ADAPTATION AND INSTALLATION OF A ROBUST STATE ESTIMATION PACKAGE IN THE EEF UTILITY

by

Michael A. Chapman

Thesis submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE
in the Bradley Department of
Electrical and Computer Engineering

Lamine M. Mili (chair)
Robert P. Broadwater
Hugh F. VanLandingham

February 8, 1999
Blacksburg, Virginia

Keywords: Power systems, Robust state estimation,
EEF Utility, Real-time installation, SHGM-estimator

ADAPTATION AND INSTALLATION OF A ROBUST STATE ESTIMATION PACKAGE IN THE EEF UTILITY

Michael A. Chapman

(ABSTRACT)

Robust estimation methods have been successfully applied to the problem of power system state estimation in a real-time environment. The Schweppe-type GM-estimator with the Huber psi-function (SHGM) has been fully installed in conjunction with a topology processor in the EEF utility, headquartered in Fribourg, Switzerland.

Some basic concepts of maximum likelihood estimation and robust analysis are reviewed, and applied to the development of the SHGM-estimator. The algorithms used by the topology processor and state estimator are presented, and the superior performance of the SHGM-estimator over the classic weighted least squares estimator is demonstrated on the EEF network. The measurement configuration of the EEF network has been evaluated, and suggestions for its reinforcement have been proposed.

ACKNOWLEDGEMENTS

The author wishes to acknowledge Dr. Lamine Mili for his financial support and academic advice. Appreciation is also expressed to Dr. Robert P. Broadwater and Dr. Hugh F. VanLandingham for serving as committee members, to Professor Alain Germond for attending and contributing to the defense, to the Ecole Polytechnique Fédérale de Lausanne for participation in the exchange program, and to the Entreprises Electriques Fribourgeoises for their financial support and collaboration in the project.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS vi
LIST OF TABLES vii

1 – INTRODUCTION 1

1.1. Power System State Estimation 1
1.2. Developments in State Estimation 4
1.3. Personal Contribution 6

2 – MAXIMUM LIKELIHOOD ESTIMATION 11

2.1. Definitions 11
2.2. Fisherian Concepts 12

3 – ROBUSTNESS ANALYSIS 16

3.1. Robustness Concepts 17
3.2. M-estimators 20
3.3. Leverage Points 24
3.3.1. Linearization of the Regression Model 25
3.3.2. Projection Statistics 26

4 – ALGORITHMS 29

4.1. Topology Processor 29
4.2. Updates to the SHGM-Estimator Code 34
4.3. Projection Statistics 36
4.4. Iteratively Reweigthed Least Squares (IRLS) 37
4.5. Overall SHGM-IRLS Algorithm 38

5 – SIMULATION RESULTS 41

5.1. EEF System Model 41
5.2. Observability and Local Redundancy 46
5.3. Weighted Least Squares 48

5.4. Example of a Real-Time Case 51
5.5. SHGM and the Identification of Multiple Bad Data 65
5.6. Meter Placement 68

6 – CONCLUSION **71**

7 – REFERENCES **72**

8 – APPENDIX **76**

8.1. Power system modeling 76
8.2. Regression model in power systems 78
8.3. Weighted Least Squares 80
8.4. Formation of the IRLS Jacobian 81
8.5. Equivariance Properties 84
8.6. Probability Distributions 84
8.7. Variance Calculations 86

VITA **89**

LIST OF ILLUSTRATIONS

Figure 4.1.1 – Environment of the topology processor and state estimator 30
Figure 4.1.2 – Flowchart describing the process followed by the topology processor 31
Figure 4.2.1 – π-circuit transformer model 35
Figure 5.1.1 – Monteynan island 43
Figure 5.1.2 – Botterens island 44
Figure 5.1.3 – Villarepos island 44
Figure 5.2.1 – Example 5-bus system 47
Figure 5.2.2 – Example 5-bus system with measurement (4) deleted 47
Figure 5.4.1 – Monteynan island topology in real-time case 53
Figure 5.4.3 – Botterens island topology in real-time case 61
Figure 5.6.1 – Topological details of interest in the Monteynan and Hauterive substations 69
Figure 8.1.1 – Nominal π-circuit model of a transmission line 77

LIST OF TABLES

Table 3.2.1 – ρ- and ψ-functions of some M-estimators 21

Table 3.2.2 – Performance of the mean, median and Huber at varying distributions 23

Table 5.1.1 – Projection statistics of active power flow measurements. 45

Table 5.3.1 – Weighted residuals provided by the WLS estimator. 50

Table 5.3.2 – Smearing effect and breakdown of the WLS estimator 51

Table 5.4.1 – Monteynan island active power flows 54

Table 5.4.2 – Monteynan island reactive power flows 55

Table 5.4.3 – Monteynan island power injections 56

Table 5.4.4 – Monteynan island voltages 57

Table 5.4.5 – Botterens island power flows 62

Table 5.4.6 – Botterens island power injections 62

Table 5.4.7 – Botterens island voltages 63

Table 5.4.8 – Villarepos island power flows 64

Table 5.4.9 – Villarepos island power injections 64

Table 5.4.10 – Villarepos island voltages 64

Table 5.5.1 – SHGM-estimator results for the bad data 66

Table 5.5.2 – Weighted residuals provided by the SHGM-estimator. 67

1 – INTRODUCTION

1.1. Power System State Estimation

Modern electric power systems are enormous in scale, typically spanning continents, and provide service for hundreds of millions of customers. These systems consist of numerous interconnected networks that contain various types of generators and consumers of electrical energy, which are also interconnected by high voltage equipment such as transmission lines and transformers. Each of these networks is operated and maintained by companies who are responsible for the consistently secure and economic operation of their network, as well as for the reliability of the larger power system. The famous 1965 blackout in the Northeastern United States awakened the power community and the nation to the fact that their means of providing this service would require the implementation of more thorough and advanced techniques.

Prior to this, power system operators working at the control centers had only a minimal amount of information and controls available to achieve these objectives. The only information received was that which is essential to controlling the real-time network's basic operation, such as system frequency, breaker statuses, and a minimal set of active power measurements. [12] In the late 1960's, a more advanced computer-aided system was finally devised for operators that allowed for the monitoring of switching device statuses and electric quantities such as power flows, power injections, and node voltage magnitudes, as well as logs of data and events, alarms, and graphic displays. This system is generally referred to as supervisory control and data acquisition, or SCADA. In addition to the vast amount of system information made available, research led to the development of many analytical programs that allowed for a broad scope of analysis, control, and security functions.

Originally engineers thought that the newly developed analytical tools could make direct use of the measurements provided by the SCADA, but this turned out to be quite

wrong. These functions require a coherent picture of the state of the network as a starting point, such as what is provided by a load flow. However, no real-time load flow existed, and certainly not one that dealt with all the uncertainties of a system operating in real-time. In 1969 Fred Schweppe introduced the idea of using the redundant number of measurements made available by the SCADA to statistically determine the state of the network. His proposition, the state estimator, was eventually accepted and serves the essential role of a real-time load flow. The methods described by Schweppe et al. in [34] have served as a basis for static state estimation since then. He also helped to standardize the terminology used in state estimation, as in [35].

The introduction of state estimation created a shift in the approach of control center design, transitioning from the SCADA to an environment that incorporates real-time system metering and operation with on- and off-line analytical, control, and security functions, the well-known energy management system (EMS). Schweppe likened the state estimator to a digestive system that removes impurities from the nourishment of the SCADA measurements, providing a clean flow of life giving data to the control center's man/machine brain. This brain is the EMS environment, which has been classically defined. The goal of control center design is security control under the three states of power system operation: the normal, emergency, and restorative states. There are several functions performed by a control center in the classical EMS environment, and the most difficult of these to implement are those that run in a real-time environment. The key functions include state estimation, security monitoring, on-line load flow, security analysis, supervisory control, automatic generation control, automatic voltage/VAR control, and economic dispatch control [11], [10]. These functions interact in an often complex manner, but all are aimed at providing the system operator with a coherent view of the system and/or carrying out the operator's decisions. Since all of these functions are directly dependent on state estimation, it is essential that the system operators trust the results it provides. The classical model as discussed above assumes steady-state functions; an introductory overview of dynamic-state functions is provided by [29].

Due to present and coming changes in the deregulated energy market, today's EMS is in a state of transition. Transmission and generation will be ungrouped from one another, and a probable mechanism for energy access is the "Independent System Operator" (ISO). Among the ISO's many responsibilities is the assurance of secure and efficient operation for the systems in its market. [36] This will require an EMS that monitors much larger and more complex systems, probably encompassing several thousands of buses. The requirements of the state estimator and other functions for such an EMS will be increasingly challenging.

Since power systems are dynamic, real-time systems, the question has to be asked, why not employ a dynamic state estimation? In fact, Schweppe et al. [34] originally considered that dynamic state estimation would probably replace static state estimation in the future. Instead, static state estimation has received considerably more attention and is chosen for implementation almost to the exclusion of dynamic state estimation. There are at least four reasons why: static state estimation presently fills the control center needs; the objectives of dynamic state estimation are unclear; there are many difficulties in determining the dynamic system model; and dynamic state estimation is computationally intensive. [30] Of these, the difficulties of defining a tractable, reliable model of the dynamic power system is the biggest inhibitor, due to the highly unpredictable and nonlinear nature of power systems.

Regardless of the choice of a static or dynamic state estimator, or whether it operates in the EMS of a small network or large ISO, its function is essentially unchanged. The state estimator plays the essential role of a purifier, creating a complete and reliable database for security monitoring, security analysis and the various controls of a power system. The state estimator thus employs statistical methods to act as a tunable filter between the field data measurements and security and control functions. To this end, a static state estimator requires as input a set of real-time measurements that redundantly observes the system state along with a topology processor that determines the system topological model based on the telesignalings of circuit breaker statuses.

1.2. Developments in State Estimation

Although the role of a state estimator is clear, there is much freedom of choice in its practical implementation. One of the most notable and important options is that of the statistical methodology used to purify the measured data. The power system model is well defined, and a redundant measurement set is made available to determine its operating state, but what can be used to filter measurement noise and correct the influence of bad data measurements, that would otherwise corrupt or give false impressions of the system's state? In his original work in 1969 and 1970, Schweppe proposed the classic quadratic weighted least-squares (WLS) estimator based on a linear regression to produce the state estimate. He also discussed, without elaboration, methods to detect the presence of and to identify bad data. [34] Schweppe understood the WLS's inherent sensitivity to bad data, and so in 1971 he and Merrill [21] proposed a modification of the WLS objective function, the so-called "quadratic-square root", or simply Merrill-Schweppe, estimator. This work represented the introduction into power system state estimation of non-quadratic estimators, which use WLS estimation for small residuals but are robustified against bad data by "downweighting" large residuals. The WLS and non-quadratic estimators belong to the M-estimator class presented in Section 3.2.

Alternative methods of bad data analysis were also developed in connection with the WLS solution, notably the ordered residual analysis and grouped residual analysis. Ordered residual analysis involves elimination of the largest weighted or normalized residual when bad data are detected after performing a WLS estimate, then repeating the estimation-elimination process until no bad data are detected. Grouped residual analysis is similar, except that a candidate group of large weighted or normalized residuals are removed and then replaced one at a time, obtaining the WLS solution after each replacement, until a bad data is once again detected. These methods were compared with various non-quadratic estimators by Handschin, Schweppe, et al. in [18].

The WLS estimator with ordered residual analysis, termed identification by elimination (IBE), and some non-quadratic criteria (NQC) estimators were compared again in the mid-80's by Mili, Van Cutsem, and Ribbens-Pavella [22]. They introduced and

evaluated a new technique in bad data identification, called the hypothesis test identification (HTI) method, based on an idea from Xiang et al. [40], in which hypothesis testing is performed on the individual error estimates to check whether the corresponding measurement is bad and should be eliminated. This technique is possible because in first approximation, the residuals may be expressed as linear combinations of the errors through the residual sensitivity matrix $\underline{W}$ as $\underline{r} = \underline{W}\underline{e}$ (see Appendix). Here a submatrix of $\underline{W}$ corresponding to the residuals and errors of redundant measurements is invertible, such that $\underline{e}_s = \underline{W}_{ss}^{-1}\underline{r}_s$, where the s subscript stands for suspected measurements.

In their evaluation of these identification methods, Mili et al. [22] considered the IBE method to be attractive, but weak in areas where multiple bad data interact with each other and/or local redundancy of measurements is poor, and noted that it is computationally intensive. The NQC estimators were considered to demonstrate a poor performance, with slow convergence or even divergence, and a high sensitivity to low local redundancy and multiple interacting bad data. They concluded that under general conditions, the HTI method demonstrated superior performance, even though it is also susceptible to the masking effect. This evaluation weakened to some degree interest in the application of non-quadratic estimators to state estimation for some time.

The class of M-estimators, WLS and non-quadratic alike, are based on parametric estimation theory, since they make underlying assumptions about the distribution of the measurement error made possible by the nature of the state estimation problem. Research has also been carried out exploring the application of non-parametric estimation theory, in which no *a priori* knowledge is required of the measurement errors' distribution. One such example is the research conducted by Mili, Phaniraj, and Rousseeuw [23], [27] on the least median of squares (LMS) estimator. Although this estimator can provide good estimates despite large percentages of bad data measurements when there are no dependencies among the rows of the design matrix, and initially appeared very promising in power system state estimation, it proved to have a tuning problem. When the estimator is adjusted for estimation in areas of high local redundancy of measurements, it may lose its ability to reject bad data in areas of low local redundancy, whereas if it is adjusted for estimation in areas of

low local redundancy, it may not reject multiple interacting bad data in areas of high local redundancy. To overcome this difficulty, Mili, Cheniae, et al. [24] proposed to decompose the system into small subnetworks and apply LMS to each subnetwork.

Recently interest in the non-quadratic M-estimators has been revived. Their performance in earlier comparisons was realized to be inhibited by the presence of bad data that lie on leverage points in the system (see Section 3.3). Mathematically, leverage points are outliers in the factor space of the linear regression, and for better or worse, heavily influence the solution of any M-estimator. This was realized to be the source of slow convergence or divergence in [22], and in 1996, Mili, Cheniae, et al. [25] introduced the use of projection statistics to identify leverage points and bound the influence of bad data that lie on them. The calculation of leverage points is used to define weights for M-estimators, producing the GM-estimators (G as in generalized). Certain of these, and notably the Schweppe-type Huber GM-estimator, have emerged as highly desirable tools in power system state estimation. They enjoy good convergence properties, are not sensitive to multiple interacting bad data, and are easily implemented as a modification of the relatively simple WLS-based algorithm.

It should be noted that an alternative method of dealing with leverage points has recently been proposed by Abur et al. [1] to eliminate leveraging effects of injection measurements by using matrix stretching techniques on the Jacobian matrix. The method was successfully applied to the least absolute value (LAV) estimator.

1.3. Personal Contribution

The work of this thesis was made possible by an exchange between the Virginia Polytechnic Institute and State University and the Ecole Polytechnique Fédérale de Lausanne, in Lausanne, Switzerland. During the 1996/1997 academic year, the Schweppe-type Huber GM-estimator (referred to as the SHGM-estimator), was developed for the Entreprises Electriques Fribourgeoises, known as the EEF electric utility and headquartered in the city of Fribourg, Switzerland. Contact with the EEF was made through Professor

Alain Germond and Dr. Rachid Cherkaoui. The author worked at the EPFL in conjunction with EEF engineers, notably Dr. Christian Tinguely, to develop an estimator package during a one year period.

The EEF utility services the Swiss canton of Fribourg, with an area of about 2,000km^2. The principle network consists of 33 substations interconnected by 540km of 60kV subtransmission lines, which is externally connected by one 130kV- and two 220kV-tie lines owned by the Swiss company EOS (Energie Ouest Suisse). These three tie lines, located at the Monteynan, Botterens and Villarepos substations, allow the network to be subdivided into three interconnectable islands that in turn transmit energy to loads through 18kV distribution networks.

The estimator was installed as described in [5] in the EEF energy management system. This decision was made after clear evidence was demonstrated of the superiority of the SHGM-estimator over the conventional WLS-based method, which the EEF employed previously. The program that originally implements the SHGM algorithm was developed for research at Virginia Tech by M.G. Cheniae, N. Vichare, and L. Mili. The code, written in FORTRAN, was underdeveloped for operation in real power systems and had to be upgraded to meet the specific requirements of the EEF. The upgrades included in the final version of the state estimator program are: a π-circuit transformer model, the ability to model parallel transmission lines or tie-line transformers, the inclusion of multiple bus voltage measurements, the independent treatment of active and reactive power measurements, text file output under a format specified by the EEF engineers, the creation of output files for use by an external load flow/short-circuit analysis program, modification options for the algorithm execution and output parameters, and the automatic calculation of recursive statistics based on the accumulated residuals of metered locations.

Once the algorithm was proven capable of making estimations with good convergence rates on the real EEF network, a topology processor code was written (also in FORTRAN) that transforms real-time SCADA files into a supernode model for the estimator. The topology processor and state estimator were eventually subjected to an

extensive off-line testing and debugging process, in which they were modified and updated for compatibility with the recently installed EEF SCADA system database. During this testing period, studies were performed to investigate how the topology processor handles difficult topologies, as well as how the state estimator calculates projection statistics and rejects outlying measurements under varying topologies and loading conditions. The topology processor's code and supporting files and the state estimator's output were developed with a high priority placed on minimization of the maintenance required of EEF control center personnel, and key personnel were familiarized with the interface between the SCADA and topology processor .

The simulations performed during the development and testing period showed that the iteratively reweighted least squares (IRLS) algorithm solution that implements the SHGM-estimator exhibits excellent convergence properties, with and without the presence of bad data with large gross errors. The algorithm converges in three iterations even when rejecting gross measurement errors, and the projection algorithm consistently identifies the system leverage points.

The EEF engineers did not express a strong preference concerning the modeling of extremely short lines (less than 500 m), stating that the emphasis should instead be placed on obtaining the best possible state estimate. The gain matrix remains well-conditioned in the presence of extremely short lines in the system, so that they do not present convergence problems. The elimination or inclusion of lines in the topological model was therefore decided on an individual basis. In general, short lines in the EEF system have only one flow measurement. The line might therefore be removed from the topological model for estimation and replaced for display, but in this case the state estimator cannot provide an estimate of the flow on that line. An estimate of the flow would then be determined uniquely by the measured flow and could not be rejected. Inclusion of extremely short lines in the model implies treating them as (extreme) leverage points. Since there is no problem of convergence in this case, inclusion of the lines becomes only a question of local redundancy and topological feasibility.

Short lines that connect an antenna bus to a network bus might be removed and the flow incorporated in the network bus' injection. However, these lines were modeled because in all cases the network bus injection is already unavailable. The most interesting cases of short lines occur in the Hauterive/Monteynan area and are addressed in Section 5.5. Metering studies conducted on the EEF system revealed some weaknesses in the measurement configuration. The most significant contribution of the metering studies was the determination of the need for added redundancy in the area of the Monteynan/Hauterive substations. Poor local redundancy in the Monteynan area gives rise to problems of bad data identification. Hence the SHGM-estimator is restricted to detecting the presence of bad data in the area, but cannot properly identify bad data. It should be noted that this is simply a problem of poor local redundancy and not short lines.

When processing real-time EEF data files, the SHGM-estimator confirmed a discrepancy in sign conventions at the tie line interconnections, and adjustments were made accordingly in the topology processor. Efforts were made to demonstrate the bias of the system's measuring devices. Significantly nonzero SCADA values given by certain measuring devices located on deenergized lines indicate the presence of a high bias in those devices.

Since the system's installed generation capacity is primarily hydroelectric and produces more during summer, system loading conditions were generally low in the summertime when the majority of testing was performed. Additional studies may therefore be performed during the higher loading conditions of the winter season, such as the one in Section 5.4.

A complete documentation of all files necessary for maintenance of the topology processor and state estimator package was written and submitted to the EEF research and development department and is available at the control center. An introduction to the package was formally presented for the orientation of all EEF engineers and personnel involved in the project.

The remainder of the thesis is divided into four sections. Chapter 2 introduces maximum likelihood theory, while Chapter 3 develops relevant robust theory, reviews the robust performance of some M-estimators, and applies robust estimators to the problem of leverage points in state estimation regression through projection statistics. Chapter 4 is a review of the pertinent algorithms of the topology processor and state estimator as installed in the EEF system. Chapter 5 provides simulated results to demonstrate the superior performance of the SHGM-estimator over that of the WLS estimator on the EEF network, as well as an evaluation of the network's meter placement.

2 – MAXIMUM LIKELIHOOD ESTIMATION

In power system state estimation we are interested in finding an estimate of the locations of the complex bus voltages, as well as other multivariate estimates of location and variance, in a robust manner. To do this we need to review some notions of multivariate estimation and robustness, and we begin by developing some basic concepts of univariate estimation theory, notably the idea of maximum likelihood. Parametric estimation theory was founded in the 1920's by Sir R.A. Fisher, who contributed the maximum likelihood estimators for all classes of distributions. Fisherian concepts of consistency, unbiasedness and efficiency reviewed, and equivariance properties of estimators are provided in the Appendix. For simplicity we will begin with univariate estimation, moving to multiple dimensions in later discussions after the basic concepts have been established.

2.1. Definitions

We begin by considering a sample of *m* independently and identically distributed (i.i.d.) observations $\underline{z} = \{z_1, \ldots, z_m\}$ that are drawn from a function $f(z;\underline{\theta})$,where $\underline{\theta}$ is a vector of parameters that defines the function. We say that $f(z;\underline{\theta})$ is a probability density function when it is a function of $\underline{z}$ given $\underline{\theta}$, whereas it is a likelihood function when it is a function of $\underline{\theta}$ given $\underline{z}$. Likelihood theory is based on taking a given sample and finding the parameter values to which the sample is most likely to conform. Since we are under the assumption of independence, the likelihood function can be defined as

$$L(\underline{\theta};\underline{z}) = c\, f(\underline{\theta} \mid \underline{z}) , \tag{2.1}$$

where c is an arbitrary constant which we set to 1. The likelihood function is proportional to the probability that the sample follows a distribution with the density function $f(\underline{z};\underline{\theta})$. Again, since the *m* observations are assumed independent, the result is

$$L(\underline{z};\underline{\theta}) = f(z_1,\ldots,z_m;\underline{\theta}) = \prod_{i=1}^{m} f(z_i;\underline{\theta}) , \tag{2.2}$$

and we want to find a $\underline{\theta}$ that maximizes this product, an estimator that maximizes this likelihood function called the maximum likelihood estimator $\hat{\underline{\theta}}_{ML}$. Since the exponential function is a monotonic function, the following equivalence holds:

$$\max_{\hat{\underline{\theta}}_{ML}} L(\underline{z};\underline{\theta}) \Leftrightarrow \min_{\underline{\theta}}\left[-\ln\prod_{i=1}^{m} f(z_i;\underline{\theta})\right] \Leftrightarrow \min_{\underline{\theta}}\left[\sum_{i=1}^{m} -\ln f(z_i;\underline{\theta})\right]. \tag{2.3}$$

Allowing $\underline{\theta}$ to be only of one parameter, an estimator θ of location, we define the residual of the i^{th} observation z_i as $r_i = z_i - \theta$. Then we can denote the probability density function as $f(r_i)$ because this change in notation only represents a shift in the density function. Thus the objective function in (2.3) to be minimized with respect to θ is

$$J(\theta) = \sum_{i=1}^{m} -\ln f(r_i), \tag{2.4}$$

and we set $\dfrac{\partial J(\theta)}{\partial \theta} = \dfrac{\partial J(\theta)}{\partial r}\dfrac{\partial r}{\partial \theta} = 0$.

Since we are interested in working with the residuals and $\partial r/\partial\theta = -1$, we define a function, called the psi-function, or ψ-function, as

$$\psi(r) = \frac{\partial J(\theta)}{\partial r} = \frac{d}{dr}\sum_{i=1}^{m}[-\ln f(r_i)] = \sum_{i=1}^{m} -\frac{f'(r_i)}{f(r_i)}. \tag{2.5}$$

Hence we desire to find θ such that

$$J(\theta) = \sum_{i=1}^{m}\psi(r_i) = 0. \tag{2.6}$$

Equation (2.6) is referred to as the likelihood equation, and by minimizing the objective function, we maximize the probability that $\underline{z}$ follows a distribution with a probability density function $f(r)$.

2.2. Fisherian Concepts

The concepts developed by Fisher about the consistency, unbiasedness and efficiency of an estimator lay the foundation necessary to formally develop ideas about the

robustness of an estimator. The set of *m* observations in $\underline{z}$ that follow exactly a distribution with a probability density *f*(z) and a cumulative distribution function (c.d.f.) F(z) may be viewed as *m* realizations of a random variable z. The empirical cumulative distribution function $F_m(z)$ based solely on the sample $\underline{z}$ is given by

$$F_m(z) = \frac{1}{m}\sum_{i=1}^{m}\Delta(z - z_i), \quad \Delta(z - z_i) = \begin{cases} 0, & z < z_i \\ 1, & z \geq z_i \end{cases} \tag{2.7}$$

The asymptotic cumulative distribution function is $F(z) = \int_{-\infty}^{z} f(t)\,dt$, and by virtue of the Glivenko-Cantelli [4] theorem, $F_m(z) \xrightarrow{\text{w.p.1}} F(z)$ as $m \to \infty$.

The expression $\hat{\theta}_m(F_m)$ is referred to as the functional form of an estimator if we have $\hat{\theta}_m(F_m) = \hat{\theta}(F_m)$ for all *m* and F_m [31]. Functional forms are now employed to introduce Fisherian concepts.

Consistency An estimator is said to be Fisher consistent at the distribution F if, given the entire population, the estimator estimates the true value of the parameter. Using the notation of functional form,

$$\lim_{m\to\infty} \hat{\theta}(F_m) = \hat{\theta}(F) = \theta_{true} \ . \tag{2.8}$$

Bias The bias of an estimator is a repeated sampling concept. On average an estimator $\hat{\theta}_m$ returns a given value, which is its bias. The bias of an estimator at the distribution F is defined as

$$b_m = E[\hat{\theta}_m] - \theta_{true} \ . \tag{2.9}$$

The estimator $\hat{\theta}_m$ is said to be unbiased at F if $E[\hat{\theta}_m] = \theta_{true}$, that is, if on average it returns the true value θ.

Efficiency In order to define the efficiency of an estimator, we must first define a quantity known as the Fisher information, I_f. Given the sample $\underline{z}$ and assuming the model

$\underline{z} = \theta_{true} + \underline{e}$, where $\underline{e}$ is an error vector, we see that $\underline{z}$ contains information about θ_{true}. This information is expressed by the Fisher information:

$$I_f(\theta) = \int_{-\infty}^{\infty} \left(\frac{f'(r)}{f(r)} \right)^2 f(r) dr \tag{2.10}$$

This is equivalent to $E\left[\psi^2(r)\right]$ for the ψ-function of the maximum likelihood estimator. In effect, I_f tells us about the lower bound on variance that any estimator may achieve. The Cramèr-Rao lower bound [17] is a so-called information inequality that relates I_f to the lower bound of the variance that any estimator can achieve for a given $f(r)$. If we consider any normalized estimator $\hat{\theta}_m^* = \sqrt{m}\hat{\theta}_m$, the lower bound on the variance that it can achieve is given by the Cramèr-Rao lower bound:

$$Var\left(\hat{\theta}_m^*\right) \geq \frac{1}{I_f(\theta)}, \tag{2.11}$$

that is, the variance of any normalized estimator cannot be smaller than the reciprocal of the Fisher information.

Hence we may think of I_f as a measure of the ease with which θ can be estimated. However, what estimators can attain this lower bound? The definition of an estimator's efficiency follows directly from the Cramèr-Rao lower bound:

$$0 \leq e = \frac{1/I_f}{Var\left(\sqrt{m}\hat{\theta}_m\right)} \leq 1 \tag{2.12}$$

A Fisher consistent estimator is said to be asymptotically efficient if it achieves the Cramèr-Rao lower bound:

$$\lim_{m \to \infty} Var\left(\sqrt{m}\hat{\theta}_m\right) = 1/I_f \ . \tag{2.13}$$

A Fisher consistent maximum likelihood estimator is asymptotically efficient, that is, it has an asymptotic efficiency equal to 1. Hence, the maximum likelihood estimator for a given density $f(r)$ is asymptotically the best estimator in the sense that it minimizes the asymptotic variance ; it is the most efficient estimator for that density function. It should be noted that in many cases, there is no estimator that obtains the lower bound for finite samples.

Furthermore, given two Fisher consistent estimators $\hat{\theta}'$ and $\hat{\theta}''$, the asymptotic relative efficiency of $\hat{\theta}'$ with respect to $\hat{\theta}''$ is defined as

$$e_{\hat{\theta}',\hat{\theta}''} = \lim_{m\to\infty} \frac{\mathrm{Var}\left(\sqrt{m}\hat{\theta}''_m\right)}{\mathrm{Var}\left(\sqrt{m}\hat{\theta}'_m\right)} . \tag{2.14}$$

Typically, $\hat{\theta}''$ is chosen as the maximum likelihood estimator; since it is the most efficient estimator, it may be chosen as a benchmark.

3 – ROBUSTNESS ANALYSIS

In order to analyze the robustness of different estimators, we must determine specifically what is meant by robustness. The parametric estimation theory developed by Fisher assumes *a priori* knowledge of the probability distributions of a sample, and that they are i.i.d. The maximum likelihood estimators developed under this theory, and most notably the least squares, are easily implemented, but this comes at the high price of great vulnerability to variations from the assumptions. Hence, the need was recognized for estimators that protect against the various inevitable violations of these assumptions.

One position opposite to that of Fisherian parametric theory is non-parametric estimation theory, for which no *a priori* knowledge of the probability distribution is required. Another alternative was provided when Peter Huber [19] founded robust estimation theory by recognizing that we may usually assume not an exact but an approximate probability distribution of the data set. Based on this approximately known distribution, a regression model may be fitted to the majority of the data, the rest being considered as outliers whose influence must be bounded. Along these lines, Huber [19], [20] defined robustness as being "synonymous with insensitivity to small departures from the assumptions."

In 1968, Fred Hampel developed Huber's robust theory in his doctoral dissertation [14] on distinct concepts of robustness. He studied the stability of the bias and variance of an estimator when the assumptions are affected by perturbations in the sample. Hampel's studies included [15], [16] three different kinds of robustness:

(i) Qualitative robustness, which considers small perturbations that have small effects.

(ii) Global robustness, which is a sort of quantification of the breakdown point, that is, it considers the largest possible perturbation before the estimator "breaks down".

(iii) Local robustness, which involves infinitesimal changes and is measured by the influence function (IF) and the change-of-variance function (CVF).

We are most interested in local and global concepts of robustness. Specifically, we will discuss global robustness in terms of breakdown point and maximum bias curve, and local robustness in terms of the influence function and gross error sensitivity.

These concepts of robustness are useful tools in evaluating the robustness of M-estimators. In Section 3.2 the M-estimator class is defined, and some key M-estimators are compared in terms of the robust theory developed in Section 3.1. Section 3.3 provides the application of robust estimators to the problem of leverage points in power system state estimation. The material of Section 3.1 is based largely on the discussions in [23].

3.1. Robustness Concepts

Consider a sample $\underline{z} = \{z_1,\ldots,z_m\}$ of m observations that are i.i.d. according to a c.d.f. G. Generally G is unknown and we approximate it with F. Let the c.d.f. of the estimator $\hat{\theta}_m(z_1,\ldots,z_m)$ associated with the true c.d.f. G and the assumed c.d.f. F be denoted as $L_G\{\hat{\theta}_m\}$ and $L_F\{\hat{\theta}_m\}$, respectively. Hampel's theorem states that a series of estimators $\{\hat{\theta}_m;\ m \geq 1\}$ is said to be qualitatively robust at F if a small deviation between F and G yields a small deviation between $L_G\{\hat{\theta}_m\}$ and $L_F\{\hat{\theta}_m\}$. [23] In qualitative robustness we are exploring an aspect of robust estimation theory by perturbing the ideal, defined distribution F with a true, unknown distribution G similar to F and seeing whether the estimator remains in the neighborhood of the true parameter value.

The asymptotic bias of an estimator $\hat{\theta}$, different from Fisherian bias, may be defined as

$$b = \left|\hat{\theta}(G) - \hat{\theta}(F)\right|. \tag{3.1}$$

To declare an estimator qualitatively robust implies that the maximum possible bias of $\hat{\theta}$ remains bounded for a positive fraction ε of contamination. This leads to global robustness concepts by asking what the maximum fraction ε* of arbitrarily large errors is that a given estimator can handle, that is, at what point the estimator breaks down. We may denote the distribution of outliers in G by their c.d.f. H and write the ε-contaminated model [20] as

$$G = (1-\varepsilon)F + \varepsilon H . \tag{3.2}$$

Assuming the *m* data points of the sample $\underline{z}$ are good, meaning they follow F, consider the set Z' of all possible corrupted samples $\underline{z}$' obtained by replacing an integer number ξ of the good data points with arbitrarily large values such that the fraction of contamination is ε = ξ / *m*. If $\hat{\theta}'_m$ represents the estimate of the samples in Z', the maximum bias resulting from the contamination is

$$b_{max} = \sup \left| \hat{\theta}_m - \hat{\theta}'_m \right| , \tag{3.3}$$

or

$$b_{max}(\varepsilon) = \sup_H \left| \hat{\theta}(F) - \hat{\theta}(G) \right| . \tag{3.4}$$

The breakdown point ε* of $\hat{\theta}_m$ at $\underline{z}$ is then

$$\varepsilon^* = max\{\varepsilon = \xi / m \mid b_{max} \text{ is finite}\}. \tag{3.5}$$

It turns out that in practice the breakdown point does not depend on F. Robust estimators rely on the majority of the data in G and bound the influence of outliers, which are assumed to be a minority in the data set $\underline{z}$. As a result, the largest ξ that any estimator can handle is half the number of redundant observations:

$$\xi_{max} = \left[\frac{m-n}{2} \right] , \tag{3.6}$$

so the breakdown point is given by

$$\varepsilon^*_{max} = \frac{1}{m} \left[\frac{m-n}{2} \right] . \tag{3.7}$$

Here [·] denotes the integer part, and *n* is the number of unknowns being estimated. In the one-dimensional case, *n* = 1. The asymptotic value (as m→∞) of ε^*_{max} is ½, and estimators that have large values of ε* are called high breakdown point estimators.

An estimator's local robustness at a given probability distribution is measured by the Influence Function (IF) and Change-of-Variance Function (CVF). In power system state estimation, we are concerned with effects of outliers on an estimator's bias, which is measured by the IF. The CVF measures the effects of outliers on an estimator's spread, hence we will develop only the IF. For our sample of *m* observations $\underline{z} = \{z_1,\ldots, z_{m-1}, z\}$ that follow F exactly, we perturb a selected observation z such that it may take on any real number. The influence of z on an estimator of location $\hat{\theta}_m$ is written as

$$\mathrm{IF}_m\left(z;\hat{\theta}_m,F\right)=\frac{\hat{\theta}_m\left(z_1,\ldots,z_{m-1},z\right)-\hat{\theta}_{m-1}\left(z_1,\ldots,z_{m-1}\right)}{\varepsilon}. \tag{3.8}$$

Since only one observation in the sample is perturbed, $\varepsilon = 1/m$, so

$$\mathrm{IF}_m\left(z\right)=m\left(\hat{\theta}_m-\hat{\theta}_{m-1}\right). \tag{3.9}$$

The IF can be bounded or unbounded, depending on the estimator's properties. In the asymptotic case, where the limit is taken as $\varepsilon\to 0$ of the finite case, the IF becomes the directional derivative of the functional form $\hat{\theta}(G)$ in the direction of G. It thus describes how the estimator behaves in the close neighborhood of F. The IF defines an important measure of local robustness, the gross error sensitivity γ^*. It is the supremum of the absolute value of the IF where z is taken over the range of real numbers:

$$\gamma^*=\sup_z\left|\ \mathrm{IF}\left(z;\hat{\theta}_m,F\right)\right|. \tag{3.10}$$

The gross error sensitivity is important because it is linked to the maximum bias in the area of $\varepsilon\to 0$. The bias is expressed as before, but now in terms of the IF, yielding

$$b=\left|\hat{\theta}_m-\hat{\theta}_{m-1}\right|\approx\varepsilon\left|\mathrm{IF}\right|,$$

so that

$$b_{max}\approx\varepsilon\,\gamma^*. \tag{3.11}$$

We therefore have an idea of how the maximum bias behaves near $\varepsilon = 0$, keeping in mind that as ε increases (3.11) becomes more approximate.

From the global perspective, an estimator should have a low maximum bias, while enjoying the highest possible breakdown point. On a more local level, it is not only

desirable but usually quite important for an estimator to have a bounded IF and a small gross error sensitivity at a given probability distribution. In the next section it will be shown that when estimators possess these properties, it usually comes at the expense of the estimator's efficiency at the Gaussian distribution.

3.2. M-estimators

The maximum likelihood-type estimator, or M-estimator, class was introduced by Peter Huber in 1964. It is a large class of estimators that contains maximum likelihood estimators as a subclass. An M-estimator is defined as the minimizer of the objective function $J(\theta)$. For the one-dimensional case with *m* observations in $\underline{z}$, an M-estimator is defined by minimizing the objective function

$$J(\theta) = \sum_{i=1}^{m} \rho\left(\frac{r_i}{s}\right) \tag{3.12}$$

similarly to (2.5), so that in the one-dimensional case we have

$$\sum_{i=1}^{m} \psi\left(\frac{r_i}{s}\right) = 0, \tag{3.13}$$

where s is some estimator of scale. Here we have defined $J(\theta)$ in terms of a real-valued function $\rho(r)$, which is even and non-decreasing for positive residuals, and zero at the origin. From now on, we will assume that s = 1. In order for the objective function $J(\theta)$ to have a single minimum value, $\rho(r)$ needs to be convex. From the theory of Chapter 2 we see that an M-estimator is a maximum likelihood estimator when a probability density function *f*(r) can be found from $\rho(r) = -\ln f(r)$, that is when *f*(r) is of the form $f(r) = Ke^{-\rho(r)}$ for some constant K. Any M-estimator that employs a hard rejection of outliers, that is, for which $\rho(r)$ is constant past a threshold, can therefore not be a maximum likelihood estimator.

Table 3.2.1 – ρ- and ψ-functions of some M-estimators

	M-estimator	Range	$\rho(r)$	$\psi(r)$
First subclass	Mean	$\forall r$	$\frac{1}{2}r^2$	r
	Median	$\forall r$	$\lvert r\rvert$	$\operatorname{sign}(r)$
	Huber	$\lvert r\rvert \le b$ $\lvert r\rvert > b$	$\frac{1}{2}r^2$ $b\lvert r\rvert - b^2/2$	r $b\ \operatorname{sign}(r)$
Second subclass	Muller	$\lvert r\rvert \le b$ $\lvert r\rvert > b$	$\frac{1}{2}r^2$ $\frac{1}{2}b\lvert r\rvert$	r $\frac{1}{2}b\ \operatorname{sign}(r)$
	Merrill-Schweppe	$\lvert r\rvert \le b$ $\lvert r\rvert > b$	$\frac{1}{2}r^2$ $2b^{3/2}\sqrt{\lvert r\rvert} - \frac{3}{2}b^2$	r $b^{3/2}\frac{\operatorname{sign}(r)}{\sqrt{\lvert r\rvert}}$
Third subclass	Huber type-skipped mean	$\lvert r\rvert \le b$ $\lvert r\rvert > b$	$\frac{1}{2}r^2$ $\frac{1}{2}b^2$	r 0
	Hampel	$\lvert r\rvert < a$ $a < \lvert r\rvert \le b$ $b < \lvert r\rvert \le c$ $\lvert r\rvert > c$	$\frac{1}{2}r^2$ $a\lvert r\rvert - a^2/2$ $a\left(\frac{c\lvert r\rvert - \frac{1}{2}r^2}{c-b}\right) - \frac{a^2}{2} - \frac{ab^2}{2(c-b)}$ $\frac{a}{2}(c+b-a)$	r $a\ \operatorname{sign}(r)$ $a\ \frac{(c-\lvert r\rvert)}{(c-b)}\ \operatorname{sign}(r)$ 0
	Tukey's Biweight	$\lvert r\rvert \le b$ $\lvert r\rvert > b$	$\frac{1}{6}\left[b^6 - \left(b^2 - r^2\right)^3\right]$ $b^6/6$	$r\left(b^2 - r^2\right)^2$ 0

A robustness study was conducted at Princeton [2] analyzing 68 estimators. Of these, several have been considered for application to the multivariate task of state estimation. Table 3.2.1 describes several important M-estimators. M-estimators of the first subclass are characterized by a convex $\rho(r)$ function. In the second subclass, $\rho(r)$ is not convex, but $\psi(r) \neq 0$ when $r \neq 0$. M-estimators of the third subclass do not have a convex $\rho(r)$, and $\psi(r) = 0$ for r greater than a threshold. The following are some useful properties of all M-estimators:

1) An M-estimator is Fisher consistent at F if

$$\int_{-\infty}^{\infty} \psi(r)\,dF(r) = 0 \text{ for all F, which implies } \hat{\theta}_M(F) = \theta_{true} \tag{3.14}$$

2) A Fisher consistent M-estimator is asymptotically Gaussian in that

$$L\left\{\sqrt{m}\left(\hat{\theta}_M - \theta_{true}\right)\right\} \xrightarrow[m\to\infty]{} N\left(0, \sigma_M^2\right),$$
$$\text{where } \sigma_M^2 = \frac{E\left[\Psi^2(r)\right]}{\left(E\left[\Psi'(r)\right]\right)^2} \tag{3.15}$$

The variance σ_M^2 is an asymptotic variance.

The influence function of an M-estimator is given by

$$IF(r;F) = \left.\frac{d\hat{\theta}}{d\varepsilon}\right|_{\varepsilon = 0} = \frac{\psi(r)}{E\left[\psi'(r)\right]} \quad . \tag{3.16}$$

Since the IF exists at F for all M-estimators, it may be used to define the asymptotic variance of an estimator at a given probability distribution. If the expected value of an estimator's IF at F is zero, then the expected value of IF^2 yields the estimator's asymptotic variance at F:

$$V\left(\hat{\theta};F\right) = \int_{-\infty}^{\infty} IF^2(r)\,dF(r) = E\left[IF^2(r)\right]. \tag{3.17}$$

From (3.16) and (3.17), the asymptotic variance of an M-estimator is

$$V\left(\hat{\theta};F\right) = E\left[IF^2(r)\right] = \frac{E\left[\psi^2(r)\right]}{\left(E\left[\psi'(r)\right]\right)^2}. \tag{3.18}$$

We may use this asymptotic variance to investigate the efficiencies of some key M-estimators at various distributions. We consider the mean, median and Huber estimators,

Table 3.2.2 – Performance of the mean, median and Huber at varying distributions

f(r)	M-estimator	Theoretical values		Monte Carlo values	
		$\frac{E[\psi^2(r)]}{(E[\psi'(r)])^2}$	$e_{\hat{\theta}',\hat{\theta}''}$	$\bar{\hat{\theta}}$, x10^{-3}	$V\left(\sqrt{500}\left(\hat{\theta} - \bar{\hat{\theta}}\right)\right)$
Gaussian N(0,1)	Mean	1	1	-1.47	0.989
	Median	1.571	0.637	-2.46	1.552
	Huber	1.037	0.964	-1.45	1.039
Laplacian *f*(0,100)	Mean	20 000	0.5	0.0744	19 213
	Median	10 000	1	0.0258	10 401
	Huber	10 050	0.995	0.1438	12 803
Cauchy *f*(0,1)	Mean	∞	-	33.92	644 000 000
	Median	2.467	-	3.547x10^{-4}	2.412
	Huber	2.993	-	2.396x10^{-3}	3.248

and compare their asymptotic variances at the Gaussian, Laplacian, and Cauchy distributions. At each distribution we calculate the asymptotic variance of the M-estimators and run Monte Carlo simulations to verify the results. The calculations are carried out in Section A.7, and the results are recorded in Table 3.2.2.

The results obtained at each distribution are quite different. The mean becomes very inefficient at the Laplacian and breaks down completely at the Cauchy. The median is inefficient at the Gaussian, but is the maximum likelihood estimator at the Laplacian and performs well at the Cauchy. The Huber does not achieve the best variance at any of the distributions, but is quite efficient in each case. Because the Monte Carlo simulations are not perfect, the results obtained for the Huber do not always closely reflect the asymptotic values, but they do verify its trend of good performance at different distributions.

These well-known results strongly emphasize that the performance of an estimator depends heavily on the probability density function at which it is being evaluated. Most significantly, the mean is not to be trusted for all distributions, regardless of whether outliers are considered. Highly robust M-estimators such as the median can generally be applied with confidence, although at the cost of a higher variance when the distribution is purely Gaussian. However, if we consider an ε-contaminated model of (3.2) wherein F and H are Gaussian and H has a higher variance than F, the median becomes more efficient than the mean above a very small percentage of contamination.

The Huber proves to retain a good efficiency at various distributions, including the ε-contaminated model for a large variation of ε. For the problem of power system state estimation, the Huber is an excellent choice because it filters measurement noise efficiently, while its robustness properties allow it to neatly bound the influence of outliers without totally rejecting them. It also has the advantage of a convex objective function.

3.3. Leverage Points

The general regression model is developed in Section A.2. Bad data are outliers in the regression's response space, that is, in the metered values. Leverage points are outliers in the space spanned by the row vectors of the Jacobian matrix, meaning that they do not follow the pattern of the point cloud in that space. It is shown in Section A.4 that the entries of the Jacobian matrix are determined by the topology and parameters of the power network. Leverage points occur when a line is very short; since its reactance is small, the susceptance B_{ij} of line ij's admittance will be quite large. Leverage points also occur for injection measurements when many lines are joined at a bus, because the design matrix's entry corresponding to the row of the measurement and the column of the injection bus' state variable is directly proportional to the number of adjacent lines.

There are two cases associated with leverage points. When a measurement is a leverage point and has a wrong metered value, it is a "bad leverage point" and may wreck the performance of any M-estimator. If, however, the measurement is a leverage point and

has a good metered value, it is a "good leverage point" and heavily reinforces the M-estimator's performance. In order to discuss how to identify and deal with leverage points, we place the regression model in its linearized form, then develop the multivariate estimators used for the identification process, termed the projection statistics.

3.3.1. *Linearization of the Regression Model*

The system state variables may be related to the measurements according to the nonlinear regression model

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

given by (8.4). Neugebauer and Mili [26] derived the influence functions of M-estimators in nonlinear regression, showing that the influence of the model must be bounded together with the influence of residuals. Power systems normally operate in the neighborhood of the flat voltage profile, and the Jacobian matrix of the design matrix is relatively constant in the area of this operating point if its Hessian matrix is negligible. Assuming this is so, the nonlinear model may then be approximated by performing a first-order Taylor series expansion of (8.4) about an operating point, beginning with the flat voltage profile $\underline{x}^{(0)}$, in which all voltage magnitudes are unity and all phase angles zero. For small changes about the operating point $\underline{x}^{(k)}$, the relationship is given by

$$\Delta\underline{z} = \underline{H}\left(\underline{x}^{(0)}\right)\Delta\underline{x} \,. \tag{3.19}$$

Here the matrix $\underline{H}(\underline{x}) = \partial\underline{h}(\underline{x})/\partial\underline{x}$ denotes the Jacobian matrix, which is very sparse in power systems, and increasingly so with the system's size. The rows of the normalized Jacobian matrix $\sqrt{\underline{R}^{-1}}\,\underline{H}(\underline{x})$ define a set of m vectors, $\{\underline{\ell}_i^T,\ i = 1,\ldots,m\}$. Each of these vectors defines a point in the factor space of the linear regression, and we may use them to identify and locate the presence of the system's leverage points.

The sparsity of $\underline{H}(\underline{x})$ is reflected by the characteristic of the vectors $\underline{\ell}_i$ that they contain relatively few non-zero elements. They therefore lie on a much smaller subspace of the factor space termed the relevant subspace. For active or reactive power flows, the relevant subspace is of dimension 4, corresponding to the two state variables of the buses on

the line's two ends. If the decoupled model is considered in which active power is coupled to phase angle and reactive power to voltage magnitude, the dimension is reduced to 2. For injections, a power measurement made on a bus with *k* adjacent lines has a relevant subspace of dimension 2*k*+1 in the decoupled model, and this dimension typically does not exceed 15.

3.3.2. *Projection Statistics*

Although robust M-estimators are useful tools when used to estimate the state of a power system, they fail along with classical estimators when bad leverage points are not treated. Specifically, measures must be carefully taken to identify leverage points in a power system so that the negative influence of bad data at these points may be limited and the positive influence of good leverage points kept. The following discussion closely parallels that in [25].

In order to identify outliers in the multidimensional factor space, we require multivariate estimators of location and covariance. Instead of estimating the location and spread of *m* data points defined by the set $\{z_1,\ldots,z_m\}$ of real numbers, we wish to estimate the equivalent information for *m* data points defined by the set of vectors $\{\underline{\ell}_1,\ldots,\underline{\ell}_m\}$. In this case, our intent is to estimate the center and the dispersion of the point cloud that this set forms by means of a multivariate location estimator and a covariance matrix, respectively. If we suppose that the vectors follow a multivariate normal distribution $N(\underline{\mu},\underline{\Sigma})$ with mean $\underline{\mu}$ and covariance matrix $\underline{\Sigma}$, then the maximum likelihood estimate of $\underline{\mu}$ is the sample mean,

$$\underline{\bar{\ell}} = \frac{1}{m}\sum_{i=1}^{m}\underline{\ell}_i \ , \tag{3.20}$$

and the unbiased maximum likelihood estimate of $\underline{\Sigma}$ is the sample covariance matrix,

$$\underline{C} = \frac{1}{m-1}\sum_{i=1}^{m}\left(\underline{\ell}_i - \underline{\bar{\ell}}\right)\left(\underline{\ell}_i - \underline{\bar{\ell}}\right)^T . \tag{3.21}$$

Both $\underline{\bar{\ell}}$ and $\underline{C}$ are affine equivariant, meaning they are independent of the coordinate system chosen. They are employed by the conventional method of identifying outliers to the point

cloud. This is the Mahalanobis distance MD_i, a measure of the distance of point $\underline{\ell}_i$ with respect to the bulk of the point cloud, defined as

$$MD_i = \sqrt{\left(\underline{\ell}_i - \bar{\underline{\ell}}\right)^T \underline{C}^{-1} \left(\underline{\ell}_i - \bar{\underline{\ell}}\right)} \ . \tag{3.22}$$

An ellipsoid centered at $\bar{\underline{\ell}}$ is formed by the set of points $\underline{\ell}$ satisfying $MD_i^2(\underline{\ell}) \le b$, whose principal axes are defined by the eigenvectors of $\underline{C}$, $\underline{v}_j$, j=1,…,*n*. Under the Gaussian model, the MD_i^2 have approximately a chi-squared (χ_n^2) distribution with *n* degrees of freedom. Thus, the classical method of identifying point $\underline{\ell}_i$ as an outlier is to test its Mahalanobis distance against a cutoff value, such as $MD_i \ge \left(\chi^2_{n,0.975}\right)^{1/2}$.

Just as in the one-dimensional case, the sample mean and sample covariance matrix are not robust, and the Mahalanobis distance is prone to the masking effect of multiple outliers, especially outliers that lie in clusters. Robust estimators are thus sought as an alternative to identifying outliers. Donoho [9] and Stahel [38] independently showed that the MD_i can alternatively be expressed as

$$MD_i = \underset{\forall \underline{v}, \|\underline{v}\| = 1}{max} \frac{\left| \underline{\ell}_i^T \underline{v} - L\left(\underline{\ell}_1^T \underline{v}, \cdots, \underline{\ell}_m^T \underline{v}\right) \right|}{S\left(\underline{\ell}_1^T \underline{v}, \cdots, \underline{\ell}_m^T \underline{v}\right)} \ , \tag{3.23}$$

where L and S represent the sample mean and sample standard deviation of the projections of the $\underline{\ell}_i$ onto the direction of the vector $\underline{v}$, and the equality holds only when all possible directions of $\underline{v}$ are considered. The distance MD_i can be robustified by choosing robust estimators to replace L and S. Noting that it is not practical to investigate all possible projection directions, Gasko and Donoho [13] propose a projection algorithm wherein only certain directions are explored. The selected directions are those which pass through the coordinatewise median $\underline{M}$ and each of the data points $\underline{\ell}_i$, that is, $\underline{v} = \underline{\ell}_i - \underline{M}$, where

$$\underline{M} = \left[\underset{j}{\text{median}}\, \ell_{j1}, \ldots, \underset{j}{\text{median}}\, \ell_{jm} \right]^T \ . \tag{3.24}$$

For the power system problem, $\underline{M}$ is constrained to be the origin. One good choice for replacing S is the robust median-absolute-deviation (MAD), defined as

$$\mathrm{MAD} = 1.4826 \ \underset{i}{\mathrm{med}} \left| \underline{\ell}_i^T \underline{v} \right| . \tag{3.25}$$

The following slightly modified version of a highly robust estimator of scale developed by Rousseeuw and Croux [32], [8] may also be used to replace S:

$$S'_m = 1.1926 \left(\underset{i}{\mathrm{lomed}} \ \underset{j \neq i}{\mathrm{lomed}} \left| \underline{\ell}_i^T \underline{v} + \underline{\ell}_j^T \underline{v} \right| \right) . \tag{3.26}$$

Here, lomed is a low median defined as the integer part of the [(*m*+1)/2]-th order statistic out of *m* numbers. The distance resulting from the recalculation of (3.23) using this information is called the *projection statistic*, denoted by PS_i and formally defined by

$$PS_i = \max_{\underline{v}} \frac{\underline{\ell}_i^T \underline{v}}{S'_m} , \tag{3.27}$$

for $\underline{v} = \underline{\ell}_j$, j = 1,…, *m*. Once calculated, the projection statistics are easily used to robustify the M-estimators, as described in Section 4.3.

4 – ALGORITHMS

This chapter describes the algorithms used to implement the topology processor and the SHGM-estimator. The topology processor code introduced in Section 4.1 was written specifically for the EEF system and is a contribution of the thesis. Section 4.2 reviews the changes made to the update the state estimator program for use on the EEF network. The identification of leverage points using the projection statistics is and the IRLS algorithm are reviewed in Sections 4.3 and 4.4, respectively. An overall pseudocode describng the SHGM algorithm is provided in Section 4.5.

4.1. Topology Processor

We have said that the security functions used in the EMS cannot be directly connected to the SCADA, because they require the output a state estimator. Neither can the state estimation algorithm be directly connected to the SCADA, for it requires a topology processor to develop a coherent topological model complete with accurate meter placements. The function of a topology processor is therefore to deliver an updated, consistent model of the system in terms of topology and measurements, based on known system connections and parameters and real-time SCADA input. It must divide the network into electrical islands if necessary, provide the necessary system parameters, and place measurements in their topological locations. It may also serve other important but less essential functions, such as posting warnings for circuit breaker and/or isolator conflicts, or various anomalies in the fields of the SCADA database. Figure 4.1.1 depicts the computer environment of the topology processor in relation to system files, static supporting files that contain information about the system's topology and network parameters, and the state estimator.

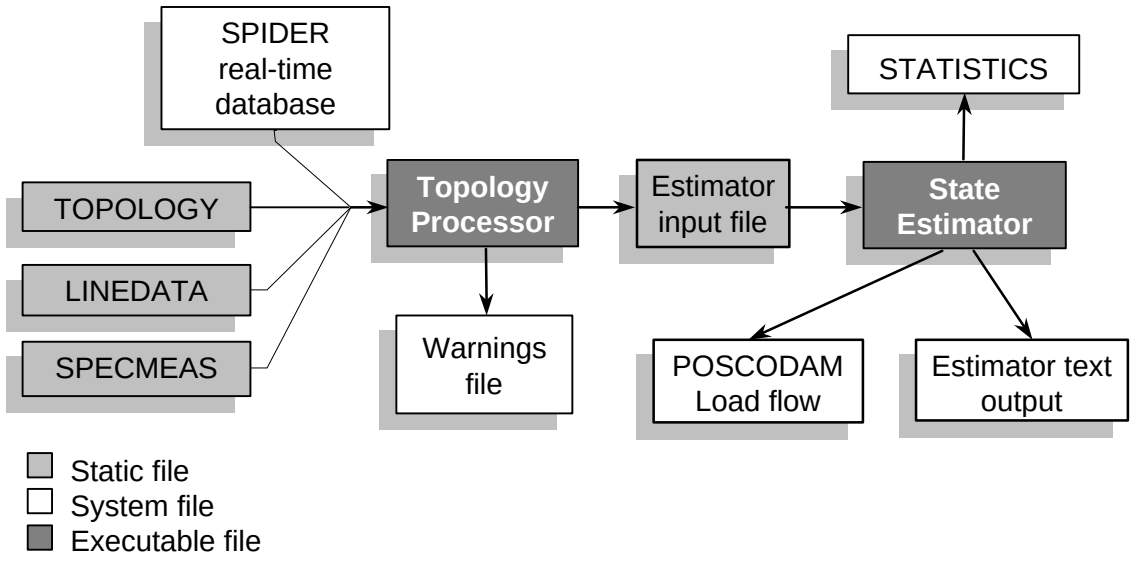

Figure 4.1.1 – Environment of the topology processor and state estimator

The standard model for topology processors was proposed by Sasson et al. [33]. However, a specialized topology processor program was developed for compatibility issues and to deal with special cases of the EEF network. Part of the logic proposed in [33] was adapted for the identification of system islands. The topology processor follows the process shown in Figure 4.1.2 to produce the input to the state estimator as indicated in Figure 4.1.1.

The program first reads the information contained in static files. A static file is considered as one that is updated manually by control center personnel, with the exception of the temporary file between the topology processor and state estimator. System files are considered as those that involve updated real-time information. TOPOLOGY contains a list of the elements (circuit-breakers, isolators, lines, transformers, and generators) connected to the busbars of each substation, which are required to determine the system's nodal topology. LINEDATA contains the per-unit parameters of the system lines and transformers. SPECMEAS contains the location and types of several known special measurements within the system. SPIDER is an acronym for the SCADA system the EEF has recently installed. Its real-time file contains the circuit-breaker statuses and measurements provided by the SCADA and is the final input to the topology processor. Since the batch file that executes

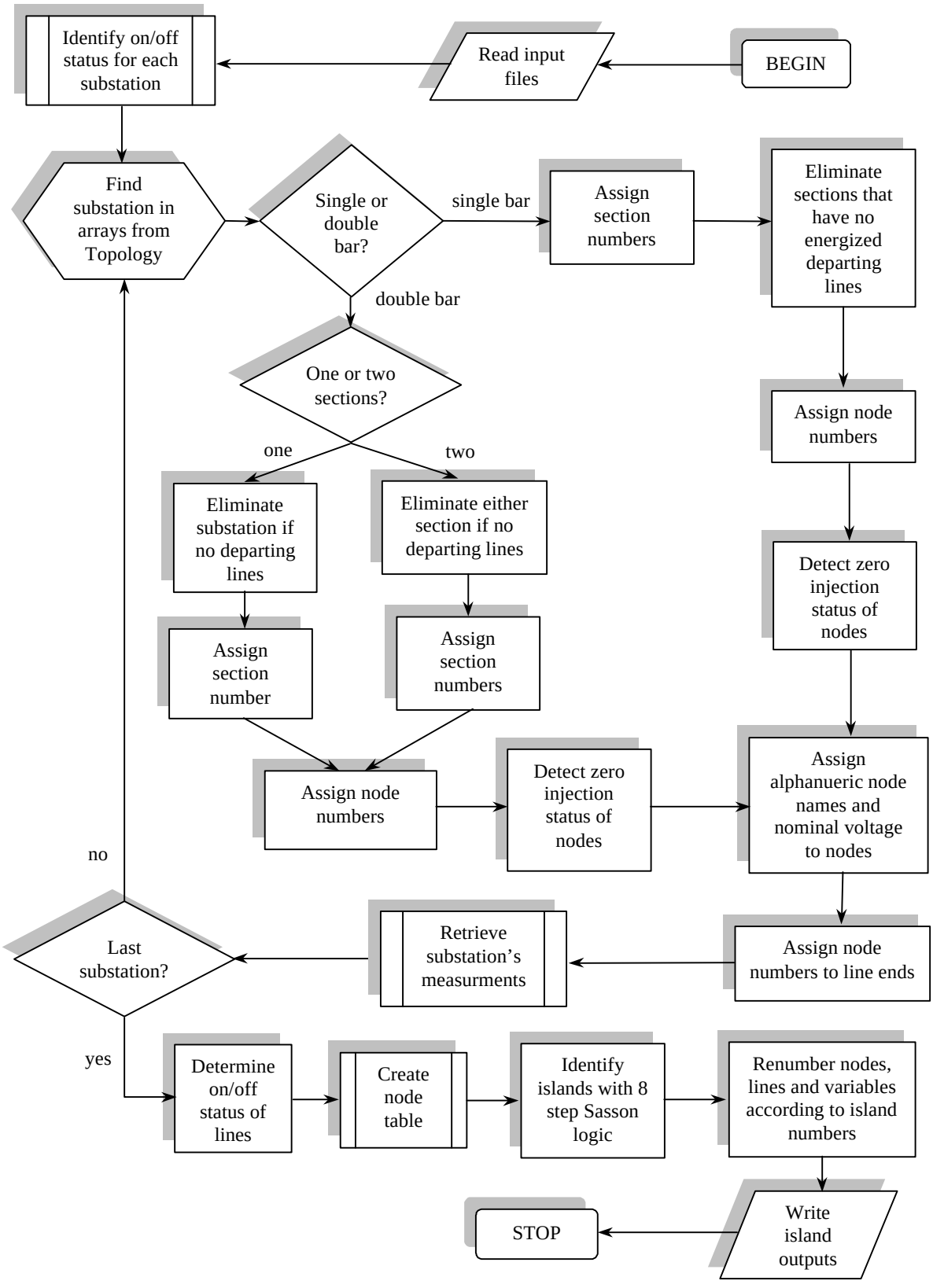

Figure 4.1.2 – Flowchart describing the process followed by the topology processor

the topology processor and state estimator programs is run every 5 minutes, the SPIDER file is made available every 3 minutes to guarantee that fresh data is always available. The data in this file is a snapshot of the available system information in the sense that all the entries in the file are simultaneously received from the remote terminal units.

After receiving its input, the program calls a subroutine to identify the on/off status of each element found in TOPOLOGY. If any element is not properly identified in the SPIDER database, a warning is written to an external file (see Figure 4.1.1). Additional warnings are written if abnormal circuit-breaker/isolator configurations are detected, such as a line's circuit-breaker and ground isolator both being on. These on/off statuses are then passed back to the main program, where a logic loop determines the system supernodes (or simply nodes) derived from the isolated busbars of each substation.

Two distinct classes of substations exist in the EEF system, those with single busbar configurations and those with double (parallel) busbar configurations. The two classes have distinct logics for determining the substation's nodes. In the single busbar configuration logic, the busbar breakers and isolators are used to identify how the main busbar is divided into smaller busbars, and each distinct busbar is assigned a section number. Any section that does not have an energized line or tie-line transformer connected is discarded. The remaining sections are assigned supernode numbers, and any of these which has no generator or load attached is identified as a zero injection bus. In the double busbar configuration logic, the parallel busbars may be tied or separated. If they are tied there is only one section in the substation, and if no energized line or tie-line transformer is attached, the entire substation is disregarded. If the busbars are separated, two sections are identified, and either section is discarded if no energized line or tie-line transformer is attached. In either case, node numbers are assigned to any kept section. If any node has no generation or load attached, it is identified as a zero injection node.

Once the substation's nodes have been identified, alphanumeric node names and nominal voltages are assigned based on fields in TOPOLOGY. Before continuing to the next substation, the line ends attached to each node are assigned the node's number and a

subroutine is called to locate and assign the substation's measurements from the SCADA information. The subroutine that assigns each substation's measurements searches through the substation's SCADA list and assigns power flows and busbar voltages to each line end where they exist. It also uses the information in SPECMEAS to assign additional measurements such as busbar voltages not associated with a line end. In this subroutine, voltage measurements that do not meet reasonable range requirements are not assigned. The algorithm also searches for nodes that have measured generation injection but no load attached, and assigns an injection measurement when such a node is found.

Once all nodes have been identified and all measurements assigned, the system must be divided into electrical islands. The on/off status of the lines are determined by checking the breaker statuses at the line's ends. Next a subroutine is called to create a table that contains all lines and tie-line transformers connected to each node. The eight-step logic used by Sasson et al. [33] to identify circuits (what we have called sections) in a substation through closed circuit-breaker paths has been adapted here to identify the system islands through closed lines (and tie-line transformers). The logic passes from one node to another through closed lines. When no more nodes can be reached through closed lines, an island is formed and the logic is begun again starting with a node not already considered. When all nodes have been included in an island, the process is finished.

After the system's division into islands, the nodes are renumbered for each island, so all related information such as lines and measurements are reorganized into a format for writing each island's information separately. Finally, each island's nodes, line and transformer parameters, and measurements are written island by island into a text file for input into the state estimator. The code also uses and passes along execution information read from the static input files. This information includes convergence requirements and the Huber cutoff value, the maximum number of IRLS iterations allowed, the number of iterations for which the Jacobian matrix is updated, and default standard deviations for different measurement types. In this way this information can be altered by EEF personnel without having to recompile the programs' source codes.

4.2. Updates to the SHGM-Estimator Code

The projection statistics and SHGM-IRLS algorithms had been developed by Cheniae, Vichare, and Mili for research purposes previous to the Virginia Tech/EPFL/EEF collaboration . As mentioned in Section 1.3, the FORTRAN program that implemented the algorithms was designed for an artificial environment, so modifications were made to adapt them for installation on a real system, including: a π-circuit transformer model, the ability to model parallel transmission lines or tie-line transformers, the presence of multiple bus voltage measurements, the independent treatment of active and reactive power measurements, and the automatic calculation of recursive statistics based on the accumulated residuals of the metered locations. Originally, the program included calculations to predetermine the power flows and injections based on a given complex voltage profile, then add Gaussian noise to these generated values in order to create an artificial measurement set. This preprocessing algorithm was removed and made into a separate program, whose format was updated in parallel with that of the state estimator algorithm. This program was used to create the artificial cases for study and testing, such as those seen in Chapter 5.

The π-circuit transformer model of [37], developed for use in load flows, permits the tie-line transformers to be treated identically as lines in the algorithm. These transformers have a finite number (between 18 and 33 in the EEF system) of tap positions that allow the tap ratio to be other than nominal, allowing the transformer to adjust for loading conditions. The model is given in Figure 4.2.1, where a is the reduced tap ratio and $\overline{Y}_{ij}$ is the per-unit leakage reactance. The resistance is neglected and the parameters of the equivalent π-circuit are a function of a, so that the shunt branches are asymmetric. Although this model allows the transformers to be treated as lines, sufficient information to calculate the parameters $\overline{Y}_{ij}$ and a are required from the topology processor, and the transformers are recognized in the input as branches distinctly from lines. Since the shunt branches are asymmetric, it is important to track the node number of the tap-changing side.

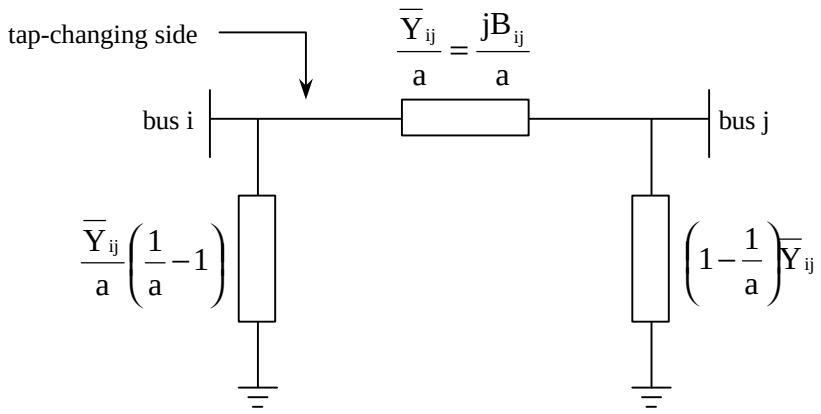

Figure 4.2.1 – π-circuit transformer model

Effort was made to include the tap position of each transformer as a state variable in the state estimation. This is not very difficult and only requires a minor extension of the Jacobian, along with careful attention to properly modify the state variable and measurement vectors. A version was successfully developed and implemented for testing, but the testing revealed that there was insufficient local measurement redundancy in the area of the tie-line transformers to add a state variable (see Section 5.4).

The capability of handling parallel lines was added to the algorithm through the assignment of an identification number to each line, then associating these line numbers with the flow measurements. The program originally grouped active and reactive power measurements as pairs, so the ability to treat them independently was added by splitting the arrays storing power measurements into separate arrays, reordering and adding indices where needed. This ability becomes important when reactive power measurements are not installed or either measurement is unavailable at a metering location. The capability of handling multiple voltage magnitude measurements at each bus was implemented by changing the indexing of the voltage measurement array. Once the program was ready to receive input from the topology processor, an array containing the standard deviations of each measuring device, that is, their corresponding entries in the covariance matrix (see the next section), was added so that they could be individually specified in TOPOLOGY.

After the state estimator was fully integrated into and tested in the real-time EEF environment, a subroutine was included to recursively calculate the sample mean and sample variance of the accumulated residuals of the metered locations. After the state estimate has been produced every five minutes, the statstical information of each meter location of each entry of the residual vector is searched in the file STATISTICS, and updated. This file is stored and the information reset at the beginning of each month so that a monthly record of the statistics is kept for analysis. An analysis of the residual samples to investigation the justification of robust estimators is considered as a future work. The results of the statistics have so far not been used, but the application concept is to alter the measurement error standard deviation in TOPOLOGY for measurements that consistently show a larger location and spread of residuals than other measurements in the network.

4.3. Projection Statistics

In Section 3.3, the calculation of the projection statistics was derived. Now we show how they robustify M-estimators, by identifying leverage points. Rousseeuw and Van Zomeren [31] reported that the projection statistics roughly follow a chi-squared distribution. Mili et al. [25] demonstrated that they behave as a chi-squared distribution with ν degrees of freedom, where ν is the dimension of the relevant subspace of $\underline{\ell}_i$. Hence any $\underline{\ell}_i$ that satisfies

$$PS_i > b_i = \chi^2_{\nu,0.975} \tag{4.1}$$

may be identified as a leverage point, since it lies outside the ellipsoid.

These cutoff values b_i help to define the projection statistic weights as

$$w_i = \min\left\{1, \left(\frac{b_i}{PS_i}\right)^2\right\}. \tag{4.2}$$

When PS_i is sufficiently large, w_i will be less than 1. These weights are used to robustify an M-estimator by their inclusion in the objective function $J(\underline{x})$ in a Schweppe-type GM-estimator [17], [18]. Such an estimator minimizes

$$J(\underline{x}) = \sum_{i=1}^{m} w_i^2 \rho\left(\frac{r_i}{\sigma_i w_i} \right) \tag{4.3}$$

and so is a solution of

$$\sum_{i=1}^{m} w_i \underline{\ell}_i \psi\left(\frac{r_i}{\sigma_i w_i} \right) = \underline{0}. \tag{4.4}$$

The weights w_i cancel in (4.4) where the residuals fall in the quadratic region of a nonquadratic M-esimator's ψ-function, because in this region $\psi(r_i / (\sigma_i w_i)) = r_i / (\sigma_i w_i)$. The result is that any small standardized residual $r_{Si} = r_i / (\sigma_i w_i)$ is not downweighted regardless of whether it is a leverage point, so that good leverage points are not downweighted. The chosen ψ-function is that of the Huber M-estimator introduced in Section 3.2, given with r_{Si} by

$$\psi(r_{Si}) = \begin{cases} r_{Si}, & |r_{Si}| \le b \\ b \ \mathrm{sign}(r_{Si}), & \text{elsewhere} \end{cases}, \tag{4.5}$$

and with the corresponding ρ-function given by

$$\rho(r_{Si}) = \begin{cases} \frac{1}{2} r_{Si}^2, & |r_{Si}| \le b \\ b|r_{Si}| - b^2/2, & \text{elsewhere} \end{cases}. \tag{4.6}$$

Hence, when $w_i < 1$ and $r_{Si} > b$, point $\underline{\ell}_i$ is considered as a bad leverage point and w_i is incorporated into the downweighting of the point's influence on the estimation results. From a practical point of view, a lower limit may be chosen for w_i to avoid numerical instabilities in the estimator's iterative solution.

4.4. Iteratively Reweigthed Least Squares (IRLS)

The IRLS algorithm was chosen to find a solution for the SHGM-estimator largely because Newton's method is more prone to numerical problems in the context of state estimation [25]. It is derived from the Schweppe-type GM-estimator as follows. In (4.4), divide and multiply the ψ-function by r_{Si}. Then in matrix form, we have

$$\underline{H}^T \underline{R}^{-1} \underline{Q} \underline{r} = \underline{0}, \tag{4.7}$$

where the weight function $q(r_{Si}) = \psi(r_{Si}) / r_{Si}$ corresponds to the weight matrix $\underline{Q} = \text{diag}(r_{Si})$, and the residual vector is given by $\underline{r} = \underline{z} - \underline{h}(\underline{x})$. The design matrix $\underline{h}(\underline{x})$ is linearized by performing a first-order Taylor series expansion about an operating point of the state vector, $\underline{x}^{(k)}$, yielding

$$\underline{h}(\underline{x}) = \underline{h}\left(\underline{x}^{(k)}\right) + \underline{H}\left(\underline{x}^{(k)}\right)\left(\underline{x} - \underline{x}^{(k)}\right). \tag{4.8}$$

Incorporating (4.8) and the fact that $\underline{r}^{(k)} = \underline{z} - \underline{h}\left(\underline{x}^{(k)}\right)$ into (4.7), we obtain

$$\Delta \underline{x}^{(k)} = \left(\underline{H}^{(k)T}\, \underline{R}^{-1}\, \underline{Q}^{(k)}\, \underline{H}^{(k)}\right)^{-1} \underline{H}^{(k)T}\, \underline{R}^{-1}\, \underline{Q}^{(k)}\, \underline{r}^{(k)}. \tag{4.9}$$

The iterations continue until a tolerance $\left|\underline{x}^{(k+1)} - \underline{x}^{(k)}\right| < \varepsilon$ is satisfied. The implementation of the robustified SHGM-estimator lies in the $\underline{Q}$ matrix. To illustrate, if $r_{Si} < b$ for i = 1,…,*m*, then $\underline{Q}$ becomes the identity matrix and the algorithm is identical to the conventional algorithm used to solve the WLS. The matrix $\left(\underline{H}^{(k)T}\, \underline{R}^{-1}\, \underline{Q}^{(k)}\, \underline{H}^{(k)}\right)^{-1}$ is referred to as the gain matrix and must be numerically solved, with a proper balance between numerical stability and computational efficiency. A simple LU decomposition and backsubstitution is found to be sufficient for small systems such as the EEF. For larger systems, sparsity techniques and possibly other numerical tools may need to be applied.

In the SHGM-estimator program installed at the EEF control center, the projection statistics weights are calculated in a subroutine whose inputs are the Jacobian matrix and its dimensions. The weights are calculated once, before the initial iteration of the IRLS solution. They could in fact be recalculated at each iteration as the Jacobian is updated.

4.5. Overall SHGM-IRLS Algorithm

The following pseudo-code provides a rough summary of the sequence of steps taken by the state SHGM-estimator as installed in the EEF EMS. M represents the number of islands, σ_i is the standard deviations equal to the square-root of the diagonal elements of the covariance matrix $\underline{R}$, max_it is the maximum number of iterations allowed the IRLS algorithm, ε_θ is the phase angle convergence tolerance, ε_V is the voltage magnitude

convergence tolerance. The vector $(\underline{A}^{-1}\underline{b})_\theta$ represents the correction vector entries corresponding to the phase angle state variables, while $(\underline{A}^{-1}\underline{b})_V$ represents the correction vector entries corresponding to the voltage magnitude state variables.

for L =1:M, **do**

input line parameters, $\underline{z}$, $\underline{x}^{(0)}$, σ_i

Calculate $\overline{Y}_{bus}$

$\underline{x}^{(0)} \leftarrow$ flat start

Calculate $\underline{H}(\underline{x}^{(0)})$

for i = 1:m, **do**

$\nu \leftarrow$ dimension (relevant subspace of $\underline{\ell}_i$)

$b_i \leftarrow \chi^2_{\nu,0.975}$

$$S'_m \leftarrow 1.1926\left(\underset{i}{\text{lomed}}\ \underset{j\neq i}{\text{lomed}}\left|\underline{\ell}_i^T\underline{v}+\underline{\ell}_j^T\underline{v}\right|\right)$$

$$PS_i \leftarrow \max_{\underline{v}}\frac{\underline{\ell}_i^T\underline{v}}{S'_m},\ \underline{v}=\underline{\ell}_j,\ j=1,\ldots,m$$

$$w_i \leftarrow \min\left\{1,\left(\frac{b_i}{PS_i}\right)^2\right\}$$

end

for k = 1:max_it, **do**

Calculate $\underline{H}(\underline{x}^{(k)})$

$\underline{r}^{(k)} \leftarrow \underline{z}-\underline{h}\left(\underline{x}^{(k)}\right)$

for i = 1:m, **do**

$$\underline{Q}(i,i)=\begin{cases}1, & r_i/(\sigma_i w_i)\le b\\ |b/r_{Si}|, & r_i/(\sigma_i w_i)>b\end{cases}$$

end

$\underline{A} \leftarrow \underline{H}^{(k)T}\underline{R}^{-1}\underline{Q}^{(k)}\underline{H}^{(k)}$

Check conditioning of $\underline{A}$

$\underline{b} \leftarrow \underline{H}^{(k)T} \underline{R}^{-1} \underline{Q}^{(k)} \underline{r}^{(k)}$

$\underline{x}^{(k+1)} \leftarrow \underline{x}^{(k)} + \underline{A}^{-1}\underline{b}$ (by LU decomposition and back substitution)

if $(||(\underline{A}^{-1}\underline{b})_\theta||_\infty < \varepsilon_\theta)$ **&** $(||(\underline{A}^{-1}\underline{b})_V||_\infty < \varepsilon_V)$, **break**

end

$\underline{r}^{(k)} \leftarrow \underline{z} - \underline{h}\left(\underline{x}^{(k)}\right)$

output $\underline{r}^{(k)}$

Calculate $\underline{h}(\underline{x}^{(k)})$

output $\underline{h}(\underline{x}^{(k)})$

output Load Flow format

input recursive statistics

Calculate new recursive statistics

output recursive statistics

Reset variables

end

5 – SIMULATION RESULTS

This chapter contains simulation results based on a typical island topology of the EEF system. In order to simulate a real-time scenario, loading conditions are specified through calculations based on the true voltage profile $\underline{x}$. Next these conditions are used to generates a coherent measurement set $\underline{z}$ based on the real EEF measurement configuration. Gaussian noise is then added to $\underline{z}$ according to the values of σ_i, and outliers are simulated by manually altering measurement values before input to the state estimator program. Hence the true and the input measurements are known, which allows for a more precise evaluation of the estimators. The program used to prepare the data is introduced in Section 4.2.

Section 5.1 provides the model of the EEF system. One of the islands in this model is then used in simulations of the WLS state estimator in Section 5.2, and the SHGM-estimator in Section 5.3. Meter placement in the system is discussed in Section 5.4.

5.1. EEF System Model

The principle EEF network consists of 33 substations interconnected by 62.5kV subtransmission lines, which is externally connected by one 130kV- and two 220kV-tie lines owned by the Swiss company EOS (Energie Ouest Suisse). The three tie lines, located at the Monteynan, Botterens and Villarepos substations, allow the network to be subdivided into three interconnectable islands that in turn transmit energy to loads through 18kV distribution networks. The EEF's standard operating practice is to divide the system into the three islands, which may be referred to as Monteynan island, Botterens island, and Villarepos island. The typical topologies and measurement configurations of these islands are depicted in Figure 5.1.1, Figure 5.1.2, and Figure 5.1.3, respectively. In a typical estimation, the state estimator receives some variation of these three islands that together typically contain 57 voltage measurements, 54 pairs of real and reactive power flow measurements, only two or three power injection measurements, and several zero injection buses, yielding a total of about 170 to 180 measurements.

The topology for Monteynan island shown in Figure 5.1.1 consists of two meshed subnetworks connected by a single link, which is made up of two transformers. This island usually consists of 20 buses, 2 transformers and 26 lines, four of which are very short with relatively small reactances. These lines are OELB-MAIG, OELB-STL2, HV11-HV31 and HV12-HV32. Table 5.1.1 shows that the projection statistics as identified by the program's projection algorithm correctly identifies the corresponding power flow measurements as leverage points.

Zero injection buses are not true measurements, but constraints in the power balance equations of the system. However, the constraints are closely approximated in the estimation by setting the injections' σ_i very small and their projection statistics weights to 1.0 so that they are not rejected as measurements. The σ_i cannot be set too small or they result in numerical instability. Three of the zero injection buses that exist in Monteynan island are not shown in Figure 5.1.1, because each has only two adjacent lines and contains no measurements. This is discussed in Section 5.4.

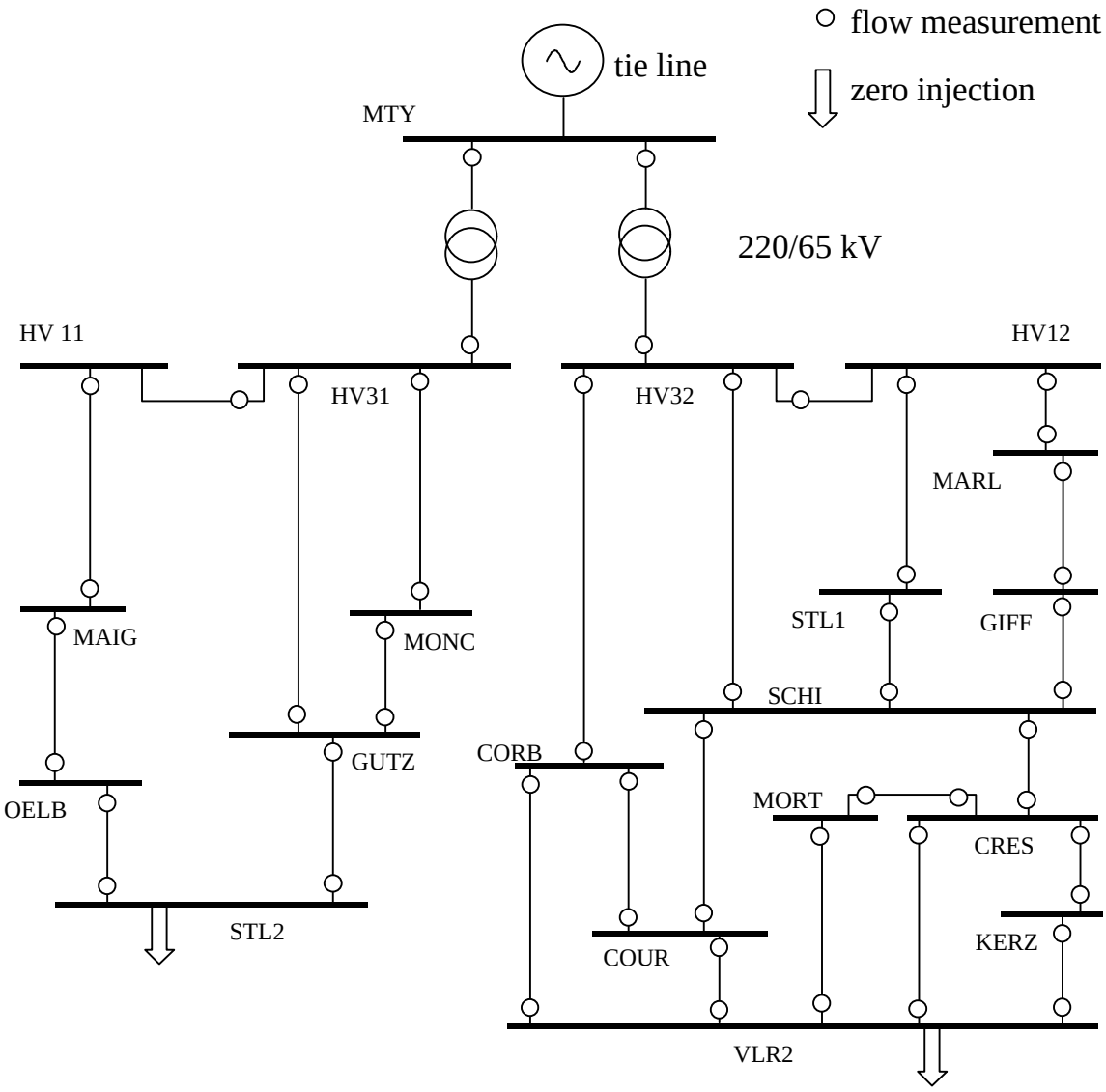

Figure 5.1.1 – Monteynan island

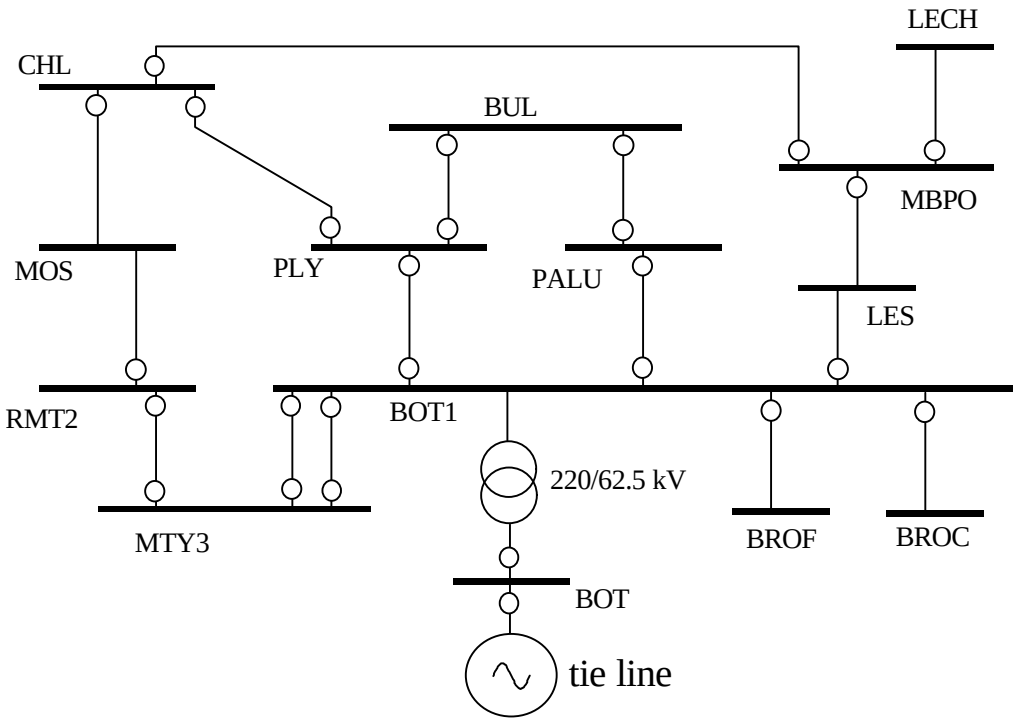

Figure 5.1.2 – Botterens island

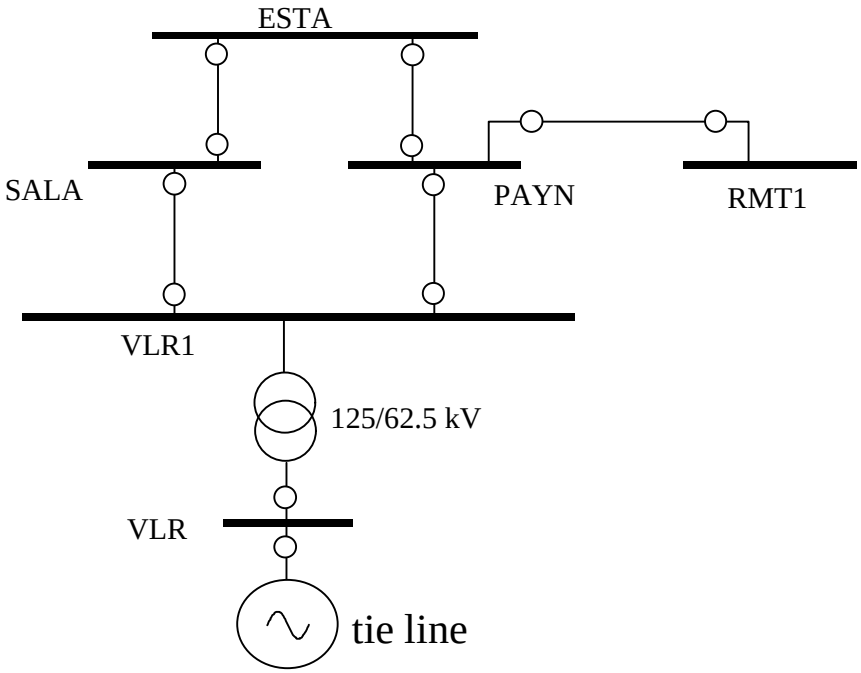

Figure 5.1.3 – Villarepos island

Table 5.1.1 – Projection statistics of active power flow measurements.

Active Power Measurement	Branch Reactance	Projection Statistic	Weight w_i
CORB - HV32	0.04234	1.866	1.000
CORB - VLR2	0.09442	3.230	1.000
CORB – COUR	0.03812	5.791	1.000
COUR - CORB	0.05179	3.836	1.000
COUR - VLR2	0.05555	6.561	1.000
COUR - SCHI	0.08349	2.347	1.000
CRES - KERZ	0.11665	1.452	1.000
CRES - MORT	0.04649	2.343	1.000
CRES - SCHI	0.04457	5.361	1.000
CRES - VLR2	0.06249	5.672	1.000
GIFF - MARL	0.08166	3.575	1.000
GIFF - SCHI	0.18870	1.394	1.000
GUTZ - STL2	0.06376	5.151	1.000
GUTZ - HV31	0.05345	2.272	1.000
GUTZ - MONC	0.01365	8.914	1.000
HV11 - MAIG	0.04806	4.606	1.000
HV12 - STL1	0.07548	5.866	1.000
HV12 - MARL	0.01924	10.106	1.000
HV32 - CORB	0.04234	2.653	1.000
HV32 – MTY	0.12914	3.047	1.000
HV31 – MTY	0.14016	2.735	1.000
HV31 - MONC	0.03671	6.618	1.000
HV31 - GUTZ	0.05345	2.272	1.000
HV31 - HV11	0.00270	76.966	0.021
HV32 - SCHI	0.15967	1.300	1.000
HV32 - HV12	0.00270	76.923	0.021
KERZ - CRES	0.11665	1.452	1.000
KERZ - VLR2	0.21621	1.381	1.000
MAIG - HV11	0.04806	4.606	1.000
MAIG - OELB	0.00795	21.639	0.265
MARL - HV12	0.01924	10.106	1.000
MARL - GIFF	0.08166	3.575	1.000
MONC - GUTZ	0.01365	8.914	1.000
MONC - HV31	0.03671	6.618	1.000
MORT - CRES	0.04649	2.343	1.000
MORT - VLR2	0.06130	4.565	1.000
MTY - HV32	0.12914	3.047	1.000
MTY - HV31	0.14016	2.735	1.000
OELB - MAIG	0.00795	21.639	0.265
OELB - STL2	0.02891	14.586	0.583
SCHI - COUR	0.08349	2.347	1.000
SCHI - HV32	0.15967	1.300	1.000
SCHI - CRES	0.04457	5.361	1.000
SCHI - STL1	0.06686	2.830	1.000
SCHI - GIFF	0.18870	1.394	1.000
STL1 - SCHI	0.06686	2.830	1.000
STL1 - HV12	0.07548	5.866	1.000
STL2 - OELB	0.02891	14.586	0.583
STL2 - GUTZ	0.06376	5.151	1.000
VLR2 - MORT	0.06130	4.565	1.000
VLR2 - CRES	0.06249	5.672	1.000
VLR2 - COUR	0.05555	6.561	1.000
VLR2 - CORB	0.09442	3.230	1.000
VLR2 - KERZ	0.21621	1.381	1.000

5.2. Observability and Local Redundancy

The concepts of observability and local redundancy in power system state estimation will be useful in the discussions of the following sections. The condition of observability has so far been left as an underlying assumption of state estimation. A system containing m measurements and n = 2N-1 state variables is observable when the measurements are spread throughout the system so that the Jacobian matrix is of full rank:

$$\text{rank}\left(\underline{\mathrm{H}}^{\mathrm{T}}\,\underline{\mathrm{H}}\right) = \text{rank}\left(\underline{\mathrm{H}}\right) = \min\left(m, n\right) = n \qquad (5.)$$

Since the gain matrix shown in (4.9) is singular when this condition is not met, the system clearly must be observable for a state estimation solution to be found.

The fundamental set of a state variable is the set of all measurements that have nonzero terms in the column of $\underline{\mathrm{H}}$ that correspond to the state variable. Hence when each state variable has a fundamental set containing at least one measurement, $\underline{\mathrm{H}}$ is of full rank and the system is observable. This implies that each of the state variables is observed by at least one measurement in an observable system. Obviously, for the purposes of state estimation, we would like to have large fundamental sets for each state variable in order to provide the best estimate for its location.

The local redundancy of a measurement is defined as the minimum number of other measurements that when deleted make that measurement a critical measurement [6], [7]. A critical measurement is any measurement that when deleted, makes the Jacobian matrix rank deficient. A critical pair of measurements is any pair that when either of the measurements are deleted, the other becomes critical. In general, a critical p-tuple is any set of p measurements that when p-1 are deleted, the last becomes a critical measurement. [6], [7] The local redundancy of a critical measurement is therefore zero, that of a measurement belonging to a critical pair is one, and in general the local redundancy of a measurement belonging to a critical p-tuple is p-1.

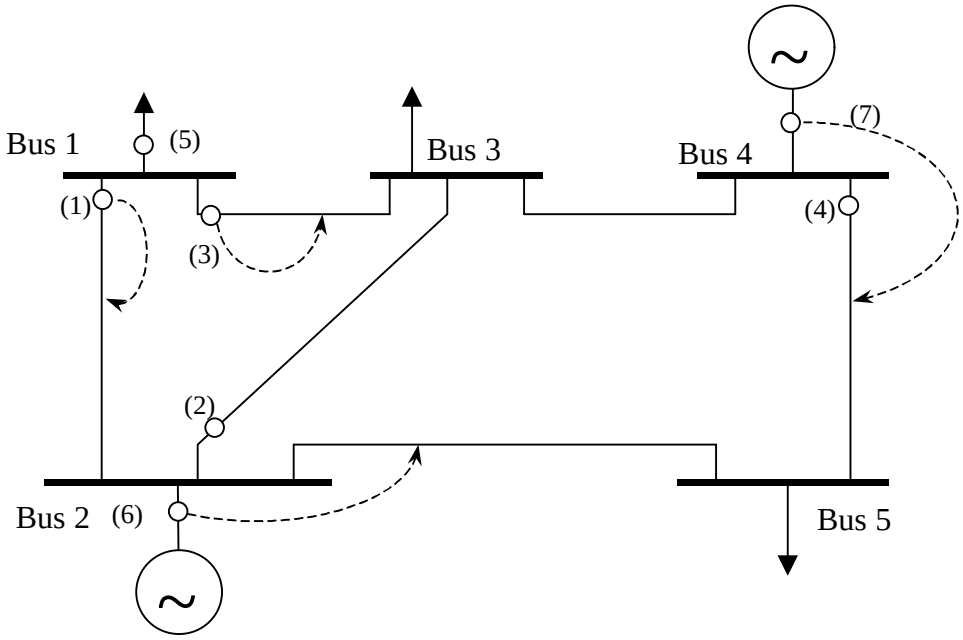

Figure 5.2.1 – Example 5-bus system

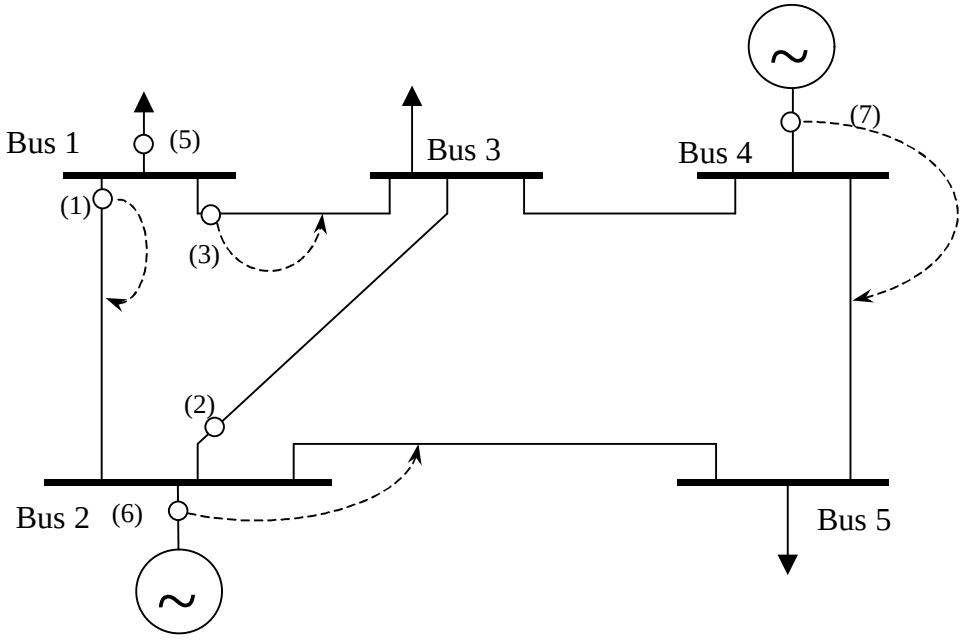

Figure 5.2.2 – Example 5-bus system with measurement (4) deleted

Figure 5.2.1 and Figure 5.2.2 illustrate the concept of critical measurements and critical pairs. If we consider the dc model, there are $m = 7$ active power measurements that observe the N-1 = 4 phase angle state variables. The set of measurements $\Im$ = {(1), (3), (6), (7)} are an example of a set of measurements that are sufficient to observe the system, that is, the other measurements may be deleted without making the Jacobian rank deficient. In order to demonstrate that the measurements in $\Im$ observe all of the state variables, arrows are drawn from each of its measurements to the branch through which a new state variable is observed. We may note that each of the measurement sets {(4), (6)}, {(4), (7)}, and {(6), (7)} form a critical pair. If measurement (4) is deleted as in Figure 5.2.2, measurements (6) and (7) both become critical measurements.

5.3. Weighted Least Squares

The WLS-based algorithm is a classic estimator and a very popular choice in power system state estimation. However, it is clear from robust estimation theory that it is not robust to outliers in the data. When data measurements deviate significantly from a Gaussian distribution, the WLS estimate immediately begins to break down. Since in power systems outlying measurements occur on a regular basis [39], it is important to protect the state estimator from breaking down. The estimators must perform in the presence of these outliers, and the low breakdown point and smearing effect of the WLS-based algorithm are easily demonstrated. The WLS-based algorithm is implemented using the IRLS algorithm of the SHGM method, with the Huber cutoff chosen at 10,000 and the projection statistics subroutine bypassed. Hence, unless $r_i / \sigma_i > 10{,}000$, the IRLS algorithm acts exactly as a WLS-based algorithm would.

The Monteynan island was chosen for demonstration using the WLS, with only three non-interacting bad data introduced. The active power flows CRES-KERZ and OELB-MAIG, and an active power injection at CORB were given arbitrarily large values, where the absolute value of the increased value is at least $5\sigma_i$. The injection at bus CORB is not usually available in the EEF network, but was chosen here to more clearly demonstrate the

smearing effect. A solution was reached in 2 iterations, and the results are recorded in Table 5.3.1.

In Table 5.3.1, the columns marked error/σ_i denote the absolute difference between the measurement's true value and the value received by the state estimator program, divided by its standard deviation so that the values will be on an even scale. The weighted residuals $r_{wi} = r_i / \sigma_i$ (σi = 0.01 p.u. for all power measurements) reflect the difference between the final estimated power flow or injection value and the "measured", or input, value. Therefore, the bad data points are identified by the value error/σ_i and should have large residuals, whereas a relatively small value of error/σ_i only reflects Gaussian noise and should not have a large weighted residual.

The non-robust performance of the WLS is evident in the neighborhood of each bad data point. The smearing effect is quite obvious in the active power flow measurements on the lines incident to CORB, i.e., flow measurements CORB-HV32 and HV32-CORB, CORB-COUR, COUR-CORB, and somewhat in CORB-VLR2 and VLR2-CORB. Some smearing also seems to be going on in the direction of COUR and VLR2, although it is difficult to say because measurements VLR2-COUR and VLR2-CORB have relatively large Gaussian noise. The WLS also fails to properly reject the bad data point at CRES-KERZ. It has a large residual at CRES-KERZ but spreads the error to the measurement at the other side of the line, rejecting a good measurement.

The same phenomenon is observed at the bad data point OELB-MAIG, although we cannot expect any M-estimator to perform well at a bad leverage point without a method to counter the leverage. In all of these cases, the WLS breaks down because in the estimation of location of the voltage magnitude the estimate is moved by the influence of one data point. In the cases of bad data at CORB and CRES, the local redundancy is high for a power system, reinforcing the fact that the WLS cannot handle even one arbitrary gross error. Table 5.3.2 provides the true, noisy, and estimated values in MW of the power measurements involved in the smearing effect.

Table 5.3.1 – Weighted residuals provided by the WLS estimator.

Measurement	Active Power		Reactive Power	
	error/σi	r_{wi}	error/σ_i	r_{wi}
CORB - HV32	2.306	21.648	.531	.239
CORB – VLR2	.607	8.779	.649	-.701
CORB – COUR	.496	13.973	.028	.133
COUR – CORB	.320	-14.169	.395	-.661
COUR – VLR2	.853	-6.985	.341	-.232
COUR - SCHI	.719	-1.035	.340	.739
CRES - KERZ	49.100	32.975	.414	.172
CRES - MORT	.915	-.202	.513	.234
CRES - SCHI	2.062	-2.964	1.238	.685
CRES – VLR2	.367	-6.027	1.441	-2.278
GIFF - MARL	1.073	-1.465	1.943	-.530
GIFF - SCHI	.334	2.050	.443	-.469
GUTZ – STL2	.171	-1.113	.484	.687
GUTZ - HV31	1.296	1.593	.409	-.374
GUTZ - MONC	1.042	.400	.069	.598
HV11 - MAIG	.745	2.827	.569	-.179
HV12 – STL1	1.122	3.128	1.897	1.142
HV12 - MARL	1.681	-.817	1.545	-.402
HV32 - CORB	1.334	-18.152	1.167	1.201
HV31 - MONC	1.091	.418	.898	1.092
HV31 - GUTZ	1.162	.864	1.216	-1.230
HV31 - MTY	.664	-.694	.526	-.402
HV31 – HV11	.820	.188	.077	-.061
HV32 – SCHI	.068	2.438	1.474	.822
HV32 - HV12	1.308	.246	.735	.028
HV32 – MTY	1.652	1.552	.772	.649
KERZ – CRES	1.544	17.424	.144	.013
KERZ – VLR2	.802	6.788	.043	1.422
MAIG – HV11	1.553	-.526	1.272	.913
MAIG – OELB	.053	-25.438	.500	.203
MARL - HV12	.536	-1.398	.718	-.415
MARL – GIFF	.425	.814	.577	-.793
MONC – GUTZ	1.010	-.368	1.794	1.132
MONC - HV31	1.787	2.459	.238	-.412
MORT – CRES	.625	-.084	1.873	-2.602
MORT – VLR2	.770	-6.058	1.960	2.339
MTY - HV32	.838	.937	.088	.047
MTY - HV31	1.297	-1.262	.873	-.990
OELB - MAIG	51.400	-26.029	.133	.394
OELB – STL2	.331	1.271	1.242	-1.440
SCHI – COUR	.043	.274	.466	.101
SCHI - HV32	.431	-2.087	.210	.458
SCHI – CRES	.785	1.692	.675	1.260
SCHI – STL1	.244	-3.196	1.547	-.187
SCHI - GIFF	.874	-2.604	.993	-.930
STL1 - SCHI	1.185	1.746	.026	-1.394
STL1 - HV12	.205	-1.803	.214	.995
STL2 - OELB	.620	-1.560	.609	.820
STL2 - GUTZ	1.515	2.457	.758	.583
VLR2 - MORT	1.129	6.410	.060	-.312
VLR2 - CRES	.776	6.422	1.774	2.595
VLR2 - COUR	2.252	8.379	.578	.473
VLR2 - CORB	1.757	-7.678	1.454	2.741
VLR2 - KERZ	.327	-7.324	.603	-.861
CORB	77.767	-34.572	2.003	.466
MTY	.449	.583	.498	.517
STL2	.003	.042	.040	-.076
VLR2	.020	-.625	.171	-.063

Table 5.3.2 – Smearing effect and breakdown of the WLS estimator

Measurement	P_{true} (MW)	$P_{metered}$ (MW)	$P_{estimated}$ (MW)
CORB	-7.233	-85.000	-50.428
CORB – COUR	-.128	-.624	-14.596
CORB - HV32	-6.690	-4.384	-26.031
CORB - VLR2	-.415	-1.022	-9.801
COUR – CORB	.128	.448	14.617
COUR - VLR2	-.404	-1.256	5.729
CRES – KERZ	2.900	52.000	19.025
CRES - VLR2	2.821	2.454	8.481
HV32 – CORB	6.701	8.034	26.186
KERZ – CRES	-2.893	-1.349	-18.773
KERZ - VLR2	-.746	-1.548	-8.336
MAIG – OELB	-2.400	-2.453	22.985
MORT - VLR2	-1.251	-2.021	4.037
OELB – MAIG	2.400	-49.000	-22.971
VLR2 – CORB	.415	2.172	9.850
VLR2 – COUR	.404	2.656	-5.723
VLR2 – KERZ	.747	1.074	8.398

5.4. Example of a Real-Time Case

In order to demonstrate the real-time operation of the SHGM-estimator in the EEF system, an example has been taken from real-time data and analyzed. This data was captured on Friday, January 8, 1999 at 7:54:31 A.M. (Swiss time), and processed by a FORTRAN program identical to the one currently installed in the EEF control center. The algorithm for each island was executed with a Huber cutoff value of b = 2.7, a voltage convergence tolerance of ε_V = 0.01 p.u., and a phase angle convergence tolerance of ε_θ = 0.1°. The two larger islands converged in three iterations and the smallest in two iterations, and the results are summarized in Table 5.4.1 through Table 5.4.10. In these tables, the metered and estimated values are in MW, Mvar, or kV, whereas the normalized residuals and Q matrix entries are based on per-unit quantities.

The standardized residuals, denoted r_{Si}, are normalized by dividing both by their σ_i and by a robust scale estimate similar to the one given in (3.26). The majority of power flows and injections are assigned a standard deviation of σ_i = 0.025 p.u., while the voltages

receive σ_i = 0.01 p.u. and zero injection measurements receive σ_i = 0.005 p.u. Power flows that are metered as 0 MW or 0 MVar receive a standard deviation of four times the normal value, because we do not place much confidence in such a metered value. Two robust scale estimates exist for the residuals of each island, one for the residuals corresponding to power measurements (S_n), and one for the voltage residuals (S_{nV}). Since the nominal power of the network is 100 MVA, an example calculation of r_{Si} in p.u. is therefore given for an active power flow by

$$r_{Si} = \frac{\left(P_{estimated} - P_{metered}\right)/100}{\sigma_i S_n} , \tag{5.1}$$

and for a 62.5 kV voltage measurement by

$$r_{Si} = \frac{\left(V_{estimated} - V_{metered}\right)/62.5}{\sigma_i S_{nV}} . \tag{5.2}$$

The system is divided into three islands similar to those shown in Section 5.1. The topology of Monteynan island is provided in Figure 5.4.1. Note that a new substation, abbreviated as WUEN, has been installed between SCHI and GIFF. At the Hauterive substation, HV31 and HV11 have been merged into HV 1, and HV32 and HV12 have been merged into HV 2. At the Monteynan substation, the zero injection buses MTY1 and MTY2 have been modeled. In addition, the zero injection bus CIG is modeled between substations COUR and CORB. The results obtained for Monteynan island demonstrate the estimator's ability to provide a good estimate when no bad data are present. It is an example of the typical results returned by the estimator during real-time operation.

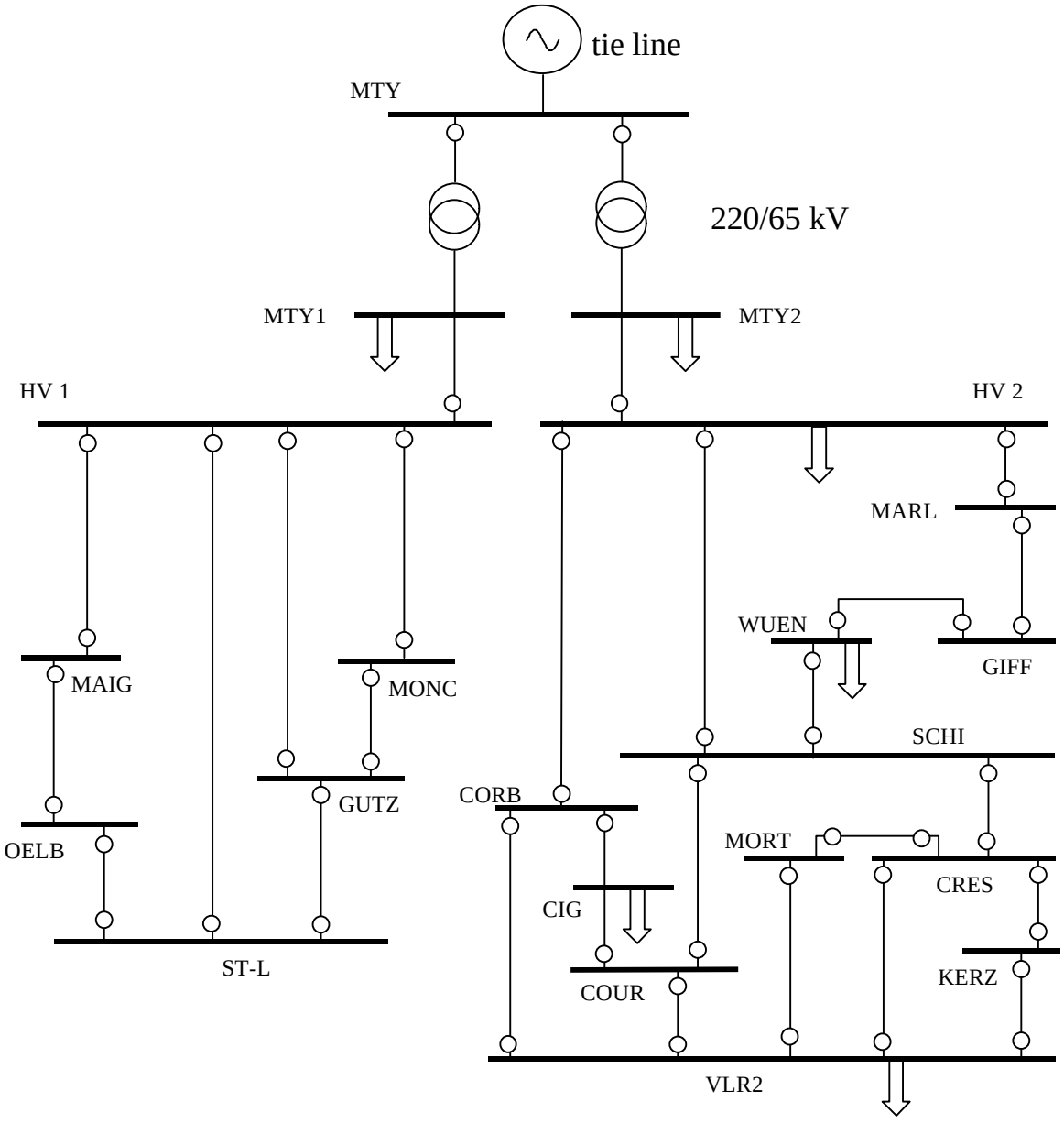

Figure 5.4.1 – Monteynan island topology in real-time case

Table 5.4.1 – Monteynan island active power flows

Measurement Location	Metered Value	Estimated Value	r_{Si}	$\underline{Q}_{i,i}$
CORB-CIG	9.12	8.83	.73	1
CORB-HV 2	-25.92	-25.62	-.75	1
CORB-VLR2	8.64	8.26	.95	1
COUR-CIG	-9.12	-8.79	-.80	1
COUR-VLR2	.00	-.36	.22	1
COUR-SCHI	-3.36	-3.55	.47	1
CRES-KERZ	2.88	2.62	.64	1
CRES-MORT	2.88	2.85	.07	1
CRES-SCHI	-8.64	-8.82	.44	1
CRES-VLR2	-1.92	-1.87	-.12	1
GIFF-MARL	-10.08	-10.96	2.16	1
GIFF-WUEN	1.92	2.30	-.93	1
GUTZ-ST-L	3.36	2.45	2.24	1
GUTZ-HV 1	-11.52	-11.15	-.90	1
GUTZ-MONC	-5.28	-5.23	-.13	1
HV 1-GUTZ	12.48	11.18	3.20	1
HV 1-MAIG	16.32	16.66	-.84	1
HV 1-STL1	9.12	10.00	-2.17	1
HV 1-MONC	14.40	14.33	.18	1
HV 1–MTY1	-68.16	-68.04	-.30	.30
HV 2-CORB	25.92	25.78	.34	1
HV 2-MARL	13.92	14.01	-.23	1
HV 2-SCHI	9.60	9.94	-.85	1
HV 2-MTY2	-48.96	-49.74	1.92	.08
KERZ-CRES	-2.40	-2.62	.53	1
KERZ-VLR2	-1.92	-2.01	.23	1
MAIG-HV 1	-15.84	-16.61	1.89	1
MAIG-OELB	5.28	5.59	-.76	.26
MARL-GIFF	11.52	11.03	1.22	1
MARL-HV 2	-13.92	-13.99	.17	1
MONC-GUTZ	5.28	5.23	.12	1
MONC-HV 1	-14.40	-14.30	-.25	1
MORT-CRES	-2.40	-2.85	1.10	1
MORT-VLR2	-4.32	-3.97	-.85	1
MTY –MTY1	68.16	68.11	.12	1
MTY –MTY2	49.92	49.78	.33	1
OELB-MAIG	-5.28	-5.59	.76	.26
OELB-ST-L	-3.36	-3.12	-.58	1
SCHI-COUR	3.60	3.55	.11	1
SCHI-CRES	9.12	8.83	.71	1
SCHI-WUEN	-1.92	-2.29	.91	1
SCHI-HV 2	-9.84	-9.87	.06	1
ST_L-GUTZ	-3.84	-2.45	-3.42	1
ST_L-HV13	-9.12	-9.97	2.09	1
ST_L-OELB	3.36	3.12	.58	1
VLR2-CORB	-8.04	-8.22	.44	1
VLR2-COUR	.45	.36	.23	1
VLR2-CRES	1.80	1.87	-.17	1
VLR2-KERZ	2.13	2.02	.28	1
VLR2-MORT	4.65	3.98	1.65	1
WUEN-GIFF	-1.62	-2.30	1.67	1
WUEN-SCHI	1.98	2.29	-.76	1

Table 5.4.2 – Monteynan island reactive power flows

Measurement Location	Metered Value	Estimated Value	r_{Si}	$Q_{i,i}$
CORB-CIG	5.28	4.88	.98	1
CORB-HV 2	-11.52	-10.41	-2.73	1
CORB-VLR2	3.60	3.09	1.26	1
COUR-CIG	-4.32	-4.85	1.31	1
COUR-VLR2	.00	.11	-.07	1
COUR-SCHI	-.96	-.52	-1.08	1
CRES-KERZ	.48	.42	.15	1
CRES-MORT	.96	1.25	-.73	1
CRES-SCHI	-2.40	-2.38	-.05	1
CRES-VLR2	-.48	-1.20	1.78	1
GIFF-MARL	-5.28	-5.03	-.61	1
GIFF-WUEN	.48	.91	-1.07	1
GUTZ-ST_L	1.20	1.22	-.05	1
GUTZ-HV 1	-5.52	-5.03	-1.21	1
GUTZ-MONC	-2.16	-2.51	.86	.87
HV 1-GUTZ	6.72	5.06	4.08	1
HV 1-MAIG	7.20	7.69	-1.20	1
HV 1-ST_L	4.32	4.60	-.70	1
HV 1-MONC	5.76	6.47	-1.75	1
HV 1-MTY1	-24.96	-26.48	3.75	.04
HV 2-CORB	11.04	10.70	.85	1
HV 2-MARL	7.20	6.17	2.53	.75
HV 2-SCHI	5.76	6.27	-1.25	1
HV 2-MTY2	-22.08	-23.09	2.48	.06
KERZ-CRES	-.48	-.49	.03	1
KERZ-VLR2	-.96	-.91	-.11	1
MAIG-HV 1	-7.20	-7.57	.92	1
MAIG-OELB	2.40	2.38	.04	1
MARL-GIFF	5.76	5.05	1.74	1
MARL-HV 2	-6.72	-6.15	-1.42	1
MONC-GUTZ	2.40	2.50	-.26	1
MONC-HV 1	-5.76	-6.43	1.65	1
MORT-CRES	-.72	-1.28	1.38	1
MORT-VLR2	-2.64	-1.36	-3.16	1
MTY –MTY1	32.64	33.82	-2.91	1
MTY –MTY2	26.88	26.77	.26	1
OELB-MAIG	-2.16	-2.39	.56	.43
OELB-ST_L	-1.20	-1.36	.40	1
SCHI-COUR	.72	.47	.63	1
SCHI-CRES	2.88	2.38	1.23	1
SCHI-WUEN	-1.44	-1.08	-.89	1
SCHI-HV 2	-5.52	-6.16	1.58	1
ST_L-GUTZ	-1.44	-1.27	-.42	1
ST_L-HV13	-3.84	-4.58	1.82	1
ST_L-OELB	1.44	1.34	.25	1
VLR2-CORB	-3.06	-3.08	.05	1
VLR2-COUR	.45	-.15	1.48	1
VLR2-CRES	.69	1.15	-1.14	1
VLR2-KERZ	.81	.77	.11	1
VLR2-MORT	1.95	1.33	1.53	1
WUEN-GIFF	-1.23	-1.02	-.52	1
WUEN-SCHI	1.14	1.04	.24	1

Table 5.4.3 – Monteynan island power injections

	Active Power				Reactive Power			
Meas. Location	Metered Value	Estimated Value	r_{Si}	$Q_{i,i}$	Metered Value	Estimated Value	r_{Si}	$Q_{i,i}$
CIG	.00	.00	.02	1	.00	-.02	.22	1
CORB	N/A	-8.54	--	--	N/A	-2.44	--	--
COUR	N/A	-12.70	--	--	N/A	-5.26	--	--
CRES	N/A	-5.22	--	--	N/A	-1.91	--	--
GIFF	N/A	-8.66	--	--	N/A	-4.12	--	--
GUTZ	N/A	-13.93	--	--	N/A	-6.32	--	--
HV 1	N/A	-15.87	--	--	N/A	-2.66	--	--
HV 2	.00	.00	.00	1	.00	.05	-.62	1
KERZ	N/A	-4.63	--	--	N/A	-1.41	--	--
MAIG	N/A	-11.02	--	--	N/A	-5.19	--	--
MARL	N/A	-2.96	--	--	N/A	-1.09	--	--
MONC	N/A	-9.07	--	--	N/A	-3.92	--	--
MORT	N/A	-6.82	--	--	N/A	-2.64	--	--
MTY1	.00	.00	-.03	1	.00	.00	.02	1
MTY2	.00	.00	.04	1	.00	.05	-.58	1
MTY	N/A	117.89	--	--	N/A	60.59	--	--
OELB	N/A	-8.71	--	--	N/A	-3.75	--	--
SCHI	N/A	.23	--	--	N/A	-4.39	--	--
ST_L	N/A	-9.30	--	--	N/A	-4.51	--	--
VLR2	.00	.01	-.08	1	.00	.02	-.20	1
WUEN	.00	-.01	.09	1	.00	.02	-.27	1

Table 5.4.4 – Monteynan island voltages

Bus	$V_{estimated}$	$\theta_{estimated}$	$V_{metered}$	r_{Si}	$Q_{i,i}$
CIG	63.35	-4.04	N/A	--	--
CORB	63.53	-3.89	63.36	-.28	1
CORB			63.62	.14	1
CORB			63.36	-.28	1
COUR	63.10	-4.25	63.36	.42	1
COUR			63.36	.42	1
COUR			63.62	.84	1
CRES	63.03	-4.29	62.83	-.32	1
CRES			63.36	.53	1
CRES			63.36	.53	1
CRES			62.83	-.32	1
GIFF	63.40	-3.88	63.89	.79	1
GIFF			63.62	.37	1
GUTZ	63.18	-5.42	63.62	.70	1
GUTZ			63.62	.70	1
GUTZ			63.10	-.14	1
HV 1	63.48	-5.14	63.89	.66	1
HV 1			63.89	.66	1
HV 1			64.15	1.08	1
HV 1			63.89	.66	1
HV 1			63.62	.24	1
HV 1			63.36	-.18	1
HV 2	64.16	-3.42	64.68	.83	1
HV 2			64.68	.83	1
HV 2			64.42	.41	1
HV 2			64.42	.41	1
HV 2			64.15	-.01	1
KERZ	62.89	-4.44	62.83	-.08	1
KERZ			62.83	-.08	1
MAIG	63.08	-5.52	62.57	-.82	1
MAIG			62.83	-.40	1

Table 5.4.4 – Monteynan island voltages (continued)

Bus	$V_{estimated}$	$\theta_{estimated}$	$V_{metered}$	r_{Si}	$Q_{i,i}$
MARL	63.99	-3.53	63.62	-.58	1
MARL			63.62	-.58	1
MONC	63.22	-5.39	62.83	-.62	1
MONC			62.83	-.62	1
MORT	62.93	-4.34	63.10	.27	1
MORT			63.36	.69	1
MTY	234.71	0.00	234.43	-.13	1
MTY1	63.57	-5.09	63.62	.09	1
MTY2	64.23	-3.39	64.42	.30	1
OELB	63.06	-5.54	63.10	.06	1
OELB			61.78	-2.05	1
SCHI	63.19	-4.09	63.10	-.15	1
SCHI			62.83	-.57	1
SCHI			62.30	-1.42	1
SCHI			62.83	-.57	1
SCHI			62.04	-1.84	1
SCHI			62.30	-1.42	1
ST-L	63.10	-5.49	63.36	.41	1
ST-L			63.36	.41	1
ST-L			63.36	.41	1
ST-L			60.98	-3.39	.80
ST-L			63.10	-.01	1
VLR2	63.10	-4.24	63.20	.15	1
VLR2			63.20	.15	1
VLR2			63.49	.63	1
VLR2			63.26	.26	1
VLR2			63.13	.05	1
VLR2			63.33	.36	1
WUEN	63.25	-4.04	63.66	.66	1
WUEN			63.49	.16	1

The topology of Botterens island for this case is shown in Figure 5.4.2. The reactive flow BOT2-BOT, from the transformer secondary to the primary, is transmitted by the SCADA as 0 MVar, whereas the flow on the primary side is measured as 22.44 MVar, and the injected the injected value at bus BOT is 22.68 MVar. The estimator corrects the zero flow at the secondary, estimating the value according to the transformer parameters, as seen in Table 5.4.5. The reactive power measurement set is reinforced in this area by the voltage magnitude measurements at BOT and BOT2. It is notable that this measurement is not downweighted by the estimator, that is, its entry in the $\underline{Q}$ matrix is 1. This is because it has been downweighted in advance by increasing the location's σ_i by a factor of 4 (since the metered value was zero).

The $\underline{Q}$ matrix entries of many of the flow measurements are quite small despite the fact that their residuals are not large. Recalling the definition of $\underline{Q}$ from Section 4.3, we may note that this is due to the fact that leverage points have been identified in the network. The projection statistics for reactive flows in this island are different for those of the active flows because the measurement configuration is different for active and reactive flows: the line flows leaving bus BUL have active but not reactive measurements (see Table 5.4.5). The flow measurement PLY-BUL has a high normalized residual, which may lead us to suspect it as a borderline bad data. In this case it is possible that the voltages, which have a smaller σ_i, have adjusted the estimated value of the reactive flow. It is also possible that the admittance parameters of these lines are not precise. For whatever reason, this measurement is an examples of a borderline case, meaning that it breaks Gaussianity but is not a gross error. This shows the advantage of the way in which the Huber ψ-function gradually downweights large residuals. Under the treatment of other M-estimators or techniques such as the WLS with residual analysis, a measurement with such a borderline residual may experience hard rejection.

The topology of Villarepos island is the same as that in Figure 5.1.3 except that there are two transformers in parallel between VLR and VLR1. The algorithm converged in two iterations and the results are recorded in Table 5.4.8 through Table 5.4.10. What is notable about this case is that the tap positions of the parallel transformers were not found by the

topology processor, and nominal values were therefore assumed. Since the voltage magnitudes of VLR and VLR1 are near the nominal values (according to their estimated values), this was a good assumption and no metered values were rejected. Normally the power injection and several bus voltages are available at bus VLR. This is indicative that the SPIDER database has been updated without updating the supporting files (see Figure 4.1.1).

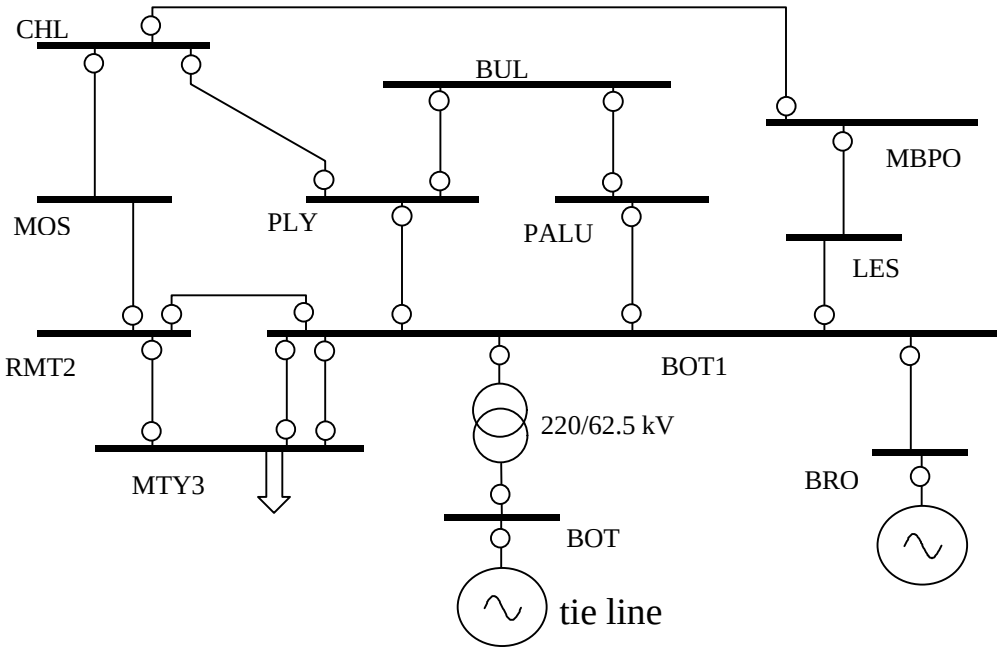

Figure 5.4.2 – Botterens island topology in real-time case

Table 5.4.5 – Botterens island power flows

	Active Power (MW)				Reactive Power (MVar)			
Measurement Location	Metered Value	Estimated Value	r_{SNi}	$Q_{i,i}$	Metered Value	Estimated Value	r_{SNi}	$Q_{i,i}$
BOT1-PLY	11.55	11.74	-.49	1	4.98	4.50	1.27	.13
BOT1-BRO	.03	.00	.08	1	.00	.00	.00	1
BOT1-PALU	14.97	14.80	.45	1	5.07	3.97	2.93	.04
BOT1-LES	-2.85	-2.73	-.32	1	2.49	2.10	1.05	1
BOT1–RMT2	7.56	7.41	.40	1	2.79	2.83	-.11	1
BOT1-MTY3	3.42	2.89	.33	1	1.08	1.02	-.34	1
BOT1-MTY3	3.42	3.30	.33	1	1.08	1.21	-.34	1
BOT1–BOT	-44.64	-44.75	.31	1	.00	-19.04	12.67	1
BOT –BOT1	44.34	44.75	-1.10	1	22.44	22.29	.40	1
BRO -BOT	N/A	.00	--	--	N/A	.00	--	--
BUL -PLY	1.13	.88	.65	1	N/A	3.63	--	--
BUL –PALU	-7.23	-6.95	-.74	1	N/A	-3.54	--	--
CHL -PLY	-1.89	-2.01	.31	1	-2.05	-2.70	1.73	1
CHL -MOS	4.41	4.43	-.06	1	.53	-.08	1.61	1
CHL –MBPO	-9.30	-9.01	-.77	1	-.57	.31	-2.35	1
LES –BOT1	N/A	2.74	--	--	N/A	2.74	--	--
LES –MBPO	N/A	2.90	--	--	N/A	2.90	--	--
MBPO-CHL	9.43	9.10	.86	1	.03	-.29	.84	1
MBPO-LES	-2.88	-2.90	.06	1	.00	1.06	-.71	1
MOS -CHL	N/A	-4.42	--	--	N/A	-4.42	--	--
MOS –RMT2	N/A	-1.23	--	--	N/A	-1.23	--	--
MTY3-BOT1	-2.88	-2.88	.00	1	-1.44	-1.15	-.77	1
MTY3-BOT1	-2.88	-3.29	.00	1	-1.44	-1.32	-.77	1
MTY3-RMT2	6.72	6.16	1.50	1	2.40	2.42	-.05	1
PALU-BOT1	-15.00	-14.77	-.60	1	-5.01	-3.91	-2.92	.04
PALU-BUL	6.93	6.96	-.07	1	2.86	3.54	-1.79	.07
PLY –BOT1	-11.34	-11.70	.96	1	-5.09	-4.46	-1.68	.11
PLY -BUL	-.72	-.88	.43	1	-.63	-3.64	8.02	.04
PLY -CHL	1.80	2.01	-.57	1	1.88	2.59	-1.91	1
RMT2-BOT1	-7.20	-7.36	.44	1	-2.88	-2.87	-.02	1
RMT2-MOS	1.44	1.23	.57	1	.96	.85	.31	1
RMT2-MTY3	-5.76	-6.13	.99	1	-1.92	-2.48	1.49	1

Table 5.4.6 – Botterens island power injections

	Active Power, MW				Reactive Power, MVar			
Meas. Location	Metered Value	Estimated Value	r_{SNi}	$Q_{i,i}$	Metered Value	Estimated Value	r_{SNi}	$Q_{i,i}$
BOT1	N/A	-7.36	--	--	N/A	-3.42	--	--
BOT	45.18	44.75	1.13	1	22.68	22.29	1.04	1
BRO	.00	.00	.02	1	.00	.00	.01	1
BUL	N/A	-6.07	--	--	N/A	.09	--	--
CHL	N/A	-6.59	--	--	N/A	-2.47	--	--
LES	N/A	5.63	--	--	N/A	-3.26	--	--
MBPO	N/A	6.20	--	--	N/A	.77	--	--
MOS	N/A	-5.65	--	--	N/A	-.93	--	--
MTY3	.00	-.01	.14	1	.00	-.05	.60	1
PALU	N/A	-7.82	--	--	N/A	-.38	--	--
PLY	N/A	-10.57	--	--	N/A	-5.51	--	--
RMT2	N/A	-12.27	--	--	N/A	-4.51	--	--

Table 5.4.7 – Botterens island voltages

Bus	$V_{estimated}$	$\theta_{estimated}$	$V_{metered}$	r_{Si}	$Q_{i,i}$
BOT	235.19	0.00	235.20	.00	1
BOT			236.10	.41	1
BOT			235.80	.28	1
BOT1	63.04	-3.43	62.70	-.54	1
BOT1			62.83	-.33	1
BOT1			62.77	-.43	1
BOT1			62.93	-.17	1
BOT1			62.83	-.33	1
BOT1			63.16	.20	1
BOT1			63.20	.25	1
BRO	63.04	-3.43	62.70	-.54	1
BUL	62.73	-3.80	63.03	.48	1
CHL	62.36	-3.91	62.17	-.29	1
CHL			62.30	-.08	1
CHL			62.21	-.24	1
LES	62.95	-3.20	--	--	--
MBPO	62.95	-3.23	62.72	-.37	1
MBPO			62.83	-.20	1
MBPO			62.79	-.25	1
MOS	62.23	-4.27	62.77	.86	1
MTY3	62.80	-3.69	63.62	1.32	1
MTY3			62.83	.05	1
MTY3			62.83	.05	1
MTY3			62.83	.05	1
PALU	62.83	-3.71	63.03	.32	1
PALU			63.03	.32	1
PLY	62.70	-3.80	62.80	.15	1
PLY			62.70	-.01	1
PLY			62.67	-.06	1
RMT2	62.33	-4.21	62.04	-.46	1
RMT2			62.04	-.46	1
RMT2			62.30	-.03	1
RMT2			62.30	-.03	1

Table 5.4.8 – Villarepos island power flows

	Active Power				Reactive Power			
Measurement Location	Metered Value	Estimated Value	r_{Si}	$\underline{Q}_{i,i}$	Metered Value	Estimated Value	r_{Si}	$\underline{Q}_{i,i}$
ESTA-PAYN	-1.44	-1.26	-.20	1	.96	2.26	-1.47	1
ESTA-SALA	-12.48	-11.75	-.83	1	-3.36	-6.74	3.82	1
PAYN-ESTA	1.44	1.26	.20	1	-1.20	-2.30	1.24	1
PAYN-RMT1	5.28	4.86	.48	1	2.64	2.55	.10	1
PAYN-VLR1	-19.20	-18.82	-.43	1	-6.24	-3.87	-2.68	1
RMT1-PAYN	-4.32	-4.82	.57	1	-2.40	-2.58	.21	1
SALA-ESTA	11.52	11.84	-.36	1	5.28	6.91	-1.84	1
SALA-VLR1	-17.76	-17.61	-.16	1	-6.72	-7.59	.99	1
VLR1-PAYN	19.92	19.18	.83	1	6.62	4.27	2.65	1
VLR1-SALA	17.73	17.68	.06	1	7.17	7.74	-.64	1
VLR1-VLR	N/A	-22.74	--	--	N/A	-22.74	--	--
VLR1-VLR	N/A	-20.72	--	--	N/A	-20.72	--	--
VLR -VLR1	20.82	22.74	.11	1	6.99	9.00	-1.37	1
VLR -VLR1	20.82	20.72	.11	1	6.99	8.20	-1.37	1

Table 5.4.9 – Villarepos island power injections

	Active Power (MW)				Reactive Power (MVar)			
Meas. Location	Metered Value	Estimated Value	r_{SNi}	$\underline{Q}_{i,i}$	Metered Value	Estimated Value	r_{SNi}	$\underline{Q}_{i,i}$
ESTA	N/A	-13.01	--	--	N/A	-4.48	--	--
PAYN	N/A	-12.70	--	--	N/A	-3.61	--	--
RMT1	N/A	-4.82	--	--	N/A	-2.58	--	--
SALA	N/A	-5.78	--	--	N/A	-.68	--	--
VLR1	N/A	-6.60	--	--	N/A	-3.18	--	--
VLR	N/A	43.46	--	--	N/A	17.21	--	--

Table 5.4.10 – Villarepos island voltages

Voltage Location	$V_{estimated}$	$\theta_{estimated}$	Measured Value	r_{Si}	$\underline{Q}_{i,i}$
ESTA	60.83	-3.57	60.98	.99	1
ESTA			60.72	-.67	1
PAYN	60.77	-3.48	60.72	-.30	1
PAYN			60.72	-.30	1
PAYN			60.72	-.30	1
RMT1	60.23	-3.70	59.93	-1.88	1
RMT1			60.46	1.44	1
SALA	61.80	-2.76	62.04	1.50	1
SALA			62.04	1.50	1
VLR	126.33	0.00	N/A	--	--
VLR1	62.21	-2.34	62.01	-1.31	1
VLR1			62.07	-.89	1
VLR1			62.24	.15	1
VLR1			62.24	.15	1
VLR1			62.21	-.06	1

5.5. SHGM and the Identification of Multiple Bad Data

On the same network, the SHGM-estimator demonstrates its ability to cleanly reject outlying measurements without the smearing effect and to reject multiple bad data, even interacting bad data. Six active/reactive pairs of P and Q bad data on power flow measurements were introduced. Three of these are at leverage points, the flow measurements on line HV31-HV11, HV32-HV12, and OELB-MAIG. Two other pairs are conforming bad data located on line COUR-SCHI, whereas the last one is also a pair of bad power flows at CRES-KERZ. The solution required 3 iterations for convergence, and the results are summarized in Table 5.5.2. Clearly all 12 bad data have been rejected with large weighted residuals. Table 5.5.1 provides the true, bad and estimated power measurements for the bad data points in MW and MVar.

The SHGM-estimator shows its ability to detect bad data and limit their influence even when they occur at leverage points, as evidenced by the bad data at HV31-HV11 and HV32-HV12. The short lines at Hauterive Stations 1 and 3 are only 100 meters in length. However, we should recognize that the estimator is unable to provide a proper estimate for these particular flow values because their branch impedances are so small. The state variables being estimated are the bus voltage magnitudes and phase angles, and any small error in their estimation results in erroneous flow values on these very short lines, whereas the surrounding flow values show that the voltages have in fact been correctly estimated. This can be explained by resorting to the dc model, which stems from a linearization of the state estimation model about the flat voltage profile ($1\angle 0^\circ$ pu). Using this model, the active and reactive power flow P_{ij} and Q_{ij} through a transmission line with a reactance X_{ij} is given by $\Delta P_{ij} = (1/X_{ij})\Delta\theta_{ij}$ and $\Delta Q_{ij} = (1/X_{ij})\Delta V_{ij}$. Hence, if $1/X_{ij}$ is very large, a small change in θ_{ij} results in a large change in the power flow P_{ij}. Similarly, a small change in V_{ij} yields a large change in Q_{ij}.

Table 5.5.1 – SHGM-estimator results for the bad data

Measurement	P_{true} (MW)	P_{bad} (MW)	$P_{estimated}$ (MW)	Q_{true} (MVar)	Q_{bad} (MVar)	$Q_{estimated}$ (MVar)
COUR – SCHI	-9.765	-36.000	-11.640	-4.766	-40.000	-6.517
CRES – KERZ	2.900	15.000	6.220	1.000	31.000	1.422
HV31 – HV11	-4.813	-150.000	-65.112	-19.773	-160.000	-71.003
HV32 – HV12	-1.640	-150.000	-69.672	27.031	180.000	13.431
OELB – MAIG	2.400	55.000	2.299	2.896	58.000	3.066
SCHI – COUR	9.800	36.000	11.692	4.833	41.000	6.602

We have three options for how to treat these short lines: merge the buses on either side of the lines' ends, treat the lines as zero-impedance branches, or keep them as leverage points. In the installed estimator on the EEF system, buses HV31 and HV11 are merged into a single bus in the topology processor, as are buses HV32 and HV12. However, the possibilities have not been fully explored, and a determination of the best method remains a future work. Here no bad data have been introduced in the Monteynan substation.

Table 5.5.2 – Weighted residuals provided by the SHGM-estimator.

Measurement	Active Power		Reactive Power	
	error/σi	r_{wi}	error/σ_i	r_{wi}
CORB - HV32	0.372	0.087	1.490	-0.780
CORB - VLR2	0.654	-0.617	0.237	-0.142
CORB – COUR	0.037	-0.304	1.612	0.661
COUR - CORB	0.044	0.297	0.071	1.054
COUR - VLR2	0.474	0.251	0.691	0.203
COUR – SCHI	26.235	-24.349	35.234	-33.477
CRES – KERZ	17.100	13.725	30.000	29.578
CRES – MORT	1.717	0.411	1.156	-0.900
CRES – SCHI	1.371	1.713	1.542	-0.201
CRES - VLR2	0.764	-2.421	0.181	-0.843
GIFF – MARL	1.724	-0.888	1.357	0.499
GIFF – SCHI	1.748	1.889	1.022	0.620
GUTZ – STL2	1.233	-0.051	0.459	0.950
GUTZ - HV31	1.710	1.210	0.323	-0.094
GUTZ – MONC	1.019	-1.039	0.243	0.103
HV11 – MAIG	0.562	0.003	0.158	-0.979
HV12 - STL1	0.343	-1.023	1.290	0.975
HV12 – MARL	0.093	-0.305	0.642	1.196
HV32 – CORB	0.708	0.247	0.361	-1.062
HV31 – MONC	0.913	-0.190	0.165	0.728
HV31 – GUTZ	0.645	1.147	0.488	-0.044
HV31 – MTY	1.550	-0.733	0.709	0.717
HV31 - HV11	145.187	-85.344	140.227	-66.354
HV32 – SCHI	0.328	1.022	0.615	-0.456
HV32 - HV12	148.360	-80.033	152.969	167.204
HV32 – MTY	1.772	0.884	0.822	-0.750
KERZ – CRES	0.157	3.196	0.522	-0.097
KERZ - VLR2	1.242	0.180	0.481	0.878
MAIG - HV11	0.007	0.550	0.535	0.303
MAIG – OELB	0.903	0.802	0.910	0.386
MARL - HV12	0.670	-0.457	1.693	1.146
MARL - GIFF	0.582	-0.257	0.228	0.671
MONC - GUTZ	0.903	-0.883	0.274	0.421
MONC - HV31	0.109	-0.611	0.419	-0.112
MORT – CRES	0.110	1.190	0.430	0.182
MORT - VLR2	1.055	-1.585	0.988	-1.251
MTY - HV32	0.007	0.897	2.094	-0.407
MTY - HV31	0.040	-0.774	0.068	-0.107
OELB - MAIG	52.600	52.702	55.104	55.631
OELB - STL2	1.522	0.336	1.040	-0.381
SCHI – COUR	26.200	24.296	36.167	34.392
SCHI - HV32	0.499	-0.199	1.228	1.111
SCHI – CRES	0.788	-1.134	0.603	-1.938
SCHI – STL1	0.224	-0.020	0.690	-0.984
SCHI – GIFF	0.489	0.349	0.409	0.074
STL1 – SCHI	1.036	-1.240	0.155	0.162
STL1 - HV12	0.165	0.514	0.144	0.489
STL2 - OELB	1.162	2.347	0.896	0.534
STL2 - GUTZ	0.767	-0.418	1.060	-0.331
VLR2 - MORT	0.243	0.287	1.233	1.515
VLR2 - CRES	0.213	1.441	1.222	1.900
VLR2 - COUR	0.071	-0.655	0.368	-0.505
VLR2 - CORB	0.880	-0.917	1.498	1.907
VLR2 - KERZ	1.745	0.320	0.494	0.164
MTY	1.128	1.204	1.219	2.868
STL2	0.003	-0.013	0.040	-0.008
VLR2	0.020	-0.038	0.171	-0.064

5.6. Meter Placement

Metering studies conducted on the EEF system revealed some weaknesses in the measurement configuration. Certainly the nearly complete lack of power injection measurements hampers the local redundancy throughout the system. They are unavailable because the power flows across transformers connecting the 62.5 kV and 18 kV networks are not measured. Since these transformer flows are to be modeled as 62.5 kV bus injections, the sum of power injections is made impossible at most of these buses. It was well known to the EEF engineers that the local measurement redundancy is poor or nonexistent for many of the bus voltage magnitudes of the southern network (Botterens island), as can be seen from the measurement configuration of Figure 5.1.2. The tap position of the 125/62.5 kV transformers in the Villarepos substation were not available throughout most of the testing. When they became available, estimation of the bus voltage magnitudes of Villarepos island improved dramatically.

The most significant contribution of the metering studies was the determination of the need for added redundancy in the area of the Monteynan and Hauterive substations. The poor local redundancy, combined with multiple zero-impedance branches and the presence of zero injection buses, make bad data detection in this area nearly impossible under the normal operating topology. Figure 5.1.1 provides only an approximate topological model of the normal operating topology in this area. Figure 5.6.1 shows some of the system details seen by the topology processor. Note that there are buses directly connecting Hauterive Station 1 and Hauterive Station 3 instead of the lines that exist in the simulations of the last section. This shows the substations as they are currently modeled, with the lines neglected and the buses merged.

In Figure 5.6.1 we see additional buses in the Monteynan substation at the 62.5 kV level. These are zero injection buses, and the lines connecting Hauterive to Monteynan are only 300 meters in length. Hence, when the circuit breaker configuration creates the topology of Figure 5.1.1, these zero impedance lines should not be modeled because they decrease the local redundancy. Unfortunately, this is not an option because from a practical

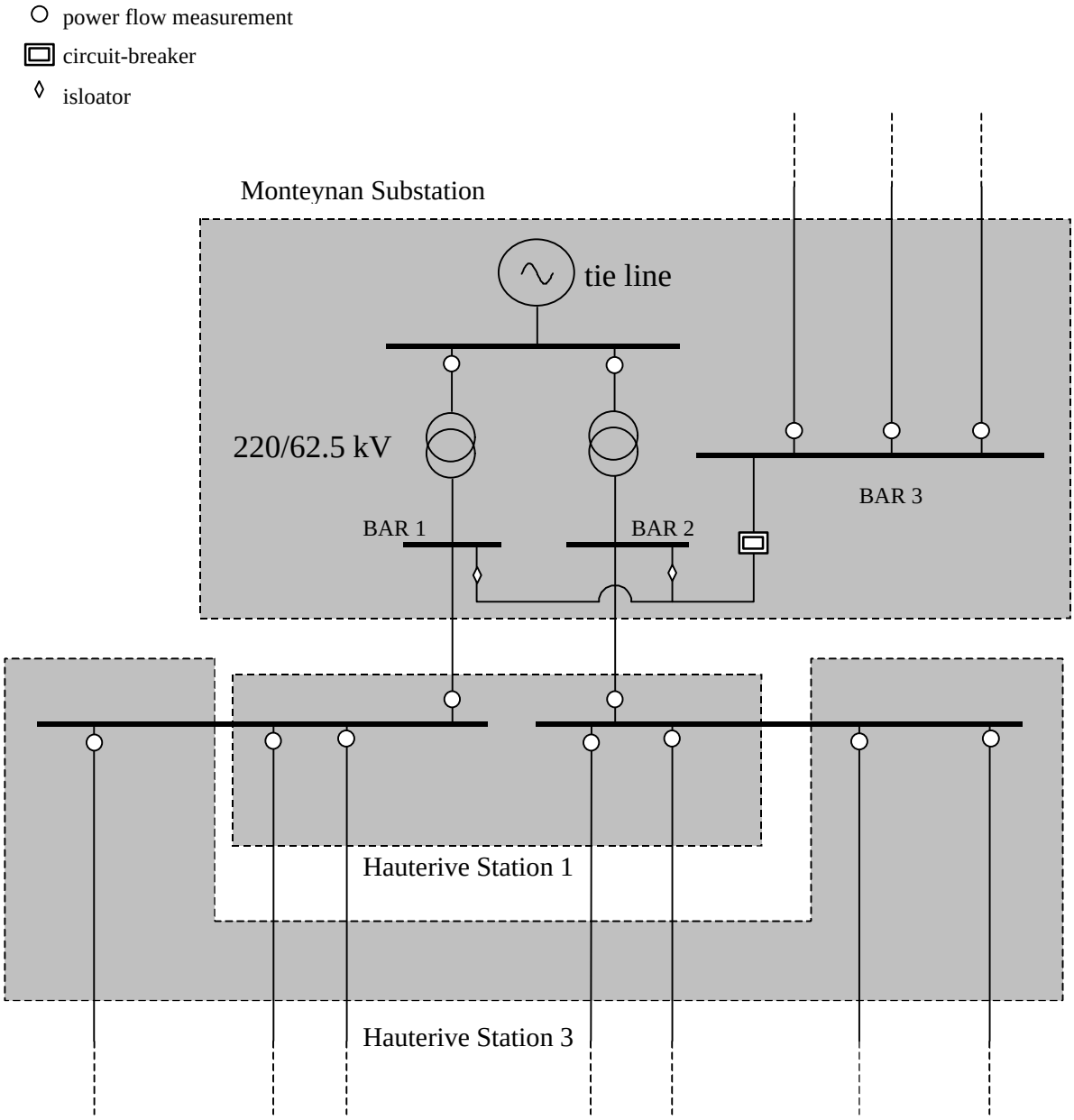

Figure 5.6.1 – Topological details of interest in the Monteynan and Hauterive substations

standpoint they may not be excluded from the topological model. The EEF operating practice commonly calls for power to be transferred to Botterens island from the Monteynan tie-line through busbar 3. It may be therefore be connected to either busbar 1 or busbar 2 without being connected to the Hauterive Stations, making it impossible to remove these lines from the system topological model.

Two options remain available for the treatment of these lines: allow the state estimator to identify them as leverage points, or model them as zero-impedance lines and use the Hauterive-Monteynan power flows as state variables. Simulations showed early in the testing process that the SHGM-estimator is unable to correctly identify bad data in this region. Further simulations confirmed that this is due to the area's poor local redundancy. The local redundancy of a measurement is formally defined as the minimum number of other measurements that when deleted makes that measurement a critical measurement. [2] A critical measurement is any measurement that when deleted, makes the Jacobian matrix rank deficient. In other words, a critical measurement is the only measurement available to observe a particular state variable, and has a local redundancy of zero. Critical measurements and low local redundancy are of particular interest in state estimation because the number of measurements observing a state variable is small, causing the state estimator to easily break down in that area.

It is clear from Figure 5.6.1 that the measurements observing busbar 1, busbar 2, and the tie-line bus in Monteynan substation have very low local redundancy, even though the zero injections at busbars 1 and 2 are treated as measurements. Since we may not increase the local redundancy by eliminating the Monteynan-Hauterive lines and the zero injection buses, active and reactive measurement pairs at the transformer secondaries and at the tie-line injection have been recommended to the EEF to improve the estimator's performance in this area. Fortunately, voltage and current measurements already exist at the transformer secondaries, so the only equipment lacking for the installation of power measurements are the power converters. Measurement of the tie-line injection might also be made available to the SCADA database. A formal report is being prepared for the EEF engineering department in order to motivate the placement of measurements in this area.

6 – CONCLUSION

The SHGM-estimator has shown that it outperforms the conventional WLS-based method in that it exhibits excellent convergence properties while rejecting multiple bad data. It is not prone to the smearing and masking effects so characteristic of the WLS methods, making multiple bad data identification rather straightforward. The SHGM possesses these capabilities because the influence of residual as well as the influence of position are bounded through the Huber ψ-function together with the projection statistic. It has been upgraded and installed in the real-time EEF system in conjunction with a topology processor, where it has maintained its performance expectations.

Some issues remain as future work for the EEF state estimator package. The standard deviations of the metering devices have only been tentatively assigned. The standard deviation given to a measurement is not an absolute measure of the spread of its error distribution, but rather a measure of the confidence given to it relative to the other measuring devices in the system. Additional efforts should therefore be made to more accurately determine the measurements' biases so that those devices with high bias can be assigned larger standard deviations. A study of the measurements' recursive residual statistics should also be undertaken as an effort to determine the standard deviations, as well as a study of the statistical samples for analysis of the estimators used on the samples.

The short lines connecting Hauterive Stations 1 and 3 are currently eliminated and the busbars at the lines' ends merged. This solution is the simplest but perhaps not the best. Further analysis is required to properly determine the best way to model the problem. Leaving the lines in the topological model and allowing the SHGM-estimator to treat them as leverage points allows us to use the single flow measurement on each line to reinforce the state estimate when the measurements are good. Treating the lines as zero-impedance branches and adding the power flows as state variables may create critical measurements. Reinforcement of the local redundancy near the Monteynan tie-line will greatly improve estimation, allowing the SHGM-estimator to properly identify bad data in that area.

7 – REFERENCES

[1] Abur, F.H. Magnago, F.L. Alvarado, “Elimination of Leverage Measurements Via Matrix Stretching”, *International Journal of Electrical Power & Energy Systems*, Vol.19, No.8, Nov. 1997, pp. 557-562.

[2] P. Albertos, C. Alvarez, “An On-Line Algorithm for Power System Observability Determination”, *IFAC Symp. on Software for Computer Control*, Madrid, October, 1982, pp. 433-439.

[3] D.F. Andrews, P.J. Bickel, F.R. Hampel, P.J. Huber, W.H. Rogers, J.W. Tukey, *Robust Estimates of Location: Survey and Advances*, Princeton University Press, 1972.

[4] P. Billingsley, *Probability and Measure*, John Wiley, 1986.

[5] M. Chapman, L. Mili, C. Tinguely, R. Cherkaoui, "Applications of a Robust State Estimator to the EEF Power System", *LESCOPE 98. 1998 Large Engineering Systems Conference on Power Engineering. Conference Proceedings. Theme: Powering up to 2000*, Large Engineering Systems, Halifax, NS, Canada; 1998; ix+324, pp. 20-24.

[6] K.A. Clements, G.R. Krumpholz, P.W. Davis, “State Estimator Measurement System Reliability Evaluation – An Efficient Algorithm Based on Topological Observability Theory”, *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-101, No. 4, pp. 997-1004, April 1982.

[7] K.A. Clements, G.R. Krumpholz, P.W. Davis, “Power System State Estimation With Measurement Deficiency: An Observability/Measurement Placement Algorithm”, *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-102, No. 7, pp. 2012-2020, July 1983.

[8] C. Croux, P.J. Rousseeuw, “Time-efficient Algorithms for Two Highly Robust Estimators of Scale”, *Computational Statistics*, Vol. 1, eds. Y. Dodge and J. Whittaker, Heidelberg: Physica-Verlag, 1992, pp. 411-428.

[9] D.L. Donoho, “Breakdown Properties of Multivariate Location Estimators”, qualifying paper, Harvard University, Boston, MA, 1982.

[10] T.E. Dy Liacco, “An Overview of Power System Control Centers”, *Energy Control Center Design, IEEE Tutorial Course*, TU0010-9PWR, 1977.

[11] T.E. Dy Liacco, “System Security: The Computer’s Role”, *IEEE Spectrum*, June 1978, pp. 43-50.

[12] T.E. Dy Liacco, "The Role and Implementation of State Estimation in an Energy Management System", *Electrical Power & Energy Systems*, Vol. 12, No. 2, April 1990, pp. 75-79.

[13] M. Gasko and D. Donoho, “Influential Observation in Data Analysis”, American Statistical Association, Proceedings of the Business and Economic Statistics Section, 1982, pp. 104-110.

[14] F.R. Hampel, *Contributions to the Theory of Robust Estimation*, Ph.D. Thesis, University of California, Berkeley, 1968.

[15] F.R. Hampel, "A General Qualitative Definition of Robustness", *Annals of Mathematical Statistics*, Vol. 42, 1971, pp. 1887-1896.

[16] F.R. Hampel, "The Influence Curve and its Role in Robust Estimation", *Journal of the American Statistical Association*, Vol. 69, 1974, pp. 383-393.

[17] F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw, W.A. Stahel, *Robust Statistics: The Approach Based on Influence Functions*, John Wiley, 1986.

[18] E. Handschin, F.C. Schweppe, J. Kohlas, A. Fiechter, “Bad Data Analysis for Power System State Estimation”, *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-94, No. 2, Mar./Apr. 1975.

[19] P.J. Huber, “Robust Estimation of a Location Parameter”, *Annals of Mathematical Statistics*, Vol. 35, 1964, pp. 73-101.

[20] P.J. Huber, *Robust Statistics*, John Wiley, 1981.

[21] H.M. Merrill, F.C. Schweppe, “Bad Data Suppression in Power System Static State Estimation”, *IEEE on Power Apparatus and Systems*, Vol. PAS-90, No. 6, Nov./Dec. 1971, pp. 2718-2725.

[22] L. Mili, T. Van Cutsem, M. Ribbens-Pavella, “Bad Data Identification Methods in Power System State Estimation – A Comparative Study”, *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-104, No.11, Nov. 1985.

[23] L. Mili, V. Phaniraj, P.J. Rousseeuw, "Robust Estimation Theory for Bad Data Diagnostics in Electric Power Systems", *Advances in Control and Dynamic Systems*, Vol. 37 (ed. C.T. Leondes) Academic Press, New York, 1990, pp. 271-325.

[24] L. Mili, M.G. Cheniae, P.J. Rousseeuw, "Robust state estimation of electric power systems", *IEEE Transactions on Circuits and Systems Part I, Fundamental Theory and Applications*, Vol. 41, May 1994, pp. 349-358.

[25] L. Mili, M. Cheniae, N. S. Vichare, P. J. Rousseeuw, "Robust State Estimation Based on Projection Statistics", *IEEE Transactions on Power Systems*, Vol. 11, No. 2, May 1996.

[26] S.P. Neugebauer, L. Mili, "Local Robustness of Nonlinear Regression M-estimators", *Proceedings of the 1997 IEEE/EURASIP Workshop on Nonlinear Signal and Image Processing*, Mackinae Island, Michigan, Sept. 8-10, 1997.

[27] V. Phaniraj, *Robust State Estimation in Power Systems*, Ph.D. Thesis, Virginia Tech, 1991.

[28] W. H. Press, B. P. Flannery, S. A. Teukolksy, W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, 1989.

[29] M. Ribbens-Pavella, T. Van Cutsem, P. Rousseaux, "On-Line Stability and Dynamic Security Assessment of Electric Power Systems", *Proc. Of the IFAC Symposium on Planning and Operation of Electric Energy Systems*, Rio de Janeiro, Brazil, July 22-25, 1985.

[30] P. Rousseaux, Th. Van Cutsem, T.E. Dy Liacco, "Whither Dynamic State Estimation?", *Electrical Power & Energy Systems*, Vol. 12, No. 2, April 1990, pp. 104-116.

[31] P.J. Rousseeuw, B.C. Van Zomeren, "Robust Distances: Simulations and Cutoff Values", *Directions in Robust Statistics and Diagnostics*, Part II, edited by W. Stahel and S. Weisberg, Springer-Verlag, 1992, pp. 411-428.

[32] P.J. Rousseeuw, C. Croux, "Alternatives to the Median Absolute Deviation", *Journal of the American Statistical Association*, Vol. 88, No. 424, Dec. 1993, pp. 1273-1283.

[33] A. M. Sasson, S. T. Ehrmann and P. Lynch, and L. S. Van Slyck, "Automatic Power System Network Topology Determination," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-92, March/April 1973, pp. 610-618.

[34] F.C. Schweppe, J. Wildes, D.B. Rom, “Power System Static-State Estimation, Parts I, II and III”, *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-89, No. 1, January 1970, pp. 120-135.

[35] F.C. Schweppe, E.J. Handschin, “Static State Estimation in Electric Power Systems”, *Proceedings of the IEEE*, Vol. 62, No. 7, July 1974.

[36] D. Shirmohammadi, B. Wollenberg, A. Vojdani, et. al. “Transmission Dispatch and Congestion Management in the Emerging Energy Market Structures”, *IEEE Transactions on Power Systems*, Vol. 13, No. 4, Nov. 1998, pp. 1466-1474.

[37] G.W. Stagg, A.H. El-Abiad, *Computer Methods in Power System Analysis*, McGraw-Hill, 1968.

[38] W.A. Stahel, “Robuste Schätzungen: Infinitesimale Optimalität und Schätzungen von Kovarianzmatrizen”, PhD Thesis, E.T.H Zürich, Switzerland, 1981.

[39] L. S. VanSlyck, J. J. Allemong, “Operating Experience with the AEP State Estimator,” *IEEE Transactions on Power Systems*, Vol. 3, No. 2, May 1988.

[40] N. Xiang, S. Wang, E. Yu, “A New Approach for Detection and Identification of Multiple Bad Data in Power System State Estimation”, *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-101, No. 2, Feb. 1982, pp. 454-462.

8 – APPENDIX

8.1. Power system modeling

Modern power systems are ideally operated using balanced three-phase sinusoidal signals. Three-phase means that the power generated by a machine, transmitted through a network or consumed by a load is realized in three physically separate though electrically coupled wirings, which act together to provide an effective power. Balanced means that the three phases' sinusoidal signals have the same frequency and magnitude, and are each separated in phase by 120° of the sinusoids' period. The balance of the three phases is a valid approximation when phase permutation is used and the power system is in a normal state of operation.

When the power system is in steady-state operation, the frequency of the signal is constant and we can adapt what is termed phasor notation. As an example, a current which is normally represented as $I(t,\theta) = I\cos(\omega_s t + \theta)$ may be represented as the complex number $\bar{I} = Ie^{j\theta}$. Here we have dropped the angle term $\omega_s t$ common between all three phases, and described the current signal in terms of its magnitude and phase angle. In addition, the balance allows steady-state modeling of the system to be represented using a single-phase equivalent model of the three phases, and to be drawn concisely using the so-called one-line diagram. Under this notation, the three phases of a machine, transmission line, transformer, load, etc. are represented by the total power, a single voltage, a single current, and a single impedance, all in complex notation.

Simplified power system network models are comprised primarily of nodes, also called buses, interconnected by branches. At each bus, there is a net power generation and/or consumption, referred to as the bus' power injection. Branches may be either transmission lines or transformers. The power flowing across a branch can be expressed in terms of the voltage and impedance characteristics of the branch. First let us consider

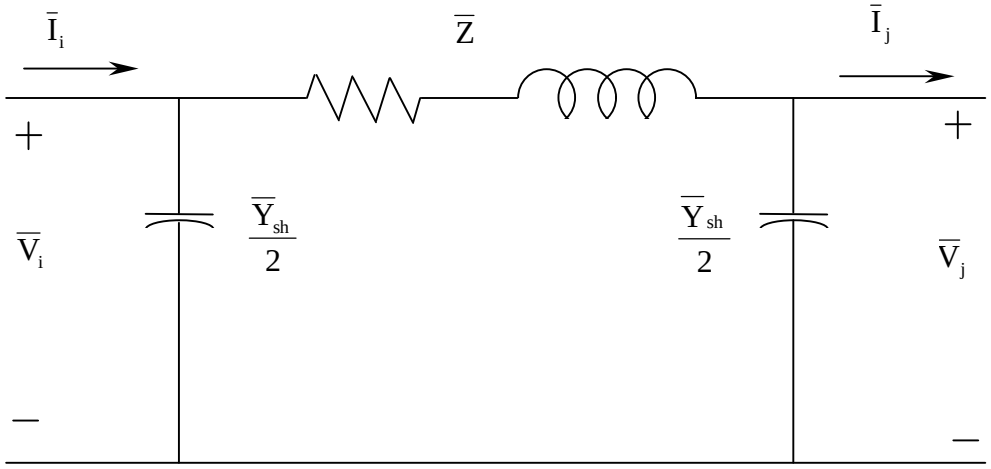

Figure 8.1.1 – Nominal π-circuit model of a transmission line

transmission lines. Small to medium length transmission lines, up to about 240 km, may be sufficiently represented by resistance and inductance as lumped parameters, with the shunt capacitance of the line divided into two equal parts placed at the sending and receiving ends of the line, running from the bus to ground. This model, depicted in Figure 8.1.1, is the well-known nominal π circuit model.

From the π-circuit model, we can see that the power leaving bus i toward bus j is the terminal power, $\bar{S}_{ij} = P_{ij} + jQ_{ij} = \bar{V}_i \bar{I}_i^*$, where * is the complex conjugate operator. Given the parameters $\bar{Z}$ and $\bar{Y}_{sh}$, we may calculate P_{ij} and Q_{ij} in the following manner:

$$\bar{Y}_{ij} = G_{ij} + jB_{ij} = \frac{1}{R_{ij} + X_{ij}} = \frac{1}{\bar{Z}_{ij}}, \quad \bar{Y}_{shij} = jB_{shij}, \quad \theta_{ij} = \theta_i - \theta_j$$

$$\begin{aligned}
\bar{S}_{ij} &= \bar{V}_i \bar{I}_i^* = \bar{V}_i \left[\bar{V}_i \left(\tfrac{1}{2} \bar{Y}_{shij} \right) + \left(\bar{V}_i - \bar{V}_j \right) \left(G_{ij} + jB_{ij} \right) \right]^* \\
&= \bar{V}_i \left[\bar{V}_i^* \left(-j \tfrac{1}{2} B_{shij} \right) + \left(\bar{V}_i^* - \bar{V}_j^* \right) \left(G_{ij} - jB_{ij} \right) \right] \\
&= -j \tfrac{1}{2} V_i^2 B_{shij} + \left(V_i^2 - V_i V_j e^{j\theta_{ij}} \right) \left(G_{ij} - jB_{ij} \right) \\
&= -j \tfrac{1}{2} V_i^2 B_{shij} + \left[V_i^2 - V_i V_j \cos(\theta_{ij}) - jV_i V_j \sin(\theta_{ij}) \right] \left(G_{ij} - jB_{ij} \right)
\end{aligned}$$

Let $\alpha = V_i^2 - V_i V_j \cos\theta_{ij}$, and $\beta = -V_i V_j \sin\theta_{ij}$, then

$$\bar{S}_{ij} = -j\tfrac{1}{2}V_i^2 B_{shij} + \alpha G_{ij} - j\alpha B_{ij} + j\beta G_{ij} + \beta B_{ij}$$

$$\Rightarrow \begin{cases} P_{ij} = V_i^2 G_{ij} - V_i V_j \left(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}\right) \\ Q_{ij} = -V_i^2\left(B_{ij} + \tfrac{1}{2}B_{shij}\right) - V_i V_j \left(G_{ij}\sin\theta_{ij} - B_{ij}\cos\theta_{ij}\right) \end{cases} \tag{8.1}$$

The total power injection at a given bus is the sum of the power flows from the lines adjacent to the bus. Hence, if we define the total complex power injection at bus *i* as $\bar{S}_i$, then we get

$$\bar{S}_i = \sum_{j=\ell}^{k} \bar{S}_{ij} = \sum_{j=\ell}^{k} \left(P_{ij} + jQ_{ij}\right) = P_i + jQ_i ,$$

such that

$$\begin{aligned} P_i &= \sum_{j=\ell}^{k} P_{ij} = \sum_{j=\ell}^{k} \left[V_i^2 G_{ij} - V_i V_j \left(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}\right)\right] \\ Q_i &= \sum_{j=\ell}^{k} Q_{ij} = -\sum_{j=\ell}^{k} \left[V_i^2\left(B_{ij} + \tfrac{1}{2}B_{shij}\right) + V_i V_j \left(G_{ij}\sin\theta_{ij} - B_{ij}\cos\theta_{ij}\right)\right] \end{aligned} \tag{8.2}$$

where $\ell,\ldots,k$ are the bus numbers at the receiving ends of the lines adjacent to bus *i*.

The real and/or reactive power flow may be independently measured on either or both ends of a transmission line or transformer. Similarly, the real and/or reactive net power injection may be independently measured at a given bus. Hence, the measurement configuration may vary widely. The real power may be measured without the reactive power, although it is rare.

8.2. Regression model in power systems

The state of a power network may be defined as the flow of complex power through the network's branches and the net injection of complex power at each bus. We see then from (8.1) and (8.2) that the state of a network at a given moment may be completely described by the set of all complex bus voltages,

$$\left\{V_1 e^{j\theta_1}, \cdots, V_N e^{j\theta_N}\right\}, \tag{8.3}$$

assuming the connections between nodes and the parameters of the branches are completely known. For a system with N buses, there are therefore 2N-1 unknown *state variables* to be determined in order to describe the state of the system, because we choose one bus' phase angle as a reference. As mentioned in Chapter 1, power systems typically have far more measurements, say *m* measurements, available than state variables, which allow us to employ statistical methods to estimate the state variables of the power system.

A measurement vector $\underline{z}$ may be created which contains *m* measurements from the power system. Measurements include real and reactive power line flows and bus injections, voltage magnitudes at buses, tap ratios for transformers, and more recently, phase angle measurements. Those used in this thesis are power flows and injections and voltage magnitudes. The 2N-1 state variables constitute the state vector $\underline{x}$, which may be related to the measurements $\underline{z}$ in a regression model:

$$\underline{z} = \underline{h}(\underline{x}) + \underline{e} \tag{8.4}$$

The vector $\underline{e}$ is the error vector, which contains *m* variables: $e_1,\ldots, e_m$. Its purpose is to account for the uncertainty in the measurements and the model. Hence, $e_1,\ldots, e_m$ are random variables assumed to have a zero mean and a known diagonal covariance matrix, $\underline{R} = \text{diag}(\sigma_1^2,\ldots,\sigma_m^2)$.

Let us step away from the power system problem and consider a general regression model. Regression implies that the data in $\underline{z}$ have a pattern, which is represented by the vector $\underline{h}(\underline{x})$, formulated as follows. Suppose that we have *m* observations that are related to two state variables, x_1 and x_2, through a linear regression model given by

$$\begin{aligned} z_1 &= h_{11}x_1 + x_2 + e_1 \\ z_2 &= h_{12}x_1 + x_2 + e_2 \\ &\vdots \\ z_m &= h_{1m}x_1 + x_2 + e_m \end{aligned} \tag{8.5}$$

In the case where we have *n* state variables,

$$z_i = h_{1i}x_1 + h_{2i}x_2 + \cdots + h_{n-1,i}x_{n-1} + x_n + e_i \quad , \; i = 1,\ldots,m \tag{8.6}$$

The n-1 variables h_{1i}, h_{2i},…,$h_{n-1,i}$ are termed the independent, or explanatory, variables, while z_i is termed the response, or dependent, variable. A data point is ($h_{i,1}$,$h_{i,2}$, … ,$h_{i,n-1}$, z_i), and lies in an n-dimensional space, which is made up of two subspaces: the response space and the factor space, also called the design space. The response space is one-dimensional and contains the response variable. The factor space is made up of the n-1 h-dimensions. The slopes of the hyperplane in $\Re^n$ are x_1,…,x_{n-1}, while the intercept is x_n. Hence, the task is to find $\underline{x}$ such that the hyperplanes best relate, or fit, the observations to the independent variables. In matrix form, we have

$$\underline{z} = \underline{H}\underline{x} + \underline{e}\text{, where } \underline{H} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1,n-1} & 1 \\ h_{21} & h_{22} & \cdots & h_{2,n-1} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ h_{m1} & h_{m2} & \cdots & h_{m,n-1} & 1 \end{bmatrix}. \tag{8.7}$$

$\underline{H}$ is called the design matrix, and $\underline{x}$ the parameter vector.

In the power system problem, implying that the regression is constrained to pass through the origin, the last column of $\underline{H}$ is unnecessary. The task of state estimation is thus to use statistical methods to find $\underline{x}$ such that the measurements given by $\underline{z}$ best fit the known system model provided by the design matrix.

8.3. Weighted Least Squares

Given the regression model of (8.4), we might assume that $\underline{e} \sim N(\underline{0},\underline{R})$, where $\underline{R}$ is the variance-covariance matrix, or simply covariance matrix, of $\underline{e}$, which is given by $\underline{R}$ = diag(σ_1^2,…,σ_m^2). A more formal definition may be provided using the expected value:

$$\underline{R} = E\left[(\underline{e} - E[\underline{e}])(\underline{e} - E[\underline{e}])^T\right] = E\left[\underline{e}\underline{e}^T\right], \text{ where } E[\underline{e}] = 0\,. \tag{8.8}$$

As in Chapters 2 and 3, we have the objective function

$$J(\underline{x}) = \frac{1}{2}\sum_{i=1}^{m}\left(\frac{r_i}{\sigma_i}\right)^2 = \frac{1}{2}\underline{r}^T\underline{R}^{-1}\underline{r} = (\underline{z} - \underline{H}\underline{x})^T\underline{R}^{-1}(\underline{z} - \underline{H}\underline{x}),$$

where $1/\sigma_i^2$ acts as a weight. Minimizing $J(\underline{x})$ yields

$$\frac{\partial J(\underline{x})}{\partial \underline{x}} = \underline{0} = -\underline{H}^T\underline{R}^{-1}(\underline{z} - \underline{H}\underline{x})\,. \tag{8.9}$$

We would like to find the estimate $\underline{\hat{x}}$ such that, by the pseudo-inverse,

$$\underline{\hat{x}} = \left(\underline{H}^T \underline{R}^{-1} \underline{H}\right)^{-1} \underline{H}^T \underline{R}^{-1} \underline{z} . \tag{8.10}$$

This, along with (8.4), leads us to

$$\underline{\hat{z}} = \underline{H}\underline{\hat{x}} = \underline{H}\left(\underline{H}^T \underline{R}^{-1} \underline{H}\right)^{-1} \underline{H}^T \underline{R}^{-1} \underline{z} = \underline{S}\underline{z} , \tag{8.11}$$

where $\underline{S}_{mxm}$ is called the hat matrix. It allows us to form the very important residual sensitivity matrix $\underline{W}$ that relates the residuals to the errors:

$$\underline{r} = \underline{z} - \underline{\hat{z}} = \underline{z} - \underline{S}\underline{z} = \left(\underline{I} - \underline{S}\right)\underline{z} = \underline{W}\underline{z} , \quad \text{yielding} \quad \underline{r} = \underline{W}\underline{e} . \tag{8.12}$$

It is important to note that the matrix $\underline{W}$ is singular, hence there are multiple *m*-vectors $\underline{b}$ which can satisfy $\underline{r} = \underline{W}\underline{b}$. We see that the errors are considered as linear combinations of the residuals. This becomes important both in WLS residual analysis and in explaining certain properties of the WLS in the presence of gross errors.

Most notable of these properties are the spreading effect and the masking effect. The spreading effect stems from the fact that all residuals are a function of each error. One large error can therefore result in several large residuals, effectively spreading the error over several observations such that it is difficult or impossible to properly identify the faulty observation. The masking effect occurs when linear combinations of large errors result in small residuals, masking the large errors' presence.

8.4. Formation of the IRLS Jacobian

The Jacobian matrix, used in the first-order Taylor-series linearization of the regression model, is defined as $\underline{H}\left(\underline{x}\right) = \partial \underline{h}\left(\underline{x}\right) / \partial \underline{x}$. Recall that in power systems the regression model is without intercept. The state vector $\underline{x}$ of a system with N buses is

$$\underline{x} = \left[\theta_2 \cdots \theta_N \; V_1 \cdots V_N\right]^T , \tag{8.13}$$

and the measurement vector typically contains real and reactive power flows ($\underline{P}_{flow}$ and $\underline{Q}_{flow}$), real and reactive power injections ($\underline{P}_{inj}$ and $\underline{Q}_{inj}$), and voltage magnitude measurements ($\underline{V}_{meas}$). With these, the Jacobian may be expressed with the state vector as

$$\underline{H}\,\underline{x} = \begin{bmatrix} \frac{\partial \underline{P}_{flow}}{\partial \underline{\theta}} & \frac{\partial \underline{P}_{flow}}{\partial \underline{V}} \\ \frac{\partial \underline{P}_{inj}}{\partial \underline{\theta}} & \frac{\partial \underline{P}_{inj}}{\partial \underline{V}} \\ \frac{\partial \underline{Q}_{flow}}{\partial \underline{\theta}} & \frac{\partial \underline{Q}_{flow}}{\partial \underline{V}} \\ \frac{\partial \underline{Q}_{inj}}{\partial \underline{\theta}} & \frac{\partial \underline{Q}_{inj}}{\partial \underline{V}} \\ \frac{\partial \underline{V}_{meas}}{\partial \underline{\theta}} & \frac{\partial \underline{V}_{meas}}{\partial \underline{V}} \end{bmatrix}_{mx(2N-1)} \begin{bmatrix} \underline{\theta} \\ \underline{V} \end{bmatrix}_{(2N-1)x1}, \tag{8.14}$$

where $\underline{P}_{flow}$, $\underline{P}_{inj}$, $\underline{Q}_{flow}$, $\underline{Q}_{inj}$, and $\underline{V}_{meas}$ are the subvectors making up $\underline{z}$, such that

$$\underline{z} = \left[\underline{P}^T_{flow}\ \underline{P}^T_{inj}\ \underline{Q}^T_{flow}\ \underline{Q}^T_{inj}\ \underline{V}^T_{meas}\right]^T_{1xm}. \tag{8.15}$$

The partial derivatives are readily available from (8.1) and (8.2). For power flows from bus i to bus j,

$$\begin{aligned}
\frac{\partial P_{ij}}{\partial \theta_i} &= V_i V_j \left(G_{ij}\, sin\,\theta_{ij} - B_{ij}\, cos\,\theta_{ij}\right) = -\frac{\partial P_{ij}}{\partial \theta_j} \\
\frac{\partial P_{ij}}{\partial V_i} &= 2V_i G_{ij} - V_j \left(G_{ij}\, cos\,\theta_{ij} + B_{ij}\, sin\,\theta_{ij}\right) \\
\frac{\partial P_{ij}}{\partial V_j} &= -V_i \left(G_{ij}\, cos\,\theta_{ij} + B_{ij}\, sin\,\theta_{ij}\right) \\
\frac{\partial Q_{ij}}{\partial \theta_i} &= -V_i V_j \left(G_{ij}\, cos\,\theta_{ij} + B_{ij}\, sin\,\theta_{ij}\right) = -\frac{\partial Q_{ij}}{\partial \theta_j} \\
\frac{\partial Q_{ij}}{\partial V_i} &= -2V_i \left(B_{ij} + \tfrac{1}{2} B_{shij}\right) - V_j \left(G_{ij}\, sin\,\theta_{ij} - B_{ij}\, cos\,\theta_{ij}\right) \\
\frac{\partial Q_{ij}}{\partial V_j} &= -V_i \left(G_{ij}\, sin\,\theta_{ij} - B_{ij}\, cos\,\theta_{ij}\right)
\end{aligned} \tag{8.16}$$

It follows for the injection at bus i, with its adjacent buses j = l,…, k , that

$$\frac{\partial P_i}{\partial V_i} = \sum_{j=\ell}^{k} \frac{\partial P_{ij}}{\partial V_i} \qquad \frac{\partial P_i}{\partial \theta_i} = \sum_{j=\ell}^{k} \frac{\partial P_{ij}}{\partial \theta_i}$$
$$, \qquad , \qquad (8.17)$$
$$\frac{\partial Q_i}{\partial V_i} = \sum_{j=\ell}^{k} \frac{\partial Q_{ij}}{\partial V_i} \qquad \frac{\partial Q_i}{\partial \theta_i} = \sum_{j=\ell}^{k} \frac{\partial Q_{ij}}{\partial \theta_i}$$

$$\frac{\partial P_i}{\partial V_j} = -\frac{\partial P_{ij}}{\partial V_i},\ j = \ell,\ldots,k, \text{ and } \frac{\partial Q_i}{\partial V_j} = -\frac{\partial Q_{ij}}{\partial V_i},\ j = \ell,\ldots,k. \qquad (8.18)$$

Note that the adjacent bus numbers j = ℓ,…,k must be the same for real power measurements and reactive power measurements, since the constraint equation governing (8.2) holds regardless of the presence of power flow measurements on the adjacent lines.

In the case of a transformer as described in Section 4.2, where a is the tap position, (8.16) becomes

$$\begin{aligned}
\frac{\partial P_{ij}}{\partial a} &= V_i V_j \frac{B_{ij}}{a^2} \sin(\theta_{ij}) = \frac{\partial P_i}{\partial a} \\
\frac{\partial P_{ji}}{\partial a} &= V_i V_j \frac{B_{ij}}{a^2} \sin(\theta_{ji}) = \frac{\partial P_j}{\partial a} \\
\frac{\partial Q_{ij}}{\partial a} &= 2\frac{B_{ij}}{a^3} V_i^2 - V_i V_j \frac{B_{ij}}{a^2} \cos(\theta_{ij}) = \frac{\partial Q_i}{\partial a} \\
\frac{\partial Q_{ji}}{\partial a} &= -V_i V_j \frac{B_{ij}}{a^2} \cos(\theta_{ij}) = \frac{\partial Q_j}{\partial a} \\
\frac{\partial a}{\partial a} &= 1
\end{aligned} \qquad (8.19)$$

For the voltage measurement vector $\underline{V}_{meas}$, the partial derivatives are quite simple:

$$\frac{\partial V_i}{\partial \theta_j} = 0, \forall j \text{ , and } \frac{\partial V_i}{\partial V_j} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \qquad (8.20)$$

Hence we have derived the formulation of all entries of the Jacobian. Naturally, to form $\underline{H}^{(k)}$ during iteration k of the IRLS algorithm, the values of $\underline{x}^{(k)} = [\underline{\theta}^{(k)}\ \underline{V}^{(k)}]$ are used. The Jacobian is a function of $\underline{x}$ and so may change as $\underline{x}^{(k)}$ converges. However, as $\underline{x}^{(k)}$ nears the solution, it changes very little, thus it may become wasteful to continue updating the

Jacobian after a certain number of iterations of the IRLS. In fact, it might be blocked after as early as two iterations.

8.5. Equivariance Properties

The estimators we will be interested in need to enjoy equivariance properties. We are specifically interested in the following properties:

1) An estimator $\hat{\theta}_m$ is said to be location equivariant if

$$\hat{\theta}_m(z_1 + v, \ldots, z_m + v) = v + \hat{\theta}_m(z_1, \ldots, z_m), \ \forall v \in \Re \tag{8.21}$$

2) An estimator $\hat{\theta}_m$ is said to be scale equivariant if

$$\hat{\theta}_m(cz_1, \ldots, cz_m) = c\hat{\theta}_m(z_1, \ldots, z_m), \ \forall c \in \Re, \ c \neq 0 \tag{8.22}$$

We are also interested in equivariance properties of multivariate estimators. Letting $\forall \_ \in \ ^n$ $\underline{A}$ be an , non-singular matrix, we make the following definitions:

3) A multivariate location estimator T is said to be translation equivariant if

$$T(\underline{\ell}_1 + \underline{b} \ \ldots \ \underline{\ell}_m + \underline{b}) = \underline{b} + T(\underline{\ell}_1 \ \ldots, \underline{\ell}_m). \tag{8.23}$$

4) A multivariate location estimator T is said to be affine equivariant if

$$T(\underline{A}\underline{\ell}_1 + \underline{b}, \ldots, \underline{A}\underline{\ell}_m + \underline{b}) = \underline{A}T(\underline{\ell}_1, \ldots, \underline{\ell}_m) + \underline{b}. \tag{8.24}$$

5) A multivariate covariance estimator $\underline{C}$ is said to be affine equivariant if

$$\underline{C}(\{\underline{A}\underline{\ell}_1 + \underline{b}, \ldots, \underline{A}\underline{\ell}_m + \underline{b}\}) = \underline{A}\,\underline{C}(\{\underline{\ell}_1, \ldots, \underline{\ell}_m\})\underline{A}^T. \tag{8.25}$$

8.6. Probability Distributions

This section provides a reference for the probability distributions discussed in Chapter 3. The Gaussian, or normal, probability distribution function, with mean μ and variance σ^2, is written as

$$\varphi(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}. \tag{8.26}$$

Since a closed form of its cumulative distribution function $\Phi(x) = \int_{-\infty}^{x} \varphi(t)dt$ does not exist, $\Phi(x)$ is determined numerically, and usually through tables.

The Laplacian, or double exponential, probability distribution function, with mean $\mu = E[x] = b$, and variance $E[(x-\mu)^2] = 2a^2$, is given by

$$f(x) = \frac{1}{2a} e^{-\frac{|x-b|}{a}}. \tag{8.27}$$

Laplacian random variables can be transformed from random variables U following the uniform distribution $U_{[0,1]}$ through the inverse cumulative distribution function (where b = 0), given by

$$x = F^{-1}(y) = \begin{cases} -a \ln(2 - 2y), & y \geq \frac{1}{2} \\ a \ln(2y), & y < \frac{1}{2} \end{cases}. \tag{8.28}$$

The Cauchy probability distribution function, whose mean and variance are undefined, is given by

$$f(x) = \frac{1}{\pi a}\left(\frac{1}{1 + \frac{(x-b)^2}{a^2}}\right). \tag{8.29}$$

The parameters a and b measure the scale and location of the distribution, respectively. Cauchy random variables can be generated from $U_{[0,1]}$ through its inverse cumulative distribution function, which when b = 0 and a = 1 is given by

$$x = F^{-1}(y) = tan\left(\pi y - \frac{\pi}{2}\right). \tag{8.30}$$

The chi-squared distribution, based on the gamma distribution, is defined as

$$f(x) = \begin{cases} \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \; x^{\nu/2-1} \, e^{-x/2}, & x > 0 \\ 0, & \text{elsewhere} \end{cases} \tag{8.31}$$

where ν is a positive integer, and the gamma function is defined as

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\,dx, \qquad \text{for } \alpha > 1. \tag{8.32}$$

8.7. Variance Calculations

We begin with the standard normal distribution $\Phi = N(0,1)$ for convenience, and calculate the asymptotic variance by (3.18). The mean's ψ-function makes it easy to compute the variance, but it can be shown that for Fisher consistent maximum likelihood estimators, $E[\psi^2(r)] = E[\psi'(r)]$, that is, we obtain $\mathrm{Var}(\hat{\theta}; \Phi) = 1$, the Cramèr-Rao lower bound for variance at the normal. For the median, we have

$$E[\psi^2(r)] = \int_{-\infty}^{\infty} (\mathrm{sign}(r))^2 \varphi(r)\,dr = 1$$

and

$$E[\psi'(r)] \int_{-\infty}^{\infty} 2\delta(0)\varphi(r)\,dr = 2\varphi(0)\ ,$$

so that $\mathrm{Var}(\hat{\theta}; \Phi) = \left(\left(2/\sqrt{2\pi}\right)^2 \right)^{-1} \approx 1.5708$. For the Huber, the variance is a function of its cutoff value, and we choose a typical value of 1.5. Hence, using symmetry, some calculus, and a statistical table for area under the standard normal curve,

$$E[\psi^2(r)] = 2\int_{-\infty}^{-b} b^2 \varphi(r)\,dr + \int_{-b}^{b} r^2 \varphi(r)\,dr = 2b^2\Phi(-b) + \left(1/\sqrt{2\pi}\right)\int_{-b}^{b} r^2 e^{-r^2/2}\,dr$$

$$= 2(1.5)^2(0.0668) + 0.4778 = 0.7784$$

$$E[\psi'(r)] = \int_{-b}^{b} \varphi(r)\,dr = \Phi(b) - \Phi(-b) = 0.9332 - 0.0668 = 0.8664$$

Its variance at Φ is thus $\mathrm{Var}(\hat{\theta}; \Phi) \approx 1.037$.

By (2.14), the median and Huber's asymptotic efficiencies relative to the mean are thus 0.637 and 0.964, respectively. We conclude that the mean is of course the best bet at a Gaussian distribution and the median suffers a 36% loss in efficiency, while the Huber has a good efficiency, losing only 4%. Again the Huber's efficiency is dependent on its cutoff value b, and its variance decreases with increasing values of b.

The Laplacian distribution is chosen with a zero location and a large variance of $2a^2 = 100$. The mean's asymptotic variance is $\mathrm{Var}(\hat{\theta};F) = E[\psi^2(r)] \;/\; (E[\psi'(r)])^2 = 20\,000$. The median, which is the maximum likelihood estimator at the Laplacian, is asymptotically efficient with a variance of $\mathrm{Var}(\hat{\theta};F) = 1 / (1/a)^2 = 10\,000$. The Huber performs extremely well, with an asymptotic variance of

$$\mathrm{Var}(\hat{\theta};F) = \frac{\int_{-b}^{b} \frac{r^2}{2a} e^{-|r|/a}\,dr + b^2\left(1 - \int_{-b}^{b} f(r)dr\right)}{\int_{-b}^{b} f(r)dr} \approx \frac{2.2276}{(0.014888)^2} = 10\,050\,.$$

Hence the relative asymptotic efficiencies of the mean and Huber with respect to the median are 0.5 and 0.995, respectively.

The Cauchy distribution is chosen with a zero location and scale parameter of a = 1. At any Cauchy distribution, the mean's averages $\bar{q}$ will fluctuate about the distribution's center, while the asymptotic variance is infinite. The infinite variance is seen from the fact that $E[\psi'(r)] = 1$ and

$$E[\psi^2(r)] = \int_{-\infty}^{\infty} \frac{r^2}{\pi} \frac{dr}{1+r^2} = \infty\,.$$

The median has an asymptotic variance of $\mathrm{Var}(\hat{\theta};F) = 1 / (2\,f(0))^2 = 1 / (2/\pi)^2 \approx 2.467$. The Huber's is given by

$$\mathrm{Var}(\hat{\theta};F) = \frac{\int_{-b}^{b} \frac{r^2}{\pi} \frac{dr}{1+r^2} + b^2\left(1 - \int_{-b}^{b} f(r)dr\right)}{\int_{-b}^{b} f(r)dr} \approx \frac{1.1715}{(0.6256659)^2} = 2.993\,.$$

Hence, the relative asymptotic efficiency of the mean is zero, whereas the median and the Huber perform similarly to each other.

The above results are now verified through Monte Carlo simulations. A random number generator using the Box-Muller method as described in [28] is used to generate 1000 samples containing 500 normally distributed numbers. After this, an estimate of the location $\hat{\theta}_i$ of each sample is obtained for each of the M-estimators through one iteration of

the IRLS algorithm. The average of the estimators is determined by $\bar{\hat{\theta}} = \frac{1}{1000}\sum_{i=1}^{1000}\hat{\theta}_i$, and the variance by $\mathrm{Var}\left(\sqrt{500}\left(\hat{\theta} - \bar{\hat{\theta}}\right)\right) = \frac{500}{999}\sum_{i=1}^{1000}\left(\hat{\theta} - \bar{\hat{\theta}}\right)^2$.

For the Laplacian and Cauchy simulations, a random number generator was used to create 1000 samples, each containing 500 uniform, i.i.d. random variables, and these were transformed into Laplacian and Cauchy variables through (8.28) and (8.30). The Laplacian samples were transformed so as to have a variance corresponding to the parameter a = 100, while the Cauchy samples correspond to a spread with the parameter a = 1. The same calculations of $\bar{\hat{\theta}}$ and $\mathrm{Var}\left(\sqrt{500}\left(\hat{\theta} - \bar{\hat{\theta}}\right)\right)$ were made as for the standard normal samples, and the obtained results are recorded in Table 3.2.2.

VITA

Michael A. Chapman was born in Spartanburg, South Carolina on April 23, 1973. He received his B.S. degree in Electrical Engineering from the Virginia Polytechnic Institute and State University, Blacksburg, VA in 1996. He performed research and learned the French language at the Ecole Polytechnique Fédérale de Lausanne in Lausanne, Switzerland during the 1996-1997 school year and in Fribourg, Switzerland. During the summer of 1998 he gained industrial experience in an internship with the Entreprises Electriques Fribourgeoises utility, where he installed a state estimator package. He is currently in the process of completing the requirements of M.S. degree in Electrical Engineering at the Virginia Polytechnic Institute and State University, his area of research being robust estimators applied to power system state estimation.

Michael A. Chapman