}$ AND $\mathbf{Y}$ MODELS OF THE LINE - $\mathcal{S}_\theta = 18$ AND $N = 3e^6$.

Model	T		Z		Y	
IT	.1	1	.1	1	.1	1
e_{ψ_1}	4	7.2	34	150	80	99
e_{ψ_2}	14	21	36	320	67	99
e_{ψ_3}	1.8	2.8	12	97	80	99
e_{ψ_4}	5.5	7.7	43	120	69	99
e_{ψ_5}	24	25	1800	400	460	130
e_{ψ_6}	180	190	13000	3100	2900	160

TABLE XI
ABSOLUTE RELATIVE ERROR (IN %) OBTAINED WITH THE ESTIMATION METHOD PROPOSED IN [13] FOR $N = 3e^6$.

Line type	Coaxial		Overhead	
IT	.1	1	.1	1
$e_{\theta_1^{\mathbf{Y}}}$	99	100	76	99
$e_{\theta_2^{\mathbf{Y}}}$	100	100	62	99
$e_{\theta_3^{\mathbf{Y}}}$	99	100	77	99
$e_{\theta_4^{\mathbf{Y}}}$	97	100	61	99
$e_{\theta_5^{\mathbf{Y}}}$	230	73	240	590
$e_{\theta_6^{\mathbf{Y}}}$	/	/	6900	2700

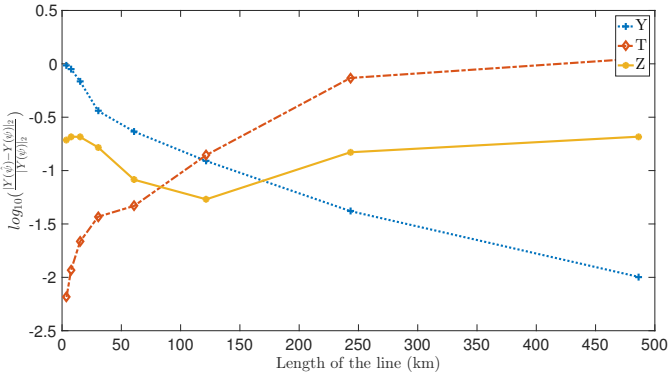

Fig. 8. Evolution of the quality of $\mathbf{Y}$ estimation with the length of the line for different modelization of the line.

TABLE XII
TRUE VALUE OF THE PARAMETERS OF COAXIAL LINE ($\mathcal{S}_\theta = 5$).

θ_1	2.8e-04	$\theta_1^{\mathbf{Y}}$	7.3e+02	$\theta_1^{\mathbf{Z}}$	2.8e-04
θ_2	5.2e-04	$\theta_2^{\mathbf{Y}}$	3.7e+02	$\theta_2^{\mathbf{Z}}$	5.2e-04
θ_3	1.2e-04	$\theta_3^{\mathbf{Y}}$	-1.3e+03	$\theta_3^{\mathbf{Z}}$	1.2e-04
θ_4	-1.2e-04	$\theta_4^{\mathbf{Y}}$	2.2e+02	$\theta_4^{\mathbf{Z}}$	-1.2e-04
θ_5	5.3e-02	$\theta_5^{\mathbf{Y}}$	5.3e-02	$\theta_5^{\mathbf{Z}}$	1.9e+01

TABLE XIII
TRUE VALUE OF THE PARAMETERS OF OVERHEAD LINE ($\mathcal{S}_\theta = 6$).

$\theta 1$	1.8e-03	$\theta_1^{\mathbf{Y}}$	8.6e+01	$\theta_1^{\mathbf{Z}}$	1.8e-03
$\theta 2$	5.1e-03	$\theta_2^{\mathbf{Y}}$	-2.7e+01	$\theta_2^{\mathbf{Z}}$	5.1e-03
$\theta 3$	6.3e-04	$\theta_3^{\mathbf{Y}}$	-2.3e+02	$\theta_3^{\mathbf{Z}}$	6.3e-04
$\theta 4$	2.1e-03	$\theta_4^{\mathbf{Y}}$	6.4e+01	$\theta_4^{\mathbf{Z}}$	2.1e-03
$\theta 5$	5.0e-03	$\theta_5^{\mathbf{Y}}$	5.0e-03	$\theta_5^{\mathbf{Z}}$	2.1e+02
$\theta 6$	-6.6e-04	$\theta_6^{\mathbf{Y}}$	-6.6e-04	$\theta_6^{\mathbf{Z}}$	3.2e+01

V. CONCLUSION

We presented methods to estimate parameters of three-phase untransposed lines based on different line models. From the comparison of the different methods and line models we can draw the following important conclusions: 1) The $\mathbf{T}$ model provides more accurate and robust estimates of the line parameters than the $\mathbf{Y}$ and $\mathbf{Z}$ models for short lines. However, we also observed that the accuracy and robustness of a particular model depends on the length of the line. In particular it may be beneficial to use the $\mathbf{Y}$ model for longer lines. 2) For a given line model and given length, different estimation methods give rise to different accuracy and have different computational complexities. In particular, for the $\mathbf{T}$ model we observe that the WLS method almost always marginally outperforms the other estimation methods in terms of accuracy and has a reasonable computational complexity. The WTLS method, although being asymptotically optimal, provides less accurate estimates than the WLS for finite number of samples and for the specific scenarios considered in the thesis. Moreover, it is prohibitive in terms of computational complexity because it requires to solve a non-convex optimization problem. The OLS method has a reasonable accuracy with the least computational complexity. Finally, we observe that the OTLS methods in most of the cases considered yield unusable estimates of the parameters because of slow convergence rates.

There are several interesting avenues for future research. We believe that combinations of different line models could give rise to better and more stable estimates of line parameters. Furthermore, state-of-the-art machine learning methods can also be envisioned to select the correct line model(s) to be used for specific estimation scenarios.

ACKNOWLEDGEMENTS

First and foremost I would like to express my deep acknowledgements to Prof. J-Y. Le Boudec who gave me the opportunity to work on this captivating subject in his laboratory. His guidance has been fundamental for me and he never failed to be present when I requested his help.

I wish also to emphasize my gratefulness to Phd. Arpan Mukhopadhyay for the many significant discussions about this work and for the great help he gave me to write this document. This work would not be the same without his essential advices.

I will not forget to express my best considerations to Prof. M. Paolone whose power systems expertise has been one of the main building block of this work.

I would like also to thank particularly my father for the many scientific and personal advices he provided me during these 5 last years.

Finally, I express my gratitude to all those who supported me during my studies, and who are not explicitly mentioned here.

APPENDIX

A. Noise derivation

Using Moment Generating Function (MGF) of $\tilde{\phi} \sim \mathcal{N}(\phi, \sigma_\phi)$ we have:

$$M_{\tilde{\phi}}(t) = E\left[e^{\tilde{\phi}t}\right] \tag{50}$$
$$= e^{\phi t + \frac{\sigma_\phi^2 t^2}{2}} \tag{51}$$

Which easily leads to:

$$E\left[\cos(\tilde{\phi})\right] = E\left[\mathcal{R}\left(e^{\tilde{\phi}j}\right)\right] \tag{52}$$
$$= \mathcal{R}\left(E\left[e^{\tilde{\phi}j}\right]\right) \tag{53}$$
$$= \mathcal{R}\left(e^{\phi j - \frac{\sigma_\phi^2}{2}}\right) \tag{54}$$
$$= \cos(\phi)e^{-\frac{\sigma_\phi^2}{2}} \tag{55}$$

and similarly to:

$$E\left[\sin(\tilde{\phi})\right] = E\left[\mathcal{I}\left(e^{\tilde{\phi}j}\right)\right] \tag{56}$$
$$= \mathcal{I}\left(E\left[e^{\tilde{\phi}j}\right]\right) \tag{57}$$
$$= \mathcal{I}\left(e^{\phi j - \frac{\sigma_\phi^2}{2}}\right) \tag{58}$$
$$= \sin(\phi)e^{-\frac{\sigma_\phi^2}{2}} \tag{59}$$

.

The expected value of the real and imaginary part of the noise can then be derived as follows:

$$E[\Delta_r] = E\left[\tilde{\rho}\cos(\tilde{\phi}) - \rho\cos(\phi)\right] \tag{60}$$
$$= \rho E\left[\cos(\tilde{\phi})\right] - \rho\cos(\phi) \tag{61}$$
$$= \rho\cos(\phi)e^{-\frac{\sigma_\phi^2}{2}} - \rho\cos(\phi) \tag{62}$$
$$= \rho\cos(\phi)\left(e^{-\frac{1}{2}\sigma_\phi^2} - 1\right) \tag{63}$$
$$E\left[\Delta_i\right] = E\left[\tilde{\rho}\sin(\tilde{\phi}) - \rho\sin(\phi)\right] \tag{64}$$
$$= \rho E\left[\sin(\tilde{\phi})\right] - \rho\cos(\phi) \tag{65}$$
$$= \rho\sin(\phi)e^{-\frac{\sigma_\phi^2}{2}} - \rho\cos(\phi) \tag{66}$$
$$= \rho\sin(\phi)\left(e^{-\frac{1}{2}\sigma_\phi^2} - 1\right) \tag{67}$$

.

Finally the second moment of the real and imaginary part of the noise can then be derived as follows:

$$E[\Delta_r^2] = E\left[\left(\tilde{\rho}\cos(\tilde{\phi}) - \rho\cos(\phi)\right)^2\right] \tag{68}$$
$$= E\left[\tilde{\rho}^2\cos(\tilde{\phi})^2 + \rho^2\cos(\phi)^2 - 2\tilde{\rho}\rho\cos(\tilde{\phi})\cos(\phi)\right] \tag{69}$$
$$= E\left[\tilde{\rho}^2\right]E\left[\cos(\tilde{\phi})^2\right] + \rho^2\cos(\phi)^2$$
$$- 2\rho\cos(\phi)E\left[\tilde{\rho}\right]E\left[\cos(\tilde{\phi})\right] \tag{70}$$

, where we have:

$$E\left[\tilde{\rho}^2\right] = Var\left[\tilde{\rho}\right] + E\left[\tilde{\rho}\right]^2 \tag{71}$$
$$= \frac{\alpha^2\rho^2}{9} + \rho^2 \tag{72}$$
$$= \rho^2\left(\frac{\alpha^2}{9} + 1\right) \tag{73}$$

$$E\left[\cos(\tilde{\phi})^2\right] = E\left[\frac{\cos(2\tilde{\phi}) + 1}{2}\right] \tag{74}$$
$$= \frac{1}{2}\left(1 + E\left[\cos(2\tilde{\phi}\right]\right) \tag{75}$$
$$= \frac{1}{2}\left(1 + E\left[\mathcal{R}(e^{2\tilde{\phi}j}\right]\right) \tag{76}$$
$$= \frac{1}{2}\left(1 + \mathcal{R}\left(E\left[e^{2\tilde{\phi}j}\right]\right)\right) \tag{77}$$
$$= \frac{1}{2}\left(1 + \mathcal{R}\left(e^{2\phi j - \frac{4\sigma_\phi^2}{2}}\right)\right) \tag{78}$$
$$= \frac{1}{2}\left(1 + \cos(2\phi)e^{-2\sigma_\phi^2}\right) \tag{79}$$

. Then, inserting (73), (79) and (55) in (70) we get:

$$E\left[\Delta_r^2\right] = \frac{1}{2}\left(1 + \frac{\alpha^2}{9}\right)\rho^2\left(1 + e^{-2\sigma_\phi^2}\cos(2\phi)\right)$$
$$+ \rho^2\cos^2(\phi)\left(1 - 2e^{(-\frac{1}{2}\sigma_\phi^2)}\right) \tag{80}$$

. In a very similar way we easily get that:

$$E\left[\Delta_i^2\right] = \frac{1}{2}\left(1 + \frac{\alpha^2}{9}\right)\rho^2\left(1 - e^{-2\sigma_\theta^2}\cos(2\theta)\right)$$
$$+ \rho^2 sin^2(\phi)\left(1 - 2e^{(-\frac{1}{2}\sigma_\phi^2)}\right) \tag{81}$$

Finally, by using MGF of θ and trigonometric formulas, the covariance between imaginary and real parts of the noise can be easily derived as follows:

$$E\left[\Delta_r\Delta_i\right] = E\left[\left(\tilde{\rho}\cos(\tilde{\phi}) - \rho\cos(\phi)\right)\left(\tilde{\rho}\sin(\tilde{\phi}) - \rho\sin(\phi)\right)\right] \tag{82}$$
$$= E\left[\tilde{\rho}\cos(\tilde{\phi})\tilde{\rho}\sin(\tilde{\phi}) - \tilde{\rho}\cos(\tilde{\phi})\rho\sin(\phi)\right.$$
$$\left.-\rho\cos(\phi)\tilde{\rho}\sin(\tilde{\phi}) + \rho\cos(\phi)\rho\sin(\phi)\right] \tag{83}$$
$$= E\left[\tilde{\rho}^2\right]\frac{1}{2}E\left[\sin(2\tilde{\phi})\right] - E\left[\tilde{\rho}\right]E\left[\cos(\tilde{\phi})\right]\rho\sin(\phi)$$
$$- E\left[\tilde{\rho}\right]E\left[\sin(\tilde{\phi})\right]\rho\cos(\phi) + \frac{\rho^2}{2}\sin(2\phi) \tag{84}$$
$$= \frac{1}{2}\sin(2\phi)\left[\left(1 + \frac{\alpha^2}{9}\right)\rho^2 e^{-2\sigma_\phi^2} - 2\rho^2 e^{-\frac{1}{2}\sigma_\phi^2} + \rho^2\right] \tag{85}$$

B. OLS bias under EIV model

To ease notation we denote $\mathbf{H}(\tilde{x})$ as $\tilde{\mathbf{H}}$ and $\tilde{d} = \tilde{y} - \gamma(\tilde{x})$. When the observations are assumed to be independent (44) becomes

$$\hat{\theta}_{ols} = \left(\tilde{\mathbf{H}}^T\tilde{\mathbf{H}}\right)^{-1}\tilde{\mathbf{H}}^T\tilde{d}$$

. This equation can be rewritten as follows:

$$\hat{\theta}_{ols} = \left(\sum_{i=1}^{N}\mathbf{H}(\tilde{x}_i)^T\mathbf{H}(\tilde{x}_i)\right)^{-1}\left(\sum_{i=1}^{N}\mathbf{H}(\tilde{x}_i)^T\tilde{d}_i\right)$$

, where

$$\tilde{d}_i = y_i - \gamma(x_i) + \Delta_{d_i} \quad \Delta_{d_i} \sim \mathcal{N}\left(0, Q_{d_i}\right)$$

. Under this model we can compute the asymptotic bias of the OLS estimator as follows:

$$\lim_{N\to\infty}\hat{\theta}_{ols} = \lim_{N\to\infty}\left(\sum_{i=1}^{N}\mathbf{H}(\tilde{x}_i)^T\mathbf{H}(\tilde{x}_i)\right)^{-1}\left(\sum_{i=1}^{N}\mathbf{H}(\tilde{x}_i)^T\tilde{d}_i\right) \quad (86)$$

$$= \lim_{N\to\infty}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{H}(\tilde{x}_i)^T\mathbf{H}(\tilde{x}_i)\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{H}(\tilde{x}_i)^T\tilde{d}_i\right) \quad (87)$$

By the continuous mapping theorem and Slutsky's theorem we have that, defining a_N and b_N to be two series of random variables:

$$a_N \xrightarrow{p} a \quad \text{and} \quad b_N \xrightarrow{p} b$$

. It follows by continuous mapping theorem that:

$$(a_N)^{-1} \xrightarrow{p} (a)^{-1}$$

. Moreover if a and b are two constant values, applying Slutsky's theorem gives:

$$(a_N)^{-1} b_N \xrightarrow{d} (a)^{-1} b$$

but the product $(a)^{-1} b$ is also a constant value and so

$$(a_N)^{-1} b_N \xrightarrow{p} (a)^{-1} b$$

.

If we redefine (without loss of generality) the $\mathbf{H} : \mathbb{R}^{12} \to \mathbb{R}^{12\times\mathcal{S}_\theta}$ mapping as:

$$\mathbf{H}(x)^T = [\bar{\mathbf{\Omega}}_1 x \dots \bar{\mathbf{\Omega}}_{12} x]$$

, where $x \in \mathbb{R}^{12}$ and $\bar{\mathbf{\Omega}}_j \in \mathbb{R}^{\mathcal{S}_\theta\times 12}, \forall j \in [1,12]$. If we also define $\mathbf{H}(x)_j = \bar{\mathbf{\Omega}}_j x, \forall j \in [1,12]$ and the associated noise vector $\Delta_{\mathbf{H}_j(x)} = \bar{\mathbf{\Omega}}_j\Delta_x \sim \mathcal{N}\left(0, \bar{\mathbf{\Omega}}_j Q_x \bar{\mathbf{\Omega}}_j^T\right)$. For simplicity we also define $Q_{\mathbf{H}_j(x_i)} = \bar{\mathbf{\Omega}}_j Q_x \bar{\mathbf{\Omega}}_j^T$ Using the previous result and the new defined variables we can define

$$a_N = \left(\frac{1}{12N}\sum_{i=1}^{N}\mathbf{H}\tilde{x}_i^T\mathbf{H}(\tilde{x}_i)\right) \quad (88)$$

$$= \frac{1}{12N}\sum_{i=1}^{N}\sum_{j=1}^{12}\bar{\mathbf{\Omega}}_j\tilde{x}_i\left(\tilde{x}_i\right)^T\bar{\mathbf{\Omega}}_j^T \quad (89)$$

$$= \frac{1}{12N}\sum_{i=1}^{N}\sum_{j=1}^{12}\Big(\bar{\mathbf{\Omega}}_j x_i {x_i}^T\bar{\mathbf{\Omega}}_j^T + \bar{\mathbf{\Omega}}_j\Delta_{x_i}\Delta_{x_i}^T\bar{\mathbf{\Omega}}_j^T + \bar{\mathbf{\Omega}}_j x_i\Delta_{x_i}^T\bar{\mathbf{\Omega}}_j^T + \bar{\mathbf{\Omega}}_j\Delta_{x_i}{x_i}^T\bar{\mathbf{\Omega}}_j^T\Big) \quad (90)$$

$$= \frac{1}{12N}\sum_{i=1}^{N}\sum_{j=1}^{12}\Big(\mathbf{H}_j(x_i)^T\mathbf{H}_j(x_i) + \mathbf{H}_j(x_i)^T\Delta_{\mathbf{H}_j(x_i)} + \Delta_{\mathbf{H}_j(x_i)}^T\mathbf{H}_j(x_i) + \Delta_{\mathbf{H}_j(x_i)}^T\Delta_{\mathbf{H}_j(x_i)}\Big) \quad (91)$$

$$\xrightarrow{p} \frac{1}{12N}\sum_{i=1}^{N}\sum_{j=1}^{12}\Big(\mathbf{H}_j(x_i)^T\mathbf{H}_j(x_i) + Q_{\mathbf{H}_j(x_i)}\Big) \quad (92)$$

This holds because by law of large numbers

$$\sum_{i=1}^{N}\sum_{j=1}^{12}\frac{\Delta_{\mathbf{H}_j(x_i)}}{12N} \xrightarrow{p} 0$$

and also thanks to Hoeffding bound on random sub-gaussian variables (proposition 2.1 in [22]). The latter gives:

$$\sum_{i=1}^{N}\sum_{j=1}^{12}\frac{\Delta_{\mathbf{H}_j(x_i)}^T\Delta_{\mathbf{H}_j(x_i)}}{12N} \xrightarrow{p} \sum_{i=1}^{N}\sum_{j=1}^{12}\frac{Q_{\mathbf{H}_j(x_i)}}{12N}$$

. And also define:

$$b_N = \left(\frac{1}{12N}\sum_{i=1}^{N}\mathbf{H}(\tilde{x}_i)^T\tilde{d}_i\right) \quad (93)$$

$$= \left(\frac{1}{12N}\sum_{i=1}^{N}\left(\mathbf{H}(x_i) + \Delta_{\mathbf{H}(x_i)}\right)^T\left(\mathbf{H}(x_i)\theta + \Delta_{d_i}\right)\right) \quad (94)$$

$$\xrightarrow{p} \frac{1}{12N}\sum_{i=1}^{N}\mathbf{H}(x_i)^T\left(\mathbf{H}(x_i)\theta\right) \quad (95)$$

. Defining the average covariance matrix of H matrix of i^{th} observation as

$$Q_{\mathbf{H}(x_i)} = \frac{1}{12}\sum_{j=1}^{12}Q_{\mathbf{H}_j(x_i)}$$

, the asymptotic bias of the OLS estimator is

$$\hat{\theta}_{ols} \xrightarrow{p} \left(\sum_{i=1}^{N}\left(\mathbf{H}(x_i)^T\mathbf{H}(x_i) + Q_{\mathbf{H}(x_i)}\right)\right)^{-1}\sum_{i=1}^{N}\mathbf{H}(x_i)^T\mathbf{H}(x_i)\theta \quad (96)$$

$$\xrightarrow{p} \left(\mathbf{I}_{\mathcal{S}_\theta} - \left(\sum_{i=1}^{N}\mathbf{H}(l_i^r)^T\mathbf{H}(x_i) + Q_{\mathbf{H}(x_i)}\right)^{-1}\sum_{i=1}^{N}Q_{\mathbf{H}(x_i)}\right)\theta \quad (97)$$

.

If we define $\mathbf{H}_{wls}(x_i)$ and $Q^{-1}_{\mathbf{H}_{wls}(x_i)}$ as follow:

$$\mathbf{H}_{wls}(x_i) = Q_{x_i}^{-1/2}\mathbf{H}(x_i) \quad (98)$$

$$Q^{-1}_{\mathbf{H}_{wls}(x_i)} = \bar{\mathbf{\Omega}}_j Q_{x_i}^{-1/2} Q_{x_i} {Q_{x_i}^{-1/2}}^T\bar{\mathbf{\Omega}}_j^T \quad (99)$$

Then we can also easily show that the bias of the WLS estimation methods under EIV model is equal to

$$\hat{\theta}_{ols} \xrightarrow{p} \left(\sum_{i=1}^{N}\left(\mathbf{H}_{wls}(x_i)^T\mathbf{H}_{wls}(x_i) + Q_{\mathbf{H}_{wls}(x_i)}\right)\right)^{-1}\left(\sum_{i=1}^{N}\mathbf{H}_{wls}(x_i)^T\mathbf{H}_{wls}(x_i)\theta\right) \quad (100)$$

$$\xrightarrow{p} \left(\mathbf{I}_{\mathcal{S}_\theta} - \left(\sum_{i=1}^{N}\mathbf{H}_{wls}(x_i)^T\mathbf{H}(x_i) + Q_{\mathbf{H_{wls}}(x_i)}\right)^{-1}\sum_{i=1}^{N}Q_{\mathbf{H}_{wls}(x_i)}\right)\theta \quad (101)$$

.

C. Fisher Information Matrix

The derivation of the fisher information matrix of the line's parameters estimation problem can be computed as follows:

$$\mathcal{I}(\theta, x) = \begin{bmatrix} E[\nabla_x\nabla_x^T] & E[\nabla_x\nabla_\theta^T] \\ E[\nabla_\theta\nabla_x^T] & E[\nabla_\theta\nabla_\theta^T] \end{bmatrix}$$

Where ∇_x stands for $\nabla_x \log(f(\tilde{x}, \tilde{x}|\theta, x))$ and ∇_θ stands for $\nabla_\theta \log(f(\tilde{x}, \tilde{y}|\theta, x))$. It can be shown that

$$\nabla_x log\left(f(\tilde{x}, \tilde{y}|\theta, x)\right) = Q_x^{-1}(\tilde{x} - x) + \bar{\mathbf{B}}(\theta)Q_y^{-1}\left(\tilde{y} - \bar{\mathbf{B}}(\theta)x\right) \quad (102)$$

$$\nabla_\theta log\left(f(\tilde{x}, \tilde{y}|\theta, x)\right) = \mathbf{H}(x)^T Q_y^{-1}\left(\tilde{y} - \mathbf{H}(x)\theta - \gamma(x)\right) \quad (103)$$

We can express these expressions in the following way:

$$\nabla_x log\left(f(\tilde{x},\tilde{y}|\theta,x)\right) = Q_x^{-1}(\tilde{x}-E\left[\tilde{x}\right]) + \bar{\mathbf{B}}(\theta)Q_y^{-1}\left(\tilde{y}-E\left[\tilde{y}\right]\right) \quad (104)$$

$$\nabla_\theta log\left(f(\tilde{x},\tilde{y}|\theta,x)\right) = \mathbf{H}(x)^T Q_y^{-1}\left(\tilde{y}-E\left[\tilde{y}\right]\right) \quad (105)$$

Because $\tilde{x}$ and $\tilde{y}$ are independent random variables, the following equalities hold:

$$\begin{aligned} E\left[\nabla_x \nabla_x^T\right] =& E\left[Q_x^{-1}\left(\tilde{x}-E[\tilde{x}]\right)\left(\tilde{x}-E[\tilde{x}]\right)^T {Q_x^{-1}}^T + \right. \\ & \bar{\mathbf{B}}(\theta)Q_y^{-1}\left(\tilde{y}-E\left[\tilde{y}\right]\right) \\ & \left.\left(\bar{\mathbf{B}}(\theta)Q_y^{-1}\left(\tilde{y}-E\left[\tilde{y}\right]\right)\right)^T\right] \quad (106) \\ =& Q_x^{-1} + \bar{\mathbf{B}}Q_y^{-1}\bar{\mathbf{B}}(\theta)^T \quad (107) \end{aligned}$$

. We can obtain the other term similarly, it finally yields:

$$\mathcal{I}(\theta,x) = \begin{bmatrix} Q_x^{-1} + \bar{\mathbf{B}}Q_y^{-1}\bar{\mathbf{B}}^T & \bar{\mathbf{B}}Q_y^{-1}\mathbf{H}(x)] \\ \mathbf{H}(x)^T Q_y^{-1}\bar{\mathbf{B}}^T & \mathbf{H}(x)^T Q_y^{-1}\mathbf{H}(x) \end{bmatrix}$$

.

D. Estimation Methods Comparison

Fig. 9. Relative errors of different estimation methods

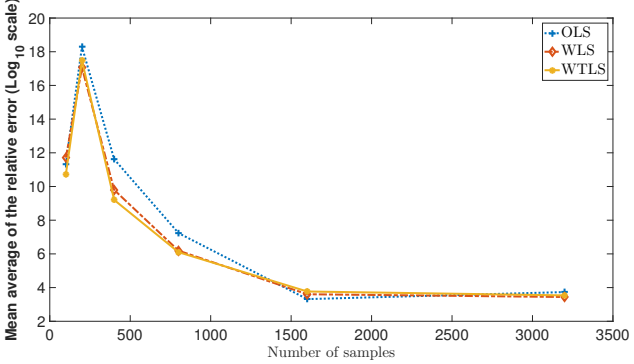

Fig. 10. $\mathcal{S}_\theta = 15$ - IT class = 0.1

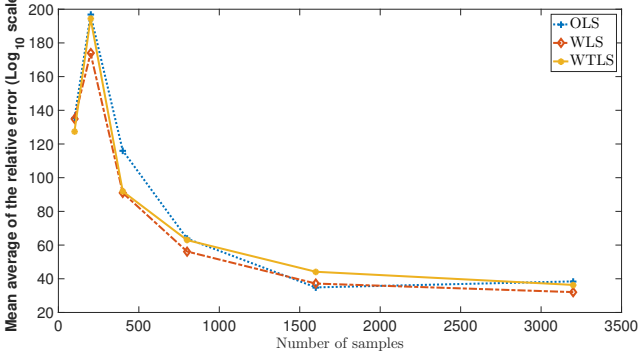

Fig. 11. $\mathcal{S}_\theta = 18$ - IT class = 1

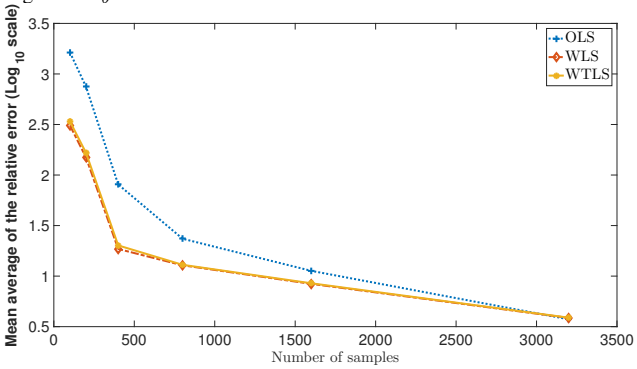

Fig. 12. $\mathcal{S}_\theta = 15$ - IT class = 0.1

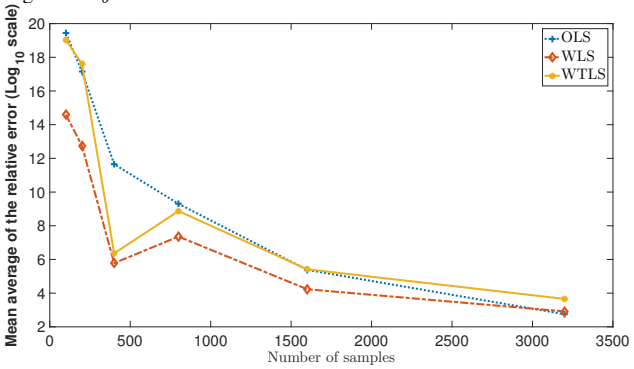

Fig. 13. $\mathcal{S}_\theta = 18$ - IT class = 1

E. Cramer-Rao Bound Comparison

TABLE XIV
COMPARISON BETWEEN T, Z AND Y IN TERM OF CRB - $\mathcal{S}_\theta = 15$ AND $N = 1000$

Model	T		Z		Y	
IT	.1	1	.1	1	.1	1
$\mathrm{tr}(\Sigma_{CR})$	3.2e-05	3.5e-03	3.6e-05	4.0e-03	2.5e+08	2.7e+10

TABLE XV
COMPARISON BETWEEN T, Z AND Y IN TERM OF CRB - $\mathcal{S}_\theta = 6$ AND $N = 1000$

Model	T		Z		Y	
IT	.1	1	.1	1	.1	1
$\mathrm{tr}(\Sigma_{CR})$	3.3e-06	3.7e-04	1.7e+04	1.9e+06	4.9e+01	5.5e+03

TABLE XVI
COMPARISON BETWEEN T, Z AND Y IN TERM OF CRB - $\mathcal{S}_\theta = 5$ AND $N = 1000$

Model	T		Z		Y	
IT	.1	1	.1	1	.1	1
$\mathrm{tr}(\Sigma_{CR})$	2.7e-07	3.1e-05	6.8e-07	7.9e-05	2.1e+06	2.4e+08

F. Coefficient pseudocode

Algorithm 3 $\mathbf{H_T}$ mapping

procedure $\mathbf{H_T}(l)$
  for $i = 1 : 18$ **do**
    $\theta \leftarrow e_i^{(18)}$
    for $j = 1 : 12$ **do**
      $l \leftarrow e_j^{(12)}$
      $\mathbf{\Omega}_i(:, j) \leftarrow \bar{\mathbf{T}}(\theta) l - \gamma(l)$
    end for
    $\mathbf{H_T}(l)(:, i) \leftarrow \mathbf{\Omega}_i l$
  end for
  return $\mathbf{H_T}(l)$
end procedure

Remark. *The pseudo code can be very simply generalized for* $\mathbf{Z}$ *and* $\mathbf{Y}$ *line models.*

G. $\mathbf{T}, \mathbf{Y}, \mathbf{Z}$ comparison

TABLE XVII
COMPARISON BETWEEN T, Z AND Y IN TERM OF RELATIVE ERROR (IN %) - $\mathcal{S}_\theta = 6$ AND $N = 3e^6$

Model	T		Z		Y	
IT	.1	1	.1	1	.1	1
$e_{\theta_1^{\mathbf{Y}}}$	0.6	1.4	2.3	80	82	99
$e_{\theta_2^{\mathbf{Y}}}$	0.81	2.1	5.5	51	72	99
$e_{\theta_3^{\mathbf{Y}}}$	0.62	0.52	0.73	22	83	99
$e_{\theta_4^{\mathbf{Y}}}$	1.4	2.5	1.6	36	70	99
$e_{\theta_5^{\mathbf{Y}}}$	17	25	1800	390	7000	630
$e_{\theta_6^{\mathbf{Y}}}$	130	190	13000	3000	52000	4100

TABLE XVIII
COMPARISON BETWEEN T, Z AND Y IN TERM OF RELATIVE ERROR (IN %) - $\mathcal{S}_\theta = 15$ AND $N = 3e^6$

Model	T		Z		Y	
IT	.1	1	.1	1	.1	1
$e_{\theta_1^{\mathbf{Y}}}$	24	42	67	96	100	100
$e_{\theta_2^{\mathbf{Y}}}$	59	86	140	100	100	100
$e_{\theta_3^{\mathbf{Y}}}$	17	31	18	100	99	100
$e_{\theta_4^{\mathbf{Y}}}$	73	130	130	100	98	100
$e_{\theta_5^{\mathbf{Y}}}$	0.004	0.001	0.0037	0.18	76	87

TABLE XIX
COMPARISON BETWEEN T, Z AND Y IN TERM OF RELATIVE ERROR (IN %) - $\mathcal{S}_\theta = 5$ AND $N = 3e^6$

Model	T		Z		Y	
IT	.1	1	.1	1	.1	1
$e_{\theta_1^{\mathbf{Y}}}$	5.7	32	66	140	100	100
$e_{\theta_2^{\mathbf{Y}}}$	11	69	130	59	100	100
$e_{\theta_3^{\mathbf{Y}}}$	3.7	13	13	130	99	100
$e_{\theta_4^{\mathbf{Y}}}$	16	73	100	190	98	100
$e_{\theta_5^{\mathbf{Y}}}$	0.0033	0.01	0.0028	0.16	53	38

REFERENCES

[1] A. Bernstein, L. Reyes-Chamorro, J.-Y. Le Boudec, and M. Paolone, "A composable method for real-time control of active distribution networks with explicit power setpoints. part I: Framework," *Electric Power Systems Research*, vol. 125, pp. 254–264, 2015.

[2] Q. Gemine, E. Karangelos, D. Ernst, and B. Cornélusse, "Active network management: planning under uncertainty for exploiting load modulation," in *Bulk Power System Dynamics and Control-IX Optimization, Security and Control of the Emerging Power Grid (IREP), 2013 IREP Symposium*, pp. 1–9, IEEE, 2013.

[3] A. Borghetti, M. Bosetti, S. Grillo, S. Massucco, C. A. Nucci, M. Paolone, and F. Silvestro, "Short-term scheduling and control of active distribution systems with high penetration of renewable resources," *IEEE Systems Journal*, vol. 4, no. 3, pp. 313–322, 2010.

[4] L. F. Ochoa, C. J. Dent, and G. P. Harrison, "Distribution network capacity assessment: Variable dg and active networks," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 87–95, 2010.

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

[6] K. Christakou, J.-Y. LeBoudec, M. Paolone, and D.-C. Tomozei, "Efficient computation of sensitivity coefficients of node voltages and line currents in unbalanced radial electrical distribution networks," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 741–750, 2013.

[7] L. G. Perez and A. J. Urdaneta, "Optimal computation of distance relays second zone timing in a mixed protection scheme with directional overcurrent relays," *IEEE Transactions on Power Delivery*, vol. 16, no. 3, pp. 385–388, 2001.

[8] J. D. Glover, M. S. Sarma, and T. Overbye, *Power System Analysis & Design, SI Version*. Cengage Learning, 2012.

[9] A. Egozi, "Subscriber line impedance measurement device and method," Nov. 7 1995. US Patent 5,465,287.

[10] K. Dasgupta and S. Soman, "Line parameter estimation using phasor measurements by the total least squares approach," in *Power and Energy Society General Meeting (PES), 2013 IEEE*, pp. 1–5, IEEE, 2013.

[11] K. Dasgupta and S. Soman, "Line parameter estimation using phasor measurements by the total least squares approach," in *Power and Energy Society General Meeting (PES), 2013 IEEE*, pp. 1–5, IEEE, 2013.

[12] X. Zhao, H. Zhou, D. Shi, H. Zhao, C. Jing, and C. Jones, "On-line pmu-based transmission line parameter identification," *CSEE Journal of Power and Energy Systems*, vol. 1, no. 2, pp. 68–74, 2015.

[13] D. Shi, D. J. Tylavsky, K. M. Koellner, N. Logic, and D. E. Wheeler, "Transmission line parameter identification using pmu measurements," *International Transactions on Electrical Energy Systems*, vol. 21, no. 4, pp. 1574–1588, 2011.

[14] H. Zhou, X. Zhao, D. Shi, H. Zhao, and C. Jing, "Transmission line sequence impedances identification using pmu measurements,"
[15] D. Shi, D. J. Tylavsky, N. Logic, and K. M. Koellner, "Identification of short transmission-line parameters from synchrophasor measurements," in *Power Symposium, 2008. NAPS'08. 40th North American*, pp. 1–8, IEEE, 2008.
[16] S. Azizi, M. Sanaye-Pasand, and M. Paolone, "Locating faults on untransposed, meshed transmission networks using a limited number of synchrophasor measurements," *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4462 – 4472, 2016.
[17] C. R. Paul, *Analysis of multiconductor transmission lines*. John Wiley & Sons, 2008.
[18] S. Sarri, "Methods and performance assessment of pmu-based real-time state estimation of active distribution networks," 2016.
[19] M. Paolone, J.-Y. Le Boudec, S. Sarri, and L. Zanni, "Static and recursive pmu-based state estimation processes for transmission and distribution power grids," *Advanced Techniques for Power System Modelling, Control and Stability Analysis*, pp. 189–239, 2015.
[20] *The Total Least Squares Problem: Computational Aspects and Analysis*. SIAM, 1991.
[21] I. Markovsky and S. V. Huffel, "Overview of total least-squares methods," *Signal Processing*, vol. 87, no. 10, pp. 2283 – 2302, 2007. Special Section: Total Least Squares and Errors-in-Variables Modeling.
[22] M. Wainwright, *Basic tail and concentration bounds 2*.