UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS

Júlio Augusto Druzina Massignan

A Bayesian Perspective for Distribution System State Estimation: Theoretical and Practical Considerations

São Carlos

2021

Júlio Augusto Druzina Massignan

A Bayesian Perspective for Distribution System State Estimation: Theoretical and Practical Considerations

Thesis presented to the School of Engineering of São Carlos of the University of São Paulo, to obtain the degree of Doctor of Science - Electrical Engineering

Concentration Area: Electric Power Systems

Supervisor: Prof. Dr. João Bosco Augusto London Junior

São Carlos

2021

Note: Trata-se da versão corrigida da tese. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-Graduação de Engenharia Elétrica.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da EESC/USP com os dados inseridos pelo(a) autor(a).

Massignan, Julio Augusto Druzina
M417a A Bayesian Perspective for Distribution System State Estimation: Theoretical and Practical Considerations / Julio Augusto Druzina Massignan; orientador João Bosco Augusto London Junior. São Carlos, 2021.

Tese (Doutorado) - Programa de Pós-Graduação em Engenharia Elétrica e Área de Concentração em Sistemas Elétricos de Potência -- Escola de Engenharia de São Carlos da Universidade de São Paulo, 2021.

1. State Estimation. 2. Distribution Systems. 3. Power Systems. 4. Bayesian Inference. 5. Information Fusion. 6. Decomposition Methods. 7. Information Theory. 8. Smart Grids. I. Título.

Eduardo Graziosi Silva - CRB - 8/8907

FOLHA DE JULGAMENTO

Candidato: Engenheiro **JÚLIO AUGUSTO DRUZINA MASSIGNAN.**

Título da tese: "Uma Perspectiva Bayesiana para Estimação de Estado em Sistemas de Distribuição: Aspectos Teóricos e Práticos".

Data da defesa: 05/08/2021.

Comissão Julgadora	**Resultado**
Prof. Associado **João Bosco Augusto London Junior (Orientador)** (Escola de Engenharia de São Carlos/EESC-USP)	Aprovado
Prof. Dr. **Djalma Mosqueira Falcão** (Universidade Federal do Rio de Janeiro/UFRJ)	Aprovado
Prof. Dr. **Madson Cortes de Almeida** (Universidade Estadual de Campinas/UNICAMP)	Aprovado
Prof. Dr. **Vladimiro Henrique Barrosa Pinto de Miranda** (Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência/INESC TEC)	Aprovado
Prof. Dr. **Paolo Attillio Pegoraro** (Universidade de Cagliari/Itália)	Aprovado

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:
Prof. Dr. **João Bosco Augusto London Junior**

Presidente da Comissão de Pós-Graduação:
Prof. Titular **Murilo Araujo Romero**

This Thesis is dedicated to my supportive and loving parents,
Eduardo and Mara,
my dear brothers
João Paulo, Luiz Fernando and Eduardo Junior
an to my beloved wife
Geise.

ACKNOWLEDGEMENTS

Thanks to my family, my father Eduardo Massignan, my reference of wisdom and character, my mother Mara Lucia, my source of love and happiness, and my brothers João Paulo, Luiz Fernando, Eduardo Junior, for their unconditional love, the support they gave me, the many teachings, for believing in my potential, and for all our many laughs together.

Thanks to my beloved wife Geise, whose affection, support, patience and trust were my safe harbour in all these years, and with whom I am very happy to share special moments. Thanks also to Ivonete, her mother, Joyce, her sister, a second family that I felt welcomed in.

Thanks to Prof. João Bosco Augusto London Jr. that provided me many opportunities and helped me to succeed, for his exceptional support and guidance throughout my career, for providing the resources and space for my research, for his friendship and many councils all these years, and for being the caring person he is, someone I can call a true mentor in my life.

Thanks to Prof. Vladimiro Miranda that expanded my research horizon, reinforced my confidence, for the many histories and wisdom shared, and for his true enthusiasm for science.

Thanks to Prof. Carlos Maciel and Prof. Benvindo Pereira for the many discussions, exploring new concepts, and for their collaborative support and friendship in all these years.

Thanks to all involved in the PD ANEEL 2866-0504/2018 and to COPEL Distribuição S/A, whose opportunities enhanced my research and enabled the practical applications, and specially to Marcos Camillo, Rodrigo Fannuchi, Luciano Bento and Profa. Telma Soares.

Thanks to my great friends Michel Bessani, Guilherme Fuzato and Felipe Pastori for the many laughs, shared perspectives and bike rides that were essential for keeping myself up.

Thanks to my friends of the University Taylon Landgraf, Luana Locatelli, José Paulo, Gustavo Hebling, Francisco Lemes, Etiane Ponciano, Nayara, Vitor Melo, Miguel Modesto, Gabriel, Jéssica Freire, Jéssica Santos, Diego Dourado, Elian, Maurício Escalante, Breno Bretas, José Nuno, Murilo Bento, Leandro Tolomeu, Camilla Fantin, Guliherme Borges, Alexandre Sohn, Daniel Siqueira, Ana Cecília, Renato Oliveira.

Thanks to my great friends from INESC-TEC and Porto Luiz Eduardo, Fernando Ribeiro, Shabnam Pesteh, Patricia Vale, João Victor Costa, Igor Resende, Mohamed Lotfi, Rogério Rocha, Matthew Gough, Sérgio, Diego Piserà, Simone Galisai, Morteza, Sérgio Santos, Kamal, Paula Raissa, Clarissa and Fabrício Barros.

Also to my friends from life Lauro Ito, Rodrigo Sampaio, Caio César, João Ricardo, Lucas Nishioka, Leandro Rissato, Gustavo Siqueira, Brhuno Ribeiro, Gabriel Faustino, Jefferson Amadeu, Carlos Gehring, Vinícuis Rosseti, Diego Medeiros, Carlos Silva, Alexandre Nonato, Leonardo Hernandes, Lucas Carrer, Thiago Martinelli, Adilson Barros, Wanderson Alves, Lean-

dro Marcelino, Rogério Souza, Fábio Urban, Jonas Ansoni, Willian Vilela, Luis Marquiti, Mario Sato, Viviane Kim, Marcos, and the many others that are special but I fail to remind here.

Thanks to the University of São Paulo and the Department of Electrical and Computing Engineering of the School of Engineering of São Carlos, especially to Marisa, Daniel and Jussara.

And thanks to CNPq (grant no. 142220/2016-7) and FAPESP (grants no. 2016/19646-6 and 2018/00214-4) for the financial support that made this research, and so many others, possible.

"Não se pode ter sol na eira e chuva no nabal."
Portuguese Saying

"Whether you can observe a thing or not depends on the theory which you use. It is the theory which decides what can be observed."
Albert Einstein

ABSTRACT

MASSIGNAN, J. A. D. **A Bayesian Perspective for Distribution System State Estimation: Theoretical and Practical Considerations**. 2021. 237p. Thesis (Doctor of Science) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2021.

Massive and heterogeneous data sources are becoming incrementally available at power distribution networks, due to enhancements on traditional SCADA monitoring, including advanced metering infrastructures, and installing new sensors such as phasor measurement units. Data management then becomes a crucial process for operation and control of such power networks, for processing such diverse information, performing network assessment and optimizing decisions. Within such a perspective, distributions operators rely on state estimation applications, bridging the information from measured data with detailed physics-based models of power grids.

This Thesis extends the concepts of distribution system state estimation under a Bayesian Inference perspective, exploring a probabilistic interpretation for the state variables and associated randomness instead of only seeking to calculate a fixed state vector. This work employs this conceptual framework under three distinct and novel applications in the context of electrical distribution networks: dealing with non-Gaussian noise models under a Correntropy Extended Kalman Filter in power system state estimation, both from measurement noise and state behaviour; the proposition of a Bayesian information fusion to merge data gathered from pseudo-measurements, smart meters, SCADA and phasor measurements in distribution networks; the exploration of scalability of the three-phase unbalanced state estimation under a multiarea procedure based on Bayesian spatial fusion.

Besides, a high-resolution and detailed model for distribution network is presented in the form of a generic component model based on a two-port admittance matrix formulation, improving the resolution of digital twin models for three-phase, unbalanced and asymmetrical distribution networks from the high voltage substations and primary feeders to low voltage secondary circuits. The proposed Thesis also employs an orthogonal formulation and sparsity treatments to overcome numerical conditioning issues, a well-known challenge for classical state estimation formulations, while enhancing computational efficiency to ensure real-time performance.

The developed algorithms and frameworks are evaluated on the IEEE test feeders and by the application of the proposed methods on real Brazilian test systems (both at distribution and transmission levels). The results corroborate the crucial task of including temporal characteristics on the state estimation while dealing with more generic noise characteristics under a kernel density concept, while properly tunning the kernel's bandwidths under different system's transitions. The use of information fusion shows itself as an essential practical resource to deal with different sampling and updating rates of the diverse set of measurements employed in distribution networks, especially when abrupt transitions are present while improving computational performance. Besides, the multiarea decomposition methods, along with sparse orthogonal formulations, are

prominent in ensuring scalability and numerical stability of the estimation as a whole, a crucial practical contribution for large-scale distribution networks assessment.

Keywords: State Estimation, Distribution Systems, Power Systems, Bayesian Inference, Information Fusion, Decomposition Methods, Information Theory, Smart Grids.

RESUMO

MASSIGNAN, J. A. D. **Uma Perspectiva Bayesiana para Estimação de Estado em Sistemas de Distribuição: Aspectos Teóricos e Práticos**. 2021. 237p. Tese de Doutorado - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2021.

Massivas e heterogêneas fontes de dados estão se tornando cada vez mais disponíveis em redes de distribuição de energia, devido a aprimoramentos nos tradicionais sistemas de monitoramentos SCADA, incluindo infraestrutura de medição avançada, e instalando novos sensores como as unidades de medição fasorial. Gerenciamento de dados torna-se então um processo crucial para operação e controle das redes elétricas, para o processamento dessas diversas informações, realizar avaliação da rede elétrica e otimizando a tomada de decisões. Dentro desta perspectiva, os operadores dos sistemas de distribuição dependem de aplicações de estimação de estado, conectando a informação obtida de dados telemedidos com detalhados modelos físicos das redes elétricas.

A presente Tese estende os conceitos de estimação de estado para sistemas de distribuição sob a perspectiva de Inferência Bayesiana, explorando uma interpretação probabilística para as variáveis de estado e incerteza associada ao invés de buscar o cálculo de um vetor de estado fixo e determinístico. Este trabalho emprega este arcabouço conceitual em três distintas e inovadoras aplicações no contexto de redes de distribuição de energia elétrica: tratamento de modelos de ruído não Gaussianos através de um Filtro de Kalman Estendido por Correntropia em estimação de estado para sistemas elétricos, tanto ruido de medida como comportamento estocástico do estado; a proposta de Fusão Bayesiana de Informações para mesclar dados provenientes de pseudo medidas, medidores inteligentes, SCADA e medidas fasoriais em sistemas de distribuição; a exploração de escalabilidade de estimação de estado trifásica e desbalanceada em procedimentos multiárea e baseados em fusão espacial Bayesiana.

Além disto, modelos detalhados e de alta resolução para redes de distribuição são apresentados na forma de um modelo genérico baseado na formulação da matriz de admitância de quadripólos, aumentando a resolução de modelos *digital twin* para redes elétricas trifásicas, desbalanceadas e assimétricas, desde as subestações de alta tensão e alimentadores primários até os circuitos de baixa tensão da rede secundária. A Tese proposta também emprega uma formulação ortogonal e tratamentos de esparsidade para superar problemas de condicionamento numérico, um conhecido desafio para formulações clássicas de estimação de estado, ao mesmo tempo que aumentando eficiência computacional para aplicações em tempo real.

Os arcabouços e algoritmos desenvolvidos são avaliados em alimentadores de teste do IEEE e pela aplicação das metodologias em alimentadores reais brasileiros (tanto em sistemas de

distribuição como de transmissão). Os resultados corroboram com a necessidade crucial de se incluir características temporais sob conceitos de estimação de *kernel*, em conjunto com ajustes adequados das bandas para representar transições sistêmicas. O uso de fusão de informações demonstra-se como prática essencial para lidar com as diferentes taxas de amostragem e atualização dos conjuntos de medidas em toda sua diversidade nos sistemas de distribuição, especialmente quando transições abruptas se fazem presentes e ao mesmo tempo aprimorando a performance computacional. Além disto, técnicas de decomposição multiárea, em conjunto com formulações ortogonais esparsas, São proeminentes em garantir escalabilidade e estabilidade numérica do processo de estimação como um todo, uma contribuição prática crucial para avaliação de redes de distribuição de larga escala.

Palavras-chave: Estimação de Estado, Sistemas de Distribuição, Sistemas ELétricos, Inferência Bayesiana, Fusão de Informações, Métodos de Decomposição, Teoria da Informação, Redes Inteligentes.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AI	Artificial Intelligence
AMB	Admittance Matrix Based
BC	Branch Current
CIM	Correntropy Induced Metric
DS	Distribution System
DSE	Dynamic State Estimation
DSO	Distribution System Operator
DSSE	Distribution System State Estimation
EKF	Extended Kalman Filter
FASE	Forecasting-Aided State Estimation
GMLE	Generalized Maximum Likelihood Estimation
GE	Gross Error
IED	Intelligent Electronic Device
IoT	Internet of Things
IEKF	Iterated Extended Kalman Filter
KF	Kalman Filter
LMS	Least Median Squared
LRNT	Largest Normalized Residual Test
MAE	Mean Absolute Error
MAP	Maximum a Posteriori
MASE	Multiarea State Estimation
MCMC	Monte Carlo Markov Chain
MCC	Maximum Correntropy Criterion
MCEKF	Maximum Correntropy Extended Kalman Filter

PMU	Phasor Measurement Unit
pdf	probability distribution function
RTU	Remote Terminal Unit
SCADA	Supervisory Control and Data Acquisition
SHGM	Schweppe-Huber Generalized M-Estimator
UI	Undetectablity Index
WLS	Weighted Least Squares
WLAV	Weighted Least Absolute Value

CONTENTS

1 INTRODUCTION

1.1 Motivations

Energy management is a decisive technology on modern electric power systems, and it enables secure and optimized operation of this complex infrastructure. But, what can you manage, if you cannot measure it? Or, in the case of power systems, estimate it? The crucial processes of energy management and monitoring are in the spotlight of digitalization to accomplish the decarbonization goals, since they can, for instance, optimize the integration of renewable and distributed energy resources to meet the rising electricity demand and enable novel market dynamics in the electricity sector.

The International Energy Agency and the European Green Deal point to digital technologies as a critical enabler for realizing sustainability goals. For instance, Artificial Intelligence (AI), 5G, cloud and edge computing and the internet of things all together are the heart of data-driven innovation, providing monitoring and optimizing energy usage in the medium and low-voltage electricity distribution networks (EU Commission, 2019; IEA, 2019). The digitalization of distribution grids in this new context means a cyber-physical system accommodating high levels of renewable energy resources, electric vehicles, energy storage and local markets, monitored by widespread sensors, and supported by advanced analytics for decision-making (Kezunovic et al., 2020; Kadurek et al., 2014; Lopes et al., 2020; Haque et al., 2019; Miranda et al., 2019).

In the above context, real-time monitoring of power systems has a pivotal role in ensuring the energy management, intimately connected to the digitalization of the electricity sector. This is a fundamental need, especially at the distribution level of the power systems, where a vast and diverse set of new green technologies are expected to be included. The computational tool responsible for the real-time monitoring of power system is the **State Estimator**, which comprises the objective, in a succinct overview, of obtaining the state of the power grid by processing the information gathered from telemetry sensors along with a mathematical model of the electrical networks.

Thereby, the state estimation problem emerges as a bridge between measured information and physics-based network model to accurately assess the network operational condition in real-time (Abur; Gómez-Expósito, 2004; Monticelli, 1999). The conceptual background for power system state estimation was mainly proposed in the 1970s, 1980s and 1990s, with few modifications on the fundamental concepts since then, but with a large number of enhancements of complementary features such as alternative formulations, integration of new types of sensors to the Supervisory, Control and Data Acquisition (SCADA) systems, increasing computational performance, and extending models to capture the power system dynamics. Even though the

state estimator consolidated itself as an essential computational tool in transmission systems operation, this is not the primary reality on distribution systems.

Historically, one of the main challenges for the real-time determination of the distribution system operating state is the lack of real-time measurements in distribution feeders. Therefore, the operating condition of distribution feeders is usually determined through statistical characterization of their loads performed by a process called load aggregation (Feng; Yang; Peterson, 2012; Massignan et al., 2018; Džafić et al., 2013). This process is based on customer monthly energy consumption (kWh), customer classification, e.g., residential, commercial and industrial, and typical load profiles for each customer class, typically called pseudo measurements. As such profiles are only a rough approximation to given load demand, the quality of the feeder operating state obtained is low and does not have similar precision for distribution system monitoring compared to measured values along the feeder.

In recent years, smart grids initiatives have been developed and created new data sources at unprecedented volumes in distribution systems (e.g., phasor measurement units, intelligent electronic devices and smart meters). These initiatives have motivated the proposition and development of distribution system state estimators to provide the real-time monitoring capabilities required by the many new automatic functionalities. Several works in distribution system state estimation pointed some crucial challenges yet to be dealt with, from both a theoretical as well as practical requirements (Gómez-Exposito et al., 2011a; Lefebvre; Prévost; Lenoir, 2014; Della Giustina et al., 2014; Liao; Milanovic, 2016; Primadianto; Lu, 2017; Ahmad et al., 2018; Dehghanpour et al., 2019):

1. Incorporation of new equipment related to the automation of distribution systems, as the inclusion of smart meters and synchronized phasor measurements in the state estimation process;
2. Improvements on the models employed in the estimation process, such as new statistical models for the measurement set, including uncertainty in the parameters and topology of the network, and extending models to represent the low voltage secondary systems;
3. Computational aspects for real-time monitoring of the wide area and large scale of distribution networks, along with the possibility of decentralization of monitoring and control of microgrids;
4. Increasing the integration among different operators from transmission and distribution systems, especially regarding the exchange of information, to increase operational flexibility.

Such challenges are consonant to the paradigm shift in distribution system operation under the context of *Smart Grids* initiatives, which also relies on distribution system state estimation to provide accurate and detailed information about the network in real-time. For instance, a significant shift is the advanced metering infrastructure, through the deployment of

many smart meters and providing a high volume of information at the low voltage networks, since they are often installed directly at the consumer units. This deployment increases the amount of measured information about the network condition, reducing the state estimation dependence on pseudo measurements and sensibly enhancing accuracy.

Although it has been the topic of many recent scientific research efforts, the state estimation problem for distribution systems still presents many enhancement opportunities and open challenges. Due to the complexity and scale of real distribution systems, the developed methods do not always comply with practical requirements. Often multiple different methodologies are integrated to deal with specific parts of the distribution networks, without a unified theoretical framework to bring them together in a coherent manner. The recent advances in advanced data analytics and artificial intelligence also enlarge the knowledge frontier for the real-time operation of power systems. In view of the above considerations, this Thesis brings a new perspective to improve distribution system state estimation by employing Bayesian Inference concepts to tackle such practical and theoretical issues.

1.2 Objective and Contributions

The main objective of this Doctoral Thesis was to develop and implement software for Distribution System State Estimation, which can be applied to large scale distribution networks, deal with heterogeneous sources of information and their particularities, and enhance accuracy while dealing with Non-Gaussian noise characteristics. The developed Distribution System State Estimators are featured to deal with three-phase, unbalanced, asymmetrical, untransposed, radial or meshed networks without loss of precision.

The main contributions of this Thesis can be devised as:

1. a new probabilistic perspective for power system state estimation based on the Bayesian Inference framework, primarily directed to deal with distribution systems practical challenges;

2. the development of a Bayesian Information Fusion methodology to deal with the lack of synchronism and different sampling rates of measurements in modern distribution systems, processing pseudo measurements, smart meters, SCADA measurements and phasor measurement units in a single framework;

3. the proposition of a Multiarea State Estimation in the form of Bayesian Spatial Fusion method to deal with large scale distribution networks, by a decomposition of the distribution grid into sub-areas and processed in a hierarchical architecture;

4. the application of Information Theory concepts to deal with Non-Gaussian noise characteristics in power system state estimation, as a smoothed Bayesian non-Gaussian estimator, able to capture dynamic features of the estimation aided by a first-order state-space model.

Along with such contributions, some minors were also derived from this work, that can be summarized as:

- A new matrix formulation for the two-port admittance matrix model for representing power flow equations in a diverse set of components in distribution grids without loss of generalization, such as substations, primary feeders and low voltage networks;
- The application of orthogonal methods in the state estimation problem for distribution systems, associated to sparsity concepts to increase computational efficiency, to deal with numerical ill-conditioning of practically sized distribution networks;
- The introduction of a concept of sampling layers for dealing with multiple unsynchronized and different sampling rates of the diverse sources of information in modern distribution systems, as pseudo measurements, smart meters, SCADA measurements and phasor measurement units;
- The exploration of the concept of Bayesian credibility intervals to provide more meaning to the estimates and as detection threshold for state transitions due to systemic changes on the distribution networks, as in the case of sudden load changes and generation intermittency;
- The demonstration of scalability on real and practical sized distribution systems, dealing with different parts of the distribution grid, primary feeders, substations and low voltage circuits, in a new hierarchical estimation architecture;
- A new method to update Parzen windows in the context of Maximum Correntropy estimation for suppressing suspect samples that may be contaminated with gross errors or also in case of different kind of state transitions, such as sudden load changes, contingencies and the presence of gross errors.

1.3 Thesis Structure

Following this Introduction, the structure of this Thesis is summarized as follows:

Chapter 2 presents the literature review of the leading research efforts on distribution system state estimation. The intent is to provide an overview of fundamental challenges and novel advances for distribution system state estimation.

Chapter 3 introduces the new Bayesian perspective for power system state estimation. The exploration of the Bayesian Inference paradigm moves the conceptual interpretation of the state variables in the pure probabilistic sense, which is the main underlying framework employed by this Thesis.

Chapter 4 introduces the three-phase unbalanced component models proposed in this Thesis based on the proposed generic two-port admittance matrix formulation for multiple phases and electrical quantities measured.

Chapter 5 demonstrates a tracking state estimation algorithm based on Information Theory concepts to deal with non-Gaussian noise characteristics in a state-space formulation. In the Bayesian sense, the method consists of a smoothed version for non-parametric models in the variational sense, based on kernel density estimation principles along with a novel Kalman Filter formulation.

Chapter 6 presents a Bayesian information fusion method for distribution systems to deal with temporal aspects of the metering systems and time-varying systems. The concept of sampling layers treats each set of sensors separately according to the respective sampling and updating rates.

Chapter 7 lays out a Bayesian spatial fusion concept to support multiarea state estimation in large scale distribution networks. The decomposition of sub-areas is supported by the inclusion of prior knowledge that guarantees observability for local estimation.

The main conclusions and final considerations are provided in Chapter 8. Different future work possibilities are also highlighted, showing the diverse set of possible research paths associated with this Thesis.

Appendix A provides the theoretical background regarding power system state estimation. This chapter intends to provide a quick reference for readers with the main concepts and basic theory of traditional algorithms for distribution system state estimation.

Appendix B presents details about the sparse orthogonal formulation to perform state estimation employed in this work. This content is essential to ensure numerical and computational efficiency to deal with ill-conditioned distribution systems.

Appendix C introduces an application example of the Bayesian Inference concepts in the form of a Linear Bayesian Information Fusion procedure for integrating PMUs and SCADA state estimators according to the respective sampling rates.

Appendix D provides detailed models for most distribution system components, such as distribution circuits, transformers, voltage regulators, shunt capacitors and loads.

Finally, the Annex presents the main results in terms of publications obtained during the Doctoring period.

2 LITERATURE REVIEW

This chapter presents a literature review regarding power system state estimation, emphasizing applications and methods developed to deal with distribution networks. The objective is to provide an overview of the research efforts developed across the years and the many challenges still open, both in a theoretical and practical implementation. Therefore, clarifying the literature gap and contributions achieved by this work.

2.1 Power System State Estimation

Since its conception in the 1970s by the seminal work of Schweppe (Schweppe; Wildes, 1970), power system state estimation became an essential part of energy management systems. It extended several research and application efforts across the years and established state estimation as a crucial computational tool for the real-time operation of transmission systems (Abur; Gómez-Expósito, 2004; Monticelli, 1999; Zhao et al., 2020). Power system state estimation is the computational tool that enables, in real-time, obtaining the state variables of a power system, thus fully characterizing its operational condition. It consists of a network model that comprises component parameters and models, the updated topology of the power system, and redundant measurements of different electrical quantities of a power system, such as active and reactive powers, voltage magnitudes, or even voltage phasors and current phasors. It typically encompasses a static analysis of the power system associated with statistical models to represent measured information in real-time and may also incorporate dynamic features of the power system. The measurement system is comprised by the Supervisory, Control and Data Acquisition (SCADA) systems, that gather measured information from instrumentation devices and sensors spread across the electrical networks, mainly through automation protocols and communication with Remote Terminal Units (RTUs) and Intelligent Electronic Devices (IEDs) (Thomas, 2015).

The main definitions and concepts in power system state estimation were consolidated shortly after the proposition, fomenting new technological aspects throughout the years and mainly based on transmission system operators experience and requirements. Complementary functionalities were proposed for the state estimator: the network and topology configurator; observability analysis; the state estimation numerical process; and a bad data processing final validation (Abur; Gómez-Expósito, 2004; Monticelli, 1999). Each of these aspects were tackled with different theoretical perspectives, such as: decentralized state estimation with Multiarea concepts (Falcao; Arias, 1994; Gómez-Exposito et al., 2011b); enhancing robustness against bad data with non-quadratic functions (Mili; Phaniraj; Rousseeuw, 1991; Celik; Abur, 1992; Mili et al., 1996; Baldick et al., 1997; Pires; Costa; Mili, 1999; Miranda; Santos; Pereira, 2009); including temporal relations in the estimation process (Debs; Larson, 1970; Falcao; Cooke; Brameller, 1982; Bretas, 1989; Zhao et al., 2019); incorporating high sampling rate

measurements (Zhou et al., 2006; Zhu; Abur, 2007; Simões Costa; Albuquerque; Bez, 2013); introducing artificial intelligence support in the estimation process (Zhang; Wang; Giannakis, 2019; Zamzam; Sidiropoulos, 2020); improving observability with optimal meter placement (London Jr; Alberto; Bretas, 2007; Vigliassi et al., 2019); and many others. Appendix A presents details about the main concepts and formulations of power system state estimation, and readers may refer to it for additional theoretical background information.

Despite such rapid development and consolidation for transmission systems, there was a delayed interest toward distribution systems, which began only in the 1990s, with specialized methods for distribution system state estimation. The first research works began to explore distribution systems, and several characteristics reduced the application of the consolidated methods for transmission systems, requiring new methods and models more suitable to accommodate the distribution system peculiarities. Some of these peculiarities are well-known challenges for distribution system state estimation, such as (Hansen; Debs, 1995; Baran, 2012; Singh; Pal; Jabr, 2009; Primadianto; Lu, 2017; Ahmad et al., 2018; Švenda; Strezoski; Kanjuh, 2017; Huang; Lu; Lo, 2015; Pau; Pegoraro; Sulis, 2013; Lefebvre; Prévost; Lenoir, 2014; Massignan et al., 2018):

- diversity of connections, with single, two, and three-phase circuits, and transformer connections;
- unbalanced loads;
- short distance and untransposed lines with high resistance/reactance ratio;
- typically very large scale radial networks;
- presence of discrete variables, such as switching devices statuses and voltage regulators not directly monitored;
- reduced number of real time measurements, usually located at the substation and some particular components along the feeders;
- presence of current magnitude measurements instead of active and reactive power measurements;
- as a consequence of the above, the distribution system state estimation often is numerically ill-conditioned;

Although each approach for distribution systems state estimation has its own particularities and contributions, there are three primary research lines regarding the estimation algorithms, which encompass different definitions of electrical quantities as the state variables (Primadianto; Lu, 2017):

1. *Weighted Least Squares (WLS) State Estimator*: the three-phase version of the conventional state estimator for transmission systems, initially applied for distribution networks in (Baran; Kelley, 1994; Whei-Min Lin; Jen-Hao Teng, 1996), that uses the complex nodal voltages in polar coordinates as state variables;

2. *Admitance Matrix Based (AMB) State Estimator*: initially proposed in (Lu; Teng; Liu, 1995), this estimator is based on the nodal admittance matrix of the electrical network and in transformations of active and reactive power measurements in equivalent current phasor measurements, in rectangular coordinates, and also the complex nodal voltages in rectangular coordinates as state variables;

3. *Branch Current (BC) State Estimator*: initially proposed in (Baran; Kelley, 1995), this estimator is also based on equivalent current measurements in rectangular coordinates, but it uses the branch currents in rectangular coordinates as state variable, what makes the measurement model basically build upon an adjacency matrix of the system while voltage phasors are updated by an additional step based on a forward sweep method.

In addition, Figure 1 organizes such main topics in a timeline for the appearance and consolidation of transmission system state estimation and distribution system state estimation, highlighting the main research topics discussed in each period.

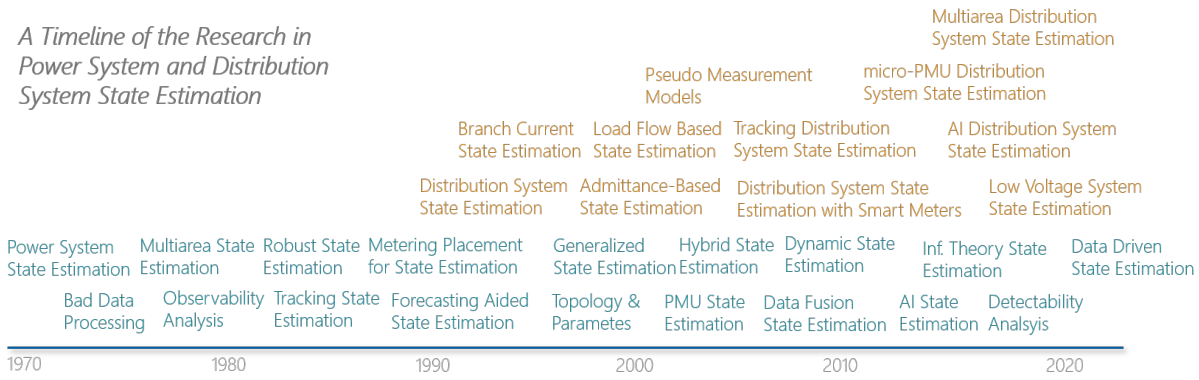


Figure 1 – Brief timeline of transmission systems and distribution system state estimation.

In general, the approaches and methodologies for real-time monitoring of distribution systems, based on state estimation, were also divided into different interconnected processes, similarly to transmission systems: Network Configurator; Observability Analysis; State Estimation; Bad Data Processing; and also often including a fifth process related to the pseudo measurements models, referred as Load Estimation. The flowchart in Figure 2 illustrates these processes and their relations and a typical information flow available for the state estimation in distribution systems. Like the state estimators, it is noteworthy that specific approaches were developed in the literature for the Network Configurator, Load Estimator, Observability Analysis,

and Bad Data Processing for distribution systems. However, they will not be highlighted in this review, keeping the main focus on state estimation itself.

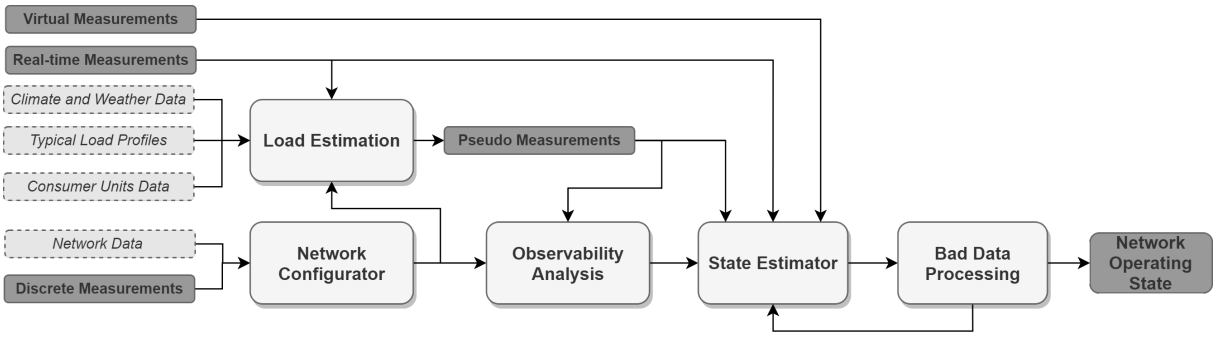


Figure 2 – Process flowchart of a typical Distribution System State Estimation framework and its main inputs.

The low amount of real-time measurements at the distribution systems, almost lacking real-time information, is the primary limiting factor for the widespread deployment of state estimators in distribution networks. This way, the operation centres typically use complementary tools to monitor their networks in real-time, mainly including several pseudo measurements in the state estimation process. Such pseudo measurements are based on the estimates of the active and reactive power injections of unmonitored nodes of the distribution systems (distribution transformer loads). This estimation process, conceptually near the state estimation problem, lead to the so-called Load Estimators (or load allocator) that can be used as a direct real-time monitoring tool for distribution networks or as pseudo measurements generator to aid the state estimators. This latter approach will not be discussed in this work, and more information regarding load estimation may be referred to in (Ghosh; Lubkeman; Jones, 1997; Youman Deng; Ying He; Boming Zhang, 2002; Džafić et al., 2013; Massignan et al., 2018).

Besides, different parts of the distribution systems also present their particularities regarding steady-state analysis and state estimation, which can be divided in the following (Kersting, 2001; Sallam; Malik, 2018):

- **High Voltage Substation**: comprises the power transformer with both high voltage and medium voltage nodes, as well as feeder's bays. Typically contains a high number of measurements from the Supervisory Control and Data Acquisition (SCADA) systems, and Phasor Measurement Units (PMUs) with a small number of state variables;
- **Medium Voltage Primary Feeders**: comprise the medium-voltage circuits and components typically spread across a large area, including urban and rural feeders. Contain a small number of measurements, typically from the SCADA system at specific equipment as automatic reclosers, voltage regulators, and shunt capacitors, or at particular consumers directly connected at medium voltage level. Moreover, it requires a large number of state

variables to be properly represented. Observability of such networks may be obtained through typical load curves in the form of load pseudo measurements or through new procedures capable of incorporating and processing new measurement technologies, such as Intelligent Electronic Devices (IEDs);

- **Low Voltage Secondary Circuits**: comprise the final connection with consumers at the low voltage level, with a diverse set of connections, grounding systems, and neutral conductors. Typical load profiles usually characterize the pseudo measurements. In the context of Advanced Metering Infrastructure (AMI) may present smart meters installed in all loads and at the local controller of the power transformers. They also present a relatively large number of state variables.

Figure 3 illustrates the typical structure and different components of a distribution network. A diverse set of equipment and components comprise the network, and different connections can lead to structural asymmetries in the distribution system. Meanwhile, different load connections and distinct load values in each phase result in additional unbalance.

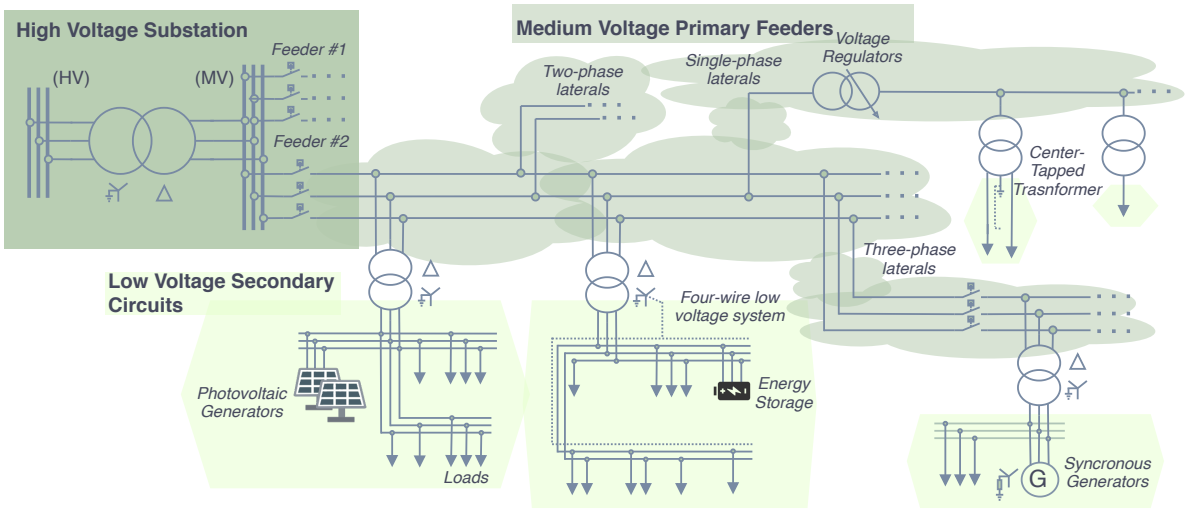


Figure 3 – Three-phase diagram of a generic distribution system comprising a high voltage substation, medium voltage primary feeders, and low voltage secondary circuits.

The advances in the Smart Grid concepts, associated with increasing computing power and new digital technologies, starting in the 2000s and emphasized in the last decade, gave new thrust for more research, applications and pilot projects of distribution system state estimation. The interest in deploying new measurements to the distribution systems reignited the interest in state estimation at the operation centres. These new measurements became available with the growing level of automation of distribution systems, fomented by the consolidation of standards like IEC 61850, installation of new controllers, an increase of flexibility with more switching devices, the inclusion of PMUs, and also by the evolution of the Smart Grids concept (Gómez-Exposito et al., 2011a; Della Giustina et al., 2014; Lefebvre; Prévost; Lenoir, 2014). In this sense,

it is noteworthy that the state estimator has a crucial role for distribution operators to take a more active posture in controlling distribution networks by transforming the data gathered from the many measurements across the network into tangible knowledge for decision making by the operators.

As for the state-of-the-art of the state estimation problem for distribution systems, the following topics are highlighted as the main challenges for research according to (Gómez-Exposito et al., 2011a; Lefebvre; Prévost; Lenoir, 2014; Della Giustina et al., 2014; Liao; Milanovic, 2016; Primadianto; Lu, 2017; Ahmad et al., 2018; Wang et al., 2019; Dehghanpour et al., 2019):

1. The incorporation of new measurement information from automation systems of distribution networks, along with the inclusion of smart meters and synchronized phasor measurements into the state estimation process;
2. Improvement of the mathematical models employed in the state estimation problem, as new probabilistic models for pseudo measurements, modelling uncertainty of network topology and parameters, and also comprehending low voltage circuits into the state estimation;
3. Computational aspects for wide-area monitoring, related to the large scale of distribution networks and also the decentralization of the control and operation of distribution networks;
4. The increased integration among distribution system operators and transmission system operators, both at the level of information exchange as increased flexibility requirements from novel market strategies;
5. The technical alignment with modern techniques of data science and machine learning algorithms, providing more knowledge about the grid in real-time as well as recognizing patterns in the system condition;

In the above context, new methods and algorithms for state estimation were proposed for distribution systems. Among such methods, the following topics are highlighted:

- new numerical methods for the solution of the state estimation problem and the extension of the state vector, such as: application of heuristics based on particle swarm optimization to solve the estimation problem in (Nanchian; Majumdar; Pal, 2017b); non-linear mixed integer programming to deal with discrete variables such as voltage regulators taps in (Nanchian; Majumdar; Pal, 2017a); extension of optimal power flow methods to state estimation (Duque et al., 2017); extensions of the Hamiltonian graph cycle theories for dealing with meshed distribution networks (Leite; Mantovani, 2016); the application of intelligent systems as neural networks (Manitsas et al., 2012) and also *autoencoders* (Barbeiro et al., 2015) to support the state estimation problem; with deep learning approaches for state estimation problem as in (Mestav; Luengo-Rozas; Tong, 2019); applying sparse

and orthogonal methods to deal with ill-conditioning as in (Hebling et al., 2020); and also in association with long short-term recurrent neural networks as surrogates functions for the state estimation problem in (Cao et al., 2020); estimation methods based on symmetric components in (Fernandes; Venkatesh; Almeida, 2021);

- new formulations and statistical theory for measurement model: closed-loop formulations for state estimation and pseudo measurements model (Hayes; Gruber; Prodanovic, 2015); extension of the state variables to accommodate active power injections (pseudo measurements) in the state estimation process (Džafić et al., 2013; Rousseaux et al., 2015); representation of correlation among measurements and pseudo measurements (Muscas et al., 2014); non-Gaussian statistical models for the pseudo measurements (Singh et al., 2011); non Gaussian measurement models with Monte Carlo based estimation methods (Pegoraro et al., 2017); state estimation based on new robust theories as M-estimators (Zhao et al., 2020), least absolute value (Göl; Abur, 2014) and information theoretic learning (Freitas; Costa; Miranda, 2017; Pesteh et al., 2019); and closed-loop dynamic state estimation with Kalman Filter approaches (Picallo; Bolognani; Dörfler, 2020); applications of artificial intelligence to support state estimation methods (Zhang; Wang; Giannakis, 2019; Zamzam; Sidiropoulos, 2020);

- variations of state estimators to accommodate new measurement technologies: inclusion of synchronized phasor measurements from PMUs (also referred as micro-PMUs for Distribution Systems) (Pau; Pegoraro; Sulis, 2013; Silva; Laburu; de Almeida, 2017) in distribution systems; approaches that include smart meters into state estimation for low voltage systems (Huang; Lu; Lo, 2015; Angioni et al., 2016a; Melo et al., 2016; Ni et al., 2018); approaches that include smart meters into state estimation for primary medium voltage feeders (Chen et al., 2016; Al-wakeel; Wu; Jenkins, 2016; Ni et al., 2018); treatment of the lack of synchronism and different time scales among smart meters and SCADA or PMU measurements (Alimardani et al., 2015; Gómez-Expósito; Gómez-Quiles; Džafić, 2015a);

- scalability and practical perspectives for distribution system state estimation: incorporation of practical perspectives into the state estimation problem (Švenda; Strezoski; Kanjuh, 2017; Fantin, 2016); propositions of large scale architectures based on multiarea state estimation (Nusrat et al., 2015; Pau et al., 2017); extension of new computing environments for real-time monitoring such as IoT and Cloud-based state estimation (Gonzaga et al., 2018; Pau et al., 2019; Meloni et al., 2018);

- solutions of correlated problems to state estimation: topology estimation methods (Singh et al., 2011); optimal meter placement (Xygkis; Korres, 2017); and bad data processing (Bretas et al., 2017; Krsman; Sarić, 2017).

Among the many exciting and promising research topics, this work explores mainly three lines inside the state-of-the-art distribution system state estimation, which will be presented in details in the following. The perspective for this choice is that both practical and theoretical contributions can be explored in these topics.

1. The problem of the lack of synchronism in the measurement set, due to different sampling and updating rates among the diverse set of sensors of distribution systems;

2. The large scale of distribution networks, alongside numerical and computational challenges associated with distribution system state estimation;

3. The need for suitable methods for non-Gaussian characteristic of measurement error, as well as for the state of the system.

2.2 Diversity of Sampling Rates and Temporal Aspects

The lack of synchronism in the measurement set arises from the different sampling and updating rates among the sensors installed in the electrical network. Essentially, by sampling the electrical quantities in different periods, each group of sensors monitors different operational conditions, different states, perceiving events on the grid that can be faster or slower according to the type of measurement.

The first attempts to deal with different sampling rates in the state estimation for DSs were based on the so-called hybrid approaches. Typically, these approaches assign different weights according to the sampling rates of measurements and introduce them directly in specialized DSSE algorithms, as in (Pau; Pegoraro; Sulis, 2013), and (Almeida; Ochoa, 2017). Indeed the majority of state estimators proposed for distribution systems (Lu; Teng; Liu, 1995; Baran, 2012), often do no tackle the problem of different sampling rates, mixing all measurements in the same measurement vector, thus pertaining to this class of hybrid algorithms.

The work in (Pau; Pegoraro; Sulis, 2013) deals with including Phasor Measurement units in the BC estimator formulation, with a hybrid perspective, including PMUs directly into the measurement vector model for the state estimator. The work tackles both estimation procedures based on polar and rectangular coordinates, exploiting the fast computation characteristic of the BC estimator under radial and weakly meshed distribution networks. The authors present a classical result on incorporating phasor measurements in the estimation process that their increased accuracy on measuring complementary information enhances the accuracy of the estimation process. As practice in hybrid state estimation, the authors neglect different sampling rates and lack of synchronism among measurements.

The work in (Fantin, 2016), that tackled such efforts to include PMUs in the WLS estimator for distribution systems, using a hybrid state estimator as well. Practical considerations

of three-phase unbalanced models are incorporated in the model, and the work brings an observability analysis to support three-phase systems with PMUs. Despite the effort in representing practical aspects, the work does not deal with scalability and ill-conditioning of distribution system state estimation. Besides, temporal aspects are also neglected in this work.

Another exciting work that employs the WLS estimator assigns larger measurement variances to compensate for the lack of synchronism of SCADA measurements is presented in (Alimardani et al., 2015). The work introduces additional variance to measured values to accommodate possible load variation and additional bias on the measurement model. The additional bias is based on the mean value of load variation in past estimations, which is not representative in case of sudden load changes or generation intermittency, far from the mean load variation. Besides, the additional variance reduces the estimation model's statistical consistency, which may hamper, for instance, bad data processing and artificially enlarges the estimator variance in stationary conditions. A similiar approach is also proposed in (LIU et al., 2012) specifically for smart-meters. Also, such works do not address different sampling rates, neither a diverse set of sources of information as expected in distribution systems.

In another perspective, the work in (Gómez-Expósito; Gómez-Quiles; Džafić, 2015b) employs an interpolation approach to deal with two time scales in distribution system state estimation, the first scale from SCADA and the second one from pseudo measurements. The work interpolates pseudo measurements value to increase its sampling rate, matching the updating time of SCADA measurements. A WLS state estimator is employed considering updates in the latest measured SCADA values and the set of interpolated pseudo measurements. Such application can reduce bias of estimation if the system follows the expected load curve, but it is susceptible to reduced pseudo measurement quality and may lose accuracy on sudden load changes. Besides, the paper does not address any faster sampling rates such as PMUs or smart meter in the networks, and it lacks a three-phase model for the application, reducing its practical perspective.

A different approach in (Ni et al., 2018) deals with aggregating smart meter data into the estimation. This aggregation is a common practice to surpass real time-communication issues for a large number of smart meters, thus only dealing with aggregated values at the medium voltage transformers. It introduces a pre-processing stage to compensate for unsynchronized smart meter data during the aggregation. The work presents an increase in the accuracy of a WLS estimator. However, it does not address the different temporal scales between SCADA and the aggregated smart meters, employing a hybrid approach in its essence.

Although these hybrid methods exhibited good results in theory, they face practical issues since they neglect the temporal aspect of the estimation process, mixing measurements gathered at different instants in a single set (Simões Costa; Albuquerque; Bez, 2013; Zhao et al., 2016). The problem arises from the fact that essentially each different measurement set, for instance, SCADA and PMUs, are sampling electrical quantities at different instants. This intrinsically adds

systematic error to the filtering process, from the fact that each type of measurements provides information about different instants of the grid condition. It is noteworthy that this approach is the most common in the literature and applications, and few authors deal with practical issues of different sampling rates so far.

Such limitation encouraged distribution system state estimation that could model the temporal aspect of the measurements and state. The first formulations of dynamic state estimation in transmission systems were proposed shortly after the consolidation of static state estimation in transmission system operation centres (Debs; Larson, 1970; Falcao; Cooke; Brameller, 1982; Bretas, 1989; Zhao et al., 2019). However, the computational burden of the algorithms, and the lack of fast updating measurements in the SCADA system, limited the practical applications of these first attempts in realistic power systems (Rousseaux; Van Cutsem; Dy Liacco, 1990). Only in the late 2000s, dynamic state estimators regained significant interest, with the advent of faster computer architectures, by the allocation of new instrumentation and communication technologies with faster updating and sampling rates, and finally, due to the acute power system operational requirements carried by a more volatile environment with renewable energy resources (Zhao et al., 2019; Zhao et al., 2020).

Regarding distribution networks, the Forecasting-Aided State Estimation (FASE), a type of dynamic state estimation method, established the ground base for modelling these temporal relations (Coutto Filho; Souza, 2009; Zhao et al., 2019). FASE is a recursive estimation method based on several measurement snapshots in a time sequence in the form of state-space equations. The standard dynamic state estimation approach of the FASE procedure is based on the Kalman Filter algorithm for state estimation, widely explored in many research efforts, such as in (Huang; Lu; Lo, 2015; Sarri et al., 2016; Carquex; Rosenberg; Bhattacharya, 2018; Louis et al., 2020), especially in the context of monitoring active distribution systems. Few of them, however, deal with practical aspects from the diversity of sampling rates and its effects on the estimation process, primarily relying only on a single type of measurements, such as the Unscented Kalman Filter in (Nguyen et al., 2013) with advanced SCADA measurements, or with linear formulations relying solely on PMUs, with the discrete Kalman Filter in (Sarri et al., 2016; Kettner; Paolone, 2017).

The work in (Huang; Lu; Lo, 2015) explores the synergy between advanced metering infrastructure and SCADA measurements with a FASE procedure. It employs previous day information to build the state transition matrices and load forecasting model, while the arrival of new measurement updates the state variables from SCADA and smart meter data every 15 minutes. The Iterated Extended Kalman Filter and the Unscented Kalman Filter are compared in a quasi dynamic framework. The results show fundamental tracking capabilities for real-time operation of both low voltage and medium voltage systems, increasing accuracy when compared with the standard static approach. However, the method neglects full three-phase coupled models, and the results also show a large variability of the load values that are not properly captured by

the method, which may be related to the smoothing properties of the Kalman Filter approach.

In a different perspective, (Carquex; Rosenberg; Bhattacharya, 2018) propose an Ensemble Kalman Filter to mix PMU information with pseudo measurement information. The method does not compute a fully model-based procedure to perform the estimation. Instead, an ensemble is performed with empirical covariance matrices and a load flow calculation. The Ensemble consists of assimilating pseudo measurements' information and PMUs independently, with the power flow calculation results. A trade-off is made to reduce the complexity of measurement equations modelling into the estimation, using conventional power flow tools, which may hamper accuracy. The authors show an interesting aspect of introducing faster PMU samples, increasing the performance gain on the estimation. The method, however, does not include other sources of information in the estimation process. Besides, from a practical perspective, the work neglects three-phase network models and measurements information in different phases, improving the Ensemble, and the computational timescale assumed is far from real-time applications.

A drawback from the Kalman Filter approaches is they rely on the accuracy of the transition matrix in the forecasting stage (Zhao et al., 2019), which is largely affected by the occurrence of system abrupt changes, such as load switching, contingencies, trends and intermittences (Zhao et al., 2019; Massignan; London Jr; Miranda, 2020). Such a problem is aggravated, especially when dealing with phasor measurement units that can capture several types of transient responses of the system, including different sorts of load changes (sudden or trends). Besides, proper tuning of the Kalman Filter approaches is difficult to achieve, and the association of these methods with different anomaly detection methods also reduce their generalizations.

In a similar perspective, data fusion has been proposed for integrating fast sampling measurements. It consists of a two-step method for merging the results of independent estimators, based on the statistical error propagation theory (Simões Costa; Albuquerque; Bez, 2013). Different works propose new objective functions towards dealing with the lack of synchronism among measurements, such as the weighted least absolute value (WLAV) (Gol; Abur, 2015), the generalized maximum likelihood estimators (GMLE) in (Zhao; Mili, 2018a) and also information theory concepts in (Freitas; Costa; Miranda, 2017). Such methods are also promising to deal with non-stationary changes in the system, as in the case of sudden load changes, since the fusion method works with a non-causal state-space, that is, dismisses the transition matrix to model temporal relations, focusing on a model to mix the different sources of data, as presented in (Massignan et al., 2019). So far, such methods were applied in transmission systems with a certain redundancy, far from the reality of distribution systems. Besides, none of them tackled larger sampling differences than between SCADA and PMUs.

In this context, this thesis explores the following novel perspectives to aid the distribution system real-time monitoring tools:

- Incorporation of multiple sources of information from pseudo measurements, smart meters, SCADA measurements and Phasor measurement units, while dealing with the lack of synchronism and their different sampling rates, through formulations that adequately capture temporal relations on the state estimation process;
- tracking distribution system changes captured by fast samples on the state estimation process to aid operators with more knowledge about their grid condition.

2.3 Scalability of Distribution System State Estimation

Besides the challenge of dealing with multiple time scales and sampling rates from diverse sources of information in distribution systems, another critical practical challenge is the large scale of distribution networks, which may hamper real-time processing requirements. Distribution systems are networks spread into vast areas, usually with hundreds of different primary feeders that correspond to thousands of three-phase unbalanced buses. The inclusion of new sensors at the level of the low voltage circuits further requires an extension of the network models until the consumer units, which may lead to hundreds of thousands of variables. In this sense, a large scale comprises both the number of nodes in each feeder, as typically dealt in the standard IEEE benchmarks and numerous feeders, from various substations and may share interconnections.

In this context, two characteristics are highlighted as requirements for distribution systems state estimation: computational performance to deal with the complexity of distribution systems; and maintaining accuracy while dealing with the scalability of practical distribution networks. In conventional state estimators, the measured information is processed in a centralized manner. Given the dimension and natural subdivision of distribution systems into different feeders, the multiarea state estimation approaches are an exciting solution for the scalability problem (Primadianto; Lu, 2017; Ahmad et al., 2018). Right after the conception of state estimation in the 1970s started the idea to operate the transmission systems with a decomposition into areas and in a decentralized manner (Cutsem; Ribbens-Pavella, 1983). The primary motivation was the presence of different operators for each region of the power systems and improved computational efficiency of the estimators. This way, along the years, several estimators were proposed using the concept of decomposition of the power systems into areas for the transmission systems (Gómez-Exposito et al., 2011b). Although this topic has been brought by the leading state-of-the-art and review research papers (Primadianto; Lu, 2017; Ahmad et al., 2018), it is a consensus among the different authors that few works have effectively dealt with such concepts in their implementations (Muscas et al., 2015; Pau et al., 2017).

In (Garcia; Grenard, 2011), a multiarea distribution system state estimation approach is presented from a scalability perspective. The work seeks for a methodology to deal with the large scale of distribution networks. The conventional state estimation is explored under a computational scalability perspective. This is one of the main motivations to seek multiarea state

estimators for large scale distribution networks. The authors also show enhancements on the numerical conditioning of the problem by performing local estimations in smaller networks. Besides the obtained computational gain, the authors dismiss important practical scalability issues, such as the need for three-phase models and dismiss the secondary low voltage networks, which may not sustain the method's real-time feasibility. Another essential and common requirement is that the method requires observability for each local area.

In another architecture perspective, in (Nusrat; Irving; Taylor, 2011), a distributed computing multiarea approach is applied for distribution networks. Besides the computational and numerical advantages already pointed, this approach follows a trend on the decentralized operation of distribution networks. However, the method requires many measurements on the distribution grid and does not consider the three-phase unbalanced nature of distribution systems, a common limitation in these first propositions of multiarea state estimators for distribution systems. The proposed solution method is based on a heuristic differential evolutive algorithm, which increases information exchange among processors and the iterations in the estimation process, usually based on fast convergence with theoretical convergence proof such as the Newton method.

In (Nusrat et al., 2015), the authors continue their work but now exploring the solution Newton method with a more conventional solution approach. The method also enhances the decomposition of the distribution network by using an overlapping zone concept in the distributed estimation process, extending the areas to share one level of state variables of adjacent areas. However, still without dealing with the three-phase unbalanced nature of distribution systems, and with an iteration-based exchange of boundary variables that increases the exchange of data and the number of iterations for convergence.

The work in (Muscas et al., 2015), which started a solid track of researches in the topic by the authors, introduces the use of specialized distribution system state estimation to deal with the unbalanced nature of the networks in the local estimation stage, the BC estimator. Due to its computational efficiency to deal with primary medium voltage feeders, the method enables fast processing of the local estimation step. This research also incorporates correlations that may arise between boundary nodes in adjacent areas. Among the limitations, the scalability of the method has not been addressed, focusing the applications only in primary feeders and dismissing the secondary low voltage networks. Besides, the proposition requires that each area is observable and does not tackle with decentralized operation perspectives.

In a similar perspective, in (Angioni et al., 2016b) another approach with the BC estimator is presented. The work considers automation equipment and communication protocols in a practical implementation tested in a power systems real-time digital simulator. Different characteristics are discussed for the integration at SCADA level and state estimator processing, especially latency and data exchange. Despite the exciting results, the authors do not tackle the significant scalability problem, employing a reduced test scenario, which hampers its practically

sized networks.

In (Pau et al., 2017), another effort from the same research group as in (Muscas et al., 2015), the authors enable sharing measurement data among adjacent areas in the coordination step of the hierarchical multiarea state estimator, enhancing the achievable accuracy in a two-step method. Authors also reduce communication requirements among adjacent areas by simply exchanging the estimated voltages at the boundary nodes. Nonetheless, the method disregards the low voltage systems, limiting the scalability conclusions for real-time operation. An interesting aspect of such a method is that there is an accuracy gap between the multiarea and the centralized algorithm, showing that the coordination approach may still be further enhanced.

In a different perspective than the hierarchical estimators, a decentralized multilevel approach is proposed in (Pau et al., 2019) based on cloud computing and IoT integration to deal with the different parts of the distribution system, the medium and low voltage systems together. The distributed approach consists of exchanging the results of low voltage, where coordinators are installed at the substations, to the upstream medium-voltage feeders. The algorithm is proposed under a cloud computing framework, making it attractive for integrating aggregators and decentralized local markets. The limitation of such an approach is that it requires a physical structure of data concentrators to be installed in all medium voltage/low voltage transformers, and latency on cloud access was not considered, which may hamper real-time integration with operation centres. Nonetheless, the algorithms employed are very scalable and present high computational performance.

Amidst such interesting researches that captures the scalability of the state estimation process, two open challenges are addressed in this thesis:

- Improving multiarea distribution system state estimation to deal with the area decomposition and local observability issues, encompassing multiple primary feeders and various low voltage networks in single implementation;

- Including practical characteristics of unbalanced and asymmetrical distribution networks, while solving the three-phase state estimation problem without loss of precision and approximations assumptions, even for meshed networks, for large scale networks comprising high voltage, medium voltage and low voltage systems.

2.4 Probabilistic Characterization for State Estimation

The third research topic brought in the spotlight by this thesis consists of dealing with non-Gaussian probabilistic models for the measurement noise characteristic. This context often emerges in power systems state estimation in the context of robust state estimation since gross errors can be interpreted as a particular case o non-Gaussian heavy-tailed probability distributions for the measurement noise. Regarding distribution systems, the predominance of

pseudo measurements also favours the exploration of such models since they are often based on forecasts with a reduced quality compared to real-time measurements, introducing additional uncertainty to the process.

This probabilistic context characterization of load models have been explored under different perspectives, employing Beta, Gamma, and Weibull distribution (Ghosh; Lubkeman; Jones, 1997), nonparametric regression based on artificial neural networks (Manitsas et al., 2012) and also by employing flexible Gaussian Mixture Models (Singh; Pal; Jabr, 2010). Such a method essentially seeks for improving the fitness of load models with historical and measured data, to be further employed as pseudo measurements in distribution system state estimation. However, they do not encompass such hypothesis of non-Gaussian models in the state estimation theoretical framework and solution. For instance, the work in (Ghosh; Lubkeman; Jones, 1997) accommodates only the expected values and variances of different probabilities in a probabilistic load flow method as the state estimation procedure. In (Singh; Pal; Jabr, 2010; Manitsas et al., 2012), a Gaussian mixture and artificial neural network characterization are employed for the pseudo measurement models. However, during the state estimation procedure, only a marginal Gaussian distribution is considered, that is, a single element of the mixture that is closer to the expected value at a certain instant. Thus, despite exploring more robust and flexible characterizations of the pseudo-measurement values, such methods do not include them or enhance the state estimation and formulation.

In this sense, robust estimation appears as an appealing topic to deal with non-Gaussian noise. For instance, formulations based on the weighted least absolute value (Celik; Abur, 1992), using non-quadratic loss functions (Baldick et al., 1997) and by the projection statistics method (Mili et al., 1996). Such methods often aim at increasing robustness against outliers and have been recently considered to support distribution system state estimation (Zhao et al., 2020). Similarly to the propositions for transmission systems, by focusing on outliers, essentially, these methods do not change the primary hypothesis of the noise characteristics and practically enhance the estimation accuracy under heavy-tailed probability distributions.

In this sense, a kernel density estimation formulation presented in (Miranda; Santos; Pereira, 2009) provides a complementary perspective for the state estimation problem. By introducing a kernel model for the measurement noise characteristic is possible to increase generalization for other distributions. The method employs a non-quadratic formulation to support the estimator objective function in a more suitable form to deal with the kernel models, based on the Correntropy concept from Information Theory. Such concept has been explored under the power system state estimation problem in (Freitas; Costa; Miranda, 2017; Freitas; Simões Costa; Miranda, 2020), exploiting the Parzen Window concept to enhance robustness against outlier and from an algorithmic perspective to integrate the bad data processing inside the static state estimation process flow. More recently, important results have shown favourable properties of Generalized Correntropy associated with the Interior Point Method to deal with

challenging situations of bad data processing (Pesteh et al., 2019; Pesteh; Moayyed; Miranda, 2020). However, the kernel density estimation perspective has been focused on applications for transmission systems and the goal of increasing robustness against outliers.

One of the few works that address such a challenge of non-Gaussian models in the distribution system is the work in (Pegoraro et al., 2017). It explores the Bayesian Inference background in order to provide estimates under non-Gaussian error distributions. A Monte-Carlo-Markov-Chain (MCMC) method is employed along with the Metropolis-Hastings sampling algorithm to perform posterior inference with non-Gaussian measurement noise. Despite a promising characteristic for generalizing the statistical description of the noise characteristic, the work demands a high number of samples to converge, and the authors do not provide sensitivity for different initial samples. This aspect results in a large computational time for small test cases with four nodes, which would be intractable for real-time applications in larger systems.

This leads to the main theoretical background employed by this thesis, the exploration of the Bayesian Inference framework applied in distribution system state estimation. Despite a consolidated topic in statistics, few works explore the Bayesian framework in power system state estimation (Lourenco; Costa; Clements, 2004; Pegoraro et al., 2017; Mestav; Luengo-Rozas; Tong, 2019). The Bayesian perspective has been brought in the first propositions of state estimation already in the 1970s (Schweppe; Wildes, 1970; Schweppe; Handschin, 1974). Despite its strong probability theory background, it never got the attention of future research due to additional challenges related to the tractability of required computations and model complexity, as occurred in other different scientific areas. More recently, such methods have reignited interest in many scientific areas, ranging from biostatistics, machine learning, astrophysics, and the pursuit for causal models (Congdon, 2007; Sivia; Skilling, 2006; Puga; Krzywinski; Altman, 2015).

In the context of distribution system real-time monitoring, the interesting work in (Singh et al., 2011) brings the Bayesian Inference framework to deal with the detection of topology changes inside substations and distribution networks. The perspective of Bayesian Inference is employed as model verification to track the most probable topology of the network, based on previously stored possibilities and the results of a WLS estimator. High accuracy of topology identification is obtained by successive iterations of the method upon convergence to the most probable network. However, the authors do not address the response of the method against different transition that may occur on the system, which may mislead the interpretation of a topology change with different levels of load variations.

Regarding the process itself of distribution system state estimation, a recent work brings a Deep Learning based approach, supported by Bayesian inference as an estimator that seeks the expectancy of then state variables given the measurements (Mestav; Luengo-Rozas; Tong, 2019). A regression learning approach based on deep neural networks provides a data-driven net power injections model based on a historical training set. By associating a Monte Carlo

method with a Newton Raphson power flow, a second model for state estimation can be learned and associated with the first one. Estimation consists of computing the expectancies of the learned models according to measured data in real-time. Also, bad data processing is performed as a pre-processing stage, sampling the learned models with the prior knowledge for the state variables, and then associating conditional probabilities for the measurements and performing a hypothesis test. Despite the method provides interesting results, it address a high computational burden, a strong influence of parameter tuning and a significant influence of prior knowledge on the Bayesian approach.

A different and recent work in (Dobbe et al., 2020) presents a Bayesian approach for linear three-phase state estimation that works together with a forecasting stage based on the Gaussian process. A linearised load flow method evaluates historical active and reactive power data to forecast possible voltage levels in a distribution system. A linear least-squares estimator then updates the voltage phasors according to few PMUs measurements installed in the network. The Bayesian framework is employed as an innovation method to update previous forecasted values but restricted to PMU measurements. Despite a good integration of different levels of information, linearized power flow models may reduce the accuracy.

In a different perspective, (Zhou et al., 2020a) explores the harmonic state estimation problem with sparse Bayesian learning from Smart Meter and PMU data. The method is based on load flow calculation results aided by a regression analysis to incorporate PMU information in the estimation process, focusing on a frequency-based distribution system model. It employs interpolated smart meter data to accommodate the faster samples of PMUs, which may hamper accuracy under sudden changes. The Bayesian inference background is employed basically as regression analysis, without further investigation of its conceptual characteristics.

Finally, such direction points to two research gaps that will be addressed in this thesis:

- Dealing with non-Gaussian noise characteristics in distribution system state estimation while maintaining computational tractability for real-time applications, not only related to pseudo measurements and gross errors but also able to deal with different uncertainties from smart grid scenarios, such as intermittency and more considerable load variability from sudden changes;

- Exploring the conceptual gap on Bayesian Inference framework for power system state estimation, providing additional support and new concepts from a probabilistic perspective.

2.5 Concluding Remarks

Although the essential and recent research development, there is no consensus about which methodology is the most adequate for distribution system state estimation. While the first propositions aims to enable the numerical solution of the state estimation problem and

improve the quality of the estimation, the recent challenges are brought by emerging smart grid technologies. Besides, many practical aspects of real-life distribution networks increase the complexity for a single methodology to emerge as the most adequate.

In this context, the literature review shows there are many opportunities related to distribution system state estimation yet to be explored, both at the practical reality of nowadays passive distribution networks and the upcoming future of the smart grids. In this sense, this thesis explores the following literature gaps primarily:

- Incorporation of different sources of information from pseudo measurements, smart meters, SCADA measurements and Phasor measurement units, while dealing with the lack of synchronism and different sampling rates.
- Improving multiarea distribution system state estimation to deal with multiple primary feeders and various low voltage networks in a single implementation, dealing with practical characteristics of unbalanced and asymmetrical ill-conditioned distribution networks;
- Dealing with non-Gaussian noise characteristics in distribution system state estimation while maintaining computational tractability for real-time applications;
- Including practical characteristics of unbalanced and asymmetrical distribution networks, solving the three-phase state estimation problem without loss of accuracy and approximations assumptions, even in meshed networks, for large scale networks;
- The conceptual gap on Bayesian Inference framework and tools for power system state estimation.

It is noteworthy that this thesis's main effort and contribution rely on the proposition of a computational tool able to deal with practical challenges of distribution systems, fully representing three-phase component models, and with a high-resolution network model comprising the substations, primary feeders and low voltage networks. The Bayesian Inference framework provides the theoretical support for different applications related to the above literature gaps. Additional challenges are also addressed, such as numerical stability, computational efficiency and detailed modelling of a diverse set of distribution system components.

3 BAYESIAN INFERENCE PERSPECTIVE FOR POWER SYSTEM STATE ESTIMATION

This chapter presents the main theoretical concepts applied in this Thesis, providing a new perspective for power system state estimation under Bayesian reasoning. The objective is to provide an overview of the Probability Theory and Bayesian Inference concepts and how they layer over the power system state estimation. Three novel applications related to power systems state estimation are devised from the Bayesian Inference framework in this Thesis, exploring new features from the perspective of information fusion, spatial fusion and dealing with non-Gaussian noise characteristics. All three concepts are further explored in the subsequent chapters corresponding to the practical application in distribution systems.

3.1 Probability Theory and Bayesian Inference Concepts

The objective of probabilistic methods is to capture the effects of randomness in mathematical models that may represent a particular phenomenon or system or to evaluate more abstract concepts. Typically, the purpose is to predict future outcomes, perform regression of unknown variables, explore stochastic behaviour, and quantify uncertainty. Conversely, a deterministic approach assesses an exact solution without uncertainty on the models and variables. This latter brings causal relations as the sole source of information to explore a particular system or phenomenon, which typically implies assuming immutable boundary conditions and a closed system, far from the reality of many engineered systems subject to all sorts of uncertainties.

There is a philosophical division between two major approaches for the concept of randomness and probabilities in their fundamental essence in modern probabilistic theory. Such division follows the two fundamental principles to interpret and represent randomness, the *Frequentist* and the *Bayesian*. Both approaches have their particular perspectives and complementary concepts and should not be interpreted as opposite sides, one being right and the other the wrong one, but simply as different approaches to the same phenomenon of randomness characterized by probabilities (Bayarri; Berger, 2004; Puga; Krzywinski; Altman, 2015).

The *frequentist* (sometimes refereed as objective) approach interprets probability as a measure of repetition of specific outcomes within the range of possible values, characterized by fixed but unknown parameters. Such interpretation yields the estimation problem as a search for fixed parameters of interest considering a set of observations and based on the relative frequency obtained by the repetition of several trials or samples. Regarding parametric models, inference seeks to find the fixed parameters encompassed in a likelihood function to capture the best goodness-of-fit that explain the underlying random effects observed.

The *Bayesian* (sometimes refereed as subjective) approach represents a degree of belief, assigning probabilities for both hypothesis and data. The parameters are assigned as random

variables in the Bayesian perspective, along with a degree of uncertainty and a probability distribution. The inference is performed as a descriptive feature from a posterior distribution, drawn from both likelihood model obtained from the data and a prior distribution for the parameter, that may incorporate previous or expert knowledge for the random parameters.

Besides such philosophical arguments, the Probability Theory provides a common mathematical background for both *bayesian* and *frequentist* approaches by formalizing random variables, and the fundamental theoretical statistics quantify randomness and uncertainty. A brief description of the most important concepts are described in the sequence. The concept of a random variable is defined under a sample space Ω that represent all possible outcomes of a particular random phenomenon $\omega \in \Omega$. In the strict probabilistic sense, a random variable X can be defined as a function, continuous or discrete, that relates outcomes from a sample space to a real number, $X : \Omega \to \mathbb{R}$. A measured outcome from such sample space is called an event E, defined as a subset of the sample space. Finally, the probability is defined as a real-valued function associated to events of the sample space $P : E \to \mathbb{R}$, that attends the Kolmogorov axioms (non-negativity, unit measure and additivity on σ-algebra). Together, the sample space, the events and the probability define a probability space (Ω, E, P), where the random variables are characterized (Blitzstein; Hwang, 2019; Ash, 2008).

Different descriptive features are associated with a random variable in order to characterize it. The first, and most important concept, is the distribution function (also referred to as cumulative distribution function). It consists of a function defined on the sample space that has the following properties:

Theorem: A function $F : \mathbb{R} \to [0,1]$ is a distribution function of some random variable if and only if:

1. F is non-decreasing
2. $lim_{x \to -\infty} F(x) = 0$ and $lim_{x \to \infty} F(x) = 1$
3. F is right-continuous, i.e. $lim_{y \downarrow x} F(y) = F(x)$

The distribution function may be parametric, where an analytically tractable mathematical function fulfils the above properties, or non-parametric, defined as a probability metric without a closed-form function. The probability density function $f(x)$ derives from the above definition (in the case of discrete variables is also called probability mass function) as the following:

$$F(x) = \int_{\infty}^{x} f(t)\, dt \tag{3.1}$$

Another important notion regards conditional probability, which can be interpreted by the relation between different random variables in the situation whenever a hypothesis is given

about one of them, and the other is treated as a variable. It encompasses their joint probability function $f(x,z)$ and the marginal density of the evidence $f(z)$.

$$f(x/z) = \frac{f(x,z)}{f(z)} \tag{3.2}$$

In addition, essential features of the random variables and the associated probability distribution functions are employed to characterize them, such as the expected value, covariance, and the moment generating function. Finally, two crucial and classic theorems arise from Probability Theory and from the definitions of conditional probability and the sample space:

Theorem (Law of total Probability): This theorem relates marginal distributions to conditional probabilities, while the marginal distribution can be represented by a set of disjoint events and their conditional probabilities.

$$f(z) = \int f(x,z)dx = \int f_{Z|X}(z|x)f(x)dx \tag{3.3}$$

Theorem (Bayes's Theorem): This theorem relates the *a posteriori* density of a random variable with observed evidence and *a priori* knowledge ($\pi(x)$).

$$f_{X|Z}(x|z) = \frac{f_{Z|X}(z|x)\pi(x)}{\int f_{Z|X}(z|x)\pi(x)dx} = \frac{f_{Z|X}(z|x)\pi(x)}{f(z)} \propto f_{Z|X}(z|x)\pi(x) \tag{3.4}$$

The *likelihood function* $f_{Z|X}(z|x)$ captures the uncertainty associated with the observations (samples or measurements) for a given set of unknown parameters. It encompasses a notion of goodness-of-fit between observed data and a model related to the unknown values (or parameters). In essence, it encompasses a hypothesis that is being tested with given observations, and the variables of interest are part of such hypothesis and not of the random outcomes inside the likelihood function.

The *prior distribution* $\pi(x)$ represents the probability distribution function associated with the unknown variables (or parameters), now modelled as random variables instead of simply assuming them as a fixed but unknown part of the hypothesis. In essence, it captures uncertainty regarding the parameters of interest encompassed as random variables. In the Bayesian sense, it can be constructed in the form of a non-informative prior whenever any information regarding the parameters is known, and the likelihood (data) will provide such characterization. Examples of such non-informative priors are improper constant priors, Jeffrey priors, or distributions with very high variances. Conversely, the priors may also bring additional information based on previous or expert knowledge or even a systematic description of the uncertainty associated with the parameters of interest.

The *evidence* $f(z)$ consists on the marginal probability of occurrence of the sampled observations. In the context of Bayesian Inference is often omitted since it consists of a constant

value that scales the product between the likelihood and prior, so the posterior attends the properties of a distribution function.

Finally, the *posterior distribution* $f_{X|Z}(x|z)$ is a conditional distribution function for the unknown values updated by the gathered data and considering the embedded prior knowledge. Note that if any data is sampled, the prior knowledge keeps unchanged. Conversely, if any prior knowledge is considered, the sample size will be a dominant factor in the hypothesis's confidence. This aspect draws an essential characteristic of the Bayesian reasoning, the ability to increase confidence based on prior knowledge, a unique feature whenever sample sizes are small and complementary information, with associated uncertainty, helps the inference problem. Note that if x and z are independent, the posterior equal to the prior since the data does not provide additional evidence to support or dismiss the hypothesis. This notion induces Bayesian Inference as a powerful tool to evaluate causal inference.

Such theorems derive the fundamental concepts behind the Bayesian Inference framework, which has been a very active research topic on many scientific applications (Congdon, 2007; Sivia; Skilling, 2006). Such a broad framework encompasses different features to perform inference according to derivations from the above basic principles.

Some examples of such features are: the use of hyper-priors to extend random variables components to multiple parameters of prior distributions; the application of numerical integration method such as the Monte Carlo Markov Chain method in order to evaluate analytically intractable probabilistic models; the exploration of Bayesian networks to map causal relations in models, by representing conditional relations among variables based on observed data; the application of Bayesian hypothesis tests in the form of credibility intervals, Bayes factors or model assessment through Bayesian Information criterion; learning spatial correlations trough autoregressive models; the evaluation of predictive posterior distributions for the data, as a way to perform forecasting or artificially sampling new observations to complement the characterization of the previously gathered data; and also the assessment of risk functions in the decision theory perspective; among others. More information on different aspects of Bayesian Inference can be referred to in (Congdon, 2007; Sivia; Skilling, 2006; Rencher, 2003; Bayarri; Berger, 2004; Albert, 2009; Puga; Krzywinski; Altman, 2015; Makowski; S. Ben-Shachar; Lüdecke, 2019).

3.2 Perspective for Power System State Estimation

The problem of power system state estimation emerges as a bridge between measured data from sensors installed across the electric power systems and a network model that captures the relations among measured electrical quantities and the network state. It is traditionally formulated as static analysis, intending to obtain steady-state features of the network (Monticelli, 1999; Abur; Gómez-Expósito, 2004; Bretas N. Bretas,). It has a strong relation with power flow analysis, another vital tool to assess grid condition under known loading and generation scenarios. Although both problems frequently share similar network models, they have some

conceptual differences. The main one regards the input data. In the state estimation problem, the steady-state is evaluated from measured values, assumed with some degree of uncertainty, and able to deal with redundancy. In power flow analysis, a specified loading/generation condition is evaluated, without any redundancy, a determined problem. The state estimation process supports real-time evaluations if the system operates in one of the following conditions: *(1) Normal and secure state* ; *(2) Alert state*; or *(3) Emergency state* (Monticelli, 1999).

The static state estimation problem for a power system with m measurements and n state variables is usually formulated using the following measurement model (Abur; Gómez-Expósito, 2004) to obtain the steady-state condition of the network (system state):

$$z = h(x) + e \tag{3.5}$$

where z is the $(m \text{ x } 1)$ measurement vector; x is the $(n \text{ x } 1)$ vector of state variables (usually the complex nodal voltages); $h(x)$ is the $(m \text{ x } 1)$ set of nonlinear equations that relates the measurements with the state variables; and e is the $(m \text{ x } 1)$ noise vector, typically assumed as independent normally distributed random variables, with zero mean and known covariance matrix R.

Although the static perspective is the traditional state estimation formulation, the increased interest in dynamic state estimation models has reignited in recent years. The authors in (Zhao et al., 2019) encompass a formal difference among types of dynamic state estimators, primarily related to the definition of the state vector, and are briefly discussed in Appendix A. The fundamental change is to encompass temporal relations among the state variables, represented in a first-order state-space model, by the following:

$$\begin{aligned} x_t &= f(x_{t-1}) + \omega \\ z_t &= h(x_t) + e \end{aligned}$$

where, the first equation is known as the process model, with x_{t-1} the state vector in a previous instant $(t-1)$, $f(x_{t-1})$ is the process equation that relates temporal transitions on the state variables, ω is the process noise, typically assumed as normally distributed random variables, with zero mean and known covariance matrix Q, and the second equation is the measurement model presented previously.

Dynamic state estimation extends the state estimation problem under a state-space by introducing temporal relations among state variables and incorporating system dynamics into the estimation problem. This introduction of the relation between two different instants adds a random component to the state. From the Bayesian perspective, it consists of introducing a probabilistic model for the state vector, that is, modelling the state of the network as a random variable. It changes the traditional overview of the estimation process as a "snapshot" of the network condition towards a probabilistic characterization of the state. It encompasses the fact that different state values can occur between updates of the slower measurements, observed by

the faster ones (Massignan et al., 2019; Zhao et al., 2019). This thesis extends the temporal abstraction of the first-order state-space model to accommodate a diverse set of conditional relations for the state variables under a Bayesian Inference framework.

Such perspective induces a model where the state variables are interpreted as random variables in the Bayesian Inference perspective and not as a fixed and determined vector. It does not imply a lack of determinism in an electrical power system, much less in a lack of causality in such an electrical network. Instead, it brings the estimation problem to seek a state vector among a set of possible values closer to observed values from measurements and expert knowledge about the system behaviour. The reasoning comes from the fact that it is not possible to know the actual state of the power system in practical conditions, but only a marginal overview captured by the measurements.

This way, equation (3.4) is repeated here in the context of power system state estimaiton. A prior distribution $\pi(x)$ is assumed for the state variables, and the state estimation is defined as an inference problem regarding the posterior distribution obtained by Bayes's Theorem:

$$f_{X|Z}(x \mid z) = \frac{f_{Z|X}(z \mid x)\pi(x)}{f_Z(z)} \propto f_{Z|X}(z \mid x)\pi(x) \tag{3.6}$$

where, $f_{Z|X}(z \mid x)$ is the likelihood function obtained from the measurement model, $\pi(x)$ is the prior distribution for the state variables, $f_Z(z)$ is the evidence of the model, often omitted since it is a constant and scaling value.

Note that not necessarily sequential temporal relations are captured by the posterior in the proposed model, as performed in the dynamic state estimation approach. The prior can be constructed according to different perspectives, including the previous instant as done in the dynamic state estimation. Nevertheless, it also encompasses the possibility to enhance estimation with different sources of information. In this sense, this prior knowledge can be informative or non-informative priors designated in the Bayesian Inference framework (Sivia; Skilling, 2006; Congdon, 2007). The first disregards any knowledge about the state and system behaviour, while the second includes additional prior knowledge (that may come from historical context or expert knowledge). Also, it represents a causal inference model, since in power system state estimation, the observed values are electrical quantities physically related to the state variables, typically the complex nodal voltages in all nodes of the electrical network.

Estimation consists of performing inference in the posterior distribution, rather than only retracting a state vector from the measurement model and likelihood function, as done in the traditional state estimation sense. Two main perspectives may provide the estimation results: finding the full posterior distribution or finding pivotal quantities of such distribution. The first one, despite exact in the probabilistic sense, might be intractable if nonlinear and non-parametric models are considered and are only suitable for well-behaved models, which often imply simplifications of real-world applications. Obtaining a non-parametric posterior also requires special numerical integration methods that may be burdensome for real-time applications.

The computational complexity and sampling spaces increases regarding the models' scale, which reduces the applicability of sampling methods like the Monte Carlo Markov Chain method.

The second approach yields the estimation of quantities of interest from the posterior distribution that, despite not providing a complete description of uncertainty, enables the extraction of more tangible features in the sense of understanding the system behaviour and quantifying uncertainty. This leads to the perspective explored in this Thesis, performing inference into a pivotal quantity of the posterior and not the posterior distribution itself. Even though a more accurate inference is performed with the full posterior distribution, as an engineering problem, the development of state estimators for power systems must keep the real-time requirement in mind.

In the proposed perspective, based on the Bayesian Inference framework, the estimation process consists of updating the state variables values by a Maximum a Posteriori (MAP) estimation given by (3.7) (Congdon, 2007).

$$\hat{x} = \arg\max_{x} f_{X|Z}(x \mid z) \rightarrow \arg\max_{x} f_{Z|X}(z \mid x)\pi(x) \tag{3.7}$$

MAP estimation, in essence, provides the value that yields the maximum value posterior probability density function, that is, the mode of the posterior distribution. According to the assumptions made regarding the prior and likelihood functions presented in the following, the fundamental aspect of MAP also enables a generalization of different state estimation methods employed in power systems, according to the assumptions made regarding the prior and likelihood functions presented in the following. This is important since the underlying distribution functions assumed can often be omitted or neglected without fully understanding the consequences in the probabilistic sense. Furthermore, under a common theoretical background, the interpretation of such different estimators enables an extension of concepts and new methods, providing additional flexibility to deal with practical challenges in power system real-time monitoring.

3.2.1 From Bayesian Inference to the Static State Estimation

The first theoretical result from this probabilistic perspective for the state estimation problem lies in clarifying its relation to the traditional Maximum Likelihood Estimation (MLE) approach employed in power system static state estimation. It is straightforward that the MLE is a particular case of MAP estimation whenever the prior distribution assumes a constant value, such as uniform distribution over an enlarged version of the domain of the state variables.

In essence, the Maximum Likelihood principle consists of the particular case of MAP estimation with an improper prior as constant c. In such a case, the prior and likelihood are given by the following expressions:

$$\pi(x) \sim c$$
$$f_{Z|X}(z \mid x) \propto \prod_{i=1}^{m} e^{-\frac{(z_i - h_i(x))^2}{2R_{ii}}} \tag{3.8}$$

In such a case, all values of $x \in \mathbb{R}^n$ are assumed to have an equal probability of occurrence. The above relation enables a further understanding of an underlying assumption from traditional state estimation formulation often omitted or sometimes neglected. The fact that the prior distribution resembles an improper uniform distribution implies absurd prior knowledge about the state, such as negative voltage magnitudes have equal probability of occurrence in such model, or that phase angles larger outside the range $[-\pi, \pi]$ are equivalent to the values inside this range. Such examples are extreme cases, but common information to power engineers about the system's typical behaviour in steady-state conditions is neglected. For instance, that is not probable that higher values than 1.2 p.u. or lesser than 0.8 p.u of voltage magnitude to be steady-state values on the network due to over-voltage and under-voltage protection settings.

This omission, however, does not tend to hamper the static power system state estimation in most cases, which will be fully guided by the likelihood function instead, but it also does not provide any additional information regarding the system. This subjective notion of the state variables as a random variable may not be clearly stated in traditional MLE but certainly reduces the possibility to accommodate complementary information about the power systems.

In this sense, traditional state estimation assumes a Gaussian characteristic for the measurement noise, which yields its likelihood function. The posterior distribution will be a Gaussian distribution in the linear measurement model, but not in the nonlinear case, and obtained by the following objective function:

$$\hat{x} = \arg\max_x f_{X|Z}(x_{SL}|z_{SL}) \propto \frac{1}{\sqrt{(2\pi)^m |R|}} \prod_{i=1}^{m} e^{-\frac{(z_i - h_i(x))^2}{2R_{ii}}} c \tag{3.9}$$

The above maximization yields the traditional Weighted Least Squares (WLS) algorithm, following the unconstrained minimization problem below. More details about this method and deduction are provided in Appendix A as a reference.

$$\hat{x} = \arg\min_x J(x) = \sum_{i=1}^{m} \frac{(z_i - h_i(x))^2}{\sigma_i^2} = [z - h(x)]^T R^{-1} [z - h(x)] \tag{3.10}$$

3.2.2 From Bayesian Inference to the Kalman Filter State Estimation

Besides the widely employed static state estimation perspective, dynamic state estimation provides an attractive increment of information about the state variables, as stated previously. The first-order state-space model incorporates a significant result that has a direct interpretation from a Bayesian Inference perspective.

In essence, the Kalman Filter approach consists of a particular case of MAP estimation with a Gaussian prior obtained by the previously estimated value, recursively (Barker; Brown; Martin, 1995; Massignan et al., 2019). The likelihood is also assumed as a Gaussian distribution, maintaining the assumption regarding the measurement model. The posterior distribution is also Gaussian for linear models, and the MAP estimation yields the standard Kalman Filter approach. The proof is obtained straightforwardly from the Gaussian-Conjugacy in the linear measurement model, which yields the following relation:

$$\begin{aligned} \pi(x) &\propto \prod_{j=1}^{n}\prod_{k=1}^{n} e^{-(x_j-\hat{x}_{t-1,k})P^{-1}_{t-1,jk}(x_k-\hat{x}_{t-1,j})} \\ f_{Z|X}(z|x)(x) &\propto \prod_{i=1}^{m} e^{-\frac{(z_i-H_{i:}x)^2}{R_{ii}}} \end{aligned} \tag{3.11}$$

where, H is the Jacobian matrix of the measurement model, and $H_{i:}$ the respective i-th row.

Since the posterior is also Gaussian, due to the Conjugate linear model, the mode equals the expected values, and the MAP estimate is given by:

$$\hat{x} = \arg\max_x f_{X|Z}(x \mid z) = \prod_{i=1}^{m} e^{-\frac{(z_i-H_{i:}x)^2}{R_{ii}}} \prod_{j=1}^{n}\prod_{k=1}^{n} e^{-(x_j-\hat{x}_{t-1,k})P^{-1}_{t-1,jk}(x_k-\hat{x}_{t-1,j})} \tag{3.12}$$

The above maximization yields the following unconstrained minimization problem:

$$\hat{x} = \arg\min_x J(x) = [z-Hx]^T R^{-1}[z-Hx] + [x-x_{t-1}]^T P^{-1}_{t-1}[x-x_{t-1}] \tag{3.13}$$

Applying the Woodbury formula for the sum of inverses and working algebraically with the expected value of the posterior yields the Kalman estimation step and Kalman Gain Matrix K_t as:

$$\begin{aligned} K_t &= P_{t|t-1}H^T\left(HP_{t|t-1}H^T + R\right)^{-1} \\ \hat{x} &= \hat{x}_{\,t|t-1} + K_t\left(z_t - H\hat{x}_{\,t|t-1}\right) P_t \;\; = (I-K_tH)\,P_{t|t-1}(I-K_tH)^T + K_tRK_t^T \end{aligned} \tag{3.14}$$

where K_t is the Kalman Filter Gain matrix, I is the identity matrix and P is the calculated state covariance matrix, t is the current time step, and *t|t-1* stands for the conditional relation between two time-steps.

The Kalman Filter is recursive, in the sense that the calculated posterior yields the prior distribution for the next instant, which yield the prediction step:

$$\begin{aligned} \hat{x}_{\,t|t-1} &= \hat{x}_{\,t-1} \\ P_{t|t-1} &= (P^{-1}_{t-1} + Q^{-1})^{-1} \end{aligned} \tag{3.15}$$

It is noteworthy that the above deduction and posterior probability is only valid for the assumptions of the Gaussian process under linear models. The Kalman Filter does not provide

the full posterior probability distribution for nonlinear models, but it does provide the MAP estimation. In this sense, many variations of the above Kalman Filter algorithm are provided in the literature. For instance, the Linear Kalman Filter (Sarri et al., 2016), the Extended Kalman Filter and the Iterated Kalman Filters provide a nonlinear version for the above algorithm (Fan; Wehbe, 2013), while the Unscented Kalman Filter (Valverde; Terzija, 2011), the Cubature Kalman Filter (Sharma; Srivastava; Chakrabarti, 2017) and the Ensemble Kalman Filter (Zhou et al., 2015) are based on sampling transformations. Essentially, these approaches relate to a Bayesian approach since they all rely on a state-space model that encompasses a random variable component to the state variables.

3.2.3 From Bayesian Inference to Tikhonov regression

One particular type of estimator that employs a Conjugate-Gaussian is the Tikhonov regression, from which the so-called regularized estimator (Nocedal; Wright, 2006) explores this perspective to reduce numerical ill-conditioning. Estimation is performed under the following unconstrained optimization method:

$$\hat{x} = \arg\min_{x} [z - h(x)]^T R^{-1} [z - h(x)] + [x - x_0]^T Q^{-1} [x - x_0] \tag{3.16}$$

The above model implies the following choices of prior and likelihood as Multivariate Gaussian distribution, much like the Kalman Filter approach, but without necessarily relying on the previously estimated value recursively. In most applications, the matrix Q is chosen as a multiple of the identity matrix ($Q = \alpha I$), also referred to as Ridge regression. The main goal is to increase numerical conditioning, so the choice of α is typically a small value.

The Bayesian interpretation is the same as the above for the Kalman Filter and is devised in the prior and likelihood described below:

$$\begin{aligned} \pi(x) &\propto \prod_{j=1}^{n} \prod_{k=1}^{n} e^{-(x_j - x_{0,k}) Q_{jk}^{-1} (x_k - x_{0,j})} \\ f_{Z|X}(z|x)(x) &\propto \prod_{i=1}^{m} e^{-\frac{(z_i - h_i(x))^2}{R_{ii}}} \end{aligned} \tag{3.17}$$

The Levenberg–Marquardt algorithm (Nocedal; Wright, 2006) obtains the solution based on the nonlinear optimization theoretical background. This method has been employed in data fusion algorithms to merge the results of two independent estimators accordingly. In this case, the individual variance and the estimated state of each independent estimators are fused in a Tikhonov regression-like method (Simões Costa; Albuquerque; Bez, 2013).

3.2.4 From Bayesian Inference to LASSO

Another estimator of interest is the Least Absolute Shrinkage and Selection Operator (LASSO) estimator, which changes the prior distribution for a Laplace distribution instead of

the Gaussian model (Akrami; Asif; Mohsenian-Rad, 2020). The objective is to reduce the effect of state vector changes and has been mostly exploited for sparse learning algorithms to choose which variables are most probable to enforce changes on the system. The following optimization problem is formulated for the LASSO estimator:

$$\hat{x} = \arg\min_{x} [z - h(x)]^T R^{-1} [z - h(x)] + \lambda |[x - x_0]|_1 \tag{3.18}$$

where λ is a tuning parameter that controls the effect of the L1 norm in the second term, and enforces or avoids changes on the state variables values.

This type of estimator has not been widely employed in power system state estimation since it enforces sparsity on the state variables as if some states kept a fixed value disregarding the measurement information. In the Bayesian interpretation, the following prior and likelihood function is related to the above model under a MAP estimate perspective:

$$\begin{aligned} \pi(x) &\propto \prod_{i=1}^{n} e^{-\frac{|x_i - x_{0,i}|}{b_i}} \\ f_{Z|X}(z|x)(x) &\propto \prod_{i=1}^{m} e^{-\frac{(z_i - h_i(x))^2}{R_{ii}}} \end{aligned} \tag{3.19}$$

This sparsity enforcement for the state variables is not a typical feature exploited for power systems. However, the application of a Laplace likelihood function yields a very well-known robust state estimator, described in the following.

3.2.5 From Bayesian Inference to WLAV

Following the principle of the Laplace distribution, the Weighted Least Absolute Value (WLAV) estimator also have an important role in power system state estimation (Celik; Abur, 1992). In this case, the formulation of such an estimator comprises the following objective function:

$$\hat{x} = \arg\min_{x} |R^{-1/2} [z - h(x)]| \tag{3.20}$$

Solving the above optimization problem, typically employing linear programming techniques, yields the median of the sample distribution. This type of solution exploits that not all residuals will be active as a base during the linear programming algorithm, essentially selecting a reduced set of measurements to influence the final state estimation. In the Bayesian sense, such formulation encompasses a constant prior (as assumed in static state estimation) and a Laplace likelihood (Li; Arce, 2004), as the following:

$$\begin{aligned} \pi(x) &\propto c \\ f_{Z|X}(z|x)(x) &\propto \prod_{i=1}^{n} e^{-\frac{|z_i - h_i(x)|}{\sqrt{R_{ii}}}} \end{aligned} \tag{3.21}$$

The Laplace likelihood function assumption is as hard to prove as the Gaussian assumption in a practical application. However, this estimator presents interesting properties regarding robustness against bad data. The basic idea is that reducing the number of measurements that impacts the final estimate increases the robustness against bad data. During the search for the measurements that will compose the final set, the linear programming methods tend to neglect measurements contaminated with gross errors (or outliers), as they will present high residuals during the estimation process. This increases the robustness of the method and increases accuracy in the presence of bad data. However, it tends to increase estimator variance in normal situations without such contamination. A compromise solution typically employs a mid-term solution between the WLAV and the WLS estimator, as will be described in the following.

3.2.6 From Bayesian Inference to SHGM

Following the idea of the Laplace distribution and its robustness against bad data, but still trying to maintain a low variance estimator, the Schweppe-Huber Generalized M-estimator (SHGM) (Mili et al., 1996) employs the following objective function:

$$\hat{x} = \arg\min_{x} \sum_{i=1}^{m} \rho_i(z_i - h_i(x)) \tag{3.22}$$

$$\rho_i(r_i) = \begin{cases} \frac{1}{2}(z_i - h_i(x))^2 & if |z_i - h_i(x)| \leq a\omega_i \\ a\omega_i |z_i - h_i(x)| - \frac{1}{2}a^2\omega_i^2 & \end{cases} \tag{3.23}$$

The above objective function consists of trimmed version of the weighted least squares approach according to the tuning a and weighting ω_i factors. Once again, as in the static formulation, any knowledge regarding the state variables is included in the estimation process, also yielding an improper constant prior for the state variables.The function $\rho_i(z_i - h_i(x))$ generalizes the estimation process under the spectrum of the M-estimators, with the Huber loss function typically employed, as described above.

$$\begin{aligned} &\pi(x) \propto c \\ &f_{Z|X}(z|x)(x) \propto \prod_{i=1}^{n} e^{-\rho_i(z_i - h_i(x))} \end{aligned} \tag{3.24}$$

Finally, the solution of the above problem is obtained by either small numerical modifications of the Newton method or by the introduction of additional steps based on linear programming (Mili et al., 1996; Abur; Gómez-Expósito, 2004). This estimator typically yields good results in the presence of bad data while maintaining accuracy under normal situations.

3.3 A Bayesian Framework for Power System State Estimation

The fundamental change of perspective in this Thesis, inspired by Bayesian Inference concepts, is to move towards a full probabilistic representation of the state estimation problem.

This may seem redundant to state estimation efforts that already include random noise into account with probabilistic models for the measurements. However, such traditional approaches do not encompass the same level of probabilistic modelling for the state and the many parameters involved in the estimation problem.

As shown before, maximum likelihood is a particular case of Bayesian Inference. Besides, other commonly employed estimators also have a full interpretation under the Bayesian framework. This way, this work explores new aspects from a fundamental background that already encompass theoretically the main approaches employed in power system state estimation. The intent is to leverage the versatility of probabilistic concepts under a Bayesian framework to propose novel methods in power system state estimation, with a particular focus on distribution systems. Distribution system often relies more on a diverse set of complementary information (pseudo measurements) rather than real-time measurements to perform state estimation, due to the lack of sensors in their networks. Besides, they comprise large scale networks that may require additional methods based on exchanging information between areas to maintain scalability. Finally, the high variability and stochastic behaviour associated with the complementary information often require more generalized models instead of a Gaussian parametric model.

The proposition of extending Bayesian Inference for Distribution System State Estimation follows the need of complementary information, but in this case, in the form of prior distributions for the state variables. The influence of such prior distributions may be tuned, increasing or reducing the effect of the prior distribution, which brings additional versatility to the estimation process by moving the prior information between pure non-informative prior to informative one.

In this context, this Thesis explores three applications from the above Bayesian perspective for power system state estimation. They tackle practical limitations and current challenges of the Distribution Systems, and also for power systems in general, and that consists of the main practical contributions of this Thesis, presented in the context of distribution systems:

- *Bayesian Information Fusion*: prior knowledge comes from different sets of sensors, according to their intrinsic temporal characteristics, related to sampling and updating rate. In essence, the measurement vector is decomposed according to updating rates. Regarding distribution systems, the information typically associated with the pseudo measurements can be processed as a prior distribution for the state variables, and then updates on the state variables are obtained by a MAP procedure only with the acquired measurements at a given instant, without the need of mixing measurements from different instants in the same measurement vector;

- *Bayesian Spatial Fusion*: in this second application, the aspect of scalability of power system state estimation is explored under the concept of decomposition methods in the form of a hierarchical multiarea state estimation procedure. The formulation of traditional multiarea state estimation requires that each local area is observable. To complement such

hypothesis, a prior distribution is employed to aid the decomposition of areas, and the prior knowledge is iteratively updated during the coordination step;

- *Estimation with Non-Gaussian Characteristics*: finally, the fundamental hypothesis of a Gaussian model is surpassed by a semi-parametric model using kernel density estimation. A kernel smoothing on the posterior distribution bridges the Bayesian concepts of MAP estimation and information theory concepts, as demonstrated in (Chen; Principe, 2012). Both state variables and measurement noise are associated with kernel models, and estimation is performed as a Maximum Correntropy Kalman Filter, also proposed as part of this Thesis.

The Appendix C presents an illustrative example of the proposed Bayesian Framework, to deal with Information Fusion between SCADA measurements and fast sampled PMUs using linear models. A small example in a four nodes three-phase test system also provides the basic features of such interpretation and how the priors may increase estimation performance and integrate different measurement sets according to their sampling rates.

3.4 Concluding Remarks

State estimation carries a fundamental probabilistic formalization for power systems analysis by dealing with the uncertainty feature on the real-time operational condition of power grids. It bridges the detailed, non-linear, and physics-based model of electric power networks with observed electrical quantities gathered in real-time, by different noisy measurements spread across the power system. Nonetheless, such a perspective has fundamentally thought about the uncertainty associated with the measurements, and few works extrapolated uncertainty as a part of the state variables.

Briefly, the measurements are uncertain due to noise, but the state vector is often assumed as fixed deterministic variables to be found. The Maximum likelihood principle, or variations of it, provides the probabilistic characterization of such noise, and the estimation problem often deals with by the optimization side. Neglecting complementary probabilistic features for the state means assuming an underlying hypothesis during the inference problem, yielding the assumption of an improper constant prior to the state variables. This results in probabilities for extreme and unrealistic state values. The estimation is then entirely dominated by the likelihood function and does not exploit any additional feature or prior expert knowledge about the state, neither upper and lower physical bounds of the physical system. Besides, many approaches that focus on increasing robustness indeed only assume a different parametric model for the noise and likelihood function.

In a complementary perspective, Bayesian Inference naturally captures a broader probabilistic notion of the estimation and such aspects. By extending the interpretation of the state vector as a random variable, the estimation problem reduces to finding one of such possible val-

ues, and not necessarily the exact deterministic one. This way is possible to introduce additional and valuable information to perform estimation, especially under scenarios with constrained or reduced measurements, as in the case of distribution systems.

The following Chapters present the conceptualization of the Bayesian perspective in different applications of interest. Initially, three-phase modelling of distribution networks based on two-port admittance matrices is devised, which can help software implementation of generic detailed distribution system models. In the sequence, three different applications studied in this thesis are presented and evaluated in details. The first introduces kernel density principles in the estimation problem and exploits Information Theory concepts with a Maximum Correntropy Kalman Filter. The second, dealing with the diversity of sampling and updating rates of measurements employed in distribution systems, with a Bayesian information fusion method and the concept of sampling layers. Finally, scalability issues are discussed under a Bayesian spatial fusion method, a type of multiarea state estimation that incorporates prior knowledge in the decomposition of the power grids.

4 DISTRIBUTION SYSTEM MODELLING FOR STATE ESTIMATION

4.1 Introduction

At the core of any computational analysis of power systems, resides the electrical network model. It consists of a mathematical representation of the relations among different electrical quantities from the power grid and must capture the physical aspects of the power system. From a steady state perspective, the network model faces the challenge of capturing the algebraic relations among the electrical quantities from the network. The quantities of interest may comprise electrical currents, voltage phasors, active and reactive power flows and loads, active and reactive injections from generators, discrete controller's conditions, the energy available at storage devices, and so on. For distributions systems such challenge easily grows, as the three-phase unbalanced nature and the large scale of these networks result in a diverse set of possible connections, phase coupling, asymmetries, and topologies of the system.

This chapter introduces the basic concepts regarding three-phase network models for distribution systems steady state analysis, emphasizing a state estimation perspective. A general two-port branch model is formulated based on the admittance matrix of each component of the distribution system. General equations to calculate current and power flows in the network as well as their derivatives are presented. Finally, different types of equipment are exemplified along with the respective particularities of their admittance matrix models.

4.2 Three-Phase Two-port Models

This section introduces the formulation of a generic three-phase unbalanced model able to capture the main electrical quantities for steady state analysis. Power flow equations are initially derived since they comprise the majority of information in this type of analysis . In the sequence, a generic two-port model based on admittance matrices is formulated to simplify the equations of a diverse set of possible connections, the number of phases, grounding conductors, and asymmetries from different equipment.

4.2.1 Three-phase two-port admittance model

The three-phase network model may also be derived by representing each network component by its two-port admittance matrix, using direct concepts from the classic linear circuit theory, as illustrated in Figure 4. Different types of equipment and connections can be incorporated in the network model by properly defining their contribution to the respective admittance matrix, which represent an implementation advantage when representing complex distribution networks with several distinct components (Primadianto; Lu, 2017; Lefebvre; Prévost; Lenoir, 2014; Hebling et al., 2020).

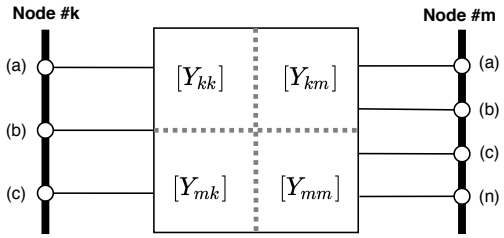


Figure 4 – Generic component model based on a two-port admittance matrix. Four submatrices relate the voltage and current phasors between both terminals, according to the type of component.

Each element is modeled by the following two-port admittance system:

$$\begin{pmatrix} \underline{\dot{I}_{km}} \\ \underline{\dot{I}_{mk}} \end{pmatrix} = \begin{pmatrix} Y_{kk} & Y_{km} \\ Y_{mk} & Y_{mm} \end{pmatrix} \begin{pmatrix} \underline{\dot{V}_{k}} \\ \underline{\dot{V}_{m}} \end{pmatrix} \tag{4.1}$$

where, k and m denote the terminals of the branch element, $\underline{\dot{I}_{km}}$ and $\underline{\dot{I}_{mk}}$ are the current phasor injected in each terminal of the branch element, that flows from terminal k to m, and from m to k, respectively, $\underline{\dot{V}_{k}}$ and $\underline{\dot{V}_{m}}$ the voltage phasor of each terminal, and Y_{kk}, Y_{km}, Y_{mk} and Y_{mm} are the admittance sub-matrices that represent each physical component and the respective parameters.

The power flow equations can be written for the respective branch as the following matrix expression for the terminal k:

$$\underline{S_{km}} = \underline{P_{km}} + j\underline{Q_{km}} = \underline{\dot{V}_{k}} \odot (\underline{\dot{I}_{km}})^{*} = \underline{\dot{V}_{k}} \odot (Y_{kk}\underline{\dot{V}_{k}} + Y_{km}\underline{\dot{V}_{m}})^{*} \tag{4.2}$$

where, $\underline{S_{km}}$ is a vector with the per-phase complex power at terminal k (active and reactive power flows, $\underline{P_{km}}$ and $\underline{Q_{km}}$), $\odot$ denotes the Hadamard product (element-wise), and * denotes the complex conjugate. As an example, for a three-phase terminal k with phases abc, the following equation can be written trough the matrix notation of the two-port model and the above product:

$$\underline{S_{km}} = \underline{\dot{V}_{k}} \odot (\underline{\dot{I}_{km}})^{*} = \begin{pmatrix} \dot{V}_{k}^{a}(\dot{I}_{km}^{a})^{*} \\ \dot{V}_{k}^{b}(\dot{I}_{km}^{b})^{*} \\ \dot{V}_{k}^{c}(\dot{I}_{km}^{c})^{*} \end{pmatrix} = diag(\underline{\dot{V}})(\underline{\dot{I}_{km}})^{*} = \begin{pmatrix} \dot{V}_{k}^{a} & 0 & 0 \\ 0 & \dot{V}_{k}^{b} & 0 \\ 0 & 0 & \dot{V}_{k}^{c} \end{pmatrix} \begin{pmatrix} (\dot{I}_{km}^{a})^{*} \\ (\dot{I}_{km}^{b})^{*} \\ (\dot{I}_{km}^{c})^{*} \end{pmatrix} \tag{4.3}$$

Besides, the voltage and current phasors can also be written with the following notation to split magnitude and phase angle, for instance:

$$\underline{\dot{V}_{k}} = \underline{V_{k}} \odot \underline{\theta_{k}} = diag(\underline{V_{k}})\underline{\theta_{k}} = [V_{k}]\underline{\theta_{k}} = \begin{pmatrix} V_{k}^{a} & 0 & 0 \\ 0 & V_{k}^{b} & 0 \\ 0 & 0 & V_{k}^{c} \end{pmatrix} \begin{pmatrix} e^{j\theta_{k}^{a}} \\ e^{j\theta_{k}^{b}} \\ e^{j\theta_{k}^{c}} \end{pmatrix} \tag{4.4}$$

where $[.]$ denotes only in this section, for the sake of notation simplification, the diagonal matrix operator $diag()$, which forms a diagonal matrix with the vector elements, $\underline{V_{k}}$ denotes the voltage

magnitude vector in each phase of the node k of the two-port model, and $\underline{\theta_k}$ is the vector composed by the complex exponentials of the voltage phase angle in each phase of the node k of the two-port model.

Expanding the power flow matrix expression, it is possible to obtain the classical active and reactive power flow equations for a three-phase component (Arrillaga; Harker, 1978):

$$P_{km}^{i} = V_k^i \sum_{j\in\Phi_{km}} \left(V_k^j \left(g_{kk}^{ij} \cos(\theta_k^i - \theta_k^j) + b_{kk}^{ij} \sin(\theta_k^i - \theta_k^j) \right) \right.$$
$$\left. - V_m^j \left(g_{km}^{ij} \cos(\theta_k^i - \theta_m^j) + b_{km}^{ij} \sin(\theta_k^i - \theta_m^j) \right) \right) \tag{4.5}$$

$$Q_{km}^{i} = V_k^i \sum_{j\in\Phi_{km}} \left(V_k^j \left(g_{kk}^{ij} \sin(\theta_k^i - \theta_k^j) - b_{kk}^{ij} \cos(\theta_k^i - \theta_k^j) \right) \right. \tag{4.6}$$
$$\left. - V_m^j \left(g_{km}^{ij} \sin(\theta_k^i - \theta_m^j) - b_{km}^{ij} \cos(\theta_k^i - \theta_m^j) \right) \right) \tag{4.7}$$

where, i and j denote different phases of the component; g_{kk}^{ij} and b_{kk}^{ij} are the real and imaginary parts of the ij element from the two-port model submatrix Y_{kk}; g_{km}^{ij} and b_{km}^{ij} are the real and imaginary parts of the ij element from the two-port model submatrix Y_{km}; Φ_{km} is the set of existing phases at branch k-m.

Regarding voltage phasors as state variables, these are typically represented in polar coordinates since there is a direct interpretation of both magnitude and phase angles, and as done in transmission systems. There is, however, the possibility of working in rectangular coordinates to represent the state variables, which may benefit estimators designed to deal with phasor measurement units.

The next subsections describe in details the matrix equations for active and reactive power flows and their respective derivatives for each two-port model. Distinct measurements for a diverse set of components may be described using such equations, only by changing the respective two-port admittance matrices according to the type of component and connections.

4.2.2 Polar Coordinates

Separating the voltage phasor in its exponential representation according to the following expression:

$$\underline{\dot{V}_k} = \underline{V_k} \odot \underline{\theta_k} = diag(\underline{V_k})\underline{\theta_k} = [V_k]\underline{\theta_k} \tag{4.8}$$

where $\underline{V_k}$ is a vector composed by the voltage magnitude in the $i-th$ phase of the system ($V_k^i = |\dot{V}_k^i|$) , and $\underline{\theta_k}$ is a vector composed by the complex exponential part of the angle of the phasor ($\theta_k^i = e^{j\theta_k^i}$). Working algebraically, it is possible to obtain the following expression for the power flow equations:

$$\underline{S_{km}} = [V_k][\theta_k]Y_{kk}^*[V_k]\underline{\theta_k}^* + [V_k][\theta_k]Y_{km}^*[V_m]\underline{\theta_m}^* \tag{4.9}$$

The above equation is used to build active and reactive power flow models for steady state analysis. The derivatives of such an equation regarding the state variables (complex nodal voltages in polar coordinates), in the same matrix representation, are:

$$\frac{\partial \underline{S_{km}}}{\partial V_k^i} = J^{ii}[\theta_k]Y_{kk}^*[V_k]\underline{\theta_k}^* + [V_k][\theta_k]Y_{kk}^*J^{ii}\theta_k^* + J^{ii}[\theta_k]Y_{km}^*[V_m]\underline{\theta_m}^*$$

$$\frac{\partial \underline{S_{km}}}{\partial V_m^i} = [V_k][\theta_k]Y_{km}^*J^{ii}\underline{\theta_m}^* \tag{4.10}$$

$$\begin{aligned}\frac{\partial \underline{S_{km}}}{\partial \theta_k^i} = {} & je^{\theta_k^i}[V_k]J^{ii}Y_{kk}^*[V_k]\underline{\theta_k}^* \\ & -j[V_k][\theta_k]Y_{kk}^*[V_k]J^{ii}\underline{\theta_k}^* \\ & +je^{\theta_k^i}[V_k]J^{ii}Y_{km}^*[V_m]\underline{\theta_m}^*\end{aligned} \tag{4.11}$$

$$\frac{\partial \underline{S_{km}}}{\partial \theta_m^i} = -j[V_k][\theta_k]Y_{km}^*[V_m]J^{ii}\underline{\theta_m}^* \tag{4.12}$$

where, J^{ii} is a matrix with the same size as the number of phases and only the ii element equal one and the others equal zero. The above equation yields derivatives for all phases of the active and reactive power flow at once.

At first glance, such equations may appear complex, but note that they represent general equations for multiple phase systems and different types of components. They also are in full vector and matrix format, which facilitates the implementation of generic models in a plug-and-play manner for different types of components.

Similarly, it is possible to obtain the power flow equations in the opposite terminal m of the two-port model:

$$\underline{S_{mk}} = [V_m][\theta_m]Y_{mk}^*[V_k]\underline{\theta_k}^* + [V_m][\theta_m]Y_{mm}^*[V_m]\underline{\theta_m}^* \tag{4.13}$$

And also its derivatives:

$$\frac{\partial \underline{S_{mk}}}{\partial V_k^i} = [V_m][\theta_m]Y_{mk}^*J^{ii}\underline{\theta_k}^* \tag{4.14}$$

$$\frac{\partial \underline{S_{mk}}}{\partial V_m^i} = J^{ii}[\theta_m]Y_{mk}^*[V_k]\underline{\theta_k}^* + J^{ii}[\theta_m]Y_{mm}^*[V_m]\underline{\theta_m}^* + [V_m][\theta_m]Y_m^*J^{ii}\underline{\theta_m}^* \tag{4.15}$$

$$\frac{\partial \underline{S_{mk}}}{\partial \theta_k^i} = -j[V_m][\theta_m]Y_{mk}^*[V_k]J^{ii}\underline{\theta_k}^* \tag{4.16}$$

$$\frac{\partial \underline{S_{mk}}}{\partial \theta_m^i} = je^{\theta_m^i}[V_m]J^{ii}Y_{mk}^*[V_k]\underline{\theta_k}^* + je^{\theta_m^i}[V_m]J^{ii}Y_{mm}^*[V_m]\underline{\theta_m}^* - j[V_m][\theta_m]Y_{mm}^*[V_m]J^{ii}\underline{\theta_m}^* \tag{4.17}$$

Current magnitudes and current phase angles can also be modelled with this formulation. Typically, distribution feeders present current magnitude measurements that are often neglected on transmission system state estimation to avoid numerical issues. However, in distribution systems, such practice reduces the already compromised observability and current magnitude measurements are often modelled to increase real time information about the feeders. Current phase angles are also a possible measurement by syncronized phasore measurement units, although not so common in typical distribution systems so far.

$$I_{km}^{i} = \frac{\sqrt{{P_{km}^{i}}^{2} + {Q_{km}^{i}}^{2}}}{V_{k}^{i}} \tag{4.18}$$

$$\delta_{km}^{i} = \theta_{k}^{i} - \arctan\left(\frac{Q_{km}^{i}}{P_{km}^{i}}\right) \tag{4.19}$$

The derivatives can be obtained by applying the chain rule in the above equations together with the previous equations shown for the derivatives of the power flows. They can be refereed to in (Zhu; Abur, 2007; Korres; Manousakis, 2012). It is noteworthy that, in polar coordinates, the derivative of current equations may present some numerical issues, especially for lightly loaded circuits and at initialization of algorithms for steady state analysis (Zhu; Abur, 2007; Korres; Manousakis, 2012).

4.2.3 Rectangular Coordinates

Based on the voltage phasor in rectangular coordinates, the following expression for the power flow equation is obtained:

$$\begin{aligned}\underline{S_{km}} =& [V_{re,k}]Y_{kk}^{*}\underline{V_{re,k}} + [V_{re,k}]Y_{km}^{*}\underline{V_{re,m}} + [V_{im,k}]Y_{kk}^{*}\underline{V_{im,k}} + [V_{im,k}]Y_{km}^{*}\underline{V_{im,m}} \\ & - j[V_{re,k}]Y_{kk}^{*}\underline{V_{im,k}} - j[V_{re,k}]Y_{km}^{*}\underline{V_{im,m}} + j[V_{im,k}]Y_{kk}^{*}\underline{V_{re,k}} + j[V_{im,k}]Y_{km}^{*}\underline{V_{re,m}}\end{aligned} \tag{4.20}$$

where $\underline{V_{re,k}}$ is a vector composed by the real part of the voltage phasor in each $i-th$ phase of the system ($V_{re,k}^{i} = \Re\{\dot{V}_{k}^{i}\}$), and $\underline{V_{im,k}}$ is a vector composed by the imaginary part of the voltage phasor in each $i-th$ phase of the system ($V_{im,k}^{i} = \Im\{\dot{V}_{k}^{i}\}$). The above equation is used to build an active and reactive power flow model for steady state analysis. The derivatives of such equation regarding the state variables (complex nodal voltages in rectangular coordinates) in the same matrix representation are:

$$\frac{\partial \underline{S_{km}}}{\partial V_{re,k}^{i}} = [J^{i}](Y_{kk}^{*}\underline{V_{re,k}} + Y_{km}^{*}\underline{V_{re,m}} - jY_{kk}^{*}\underline{V_{im,k}} - jY_{km}^{*}\underline{V_{im,m}}) + ([V_{re,k}]Y_{kk}^{*} + j[V_{im,k}]Y_{kk}^{*})\underline{J^{i}} \tag{4.21}$$

$$\frac{\partial \underline{S_{km}}}{\partial V_{re,m}^{i}} = [V_{re,k}]Y_{km}^{*}\underline{J^{i}} + j[V_{im,k}]Y_{km}^{*}\underline{J^{i}} \tag{4.22}$$

$$\frac{\partial \underline{S_{km}}}{\partial V^i_{im,k}} = [J^i](jY^*_{kk}\underline{V_{re,k}} + Y^*_{kk}\underline{V_{im,k}} + jY^*_{km}\underline{V_{re,m}} + Y^*_{km}\underline{V_{im,m}}) + ([V_{im,k}]Y^*_{kk} - j[V_{re,k}]Y^*_{kk})\underline{J^i} \quad (4.23)$$

$$\frac{\partial \underline{S_{km}}}{\partial V^i_{im,m}} = -j[V_{re,k}]Y^*_{km}\underline{J^i} + [V_{im,k}]Y^*_{km}\underline{J^i} \quad (4.24)$$

where, $\underline{J^i}$ is a vector with the same size as the number of phases and only the $i-th$ element equal one and the others equal zero.

Similarly, it is possible to obtain the power flow equations in the opposite terminal m of the two-port model:

$$\begin{aligned}\underline{S_{mk}} =&[V_{re,m}]Y^*_{mk}\underline{V_{re,k}} + [V_{re,m}]Y^*_{mm}\underline{V_{re,m}} + [V_{im,m}]Y^*_{mk}\underline{V_{im,k}} + [V_{im,m}]Y^*_{mm}\underline{V_{im,m}} \\ &+ j[V_{im,m}]Y^*_{mk}\underline{V_{re,k}} - j[V_{re,m}]Y^*_{mk}\underline{V_{im,k}} + j[V_{im,m}]Y^*_{mm}\underline{V_{re,m}} - j[V_{re,m}]Y^*_{mm}\underline{V_{im,m}}\end{aligned} \quad (4.25)$$

And also its derivatives:

$$\frac{\partial \underline{S_{mk}}}{\partial V^i_{re,k}} = [V_{re,m}]Y^*_{mk}\underline{J^i} + j[V_{im,m}]Y^*_{mk}\underline{J^i} \quad (4.26)$$

$$\frac{\partial \underline{S_{mk}}}{\partial V^i_{re,m}} = [J^i](Y^*_{mk}\underline{V_{re,k}} - jY^*_{mk}\underline{V_{im,k}} + Y^*_{mm}\underline{V_{re,m}} - jY^*_{mm}\underline{V_{im,m}}) + ([V_{re,m}]Y^*_{mm} + j[V_{im,m}]Y^*_{mm})\underline{J^i} \quad (4.27)$$

$$\frac{\partial \underline{S_{mk}}}{\partial V^i_{im,k}} = [V_{im,m}]Y^*_{mk}\underline{J^i} - j[V_{re,m}]Y^*_{mk}\underline{J^i} \quad (4.28)$$

$$\frac{\partial \underline{S_{mk}}}{\partial V^i_{im,k}} = [J^i](jY^*_{mk}\underline{V_{re,k}} + Y^*_{mk}\underline{V_{im,k}} + jY^*_{mm}\underline{V_{re,m}} + Y^*_{mm}\underline{V_{im,m}}) + ([V_{im,m}]Y^*_{mm} - j[V_{re,m}]Y^*_{mm})\underline{J^i} \quad (4.29)$$

The current real and imaginary parts may also be modelled with this formulation. By dealing both with the current phasor as well as the state variables in rectangular coordinates is possible to build a linear model.

$$\underline{I_{re,km}} = \Re\{Y_{kk}\}\underline{V_{re,k}} - \Im\{Y_{kk}\}\underline{V_{im,k}} + \Re\{Y_{km}\}\underline{V_{re,m}} - \Im\{Y_{km}\}\underline{V_{im,m}} \quad (4.30)$$

$$\underline{I_{im,km}} = \Im\{Y_{kk}\}\underline{V_{re,k}} + \Re\{Y_{kk}\}\underline{V_{im,k}} + \Im\{Y_{km}\}\underline{V_{re,m}} + \Re\{Y_{km}\}\underline{V_{im,m}} \quad (4.31)$$

$$\underline{I_{re,mk}} = \Re\{Y_{mk}\}\underline{V_{re,k}} - \Im\{Y_{mk}\}\underline{V_{im,k}} + \Re\{Y_{mm}\}\underline{V_{re,m}} - \Im\{Y_{mm}\}\underline{V_{im,m}} \quad (4.32)$$

$$\underline{I_{im,mk}} = \Im\{Y_{mk}\}\underline{V_{re,k}} + \Re\{Y_{mk}\}\underline{V_{im,k}} + \Im\{Y_{mm}\}\underline{V_{re,m}} + \Re\{Y_{mm}\}\underline{V_{im,m}} \quad (4.33)$$

The derivatives can be easily obtained from the real and imaginary parts of the two-port model admittance submatrices. This is an important aspect in distribution systems, since it motivated the pursuit of current-based state estimation formulations (Primadianto; Lu, 2017; Pau; Pegoraro; Sulis, 2013; Almeida; Ochoa, 2017). This approach is also often employed when dealing with PMUs, in order to simplify the representation and take computational advantage without losing accuracy (Pau; Pegoraro; Sulis, 2013; Almeida; Ochoa, 2017).

Appendix D presents a detailed description of the main components and respective two-port admittance models, that capture high-resolution details of the majority of traditional equipment installed across distribution networks. The matrix representation of each component along the physical description and theoretical foundations for each representation are also provided.

4.3 Concluding Remarks

Power system computational analysis is essential for operating and planning of electrical distribution networks. In their essence, distribution systems are unbalanced and asymmetrical electrical networks, spread across large areas connecting a diverse set of components with final consumers. The challenges associated with suitable distribution systems modelling, add complexity to any analysis performed. They include the increase of variables that three-phase representation carries and also encompasses detailment of component representation. From the perspective of distribution system state estimation, the challenges rely on accurate network models capable of dealing with measurements and information from all different elements of the distribution systems to translate into adherent models that capture the reality nature of such systems.

This chapter presented the fundamental equations to perform steady state analysis in distribution networks based on the power flow of the network. The model comprises a three-phase generic representation of different components and the main electrical quantities of such models, as active and reactive power flows and voltage and current phasors. A two-port admittance model facilitates the implementation of the steady state models of the distribution system components, under a matrix implementation.

Details of the two-port models of the main components of substations, primary feeders, and secondary low voltage circuits are shown in the Appendix section, such as power transformers, distribution circuits, shunt capacitors, and loads. All these aspects, if adequately addressed within the network models, can increase awareness of the system, providing accurate information for all components in distribution systems. It is a crucial feature for state estimators to act more actively in the operation of the modern distribution systems. This can be further explored into specific future works, such as extending the implementations towards four-wire, detailed neutral and grounding circuits, piecewise linear representation of power electronics converters, parallel and coupled circuits, and three-phase representation of transmission systems as well.

5 TRACKING POWER SYSTEM STATE EVOLUTION WITH MAXIMUM-CORRENTROPY-BASED EXTENDED KALMAN FILTER

The content of this chapter have been published on the following Journal:

- J.A.D. Massignan, J.B.A. London Junior, V. Miranda,"Tracking Power System State Evolution with Maximum-correntropy-based Extended Kalman Filter". *Journal of Modern Power Systems and Clean Energy*, v. 8, p. 616-626, 2020.

5.1 Introduction

This chapter is an effort towards a new interpretation of the state estimation problem that can encompass both the statistical robustness as well as flexibility to overcome system transitions in a single framework. In this sense, by exploiting Information Theory concepts and modelling the noise characteristic through a kernel density estimation approach, this paper proposes a Maximum Correntropy Extended Kalman Filter (MCEKF) to track the state of power systems. As shown in the initial work of Chen (Chen et al., 2017), by incorporating Information Theory concepts within the traditional Kalman Filter it is possible to increase accuracy without relying on the assumption of a Gaussian distribution for the measurement and process noise.

This chapter extends these efforts with a new MCEKF version, enabling the incorporation of both SCADA and PMU measurements in the estimation process, a more practical approach to represent realistic metering systems. It also presents a novel way to deal with suspect samples and system transitions, by performing a new special adjustment on the kernel size in the estimation process during the occurrence of such events. The main contributions can be summarized as:

- a new concept for the state variables, namely state evolution, formulated as a nonparametric probabilistic model for the state, that encompasses the possibility of different state values and transitions that may occur;
- the application of the Maximum Correntropy within an Extended Kalman Filter for the tracking state estimation of power systems, a type of quasi-steady state estimator, based on a Modified-Newton algorithm, able to deal with nonlinear models from both SCADA and PMUs;
- a novel method to suppress the effect of suspect samples and abrupt system transitions through a special adjustment of the kernel Parzen window sizes.

A few theoretical topics, presented in previous sections, will be revisited in the following sections to provide a quick background of the fundamental details.

5.2 Theoretical Background

5.2.1 Power System Static State Estimation

A nonlinear measurement model formulates the static state estimation problem for a power system with m measurements and n state variables (Abur; Gómez-Expósito, 2004; Zhao et al., 2019):

$$z = h(x) + e \tag{5.1}$$

where z is the measurement vector *(m x 1)*; x is the state vector *(n x 1)*; *h(x)* is the set of nonlinear equations that relates the measured electrical quantities with the state; e is the random error vector *(m x 1)* that represents noise in the measurements. Traditional state estimation assumes the noise vector as multivariate Gaussian independent random variables, with zero mean and known covariance matrix *R (m x m)*. The state vector is typically composed by the complex nodal voltage of the system, in this work represented in rectangular coordinates.

The maximization of the likelihood function results in the Weighted Least Squares (WLS) criterion to find the estimated state, referred in this work as WLS Snapshot state estimator. Under an independent multivariate Gaussian noise assumption, the following optimization problem is written.

$$\max l(e|x) = \prod_{i=1}^{m} f(e_i|x) = \prod_{i=1}^{m} \frac{1}{2\pi R_{ii}^2} e^{-\frac{(z_i - h_i(x))^2}{2R_{ii}^2}} \tag{5.2}$$

Working with the above equation, the estimated state $\hat{x}$ can then be obtained by the unconstrained minimization problem:

$$\hat{x} = \min [z - h(x)]^T R^{-1} [z - h(x)] \tag{5.3}$$

5.2.2 Tracking State Estimation

The fast sampling of PMUs (from 1 to 120 samples per second) and synchronization via GPS time stamps is a strong motivation for the formulation of state estimation as a time-dependent problem. Besides the PMU measurements, there are also measurements gathered from the SCADA system, thereby still requiring a nonlinear measurement model in order to include all the information available in the estimation. The authors in (Zhao et al., 2019) provide a comprehensive definition of different methods and objectives to include temporal aspects in the estimation problem. Basically, they can be separated in:

- **dynamical estimators**, which encompass differential equations, generators models, speed and excitation controllers of synchronous machines, and aim at estimating the dynamic state of the network and internal variables of loads and generators;

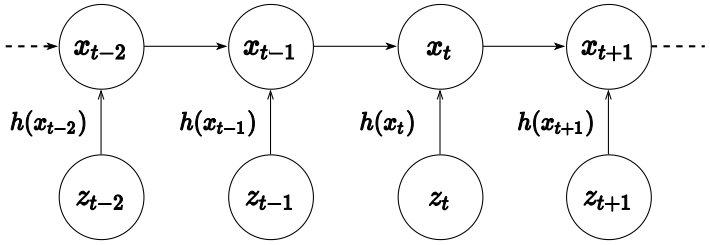


Figure 5 – Hidden markov model representation of the state space model for the tracking state estimation problem.

- **forecasting-aided state estimators** that use a forecasting model to predict the states in a subsequent instant, through smoothing functions for the state transition, aiming at the algebraic state of the network, the complex nodal voltages in the network;

- **tracking state estimators**, a particular case of the previous one where variations are assumed small enough, so the prediction corresponds to the last estimated state.

This work explores the tracking state estimation problem, since it is one of the most affected by abrupt state changes (Zhao et al., 2019). This approach may lead to fewer modifications in static estimators (and network database) currently implemented in control centers to achieve the tracking model, while obtaining significant gains in terms of accuracy, which is a practical advantage. The tracking state estimation consists in the following discrete time-variant nonlinear model:

$$\begin{aligned} x_t &= x_{t-1} + \omega_t \\ z_t &= h(x_t) + e_t \end{aligned} \tag{5.4}$$

where ω_t is the random error that represents process noise with zero mean and known covariance matrix Q (nxn). Traditional tracking state estimation also assumes the noise vectors as multivariate Gaussian independent random variables, and the state space corresponds to a Hidden Markov Model, as described in (5.4) and illustrated in Fig. 5.

A notable conceptual trait in this formulation is that the state variables are now modeled with a random component (ω_t). Despite the fact that static state estimation in (5.1) carries the notion of a state variance, as the resulting estimator variance, it does not encompass the hypothesis of different possible values for the state within its formulation and during the solution as an inference problem. The solution of the estimation problem in such state space model corresponds to the well-known Kalman Filter algorithm, based on the WLS criterion (Abur; Gómez-Expósito, 2004; Zhao et al., 2019). It consists of a two-step recursive algorithm, with the following formulation for the tracking state estimation model:

Prediction Step: The prior mean and covariance matrix are given by:

$$\hat{x}_{t|t-1} = \hat{x}_{t-1}$$
$$P_{t|t-1} = P_{t-1} + Q \tag{5.5}$$

Update Step: The Kalman Filter gain, the posterior state and posterior covariance are then updated:

$$K_t = P_{t|t-1} H^T \left(H P_{t|t-1} H^T + R \right)^{-1} \tag{5.6}$$

$$\hat{x}_t = \hat{x}_{t|t-1} + K_t \left(z_t - H\hat{x}_{t|t-1} \right) P_t = (I - K_t H) P_{t|t-1} (I - K_t H)^T + K_t R K_t^T \tag{5.7}$$

where, H is the Jacobian matrix of the measurement model, K_t is the Kalman Filter Gain matrix, I is the identity matrix and P is the calculated state covariance matrix, t is the current time step, and *t|t-1* stands for the conditional relation between two time-steps (*t* given *t-1*). The nonlinear version of the above model corresponds to the WLS Extended Kalman Filter (EKF) that performs linearization of the measurement equations, and the Iterated EKF (IEKF) that updates such linearization iteratively, thus presenting more accurate results.

5.2.3 Information Theory Concepts

Based on Information Theory, the Correntropy concept arises as a measure of similarity between two probability distribution functions (pdf) (Chen et al., 2017; Miranda; Santos; Pereira, 2009; Liu; Pokharel; Principe, 2006). It has a strong relation to the Entropy concept, which measures the information content of a pdf. The Correntropy between two random variables A and B is defined in (5.8) for a finite number of samples.

$$V(A,B) = E\left[k_\sigma(A-B)\right] = \frac{1}{N}\sum_{i=1}^{N} k_\sigma(A_i - B_i) \tag{5.8}$$

where $k(\dot{})$ represents a kernel for the random variable in the sampled value; σ is the kernel bandwidth (or size of the Parzen window in the Parzen-Rosenblatt estimator (Freitas; Costa; Miranda, 2017)); and the subscript i represents limited number N of samples of the random variables. The concept of Correntropy induces a distance function to measure the similarity of two pdfs, the Correntropy Induced Metric (CIM). Depending upon the choice of kernel and size of the bandwidth, this metric may vary from the L_{inf} to L_0 norms. Because of such property, it can be tuned to become insensitive to outliers, pushing it towards the L_0 norm (indifference in distance) (Pesteh et al., 2019).

Finally, the CIM distance motivated the creation of the Maximum Correntropy Criterion (MCC) to perform inference. For instance, in a regression analysis, it aims at finding parameter values that maximize the similarity among a set of observations ($r_i = A_i - B_i$).

$$\hat{W} = \underset{W\in\Omega}{\text{argmax}} \frac{1}{N}\sum_{i=1}^{N} k_\sigma(r_i) \tag{5.9}$$

where, $\hat{W}$ is the optimal parameter value, and Ω is the feasible set for the parameters.

5.3 Tracking State Evolution with a Maximum Correntropy Extended Kalman Filter

5.3.1 State Evolution Concept in Power Systems

The temporal aspect of the state estimation problem and the probabilistic interpretation of the state variables provide a conceptual framework to evaluate the network condition under both measurement and system uncertainties. However, it has a limitation related to the parametric model typically assumed for the process and measurement noise under conventional Kalman Filter formulations. Many efforts have been made towards exploring Non-Gaussian noise situations (Zhao et al., 2019; Wang et al., 2019; Zhao; Mili, 2018a), such as in the case of bad data, cyber attacks or even particular noise characteristic of PMUs. This work proposes a new interpretation regarding the process noise towards a generalist non-Gaussian model based on the kernel density estimation principle. The notion that, during a sequence of observations from the measurements, the state variables can present different possible values induces the proposed concept of State Evolution. Such different values can be related to system's transitions, and that can be:

1. Systemic: due to sudden load variations, generation dispatch or controller actions;
2. Structural: due to contingencies, switching operations or changes in the network;
3. Random: due to intermittency or failures;
4. Induced: due to cyber attacks or unsupervised switching.

An illustration of such transitions in a power system, observed by a measurement set, monitored by a state estimation process is illustrated in Fig. 6. Due to contingencies, changes in the controllers, natural load variation, generation intermittency, each algebraic state of the network may assume different values over time. The state evolution aims at capturing such different possible values within a single non-Gaussian pdf for the state.

A set of state vectors defines the state evolution concept, each element related to a respective set of observations at the instant t, during an observation window Δt.

$$x = \{x_0, x_1, \ldots, x_{t-1}, x_t, x_{t+1}, \ldots x_{\Delta t}\} = x_t|_{t \in \Delta t} \tag{5.10}$$

Associated with this set of state vectors, the underlying pdf of the state can be obtained in a nonparametric manner with the introduction of a kernel density estimator. The Parzen-Rosenblatt method (Freitas; Costa; Miranda, 2017) for kernel density estimation then obtains the empirical pdf for the state through (5.11). Fig. 7 illustrates this concept for a single state variable during different conditions of the network and the respective empirical probability distribution function.

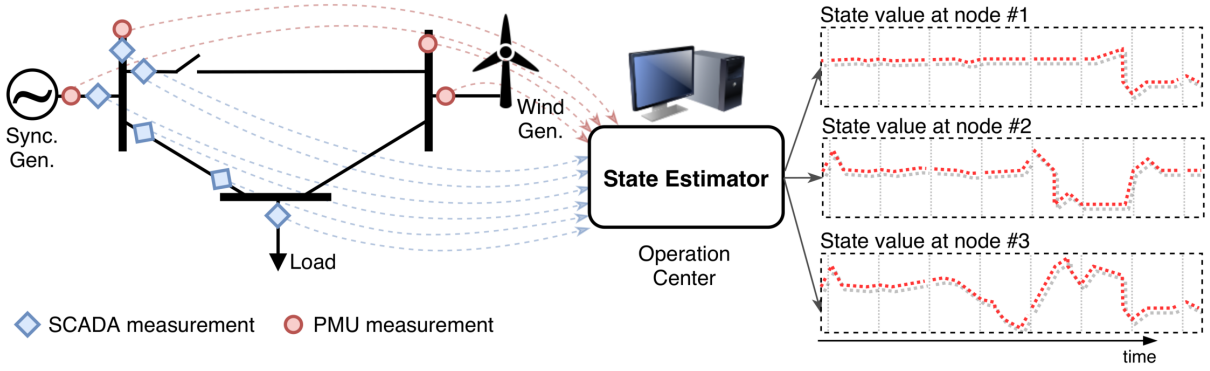


Figure 6 – Example of a power system and respective estimated state during different system transitions, monitored by SCADA and PMU measurements.

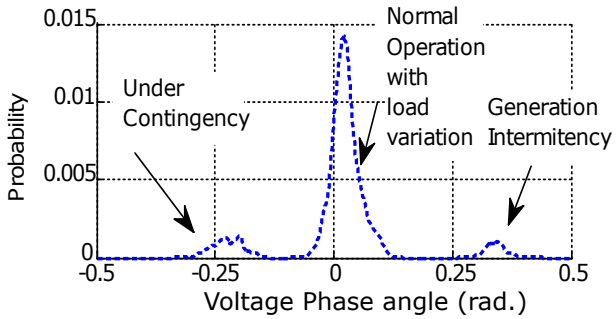


Figure 7 – State evolution concept for a single state variable (e.g. voltage phase angle), under different conditions, obtained by kernel density estimation.

$$x \sim \frac{1}{\Delta t} \sum_{t \in \Delta t} k_\sigma(x_t) \tag{5.11}$$

where, $k(\dot{})$ represents a kernel for the random variable in the sampled value; σ is the kernel bandwidth (or size of the Parzen window in the Parzen-Rosenblatt estimator (Freitas; Costa; Miranda, 2017)).

Since the goal is capturing the state transitions, through the state space model in (5.4), an observation kernel is associated with the process noise.

$$\omega_t \sim k_\sigma(x_t - x_{t-1}) \tag{5.12}$$

Along with the state vector modelled as a nonparametric random variable, this work also revises the hypothesis that the measurement noise follows a Gaussian distribution. In the case of PMUs, Zhao (Wang et al., 2018) has shown with experimental results that such a hypothesis does not hold and more robust models are a requirement for PMU-based state estimators. In a different perspective, Miranda (Miranda; Santos; Pereira, 2009) showed that the presence of gross errors (bad data) consists of a non-Gaussian noise situation. To encompass such general

probabilistic model for the measurement model, in the same way as for the state vector, the measurement noise is modelled under a kernel assumption related to each observation.

5.3.2 Extended Kalman Filter based on a Maximum Correntropy Criterion

The MCC Kalman Filter was first introduced by Chen (Chen et al., 2017), and it aims at performing inference under a state space model without the assumption of Gaussian noise. The only assumption about the noise vectors is the second-order moment is known and given by the following:

$$B_P^{-1} x_t = B_P^{-1} (x_{t-1} + \omega_t) \tag{5.13}$$

$$B_R^{-1} z_t = B_R^{-1} (h(x_t) + e_t) \tag{5.14}$$

where, B_P and B_R are the Cholesky decomposition of the process and measurement covariance matrices, respectively. The state space in (5.4) is rescaled by the respective decomposition.

The MCEKF was then devised in (Zhang; Wang; Giannakis, 2019) for nonlinear models in the context of power systems. This chapter applies the maximum Correntropy principle for the tracking state estimation in power system, which yields the following optimization problem:

$$\max J_{MCC} = \sum_{i=1}^{m} k_\sigma (z_{i,t} - h_i(x_t)) + \sum_{j=1}^{n} k_\sigma (x_{j,t} - x_{j,t-1}) \tag{5.15}$$

A Gaussian kernel, shown in (5.16), is assumed to compose the process and measurement noise. It is noteworthy that the assumption of the kernel function does not mean the same assumption for the pdf of the random variable.

$$k_\sigma (r_i) = e^{-\frac{r_i^2}{2\sigma^2}} \tag{5.16}$$

Thereby, the estimation problem consists in the following optimization problem:

$$\max J_{MCC}(x_t) = \sum_{i=1}^{m} e^{-\frac{\left(z_{i,t} - h_i(x_t)\right)^2}{2\sigma^2}} + \sum_{j=1}^{n} e^{-\frac{\left(x_{j,t} - x_{j,t-1}\right)^2}{2\sigma^2}} \tag{5.17}$$

5.3.3 Numerical Solution Method

Since it is a nonlinear optimization problem, an iterative numerical procedure must be employed to find the solution, that is, the state vector values at instant t. In this work a Modified-Newton algorithm (Nocedal; Wright, 2006) is implemented to find the solution. It is noteworthy that it consists in a different solution method than the first MCEKF proposed in (Zhang; Wang; Giannakis, 2019) that applies a fixed-point iterative method, while this chapter

employs a nonlinear numerical optimization method. The main advantage is solving the problem under a fully optimization framework, with a globally convergent method (Nocedal; Wright, 2006). An initial condition for the state $x_t^k = x_t^0$ starts the method at iteration k. The gradient function is calculated as:

$$\frac{\partial}{\partial x_j} J_{MCC}(x_t^k) = \sum_{i=1}^{m} \frac{e^{-\frac{\left(z_{i,t} - h_i\left(x_t^k\right)\right)^2}{2\sigma^2}}}{\sigma^2} \left(z_{i,t} - h_i\left(x_t^k\right)\right) H_{ij} - \frac{e^{-\frac{\left(x_{j,t}^k - x_{j,t-1}\right)^2}{2\sigma^2}}}{\sigma^2} \left(x_{j,t}^k - x_{j,t-1}\right) \quad (5.18)$$

where, $H_{ij} = \frac{dh_i(x_t^k)}{dx_j}$ the elements of the Jacobian matrix of the nonlinear measurement model calculated at a specific point x_t^k. This can be rewritten in a matrix representation as:

$$\nabla J_{MCC}\left(x_t^k\right) = \begin{bmatrix} I \\ H \end{bmatrix}^T D \begin{bmatrix} x_{t-1} - x_t^k \\ z_t - h(x_t^k) \end{bmatrix} \quad (5.19)$$

where, I is an identity matrix (nxn), and

$$D_{ii}(x_t) = \frac{e^{-\frac{\left(x_{i,t}^k - x_{i,t-1}\right)^2}{2\sigma^2}}}{\sigma^2},\ if\ i < n$$
$$D_{ii}(x_t) = \frac{e^{-\frac{\left(z_{i,t} - h_i(x_t^k)\right)^2}{2\sigma^2}}}{\sigma^2}, otherwise \quad (5.20)$$

The Hessian matrix is calculated by the following:

$$\nabla^2 J_{MCC}\left(x_t^k\right) = -\begin{bmatrix} I \\ H \end{bmatrix}^T D\left[I - R_p\right] \begin{bmatrix} I \\ H \end{bmatrix} + \sum_{i=1}^{m} D_{ii}(x_t) \frac{\left(z_{i,t} - h_i(x_t^k)\right)}{2\sigma^2} \frac{\partial^2 h_i(x_t^k)}{\partial x^2} \quad (5.21)$$

where,

$$R_{pii} = \frac{1}{\sigma^2}\left(x_{i,t}^k - x_{i,t-1}\right)^2,\ if\ i < n \quad (5.22)$$
$$R_{pii} = \frac{1}{\sigma^2}\left(z_i - h_i(x_t^k)\right)^2,\ otherwise \quad (5.23)$$

In order to create ascending directions (remind that the solution is the maximum Correntropy) the optimization method requires an approximation of the Hessian matrix (M_k), that keeps negative definite during the entire optimization process. This way, a good approximation candidate is given by (5.24), neglecting the second-order derivatives terms in (5.21).

$$\nabla^2 J_{MCC}\left(x_t^k\right) \approx M_k = -\begin{bmatrix} I \\ H \end{bmatrix}^T D\left[I - R_p\right] \begin{bmatrix} I \\ H \end{bmatrix} < 0 \quad (5.24)$$

With the gradient vector and the Hessian matrix approximation is possible to calculate the directions p_k to be followed by the maximization algorithm at each iteration k.

$$M_k \, p_k = \nabla \, J_{MCC}\left(x_t^k\right) \tag{5.25}$$

A backtracking algorithm is used to find step-size α_k and satisfy the Armijo condition (Nocedal; Wright, 2006) (the equivalent version for maximization problems), resulting in a sequence of iterations to find the optimal value $\hat{x}_t$.

$$x_t^{k+1} = x_t^k + \alpha_k p_k \tag{5.26}$$

The convergence is given by the difference of the state vector values in two successive iterations and a pre-specified tolerance (in this work $1.0E-6$). This iterative solution corresponds to the Prediction Step of the MCEKF. With the estimated state, the process covariance matrix is also updated in the Update Step of the MCEKF. It can be deduced as follows:

$$P_t = \left(P_{t-1}^{-1} + H^T R^{-1} H\right)^{-1} \tag{5.27}$$

5.3.4 Suppression of Suspect Samples and System Transitions through Parzen Window Adjustment

As shown in (Chen et al., 2017; Miranda; Santos; Pereira, 2009; Liu; Pokharel; Principe, 2006; Pesteh et al., 2019), the accuracy strongly depends on the strategy to choose the size of the Parzen windows or kernel bandwidth. The current practice on Correntropy-based estimation is a reduction of the kernel bandwidth towards an indifference of the suspect samples (for instance, outliers). This approach of successively reducing the size of the Parzen windows to achieve convergence when dealing with Correntropy is a method with widespread use, firstly proposed for training mappers under Correntropy and Entropy cost criteria (Erdogmus; Principe, 2002). This process, sometimes referred to as "kernel annealing", was also adopted for the power system state estimation (Mohiuddin; Qi, 2019; Wu et al., 2011). The goal was to suppress the effect of gross errors in the estimation, encompassed as a particular case of non-Gaussian errors.

The effect of reduced bandwidth kernel is well-known and conceptually proved in (Liu; Pokharel; Principe, 2006). In this situation, Correntropy reduces itself to the conditional expected value of the residue equal zero for that sample. However, this approach has some difficulties, from the algorithmic optimization point of view, one being that success depends on the iteration starting point, with flat start voltages no longer being an universal good point. Another one is the following: in order to create ascending directions (remind that the solution is the maximum Correntropy) the optimization method requires the Hessian (or its approximation) to be negative definite during the entire optimization process and at the optimal value of x_t. as shown in (5.24). Hence, $[I - R_p] > 0$, what implies the following lower bound for the Parzen window size:

$$\left(x_{i,t}^{k} - x_{i,t-1}\right)^2 < \sigma^2 \tag{5.28}$$

$$\left(z_i - h(x_t^k)\right)^2 < \sigma^2 \tag{5.29}$$

In practice, the reduction in the bandwidth cannot be severe, and is limited to a minimum size σ_{min}. Thereby such annealing strategy can create a bias on the suspect samples, even if it is small, in the final estimate.

This chapter introduces a new strategy to update the Parzen windows of suspect measurements, by *enlarging* the bandwidth of the suspect samples while maintaining the size of the non-suspect samples. Shrinking the kernel was meant to place an outlier in a region with metric similar to L_0 (indifference). But a Gaussian kernel with σ enough large, when compared with the remainder, reaches the same effect of indifference, because it has a smooth slope close to zero, so its value is almost constant across the domain, and it does not introduce any spikes in the optimization landscape. We take advantage of this property.

Note that, under the MCC Extended Kalman Filter, the suspect samples can be either caused by gross errors and miss calibrated measurements (assumption of white noise is false) or by large transitions on the system (encompassed in the state evolution concept). The effect is shown directly on the objective function of the MCEKF, given a suspect set $k \in S$:

$$\lim_{\sigma_k \to \infty} J_{MCC} = \sum_{i=1}^{m} e^{-\frac{\left(z_{i,t} - h_i(x_t)\right)^2}{2{\sigma_i}^2}} + \sum_{j=1}^{n} e^{-\frac{\left(x_{j,t} - x_{j,t-1}\right)^2}{2{\sigma_j}^2}} \tag{5.30}$$

Expanding the summation for the suspect samples:

$$\lim_{\sigma_k \to \infty} J_{MCC} = \cdots + \sum_{k \in S} \lim_{\sigma_k \to \infty} e^{-\frac{\left(z_{k,t} - h_i(x_t)\right)^2}{2{\sigma_k}^2}} + \ldots$$

$$\lim_{\sigma_k \to \infty} J_{MCC} = \cdots + \sum_{k \in S} e^0 + \ldots \tag{5.31}$$

And also,

$$\lim_{\sigma_k \to \infty} \frac{\partial}{\partial x_j} J_{MCC} = \cdots + \sum_{k \in S} \lim_{\sigma_k \to \infty} \frac{e^{-\frac{\left(z_{k,t} - h_i(x_t)\right)^2}{2\sigma^2}}}{\sigma^2} \left(z_{k,t} - h_i(x_t)\right) H_{kj} + \ldots$$

$$\lim_{\sigma_k \to \infty} \frac{\partial}{\partial x_j} J_{MCC} = \cdots + \sum_{k \in S} 0 + \ldots \tag{5.32}$$

The portion of the derivatives related to samples in the suspect set equals zero. Hence this does not affect at all the rest of the optimization process and estimated values. In this situation only the other measurements and the process equations are used to find the state estimates. Fig 8 shows an illustrative example of kernel density estimation and the effect of reducing and

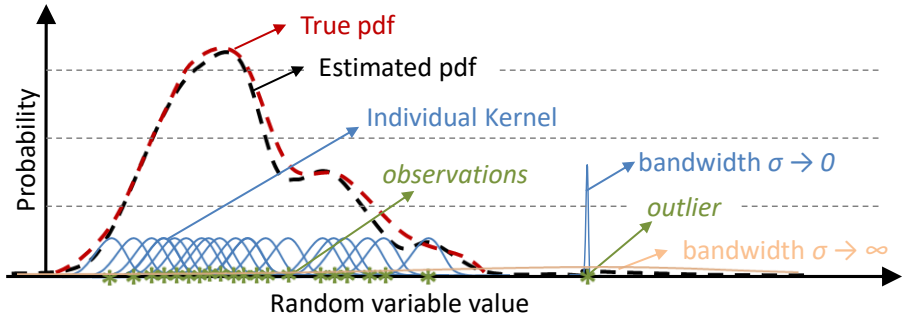


Figure 8 – Effect of different sizes of bandwidth (Parzen window size) for suspect sample on the kernel density estimator.

enlarging the Parzen window size for suspect samples. An outlier is associated with a large bandwidth and the other observations keep their initial window sizes. The effect of suspect samples with large window sizes spreads along the real line (from $-\inf$ to $+\inf$), and does not influence the estimated pdf.

If the detected event is a systemic transition, such as sudden load change, then the suspect set becomes the process noise equations. In this scenario, only the Parzen Windows related to the state space model are enlarged. A similar demonstration obtains the following equation in this situation:

$$\lim_{\sigma_j \to \infty} J_{MCC} = \sum_{i=1}^{m} e^{-\frac{\left(z_{i,t} - h_i(x_t)\right)^2}{2\sigma_i^2}} + \sum_{j=1}^{n} e^0 \tag{5.33}$$

The portion of the derivatives related to process equation on the state space then equal to zero. Hence this does not affect at all the rest of the optimization process and the estimated values. It has an effect of momentarily neglecting the state transition equations, thus the estimates rely only on the current observed values, similar to the snapshot WLS state estimator. In practice this is equivalent to breaking the time dependency on the Markov Chain, the same effect observed in (Massignan et al., 2019) for treating non-stationary events. Another situation is simultaneous process and measurement noise due to structural transitions, such as caused by topological and parameter errors. In such situations, both the process noise equations and the adjacent measurements of the affected part of the network become the suspect set and have their Parzen window sizes enlarged.

A proper method for detecting and identifying suspect measurements is required. This chapter will not treat these methods in detail, but can be obtained by a residual analysis, such as in (Coutto Filho; Souza; Guimaraens, 2014) for bad data detection, or with other techniques based on artificial intelligence such as in (Miranda et al., 2019) for system transitions. In this work, the standard normalized residual analysis (Coutto Filho; Souza; Guimaraens, 2014) triggers the identification of suspect samples for gross errors, while the system transitions are assumed as known. It is noteworthy that there are other and perhaps even more suitable ways to detect and

identify suspect samples in the context of Information Theory, such as using the Generalized Correntropy along with the Interior Point Method (Pesteh et al., 2019; Pesteh; Moayyed; Miranda, 2020).

The algorithm consists basically on the MCC Kalman Filter recursive equations and the Parzen window update strategy. The window size updating consists in multiplying the bandwidth by a large value (in this work 1.0e4).

Algorithm 1: MCEKF WITH PARZEN WINDOW UPDATE

Input : Network Data, Network Topology, Measured Data, x_0, P_0
Output : System State

Begin
  Initialize Parzen Windows σ_k =10.0
  Set x_{t-1} as x_0 and P_{t-1} as P_0
  Repeat
    Rescale the state space according to (17) and (18)
    Calculate x_t solving the MCC optimization in (21)
    Update posterior state covariance according to (30)
    Identify the suspect set S
    Update the Parzen Window size according to the type of event:
    Bad Data (gross error):
    For each z_k with $k \in S$ $\sigma_k = 10000\sigma_k$
    Systemic Transition:
    For each x_j in x_t $\sigma_j^t + 1 = 10000\sigma_k$
    Structural Transition:
    For each x_j in x_t $\sigma_j^{t+1} = 10000\sigma_k$
    For each z_k with $k \in S$ $\sigma_k^{t+1} = 10000\sigma_k$
    Update to next instant t
  Until $t < \Delta t$;
End

5.4 Simulation Methodology

Monte Carlo simulations were performed to evaluate and validate the proposed MCEKF. A sequence of load flow conditions creates the reference values for the state variables $x_{(j,t)}^{ref}$ and measurements $z_{(i,t)}^{ref}$. The simulation consists in including random noise in the reference load flow values to obtain the measured values and then perform state estimation (Singh; Pal; Vinter, 2009). The noise characteristic is given based on the precision of the measurement (2% for SCADA and 0.1% for PMUs) and the following equation obtains the measured values:

$$z_{i,t} = z_{i,t}^{ref} + u_i \left| z_{i,t}^{ref} \right| pr_i/3 \tag{5.34}$$

where u_i is the underlying pdf of the simulated noise added to the reference values. This equation provides the diagonal elements of the measurement noise covariance matrix $R_{ii} = pr_i|z_{(i,t)}|/3$, and pr_i is the metering device precision (Castillo et al., 2011).

Different instants and the respective conditions are monitored by PMUs with a 60 samples per second rate, while SCADA measurements are updated at a rate of a sample per second. Each sample represents a different instant t and the respective measured values. Thus, the simulation consists of a sequence of quasi-stationary conditions with the estimation process triggered by the latest information available, sometimes monitored by SCADA and others by PMUs, with different time scales and random noise. The Mean Absolute Error (MAE) evaluates the accuracy of the estimated state variables through (5.35).

$$\mathrm{MAE}_t = \frac{1}{n_{trials}} \sum_{trial=1}^{n_{trials}} \sum_{i=1}^{n} \left| \hat{x}_{i,t}^{trial} - x_{i,t}^{ref} \right|, \tag{5.35}$$

where $\hat{x}_{i,t}$ is the i-th state variable estimated. The MCEKF is then compared with two approaches based on the maximum likelihood principle: the WLS Snapshot; the traditional WLS Extended Kalman Filter.

5.5 Application in Transmission Systems

This section presents the application reults of the MCEKF method in a typical deployment of transmission systems monitored by SCADA system and PMUs.

5.5.1 Effect of Non-Gaussian Measurement Noise

The first test was performed with the IEEE14 test system to show aspects of non-Gaussian measurement noise, under a stationary condition, that is without process noise, using the measurement set presented in Fig. 9. A Monte Carlo simulation with the addition of non-Gaussian noise, in all SCADA and PMU measurements, evaluates the performance of the MCEKF, during a sequence of 600 samples (equivalent to 10 seconds). The noise characteristic added to the ference values to obatain measured values in the simulation follows the Gaussian mixture in (5.36). This emulates a deviation from standard calibration (white noise with Gaussin distribution).

$$u_i \sim 0.7\mathcal{N}(0,1) + 0.2\mathcal{N}(3,3) + 0.1\mathcal{N}(0,20) \tag{5.36}$$

Fig. 10 shows the performance metric over time in order to illustrate the effect of including the temporal aspect in the estimation. Note that the snapshot approach does not improve the estimation over time since it performs an independent estimation each instant only with the respective sampled values, either from SCADA or PMUs. The traditional WLS EKF does present an improvement by including the temporal relation of the state; however, the MCEKF approach increases even further the accuracy, since it is more suitable to treat non-Gaussian noise characteristics.

Table 1 presents the comparison in terms of the overall MAE performance index. A different set of fixed values for the Parzen Window sizes is also compared. For the smaller

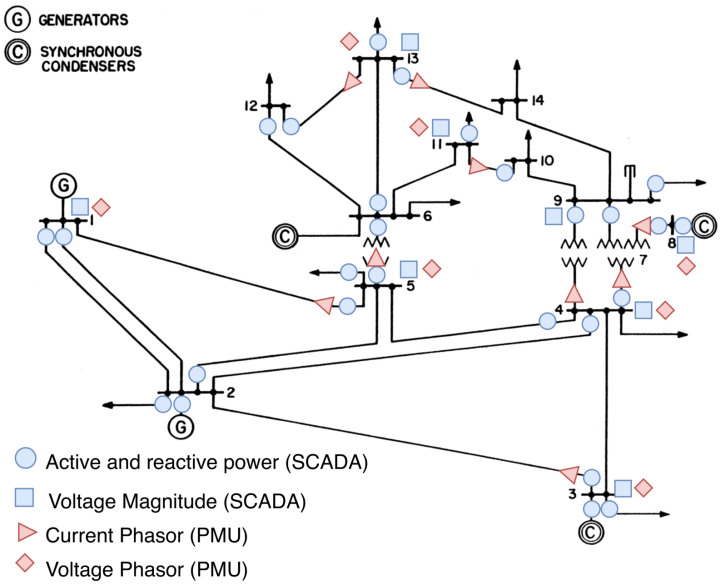


Figure 9 – IEEE14 test system and respective SCADA and PMU measurement set for the simulations.

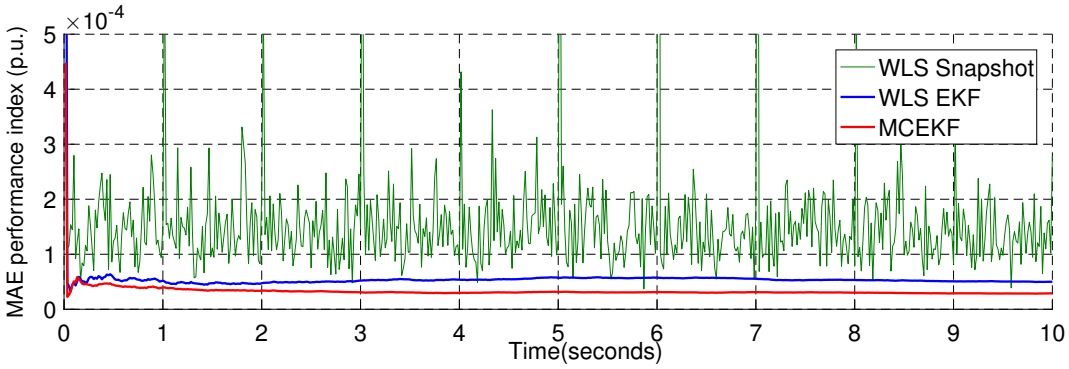


Figure 10 – MAE Performance index over the time.

window size, the lower bound is reached, and the estimation loses its performance. Another aspect is that by choosing higher Parzen Window sizes, the MCEKF gets closer to the traditional WLS based EKF, a known property on Correntropy estimation that it becomes the L_2-norm (similar to the WLS criterion) for high Parzen window size. Only with the Parzen Window update strategy, by suppressing the suspect set based on the largest residual, that more accurate estimation is achieved in this case.

5.5.2 Effect of Non-Gaussian Process Noise

The second test intends to evaluate the effect of process noise on the estimation. Process noise was added in the form of random load variation. In order to obtain a sensitivity of the effect of process noise, the percentage of load variation was increased gradually on this test, from 0.0% to 10%. Fig. 11 presents the accuracy in terms of the empirical cumulative distribution function

Table 1 – Comparison of the MAE performance index for the real and imaginary parts of the state variables

Method	Parzen Window	MAE performance index (p.u.)	
		Real V	Imag. V
WLS Snapshot	-	2.5121 E-04	0.5899 E-04
WLS EKF	-	0.9387 E-04	0.1941 E-04
MCEKF	σ_i = 0.1 p.u.	14.6335 E-04	23.0463 E-04
	σ_i = 1.0 p.u.	1.6242 E-04	0.2897 E-04
	σ_i = 10.0 p.u.	0.9119 E-04	0.1813 E-04
	σ_i = 100.0 p.u.	0.9387 E-04	0.1941 E-04
	With Parzen Update Strategy	0.5545 E-04	0.1024 E-04

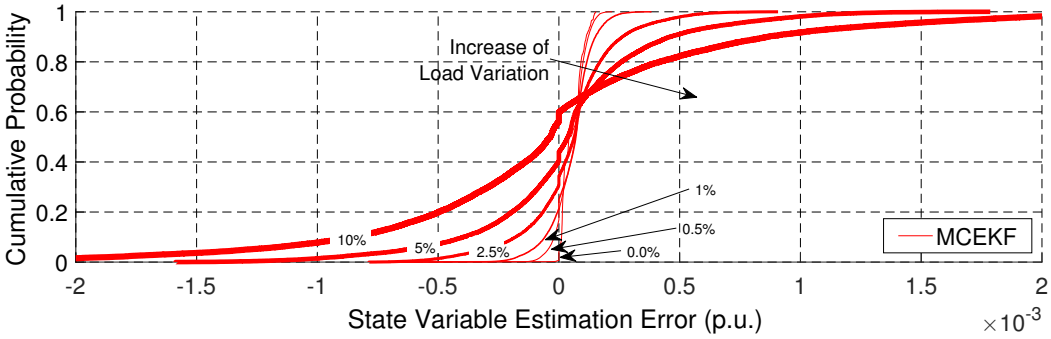


Figure 11 – Cumulative probability distribution of the estimation error for the MCEKF under different load variation (process noise) levels.

Table 2 – MAE performance index for the state variables with different levels of load variation

Load Variation	MAE performance index (p.u.)		
	WLS Snapshot	WLS EKF	MCEKF
0.5 %	1.5510 E-04	1.0368 E-04	0.6748 E-04
1.0 %	1.5510 E-04	1.2205 E-04	0.8533 E-04
2.5 %	1.5509 E-04	1.8705 E-04	1.4698 E-04
5.0 %	1.5507 E-04	3.0779 E-04	2.6412 E-04
10.0%	1.5503 E-04	5.6942 E-04	5.2730 E-04

of teh estimation error obtained by the MCEKF. Table 2 summarizes the results for the three estimators.

The effect of the process noise is a progressive increase in the estimator variance, up to a point that the most reliable estimator becomes the WLS snapshot, instead of the WLS EKF and MCEKF approaches. It is noteworthy that large variations up to 5% of the load, in the short interval of time between two consecutive PMU samples (1 cycle), can be considered as a very abrupt and abnormal condition. Furthermore, such a scenario implies on a sensitive reduction of the premise that the last observation is a good estimation for the current one, the underlying hypothesis of the Markov Chain in the tracking state space model in (4). In a typical transmission system, such scenario can be related to very large load variations or contingencies. In modern

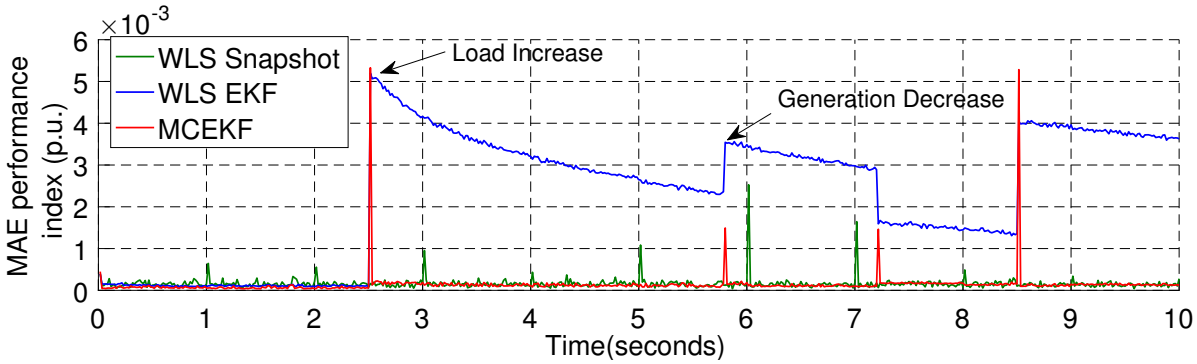


Figure 12 – MAE Performance index over the time.

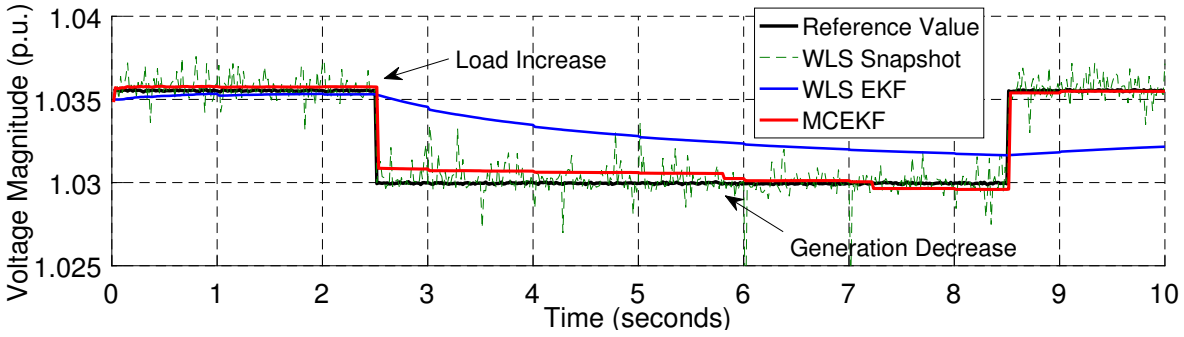


Figure 13 – Estimated voltage magnitude at node #14 during a load increase from t=2.5s to t=8.5s and generation decrease from t=5.8 to t=7.2.

power systems, with the increase of intermittent distributed energy resources, such scenario can occur more frequently. However, despite random variation may occur, it is more likely that such large and abrupt changes happen less often in the time spam of the tracking state estimation computing horizon.

This motivated the pursuit for a window size tuning strategy able to suppress the effect of such large and abrupt variations. The proposition is suppressing the negative effect that such a systemic transition has in the estimation accuracy by enlarging the Parzen window size of the process noise equation, as shown in Section 5.3.4. The reasoning behind such strategy is that an abrupt change on the system state can be interpreted as a particular case of process noise with a heavy tailed pdf. Hence the process equations become the suspect set and their Parzen window size are enlarged. In order to evaluate this strategy, we consider 0.5% of random load variation and two sudden load changes in the simulation: (1) an increase of all loads by 10% from t=2.5s to t=8.5s; and (2) a decrease of generation by 30% at node #2 from t=5.8 to t=7.2. Fig. 12 presents a comparison of performance of the estimators regarding this scenario with a sudden load change. The effect is also compared for one of the estimated voltage magnitudes at nodes #2 and #14 in Fig. 13.

The effect of updating the Parzen windows related to the previous state consists on pushing momentarily the MCEKF towards the WLS snapshot. It counters the effect of the system transition, by breaking the temporal relation of the Markov Chain within the method. Otherwise it would propagate a false assumption that the previous instant is a good approximation for the next one. Besides, it keeps a recursively improvement for estimating the steady state when the transition ends.

5.5.3 Suppression of Gross Errors

This third test, performed with the IEEE 118-bus test system, addresses the effect of gross errors and different measurement noise pdfs. In addition, the simulation considers process noise with uniform load variation of 0.5%. The test assumes the non-Gaussian noise for the SCADA measurements in 5.36 and different characteristics for the PMUs according to the proportion and pdfs below: 80% of voltage phasor follows 5.37 and the remaining as 5.38; and 50% of the current phasor follows 5.39 and the remaining as 5.40.

$$u_i \sim 0.9\mathcal{N}(0,1) + 0.1\mathcal{N}(0,20) \tag{5.37}$$

$$u_i \sim 0.6\mathcal{N}(0,1) + 0.3\mathcal{N}(0.5,0.5) + 0.1\mathcal{N}(0,20) \tag{5.38}$$

$$u_i \sim Gamma(0.2,4) \tag{5.39}$$

$$u_i \sim Unif(-1,1) \tag{5.40}$$

The simulation consists also in a SCADA and PMU observable metering system with a 5s horizon. It includes the following gross errors to illustrate the effect of bad data:

1. Addition of 30 standard deviations in PMU measurement V_{31} (voltage phasor measurement at node 31) at t = 0.6s;

2. Reduction of 30 standard deviations in PMU measurement I_{63-59} (current phasor measurement at branch between nodes 63 and 59) from t = 1.2s to t = 1.5s;

3. Addition of 30 standard deviations in SCADA measurement P_{76-118} and Q_{76-118} (active and reactive power flow measurements at branch between nodes 76 and 118) at t = 2s;

4. Addition of 30 standard deviations in PMU measurement I_{63-59} and I_{49-66} (current phasor measurement at branch between nodes 63 and 59, and at branch between 49 and 66) from t = 2.5s to t = 2.8 s;

5. Addition of 30 standard deviations in SCADA measurement V_{100} (voltage magnitude measurement at node 100) at t = 3s;

6. All the above simultaneously from t = 3.5s to t = 4.5s.

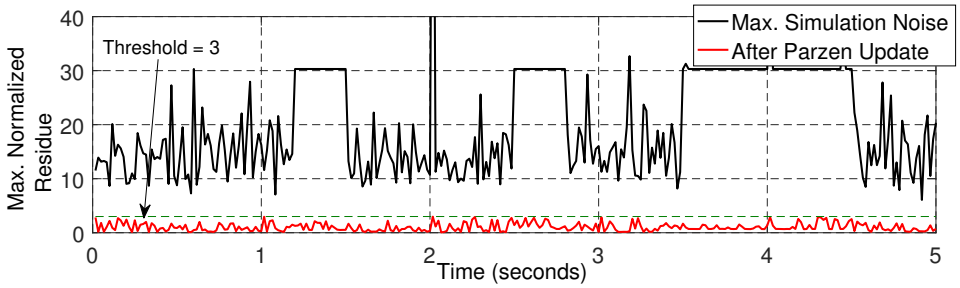


Figure 14 – Largest normalized residue after updating the Parzen window size for supression of the effect of gross errros in the MCEKF.

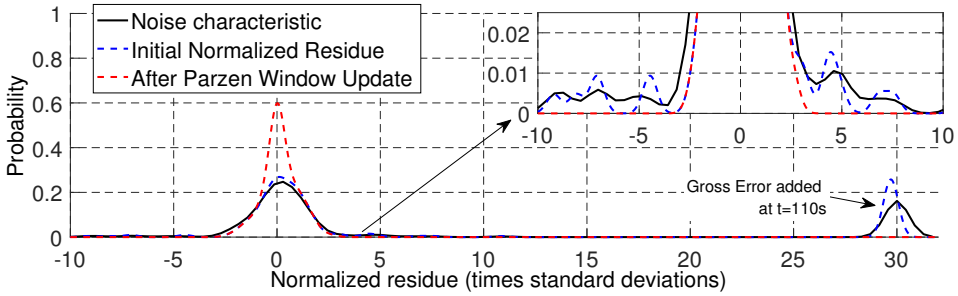


Figure 15 – Heavy tailed measurement noise with the addition of gross error and respective normalized residue obtained by the MCEKF.

The MAE performance index kept the same accuracy pattern as the previous tests, with an overall value of 3.5317 E-05. In order to show the effect of the Parzen window update, Fig. 14 shows the largest normalized residue, initially before any size update, and after all updates are done. Note that in the final estimates all normalized residues are less than the chosen threshold, thus with a reduced influence in the final estimate. It is noteworthy there are instants where the measurement noise is enough to trigger the Parzen window update, since the noise pdfs have heavy tailed distributions.

To illustrate the effect of the Parzen window update in the normalized residue, Fig. 15 shows the noise characteristic and respective calculated normalized residue for measurement V31 with the MCEKF. Not only are the gross errors properly suppressed by the Parzen window updating strategy, but also heavy tailed noise from the underlying measurement pdf. This shows that the method can sustain accurate estimates, in relatively large networks, even in the presence of a series of bad data contaminating the measurement vector.

5.5.4 Accuracy during Structural Transitions

This section evaluates the effect of structural transitions, that is, abrupt changes on the topology of the system that has a large impact on both state and measurement model. The objective is to illustrate the state evolution concept and the state tracking through normal and abnormal network conditions, with the Brazilian interconnected system (BR107). The network

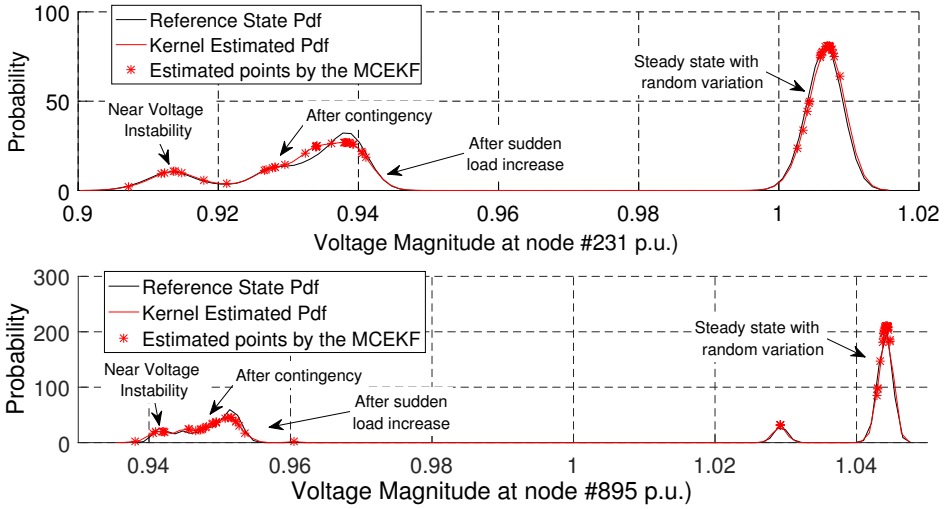


Figure 16 – Tracking state evolution pdfs for nodes 231 (minimum voltage) and 895 (terminal from the transmission line that switched off).

consists of 107 nodes along with a hypothetical set of SCADA and PMUs observable sets. The simulation emulates a voltage instability situation. It starts with 1s of steady state with random load variation of 0.5%, followed by an increase of load at a rate of 0.5% per second. Then a sudden load increase of 5% in all loads occurs at t=2.5s. After more 0.5s a contingency occurs in one of the 500 kV transmission lines that connect the south to the southeast region. Following the contingency, the load ramp increases to 1.0% per second, and voltage instability occurs in less than one second.

Fig. 16 illustrates the empirical pdf for the estimated voltage magnitude at two substations of the system. As it can be seen, the MCEKF provides an accurate estimation of the state pdf that can be used in further analysis, for instance to calculate probabilistic voltage stability margins in real time. Besides, the MCEKF presents a lot of flexibility, once by properly choosing the Parzen window sizes suppressed the negative effects of both systemic and structural transitions on the estimator accuracy. Such flexibility motivates the further exploration of improved tuning methods for the Parzen Windows.

5.5.5 Computational Aspects

Finally, the computational aspects of the previous simulations are presented in Table 3. The tests were performed using a microcomputer with a Core i7 3.60 GHz, 16GB RAM with C programming language. The MCEKF shows good convergence characteristics in the power system state estimation problem. However, computational burden must be taken into account for real time applications, especially since PMUs can have sampling rate up until 1 per cycle. The times shown in Table 3 were obtained without any concern on efficient programming and not using sparse matrix techniques or other computing efficiency resources, so they should be seen only as indicative of the feasibility of the technique.

Table 3 – Computational Aspects of the MCEKF

Test System	Mean Processing Time (ms)	Iterations with SCADA meas.	Iterations with PMU meas.
IEEE14	31.0	4 to 5	2 to 3
IEEE118	360.0	5 to 7	2 to 3
BR107	313.0	5 to 7	2 to 3

5.6 Concluding Remarks

This work proposed a tracking state estimator based on an Extended Kalman Filter under a Maximum Correntropy principle (MCEKF) that deals with both SCADA and PMU measurements. A new update strategy for Parzen window suppresses the effect of suspect samples that can be related to gross errors, process noise as well as system's transitions. The importance of expanding the static state estimation problem to include temporal aspects is emphasized in this work.

Even though many efforts have been made towards dynamical models, it is clear that currently implemented static state estimators can benefit a lot from simple modifications that leads to the tracking model. However, special care must be taken in order to account for possible system transitions, such as systemic and structural transitions. This was accomplished in this work by treating both measurement and process noise within a non-Gaussian kernel density estimation.

In this sense, the MCEKF outperforms the WLS snapshot, since it is based on a state space model, and also the conventional WLS Extended Kalman Filter for treating non-Gaussian process and measurement noise. With the proposed concept of state evolution, the state variables also gain a non-parametric interpretation that can be further used in post-processing evaluation of the condition of the network under a probabilistic framework.

The method still depends on a proper suspect sample detection and identification process. In this sense, future work comprises the development of feature extraction methods to trigger automatically the Parzen window updating strategy, based in a broader notion of the residual analysis with Bayesian Inference concepts. An important direction is also to increase the robustness of the method by adopting as solver the Interior Point method, instead of relying on Newton-Raphson iterations. Finally, the model should be extended for the dynamical estimation, treating generators, loads, and controllers to evaluate detailed transient events.

6 BAYESIAN INFERENCE APPROACH FOR INFORMATION FUSION IN DISTRIBUTION SYSTEM STATE ESTIMATION

The content of this chapter have been published to the following Journal:

- Bayesian Inference Approach for Information Fusion in Distribution System State Estimation," *IEEE Transactions on Smart Grid*, vol. 13, no. 1, pp. 526-540, Jan. 2022";

6.1 Introduction

This chapter proposes a three-phase DSSE, based on a Bayesian inference approach, which is able to perform information fusion from different kinds of sources that are usually available in DSs, extending the initial work presented in (Massignan et al., 2019). Such sources of information may consist of load profiles and forecasts (referred to as pseudo measurements), passive buses with no generation or load (virtual measurements) and different kinds of sensors, as smart meters, SCADA measurements and PMUs. They are installed across different parts of a DS, from the low voltage consumers up to the primary substation. Despite a consolidated topic in statistics, few works explore the Bayesian framework in power system state estimation, for instance in the topology estimation problem (Lourenco; Costa; Clements, 2004).

Regarding DSSE, the work in (Pegoraro et al., 2017) exploits the flexible Bayesian Inference background in order to provide estimates under non-Gaussian error distributions. A recent work in (Dobbe et al., 2020) presents a Bayesian approach for linear three-phase state estimation that works together with a forecasting stage based on Gaussian process. In a different perspective, (Zhou et al., 2020a) explores the harmonic state estimation problem with sparse Bayesian learning from Smart Meter and PMU data. This way, the contributions of this chapter reside on:

- the proposition of a three-phase DSSE based on a Bayesian fusion procedure to deal with nonlinear models of different kinds of measurements and their different temporal scales with the concept of sampling layers;
- the application of an orthogonal method associated with sparsity treatment and an ordering method into the Bayesian information fusion problem, to increase computational and numerical efficiency to deal with ill-conditioned DSs;
- the exploration of credibility intervals concept to track state transitions caused by abrupt system changes and to test if the vicinity of the prior distribution is a good approximation for the state.

A few theoretical topics, presented in previous sections, will be revisited in the following sections to provide a quick background of the fundamental details.

6.2 Information Fusion with Bayesian Inference for DSSE

The static state estimation problem for a distribution network with m measurements and n state variables consists of a nonlinear measurement model that can be formulated as (Abur; Gómez-Expósito, 2004):

$$z = h(x) + e, \tag{6.1}$$

where z is the $(m \text{ x } 1)$ measurement vector composed by the measured values, virtual measurements and pseudo measurements; x is the $(n \text{ x } 1)$ vector of state variables (usually the complex nodal voltages in polar coordinates); $h(x)$ is the $(m \text{ x } 1)$ set of nonlinear equations that relates the measurements with the state variables; and e is the $(m \text{ x } 1)$ noise vector assumed as independent normally distributed random variables, with zero mean and known covariance matrix R.

An exploration of the information that builds the measurement vector shows there is a lack of synchronism among the measurement updating (Δt), ranging from a few miliseconds (for PMUs) up to minutes (for smart meters). Thereby the measurements can be separated in subsets according to their respective sampling rates. Each subset is grouped as a sampling layer (SL), which is characterized by a common updating time interval ($\Delta t_{z_i} = \Delta t_{SL}$), given by (6.2). In addition to a common sampling rate, typically each sampling layer consists of measurements with similar relative accuracy.

$$SL_i := \{z_i \in z \mid \Delta t_{z_i} = \Delta t_{SL_i}\} \tag{6.2}$$

The following sampling layers are proposed to deal with the typical sources of information in modern DSs:

- *Pseudo measurement Layer*: includes historical data and typical load curves, as well as the virtual measurements (nodes without load or generation), that guarantee observability of the network (updated from 10 to 30 minutes);
- *Smart Meter Layer*: includes meters installed in low voltage consumers (updated from 1 to 15 minutes);
- *SCADA Layer*: encompasses the traditional SCADA system (updated from 2 to 10 seconds up to 1 minute);
- *PMU Layer*: relates to fast syncronized phasor measurements (updated at each cycle up to 100 miliseconds).

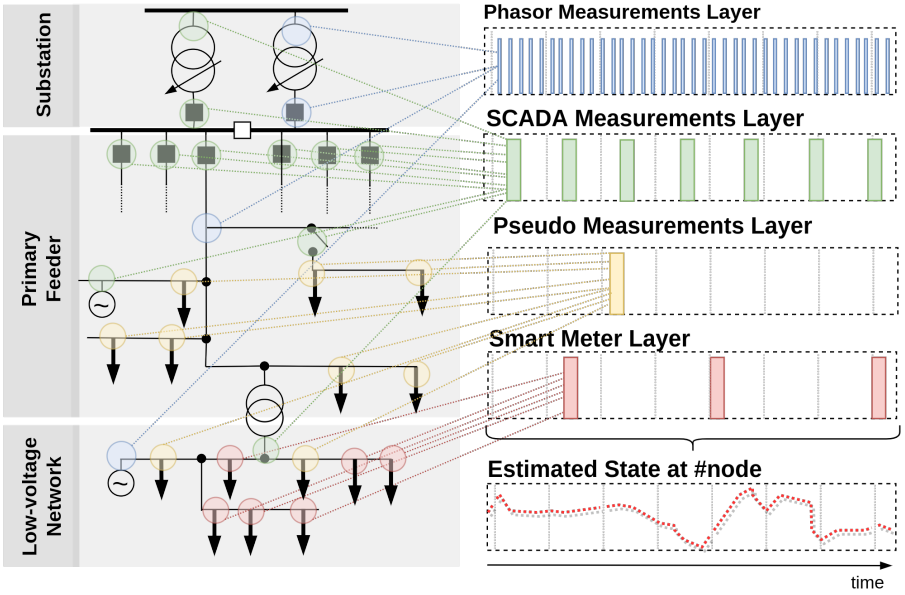


Figure 17 – Illustration of the location of different kind of measurements in DSs and the respective sampling layers according to updating time interval. The proposed sampling layer concept decomposes the measurement set according to their respective sampling rates, disregarding their location at the distribution network. An example of an estimated state illustrates the variations that may occur between measurement updates from slower sampling layers.

Fig. 17 illustrates the sampling layers of a typical set of measurements present in DSs. The challenge consists of treating a vast number of measurements, spread across all parts of a DS, and providing a reliable state vector for the DSO, denoted as $\hat{x}$, while dealing with the temporal aspect of each sampling layer. Due to the real time aspect of the DSSE, it is also important to perform it in a tractable computational time.

A two-port component model provides a general three-phase measurement model for a diverse set of DS components, such as distribution lines, transformers, voltage regulators and shunt capacitors. A detailed description of the three-phase component model was presented in Chapter 4.

6.2.1 Multi-Layered Posterior Estimation

The Bayesian Information Fusion consists on a multiple stage state estimation, triggered by each new measurement update from the respective sampling layers. It decomposes the measurement model into each separated sampling layer, and includes the relation between two layers using a hierarchical model for the state variables as follows:

$$x_{SL} = \hat{x}_{SL-1} + \omega_{SL-1} \tag{6.3}$$

$$z_{SL} = h_{SL}(x_{SL}) + e_{SL} \tag{6.4}$$

where, x_{SL} is the state vector at the current sampling layer, $\hat{x}_{SL-1}$ is the estimated value at the previous sampling layer with slower updating rate, z_{SL} is the measurement vector of the respective sampling layer SL, $h_{SL}(x_{SL})$ is the respective measurement model, e_{SL} is the measurement error vector associated with the measurements in the sampling layer, and ω_{SL} is a random component associated with the uncertainty between the time intervals of two sampling layers.

The introduction of the relation between two different layers adds a random component to the state. From the Bayesian perspective, it consists of introducing a probabilistic model for the state vector, that is, modeling the state of the network as a random variable. It changes the traditional overview of the estimation process as a "snapshot" of the network condition towards a probabilistic characterization of the state. It encompasses the fact that different state values can occur between updates of the slower measurements, observed by the faster ones. A similar approach has been applied in the tracking state estimation problem, as a state space model formulation solved by the Kalman Filter (Massignan et al., 2019; Zhao et al., 2019).

Following the Bayesian perspective, each new set of observations of a sampling layer then updates the posterior distribution for the state. The likelihood function is given by the assumed noise model for the measurements. And the prior model for the state variable is given by the previous sampling layer posterior distribution, a hierarchical prior model. The Bayes' Theorem obtains directly the posterior probability function for the state variables, according to the following:

$$f_{X|Z}(x_{SL}|z_{SL}) = \frac{f_{Z|X}(z_{SL}|x_{SL})f_X(x_{SL})}{f_Z(z_{SL})} \tag{6.5}$$

where, $f_{X|Z}(x_{SL}|z_{SL})$ is the conditional probability function of the state given the measurements in each sampling layer, $f_{Z|X}(z_{SL}|x_{SL})$ is the likelihood function of the sampling layer according to the measurement model, $f_X(x_{SL})$ is the prior distribution in the hierarchical model, and $f_Z(z_{SL})$ is the measurements probability of occurrence, a constant value that scales the posterior probability function and is often neglected.

The estimation process in each layer is triggered as soon as the respective measurements become available. It updates the state by a Maximum a Posteriori (MAP) estimation given by (6.6). The estimation result of each layer is then used as the prior distribution for the state variables in the next layers, a Bayesian hierarchical model (Congdon, 2007).

$$\hat{x}_{SL} = \arg\max_x f_{X|Y}(x_{SL} \mid z_{SL}) \tag{6.6}$$

As shown in (Massignan et al., 2019), this is much alike the FASE procedures, however, in this case it only relates different sampling layers instead of instants in a sequential Markov process. Thus the temporal relations within a specific sampling layer are discarded, what enables maintaining accuracy during systems changes, that is, when leaps in the Markov process occur.

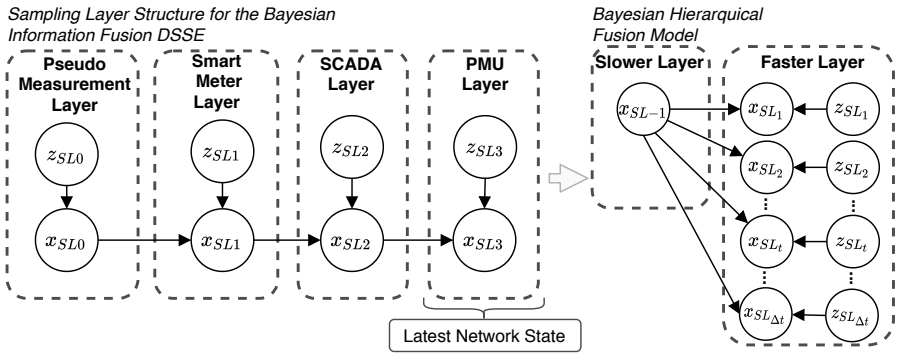


Figure 18 – Probabilistic network representation of the multi-layered posterior estimation DSSE to deal with measurements with different sampling rates (pseudo measurements, smart meters, SCADA and PMUs). Each sampling layer deals only with measured information from its group, and the final estimate is made available for the DSO and to the following sampling layers.

This difference is illustrated in Fig. 18 with a probabilistic network representation of the multi-layered information fusion process. The triggering strategy for the MAP estimate uses the latest available information of each sampling layer and the prior state distribution obtained from previous layers, as illustrated in Fig. 19.

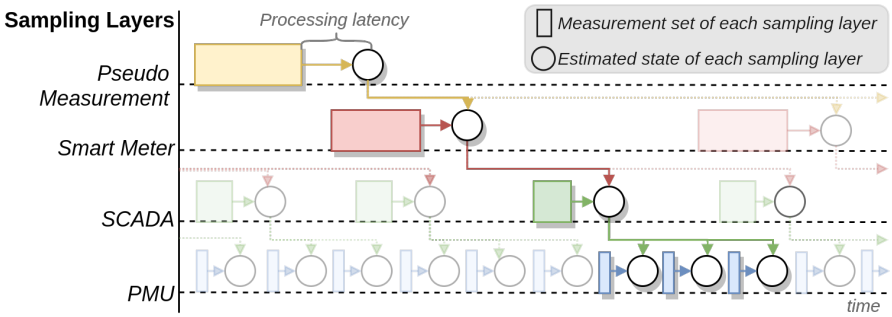


Figure 19 – Illustration of the sampling layer triggering strategy for the multi-layered posterior model. After a sampling layer is processed, the result is propagated in an updated informative prior distribution to the subsequent faster sampling layer. As soon as a new sample is available, a MAP estimation is performed to update each sampling layer's state variables. The faster layer obtains the current state of the network.

6.2.2 Multivariate Gaussian Prior Model for DSSE

In order to provide a fast solution for the state estimation problem, this work assumes a Multivariate Gaussian prior distribution for the state variables, as assumed for the measurement error in (6.1). This way is possible to maintain a tractable computational for the MAP estimate, described in (6.6), a crucial aspect for real time applications. If the measurement model were linear, such as in the case of PMUs, this approach would provide a closed-form solution for the posterior distribution.

$$Prior: \quad x_{SL} \sim \mathcal{N}(\hat{x}_{SL-1}, P_{SL-1}) \tag{6.7}$$

$$Likelihood: \quad z_{SL} \mid x_{SL} \sim \mathcal{N}(h_{SL}(x_{SL}), R_{SL}) \tag{6.8}$$

where P_{SL-1} is the state covariance matrix estimated at the previous sampling layer and R_{SL} the measurement covariance matrix of the current sampling layer, a sub-matrix of R in the original measurement model. Such prior model consists of a informative prior since it is based on knowledge obtained from previous layers. Non-informative priors may be used, but care must be taken since this increases the estimator variance.

It is noteworthy that many efforts have been made towards non-Gaussian noise characteristic in the state estimation problem (Gol; Abur, 2015; Zhao; Mili, 2018a; Massignan; London Jr; Miranda, 2020). Despite the Bayesian Inference framework can encompass different probabilistic models, this work will only attain in a Gaussian prior model. The reason is that dealing with non-Gaussian models, for the prior or the likelihood function, the Bayesian framework would require special numerical integration methods, such as the Metropolis Hastings or Gibbs sampling methods (Congdon, 2007; Pegoraro et al., 2017). This would increase the computational burden since so far such methods have not been scaled to deal with the practical size of DSs in real time.

The conjugate model assumes an approximation of the actual state behavior, at the vicinity of the results from previous sampling layers, a trade-off between accuracy and computational efficiency. Since it is an informative prior, whose values are updated as soon as new information is gathered from the previous sampling layers, such approximation tends to follow the system behavior. However, if abrupt changes occur, such hypothesis starts to lose accuracy and may hamper the estimation accuracy. This work proposes a credibility interval to track such abrupt transitions as an event detection method. It assesses if the prior distribution continues to provide good information or not, further discussed in section III.

The posterior distribution for the above conjugate model can be written as the following, by directly applying the Bayes' theorem:

$$f_{X|Z}(x_{SL}|z_{SL}) \propto \frac{1}{\sqrt{(2\pi)^m |R_{SL}|}} \prod_{i=1}^{m} e^{-\frac{(z_i - h_i(x))^2}{R_{SI,ii}}} \frac{1}{\sqrt{(2\pi)^n |P_{SL-1}|}} \prod_{j=1}^{n} \prod_{k=1}^{n} e^{-(x_{SL,j} - \hat{x}_{SL-1,k}) P^{-1}_{SL-1,jk} (x_{SL,k} - \hat{x}_{SL-1,j})} \tag{6.9}$$

The above expression can then be simplified to obtain the MAP estimate by dismissing the constant values and exploring the posterior distribution's kernel.

$$\hat{x}_{SL} = \arg\max_{x} e^{-[z_{SL} - h_{SL}(x_{SL})]^T R^{-1}_{SL} [z_{SL} - h_{SL}(x_{SL})]} e^{-(x_{SL} - \hat{x}_{SL-1})^T P^{-1}_{SL-1} (x_{SL} - \hat{x}_{SL-1})} \tag{6.10}$$

With this parametric hierarchical model to perform inference about the state variables in each sampling layer, the MAP estimate is obtained by the following nonlinear minimization problem:

$$\begin{aligned}\min_x J_{MAP}(x) =& (z_{SL} - h_{SL}(x_{SL}))^T R_{SL}^{-1} (z_{SL} - h_{SL}(x_{SL})) \\ &+ (x_{SL} - \hat{x}_{SL-1})^T P_{SL-1}^{-1} (x_{SL} - \hat{x}_{SL-1})\end{aligned} \tag{6.11}$$

The solution is obtained trough the iterative Modified-Newton method (Nocedal; Wright, 2006). An initial condition for the state variable $x_{SL} = x^k$ starts the method at iteration $k = 0$. For the first sampling layer it consists of the flat voltage profile. The following sampling layers may use the results from the previous one as initial condition for the iterative procedure (known as *hot-start*). The gradient function is calculated as:

$$\nabla J_{MAP}(x^k) = -(H^T R_{SL}^{-1} (z_{SL} - h_{SL}(x^k)) + P_{SL-1}^{-1} (x^k - \hat{x}_{SL-1})) \tag{6.12}$$

where H is the Jacobian matrix of the measurement model evaluated at each iteration and respective state vector value x^k. The Modified-Newton requires only a positive definite approximation for the Hessian matrix, obtained by the derivative of the Jacobian matrix and neglecting the higher order terms:

$$\nabla^2 J_{MAP}(x^k) \approx M^k = (H^T R_{SL}^{-1} H + P_{SL-1}^{-1}) \tag{6.13}$$

With both the gradient and the approximate Hessian matrix, the following linear system provides the direction p^k to update the state variables values. A backtracking algorithm is used to find the step-size α^k and to guarantee Wolfe's sufficient decrease conditions (Nocedal; Wright, 2006), providing the iteration update in (6.15).

$$M^k p^k = -\nabla J_{MAP}(x^k) \tag{6.14}$$

$$x^{k+1} = x^k + \alpha^k p^k \tag{6.15}$$

The convergence is given by the difference of the state vector values in two successive iterations becoming smaller than a tolerance ($1.0\text{x}10^{-6}$). The final vector x^k corresponds to the MAP estimate for the sampling layer $\hat{x}_{SL}$.

Regarding convergence, the matrix M^k is always full rank if the prior covariance matrix P_{SL-1}^{-1} is full rank. Also, since it is a sum of covariance matrices, M^k is a positive definite matrix. Therefore, the unique requirement for the method to provide an estimation is that the slowest sampling layer should be observable, which is accomplished by the virtual and pseudo measurements in the first layer. Another aspect that can be exploited are scaling factors between

two sampling layers, as done in (Massignan et al., 2019) to provide more flexibility to the MAP procedure when incorporating faster measurements.

There are works that already use the notion of a prior information in the estimation process, such as in (Simões Costa; Albuquerque; Bez, 2013). However, these works often assume a diagonal matrix for the prior model and do not relate it to the MAP problem. The posterior state covariance can also be calculated with (6.16), and is provided as the prior covariance matrix in the next sampling layer. This approach captures correlations that naturally arise from the electrical connection between nodes in the network, by fully representing off-diagonal elements computed with previous information, further increasing observability of sampling layers with few measurements.

$$P_{SL} = (H^T R_{SL}^{-1} H + P_{SL-1}^{-1})^{-1} \tag{6.16}$$

6.2.3 Application of an Orthogonal Method in the Bayesian Information Fusion Problem

To counter the associated ill-conditioning of DSs, a new strategy is proposed for the Bayesian information fusion problem. The method consists on first using an orthogonal factorization of the following matrix according to (6.17), obtaining the orthogonal matrix $\mathscr{Q}$ and the upper triangular $\mathscr{R}$.

$$\begin{pmatrix} P_{SL-1}^{-1/2} \\ R_{SL}^{-1/2} H \end{pmatrix} = \mathscr{Q}^T \mathscr{R} \tag{6.17}$$

With this orthogonal factorization, it is possible to rewrite the iterative equation (6.14) for the posterior estimation as:

$$\mathscr{R}^T \mathscr{Q} \mathscr{Q}^T \mathscr{R} p^k = \mathscr{R}^T \mathscr{Q} \begin{pmatrix} P_{SL-1}^{-1/2}(x^k - \hat{x}_{SL-1}) \\ R_{SL}^{-1/2}(z_{SL} - h_{SL}(x^k))) \end{pmatrix} \tag{6.18}$$

Since $\mathscr{Q}\mathscr{Q}^T = I$, where I is the identity matrix, and multiplying both sides by $(\mathscr{R}^T)^{-1}$, we obtain :

$$\mathscr{R} p^k = \mathscr{Q} \begin{pmatrix} P_{SL-1}^{-1/2}(x^k - \hat{x}_{SL-1}) \\ R_{SL}^{-1/2}(z_{SL} - h_{SL}(x^k))) \end{pmatrix} \tag{6.19}$$

Overall, the proposed algorithm, based on the orthogonal factorization, results in a quadratically smaller condition number than that of the correspondent original fusion problem. Stable numerical methods haven been pointed in the state estimation literature as a crucial computational enhancement (Machado; Azevedo; Monticelli, 1991; Amerongen, 1991; Freitas; Simões Costa; Miranda, 2020). This kind of orthogonal formulation also enables the application on severe ill-conditioned three-phase DSs and in the Bayesian fusion approach. Besides, by

exploring (6.18) it can be shown that $\mathscr{R}^T\mathscr{R} = P_{SL}^{-1}$. This promptly enhances the integration between different sampling layers, since $P_{SL}^{-1/2} = \mathscr{R}$. A sparsity treatment is employed to improve even further the computational efficiency and reducing *fill-ins* in the factorization process through sparse matrix ordering, by the Approximate Minimum Degree (AMD) method (Hebling et al., 2020). This work uses the sparse Multifrontal QR method of the *SuiteSparse* C-library (Davis, 2011b), an efficient implementation for large scale sparse systems.

6.3 Convergence Analysis and Observability Requirements

The proposed MAP estimation in (6.6) consists of a nonlinear unconstrained optimization problem. Two requirements are needed to guarantee convergence of the Modified-Newton method(Nocedal; Wright, 2006). The first is to ensure that the method provides only descent directions, that is:

$$\left\langle \nabla J_{MAP}(x^k)^T, p_k \right\rangle < 0 \tag{6.20}$$

As a consequence:

$$\nabla J_{MAP}(x^k)^T (M^k)^{-1} \nabla J_{MAP}(x^k) > 0 \tag{6.21}$$

Thereby, the approximation for the Hessian matrix M^k must always be definite positive, $(M^k)^{-1} > 0$ and $M^k > 0$. Remind that the first sampling layer ($SL0$), the pseudo measurement layer, is built with enough information, so it is always observable, what implies in $rank(H_{SL0}) = n$, yielding:

$$P_{SL0}^{-1} = H_{SL0}^T R_{SL0}^{-1} H_{SL0} > 0 \tag{6.22}$$

Also, for any other sampling layer:

$$H^T R_{SL}^{-1} H \geq 0 \tag{6.23}$$

The sum of a positive definite and a positive semi-definite matrix in (6.13), is positive definite, and therefore full rank.

$$H^T R_{SL1}^{-1} H + P_{SL0}^{-1} > 0 \tag{6.24}$$

By induction, any sampling layer will be algebraically observable ($rank(H^T R_{SL}^{-1} H + P_{SL-1}^{-1}) = n$) as long as the initial sampling layer is observable. It will also maintain subsequent sampling layers observable, therefore providing descent directions:

$$P_{SL}^{-1} = H^T R_{SL}^{-1} H + P_{SL-1}^{-1} > 0 \tag{6.25}$$

The Wolfe conditions of sufficient descent (Armijo condition) and curvature condition must also be satisfied (Nocedal; Wright, 2006):

$$\begin{aligned} J_{MAP}(x^k + \alpha^k \Delta x) &\leq J_{MAP}(x^k) + c_1 \alpha^k \\ \nabla J_{MAP}(x^k + \alpha^k \Delta x)^T \Delta x &\geq c_2 \nabla J_{MAP}(x^k)^T \Delta x \end{aligned} \tag{6.26}$$

where, c_1 and c_2 are constant values such $0 < c_1 < c_2 < 1$. Such conditions are satisfied in the proposed method by employing a *backtracking* technique for choosing the step-size α^k in (6.15) (Nocedal; Wright, 2006). These conditions ensure the algorithm will converge in a finite number of iterations.

However, the convergence may be compromised in the presence of numerical ill-conditioning. The matrix M^k can easily become ill-conditioned, since it is based on matrix multiplication operations. Such a problem is further aggravated in distribution systems due to the particularities of network parameters and scale. A complementary demonstration of an upper bound for the condition number of the MAP estimation method is also provided. The condition number $k(\cdot)$ of M^k is in the following order of magnitude:

$$\begin{aligned} k(M_k) &= k\left(H^T R_{SL}^{-1} H + P_{SL-1}^{-1}\right) \\ &= k\left(\begin{pmatrix} P_{SL-1}^{-1/2} & H^T R_{SL}^{-1/2} \end{pmatrix} \begin{pmatrix} P_{SL-1}^{-1/2} \\ R_{SL}^{-1/2} H \end{pmatrix}\right) \\ \approx k\left(\begin{pmatrix} P_{SL-1}^{-1/2} \\ R_{SL}^{-1/2} H \end{pmatrix}\right)^2 &\leq k\left(R_{SL}^{-1/2H}\right)^2 + k\left(P_{SL-1}^{-1/2}\right)^2 \end{aligned} \tag{6.27}$$

The major advantage of employing the orthogonal formulation is to avoid performing explicitly computations of M^k, instead of dealing with (6.17), resulting in a near quadratic improvement on numerical conditioning.

$$\begin{aligned} k(\mathscr{Q}\mathscr{R}) \approx k(\mathscr{R}) &\approx k\left(\begin{pmatrix} P_{SL-1}^{-1/2} \\ R_{SL}^{-1/2} H \end{pmatrix}\right) \\ &\leq k\left(R_{SL}^{-1/2H}\right) + k\left(P_{SL-1}^{-1/2}\right) \end{aligned} \tag{6.28}$$

Besides, it can be shown, similar to (Higham, 2002; Hebling et al., 2020), that the Multifrontal QR method, based on block householder reflections, is also backward stable in the numerical sense. It yields the following relation to assessing the error of solution of the Multifrontal QR method ($\tilde{x}$) to the real solution of (B.14) (x), according to the specified numerical tolerance and the floating-point precision $\varepsilon_{machine}$. Consequently, the obtained upper bound for the condition number of $\mathscr{R}$ also enables an evaluation of the numerical stability in the algorithm.

$$\frac{\|\tilde{x} - x\|}{\|x\|} = O(k(\mathscr{R})\varepsilon_{machine}) \tag{6.29}$$

6.4 Event Detection based on Bayesian Credibility Tests

Another aspect that can be explored from the Bayesian perspective is the use of credibility intervals (Congdon, 2007) to accommodate more meaning to the estimates. During the updates of the probability model, due to the arrival of new measurements at each sampling layer, a comparison between the posterior distribution with the prior distribution can be performed trough their credibility intervals. Therefore, for a stationary condition, the test will indicate that the inferred posterior probability is in accordance with the previous sampling layers. However, for a non-stationary condition, as load ramps, contingencies or disruptive events on the grid captured by the faster sampling rates, the test will indicate a disagreement with previous estimations. The credibility intervals can be defined as the range containing a particular percentage of probable values (Makowski; S. Ben-Shachar; Lüdecke, 2019). That is, the closed interval $[x_{SL}^{-}, x_{SL}^{+}]$ and an associated probability p_{ci}, such that:

$$\int_{x_{SL}^{-}}^{x_{SL}^{+}} f_{X|Y}(x_{SL} \mid z_{SL}) dx_{SL} = p_{ci} \tag{6.30}$$

The credibility intervals can be designed as equal-tailed intervals or as highest density intervals. The first associating the interval such that the values outside its range (larger or smaller than the interval limits) have an equal probability, and the second by incorporating the values with larger probabilities (Makowski; S. Ben-Shachar; Lüdecke, 2019). This work employs a credibility interval in the vicinity of the MAP estimated state $\hat{x}_{SL}$ of each sampling layer for the i-th state variable, using Chebyshev's inequality:

$$\hat{x}_{SL}^{i} - k^{*}\sqrt{P_{SL}^{ii}} \leq x_{SL}^{i} \leq \hat{x}_{SL}^{i} + k^{*}\sqrt{P_{SL}^{ii}} \tag{6.31}$$

where, k^{*} defines the probability level of the credibility interval, and in this work is set equal 3 yielding a 0.89 credibility interval, a stable value for any probability distribution and generally adopted in practice (Makowski; S. Ben-Shachar; Lüdecke, 2019). Two approaches may be applied to define the credibility test, comparing the intersection of credibility intervals from different instants, and providing two level of non-stationary event detection:

1. Credibility Test with previous sampling layers: provides a detection of larger non-stationary events, in the sense that the estimated state at a faster sampling layer is far from the previous layer.

2. Credibility Test within the sampling layer: provides a detection of smaller non-stationary events, in the sense that the estimated state at the latest instant is far from the previous instant within the same sampling layer;

By treating the state variables as random variables, the test evaluates if the current probable state belongs to the previous probability functions. Fig. 20 illustrates this concept, with

an hypothetical set of variables and their respective transitions mapped by a point estimate and a credibility interval.

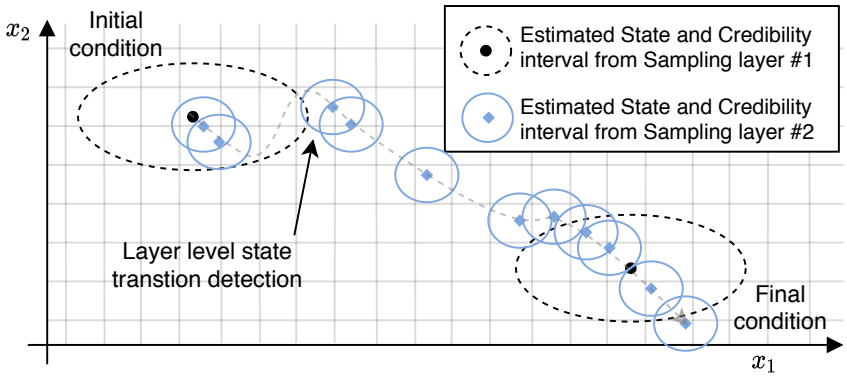


Figure 20 – An example of credibility interval concept to detect state transitions. Two variables monitored by two sampling layers, are mapped with their respective joint point estimates and credible intervals, during a change in state values from an initial to a final condition.

6.5 Simulation Results

Simulations were performed with the IEEE test feeders to validate the proposed DSSE. A load flow calculation gives the reference values for the state variables (x^{lf}) and measurements (z^{lf}). Monte Carlo simulations were performed by including random noise in the reference load flow values to emulate measured values for the DSSE (Silva, 2018). The i-th measurement value was calculated by adding a random noise with Normal distribution $u_i \sim \mathcal{N}(0, \sigma_i)$ in the corresponding i-th measurement reference value, according to (6.32).

$$z_i = z_i^{lf} + u_i \tag{6.32}$$

Different precision levels, for each measurement, generates the noise standard deviation σ_i according to (6.33) (Castillo et al., 2011):

$$\sigma_i = \left| z_i^{lf} \right| pr_i/3 \tag{6.33}$$

where, pr_i is the precision of the i-th meter. In this study it was assumed 30% for pseudo measurements, 5% for active and reactive power measurements from smart meters, 2% for active and reactive power measurements from SCADA, 1% for voltage magnitude from SCADA, 0.1% for voltage and current phasors measurements from PMUs. pseudo measurements represent the network's active and reactive loads and are obtained from typical load curves or historical data. Virtual measurements represent zero-injection nodes (nodes without load or generation and fictitious nodes) and are incorporated as power injection measurements equal to zero with the very low standard deviations.

Table 4 – Measurement set for the IEEE13 system and sampling layers.

Sampling Layer	Type of Measurement	Location
Pseudo Measurement	Active/Reactive Power	Nodes: 632, 645, 646, 634, 671, 611 and 652
	Virtual Measurement	Nodes: 680, 684, 632, 692
Smart Meter	Active/Reactive Power	Nodes: 611, 634, 675
SCADA	Active/Reactive Power	Nodes: 650 Branches: (650 - 6550R), (650R - 650), (671 - 692)
	Voltage Magnitude	Nodes: 650, 671
PMU	Current Phasor	Branches: (650 - 650R), (633-634)
	Voltage Phasor	Nodes: 650 and 633

To simulate different sampling rates, a sequence of quasi-stationary conditions obtained from successive load flow scenarios provides the reference value at each different instant t. Each instant is monitored by a different set of measurements according to their respective sampling rates and to the previously explained procedure to include noise. This work considers the following sampling rates: pseudo measurements at 15 minutes (PSEUDO sampling layer); smart meters at 1 minute (SMETER sampling layer); SCADA measurements at 2 seconds (SCADA sampling layer); and phasor measurements at 100 ms (PMU sampling layer). The Mean Absolute Error (MAE) metric evaluates the accuracy of the estimation:

$$\mathrm{MAE}_t = \frac{trial = 1}{n_{trials}} \sum_{1}^{n_{trials}} \left| \hat{x}_t^{trial} - x_t^{lf} \right|, \tag{6.34}$$

where $\hat{x}_t$ is the estimated state in t, x_t^{lf} is the reference state value in t, and n_{trials} is the number of trials during the simulation ($n_{trials} = 100$).

6.5.1 Bayesian Fusion DSSE Accuracy

The first test intends to demonstrate the effect of the MAP estimation and to evaluate accuracy aspects of the proposed DSSE with the IEEE 13 nodes test feeder. Table 4 presents the respective metering system, and the associated sampling layers. This first simulation consists in a fixed nominal loading along all observations, a stationary condition. To illustrate the effect of the MAP estimation, Fig. 21 presents the absolute error boxplot for this simulation in each sampling layer. Besides the reduction on the MAE indicator, a smaller error variance is also obtained after the execution of MAP. This because of the higher precision of the faster sampling layers, considered separately in the Bayesian Fusion approach.

The MAE indicator is also evaluated for each separate node of the system in Fig. 22. There is an increase on the state estimation accuracy for all nodes as faster samples are processed, mainly on the state variables of nodes associated with pseudo measurements. It is noteworthy that only the pseudo measurement sampling layer is fully observable, and all other layers have a small number of measurements.

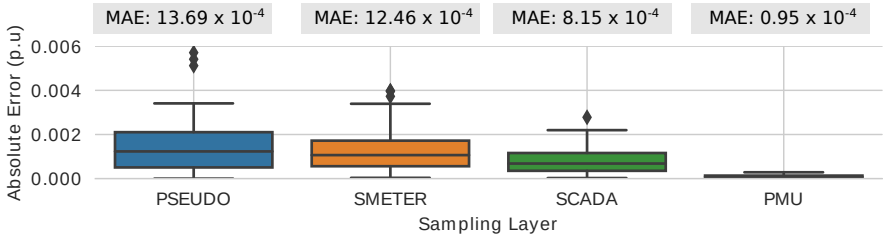


Figure 21 – Boxplot of the absolute estimation error for all state variables in each sampling layer. The MAE indicator among all simulations is also highlighted.

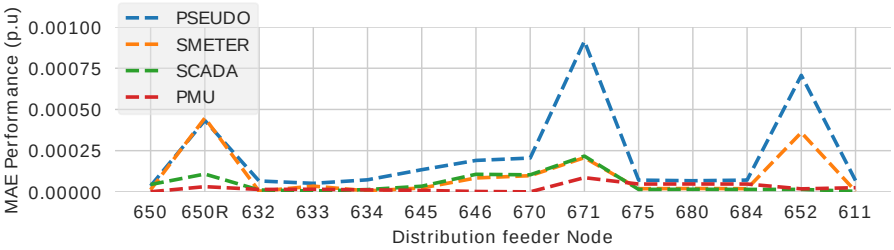


Figure 22 – MAE performance index for the estimation in each sampling layer, from the three-phase state variables in each node of the system.

Finally, Fig. 23 presents the MAE for each phase of the system in each sampling layer. Table 5 presents detailed three-phase MAE for voltage magnitude and phase angles, separately for each phase. The results corroborate the previous analysis, demonstrating the estimation improvement with information of the fastest sampling layers in all phases and state variables.

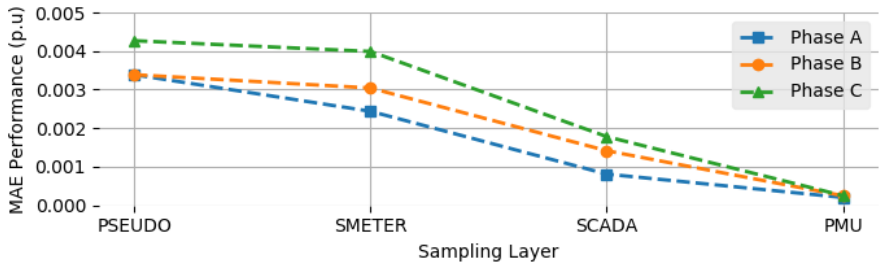


Figure 23 – MAE performance index for the estimation in each sampling layer, for each phase of the system.

6.5.2 Effect of Load Variation and Sudden Change

This second test considers the IEEE 123 nodes test feeder associated with: SCADA an PMU measurements located at the substation and at voltage regulators and switches; pseudo measurements in all loads; virtual measurements in all zero injection nodes; and few smart meters at eight nodes as special consumers. The simulation corresponds to 20 seconds with two loading scenarios: first a stationary condition; and the second including a 1% of load variation around the nominal loading and a subtle increase of 50 % in 14 randomly selected nodes. It is

Table 5 – MAE performance per phase and state variable in the IEEE13 test system for each sampling layer.

State Variable	Phase	Sampling Layer			
		PSEUDO	SMETER	SCADA	PMU
Voltage Magnitude	*Phase A*	$3.47x10^{-3}$	$3.44x10^{-3}$	$1.34x10^{-3}$	$0.34x10^{-3}$
	Phase B	$3.38x10^{-3}$	$2.82x10^{-3}$	$1.51x10^{-3}$	$0.27x10^{-3}$
	Phase C	$4.87x10^{-3}$	$4.42x10^{-3}$	$1.37x10^{-3}$	$0.29x10^{-3}$
Voltage Phase Angle	*Phase A*	$4.83x10^{-3}$	$3.01x10^{-3}$	$0.55x10^{-3}$	$0.29x10^{-3}$
	Phase B	$4.27x10^{-3}$	$4.27x10^{-3}$	$2.51x10^{-3}$	$0.49x10^{-3}$
	Phase C	$5.61x10^{-3}$	$5.25x10^{-3}$	$3.09x10^{-3}$	$0.40x10^{-3}$

a more realistic condition, since there are no guarantee the loads keep a fixed value while new samples arrive.

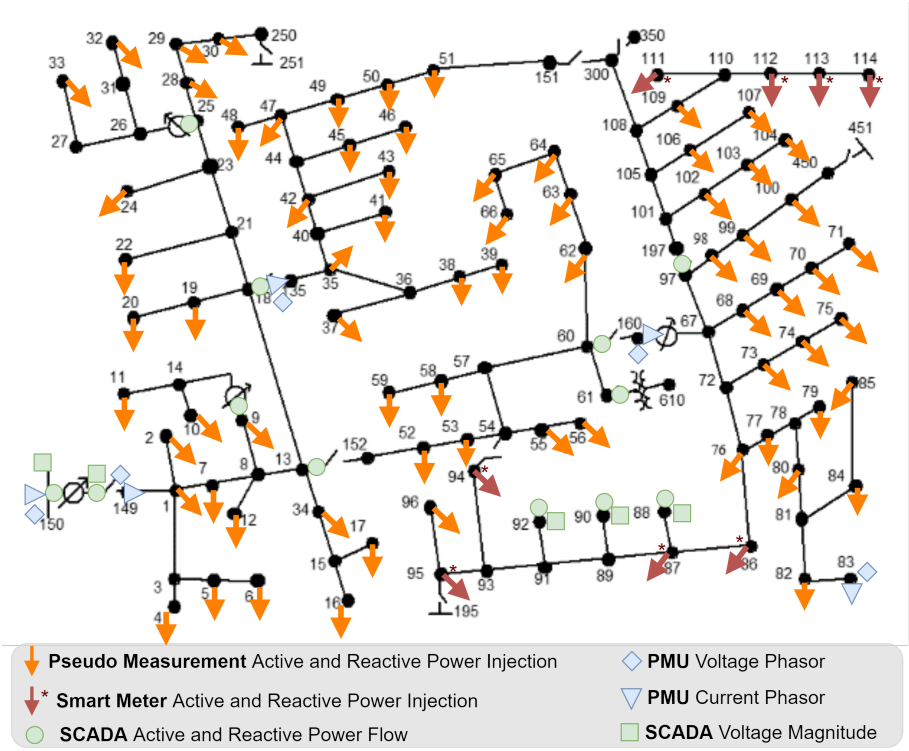


Figure 24 – IEEE 123 nodes test feeder and the respective measurement set grouped by sampling layers. Virtual measurements are considered in nodes without loads, that is, zero injection nodes.

The Bayesian Fusion DSSE is compared with a Hybrid DSSE, which simultaneously processes all measurements in a single stage via the Weighed Least Square algorithm, and with an Extended Kalman Filter (EKF) DSSE that uses recursively the estimated posterior to update the prior belief. Fig. 25 presents the comparison using the MAE performance metric for this three estimators. The MAE performance index per each phase for the three estimation methods, in the stationary condition and with the load variation and sudden load change, are presented in Table 6. Table 7 presents this comparison for different levels of load variation.

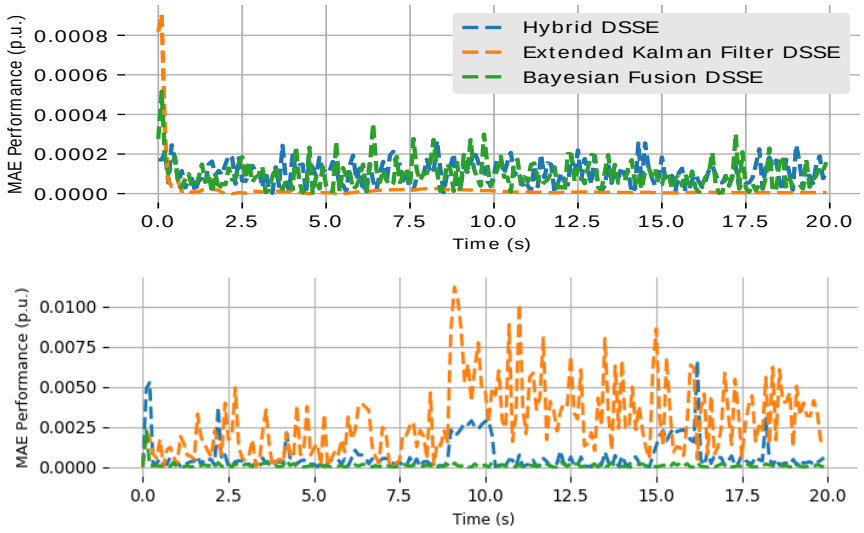


Figure 25 – Estimation performance in two loading scenarios: the first, a stationary condition of fixed load values; and the second with 2 % of load variation and subtle load increase from t=9s to t=15s.

Table 6 – Comparison of estimation performance per phase for the Hybrid, EKF and Bayesian Fusion DSSEs.

Loading Scenario	**Phase**	**DSSE method**		
		Hybrid	**Kalman Filter**	**Bayesian Fusion**
Fixed loading	*Phase A*	$1.15x10^{-4}$	$0.31x10^{-4}$	$0.82x10^{-4}$
	Phase B	$1.57x10^{-4}$	$0.34x10^{-4}$	$1.36x10^{-4}$
	Phase C	$1.43x10^{-4}$	$0.20x10^{-4}$	$1.22x10^{-4}$
With load variation	*Phase A*	$6.21x10^{-4}$	$32.47x10^{-4}$	$1.41x10^{-4}$
	Phase B	$6.61x10^{-4}$	$30.61x10^{-4}$	$2.46x10^{-4}$
	Phase C	$7.43x10^{-4}$	$31.36x10^{-4}$	$2.15x10^{-4}$
With sudden change	*Phase A*	$6.16x10^{-4}$	$56.26x10^{-4}$	$1.57x10^{-4}$
	Phase B	$10.95x10^{-4}$	$30.59x10^{-4}$	$2.42x10^{-4}$
	Phase C	$8.74x10^{-4}$	$30.91x10^{-4}$	$2.01x10^{-4}$

The EKF approach presented a very good estimation for the fixed load scenario, but lost its accuracy in the presence of load variation, specially for a subtle load increase. This is expected since the EKF has good smoothing properties, and the fixed load value is the smoothest scenario possible. When load variation was included, this premise became false, hampering its accuracy. Besides, the hybrid approach noticeably loses accuracy whenever load variation is present, since it essentially mixes measurements from different instants in the same set. Furthermore, the proposed Bayesian Fusion DSSE keeps its accuracy level even in the case of a subtle load increase, showing better results in the presence of load variation. This is related to the fact that the prior distribution gives only an initial estimate of the state vector and its covariance. Whenever faster samples are gathered, the MAP estimation uses the prior knowledge along with only measured information of this latest sample to update the state vector.

Table 7 – Comparison of estimation performance in the presence of load variation for the Hybrid, EKF and Bayesian Fusion DSSEs.

Load Variation	Hybrid DSSE	Extended Kalman Filter DSSE	Bayesian Fusion DSSE
0.0 %	$1.04x10^{-4}$	$0.20x10^{-4}$	$0.95x10^{-4}$
1.0 %	$2.7x10^{-4}$	$7.79x10^{-4}$	$1.09x10^{-4}$
2.0 %	$3.07x10^{-4}$	$16.78x10^{-4}$	$1.36x10^{-4}$
5.0 %	$8.35x10^{-4}$	$40.28x10^{-4}$	$2.65x10^{-4}$
10.0 %	$17.72x10^{-4}$	$87.73x10^{-4}$	$4.11x10^{-4}$

6.5.3 Influence of Bad Data

Another important aspect is the DSSE performance against bad data (gross errors), which was evaluated using the previous IEEE123 test setup. Initially, the influence of gross errors at the pseudo measurement layer was evaluated since pseudo measurements are prone to erroneous information in their modelling, such as forecasting error, a misconstruction of typical load curves or due to nontechnical losses. Fig. 26 illustrates the effect of bad data in the pseudo measurement layer in different sampling layers. Gross errors of 10σ (10 times the respective standard deviation) were included in randomly selected pseudo measurements, a case of multiple bad data.

The effect appears to influence the smart meter layer, which strongly depends on the pseudo measurement layer since only eight loads are monitored by smart meters. Meanwhile, the faster sampling layers from SCADA and PMU enable reducing the harmful effect on the final estimation. This reduction is due to the higher accuracy of such measurements and the inclusion of information from pseudo measurements as prior conjugate distribution instead of processing them together with the erroneous values. In essence, the pseudo measurement layer shares only its processed state information with the faster layers and the associated covariance. When SCADA and PMU measurements are processed, their accuracy is enough to improve the estimation.

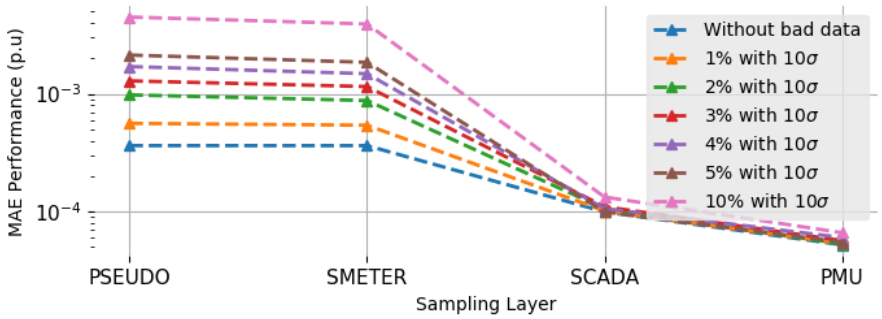


Figure 26 – Increase of bad data at the pseudo-measurement sampling layer and effect on the estimation performance at each sampling layer.

A second simulation was performed to evaluate bad data in the SCADA measurements

and their respective effects on the final PMU sampling layer. Fig. 27 illustrates the effect of bad data on the final estimation of the PMU sampling layer for all state variables in each phase (voltage magnitude and phase angle). Different values of gross errors in the range of 5σ to 15σ were included in 10% of randomly selected SCADA measurements.

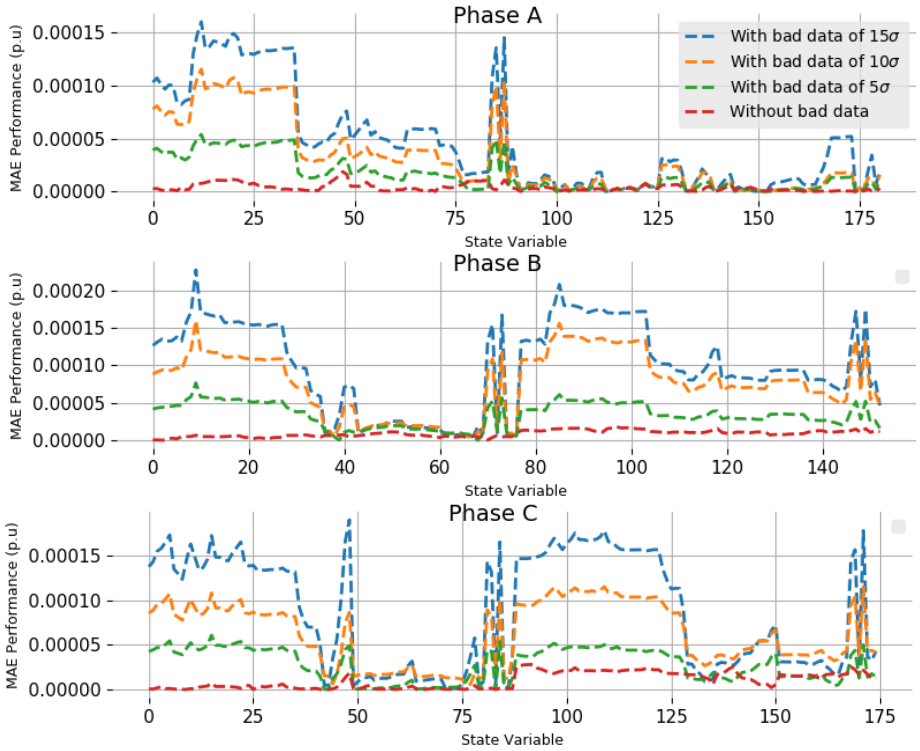


Figure 27 – Increase of bad data at the SCADA sampling layer and effect on the final estimation performance for all state variables in each phase. The first half of state variables correspond to the voltage magnitude and the second half to voltage phase angles.

In this case, bad data at the SCADA sampling layer significantly influence the final estimate obtained by the PMU layer. Once again, the conjugate model assumes only a vicinity of the estimated state from previous layers, reducing the bad data's hampering effects in the subsequent layers. The prior belief can be associated with additional flexibility to deal with the faster sampling layers information without changing the initial hypothesis about the measurement noise.

Different methods can further aid the estimation, such as employing the largest normalized residue test to detect and identify gross errors (Abur; Gómez-Expósito, 2004; Massignan et al., 2020). In this context, this final test illustrates the effect of bad data in the normalized residual analysis in each sampling layer. The test is performed using the latest estimated state variables and each sampling layer's respective measurement model. This way, faster sampling layers can aid the slower ones. The test considers the inclusion of different types of gross errors and sampling layers, exploring the effect of single and multiple gross errors, along with a situation of non-technical losses, a special case of bad data at load values from pseudo measurements or smart meters. Fig. 28 illustrates the largest normalized residue in each sampling layers.

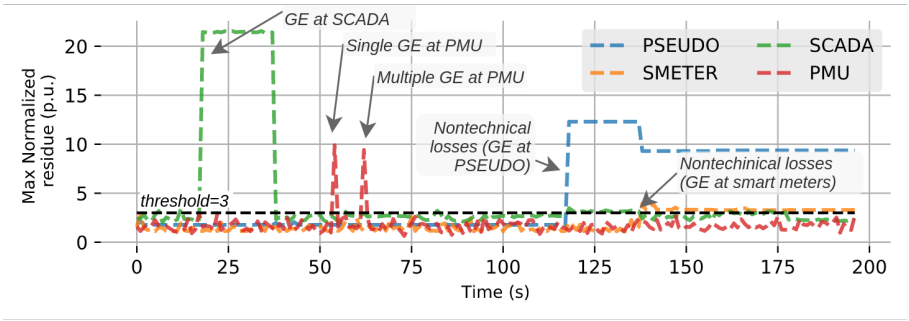


Figure 28 – Effect of bad data on the normalized residual analysis at different sampling layers in the presence of: (a) multiple gross errors of 25 σ at 5% of the SCADA sampling layer; (b) single gross error of 10 σ at one PMU; (c) multiple gross errors of 15 σ at three PMUs; (d) nontechnical losses of 50% bias in 10% of the load pseudo measurements; and (e) nontechnical losses as a no-load situation in two smart meters.

The method can properly aid several cases of gross errors at different sampling layers. It is noteworthy that further exploration of the sensitivity and possibility of identifying gross errors is required, as shown in low sensitivity situations such as on the few smart meters in this case. It primarily relates to the amount of information available and redundancy in all sampling layers. In this sense, new methods based on the Bayesian framework may complement bad data processing, such as using hyper-prior models for the noise models, especially in low redundancy scenarios such as in DSSE. Besides, the application of more robust numerical and statistical methods can aid even further the DSSE performance, such as with kernel density estimation principles in (Pesteh et al., 2019). These complementary efforts aim to provide better prior information for the estimation process, a crucial feature to increase the Bayesian fusion DSSE accuracy.

6.5.4 Smart Meters Deployment and Credibility Intervals

This test explores the Bayesian Fusion DSSE under a full deployment of smart meters in the IEEE US Low Voltage Test System (IEEE342). The SCADA measurements are located at the substation and feeder's bays, while PMU measurements, are at the high voltage nodes, as ilustrated in Figure 29. Despite the DS being observable with the smart meters at the low voltage loads, a pseudo measurement layer is kept, based on previous day information, to accommodate possible loss of communication. The test validates the proposed method under a model that comprises all parts of a DS, from the high voltage substation to a meshed low voltage circuit, and the typically associated measurements.

To evaluate the concept of credibility intervals, the simulation considers the following scenarios: (1) a normal condition with load variation of 0.5% for 5 minutes; (2) a momentary loading increase to simulate a distributed resources intermittency amidst the normal condition; (3) a medium voltage feeder contingency, where, during the first 30 seconds, the network model is not updated, with the topology momentarily incorrect; and (4) the contingency with the updated model.

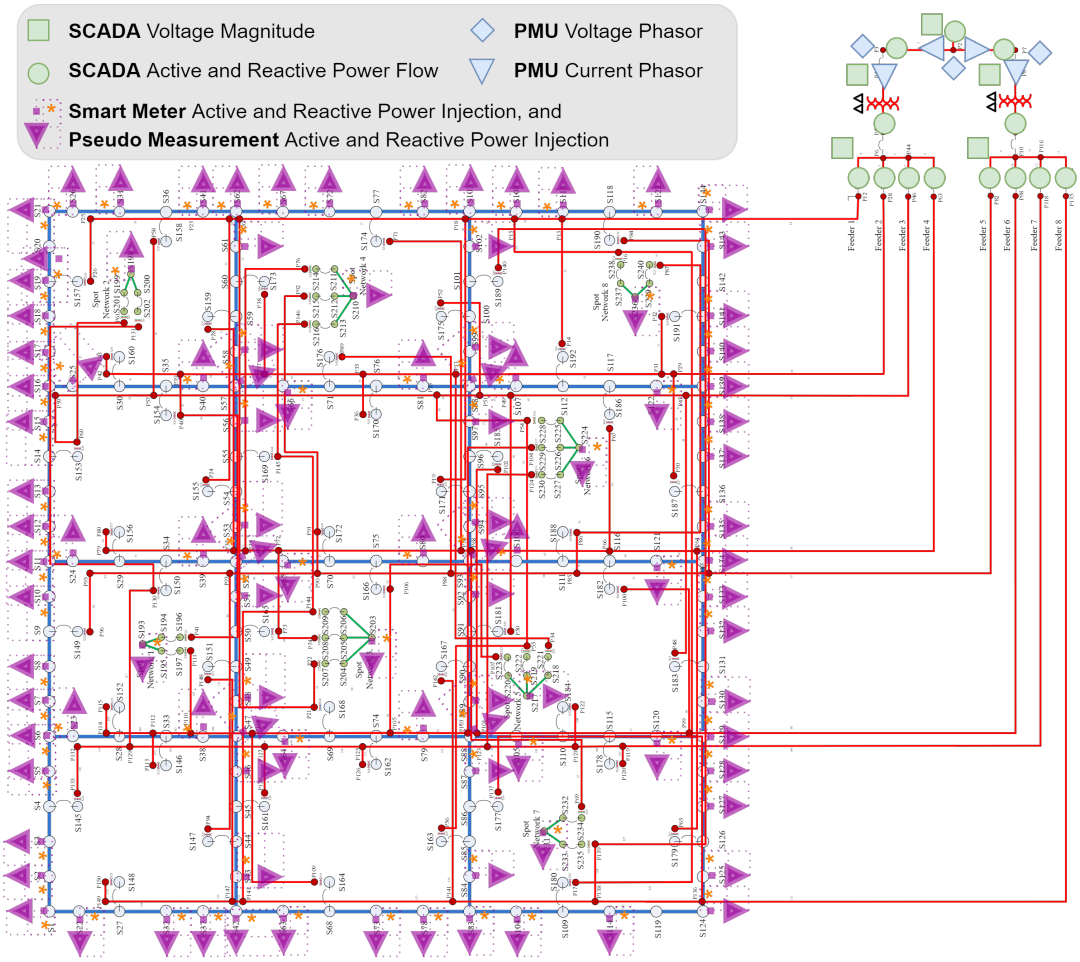


Figure 29 – IEEE US Low Voltage test system (IEEE342) and the respective measurement set grouped by sampling layers. Virtual measurements are considered in nodes without loads, that is, zero injection nodes.

Fig. 30 illustrates the temporal behavior of the estimated voltage magnitude at phase *A* of one of the low voltage nodes, during the previous described scenarios. With an enhanced observability due to the smart meters at the low voltage, the Bayesian Fusion DSSE provides accurate estimates along almost the entire sequence of events. The effect of the generation intermittency in (2) is properly assessed even tough the smart meters update their values each minute. It shows the importance to complement them with SCADA and PMU measurements, even if located at the medium and high voltage part of the DS, as in this case. This integration enables sensing faster events and increase the DSO awareness of the system. Only during the event span (3) the DSSE loses its accuracy, showing the negative effect of unmonitored contingencies. During this event, the problem is in the network model as a case of topology error, a medium voltage feeder is out-of-service while the network model says otherwise. Nonetheless, the estimator provides a direction of state change, and detects an anomaly in the system from the disagreement between the network model and measured values. After a proper update on the model in (4), the proposed DSSE regains its accuracy.

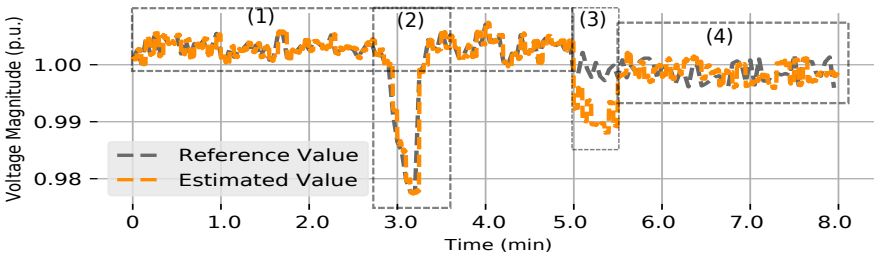


Figure 30 – Example of estimated state variable, voltage magnitude at phase A of the low voltage node # 1132, during the events on the IEEE342. Each simulated scenario is highlighted: (1) load variation; (2) generation intermitency; feeder contingency, (3) before and (4) after network model update.

The estimated values of two state variables are illustrated in Fig. 31 along with the prior and posterior credibility intervals. An important aspect of the proposed Bayesian Fusion DSSE is that during the sequence of events, the credibility test detects the transitions on the network, by properly assessing the intersection of credibility intervals with previous states. It accommodates more meaning to the estimation process, since it also carries the notion of the most probable values in the vicinity of each estimated state. The probabilistic interpretation of such result may enable its integration with different identification methods, for instance with learning algorithms based on historical data, as well as to accommodate real time risk analysis within the credibility intervals.

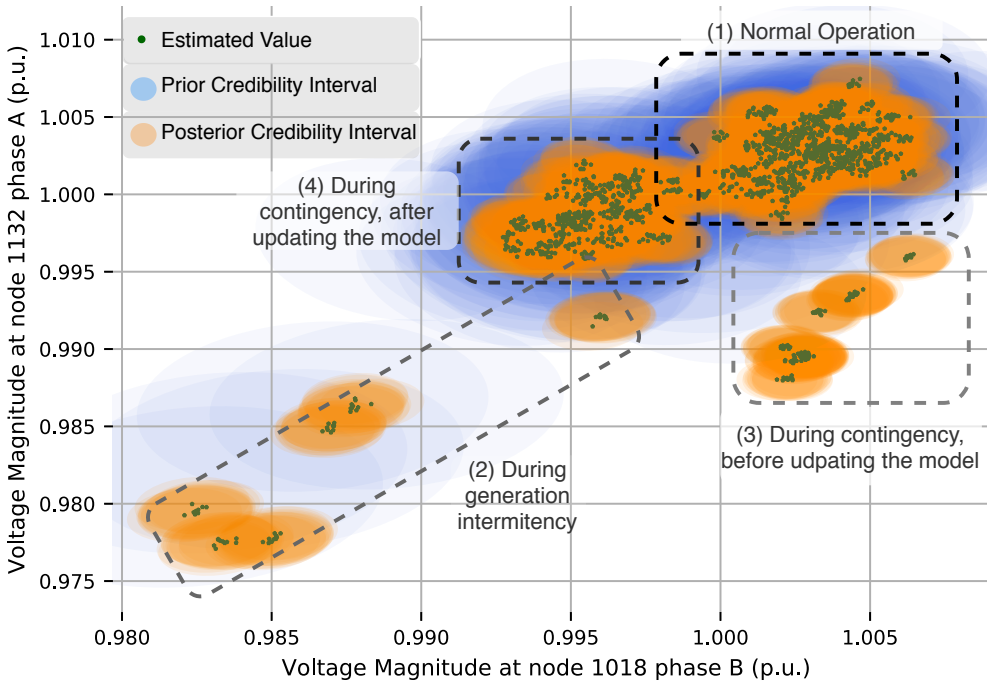


Figure 31 – Estimated state and credibility intervals for two state variables at the low voltage of the IEEE342. The Bayesian Fusion DSSE detects the occurrence of the highlighted events if the latest estimated values falls within the region of previously estimated values.

6.5.5 Validation in a Real Distribution Feeder

Finally, the Bayesian Fusion DSSE was also applied in a real Brazilian test feeder from the city of Londrina-PR, that comprises 1058 three-phase medium voltage nodes at 13.8 kV (Massignan et al., 2018). It is a traditional distribution feeder, monitored only by pseudo measurements from typical load profiles at each 15 minutes, and a few SCADA measurements at the substation (voltage magnitude and active and reactive power flow) and at a recloser (current magnitude), with updating rate of 1 minute, as illustrated in Fig. 32. Each of the 192 MV/LV transformers presents a three-phase pseudo measurement of active and reactive power injection. The evaluation in this scenario corresponds to a one-day horizon, thereby 1440 SCADA updates and 96 pseudo measurement values.

In order to validate the proposed approach, independent measurements were taken in three different transformers to compare the estimated results with the measured ones (Massignan et al., 2018). Table 8 presents an accuracy comparison between the Bayesian Fusion DSSE, the Hybrid DSSE, the Kalman Filter DSSE and the previous results obtained by a Real Time Load Monitoring, currently available at the DSO (Massignan et al., 2018), with respect to the final voltage magnitude and active and reactive load estimates gathered from the validation measurements.

It is noteworthy that all four estimators depart from the same set of initial typical load profiles, which present poor quality. It is a common practice to employ the Load Estimator results as refined pseudo measurements. However, utilities also employ such results directly as a real-time monitoring strategy depending on the application. Hence, the comparison intends to demonstrate similar setup conditions encountered in the DSO and illustrate the performance enhancement due to the concept of sampling layers.

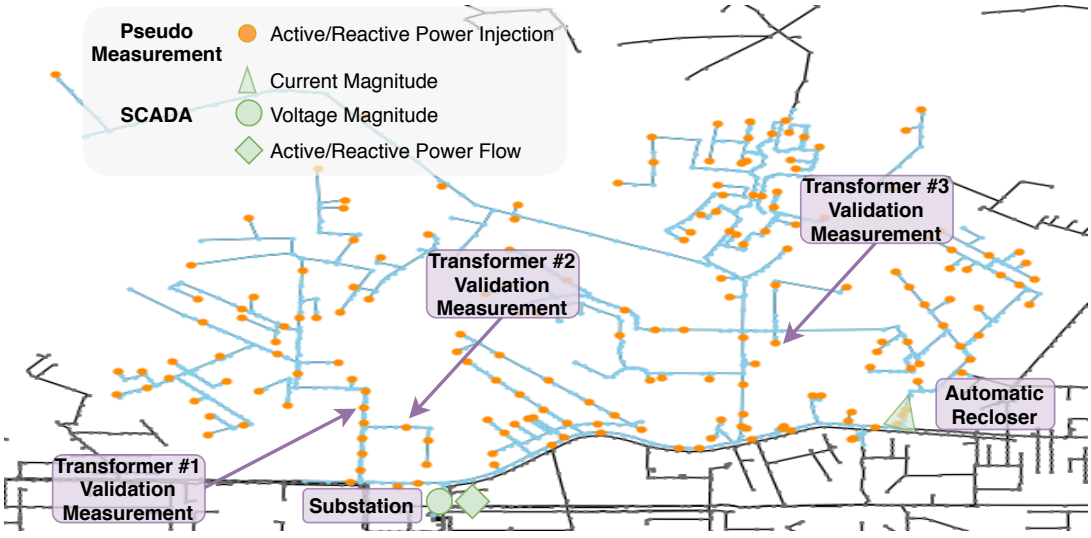


Figure 32 – Real Brazilian distribution feeder with 1058 three-phase nodes (blue). It presents mainly pseudo measurements and few SCADA measurements. In-field measurements were taken to validate the Bayesian Fusion DSSE.

The lack of measurements limits the accuracy of the estimation process according to the pseudo measurements quality, one of the main challenges for DSSEs. Interestingly, the proposed approach improved the voltage estimation accuracy, the variable of the search space in the DSSE algorithm, while the Load Estimator searches on load injection values. In this sense, to attend the requirements of smart active DS, it is imperative to increase the quality of available data for DSOs. This can be reached by incorporating smart meters and/or PMU measurements in the estimation, as shown in the previous section. It will only be effective if employed with high resolution models of the DS able to integrate different sampling rates as well as all parts of the DSs.

Regarding the Hybrid and Kalman Filter performance, similar results than the previous sections were observed. Despite the close voltage magnitude results, the same was not observed in other estimated electrical quantities (such as the active and reactive loads, and consequently the voltage phase angles). This further supports the claim that mixing information from different instants may hamper DSSE performance.

Fig. 33 illustrates the validation results, comparing an estimated and measured voltage magnitude at phase *A* of Transformer #1, along a whole day. Remind that this measured voltage was not considered in the DSSE, and it is only used for comparison purposes. The Bayesian Fusion DSSE enhance the estimations after processing the SCADA measurements, when compared with the initial estimates based on the pseudo measurement layer.

This result shows also a similar behavior of the Bayesian DSSE when compared Hybrid approach. This is a consequence of not including sequential relations for the state in the estimation model, and even though the Bayesian DSSE presents better accuracy. In the Bayesian DSSE only temporal relations among different sampling layers are captured in the model, while in the Kalman Filter two subsequent instants are related by the transition matrix, which makes the estimation more sensitive to sudden changes. Nonetheless, the other estimated electric quantities did not present such similarity, as can be inferred by the accuracy comparisons for the active and reactive loads in Table 8.

6.5.6 Numerical and Computational Aspects

The tests of the proposed Bayesian Fusion DSSE were performed using a microcomputer with a Core i7 2.10 GHz, 8GB RAM and an implementation in C programming language in a UNIX system. Table 9 presents the processing time in the test feeders for each sampling layer. Regarding the credibility interval calculation, it takes around 20 % of the total DSSE time. Thereby, the proposed Bayesian DSSE presented an adequate processing time for real time applications, for both MAP estimation and the credibility test.

Fig. 34 shows the convergence rate in the IEEE123 test system of the proposed method. The direct method for the information fusion problem with the matrix M_k is unable to converge already at the first sampling layer, due to severe ill conditioning. Conversely, the orthogonal

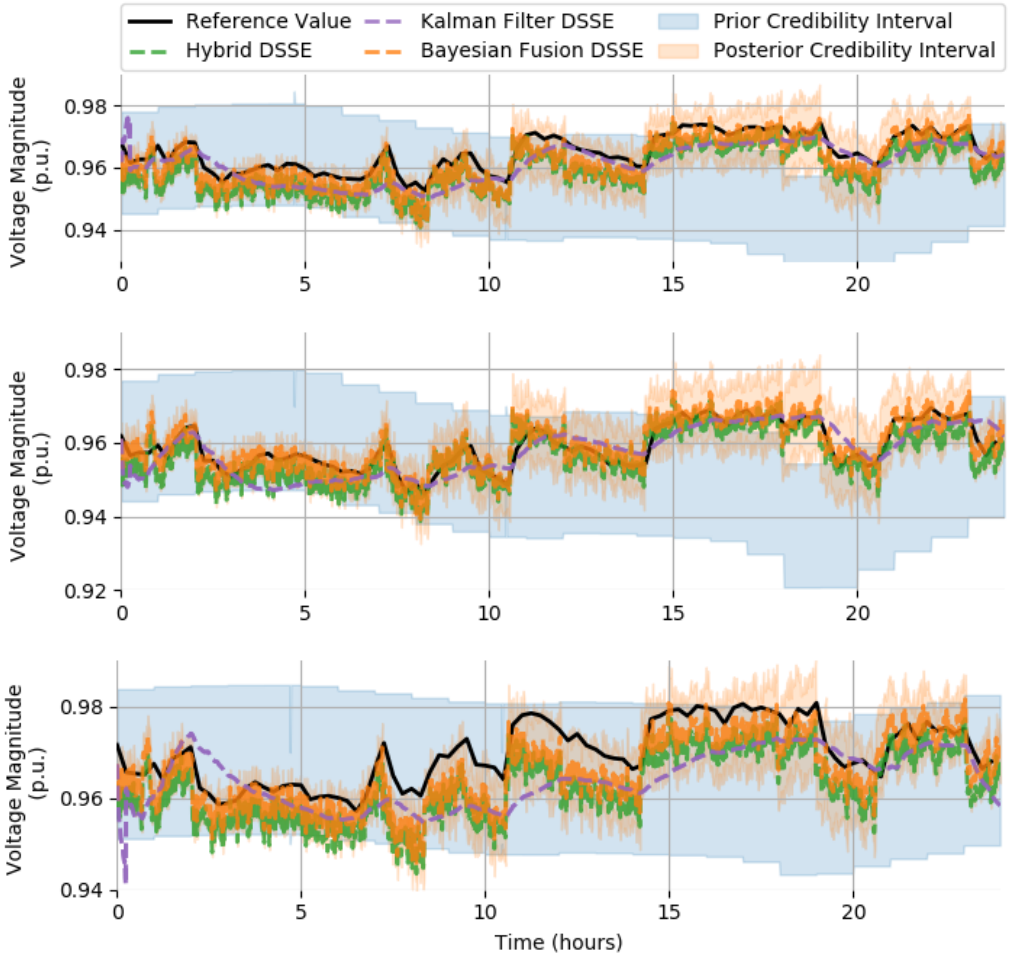


Figure 33 – Estimated voltage magnitude of Transformer 3 in the Brazilian real feeder, in phase A, phase B and phase C, respectively, and the respective validation measurement. The credibility intervals for the pseudo measurement and SCADA sampling layers are highlighted for each instant.

formulation presents a quadratic convergence characteristic of the Modified-Newton method in all sampling layers. Similar results were observed for the other test systems. The condition number in each sampling layer are presented in Fig. 35 for the IEEE123, IEEE342 and the real feeder. A near quadratic improvement on the numerical conditioning is observed with the orthogonal formulation compared to the direct approach to solving the MAP estimation with the M^k matrix. Numerical conditioning is an essential feature for any application in distribution systems, which imply special care when choosing the numerical methods to perform DSSE.

Finally, the sparsity treatment also presents the crucial numerical feature of the implementation, as ilustrated in Figure 36. For instance, a reduction of more than 87% on fill-ins was obtained by the AMD ordering in all cases. It increases numerical stability and computational efficiency by reducing unnecessary operations with zero values, and also reduces memory allocation. With fast sampling rates and higher data volumes, such as those from PMUs, computational burdens must be considered for scalability in any practical application. It is noteworthy that the

Table 8 – Validation with available real time monitoring regarding final active and reactive estimated loads.

Validation Transformer	**Percentage MAE Performance Metric**					
	Load Estimator (Massignan et al., 2018)			*Bayesian Fusion DSSE*		
	Voltage Magnitude	*Active Power*	*Reactive Power*	*Voltage Magnitude*	*Active Power*	*Reactive Power*
Transformer # 1	1.39 %	11.9 %	11.5 %	1.15 %	11.8 %	11.4 %
Transformer # 2	0.85 %	24.1 %	18.7 %	0.61 %	23.4 %	24.6 %
Transformer # 3	0.71 %	59.0 %	50.0 %	0.39 %	59.7 %	64.8 %
	Hybrid DSSE			*Kalman Filter DSSE*		
	Voltage Magnitude	*Active Power*	*Reactive Power*	*Voltage Magnitude*	*Active Power*	*Reactive Power*
Transformer # 1	1.17 %	29.0 %	23.1 %	1.49 %	71.1 %	78.5 %
Transformer # 2	0.64 %	29.8 %	26.0 %	0.82 %	76.2%	80.2 %
Transformer # 3	0.58 %	84.6 %	96.4 %	0.88 %	68.5 %	81.3 %

Table 9 – Mean processing time/number of iterations for the Bayesian Fusion DSSE sampling layers.

System	**Sampling Layers**			
	PSEUDO	*SMETER*	*SCADA*	*PMU*
IEEE13	3.2 ms / 4 its	2.4 ms / 4 its	2.4 ms / 3 its	1.2 ms / 3 its
IEEE123	22.1 ms / 4 its	7.2 ms / 3 its	7.9 ms / 3 its	2.4 ms / 3 its
IEEE342	136.8 ms / 4 its	91.8 ms / 3 its	58.4 ms / 3 its	15.9 ms / 3 its
Real 1058	236.1 ms / 4 its	(NA)	32.7 ms / 4 its	(NA)

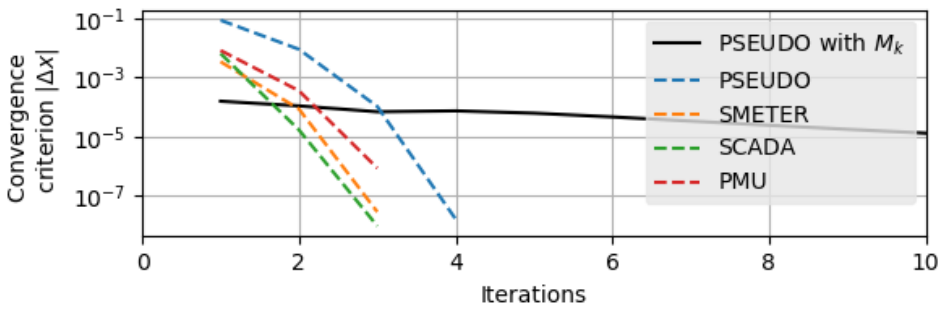


Figure 34 – Convergence rate in each sampling layer for the IEEE123 test system.

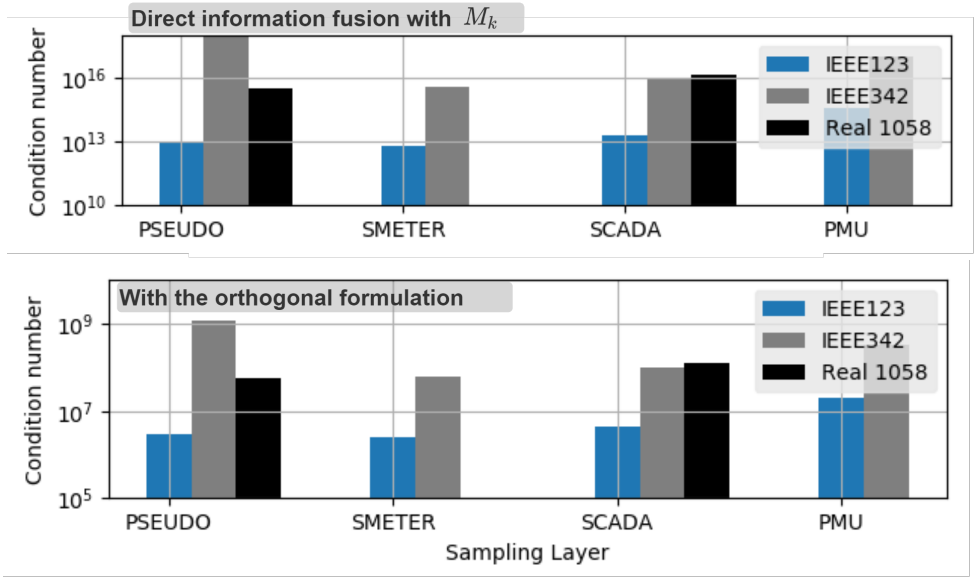


Figure 35 – Condition number of the matrix M^k and of the orthogonal factorization upper triangular $\mathscr{R}$ for the IEEE123, IEEE342 and Real feeder with 1058 nodes. A quadratic reduction is observed which translates in improved numerical stability.

implemented version of the proposed DSSE already presents a performance suitable to deal with such scenarios, and it still has a significant possibility for computational enhancements, such as using parallel computing and improving data structures and memory locality.

6.6 Concluding Remarks

This chapter presented a Bayesian inference approach to perform information fusion in DSs, gathered at different sampling rates such as typical load profiles, smart meters, SCADA and PMU measurements. The method is based on a sampling layer concept, that decomposes the measurement set according to the respective updating rates. A MAP estimation with the latest information available, based on a sparse orthogonal method, provides the estimated state, even in very ill-conditioned DSs while keeping a tractable computational time. It enables the integration of heterogeneous sources of information according to their respective accuracies and sampling rates, providing estimation with smaller mean errors at each faster sampling layer. Comparisons with the hybrid and EKF DSSE demonstrate the advantages of treating each sampling layer separately, without mixing information from different instants.

Theoretical arguments and simulation results were provided to support the following interesting features of the Bayesian fusion DSSE. By exploiting a sparse orthogonal factorization method, it is possible to apply the Bayesian Fusion DSSE in very ill-conditioned DSs while keeping a tractable computational time. The exploration of the Bayesian framework led also to the proposition of an event detection method based on credibility intervals. Thereby maintaining accuracy under small and large non-stationary events, such as load variations, subtle generation intermittency and network contingencies. Comparisons with an hybrid and an EKF

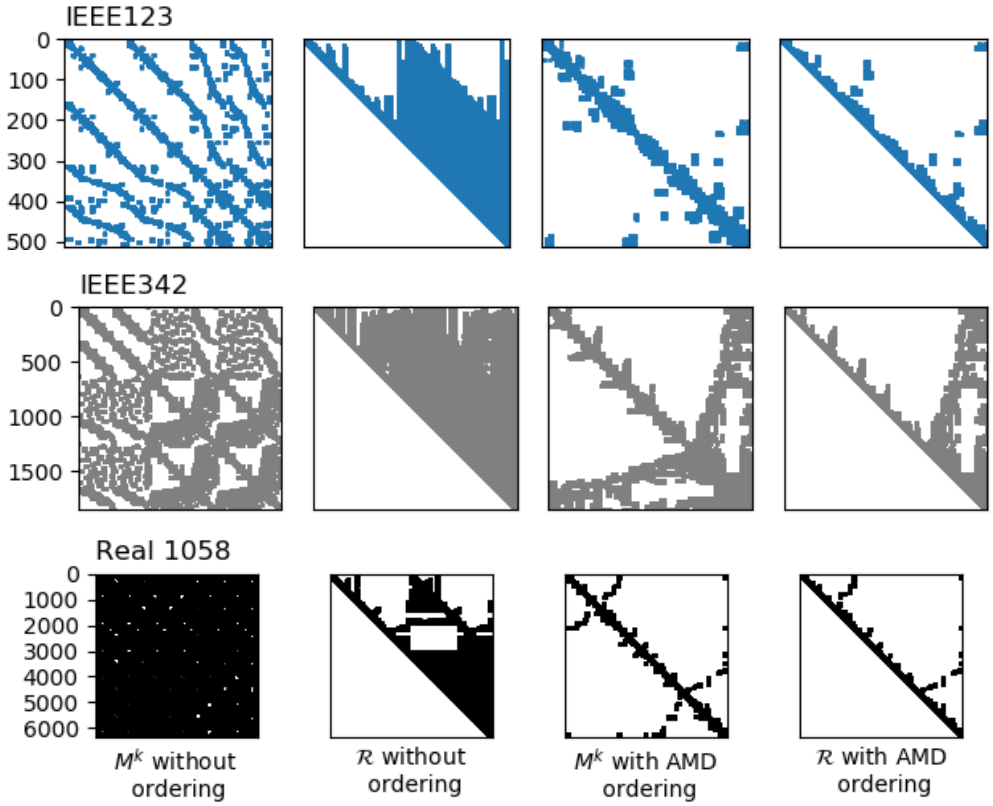


Figure 36 – Sparsity patterns of the matrix M^k and the orthogonal factorization upper triangular $\mathscr{R}$ for the IEEE123, IEEE342 and Real feeder with 1058 nodes, obtained without state variables ordering and after the approximate minimum degree ordering. A significant reduction of fill-ins improves even further the numerical stability of the implementation.

DSSE demonstrate the advantages of treating each sampling layer separately, without mixing information from different instants. For stationary operating condition, it keeps its precision even when the loads vary around their nominal values during the time interval between two SCADA data scans

Moreover, it also enables a full integration of smart meters in the DSSE problem, enhancing its performance and enabling the detection of events even at the low voltage networks with the concept of Bayesian credibility intervals. Finally the practical application in a real feeder validates the proposed approach with nowadays traditional DS feeder. Thereupon the proposed DSSE provides high resolution information about the grid condition for DSOs in real time. Future work comprises more exploration of the Bayesian framework such as the use of hyper-prior models to capture random effects of the DSs, as well as to evaluate new methods for bad data detection. The Bayesian framework may also be explored under different features of topology estimation, fault location, spatial fusion for large scale networks, anomaly detection algorithms and dynamic estimation. Another exciting direction is gathering more information to compose the prior knowledge, such as exploring short-term historical values and different

conjugate models. Besides, improving statistical robustness against non-Gaussian noise can sensibly aid the quality of the prior information, such as in the case of bad data, and can be further explored through concepts of kernel estimation and Information Theory.

7 BAYESIAN SPATIAL FUSION FOR MULTIAREA STATE ESTIMATION IN LARGE SCALE DISTRIBUTION SYSTEMS

The content of this chapter have been published in the following book chapter:

- Massignan, J. A. D.; London JR., J. B. A., P. Pegoraro, M. Pau, "Multiarea State Estimation for Distribution Systems". *Power Distribution System State Estimation*, Book Chapter, IET, 2022;

7.1 Introduction

This chapter presents a novel Multiarea perspective for DSSE based on Bayesian Inference principles. The proposed perspective enables dealing with large scale distribution systems, from the high voltage substation to the low voltage networks in a single framework. A new formulation for the Multiarea state estimation incorporates prior knowledge on the model as a spatial fusion inference problem, complementing observability. The proposed Multiarea consists of a two-step procedure, a Local Estimation followed by a Coordination Step, a hierarchical architecture consonant with modern data centres for DS operation and control.

Concerning DSSE, the Bayesian Inference background has been employed to deal with Non-Gaussian noise characteristics employing integration algorithms based on Monte Carlo sampling (Pegoraro et al., 2017). It has also been employed along with Deep Learning algorithms in the context of bad data processing along with the support of data-driven learning in (Mestav; Luengo-Rozas; Tong, 2019). Recent works also explore temporal within information fusion concepts, merging measured information according to the different sampling and updating rates from diverse measurements (Massignan et al., 2019). Regarding spatial fusion, as far as the author recall, this is the first proposition for DSSE. In this context, the contributions of this chapter reside on:

- the demonstration of scalability of state estimation in a practical large scale distribution system, with high resolution three-phase unbalanced models from the low voltage circuits until the high voltage substations;
- a novel formulation for multiarea state estimation based on bayesian inference framework, using the concept of spatial fusion, that enables decomposing the network without local area observability constraints;
- a novel coordination scheme based on information fusion to obtain a probabilistic consensus among local areas.

A few theoretical topics, presented in previous sections, will be revisited in the following sections to provide a quick background of the fundamental details.

7.2 Multiarea Distribution System State Estimation

Real time monitoring and operation of DSs require an accurate assessment of their power grid conditions, to further support decision making at the DSOs, in a cyber-physical perspective, such as fault location, service restoration, voltage control, loss reduction, and asset management (Lefebvre; Prévost; Lenoir, 2014; CAMILLO et al., 2016; Primadianto; Lu, 2017; Massignan et al., 2018; FANUCCHI et al., 2021). Due to the large scale of such networks, the decomposition of the distribution systems and the application of MASE algorithms have been an important direction to achieve computational performance and scalability. Typically MASE approaches can be Hierarchical, as will be dealt in this papper (more adequate to a central DSO to deal with large scale networks in parallel multiprocessor architectures) (Falcao; Wu; Murphy, 1995) or Distributed (more adequate to decentralized operation paradigm and cloud computing) (Gómez-Exposito et al., 2011b; Pau et al., 2019).

This enables capturing the particularities, from both system as well as the available information and measurements, of each part of the distribution networks spatially inside their respective areas, briefly described below, and illustrated in Fig. 37:

- Primary Substations: high amount of real time measurements, low number of nodes (and state variables) and typically sampled information each minute from the SCADA system;
- Primary Feeders: low amount of real time measurements, high number of nodes (and state variables) and typically sampled information each minute from the SCADA system, complemented with pseudo measurements;
- Secondary Circuits: typically represented as aggregated pseudo measurements in the primary feeders, based on typical load profiles for each consumer and electricity consumption information, or as low voltage networks with medium amount of sensors (smart meters), high number of nodes (and state variables) and typically sampled information each 5 to 30 minutes from smart meters;

Given a electric power network composed by a total of A connected sub-areas , that is, a power system that if treated in a centralized state estimation would entail $A = 1$. As a consequence of such decomposition, the following concepts related to the state variables (the three-phase complex nodal voltages in all nodes of the network) and measurement vectors (sampled electrical quantities from different sensors) arise. Depending on the level of overlapping the state variables vector in each area may be composed by (Gómez-Exposito et al., 2011b):

- x_{ik}: state variables associated to the internal nodes of each area k;

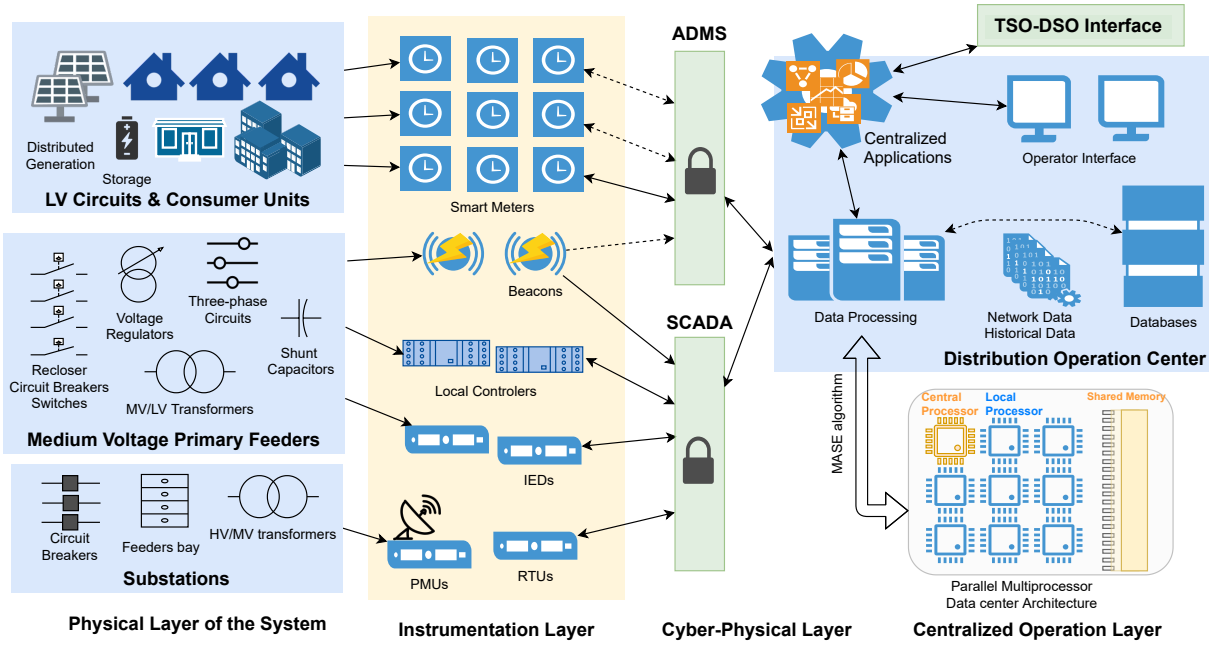


Figure 37 – Cyber-physical conceptualization of a distribution system in the context of centralized operation at the DSO. Different sources of information are gathered by instrumentation devices and exchanged by the SCADA and Advanced Distribution Management System (ADMS) to a central DSO to process them. Parallel computing architectures then support fast and scalable algorithms for decision-making.

- x_{bk}: state variables associated to the boundary nodes of each area k;
- x_{nk}: state variables associated to the internal nodes of adjacent areas to the area k (areas with extended overlapping zones).

Besides, the measurement vector can be devised for each area as:

- z_{ik}: internal measurements of each area k, that relate internal state variables x_{ik} and boundary state variables x_{bk};
- z_{bk}: boundary measurements of each area k, that relate boundary state variables x_{bk}, as well as the state variables of extended overlapping zones x_{nk};

The multiarea state estimation problem for a distribution network with m measurements and n state variables consists of a nonlinear measurement model that can be formulated as (Abur; Gómez-Expósito, 2004; Gómez-Exposito et al., 2011b; Bretas N. Bretas,):

$$z_{ik} = h_{ik}(x_{ik}, x_{bk}, x_{nk}) + e_i \tag{7.1}$$

$$z_{bk} = h_{bk}(x_{bk}, x_{nk}) + e_b \tag{7.2}$$

where z is the $(m \text{ x } 1)$ measurement vector composed by the measured values, virtual measurements and pseudo measurements; x is the $(n \text{ x } 1)$ vector of state variables (usually the complex nodal voltages); $h(x)$ is the $(m \text{ x } 1)$ set of nonlinear equations that relates the measurements with the state variables; and e is the $(m \text{ x } 1)$ noise vector assumed as independent normally distributed random variables, with zero mean and known covariance matrix R.

The above nonlinear measurement model comprise different electrical quantities monitored across the distribution systems. A three-phase unbalanced network model represents the diverse components, with their respective connections and phases, and further details are presented in previous chapters.

With the above decomposed measurement model, the Multiarea State Estimation problem is formulated according to the maximum likelihood principle yielding the weighted least squares formulation (Korres, 2011):

$$\min \quad J(x) = \frac{1}{2} r_b' . R_b^{-1} . r_b + \frac{1}{2} \sum_{k=1}^{A} r_{ik}' . R_{ik}^{-1} . r_{ik} \tag{7.3}$$

where the subscript b represents the boundary measurements for all sub-areas of the electrical network with the respective noise covariance matrix R_b, $r_{ik} = z_{ik} - h_{ik}(x_{ik})$ is the internal measurements residual for each area k and respective weighting matrix R_{ik}^{-1}, $r_b = z_b - h_b(x)$ are the boundary measurement residuals and respective weighting matrix R_b^{-1}.

It is easily shown there is a requirement of observabilitity for each area of the decomposed problem. It means that any subdivision of the power system must create observable local networks. This can be acceptable whenever a large deployment of sensors is accomplished, far from the reality of practical distribution systems (especially at the medium voltage feeders), or by including pseudo measurements in the estimation, what may hamper accuracy due to their relative poor quality. This work deals with such problem, enabling a more flexible decomposition of distribution networks, to accommodate the possibility of local areas without a complete observable measurement set.

As an example, Figure 38 illustrates such situation. The main motivation is to enable a multiarea state estimation procedure aligned with the deployment of smart meters at the low voltage secondary networks, while dealing with large scale unobservable primary feeders.

This work employs an overlapping-zone method as area decomposition, with a single overlapping node among adjacent areas. In this proposition, such decomposition is based on the natural spatial structure of distribution systems, going from the substation towards the secondary low voltage networks, and bounded by power transformers, whose terminals are both the boundary and overlapping boundaries for each area.

An illustrative example of the decomposition is illustrated in Figure 39, with details of the extended boundaries, state variables and measurements. An overlapping-zone with one node enables representing boundary measurements and injections locally. The boundary values are

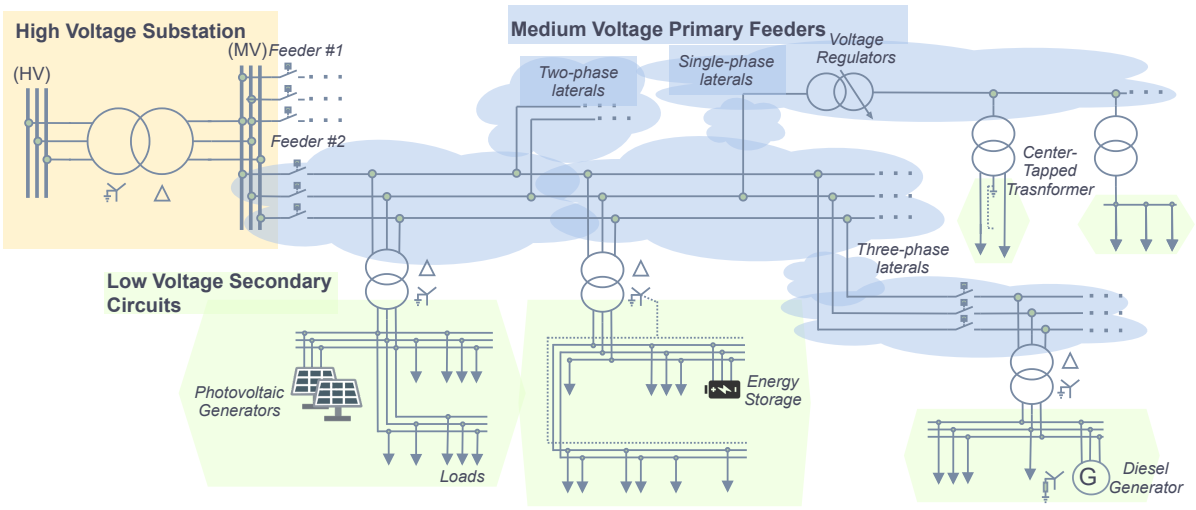


Figure 38 – Decomposition of distribution networks associated to each voltage level.

then exchanged in the Coordination Steps. The next section formulates the proposed multiarea DSSE based on Bayesian Spatial fusion principles.

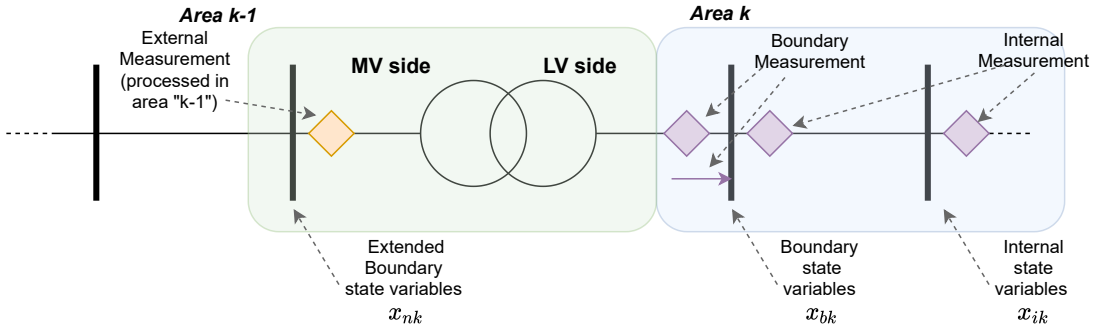


Figure 39 – Illustrative example of details of the extended boundaries in the proposed decomposition of distribution networks.

7.3 Bayesian Hierarchical Spatial Fusion for Multiarea DSSE

The Bayesian reasoning for multiarea power system state estimation extends the concept of the state vector as a random variable, in the sense that a set o possible state values may occur in the network given the measured observations and a prior knowledge about the system. Estimation then becomes performing inference on the probabilistic model of the posterior distribution instead of seeking for a fixed set of state values that yields the maximum likelihood. The proposed Bayesian Spatial Fusion comprise two-stages, as a hierarchical approach:

1. Local Estimation Step: an independent estimation step is performed for each area, employing internal, boundary measurements and extended boundary state variables previously exchanged;

2. Coordination Step: a central processor gathers boundary information from local estimation, and updates them to achieve consensual values among adjacent areas.

In the Local Estimation Step, the proposed spatial fusion extends the probabilistic model for each local area by including a prior distribution for the state variables, as the following:

$$x_{ik} = x_{ik}^{p} + \omega_{ik}^{p} \tag{7.4}$$

$$x_{bk} = x_{bk}^{p} + \omega_{bk}^{p} \tag{7.5}$$

$$x_{nk} = x_{nk}^{p} + \omega_{nk}^{p} \tag{7.6}$$

$$z_{ik} = h_{ik}(x_{ik}, x_{bk}, x_{nk}) + e_i \tag{7.7}$$

where x_{ik}^{p}, x_{bk}^{p} and x_{nk}^{p} are the expected values of the prior distribution assumed for the internal, boundary and extended boundary state variables, along with a random component to represent state variationsω_{ik}^{p}, ω_{bk}^{p} and ω_{nk}^{p}, respectively. Note that separate priors are given for the internal variables, the boundary variables and the extended boundary variables, to accommodate a reduced set of information exchange among areas, only the boundary and extended state variables and their priors are exchanged.

Applying the Bayes Theorem, yields the posterior distribution for each local area:

$$f_{X|Z}(x_{ik}, x_{bk}, x_{nk}|z_{ik}) = \frac{f_{Z|X}(z_{ik}|x_{ik}, x_{bk}, x_{nk}) f_X(x_{ik}, x_{bk}, x_{nk})}{f_Z(z_{ik})} \tag{7.8}$$

where, $f_{X|Z}(x_{ik}, x_{bk}, x_{nk}|z_{ik})$ is the conditional probability function of the state given the measurements in each local area, $f_{Z|X}(z_{ik}|x_{ik}, x_{bk}, x_{nk})$ is the likelihood function of the local area according to the measurement model, $f_X(x_{ik}, x_{bk}, x_{nk})$ is the prior distribution in the hierarchical model, and $f_Z(z_{ik})$ is the measurements probability of occurrence, a constant value that scales the posterior probability function and is often neglected.

The estimation process in each area is triggered as soon as the respective measurements become available, or if there is an update on the boundary state variables. It updates the state by a Maximum a Posteriori (MAP) estimation given by (7.9). The estimation result of each area is then used as the prior distribution for the state variables in the adjacent areas, a Bayesian hierarchical model (Congdon, 2007).

$$\hat{x}_k = \arg\max_{x} f_{X|Z}(x_{ik}, x_{bk}, x_{nk} \mid z_{ik}) \tag{7.9}$$

In this formulation, both boundary as extended boundary states are updated in the local estimation steps. Such estimation results are then exchanged with a central processor, that performs spatial fusion with the locally estimated boundary variables and boundary measurements neglected in the local stage. The essence of this coordination step it to obtain a probabilistic consensus for the boundary variables among the local areas, searching for coherent values among adjacent areas that are updated in next local estimation executions as new prior knowledge.

The coordination stage also extends the boundary measurement vector to accommodate a prior distribution for the boundary and extended boundary variables, as the following:

$$x_b = x_{bk}^{p-1} + \omega_{bk}^{p-1} \tag{7.10}$$

$$x_b = x_{nk}^{p-1} + \omega_{nk}^{p-1} \tag{7.11}$$

$$z_b = h_b(x_b) + e_b \tag{7.12}$$

where x_{bk}^{p-1} and x_{nk}^{p-1} are the expected values, and associated random characteristics ω_{bk}^{p-1} and ω_{nk}^{p-1} of the prior distribution, for the boundary and extended boundary state variables obtained in the previous local estimation execution $(p-1)$, and x_b is the boundary state vector for all areas. Note that, the extended boundaries obtained in the local estimations are also included, but in this case modelled along with the conditional relations with the boundary state variables. This introduces the results obtained by adjacent areas in the search for coherent values for the boundary states, as complementary prior information.

The Coordination step yields a common boundary state vector for all different areas. Such common boundary values are then exchanged back to the local areas to update their internal state variables using the latest boundary state variables available. The Coordination is iterative, in the sense that, boundary information keeps being exchanged between local areas as new local updates are provided, until the boundary state variables update is reduced below a numerical tolerance. Figure 40 illustrates the probabilistic model employed. Details about the state estimation solution in the local and coordination steps are described in the following subsections.

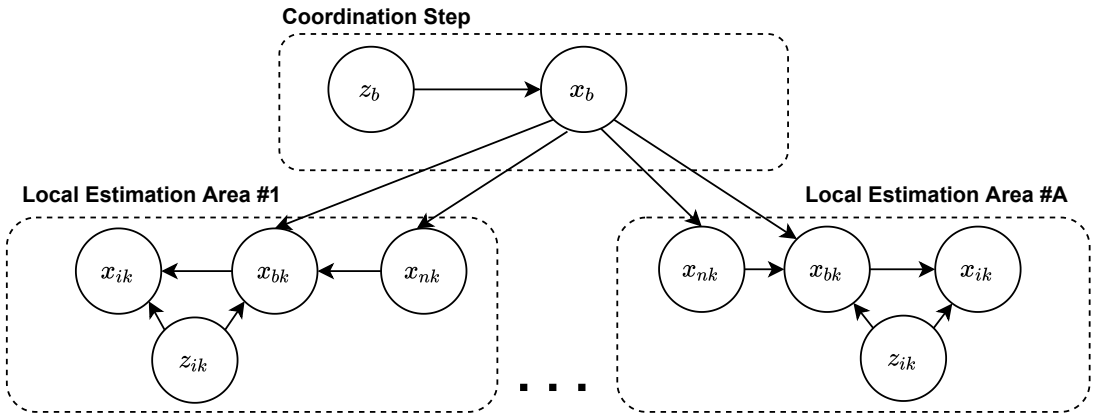


Figure 40 – Probabilistic representation of the proposed Multiarea DSSE. Prior information are included to complement local observability and also to coordinate the boundary state variables at adjacent areas.

7.3.1 Local Estimation Step

The Local estimation step provides the state vector of each area k comprised of $x_k = (x_{ik}, x_{bk}, x_{nk})$. This Thesis assumes a conjugate-Gaussian prior model for the state variables, to

accommodate prior knowledge while maintaining computational tractability. This assumption is a simplification of the real behaviour of the state variables, but fair enough since the Multiarea formulation is a decomposition of an original Gaussian process estimation. Besides, since the Coordination step reduces the effects of such approximations, since it updates iteratively the boundary state variables with the latest available information among different area. The superscript p denotes the iteration of the coordination step that provides the prior information.

$$\begin{aligned} Prior: \quad & x_{ik} \sim \mathcal{N}(x_{ik}^p, P_{ik}^p) && (7.13)\\ & x_{bk} \sim \mathcal{N}(x_{bk}^p, P_{bk}^p) && (7.14)\\ & x_{nk} \sim \mathcal{N}(x_{nk}^p, P_{nk}^p) && (7.15)\\ Likelihood: \quad & z_{ik} \mid x_{ik}, x_{bk}, x_{nk} \sim \mathcal{N}(h_{ik}(x_{ik}, x_{bk}, x_{nk}), R_{ik}) && (7.16) \end{aligned}$$

where P_{ik}^p, P_{bk}^p and P_{nk}^p are the state covariance matrices from the prior knowledge at each area, also updated in the coordination step, and R_{ik} the measurement covariance matrix for each area.

The prior knowledge in this local estimation can be represented by the following prior state vector and prior state covariance matrix, with each prior expected values and respective covariance from the above model:

$$\begin{aligned} x_k^p &= \begin{bmatrix} x_{ik}^p & x_{bk}^p & x_{nk}^p \end{bmatrix}^T \\ P_k^p &= \begin{pmatrix} P_{ik}^p & 0 & 0 \\ 0 & P_{bk}^p & 0 \\ 0 & 0 & P_{nk}^p \end{pmatrix} \end{aligned} \tag{7.17}$$

By assuming a Gaussian conjugate model for the priors and likelihood function, the following MAP estimation is obtained by the Bayes' Theorem, formulated as an unconstrained minimization problem:

$$\hat{x}_k = \min_x (z_{ik} - h_{ik}(x_k))^T R_{ik}^{-1} (z_{ik} - h_{ik}(x_k)) + (x_k - x_k^p)^T (P_k^p)^{-1} (x_k - x_k^p) \tag{7.18}$$

A nonlinear optimization algorithm, the Modified Newton method with backtracking algorithm for the step length α, provides the local estimates, with an iterative update of the state variables, according to the following equations:

$$\left(H^T R_{ik}^{-1} H + (P_k^p)^{-1}\right) \Delta x_k^{it} = H^T R_{ik}^{-1} (z_{ik} - h_{ik}(x_k) + (P_k^p)^{-1} (x_k^{it} - x_k^p) \tag{7.19}$$

$$x_k^{it+1} = x_k^{it} + \alpha \Delta x_k^{it} \tag{7.20}$$

with, H the Jacobian matrix of $h_{ik}(x_k)$ evaluated at each iteration it point x_k^{it}. The algorithm is iterative until convergence is met according to a numerical tolerance ($tol = 1.0E^{-5}$), with $|\Delta x_k^{it}|_{\text{inf}} \leq tol$.

This Thesis also employs an orthogonal formulation to improve numerical conditioning and computational performance in each local estimation step (Davis, 2011a; Hebling et al., 2020), with the following factorization.

$$\begin{pmatrix} (P_k^p)^{-1/2} \\ R_{ik}^{-1/2} H \end{pmatrix} = \mathscr{Q}^T \mathscr{R} \tag{7.21}$$

The iterative solution then becomes:

$$\mathscr{R} \Delta x_k^{it} = \mathscr{Q} \begin{pmatrix} (P_k^p)^{-1/2} (x_k^{it} - x_k^p) \\ R_{ik}^{-1/2} (z_{ik} - h_{ik}(x_k^{it})) \end{pmatrix} \tag{7.22}$$

Besides, exploring the above optimization model yields the advantage of the proposed Multiarea method, the ability to include a prior knowledge to compensate for the lack of local observability. The demonstration is straightforward by exploring the rank of the matrix being factorized. Thus, if a complete prior knowledge is given about the state variables, that is $P_k^p > 0$ and full rank, it is possible to perform separate and independent local estimation steps disregarding the full observability in the decomposition. It is noteworthy that, if the full network is not observable, then the prior knowledge will be the sole information to provide knowledge about the local state variables. This shows the importance of the coordination step, that updates the boundary information in order to include real-time information from adjacent areas into such non-observable areas.

$$H^T R_{ik}^{-1} H + (P_k^p)^{-1} > 0 \tag{7.23}$$

After convergence, the boundary and extended boundary state variables are exchanged with the central processor. Both the estimated values and posterior variance are shared with the central Coordination, as the following:

$$x_{bk}^p = \hat{x}_{bk} \tag{7.24}$$
$$x_{nk}^p = \hat{x}_{nk} \tag{7.25}$$
$$\tag{7.26}$$

$$P_{bk}^p = \left[(H^T R_{ik}^{-1} H + (P_k^p)^{-1})^{-1} \right]_{bk} \tag{7.27}$$
$$P_{nk}^p = \left[(H^T R_{ik}^{-1} H + (P_k^p)^{-1})^{-1} \right]_{nk} \tag{7.28}$$
$$\tag{7.29}$$

For the boundary covariance matrices above, only the respective elements (rows and columns) relative to the boundary and extended boundary state variables are considered, represented as a submatrix with the subscripts $[]_{bk}$ and $[]_{nk}$.

Prior knowledge may be constructed in a non-informative way, such as assigning large variances (low weights) to a standard operational condition (or a flat start situation). It can also be based on informative knowledge, from previous operational conditions, from load calculations, or sampled from typical load behavior, with tuned variances to represent a flexible uncertainty about the state. In this Thesis we assume a non-informative prior, based on a flat voltage profile, and the boundary state variables are updated by the coordination step.

7.3.2 Coordination Step

The Local Estimation Step results on the state variables of each area, and values for their own boundaries that are shared along with the central processor. The Coordination Step processes these local boundaries information in order to provide common and coherent boundary values among adjacent areas, thus providing a complete overview of the entire distribution networks, from the primary substations to the low voltage secondary networks.

The solution of the coordination is obtained by employing the remaining boundary measurements along with the boundary and extended boundary state variables estimated in the local areas. The coordination state vector comprises the full set of boundary variables. The goal in this step is to obtain coherent values among adjacent areas that will be exchanged back again to each local area.

As performed in the local areas, the Coordination step also employs a conjugate-Gaussian prior model for the boundary state variables. In this case, the state vector comprises $x_b = [x_{b1}, x_{b2}, ..., x_{bA}]^T$, the boundary values among all areas. And the prior knowledge is arranged as the following, using the local estimation results exchanged, from both the boundary and extended boundary obtained in each area:

$$\begin{aligned} x_{bk}^{p-1} &= \begin{bmatrix} x_{b1}^{p-1} & x_{b2}^{p-1} & \dots & x_{bA}^{p-1} \end{bmatrix}^T \\ P_{bk}^{p-1} &= \begin{pmatrix} P_{b1}^{p-1} & 0 & \dots & 0 \\ 0 & P_{b2}^{p-1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & P_{bA}^{p-1} \end{pmatrix} \end{aligned} \tag{7.30}$$

$$\begin{aligned} x_{nk}^{p-1} &= \begin{bmatrix} x_{n1}^{p-1} & x_{n2}^{p-1} & \dots & x_{nA}^{p-1} \end{bmatrix}^T \\ P_{nk}^{p-1} &= \begin{pmatrix} P_{n1}^{p-1} & 0 & \dots & 0 \\ 0 & P_{n2}^{p-1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & P_{nA}^{p-1} \end{pmatrix} \end{aligned} \tag{7.31}$$

The coordination step is also comprised by a MAP estimate, given by the following optimization problem:

$$\hat{x}_b = \min_x (z_b - h_b(x_b))^T R_b^{-1} (z_b - h_b(x_b)) + (x_b - x_{bk}^{p-1})^T (P_{bk}^{p-1})^{-1} (x_b - x_{bk}^{p-1}) + (x_b - x_{nk}^{p-1})^T (P_{nk}^{p-1})^{-1} (x_b - x_{nk}^{p-1}) \tag{7.32}$$

A nonlinear optimization algorithm, the Modified Newton method with backtracking algorithm for the step length α, then provides the local estimates, with an iterative update of the state variables, according to the following equations:

$$\left(H^T R_b^{-1} H + (P_{bk}^{p-1})^{-1} + (P_{nk}^{p-1})^{-1}\right) \Delta x_b^{it} = H^T R_b^{-1} (z_b - h_b(x_b) + (P_{bk}^{p-1})^{-1} (x_b^{it} - x_{bk}^{p-1}) + (P_{nk}^{p-1})^{-1} (x_b^{it} - x_{nk}^{p-1}) \tag{7.33}$$

$$x_k^{it+1} = x_k^{it} + \alpha \Delta x_k^{it} \tag{7.34}$$

in this case, H is the Jacobian matrix of the boundary measurement model $h_b(x_b)$.

If any measurement is left to be processed in the coordination step the above problem becomes linear, only the updated prior knowledge among different areas is considered in the above equations, and the boundary state variables are obtained in a single iteration. This is accomplished by the extended boundary approach, where boundary measurements can be included in each local estimation since the state vector is extended one node deep into adjacent areas.

Another aspect that can be exploited is an hierarchical weighting during the coordination step, increasing the prior weights (reducing the variance) in a *Top-Down* approach, or from areas that are observable. This way is possible to reinforce the values obtained for the boundary state variables at the first and highest voltage levels when compared to prior values from the medium voltage levels and then to the secondary voltage levels, that is, from the substation to the low voltage nodes.

Finally, the boundary state vector is exchanged back to the local estimation steps. Different strategies to trigger the coordination step may also be employed, such as waiting for local convergence or by a pre-defined number of local iterations. This exchange is performed after local convergence is achieved. Distributed approaches may apply this strategy, by updating boundary variables and performing the coordination step according to communication latency and data volume constraints, but further exploration is required to define the best exchange strategies according to practical constraints (communication bandwidth, refresh rate, data volume, etc.).

7.3.3 Implementation and Algorithm

This section presents implementation details and the algorithm of the proposed MASE approach. This Thesis employs sparsity treatments to enhance numerical and computational performance. As presented previously, the orthogonal formulation avoids the explicit matrix multiplications in the Local Estimation and Coordination Steps, which by itself represents an advancement in computational aspects for ill-conditioned distribution networks. By using

sparse representation memory and computational operations are significantly reduced, improving even further the computational performance. Besides, by employing sparse ordering algorithms reduces the total number of fill-ins in the factorization process, reducing even further ill-conditioning and improving numerical stability. The *SuiteSparse* computational library (Davis, 2011a) is employed with a sparse Multifrontal QR method based on block-Householder factorization, along with an Approximate Minimum Degree (AMD) ordering technique.

Another aspect that improve computational performance is performing local estimation in parallel. This enables enhancing the computational time by performing burdensome tasks in parallel. This Thesis employs a shared memory architecture to facilitate data integration in modern multi-core processors. The *OpenMP* programming paradigm (OpenMP, 2020) is employed to implement the local estimation in parallel, with special care in memory access and data integration, in a shared-memory architecture. A synchronization barrier triggers the coordination step after all local estimation steps converged.

The implemented algorithm is illustrated in the flowchart presented in Figure 41. The flowchart illustrates the parallel computing paradigm along with the main database integration required in practical deployment. The algorithm focuses on the exchange of information between the local steps and the coordination, as well as the computations performed.

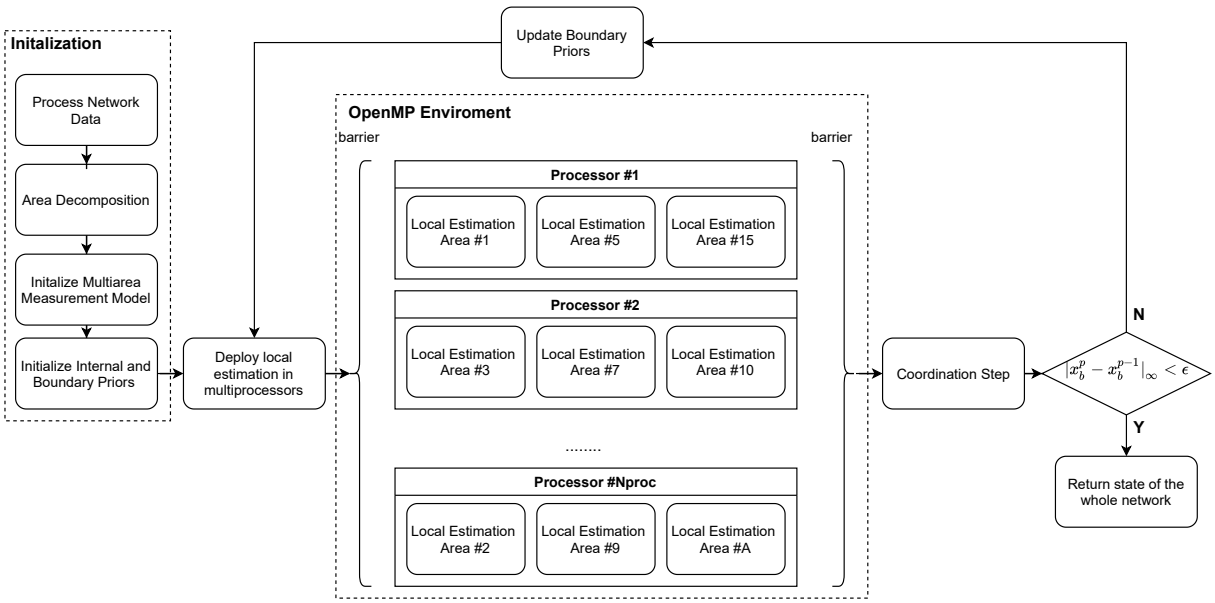


Figure 41 – Flowchart of the proposed MASE implementation in a parallel computing architecture with shared memory paradigm. Local estimations are executed in multiprocessors and coordinated sequentially by a central processor after local convergence.

7.4 Simulation Results and Application in Large Scale Distribution Systems

7.4.1 Accuracy analysis

Monte Carlo simulations were employed to represent the distribution system behaviour, and the associated measurement set, by including random noise in a load flow scenario considered as reference case. This simulation evaluates the proposed MASE performance by random sampling a noise characteristic for each measurement and associated electrical quantity, in a total number of Monte Carlo trials (100 repetitions). The noise samples are generated for each measurement, individually, according to the respective measurement standard deviation, as described by the following equation (Castillo et al., 2011):

$$z_i = z_i^{ref} + u_i(|z_i^{ref}|pr_i)/3 \tag{7.35}$$

where, z_i is the value of measurement i with a random noise and is used as input on the state estimator, z_i^{ref} is the respective value of the electrical quantity obtained in the reference scenario, that is, the load flow calculation, pr_i is the measurement precision according to its precision class (assumed as 2% for active and reactive power flow or injection measurements from SCADA systems, 5% for smart meters, 1% for voltage magnitude measurements of the SCADA systems, and 30% for pseudo measurements obtained from typical load curves), and u_i is a random variable with standard Gaussian distribution. With the emulated measurements and the respective solution from the reference load flow, the proposed MASE performance is evaluated through the Mean Absolute Error (MAE) index:

$$\text{MAE} = \frac{1}{n_{trials}} \sum_{K=1}^{n_{trials}} \left| \hat{x}^K - x^{ref} \right|, \tag{7.36}$$

where $\hat{x}$ is the estimated state in each repetition K of the Monte Carlo simulation, x^{ref} is the reference state value, and n_{trials} is the number of trials during the simulation ($n_{trials} = 100$).

The simulation results were obtained using a C/C++ implementation on UNIX platform in a microcomputer with a Core i7-9750H 2.60 GHz and 8 GB RAM. Besides, to evaluate the isolated effect of the multiarea decomposition, the measurements are assumed in a single snapshot, as a single measurement vector, such as in the case of the hybrid state estimation, that is, neglecting different sampling rates among different sensors.

The simulations were carried out with the IEEE US Low Voltage System (IEEE342) test system. It consists of a low voltage urban network with high reliability composed by spot loads in 408 V and a meshed low voltage network in 208 V, with grounded Wye connection. Eight primary feeders in 13.2 kV comprise the medium voltage system in Delta connection. Finally the 230/13.2 kV substation with two delta connected transformers and a small portion of the sub-transmission are represented. It consists of a highly complex network, where increased reliability is accomplished by a meshed low voltage network, also requiring a deeper resolution

for the entire network model. The network model and reference load flow scenario are available in (IEEE, 2018).

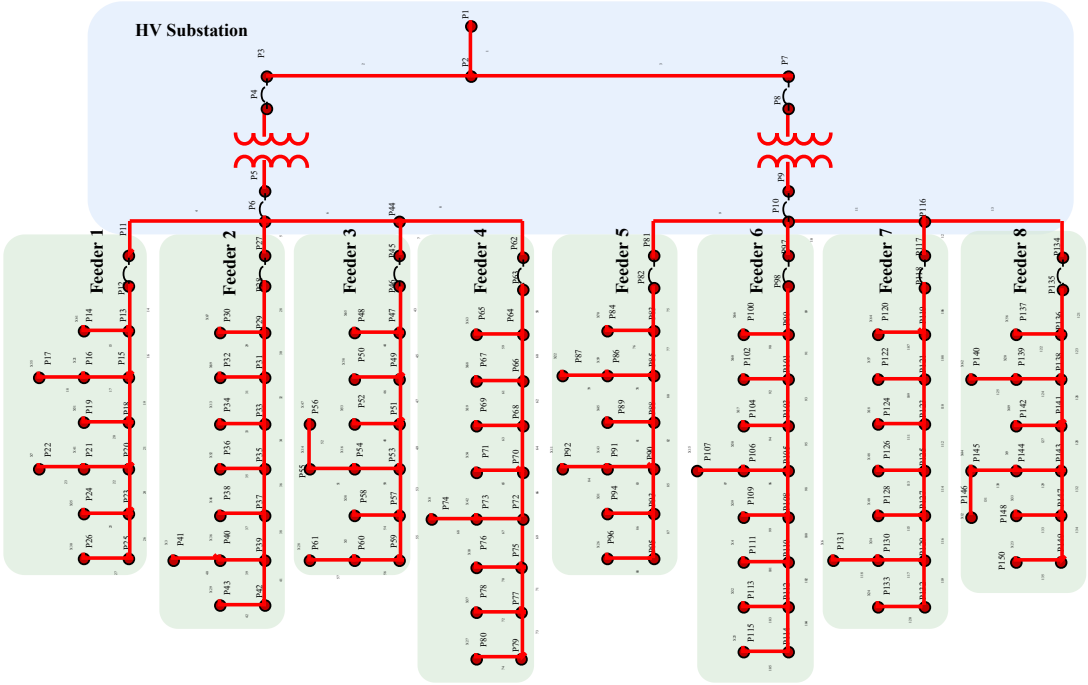


Figure 42 – Decomposition of the high voltage substation and medium voltage feeders in the IEEE US Low Voltage test feeder (IEEE, 2018).

The following metering system was considered during the simulations: high voltage substation and sub-transmission lines with active and reactive power flows and voltage magnitude measurements in all terminals from the SCADA system; primary network with active and reactive power flow and voltage magnitude measurement only at the feeder's bay, from the SCADA system; low voltage networks with active and reactive power injection and voltage magnitude obtained from smart meters. The full network also presents several virtual measurements representing nodes without loads, that is, nodes with zero active and reactive power injections, with a very low standard deviation (assumed as 1.0 E-7). Table 10 presents the amount of measurements considered in each part of the distribution system, where each metering device provides the respective electrical quantities in each phase (phase ABC).

Regarding the area decomposition, a level-based approach was devised in this simulation. This way the 230/13.2 kV substation and sub-transmission lines correspond to a particular area; each medium voltage feeder correspond to a separate area, with eight total areas for the primary network in 13.2 kV; and each spot load in the 408 V low voltage networks correspond to an individual area, and the meshed 208 V low voltage network is divided in four areas. The boundaries of the different areas are defined by the power transformers in the network, with a total of 21 areas. Figures 42 and 43 present an illustration of the area decomposition in each part of the distribution system.

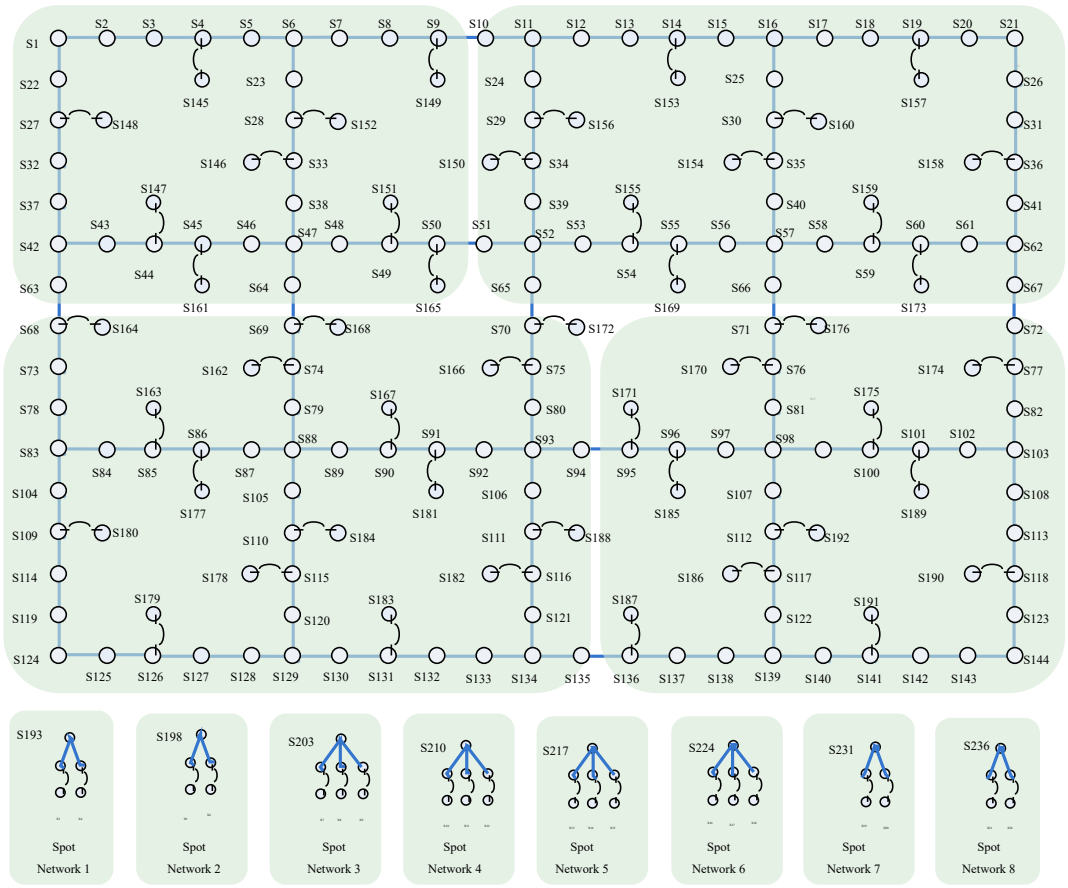


Figure 43 – Decomposition of the low voltage circuits in the IEEE US Low Voltage test feeder (IEEE, 2018).

Table 10 – Metering system location and Total amount per location in the distribution system (Each presenting the respective three-phase information)

Sources of Information	**Three-Phase Electrical Quantities**	**Substation**	**Primary Network**	**Low Voltage Network**
SCADA	Active and Reactive Power Flow	14	8	-
	Active and Reactive Power Injections	1	0	-
	Voltage Magnitudes	5	8	-
Virtual Measurements	Active and Reactive Power Injection (Zero Injection)	6	131	68
Smart Meters	Active and Reactive Power Injection and Voltage Magnitude	-	-	104

To demonstrate the solution of the multiarea state estimation problem with the proposed hierarchical approach, Figure 44 presents the estimation error for all state variables of the network (voltage magnitude and phase angle) compared to the estimated results from the centralized approach. The figure shows the results in all three phases for all nodes of the system. During the simulations each local estimation stage converged in 4 iterations. Regarding the coordination step, a total of 4 executions were needed to converge, that is, without a sensible change in the state variables for more executions of the top-down coordination. The centralized approach

converged also in 4 iterations, but with an increased computational burden. Figure 45 presents the coordination step convergence in this scenario, where after 5 iterations the method stops to update the boundary state variables. Figure 46 presents the computational time of the proposed MASE compared with the centralized approach, considering multiple threads (parallel tasks for the local estimation). Even with a single thread, the MASE presents computational gain, since it is less burdensome to deal with smaller networks on the local estimation procedure. However there is plateau of computational gain, which shows further coding optimizations may improve the performance, such as improving memory locality and optimizing communication bandwidth usage.

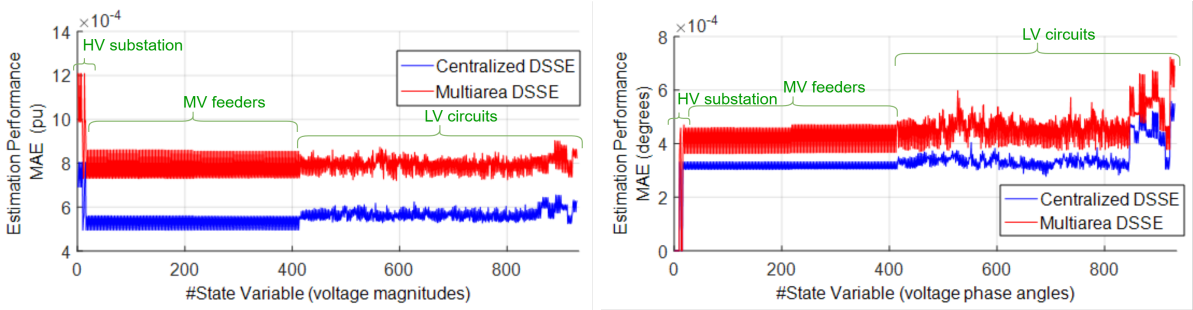


Figure 44 – Estimation performance for the state variables (voltage magnitude and phase angles) of the IEEE US low voltage system, with the centralized and the multiarea approaches.

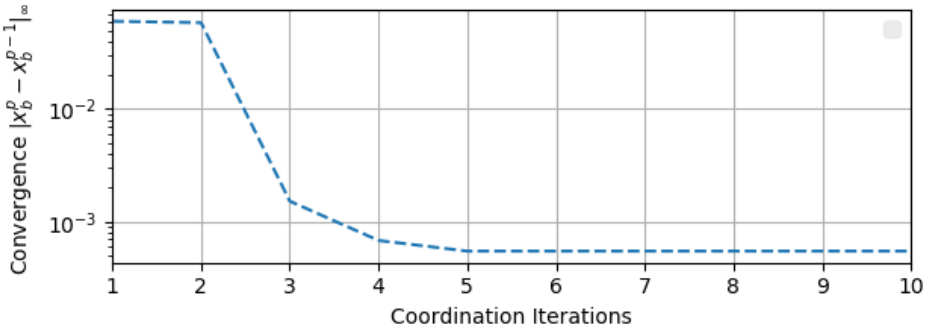


Figure 45 – Convergence of the coordination step, the outer loop of the hierarchical MASE.

Regarding accuracy of the estimation process, the hierarchical approach carries an additional bias, reducing its accuracy, due to the fact it does not use all the informations at once, as presented in Table 11. This bias consists of a trade-off between accuracy and computational performance to achieve scalability, and must be taken into account depending on the requirements for each final application.

The estimation error was also evaluated for the loads at the secondary low voltage networks (active and reactive power injections) and for the active and reactive power flows at the MV/LV transformers (13.2/0.48 and 13.2/0.208 kV). To clarify the impact on estimated electrical

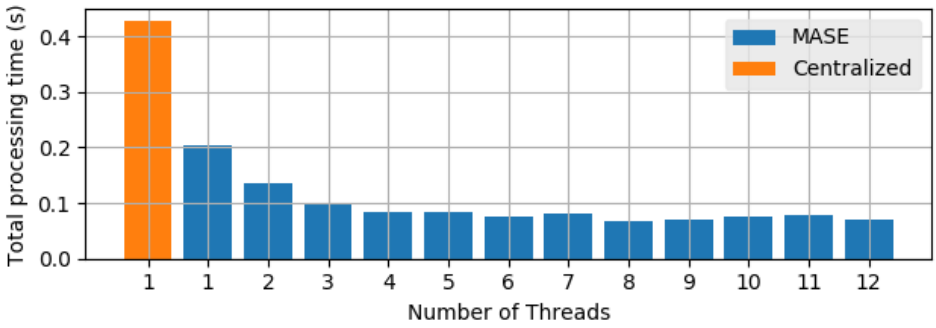


Figure 46 – Comparison of the total computational time of the DSSE methods, with different amount of threads.

Table 11 – Comparison of the estimation error in the presence of pseudo measurements and with smart meters in the proposed MASE.

Estimated Electrical Quantity	**With pseudo measurements**			**With Smart Meters**		
	Substation	*Primary Network*	*Low Voltage Network*	*Substation*	*Primary Network*	*Low Voltage Network*
Voltage Magnitude (p.u)	0,000939	0,000779	0,002069	0,000938	0,000779	0,000796
Voltage Phase-Angle (degrees)	0,01085	0,024339	0,18017	0,01085	0,024339	0,026962
Active Power Loads (%)		23,637			1,3106	
Reactive Power Loads (%)		22,843			1,3066	
Active power flow in transformers (%)		8,547			0,49068	
Reactive power flow in transformers (%)		9,8906			0,62925	

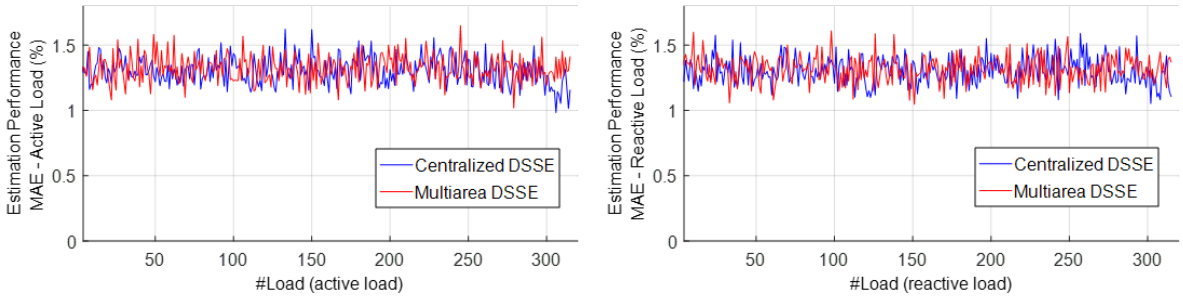


Figure 47 – Estimation performance for the power injections (active and reactive loads) of the IEEE US low voltage system, with the centralized and the multiarea approaches.

quantities on the network, Figure 47 presents the estimation error, mean absolute percentage error, for the loads in all phases and nodes of the system. Figure 48 presents the estimation error, mean absolute percentage error, for the MV/LV transformers power flows.

Both centralized and MASE approach presented similar accuracy for the load estimation, around 1.5 % of accuracy, a consonant improvement on the initial assumed precision for the smart meters (5 %). A similar result is also observed for the power transformers, but in this case with an increase on their relative precision of power flow estimation when compared to the accuracy of the loads. This fact is essentially related to the fact that the uncertainty of many

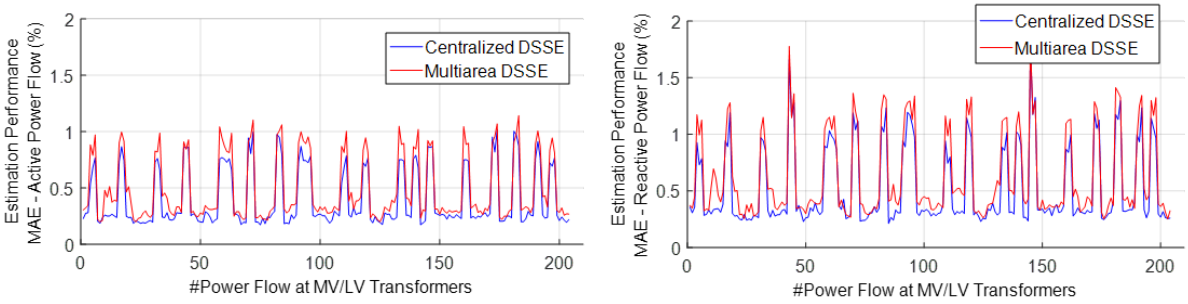


Figure 48 – Estimation performance for the active and reactive power flow at the MV/LV transformers of the IEEE US low voltage system, with the centralized and the multiarea approaches.

loads is redundantly considered all together for the calculation of the power flows on the MV/LV transformers of the meshed network. For the case of spot loads, such accuracy increase is not that expressive, since only the uncertainty of a single load monitored by the respective smart meter provides the information regarding the transformers loading. This is an advantage of properly representing each individual load instead of directly aggregating them as loads in the primary feeders, which may not consider such effect properly depending on the aggregation strategy.

It is noteworthy that the smart meters were allocated only at the consumer units (loads) and that the MV/LV transformers are not monitored by any sensors. Besides, different indicators regarding the operational condition of the network, such as power losses, voltage unbalance, voltage drop, equipment loading, among others, were also estimated with similar precision.

7.4.2 Application in Large Scale Medium Voltage System

To evaluate scalability with practically sized distribution networks, the proposed MASE was applied in a real distribution system of a Brazilian utility. The pilot region comprise an extension of 33 cities under the responsibility of a single central operator. The region comprises almost 1 million inhabitants and an area of more than 12 thousand km^2 in the south of Brazil. The distribution system comprises 48 substations, with 35 primary feeders in 34.5 kV and 241 primary feeders in 13.8 kV, illustrated in Figure 49. The size of the region was chosen to capture the spatial location of both types of feeders, since the majority of the 34.5 kV spread across different cities in long urban and rural networks. All feeders are represented according to their unbalanced and asymmetrical characteristics, and Table 12 presents a summary of the main characteristics to illustrate the composition of asymmetric connections in the primary feeders.

Regarding the measurement set available, the system is majorly monitored by pseudo measurements obtained from typical load profiles and load aggregation procedures (Massignan et al., 2018). The SCADA system comprises measurements at the primary substations, presenting voltage magnitude and active and reactive power at each feeder bay. The utility also has some special consumers telemetered, large consumers connected directly to the primary network,

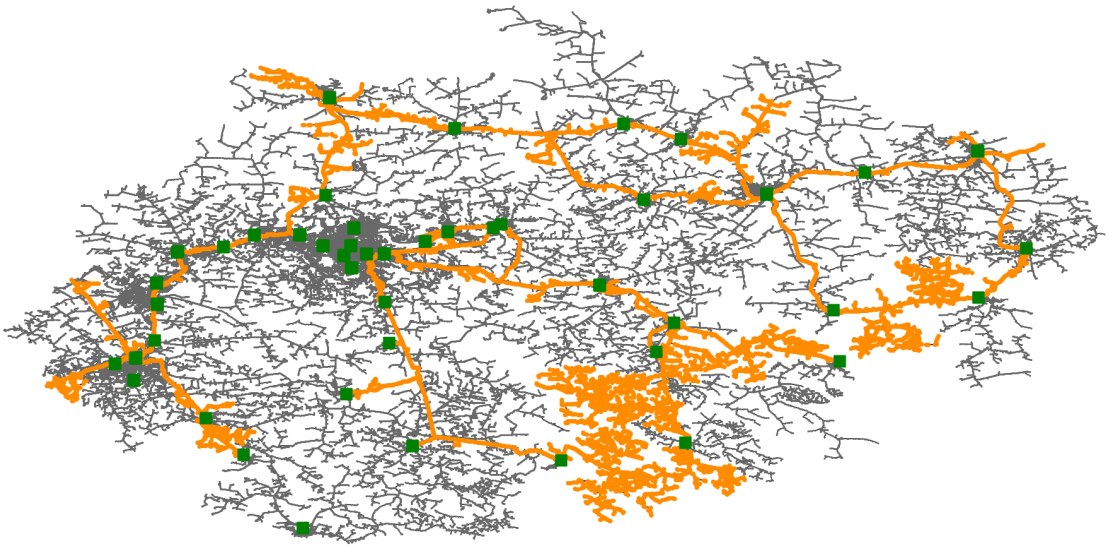


Figure 49 – Extended Large Scale Distribution Network in Brazil. The distribution networks comprise 35 feeders in 34.5 kV (orange) and 241 feeders in 13.8 kV (gray), into 48 substations (green).

Table 12 – Length and asymmetry on medium voltage distribution feeders

Voltage Level		**Phase A**	**Phase B**	Phase C	**Phases AB**	**Phases BC**	**Phase CA**	**Phases ABC**
13.8 kV	*Total length (km)*	0.40	0.12	-	41.62	2,260.83	86.40	6,740.35
	Percentage	0.004%	0.001%	0.000%	0.456%	24.763%	0.946%	73.829%
34.5 kV	*Total length (km)*	314.82	513.64	440.66	-	0.48	5.85	871.37
	Percentage	14.664%	23.926%	20.526%	0.000%	0.022%	0.272%	40.589%

Table 13 – Measurement set quantitative characteristics per phase and type.

Measurement Set		**Phase A**	**Phase B**	**Phase C**
Virtual	Active and Reactive Power Injections	220,633	222,060	222,103
Pseudo	Active and Reactive Power Injections	24,656	23,229	23,186
SCADA	Active and Reactive Power Flows	276	276	276
	Voltage Magnitude	276	276	276

along with current magnitude measurements spread across the feeders (at voltage regulators and automatic reclosers), but were not considered in this simulation. Table 13 presents quantitative details of the information and metering system (measured data, pseudo measurements and virtual measurements). The mean level of redundancy considered in the test is 1.05, a very low redundancy as typically encountered in distribution systems. The main purpose is to evaluate the scalability of the proposed approach and implementations made.

The decomposition of this system was performed in a feeder and substation basis, in the sense that each feeder is a local area, and also each substation is an upstream area. Figure 50 presents the histogram of number of variables per each local area, with the whole system presenting a total of 245,565 nodes, that together sum more than 1 million state variables (1,174,134). The largest feeder presents 4,500 nodes (approximately 27,252 state variables), and the system presents a mean number of 700 nodes per feeder. This test intends to explore details regarding computational scalability, as presented in the sequence.

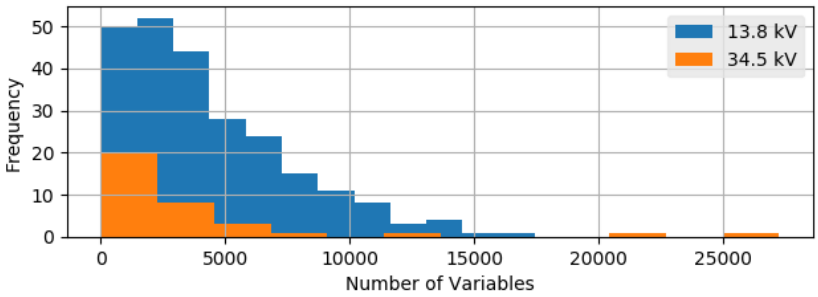


Figure 50 – Histogram of number of state variables per each local area. The 276 primary feeders present more than 1 million state variables for the system as a whole.

Initially, Figure 51 presents convergence rate of the local estimations, showing a boxplot of the reduction on the state variable changes in all 276 feeders per iteration. It illustrates a quadratic characteristic of convergence was obtained. This, however, was only possible due to the orthogonal formulation and the sparsity treatment employed. To illustrate the importance of the sparsity treatment and the associated numerical stability from the orthogonal formulation, Figure 52 illustrates the reduction of nonzero elements in the linear system of the local estimation, in equation (7.22), due to sparse ordering, and the largest singular value as a metric of numerical conditioning. The reduction on nonzero elements, from almost 25% to less than 5% in all feeders, increases computational performance, since less operations are performed to solve the estimation problem, and also increases numerical stability, since it reduces fill-ins during the factorization process.

Another important aspect for real time applications at DSOs is the processing time and memory allocation features for the estimation. Figure 53 presents the individual computational time and memory allocation for each local estimation performed. As it can be seen, less than 2 seconds are required to perform local estimation, and the majority falls below 400ms, in accordance to the real-time requirements of energy management systems. It is noteworthy that there is space for enhancements, such employing optimized general purpose routines, object-oriented programming and benchmark libraries. Regarding memory allocation, each local area present a relative small, which tends to grow linearly, with a total requirement of almost 950 MB. This, characteristic is a positive consequence of employing sparse linear algebra routines, along with reduced data structures to model the network, and can be further optimized with better software engineering practices.

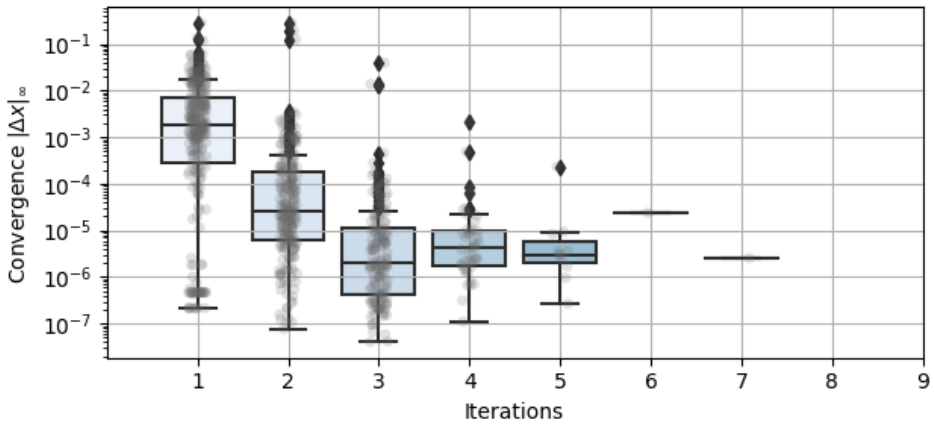


Figure 51 – Boxplot of iterations per each local estimation among the 276 feeders.

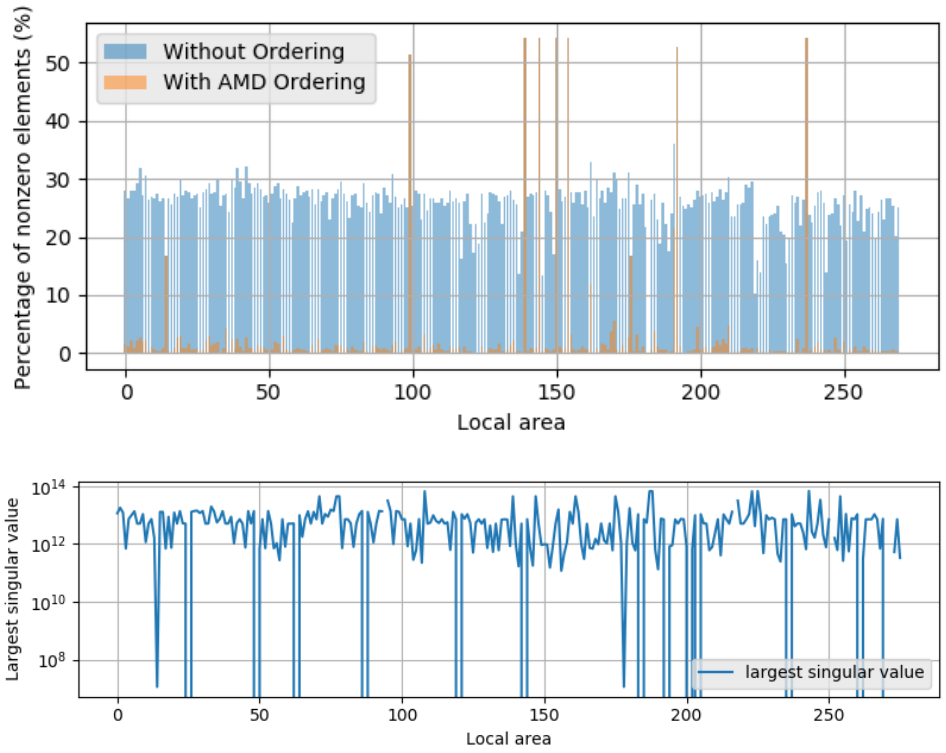


Figure 52 – Sparsity index (percentage of nonzero elements) with the AMD ordering technique, and the associated largest singular value as a metric of numerical conditioning.

Finally, the effects of parallel computing using a shared memory paradigm with OpenMP was evaluated by increasing the number of parallel threads created (as a reference the i7-9750H processor presents 6 physical processor, with hyper-threading enabled, with a total of 12 concurrent threads). Figure 54 presents the total computational time of the MASE procedure employing a different number of threads. For comparison purposes, a Backward/Forward Sweep method, with similar implementation of routines and data structures, solves a load flow problem in this large-scale network in about 3 seconds (while maintaining less than 20 ms per feeder).

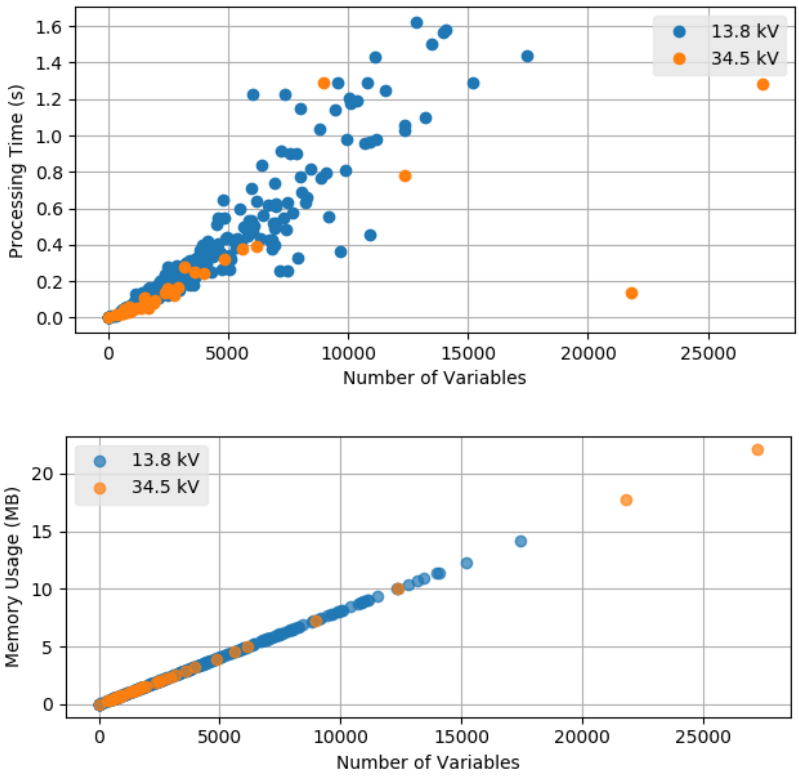


Figure 53 – Individual local area processing time and memory allocation requirements for each of the 276 feeders.

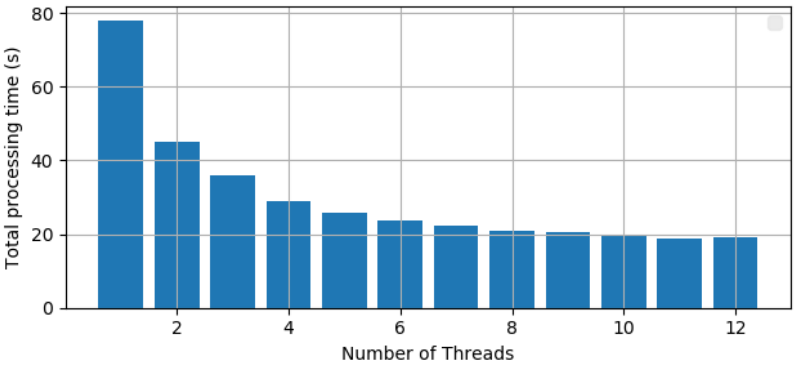


Figure 54 – Total computational time of the proposed MASE using different amount of parallel threads.

Regarding other implementations in the literature, the work in (Zhou et al., 2020b) presents computational time from 50 to 160 seconds in a single feeder with 4,500 nodes, decomposed in 4 areas, aiming at one iteration per second. Another work in (Pau et al., 2019) provides computational time for a 3,500 nodes system, decomposed in 166 areas, in less than 10ms for WLS approaches, less than 1ms for backward/forward sweep methods, and less than 300ms for a WLAV method, while a centralized WLS converges in 13 seconds. Such comparisons must be taken cautiously, since several implementation issues and characteristics affect the

performance, and coding optimization have a significant impact. However, they illustrate an important assessment to be made for any practical application, to deal with very large scale networks in less than 1 second is a desirable performance characteristic for novel real time applications. It also enable using the models and implementations for different applications that may require sampling different operational conditions of the networks, such as optimization and voltage control.

In a comparison to the Bayesian Information Fusion, presented previously, the Multiarea Bayesian DSSE employs the prior distributions for supporting spatial decomposition of the networks. This enables complementing observability, while exchanging boundary information with a central processor in a Coordination step. Despite the measured values were mixed into a single measurement set, similar to conventional state estimators, the prior knowledge could in fact accommodate sources information gathered in different instants, in a spatio-temporal manner. In essence, the extrapolation of employing simultaneously both concepts, can increase accuracy, perhaps even reducing the estimation bias below numerical tolerance, but unfortunately was not explored so far. This shows once again the strength and versatility of the Bayesian framework presented throughout this Thesis. It supports further research towards an architecture to orchestrate different features of the Bayesian framework, which is definitely an exciting direction to be pursued, and will be discussed in the final Chapter 8.

7.5 Concluding Remarks

This chapter conceived a hierarchical methodology for real time monitoring for distribution systems based on a multiarea state estimation concepts. The proposed MASE intends to ensure deployment on large scale distribution networks, maintaining numerical and computational performance. A decomposition of the distribution network according to different levels enable estimation from high-voltage substations, medium voltage primary feeders and the low voltage secondary circuits. A local estimation stage based on an efficient and numerically robust orthogonal method deals with inner ill-conditioning of the local areas and ensure high computational performance. A coordination step then updates boundary state variables among different areas.

The proposed approach presented similar accuracy than centralized approaches, with a slight bias on the estimation of state variables, but still presenting accurate estimations for meaningful electrical quantities of the distribution networks, such as loads and power flows. Conversely, a significant increase on computational performance is obtained by the decomposition strategy and also with a sensible reduction on data volume and memory requirements. This implies on faster implementations and also more flexibility on hardware requirements on distribution operation centers. Besides it also opens the possibility to employ edge computing within decentralized monitoring architectures, specially in the context of local microgrid controllers associated with local energy markets, an interesting future work direction.

It is noteworthy that the proposed MASE can be deployed into large scale distribution networks with high operational complexity, dealing with three-phase, unbalanced, in the presence of meshed networks. The methodology also considers different types of components and connections, as well as different precision levels of measurements, from SCADA, pseudo measurements and smart meters without simplifications in the nonlinear measurement model.

Future works may seek for improvements on building the prior knowledge, with the goal of reducing the optimality gap with enhanced prior knowledge. Another important direction is towards further increments on computational performance exploiting advanced computing architectures, such as cluster deployment with non-uniform memory access paradigm and also by leveraging data parallelism in graphical processing units. Besides, also regarding scalability, different approaches for estimation can be harmonically orchestrated within the proposed Bayesian formulation, what increase the possibility of employing diverse algorithms specialized in different parts of the system.

8 CONCLUSIONS AND FINAL CONSIDERATIONS

The digitalization of power grids is enabling a smooth transition towards decarbonized and flexible energy systems, ensuring higher penetration of renewable energy, creating new local energy markets and increasing safety and reliability in power system operation. An associated high level of automation ensures such digitalization by incorporating new data sources across the whole power grid, from the bulk transmission system to the deepest levels of distribution networks and the end-consumers. Nonetheless, the data by itself provides only a marginal observation of the power system condition. In this sense, state estimation emerges as a bridge between data and physics-based models, processing such information and detailed network models that capture the electric power system essence, while reducing the pervasive effects of noise.

This Thesis presents new perspectives to support power system state estimation, especially for distribution systems. The main goal is addressing a diverse set of problems from classical state estimation formulation, without relying on approximations for the measurement model or linearisation on power flow equation, while enhancing a probabilistic interpretation for the state variables. Three main challenges tackled in this work were:

- Enhancing the probabilistic characterization of the state estimation problem;
- Including temporal aspects in the estimations, such as stochastic behavior of the state, and differences in measurements sampling and updating rates;
- Dealing with scalability of state estimation algorithms for three-phase, unbalanced and large networks.

A **Bayesian Inference** perspective provides the main theoretical background that guided the developments and novel conceptualization proposed in the Thesis. It is a further step into understanding the often neglected, or misplaced, assumptions that underlay typical state estimation formulations. For example, the assumption of improper constant priors on the majority of static state estimators over reinforces the likelihood function and induces probabilities for unrealistic state values.

In essence, extending the interpretation of the state variables as random variables, inside the inference problem, enables introducing additional information to support the estimation process. This issue is often tackled under Kalman Filters but without an explicit formal generalization on the estimation formulation and only perceiving sequential instants. In this sense, this Thesis opens a fraction of such horizon under full probabilistic conceptualization of the state estimation problem. The main conclusion is that distribution system state estimation can

benefit from extensions towards including prior knowledge about the network in the formulation, both from computational performance as for accuracy under adverse scenarios. It is essential since such systems tend to present poor quality and low amount of information to aid real-time monitoring applications, increasing importance under smart grid scenarios.

Such fundamental departure for the state induces more generalist probabilistic characteristics, with non-Gaussian models, to capture such stochastic behavior of the power systems, namely in this Thesis state evolution, and also from the measurement noise. Kernel density estimation principles generalize the probabilistic models. Such extension, however, does not hold the Maximum Likelihood principle as the best estimator, as demonstrated. In this sense, **Information Theory** accommodates such a generalization in an adequate manner with a **Maximum Correntropy** principle.

Besides, a first-order state-space model was introduced to include prior characterization of state variables, based on the Kalman Filter approach, a first step into introducing temporal and probabilistic relations. The extension for non-Gaussian models then yields a Maximum Correntropy based Extended Kalman Filter, that was evaluated to treat SCADA and PMU measurements. The primary benefit of such formulation is the asymptotic response of the filter, increasing accuracy over time, a feature neglected in traditional static estimation. This, however, comes with the drawback of degrading the performance in the presence of system transitions and bad data (anomalies in the system and measurement set).

A numerically stable kernel annealing strategy suppresses such pervasive effects, maintaining the accuracy of the filter. By enlarging the kernel bandwidth of suspect samples instead of the current practice of reducing it towards zero, results in the same indifference of such samples while maintaining numerical stability of the estimation process. In a probabilistic sense, such an annealing strategy means a smoother empirical distribution function without many spikes.

Regarding the specific objective of distribution networks, this Thesis deals with the high-fidelity representation of distribution systems, employing a formulation for the **three-phase unbalanced models** based on general two-port admittance matrices. Such contribution may seem marginal, but it is an important achievement for software development since it is a well-known complexity to properly represent equipment diversity, different connections, and phase asymmetry. It dismisses often assumed simplifications, such as neglecting some components (such as transformers and voltage regulator) or, in worst cases, representing distribution networks only with a positive sequence circuit (balanced and transposed).

In this sense, the diversity of information on distribution systems is also addressed. An exploration of the measurement vector shows that information sampled in different instants is essentially mixed, representing an additional systematic error in the estimation. In this sense, a **Bayesian information fusion** approach tackles such diversity, in a computationally efficient manner. By employing a concept of sampling layers, different sources of information typically available at distribution systems (pseudo measurements, smart meters, SCADA and PMUs) are

processed separately, while transversal relations are captured with a prior distribution.

Dealing separately with each sampling layer improves accuracy compared with a hybrid estimation that mixes all information at once. Besides, by employing a fusion model, only temporal relations across different layers are captured, instead of sequential ones as in the Kalman Filter, without deprecating accuracy under system transitions. By creating an initial and observable layer based on pseudo-measurement, the subsequent layers can deal with a reduced number of real-time measurements, increasing performance. The method's performance is further boosted with a sparse orthogonal formulation, that enables dealing with three-phase unbalanced networks and fast measurement processing (such as those from PMUs), keeping real-time requirements in practically sized networks.

Finally, scalability issues were addressed throughout the Thesis, and more specifically with a novel **Multiarea State Estimation** procedure for large scale distribution networks. A hierarchical architecture was employed with parallel computing techniques to achieve higher performance and reach real time processing requirements. The Thesis also brings Bayesian Inference concepts to support the proposed method, by decomposing the power grids into areas that may not be observable. Observability is then supported by including prior information, an exciting feature to model primary feeders. Such priors are updated along with the coordination procedure, which ensures accuracy. The results illustrate a compromise solution to achieve increased performance, a trade-off of accuracy compared to centralized approaches, that enables convergence in huge systems.

A crucial feature of implementation and scalability, also explored throughout the Thesis, is applying orthogonal methods to distribution system state estimation. Such linear algebra technique is imperative to enable convergence, enhances numerical stability and increases computational performance. It is highly recommended that any practical implementation of distribution system state estimation begins with such numerical methods. Otherwise, poor convergence characteristics are likely, or even simplifications are required (both in component representation or model approximations). Other complementary features that go along the orthogonal formulation are sparsity treatments and ordering techniques to improve computational performance (both in processing time and memory requirements) and reduce the number of fill-ins (increasing even further numerical stability).

Regarding future works, the power system state estimation problem is an exciting topic since it has been increasingly revisited in the context of data-driven learning for power systems. The several peculiarities of the particular application result in dealing with high-dimensional learning, and that requires support from expert knowledge about the system. The long-term perspective is to reach a fully autonomous power grid, with closed-loop large-scale estimators associated with grid controllers, a clear research gap still open to be explored. In this sense, a closer conceptual perspective between modern signal processing theory and data analytics can aid future enhancements. Besides, the integration with technological advances must support this

endeavour, especially with increased automation in power grids, such as advanced metering infrastructure, PMUs and fast power electronics-based control.

In this context, two main research directions are devised from this Thesis, that can accommodate crucial advances for power system state estimation: Bayesian Statistical Learning and Advanced Applications in Energy Management Systems.

8.1 Bayesian Statistical Learning

The first set of future works follows the probabilistic interpretation of State Estimation. New paradigms and theoretical perspectives may emerge from Statistical Learning and Probabilistic Signal Processing. The probabilistic interpretation can be embedded as new prior and hyper-priors model, bringing additional expert knowledge directly in the state estimation formulations, while maintaining a solid and formal theoretical consistency. In this sense, a few possibilities are enumerated as possible continuations of this Thesis:

- Exploring Variational Bayesian inference to increase stochastic generalization of the process and measurement noise. This framework opens new formulations through the kullback-leibler divergence, where best probabilistic approximations for the models are searched instead of maximizing likelihood or similarity. In this sense, variational inference can be explored under two different concepts departing from this Thesis: first along numerical optimization methods in order to tune kernel bandwidth in Correntropy state estimation; and second providing a new state estimation formulation based on the evidence lower bound and parametric approximations for the measurement noise and prior distribution characteristics.

- Incrementing the probabilistic models with hyper priors for different parameters and complementary layers of information. Different priors may be assigned under an hierarchical framework to capture additional parametric uncertainty. For instance, employing priors for the measurement covariance, assumed as known in conventional state estimation, yielding new methods for detecting bad data and miss-calibration of measurements. Another exciting direction is incorporating additional variables to capture causal relations, for instance, stochastic load models, weather conditions, social behavioural, urban sensors among others. Such integration with exogenous variables can be performed in two ways, that is, supporting the state estimation procedure with additional information, or using real-time state characterization for different probabilistic analysis, like forecasting or inferring features of the power grid.

- Associating pattern recognition techniques to the state estimation problem, in order to automatically identify behaviour of the power grid. This type of evaluation is specially interesting in the context of dynamic state estimation, that tracks fundamental changes on the system along time. It could be achieved by extending the proposed state evolution

concept with high-dimensional clusterization algorithms. For instance, such type of pattern recognition technique may track topology changes according to observations, and employ structural learning, based on Bayesian Information Criterion, to evaluate the most probable topology.

- Employing data stream techniques to detect Markovian jumps and concept drifts on power system time series acquired in real time (either from SCADA or PMUs). For instance, the state space formulations may be complemented with a Dirichlet process in the Bayesian sense that keeps track of state transitions. Another direction is towards tuning anomaly detection and automatic adaptation algorithms, such as the ones based on innovation analysis, by upsampling the posteriors through Markov Chain methods.

- Finally, the statistical learning perspectives, along with recent advances on artificial intelligence, present the exciting opportunity to layer the power systems physics-based models over data-driven techniques. This enables extracting the most of both paradigms, the specificity and high dimension categorization of physics-based models, and the adaptability and robustness of data-driven techniques. Such perspective increases the goal from only estimating the state but also to recognizing the behaviour of the power grid, and besides may increment interpretability of machine learning models.

8.2 Advanced Applications in Energy Management Systems

The volume of data is flooding not only in operation centres, but also substations and new low voltage local controllers and aggregators. The technical and engineering challenges to deal with such data are immense and complex. Novel and exciting possibilities emerge in this context, such as the integration of multiple data sources, fast processing solutions, extracting complementary knowledge about the system, widening application spectrum and fully completing energy management solutions, from data to decision making. In this sense, we devise a few interesting applications that are quite reachable in future research efforts that follow the developed solution in this Thesis:

- Exploration of new numerical methods to increase even further computational efficiency. In this sense, different non-linear programming algorithms may be employed, such as the Lagrangian methods to accommodate constraints, new formulations based on the class of interior point methods, and with different sparse factorization techniques and representations. Besides, an interesting challenge is extending and tuning the state estimation algorithms for GPU computing, leveraging data parallelism paradigms and enabling faster integration with machine learning algorithms in real time as well.

- Extending the application of three-phase state estimation models for transmission systems, with the intent of increasing model resolution advanced applications. The higher degree of

model resolution may improve fault location algorithms under slight unbalanced conditions. Besides the increased number of PMUs on such networks also enables moving towards dynamic state estimation, which already extends detailed generator models but may also benefit from detailed three-phase phasor measurements, thus requiring three-phase models.

- Enabling autonomous power grid operation, supported by closed-loop architectures based on state estimation and state feedback control. In this sense, model predictive control presents itself as an interesting formulation that bridges uncertainty and input control variables in a straightforward extension of traditional state estimation formulations. The motivation is the growing interest on microgrids, that emerge the possibility of advanced control architectures on distribution networks to increase the accommodation of renewable energy, storage, electric vehicles and demand response. Besides it can also help reducing the cognitive burden of operators at control centres, aiding partial operator-in-the-loop decision making.

- Exploiting the probabilistic characterization of the state for forecasting and real time risk analysis of power systems. The probabilistic background of Bayesian Inference, along with information fusion concepts, may provide a framework to introduce additional information, such as weather and social features, while supporting different applications that deal with uncertainty and risk assessment. In this sense, new awareness indexes considering uncertainty may be created for predictive stability assessment, maximum power transfer, and resilience evaluation of the network.

- Finally, integrating the state estimation algorithms with international standards and automation protocols provides interesting engineering problems. For instance, automation and modelling standards such as IEC Common Information Model, IEC 61850, IEC 62056, IEEE C37.11 and DNP 3.0 are a few examples. This may apparently resemble only a technical problem, however this practical scenario may inspire and induce scientific challenges and novel analytics, fully conceptualizing cyber-physical architectures for power grid assessment. Regarding IEC 61850, the availability of sampled values technology and GOOSE messages, if supported by fast processing techniques, can increase response time of dynamic state estimators. Besides, advanced database pipeline and computing architectures may also be addressed in such perspective, such as cloud computing, blockchain applications and further exploring parallel multiprocessor algorithms.

BIBLIOGRAPHY

Abur, A.; Gómez-Expósito, A. **Power System State Estimation: Theory and Implementation**. Nova York: CRC Press, 2004. 327 p.

Ahmad, F. et al. Distribution system state estimation-a step towards smart grid. **Renewable and Sustainable Energy Reviews**, v. 81, p. 2659–2671, 2018. ISSN 1364-0321.

Akrami, A.; Asif, M. S.; Mohsenian-Rad, H. Sparse distribution system state estimation: An approximate solution against low observability. In: **2020 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)**. [S.l.: s.n.], 2020. p. 1–5.

Al-wakeel, A.; Wu, J.; Jenkins, N. State estimation of medium voltage distribution networks using smart meter measurements. **Applied Energy**, The Authors, v. 184, p. 207–218, 2016. ISSN 0306-2619.

Albert, J. **Bayesian computation with R**. [S.l.]: Springer Science & Business Media, 2009.

Alimardani, A. et al. Distribution system state estimation based on nonsynchronized smart meters. **IEEE Transactions on Smart Grid**, v. 6, n. 6, p. 2919–2928, 2015.

Almeida, M. C. de; Ochoa, L. F. An improved three-phase amb distribution system state estimator. **IEEE Transactions on Power Systems**, v. 32, n. 2, p. 1463–1473, 2017.

Amerongen, R. A. On the exact incorporation of virtual measurements on orthogonal-transformation based state-estimation procedures. **International Journal of Electrical Power & Energy Systems**, v. 13, n. 3, p. 167–174, 1991. ISSN 0142-0615.

Amestoy, P. R.; Davis, T. A.; DUFF, I. S. Algorithm 837: Amd, an approximate minimum degree ordering algorithm. **ACM Trans. Math. Softw.**, Association for Computing Machinery, New York, NY, USA, v. 30, n. 3, p. 381–388, set. 2004. ISSN 0098-3500.

Angioni, A. et al. Impact of pseudo-measurements from new power profiles on state estimation in low-voltage grids. **IEEE Transactions on Instrumentation and Measurement**, v. 65, n. 1, p. 70–77, 2016.

____. Real-time monitoring of distribution system based on state estimation. **IEEE Transactions on Instrumentation and Measurement**, v. 65, n. 10, p. 2234–2243, 2016.

Arrillaga, J.; Harker, B. J. Fast-decoupled three-phase load flow. **Proceedings of the Institution of Electrial Enginners**, v. 125, n. 8, p. 734–740, August 1978.

Ash, R. B. **Basic probability theory**. [S.l.]: Courier Corporation, 2008.

Baldick, R. et al. Implementing nonquadratic objective functions for state estimation and bad data rejection. **IEEE Transactions on Power Systems**, v. 12, n. 1, p. 376–382, 1997.

Baran, M. Branch current based state estimation for distribution system monitoring. p. 1–4, 2012.

Baran, M. E.; Jung, J.; McDermott, T. E. Including voltage measurements in branch current state estimation for distribution systems. In: **2009 IEEE Power Energy Society General Meeting**. [S.l.: s.n.], 2009. p. 1–5.

Baran, M. E.; Kelley, A. W. State estimation for real-time monitoring of distribution systems. **IEEE Transactions on Power Systems**, v. 9, n. 3, p. 1601–1609, 1994.

____. A branch-current-based state estimation method for distribution systems. **IEEE Transactions on Power Systems**, v. 10, n. 1, p. 483–491, 1995.

Barbeiro, P. N. P. et al. Exploiting autoencoders for three-phase state estimation in unbalanced distributions grids. **Electric Power Systems Research**, Elsevier B.V., v. 123, p. 108–118, 2015. ISSN 0378-7796.

Barker, A.; Brown, D.; Martin, W. Bayesian estimation and the Kalman filter. **Computers & Mathematics with Applications**, Pergamon, v. 30, n. 10, p. 55–77, nov 1995. ISSN 0898-1221.

Bayarri, M. J.; Berger, J. O. The Interplay of Bayesian and Frequentist Analysis. **Statistical Science**, Institute of Mathematical Statistics, v. 19, n. 1, p. 58 – 80, 2004.

Benedito, R. A. S. et al. Power system state estimation: Undetectable bad data. **International Transactions on Electrical Energy Systems**, v. 24, n. 1, p. 91–107, 2014.

Blitzstein, J. K.; Hwang, J. **Introduction to probability**. [S.l.]: Crc Press, 2019.

Bretas, A. S. et al. Multiple gross errors detection , identification and correction in three-phase distribution systems WLS state estimation : A per-phase measurement error approach. **Electric Power Systems Research**, Elsevier B.V., v. 151, p. 174–185, 2017. ISSN 0378-7796.

Bretas, N. An iterative dynamic state estimation and bad data processing. **International Journal of Electrical Power & Energy Systems**, v. 11, n. 1, p. 70–74, 1989. ISSN 0142-0615.

Bretas N. Bretas, J. B. A. L. B. C. P. . E. Y. . . A. . F. A. **Cyber-Physical Power Systems State Estimation.** [S.l.: s.n.].

Bretas, N. G. Network observability: theory and algorithms based on triangular factorisation and path graph concepts. **IEE Proceedings - Generation, Transmission and Distribution**, v. 143, n. 1, p. 123–128, 1996.

CAMILLO, M. H. et al. Combining exhaustive search and multi-objective evolutionary algorithm for service restoration in large-scale distribution systems. **Electric Power Systems Research**, v. 134, p. 1–8, 2016. ISSN 0378-7796. Disponível em: <https://www.sciencedirect.com/science/article/pii/S037877961500379X>.

Cao, Z. et al. Scalable distribution systems state estimation using long short-term memory networks as surrogates. **IEEE Access**, v. 8, p. 23359–23368, 2020.

Carquex, C.; Rosenberg, C.; Bhattacharya, K. State estimation in power distribution systems based on ensemble kalman filtering. **IEEE Transactions on Power Systems**, v. 33, n. 6, p. 6600–6610, Nov 2018. ISSN 1558-0679.

Carson, J. R. Wave propagation in overhead wires with ground return. **The Bell System Technical Journal**, v. 5, n. 4, p. 539–554, 1926.

Castillo, M. R. M. et al. Offline detection, identification, and correction of branch parameter errors based on several measurement snapshots. **IEEE Transactions on Power Systems**, v. 26, n. 2, p. 870–877, May 2011. ISSN 0885-8950.

Celik, M. K.; Abur, A. A robust wlav state estimator using transformations. **IEEE Transactions on Power Systems**, v. 7, n. 1, p. 106–113, 1992.

Chan, W.-M.; George, A. A linear time implementation of the reverse cuthill-mckee algorithm. **BIT Numerical Mathematics**, Springer, v. 20, n. 1, p. 8–14, 1980.

Chen, B. et al. Maximum correntropy kalman filter. **Automatica**, Elsevier, v. 76, p. 70–77, 2017.

Chen, B.; Principe, J. C. Maximum correntropy estimation is a smoothed map estimation. **IEEE Signal Processing Letters**, v. 19, n. 8, p. 491–494, 2012.

Chen, Q. et al. Impact of smart metering data aggregation on distribution system state estimation. **IEEE Transactions on Industrial Informatics**, v. 12, n. 4, p. 1426–1437, 2016.

Chen, T.-H.; Chang, J.-D. Open wye-open delta and open delta-open delta transformer models for rigorous distribution system analysis. In: IET. **IEE Proceedings C (Generation, Transmission and Distribution)**. [S.l.], 1992. v. 139, n. 3, p. 227–234.

Chen, T.-H. et al. Three-phase cogenerator and transformer models for distribution system analysis. **IEEE Transactions on Power Delivery**, IEEE, v. 6, n. 4, p. 1671–1681, 1991.

Cheng, C. S.; Shirmohammadi, D. A three-phase power flow method for real-time distribution system analysis. **IEEE Transactions on Power Systems**, v. 10, n. 2, p. 671–679, 1995.

Conejo, A. J. et al. **Decomposition techniques in mathematical programming: engineering and science applications**. [S.l.]: Springer Science & Business Media, 2006.

Congdon, P. **Bayesian Statistical Modelling**. [S.l.]: John Wiley & Sons, 2007. v. 704.

Costa, A. S.; Lourenco, E. M.; Clements, K. A. Power system topological observability analysis including switching branches. **IEEE Trans. Power Syst.**, v. 17, n. 2, p. 250–256, 2002.

Costa, A. S.; Salgado, R.; Haas, P. Globally convergent state estimation based on givens rotations. In: **2007 iREP Symposium - Bulk Power System Dynamics and Control - VII. Revitalizing Operational Reliability**. [S.l.: s.n.], 2007. p. 1–9.

Coutto Filho, M. B. do; Souza, J. C. S. de. Forecasting-Aided State Estimation Part I: Panorama. **IEEE Transactions on Power Systems**, v. 24, n. 4, p. 1667–1677, Nov 2009. ISSN 0885-8950.

Coutto Filho, M. B. do; Souza, J. C. S. de; Guimaraens, M. A. R. Enhanced bad data processing by phasor-aided state estimation. **IEEE Transactions on Power Systems**, v. 29, n. 5, p. 2200–2209, 2014.

Cutsem, T. V.; Ribbens-Pavella, M. Critical survey of hierarchical methods for state estimation of electric power systems. **IEEE Transactions on Power Apparatus and Systems**, PAS-102, n. 10, p. 3415–3424, 1983.

Davis, T. A. Algorithm 915, suitesparseqr: Multifrontal multithreaded rank-revealing sparse qr factorization. **ACM Trans. Math. Softw.**, Association for Computing Machinery, New York, NY, USA, v. 38, n. 1, dez. 2011. ISSN 0098-3500.

____. Multifrontal multithreaded rank-revealing sparse qr factorization. **ACM Trans. Math. Softw.**, Association for Computing Machinery, New York, NY, USA, v. 38, n. 1, dez. 2011. ISSN 0098-3500.

Debs, A. S.; Larson, R. E. A dynamic estimator for tracking the state of a power system. **IEEE Transactions on Power Apparatus and Systems**, PAS-89, n. 7, p. 1670–1678, 1970.

Dehghanpour, K. et al. A survey on state estimation techniques and challenges in smart distribution systems. **IEEE Transactions on Smart Grid**, v. 10, n. 2, p. 2312–2322, 2019.

Della Giustina, D. et al. Electrical distribution system state estimation: measurement issues and challenges. **IEEE Instrumentation and Measurement Magazine**, v. 17, n. 6, p. 36–42, 2014.

Dobbe, R. et al. Linear single- and three-phase voltage forecasting and bayesian state estimation with limited sensing. **IEEE Transactions on Power Systems**, v. 35, n. 3, p. 1674–1683, 2020.

Duque, F. G. et al. State estimator for electrical distribution systems based on an optimization model. **Electric Power Systems Research**, Elsevier B.V., v. 152, p. 122–129, 2017. ISSN 0378-7796.

Džafić, I. et al. Real time estimation of loads in radial and unsymmetrical three-phase distribution networks. **IEEE Transactions on Power Systems**, v. 28, n. 4, p. 4839–4848, 2013.

Ebrahimian, R.; Baldick, R. State estimator condition number analysis. **IEEE Transactions on Power Systems**, v. 16, n. 2, p. 273–279, 2001.

Erdogmus, D.; Principe, J. An error-entropy minimization algorithm for supervised training of nonlinear adaptive systems. **IEEE Transactions on Signal Processing**, v. 50, n. 7, p. 1780–1786, 2002.

EU Commission. **Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: The European Green Deal.** 2019. Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2019:640:FIN>.

Falcao, D.; Wu, F.; Murphy, L. Parallel and distributed state estimation. **IEEE Transactions on Power Systems**, v. 10, n. 2, p. 724–730, 1995.

Falcao, D. M.; Arias, M. A. State estimation and observability analysis based on echelon forms of the linearized measurement models. **IEEE Transactions on Power Systems**, v. 9, n. 2, p. 979–987, May 1994. ISSN 0885-8950.

Falcao, D. M.; Cooke, P. A.; Brameller, A. Power system tracking state estimation and bad data processing. **IEEE Transactions on Power Apparatus and Systems**, PAS-101, n. 2, p. 325–333, 1982.

Fan, L.; Wehbe, Y. Extended kalman filtering based real-time dynamic state and parameter estimation using pmu data. **Electric Power Systems Research**, v. 103, p. 168–177, 2013. ISSN 0378-7796.

Fantin, C. A. **Metodologia para estimação de estado trifásica em sistemas de distribuição incorporando medidas SCADA, virtuais, pseudo-medidas e medidas fasoriais sincronizadas (In portuguese)**. 2016. 130 f. Tese (Doutorado em Sistemas Elétricos de Potência) — Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2016.

FANUCCHI, R. Z. et al. A multi-objective algorithm to determine patrol sequences for out-of-service nodes in power distribution feeders. **Electric Power Systems Research**, v. 196, p. 107198, 2021. ISSN 0378-7796. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0378779621001796>.

Feng, X.; Yang, F.; Peterson, W. A practical multi-phase distribution state estimation solution incorporating smart meter and sensor data. In: **2012 IEEE Power and Energy Society General Meeting**. [S.l.: s.n.], 2012. p. 1–6.

Fernandes, T. R.; Venkatesh, B.; Almeida de M. C. Symmetrical components based state estimator for power distribution systems. **IEEE Transactions on Power Systems**, v. 36, n. 3, p. 2035–2045, 2021.

Freitas, V.; Costa, A. S.; Miranda, V. Robust state estimation based on orthogonal methods and maximum correntropy criterion. In: **2017 IEEE Manchester PowerTech**. [S.l.: s.n.], 2017. p. 1–6.

Freitas, V.; Simões Costa, A.; Miranda, V. Orthogonal method for solving maximum correntropy-based power system state estimation. **IET Generation, Transmission & Distribution**, v. 14, n. 10, p. 1930–1941, 2020.

Garcia, L. A.; Grenard, S. Scalable distribution state estimation approach for Distribution Management Systems. In: **2nd IEEE PES International Conference and Exhibition onInnovative Smart Grid Technologies (ISGT Europe)**. Manchester, Inglaterra: [s.n.], 2011. p. 1–6.

George, A.; Heath, M. T. Solution of sparse linear least squares problems using givens rotations. **Linear Algebra and its Applications**, v. 34, p. 69–83, 1980. ISSN 0024-3795.

GEORGE, A.; LIU, J.; NG, E. A data structure for sparse qr and lu factorizations. **SIAM Journal on Scientific and Statistical Computing**, v. 9, n. 1, p. 100–121, 1988.

Ghosh, A. K.; Lubkeman, D. L.; Jones, R. H. Load modeling for distribution circuit state estimation. **IEEE Transactions on Power Delivery**, v. 12, n. 2, p. 999–1005, 1997. ISSN 08858977.

Givens, W. Computation of plane unitary rotations transforming a general matrix to triangular form. **Journal of the Society for Industrial and Applied Mathematics**, Society for Industrial and Applied Mathematics, v. 6, n. 1, p. 26–50, 1958. ISSN 03684245.

Glover, J. D.; Sarma, M. S.; Overbye, T. **Power system analysis & design, SI version**. [S.l.]: Cengage Learning, 2012.

Gol, M.; Abur, A. A Hybrid State Estimator For Systems With Limited Number of PMUs. **IEEE Transactions on Power Systems**, v. 30, n. 3, p. 1511–1517, May 2015. ISSN 0885-8950.

Golub, G. H.; Van Loan, C. F. **Matrix computations**. [S.l.]: JHU press, 2012. v. 3.

Gómez-Exposito, A. et al. A multilevel state estimation paradigm for smart grids. **Proceedings of the IEEE**, v. 99, n. 6, p. 952–976, 2011. ISSN 00189219.

____. A taxonomy of multi-area state estimation methods. **Electric Power Systems Research**, v. 81, p. 1060–1069, 2011. ISSN 03787796.

Gonzaga, R. M. et al. An embedded state estimator for reducing data volume and processing in smart grids monitoring. In: **2018 Workshop on Communication Networks and Power Systems (WCNPS)**. [S.l.: s.n.], 2018. p. 1–5.

Gonzalez, X. et al. Methodology for multiarea state estimation solved by a decomposition method. **Electric Power Systems Research**, v. 123, p. 92 – 99, 2015. ISSN 0378-7796.

Gómez-Expósito, A.; Gómez-Quiles, C.; Džafić, I. State estimation in two time scales for smart distribution systems. **IEEE Transactions on Smart Grid**, v. 6, n. 1, p. 421–430, 2015.

____. State estimation in two time scales for smart distribution systems. **IEEE Transactions on Smart Grid**, v. 6, n. 1, p. 421–430, 2015.

Göl, M.; Abur, A. Lav based robust state estimation for systems measured by pmus. **IEEE Transactions on Smart Grid**, v. 5, n. 4, p. 1808–1814, 2014.

Hansen, C.; Debs, A. Power system state estimation using three-phase models. **IEEE Transactions on Power Systems**, v. 10, n. 2, p. 818–824, 1995.

Haque, A. N. M. M. et al. Integrating direct and indirect load control for congestion management in lv networks. **IEEE Transactions on Smart Grid**, v. 10, n. 1, p. 741–751, 2019.

Hayes, B. P.; Gruber, J. K.; Prodanovic, M. A closed-loop state estimation tool for mv network monitoring and operation. **IEEE Transactions on Smart Grid**, v. 6, n. 4, p. 2116–2125, 2015.

Hebling, G. M. et al. Sparse and Numerically Stable Implementation of a Distribution System State Estimation based on Multifrontal QR Factorization. **Electric Power System Research**, v. 189, p. 106734, Dez. 2020.

Higham, N. J. **Accuracy and stability of numerical algorithms**. [S.l.]: Siam, 2002. v. 80.

Householder, A. S. Unitary triangularization of a nonsymmetric matrix. **Journal of the ACM (JACM)**, ACM New York, NY, USA, v. 5, n. 4, p. 339–342, 1958.

Huang, S.; Lu, C.; Lo, Y. Evaluation of ami and scada data synergy for distribution feeder modeling. **IEEE Transactions on Smart Grid**, v. 6, n. 4, p. 1639–1647, 2015.

IEA. **Digitalization and Energy Report**. 2019. Disponível em: <https://www.iea.org/reports/digitalisation-and-energy>.

IEEE. **IEEE PES Distribution Systems Analysis Subcommittee Radial Test Feeders**. 2018. Disponível em: <http://sites.ieee.org/pes-testfeeders/>.

Jardini, J. A. et al. Daily load profiles for residential, commercial and industrial low voltage consumers. **IEEE Transactions on Power Delivery**, v. 15, n. 1, p. 375–380, Jan 2000. ISSN 0885-8977.

Jesus, P. D. O.-D.; Antunes, C. H. A detailed network model for distribution systems with high penetration of renewable generation sources. **Electric Power Systems Research**, v. 161, p. 152 – 166, 2018. ISSN 0378-7796.

Kadurek, P. et al. Aiding power system support by means of voltage control with intelligent distribution substation. **IEEE Transactions on Smart Grid**, v. 5, n. 1, p. 84–91, 2014.

Kagan, N. **Introdução aos sistemas de distribuição de energia elétrica**. [S.l.]: Editora Edgard Blücher, 2008. ISBN 9788521203551.

Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems. **Journal of Basic Engineering**, v. 82, n. 1, p. 35–45, 03 1960. ISSN 0021-9223.

Kersting, W. A three-phase unbalanced line model with grounded neutrals through a resistance. In: IEEE. **2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century**. [S.l.], 2008. p. 1–2.

Kersting, W. H. **Distribution system modeling and analysis**. Boca Raton, Florida: CRC Press, 2001.

____. Center-tapped transformer and 120-/240-v secondary models. **IEEE transactions on industry applications**, IEEE, v. 45, n. 2, p. 575–581, 2009.

Kettner, A. M.; Paolone, M. Sequential discrete kalman filter for real-time state estimation in power distribution systems: Theory and implementation. **IEEE Transactions on Instrumentation and Measurement**, v. 66, n. 9, p. 2358–2370, 2017.

Kezunovic, M. et al. Big data analytics for future electricity grids. **Electric Power Systems Research**, v. 189, p. 106788, 2020. ISSN 0378-7796.

Korres, G. N. A distributed multiarea state estimation. **IEEE Transactions on Power Systems**, v. 26, n. 1, p. 73–84, Feb 2011. ISSN 0885-8950.

Korres, G. N.; Manousakis, N. M. State estimation and observability analysis for phasor measurement unit measured systems. **IET Generation, Transmission Distribution**, v. 6, n. 9, p. 902–913, 2012.

Kron, G. Tensorial analysis of integrated transmission systems part i. the six basic reference frames. **Transactions of the American Institute of Electrical Engineers**, v. 70, n. 2, p. 1239–1248, 1951.

Krsman, V. D.; Sarić, A. T. Bad area detection and whitening transformation-based identification in three- phase distribution state estimation. **IET Generation, Transmission & Distribution**, v. 11, n. 9, p. 2351–2361, 2017.

Krumpholz, G. R.; Clements, K. A.; Davis, P. W. Power system observability: A practical algorithm using network topology. **IEEE Transactions on Power Apparatus and Systems**, PAS-99, n. 4, p. 1534–1542, 1980.

Kundur, P. Power system stability and control. In: ____. [S.l.]: McGraw Hill, 1994. p. 272–274.

Lefebvre, S.; Prévost, J.; Lenoir, L. Distribution state estimation: A necessary requirement for the smart grid. In: **2014 IEEE PES General Meeting | Conference Exposition**. [S.l.: s.n.], 2014. p. 1–5.

Leite, J. B.; Mantovani, J. R. S. Distribution system state estimation using the hamiltonian cycle theory. **IEEE Transactions on Smart Grid**, v. 7, n. 1, p. 366–375, 2016.

Li, Y.; Arce, G. R. A maximum likelihood approach to least absolute deviation regression. **EURASIP Journal on Advances in Signal Processing**, Springer, v. 2004, n. 12, p. 1–8, 2004.

Liao, H.; Milanovic, J. V. Pathway to cost-efficient state estimation of future distribution networks. In: **2016 IEEE Power and Energy Society General Meeting (PESGM)**. [S.l.: s.n.], 2016. p. 1–5.

LIU, J. et al. Trade-offs in pmu deployment for state estimation in active distribution grids. **IEEE Transactions on Smart Grid**, v. 3, n. 2, p. 915–924, 2012.

Liu, W.; Pokharel, P. P.; Principe, J. C. Correntropy: A localized similarity measure. In: IEEE. **The 2006 IEEE international joint conference on neural network proceedings**. [S.l.], 2006. p. 4919–4924.

Liu, W.-H.; Sherman, A. H. Comparative analysis of the cuthill–mckee and the reverse cuthill–mckee ordering algorithms for sparse matrices. **SIAM Journal on Numerical Analysis**, v. 13, n. 2, p. 198–213, 1976.

London Jr, J. B. A.; Alberto, L. F. C.; Bretas, N. G. Analysis of measurement-set qualitative characteristics for state-estimation purposes. **IET Generation, Transmission Distribution**, v. 1, n. 1, p. 39–45, January 2007. ISSN 1751-8687.

London Jr, J. B. A. et al. Redundancy and observability analysis of conventional and pmu measurements. **IEEE Transactions on Power Systems**, v. 24, n. 3, p. 1629–1630, Aug 2009. ISSN 0885-8950.

Lopes, J. A. P. et al. The future of power systems: Challenges, trends, and upcoming paradigms. **WIREs Energy and Environment**, v. 9, n. 3, p. e368, 2020.

Louis, A. et al. Measurement sensitivity and estimation error in distribution system state estimation using augmented complex kalman filter. **Journal of Modern Power Systems and Clean Energy**, v. 8, n. 4, p. 657–668, 2020.

Lourenco, E. M.; Costa, A. S.; Clements, K. A. Bayesian-based hypothesis testing for topology error identification in generalized state estimation. **IEEE Transactions on Power Systems**, v. 19, n. 2, p. 1206–1215, 2004.

Lu, C. N.; Teng, J. H.; Liu, W. . E. Distribution system state estimation. **IEEE Transactions on Power Systems**, v. 10, n. 1, p. 229–240, 1995.

Machado, P. A.; Azevedo, G. P. D.; Monticelli, A. J. A mixed pivoting approach to the factorization of indefinite matrices in power system state estimation. **IEEE Transactions on Power Systems**, v. 6, n. 2, p. 676–682, 1991.

Makowski, D.; S. Ben-Shachar, M.; Lüdecke, D. Describing effects and their uncertainty, existence and significance within the bayesian framework. **Journal of Open Source Software**, The Open Journal, v. 4, n. 40, p. 1541, 2019.

Manitsas, E. et al. Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling. **IEEE Transactions on Power Systems**, v. 27, n. 4, p. 1888–1896, 2012.

Massignan, J. A. D. et al. Real-time load estimation for distribution feeders. In: **2015 IEEE Eindhoven PowerTech**. [S.l.: s.n.], 2015. p. 1–6.

____. In-field validation of a real-time monitoring tool for distribution feeders. **IEEE Transactions on Power Delivery**, v. 33, n. 4, p. 1798–1808, 2018.

____. PMUs and SCADA Measurements in Power System State Estimation through Bayesian Inference. In: **2019 IEEE Milan PowerTech**. [S.l.: s.n.], 2019. p. 1–6. ISSN null.

Massignan, J. A. D.; London Jr, J.; Miranda, V. Tracking Power System State Evolution with Maximum-correntropy-based Extended Kalman Filter. **Journal of Modern Power Systems and Clean Energy**, v. 8, n. 4, p. 616–626, 2020.

Massignan, J. A. D. et al. Vulnerability of largest normalized residual test and b-test to gross errors. In: **2020 IEEE Power Energy Society General Meeting (PESGM)**. [S.l.: s.n.], 2020. p. 1–5.

Massignan, J. A. D.; Pereira, B. R.; London, J. B. A. Load flow calculation with voltage regulators bidirectional mode and distributed generation. **IEEE Transactions on Power Systems**, v. 32, n. 2, p. 1576–1577, 2017.

Melo, F. et al. LV SCADA Project: In-field Validation of a Distribution State Estimation Tool for LV Networks. In: **CIRED Workshop 2016**. Hellsinki: [s.n.], 2016. p. 1–4.

Meloni, A. et al. Cloud-based iot solution for state estimation in smart grids: Exploiting virtualization and edge-intelligence technologies. **Computer Networks**, v. 130, p. 156–165, 2018. ISSN 1389-1286.

Mestav, K. R.; Luengo-Rozas, J.; Tong, L. Bayesian state estimation for unobservable distribution systems via deep learning. **IEEE Transactions on Power Systems**, v. 34, n. 6, p. 4910–4920, 2019.

Milanovic, J. V. et al. International industry practice on power system load modeling. **IEEE Transactions on Power Systems**, v. 28, n. 3, p. 3038–3046, 2013.

Mili, L. et al. Robust state estimation based on projection statistics [of power systems]. **IEEE Transactions on Power Systems**, v. 11, n. 2, p. 1118–1127, 1996.

Mili, L.; Phaniraj, V.; Rousseeuw, P. J. Least median of squares estimation in power systems. **IEEE Transactions on Power Systems**, v. 6, n. 2, p. 511–523, May 1991. ISSN 0885-8950.

Miranda, V. et al. Through the looking glass: Seeing events in power systems dynamics. **International Journal of Electrical Power & Energy Systems**, v. 106, p. 411–419, 2019. ISSN 0142-0615.

Miranda, V.; Santos, A.; Pereira, J. State Estimation Based on Correntropy: A Proof of Concept. **IEEE Transactions on Power Systems**, v. 24, n. 4, p. 1888–1889, nov 2009. ISSN 0885-8950.

Mohiuddin, S. M.; Qi, J. Maximum correntropy extended kalman filtering for power system dynamic state estimation. In: **2019 IEEE Power Energy Society General Meeting (PESGM)**. [S.l.: s.n.], 2019. p. 1–5.

Monticelli, A. **State estimation in electric power systems: a generalized approach.** Massachussets, USA: Kluwer Academic Publishers, 1999.

Monticelli, A.; Garcia, A. Reliable bad data processing for real-time state estimation. **IEEE Transactions on Power Apparatus and Systems**, PAS-102, n. 5, p. 1126–1139, 1983.

Muscas, C. et al. Effects of measurements and pseudomeasurements correlation in distribution system state estimation. **IEEE Transactions on Instrumentation and Measurement**, v. 63, n. 12, p. 2813–2823, 2014.

____. Multiarea distribution system state estimation. **IEEE Transactions on Instrumentation and Measurement**, v. 64, n. 5, p. 1140–1148, 2015.

Nanchian, S.; Majumdar, A.; Pal, B. C. Ordinal optimization technique for three-phase distribution network state estimation including discrete variables. **IEEE Transactions on Sustainable Energy**, v. 8, n. 4, p. 1528–1535, 2017.

____. Three-phase state estimation using hybrid particle swarm optimization. **IEEE Transactions on Smart Grid**, v. 8, n. 3, p. 1035–1045, 2017.

Nguyen, P. H. et al. Dynamic state estimation for distribution networks with renewable energy integration. **Int. J. Smart Grid and Clean Energy**, Citeseer, v. 2, n. 3, p. 307–315, 2013.

Ni, F. et al. Three-phase state estimation in the medium-voltage network with aggregated smart meter data. **International Journal of Electrical Power & Energy Systems**, v. 98, p. 463–473, 2018. ISSN 0142-0615.

Nocedal, J.; Wright, S. J. **Numerical Optimization**. New York, NY, USA: Springer, 2006.

Nusrat, N.; Irving, M.; Taylor, G. Development of distributed state estimation methods to enable smart distribution management systems. p. 1691–1696, 2011.

Nusrat, N. et al. An overlapping zone-based state estimation method for distribution systems. **IEEE Transactions on Smart Grid**, v. 6, n. 4, p. 2126–2133, 2015.

OpenMP. **OpenMP API for parallel programming, version 5.1**. 2020. <http://openmp.org/wp/>.

Pau, M. et al. Design and accuracy analysis of multilevel state estimation based on smart metering infrastructure. **IEEE Transactions on Instrumentation and Measurement**, v. 68, n. 11, p. 4300–4312, 2019.

Pau, M.; Pegoraro, P. A.; Sulis, S. Efficient branch-current-based distribution system state estimation including synchronized measurements. **IEEE Transactions on Instrumentation and Measurement**, v. 62, n. 9, p. 2419–2429, 2013.

Pau, M. et al. An efficient and accurate solution for distribution system state estimation with multiarea architecture. **IEEE Transactions on Instrumentation and Measurement**, v. 66, n. 5, p. 910–919, 2017.

Pegoraro, P. A. et al. Bayesian approach for distribution system state estimation with non-gaussian uncertainty models. **IEEE Transactions on Instrumentation and Measurement**, v. 66, n. 11, p. 2957–2966, 2017.

Peng Xiao; Yu, D. C.; Wei Yan. A unified three-phase transformer model for distribution load flow calculations. **IEEE Transactions on Power Systems**, v. 21, n. 1, p. 153–159, 2006.

Pesteh, S.; Moayyed, H.; Miranda, V. Favorable properties of interior point method and generalized correntropy in power system state estimation. **Electric Power Systems Research**, v. 178, p. 106035, 2020. ISSN 0378-7796.

Pesteh, S. et al. A new interior point solver with generalized correntropy for multiple gross error suppression in state estimation. **Electric Power Systems Research**, v. 176, p. 105937, 2019. ISSN 0378-7796.

Picallo, M.; Bolognani, S.; Dörfler, F. Closing the loop: Dynamic state estimation and feedback optimization of power grids. **Electric Power Systems Research**, Elsevier, v. 189, p. 106753, 2020.

Pires, R.; Costa, A. S.; Mili, L. Iteratively reweighted least-squares state estimation through givens rotations. **IEEE Transactions on Power Systems**, v. 14, n. 4, p. 1499–1507, 1999.

Primadianto, A.; Lu, C. A review on distribution system state estimation. **IEEE Transactions on Power Systems**, v. 32, n. 5, p. 3875–3883, 2017.

Puga, J. L.; Krzywinski, M.; Altman, N. Bayesian statistics: today's predictions are tomorrow's priors. **Nature Methods**, Nature Publishing Group, v. 12, n. 5, p. 377–379, 2015.

Rencher, A. C. **Methods of multivariate analysis**. [S.l.]: John Wiley & Sons, 2003. v. 492.

Rouhani, A.; Abur, A. Linear Phasor Estimator Assisted Dynamic State Estimation. **IEEE Transactions on Smart Grid**, v. 9, n. 1, p. 211–219, jan 2018. ISSN 1949-3053.

Rousseaux, P. et al. A new formulation of state estimation in distribution systems including demand and generation states. In: **2015 IEEE Eindhoven PowerTech**. [S.l.: s.n.], 2015. p. 1–6.

Rousseaux, P.; Van Cutsem, T.; Dy Liacco, T. Whither dynamic state estimation? **International Journal of Electrical Power & Energy Systems**, v. 12, n. 2, p. 104–116, 1990. ISSN 0142-0615.

Sallam, A. A.; Malik, O. P. **Electric distribution systems**. [S.l.]: John Wiley & Sons, 2018.

Sarri, S. et al. Performance assessment of linear state estimators using synchrophasor measurements. **IEEE Transactions on Instrumentation and Measurement**, v. 65, n. 3, p. 535–548, March 2016. ISSN 0018-9456.

Schweppe, F. C.; Handschin, E. J. Static State Estimation in Electric Power Systems. **Proceedings of the IEEE**, v. 62, p. 972–982, 1974. ISSN 00189219.

Schweppe, F. C.; Wildes, J. Power System Static-State Estimation, Part I: Exact Model. **Transactions on Power Apparatus and Systems**, PAS-89, n. 1, p. 120–125, 1970.

Sharma, A.; Srivastava, S. C.; Chakrabarti, S. A cubature kalman filter based power system dynamic state estimator. **IEEE Transactions on Instrumentation and Measurement**, v. 66, n. 8, p. 2036–2045, 2017.

Silva, R. S.; Laburu, F. M.; de Almeida, M. C. On the use of micro-pmu for state estimation in distribution systems. In: **2017 IEEE Power Energy Society General Meeting**. [S.l.: s.n.], 2017. p. 1–5.

Silva, R. S. da. Specifying angular reference for three-phase distribution system state estimators. **IET Generation, Transmission & Distribution**, Institution of Engineering and Technology, v. 12, p. 1655–1663(8), April 2018. ISSN 1751-8687.

Simões Costa, A.; Albuquerque, A.; Bez, D. An estimation fusion method for including phasor measurements into power system real-time modeling. **IEEE Transactions on Power Systems**, v. 28, n. 2, p. 1910–1920, 2013.

Singh, R. et al. A recursive bayesian approach for identification of network configuration changes in distribution system state estimation. p. 1–1, 2011.

Singh, R.; Pal, B. C.; Jabr, R. A. Choice of estimator for distribution system state estimation. **IET Generation, Transmission & Distribution**, v. 3, n. 7, p. 666, 2009. ISSN 17518687.

____. Statistical Representation of Distribution System Loads Using Gaussian Mixture Model. **IEEE Transactions on Power Systems**, v. 25, n. 1, p. 29–37, 2010.

Singh, R.; Pal, B. C.; Vinter, R. B. Measurement placement in distribution system state estimation. **IEEE Transactions on Power Systems**, v. 24, n. 2, p. 668–675, 2009.

Sivia, D.; Skilling, J. **Data analysis: a Bayesian tutorial**. [S.l.]: OUP Oxford, 2006.

Teng, J. H. Using voltage measurements to improve the results of branch-current-based state estimators for distribution systems. **IEE Proceedings - Generation, Transmission and Distribution**, v. 149, n. 6, p. 667–672, Nov 2002. ISSN 1350-2360.

Thomas, J. D. M. M. S. **Power System SCADA and Smart Grids**. First. Boca Raton: CRC Press, 2015. v. 1. ISBN 9781315215372.

Valverde, G.; Terzija, V. Unscented kalman filter for power system dynamic state estimation. **IET Generation, Transmission & Distribution**, Institution of Engineering and Technology, v. 5, p. 29–37(8), January 2011. ISSN 1751-8687.

Vempati, N.; Slutsker, I.; Tinney, W. Enhancement to givens rotations for power system state estimation. **IEEE Transactions on Power Systems**, v. 6, n. 2, p. 842–849, 1991.

Vigliassi, M. P. et al. Multi-objective evolutionary algorithm in tables for placement of scada and pmu considering the concept of pareto frontier. **International Journal of Electrical Power & Energy Systems**, v. 106, p. 373–382, 2019. ISSN 0142-0615.

Wang, G. et al. Distribution system state estimation: An overview of recent developments. **Frontiers of Information Technology & Electronic Engineering**, Springer, v. 20, n. 1, p. 4–17, 2019.

Wang, S. et al. Assessing gaussian assumption of pmu measurement error using field data. **IEEE Transactions on Power Delivery**, v. 33, n. 6, p. 3233–3236, 2018.

Whei-Min Lin; Jen-Hao Teng. State estimation for distribution systems with zero-injection constraints. **IEEE Transactions on Power Systems**, v. 11, n. 1, p. 518–524, 1996.

Wu, F. F.; Monticelli, A. Network observability: Theory. **IEEE Transactions on Power Apparatus and Systems**, PAS-104, n. 5, p. 1042–1048, 1985.

Wu, W. et al. Robust state estimation method based on maximum exponential square. **IET generation, transmission & distribution**, IET, v. 5, n. 11, p. 1165–1172, 2011.

Xygkis, T. C.; Korres, G. N. Optimized measurement allocation for power distribution systems using mixed integer sdp. **IEEE Transactions on Instrumentation and Measurement**, v. 66, n. 11, p. 2967–2976, 2017.

Yannakakis, M. Computing the minimum fill-in is np-complete. **SIAM Journal on Algebraic Discrete Methods**, v. 2, n. 1, p. 77–79, 1981.

Youman Deng; Ying He; Boming Zhang. A branch-estimation-based state estimation method for radial distribution systems. **IEEE Transactions on Power Delivery**, v. 17, n. 4, p. 1057–1062, 2002.

Zamzam, A. S.; Sidiropoulos, N. D. Physics-aware neural networks for distribution system state estimation. **IEEE Transactions on Power Systems**, v. 35, n. 6, p. 4347–4356, 2020.

Zhang, L.; Wang, G.; Giannakis, G. B. Real-time power system state estimation and forecasting via deep unrolled neural networks. **IEEE Transactions on Signal Processing**, v. 67, n. 15, p. 4069–4077, 2019.

Zhao, J. et al. Power System Dynamic State Estimation: Motivations, Definitions, Methodologies, and Future Work. **IEEE Transactions on Power Systems**, v. 34, n. 4, p. 3188–3198, July 2019. ISSN 1558-0679.

____. Robust medium-voltage distribution system state estimation using multi-source data. In: **2020 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)**. [S.l.: s.n.], 2020. p. 1–5.

Zhao, J.; Mili, L. A framework for robust hybrid state estimation with unknown measurement noise statistics. **IEEE Transactions on Industrial Informatics**, v. 14, n. 5, p. 1866–1875, 2018.

____. Vulnerability of the largest normalized residual statistical test to leverage points. **IEEE Transactions on Power Systems**, v. 33, n. 4, p. 4643–4646, 2018.

Zhao, J. et al. Roles of dynamic state estimation in power system modeling, monitoring and operation. **IEEE Transactions on Power Systems**, p. 1–1, 2020.

____. Power system real-time monitoring by using pmu-based robust state estimation method. **IEEE Transactions on Smart Grid**, v. 7, n. 1, p. 300–309, 2016.

Zhou, M. et al. An Alternative for Including Phasor Measurements in State Estimators. **IEEE Transactions on Power Systems**, v. 21, n. 4, p. 1930–1937, nov 2006. ISSN 0885-8950.

Zhou, N. et al. Dynamic state estimation of a synchronous machine using pmu data: A comparative study. **IEEE Transactions on Smart Grid**, v. 6, n. 1, p. 450–460, 2015.

Zhou, W. et al. Bayesian learning-based harmonic state estimation in distribution systems with smart meter and dpmu data. **IEEE Transactions on Smart Grid**, v. 11, n. 1, p. 832–845, 2020.

Zhou, X. et al. Gradient-based multi-area distribution system state estimation. **IEEE Transactions on Smart Grid**, v. 11, n. 6, p. 5325–5338, 2020.

Zhu, J.; Abur, A. Effect of phasor measurements on the choice of reference bus for state estimation. In: **2007 IEEE Power Engineering Society General Meeting**. [S.l.: s.n.], 2007. p. 1–5.

Švenda, G.; Strezoski, V.; Kanjuh, S. Real-life distribution state estimation integrated in the distribution management system. **International Transactions on Electrical Energy Systems**, v. 27, n. 5, p. e2296, 2017. E2296 ETEP-15-0947.R1.

Appendix

APPENDIX A – POWER SYSTEM STATE ESTIMATION THEORETICAL BACKGROUND

Introduction

This Appendix presents the theoretical background for the power system state estimation, presenting the main formulation of the Weighted Least Squares state estimator, observability and bad data processing. The main concepts of multiarea state estimation for deployment in large scale networks are also presented. Followed by state-space formulations for dynamic state estimations that enlarge the conceptual framework of the estimation process by representing temporal relations among the state variables. Finally, it also presents the main algorithms developed for distribution system state estimation are provided as a reference and also to demonstrate the fundamental aspects of practical implementation. The objective is to provide the main concepts ant theoretical aspects of modern power system state estimation, both general formulations as well as specialized in three-phase unbalanced distribution systems.

Static State Estimation and the Maximum Likelihood Perspective

The context of real time monitoring is related to most of the operation and automation processes of power systems. The goal is to obtain a proper evaluation of power quality, network reliability, risk assessment, and a more efficient use of the components of the electrical networks. Thereby, to improve energy management capabilities, operators rely on the accuracy of network modelling to assess system features. Network operators must be able to evaluate the condition of the grid taking into account the simultaneous interaction among multiple components and the data gathered from measurement devices (Hansen; Debs, 1995; Abur; Gómez-Expósito, 2004; Thomas, 2015; Zhao et al., 2020).

The state estimation problem emerges from this scenario to capture the relations among measured electrical quantities and the network state. It is traditionally formulated as static analysis, that is, intending to obtain steady state features of the network. It has a strong relation with power flow analysis. Although both problems frequently share similar network models they have some conceptual differences. The main one regards the input data. In state estimation the steady state is evaluated from measured values, assumed with some degree of uncertainty, and able to deal with redundancy. Whereas in power flow analysis, a specified loading/generation condition is evaluated, without any redundancy.

Regarding the available information for real time of power systems, it can be divided according to the respective sources of data as:

- **Real Time Measurements**: measured values from the remote Terminal Units (RTUs) and

Intelligent Electronic Devices (IEDs) from the SCADA systems (also called conventional measurements), typically composed by active and reactive power flow measurements, active and reactive power injections, voltage magnitude measurements and current magnitude measurements, and also measured synchronized phasors from Phasor Measurement Units (PMUs), composed by voltage magnitude and voltage phase angles, and current magnitudes and current phase angles;

- **Pseudo Measurements**: information obtaines from a previous (or *a priori*) knowledge about the electrical network, that represent unmonitored electrical quantities in order to obtain Observability of the electrical network, that is, provide enough information to enable the mathematical solution of the state estimation problem. It generally is based on forecasting methods to accommodate momentarily loss of instrumentation devices, or in typical characterization about loads and generation. Usually in distribution systems consists of the majority of information available, representing active and reactive power injections on distribution transformers, obtained from typical load curves, consumer stratification and weather information. In comparison with the real time measurements, the pseudo measurements present lower precision, carrying large uncertainty for the state estimation process in distribution systems;

- **Virtual Measurements**: correspond to known electrical quantities of the network, mainly representing passive nodes with zero injection, that is without generation or loads, or internal nodes of equipments. The information defined as virtual measurements are formally considered as known and correct, even tough there is not an associated physical measurement. This way they are typically treated as highly accurate measurements, with low variance, or as equality constraints in the state estimation problem;

- **Logical Measurements**: measurements and information of discrete values that influence the electrical network model, such as statuses of switching devices and circuit breakers, taps of voltage regulators, status of automatic capacitor banks, and discrete variables of controllers, that can be telemetered, or also sometimes rely on updates of database in the operation centre;

- **Network Data**: information regarding the structural and physical nature of the electrical networks, composed by parameters and models for each component, and the full topology of the network with all components.

Within this perspective, the traditional state estimation problem for a power system with m measurements, and n state variables is formulated according to the following non-linear measurement model (Monticelli, 1999; Abur; Gómez-Expósito, 2004; Zhao et al., 2019):

$$z = h(x) + \varepsilon \tag{A.1}$$

where, z is the measurement vector (m x 1) composed by the measured values; x is the state vector (n x 1) typically composed by the complex nodal voltage; $h(x)$ is the set of non-linear equations (m x 1) that relates the measurement vector with the state variables; and ε is the measurement error vector (m x 1) often assumed as independent with Gaussian distribution of zero mean and known covariance matrix R (m x m).

Due to the statistical nature of the measurement model, associated with the error vector modelled as a random variable, the following equation provides the joint probability function of the above measurement model:

$$f(\varepsilon|x) = f(\varepsilon_1|x,...,\varepsilon_m|x) = \prod_{i=1}^{m} f(\varepsilon_i|x) = \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(z_i - h_i(x))^2}{2\sigma_i^2}} \tag{A.2}$$

where σ_i is the i-th measurement standard deviation, related to the precision class of the metering and instrumentation devices, that also compose the diagonal elements of the covariance matrix R.

The log-likelihood of the joint probability provides the conventional approach to obtain the state vector x, as the maximum likelihood estimation between the measurement model $h(x)$ and measured values z:

$$\log L(x) = \log f(\varepsilon|x) = -\sum_{i=1}^{m} \log\sqrt{2\pi}\sigma_i^2 - \frac{1}{2}\sum_{i=1}^{m} \frac{(z_i - h_i(x))^2}{\sigma_i^2} \tag{A.3}$$

The maximum likelihood can then be obtained by the following unconstrained optimization problem, known as Weighted Least Squares (WLS). The R^{-1} matrix is also known as weighting matrix of the state estimation problem, since it reflect the weights for the measurement residuals in the WLS criterion.

$$\min J(x) = \frac{1}{2}\sum_{i=1}^{m} \frac{(z_i - h_i(x))^2}{\sigma_i^2} = \frac{1}{2}[z - h(x)]'.R^{-1}.[z - h(x)] \tag{A.4}$$

The gradient vector of the above minimization problem is given by:

$$\nabla J(x) = -H(x)'.R^{-1}.[z - h(x)] \tag{A.5}$$

where $H(x)$ is the Jacobian matrix of the non-linear measurement model $h(x)$.

And the Hessian matrix of the problem is given by:

$$\nabla^2 J(x) = H(x)'.R^{-1}.H(x) - \sum_{i=1}^{m} \frac{(z_i - h_i(x))}{\sigma_i^2}.\frac{\partial^2 h_i(x)}{\partial x^2} \tag{A.6}$$

An important complementary aspect about the measurement model is the concept of *Observability*. A system is said to be observable if the quantity. location and type of measurements

can provide the solution of the above minimization problem. One way to represent such concept is the algebraic interpretation as used in (London Jr et al., 2009), where a system with $n$state variables is said algebraically observable if the Jacobian matrix $H(x)$ has full rank:

$$rank(H(x)) = n \tag{A.7}$$

If the optimization problem can be solved without convergence issues, the system is said to be numerically observable. There is a third observability concept based on graph theory and network topology, associated to a spanning tree related to the metering system and power system topology (Fantin, 2016).

The conventional solution for the state estimation problem, that is, obtaining the optimal value of the state vector x, is based on the linearisation of the model $h(x)$, given by equation (A.4) e employing the iterative Gauss-Newton method (Abur; Gómez-Expósito, 2004). The linearisation of $h(x)$ is obtained through the first-order terms of the Taylor expansion in the vicinity of an operating point x^k. This way, equation (A.4) is rewritten as:

$$\min J(x) = \frac{1}{2}[z-(h(x^k)+H(x^k).\Delta x)]'.R^{-1}.[z-(h(x^k)+H(x^k).\Delta x)] \tag{A.8}$$

where $\Delta x = (x^{k+1}-x^k)$ is the state vector update to be used in the next iteration of the method $k+1$.

The optimal value of x is obtained by the solution of the minimization problem in (A.8) through the first-order optimality condition:

$$H(x^k)'.R^{-1}.H(x^k).\Delta x = H(x^k)'.R^{-1}.[z-(h(x^k)] \tag{A.9}$$

where $G(x) = H(x^k)'.R^{-1}.H(x^k)$ is known as the Gain matrix of the state estimator. The solution of the above linear syste is usually obtained by direct methods for linear system, such as the factorization methods (LU and QR). The second-order optimality condition in this case is given by:

$$H(x^k)'.R^{-1}.H(x^k) > 0 \tag{A.10}$$

If the system is algebraically observable, then $H(x)$ is full rank, as previously mentioned. Besides, following the covariance matrix definition ensures that R and R^{-1} are positive definite. Thereby, a solution of (A.9) satisfies the second-order optimality condition.

The method is iterative in the sense that equation (A.9) is calculated sequentially upon convergence. Convergence is obtained whenever the values of values of x in two successive iterations presents a difference below a numerical tolerance pre-specified tol:

$$\|\Delta x\|_\infty \leq tol \tag{A.11}$$

The presented mathematical formulation for the state estimation problem for power systems consists of the general version as an unconstrained optimization problem defined by A.4. It is noteworthy there are different propositions to the above problem, by including constraints to the problem or also interchanging the objective function. Typically, equality constraints are incorporated to treat virtual measurements, that often represent nodes without any injection, that is, passive nodes without loads or generation. Regarding different objective functions, they are often associated with robust estimation methodologies with the goal of reducing the effects of possible gross errors in the estimation (also referred as bad data or outliers) (Abur; Gómez-Expósito, 2004).

Observability Analysis

A power system is said to be observable if, with a given measurement set, it is possible to obtain all of its state variables. This concept is related to the existence of a solution to the equation (A.4). Through the years some methods to evaluate the observability were developed (London Jr; Alberto; Bretas, 2007; Bretas, 1996; Wu; Monticelli, 1985). Formally, the work in (Krumpholz; Clements; Davis, 1980) presents defines the observability concept under three perspectives:

- Topological Observability: "A power system is said to be observable with respect to the given measurements if there exist an observable spanning tree associated to the metering system and power system network"
- Algebraic Observability: “A power system is defined algebraically observable with respect to the given measurements if the Jacobian Matrix ($H(x)$) has rank equal the number of state variables."
- Numerical Observability: “A power system is said numerically observable if the measurement model can be iterativelly solved for a state estimate from flat start"

The topological analysis is based on the characterization of the metering systems according the graph theory. By the definition, this analysis consists of searching for a spanning-tree, through the associated graph with the metering system G_m and the power network connectivity. A spanning tree consists of a sub-graph of G_m that is a a tree, a undirected connected graph without cycles (Krumpholz; Clements; Davis, 1980; Costa; Lourenco; Clements, 2002). The analysis seeks different combinations of spanning trees, depending on the type and location of measurements, that connects all nodes of the graph associated to the network topology, with at least one different measurement associated to each branch of the spanning tree.

The algebraic observability analysis evaluates the rank of the Jacobian matrix $H(x)$ of the measurement model. By definition, if the system is observable then the Jacobian matrix is full rank, rank equals the number of state variables n, as describe in (A.12). The notion of rank of a

matrix is related to the concept of linear dependency among its column-vectors (or row-vectors) (Golub; Van Loan, 2012). Different methods exploit such dependencies to seek for redundancy and observability features of the power system, such as the work in (London Jr; Alberto; Bretas, 2007) that exploit factorization paths of the Jacobian matrix to evaluate observability qualitative aspects.

$$rank(H(x)) = n \tag{A.12}$$

Some authors include the algebraic analysis as a particular case of numerical observability analysis (Abur; Gómez-Expósito, 2004). In this sense, it requires also complementary concepts regarding convergence of the non-linear optimization algorithms and numerical stability. Regarding the conventional Gauss-Newthon method, widely employed in the literature, the following Theorem attests its convergence:

Theorem (demonstration in (Nocedal; Wright, 2006)): Supposing the residual function $(z-h(x))$ is Lipschitz continuous and differentiable in a region of the domain where the objective function is limited $\mathscr{N} := \{x|J(x) \leq J(x_0)\}$, and that the Jacobian $H(x)$ satisfies the full rank condition in $\mathscr{N}$. Then if the iterations x^k are generated by the Gauss-Netwon method with step-length satisfying th e Wolfe condition, yields:

$$\lim_{k\to\infty} H^T(x^k)(z-h(x^k) = 0.$$

The Wolfe conditions, are decreasing conditions that ensure sufficient decrease (Armijo condition) and a curvature condition in each iteration (Nocedal; Wright, 2006):

$$\begin{aligned} J(x^k + \alpha^k \Delta x) &\leq J(x^k) + c_1 \alpha \\ \nabla J(x^k + \alpha^k \Delta x)^T \Delta x &\geq c_2 \nabla J(x^k) \Delta x, \end{aligned} \tag{A.13}$$

where, c_1 and c_2 are constants with $0 < c_1 < c_2 < 1$. They ensure a step length in each iteration towards the minimum that is not too small, with slow convergence, and not too large, that may cause divergence. Typically, are guaranteed by inexact line search methods for choosing the step-length, for instance based on *backtracking* algorithms (Nocedal; Wright, 2006).

Besides the convergence characteristic from the above theorem, another important aspect is numerical stability, specially in the solution of the linear systems associated to the Normal equations in (A.9), during each iteration of the state estimator. This concept of numerical stability is related to the possibility of performing different computations, without losing solution accuracy, while considering the limitations imposed by floating-point arithmetic in modern computer architectures. Briefly, it is strongly related to the concept of numerical conditioning, and the condition number of the Normal equations.

Bad Data Processing

One of the most essential features of state estimation is the possibility to detect and identify bad data (or gross errors), besides filtering white noise from the measurement process (Monticelli, 1999; Abur; Gómez-Expósito, 2004). These errors originate from bad calibration of the measurement equipment, current transformer saturation, communication errors and even cyber attacks. The effects of bad data in the estimation process can be very severe, reducing accuracy and even compromising functionalities that rely on the real-time state of the network.

In order to suppress such effects, the WLS estimator is associated with techniques for bad data processing (detection and identification), many of them based on the analysis of the residuals of the WLS estimator. Among them, the Largest Normalized Residual Test (LNRT) and the b-hat Test are the most popular (Monticelli; Garcia, 1983; Abur; Gómez-Expósito, 2004). They are based on the following claim typically made regarding bad data processing (Abur; Gómez-Expósito, 2004): when there is a single Gross Error (GE), the largest normalized residual (absolute value) corresponds to the bad measurement, provided that it is not critical and its removal does not create any critical measurement among the remaining ones. However, such claim is not suitable for practical applications in real-life state estimators (Zhao; Mili, 2018b; Massignan et al., 2020), since it is based on the hypothesis that only one measurement is contaminated with GE and the remaining measurements are exact (free of both white noise and GEs).

Such limitation of the WLS associated with the LRNT, motivated research on robust state estimators. These estimators do not rely on the WLS criterion and are based on different statistics to perform the estimation. Among them, the most prominent, and with extended research on the topic, are the Weighted Least Absolute Value (WLAV) in (Celik; Abur, 1992), the Least Median Squared (LMS) in (Mili; Phaniraj; Rousseeuw, 1991), and the Schweppe-Hubber Generalized M-Estimator (SHGM) in (Mili et al., 1996). In another perspective, a broader notion of the inherent error probability function presents promising features towards bad data processing, assuming Gaussian mixtures and the SHGM estimator as in (Zhao; Mili, 2018a), or with Information Theory concepts and a kernel density estimator in (Miranda; Santos; Pereira, 2009).

The bad data or gross errors (or outliers) processing step is performed after the estimated state $\hat{x}$ is obtained. It is necessary to calculate the normalised residue and, for that, the residue's covariance matrix Ω is used as follows:

$$\Omega = W^{-1} - H(\hat{x})(H(\hat{x})^T W H(\hat{x}))^{-1} H(\hat{x})^T \tag{A.14}$$

With the diagonal elements of the matrix Ω, three different indexes are calculated to guide the bad data detection and identification. First, the normalised residue is calculated for

each measurement:

$$r_i^N = \frac{|z_i - h_i(\hat{x})|}{\sqrt{\Omega_{ii}}}. \tag{A.15}$$

With the normalised residue, it's possible to calculate the index of the b-hat method, following:

$$\hat{b}_i = \frac{\sigma_{ii}}{\sqrt{\Omega_{ii}}} r_i^N, \tag{A.16}$$

where σ_{ii} is the standard deviation of the $i-th$ measurement, obtained from the diagonal of the matrix W.

The maximum normalized residue test enables the detection and identification of a measurement with GE and the identification of critical measurements, that is, if the removal of said measurement renders the system not observable. The b-hat test enables the correction of a measurement affected by gross errors.

Multiarea State Estimation

Following the presentation of the formulation of state estimation in power systems, this section presents the main concepts related to Multiarea State Estimation (MASE). The Multiarea State Estimator corresponds to the efficient application of state estimation procedures in large scale power systems, seeking computational performance gains (processing time, memory allocation, processing capacity), by exploring the fact that measurements are obtained from a wide area spread across the electrical network (Gómez-Exposito et al., 2011b). The basic approach of such estimation process consists of the separation of the power systems into sub-areas, in which a local state estimation is formulated for the internal nodes of such sub-areas and with a special treatment for frontier regions (boundaries of each area). Different architectures of MASE perform the state estimation process separately for each area, and the results refine the estimation of the frontier nodes. This section presents the nomenclature, definitions, classifications and characteristics proposed for MASE (Gómez-Exposito et al., 2011b). Such concepts are important to any practical implementation that aim in large scale applications.

Initially, given a power system by A connected sub-areas, that is, a power system that if was treated as a centralized approach it would have a single area $A = 1$. Let us denote as S_k the set of nodes that belong to the k-th area and S the set of nodes of the whole power system. Initially, is possible to define:

$$S = \bigcup_{k=1}^{A} S_k \tag{A.17}$$

According to the level of overlapping among areas, the MASE can be classified as the following, illustrated in Fig. 55:

1. *MASE without overlapping areas*: different areas does nor present any nodes or branches in common;

2. *MASE with overlapping nodes*: adjacent areas share a common node in a single level of interconnection (only the nodes at the boundaries);

3. *MASE with extended overlapping nodes*: adjacent areas share more than one common nodes in a multiple levels of interconnection (more nodes besides the ones at the boundaries);

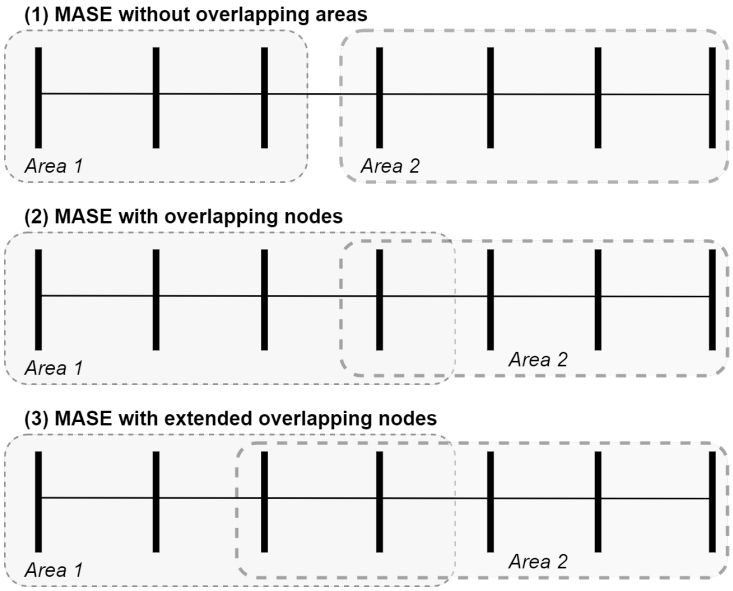


Figure 55 – Different types of overlapping zones among different areas in MASE.

Thus, depending on the level of overlapping the state variables vector in each area may be composed by:

- x_{ik}: state variables associated to the internal nodes of each area k;
- x_{bk}: state variables associated to the boundary nodes of each area k;
- x_{nk}: state variables associated to the internal nodes of adjacent areas to the area k (areas with extended overlapping zones).

Besides, the measurement vector can be devised for each area as:

- z_{ik}: internal measurements, that relates internal state variables x_{ik} and boundary state variables x_{bk};
- z_{bk}: boundary measurements, that relates boundary state variables x_{bk}, as well as the state variables of extended overlapping zones x_{nk};

The measurement model for the centralized state estimation in (A.1), can be rewritten for each area k separately according to the equations below, decomposing internal measurements and boundary measurements (Korres, 2011):

$$\begin{aligned} z_{ik} &= h_{ik}(x_{ik}) + \varepsilon_{ik} \qquad k = 1, ..., A \\ z_{bk} &= h_{bk}(x_{bk}, x_{nk}) + \varepsilon_{bk} \qquad k = 1, ..., A \end{aligned} \tag{A.18}$$

From a numerical optimization perspective, the MASE can be seen as a particular application of decomposition techniques (Gonzalez et al., 2015; Korres, 2011; Conejo et al., 2006). With the above measurement model, the state estimation problem can be formulated to minimize the WLS criterion, as previously described. Thereby, the MASE can be formulated as the following constrained optimization problem for the above measurement model (Korres, 2011):

$$\begin{aligned} \min \quad & J(x) = \frac{1}{2} r_b'.R_b^{-1}.r_b + \frac{1}{2}\sum_{k=1}^{A} r_{ik}'.R_{ik}^{-1}.r_{ik} \\ \text{s.a.} \quad & r_b - z_b + h_b(x) = 0 \end{aligned} \tag{A.19}$$

where the subscript b represents the boundary measurements for all sub-areas of the electrical network with the respective weighting matrix W_b (the inverse of the boundary measurement error covariance matrix $W_b = R_b^{-1}$), $r_{ik} = z_{ik} - h_{ik}(xik)$ is the internal measurements residual for each are area k and respective weighting matrix R_i^{-1}. Although it has been presented as a constrained problem, it can also be written as an unconstrained problem.

The multiarea state estimation problem consists of separating the above objective function into two parts: one related to the internal measurements (the local estimation process); and the second related to boundary measurements (the coordination process). Different methods have been proposed to solve the MASE problem, as presented in (Gómez-Exposito et al., 2011b). They mostly rely on the traditional state estimation WLS criterion both at the local as well as the coordination level. Some approaches apply heuristics to simplify the optimality conditions of the problem. Also, there are methods based on equality constraints among adjacent areas that are solved by variations of the Lagrangean method and other non-linear programming framework (Conejo et al., 2006).

Regarding the developed architectures, an important characteristic is the possibility to interchange information among different areas during the MASE process. As a consequence, the architectures are mostly related to the computer architecture implemented. They can be separated in two main categories (Gómez-Exposito et al., 2011b):

- **MASE with Hierarchical Architectures**: a local estimation process is performed in each area independently, followed by a centralized coordination process that synchronizes and coordinates different local estimations;

- **MASE with Distributed Architectures**: in this case there is not a centralized coordination, and the local process are performed considering an exchange of information in neighbour areas. The coordination among areas is usually performed by substituting boundary state variables with the latest estimated values, thus a relaxed version of the centralized estimation problem.

Regarding the coordination among local estimation, the Hierarchical MASE can be divided into (Gómez-Exposito et al., 2011b):

- **Coordination at the state estimation level**: local estimators provide the final converged solution from their respective areas to the central coordination. Whenever only on coordination exchange process is performed, this methodology provide sub-optimal estimates. However, this approach requires less adaptations to already available state estimators;
- **Coordination at the iteration level**: local estimators provide information for the centralized coordination step at each iteration of the local convergence process. The estimations are coordinated at each iteration, enabling faster convergence to the optimal solution of the state estimator. This approach requires adaptations at the core algorithms of traditional state estimators and demands faster communication among local estimators and the central coordinator;
- **Hybrid coordination**: consists of balanced option among the two previous processes, where at some amount of iterations the local estimations are provided for the central coordinator.

Another important aspect in any MASE is the synchronization among measurements and the processes, specially for distributed architectures. Measurement synchronization is important since the updating rate of the SCADA systems may present a large range, and, at some cases, even fail to communicate. This problem can be counter by accepting some level of sub-optimality or by improving the communication and redundancy of the SCADA system. Besides, the incorporation of synchronized phasor measurements requires special treatment across the different areas during the estimation process (Gómez-Exposito et al., 2011b). Regarding the estimation process synchronization, it comprises the updating rate and availability of information among adjacent area, or among areas and the central coordinator. In general, the central or the decentralized processes use the latest available information from adjacent areas or assume some standby time to receive missing informations.

Dynamic State Estimation

This section introduces the concept of a state space model for power system state estimation. It carries a fundamental difference from the static perspective that can be widened by

the Bayesian Inference perspective and foundations, the main theoretical background this work explores.

Since its conception in the early 1970s, power system state estimation has been focused on processing the steady-state condition of the transmissions systems, under a static analysis framework (Monticelli, 1999; Abur; Gómez-Expósito, 2004). With the penetration of renewable energy and new control equipment, the static perspective does not provide the modern energy management systems' requirements. It became imperative to capture system's dynamics for advanced automated applications, in such a manner to provide fast control actions, increase the reliability of the system, counter cyber-attacks and optimize resources across the power grid (Zhao et al., 2019; Zhao et al., 2020).

As far as this thesis explores, dynamic state estimation in electric power systems means estimating the network states, voltages magnitude and phases, for an interconnected power system, considering the time relation between states. There are many dynamic state estimation formulations, the Kalman Filter formulation being one of them (Kalman, 1960). In dynamic state estimation, due to the time relations between the states, beyond the measurement gross error, the occurrence of another anomaly may occur, such as in case of sudden large load change, yielding a larger process error (Zhao et al., 2019).

In the effort of surpassing many of the practical and theoretical challenges, different versions of the Kalman Filter were proposed, such as the Linear Kalman Filter (Sarri et al., 2016), the Extended Kalman Filter (Fan; Wehbe, 2013), the Unscented Kalman Filter (Valverde; Terzija, 2011), the Cubature Kalman Filter (Sharma; Srivastava; Chakrabarti, 2017) and the Ensemble Kalman Filter (Zhou et al., 2015). Nonetheless, such approaches consider a single stationary scenario or that known changes occurs on the system. It translates on the assumption that the system's transition matrix is always correct, what does not hold on a practical scenario. Besides, some applications focus on generator and controller responses during transient events, disregarding the bulk power system, dismissing its multivariate nature.

Within the power systems realm, dynamic state estimation has been conceptualized under different perspectives, mainly related to the definition of state variables. The fundamental difference comprises the inclusion, or not, of internal variables from generators and loads (also referred as dynamic state variables), besides the complex nodal voltages of the power network (also referred as algebraic state variables) (Zhao et al., 2019; Zhao et al., 2020). The first focus on fast transient events that may cause instabilities, while the second on slower rates of change that affect the loading condition of the power networks. Nonetheless, both represent a system subject to a stochastic phenomenon and uncertainty. The difference lies on the time scale of events and observed response of the system.

Without loss of generalization, the power system state estimation can be associated with a first-order nonlinear state space model in each instant t, given by:

$$\begin{aligned} x_t &= f(x_{t-1}) + \omega_{t-1} \\ z_t &= h(x_t) + e \end{aligned} \tag{A.20}$$

where, the first equation is known as process model, with $f(x_{t-1})$ the state vector in a previous instant $(t-1)$ and ω_{t-1} is the process noise, the second equation is the measurement model presented in the previous section. The above model can also be extended to incorporate the operational constraints of the power system components, such as generation capability and physical limitations of controllers. There is the possibility to increase the model order, incorporating additional temporal delays, however this is not a common practice for power systems.

There are three main approaches to deal with the above state-space under perspective of dynamic state estimation for power systems:

- *Tracking State Estimation*: this formulation assumes the previous instant as a good approximation for the current estimation step, by a direct relation from all state variables from a previous instant with the following;
- *Forecasting-Aided State Estimation*: this formulation introduces a forecasting stage to capture the temporal relations that may arise from algebraic variables in a linear model, employing linear regression and trend models to predict future state variables;
- *Dynamic State Estimation*: this formulation extends the state variables to accommodate internal variables for generators and controllers, formulating a differential and algebraic set of equations to model the system behaviour;

Tracking State Estimation

The tracking state estimation consists in the following discrete time-variant nonlinear model:

$$\begin{aligned} x_t &= x_{t-1} + \omega_{t-1} \\ z_t &= h(x_t) + e \end{aligned} \tag{A.21}$$

The estimated state is then obtained by a filtering process also based on the MLE principle. The solution of the estimation problem, i.e., obtaining the state variables at the current instant t ($\hat{x}_t$), has been treated in several ways, such as the ones based on Kalman Filter approaches (Sarri et al., 2016; Coutto Filho; Souza, 2009) or with augmented WLS formulation such as (Simões

Costa; Albuquerque; Bez, 2013; Zhao et al., 2016). In this case with the following optimization problem:

$$\min \quad [z_t - h(x_t)]^T R^{-1} [z_t - h(x_t)] \quad + [x_{t-1} - x_t]^T Q_{t-1}^{-1} [x_{t-1} - x_t] \tag{A.22}$$

Regarding the Kalman Filter approach, it consists of a two-step recursive algorithm, with the following formulation for the tracking state estimation model:

Prediction Step: The prior mean and covariance matrix are given by:

$$\begin{aligned} \hat{x}_{t|t-1} &= \hat{x}_{t-1} \\ P_{t|t-1} &= P_{t-1} + Q \end{aligned} \tag{A.23}$$

Update Step: The Kalman Filter gain, the posterior state and posterior covariance are then updated:

$$K_t = P_{t|t-1} H^T \left(H P_{t|t-1} H^T + R \right)^{-1} \tag{A.24}$$

$$\hat{x}_t = \hat{x}_{t|t-1} + K_t \left(z_t - H \hat{x}_{t|t-1} \right) P_t = (I - K_t H) P_{t|t-1} (I - K_t H)^T + K_t R K_t^T \tag{A.25}$$

where, H is the Jacobian matrix of the measurement model, K_t is the Kalman Filter Gain matrix, I is the identity matrix and P is the calculated state covariance matrix, t is the current time step, and $t|t\text{-}1$ stands for the conditional relation between two time-steps. The nonlinear version of the above model corresponds to the WLS Extended Kalman Filter (EKF) that performs linearization of the measurement equations, and the Iterated EKF (IEKF) that updates such linearization iteratively, thus presenting more accurate results.

Forecasting Aided State Estimation

The concept of Forecasting-Aided State Estimation (FASE) is also based on a state space model with the introduction of trends and a transition matrix to represent the relation between two successive instants. The formulation consists of the measurement model (6.1) and a forecasting model that relates the system's state at different instants. In order to prevent semantic confusion with the term dynamic state estimation (Rouhani; Abur, 2018), much common among stability research, the authors in (Coutto Filho; Souza, 2009) indicate the FASE terminology as more interesting for the state estimation area. In this sense, FASE is intended to describe "slow time evolution of the static-state, observed from multiple scans of measurements" (**??**). The state space model can be used with two different goals: to relate the current state to a previous one by the forecasting model; or to predict the next state. The first approach can be written as the following discrete time variant system:

$$\begin{aligned} x_t &= F_{t-1}.x_{t-1} + g_{t-1} + \omega_{t-1} \\ z_t &= h(x_t) + e \end{aligned} \tag{A.26}$$

where, F_{t-1} is the state transition matrix; g_{t-1} represents a forecasting trend; ω_{t-1} is the forecasting error assumed as white Gaussian noise with zero mean and covariance matrix Q_{t-1} (**??**); and the second set of equations is the measurement model at the respective instant t, in this case the nonlinear model. But if only PMUs are considered it is possible to use the linear measurement model (C.3).

The solution of the above model is also accomplished by employing the Kalman Filter approach, in a two-step algorithm similar to the previously presented, but in this case considering also the state transition model. The tracking state estimation is actually a simplified version of the forecasting aided state estimation model, assuming an identity matrix as the state transition matrix and neglecting trends.

Different methods may be employed for estimating the transition matrix, such as the exponential smoothing approach, using Holt's method, or different regression models according to learning algorithms as well. The method however is also highly affected by abrupt changes or systemic transition that modify the stochastic behaviour of the system.

Dynamic State Estimation

This formulation entails a detailed internal variables of electrical machines and components, that is, generator modelling along with their respective controllers, into the state estimation formulation. Also the load dynamics may be incorporated in the model. Such variables are also called dynamic state variables y. And also keep the algebraic state variables x of the electric power network (the complex nodal voltages). Thus it consists of a set of Differential Algebraic Equations, that describe the fast transients along with the interconnected power network in a single estimation formulation:

$$\begin{aligned} \dot{y} &= f(y,x) + \omega \\ z &= h(y,x) + e \end{aligned} \tag{A.27}$$

The solution method in this case, is obtained through numerical integration methods such as runge-Kutta methods and Euler Interpolation. Often the stochastic aspects of the algebraic state variables (complex nodal voltages of the power network). The main motivation towards this kind of formulation is encompassing measurements with fast sampling, such as the Phasor Measurement Units and sampled values from the IEC 61850 protocol, and enabling advanced

closed-loop autonomous control. More details about such type of formulation may be referred in (Zhao et al., 2020).

Distribution System State Estimation

Despite an extensive application and research on transmission systems, up to these moment the state estimators have not been widely deployed on distribution systems. In this context it is noteworthy that the distribution networks present singular characteristics to be considered into the estimation process. Therefore the state estimators consolidated in transmission system cannot be directly applied to distribution systems. (Singh; Pal; Jabr, 2009; Baran, 2012; Lefebvre; Prévost; Lenoir, 2014; Švenda; Strezoski; Kanjuh, 2017; Ahmad et al., 2018). In order to deal with such challenges, specific approaches of state estimation were developed for distribution systems, starting mainly in the beginning of the 1990s. As presented in the previous sections, regarding distribution systems three main algorithms were developed to deal with the particularities of such systems, the WLS estimator, the AMB estimator and the BC estimator. The next section presents the algorithms of these three main approaches for distribution system estimation, discussing their main theoretical concepts and particularities.

Another important aspect of distribution systems is the requirement of three-phase mathematical representation of the electrical network and its components. That is, with a greater level of details when compared to the transmission system, where the model is often represented only by its positive sequence circuit. The major effort for a proper modelling in distribution system state estimation is to capture the unbalanced and asymmetrical nature of such networks. A three-phase representation of the network enables evaluating power quality issues, and also performing inferences about the system condition in each phase and at different parts of the power grid. This section will focus on the algorithms developed to solve the distribution system state estimation problem, providing some guidance related to the three-phase aspect. Detailed information about the three-phase unbalanced and asymmetrical models are presented in Chapter 4.

Weighted Least Squares Estimator

The application of the conventional approach of power system state estimation for DSs is based directly on the theoretical formulation of the Weighted Least Squares (WLS) method presented in the previous section, and was initially proposed for DSs in (Baran; Kelley, 1994; Whei-Min Lin; Jen-Hao Teng, 1996). This approach uses the complex nodal voltages (in polar coordinates) in all buses and phases of the system as state variables. Thus, for distribution network with the number of phases n_{fases} and the number of buses n_{barras}, the state vector is given by:

$$x = \begin{bmatrix} V_\phi^p & \theta_\phi^p \end{bmatrix} \qquad \begin{matrix} p \in \{1,...,n_{barras}\} \\ \phi \in \{1,...,n_{fases}\} \end{matrix} \tag{A.28}$$

where p denotes the bus from the electric network and V and θ are the voltage magnitude and voltage phase angle at each phase ϕ. This way the amount of state variables depends on the quantity of phases in each bus (single, two or three-phase, as an example).

Besides, one of the buses is taken as angular reference for the network. In general, the voltage phase angles are assumed for each phase of the reference bus, for instance as a three-phase *abc* balanced bus:

$$\theta_a = 0^o \quad \theta_b = -120^o \quad \theta_c = 120^o \tag{A.29}$$

As presented in (Silva, 2018), the treatment of the angular reference is not a trivial question for DS state estimation and highly impacts the estimator accuracy, once, in general, the primary substation is taken as the reference bus thus affecting the entire estimated state variables for the distribution feeders. A way to counter this problem, is in the presence of synchronized phasor measurements, from PMUs, where there is not the necessity of assuming phase angle values for the reference bus. Instead, the PMUs are able to provide a direct measurement of the voltage phase angles and current phase angle of the electric network, thereby providing an angle reference for the state variables (Fantin, 2016). When such measurements are not present, the reference bus can be assumed as balanced, as previously described, a rough approximation in practical situations, or a special treatment for the reference may be included, such as by including short-circuit equivalents in the substation model as in (Silva, 2018).

Another important aspect in the WLS is the measurement model $h(x)$, that is, the mathematical model that relates the state variables with the measured electrical quantities across the network. Thus we obtain the measurement vector model $h(x)$ and the respective Jacobian matrix $H(x)$ according to the respective measurements, for instance ordered according equation (A.30) for a typical set of measurements, composed by active and reactive power injections, active and reactive power flows and voltage magnitudes.

$$h(x) = \begin{pmatrix} P_\phi^{ps}(x) \\ Q_\phi^{ps}(x) \\ P_\phi^{p}(x) \\ Q_\phi^{p}(x) \\ V_\phi^{p}(x) \end{pmatrix} \quad H(x) = \begin{pmatrix} \frac{\partial P_\phi^{ps}(x)}{\partial V_\phi^k} & \frac{\partial P_\phi^{ps}(x)}{\partial \theta_\phi^k} \\ \frac{\partial Q_\phi^{ps}(x)}{\partial V_\phi^k} & \frac{\partial Q_\phi^{ps}(x)}{\partial \theta_\phi^k} \\ \frac{\partial P_\phi^{p}(x)}{\partial V_\phi^k} & \frac{\partial P_\phi^{p}(x)}{\partial \theta_\phi^k} \\ \frac{\partial Q_\phi^{p}(x)}{\partial V_\phi^k} & \frac{\partial Q_\phi^{p}(x)}{\partial \theta_\phi^k} \\ \frac{\partial V_\phi^{p}(x)}{\partial V_\phi^k} & 0 \end{pmatrix} \begin{matrix} k = 1,...,n_{barras} \\ \phi = 1,...,n_{fases} \end{matrix} \tag{A.30}$$

where, P^{ps} and Q^{ps} are the active and reactive power flow measurements in branch ps (between nodes p and s), P^{p} and Q^{p} are the active and reactive power injections measurements at node p, V^{p} are the voltage magnitude measurements at node p, and ϕ is the respective phase of each measurement. For the WLS state estimator, the expressions that build the vector $h(x)$ are obtained from a non-linear measurement model that relates the complex nodal voltages with measured active and reactive powers. This model can be obtained by the three-phase network model that will be presented in Chapter 4, according to the components of the distribution network. Besides, the Jacobian matrix elements correspond to the partial derivatives are also obtained by the network model presented in Chapter 4.

Finally, the measurement error covariance matrix R, the inverse of the weighting matrix R^{-1}, is assumed as known and built according to the precision of each element of the measurement vector. This way, the method presented in the previous section can be applied directly for the solution of the state estimation problem. The algorithm 2 presents the main steps for the WLS state estimator implementation.

Algorithm 2: WLS STATE ESTIMATOR

Input : Network parameters, network topology, x_0, tol

Output : Estimated state vector x

Begin
  $k \leftarrow 0$ e $x_k \leftarrow x_0$
  Do
    Build measurement model: $h(x^k)$ and $H(x^k)$
    Calculate Gain matrix: $G(x^k) = H(x^k)'.R^{-1}.H(x^k)$
    Solve linear system: $G(x^k).\Delta x = H(x^k)'.R^{-1}.[z-(h(x^k)]$
    Update state variables: $x^{k+1} = x^k + \Delta x$
    k = k +1
  while $\|\Delta x^k\|_\infty \leq tol$;
end

Admittance Matrix Based Estimator

The effort for searching more computationally efficient algorithms, a new state estimator was proposed in (Lu; Teng; Liu, 1995), seeking a mathematical formulation that accomplished a constant Gain matrix throughout the iterative process. This approach proposes the complex nodal voltages in all buses and phases of the network represented in rectangular coordinates as state variables. This estimator is known as Admittance-Matrix Based (AMB) state estimator (Almeida; Ochoa, 2017). Thus, the state vector is defined as follows:

$$x = \begin{bmatrix} V_{p,\phi}^{real} & V_{p,\phi}^{imag} \end{bmatrix} \qquad \begin{matrix} p \in \{1,...,n_{barras}\} \\ \phi \in \{1,...,n_{fases}\} \end{matrix} \tag{A.31}$$

where p denotes the bus of the electric network and $V_{p,\phi}^{real}$ and $V_{p,\phi}^{imag}$ are the respective real and imaginary parts of the voltage phasor in phase ϕ. Similar to the WLS state estimator, a special treatment must be given to the reference bus in the AMB estimator.

To achieve a constant Gain matrix, besides using the voltage phasors in rectangular coordinates as state variables, this estimator converts the active and reactive power measurements (injections and power flows) into equivalent current measurements. It follows the idea that the non-linear characteristic of the measurement model is due to the non-linear relations between the complex powers and the voltage and current phasors. Thus, the active and reactive power flows and injections are transformed in equivalent current measurements, separately into their real and imaginary parts, using the voltage phasor calculated at each iteration, using:

$$\dot{I}_{ps,\phi}^{eq} = Re\{I_{ps,\phi}^{eq}\} + j.Im\{I_{ps,\phi}^{eq}\} = \left(\frac{S_{ps,\phi}^{med}}{\dot{V}_{p,\phi}^{k}}\right)^{*} \tag{A.32}$$

where, $\dot{I}_{ps,\phi}^{eq}$ is the equivalent current phasor at phase ϕ in the branch ps (between buses p and s), composed by its real and imaginary parts, $S_{ps,\phi}^{med}$ is the measured complex power flow (active power $P_{ps,\phi}^{med}$ and reactive $Q_{ps,\phi}^{med}$) in phase ϕ in the branch ps, and $\dot{V}_{p,\phi}^{k}$ is the calculated complex nodal voltage at bus p in the latest iteration of the method(k).

By obtaining equivalent current measurements, the measurement vector is composed only by complex currents in rectangular coordinates, that presents a linear relation with the complex voltages also in rectangular coordinates. As an example of such relation, an active and reactive power flow measurement in branch ps is transformed into equivalent measurement according to A.32. The measurement model in this case can be written using the admittance matrix of the branch ps as:

$$h(x) = \begin{pmatrix} Re\{I_{ps}^{eq}\} \\ Im\{I_{ps}^{eq}\} \end{pmatrix} = \begin{pmatrix} G^{pp} & G^{ps} & -B^{pp} & -B^{ps} \\ B^{pp} & B^{ps} & G^{pp} & G^{ps} \end{pmatrix} \cdot \begin{pmatrix} V_{p}^{real} \\ V_{s}^{real} \\ V_{p}^{imag} \\ V_{s}^{imag} \end{pmatrix} \tag{A.33}$$

where, G^{pp}, G^{ps}, B^{pp} and B^{ps} are the real and imaginary parts, respectively, of the admittance matrix of the branch ps. A similar treatment is given to the active and reactive power injections. Voltage magnitude measurements are also transformed in rectangular coordinates, as assumed for the state variables in the AMB state estimator. To accomplish this, equation (A.34) is applied for the voltage magnitude measurements.

$$h(x) = V_{p,\phi}^{eq} = Re\{V_{p,\phi}^{eq}\} + j.Im\{V_{p,\phi}^{eq}\} = V_{p,\phi}^{med} . \frac{\dot{V}_{p,\phi}^{k}}{|\dot{V}_{p,\phi}^{k}|} \tag{A.34}$$

where, $V_{p,\phi}^{med}$ is the measured voltage magnitude at phase ϕ in bus p,and $\dot{V}_{p,\phi}^{k}$ is the calculated values for the complex nodal voltage in bus p at phase ϕ in the latest iteration of the method k. It

is noteworthy that the voltage magnitude measurement are transformed in a pair of equivalent measurements, one for the real part and another for the imaginary part (Almeida; Ochoa, 2017).

Using equation (A.35), it is possible to obtain the elements of the Jacobian matrix of the AMB estimator, given by a linear relation. For voltage magnitude measurements the derivative are unitary elements directly related to the state variables.

$$H(x) = \begin{pmatrix} \frac{\partial Re\{I_{ps}^{eq}\}}{\partial V_p^{real}} & \frac{\partial Re\{I_{ps}^{eq}\}}{\partial V_s^{real}} & \frac{\partial Re\{I_{ps}^{eq}\}}{\partial V_p^{imag}} & \frac{\partial Re\{I_{ps}^{eq}\}}{\partial V_s^{imag}} \\ \frac{\partial Im\{I_{ps}^{eq}\}}{\partial V_p^{real}} & \frac{\partial Im\{I_{ps}^{eq}\}}{\partial V_s^{real}} & \frac{\partial Im\{I_{ps}^{eq}\}}{\partial V_p^{imag}} & \frac{\partial Im\{I_{ps}^{eq}\}}{\partial V_s^{imag}} \end{pmatrix} = \begin{pmatrix} G^{pp} & G^{ps} & -B^{pp} & -B^{ps} \\ B^{pp} & B^{ps} & G^{pp} & G^{ps} \end{pmatrix} \tag{A.35}$$

The elements of the Jacobian matrix in this case are obtained from the admittance matrices of DS components, considering also the mutual coupling among different phases. Thus, the following linear measurement model can be written for the AMB estimator:

$$z(x) = H.x + \varepsilon \tag{A.36}$$

where H is the linear model composed by the system conductance and susceptance matrices that relates the equivalent current measurements with the state variables x (the complex nodal voltages in rectangular coordinates). The measurement vector z in the above formulation is given as a function of x, since they are transformed according to the calculated state variables values at each iteration.

The measurement error covariance matrix R also needs to be calculated according to the transformations made in the measurement model. It is obtained by the calculation of the equivalent measurements variance using error propagation theory (Almeida; Ochoa, 2017). With the equivalent measurements variances, the measurement error covariance matrix is updated and also the weighting matrix of the AMB estimator. For equivalent current measurements, and for equivalent voltage magnitude measurements, the equivalent variances are calculated by equations (A.37) and (A.39), respectively, using pre-specified complex voltage values for bus p and respective phases ϕ, denoted by $e_p + i.f_p$, from previous or historical estimations or from power flow studies.

$$\begin{pmatrix} \sigma^2_{I^{real}_{ps,\phi}} & \sigma^2_{I^{real}_{ps,\phi} I^{imag}_{ps,\phi}} \\ \sigma^2_{I^{real}_{ps,\phi} I^{imag}_{ps,\phi}} & \sigma^2_{I^{imag}_{ps,\phi}} \end{pmatrix} = \begin{pmatrix} c^2.\sigma^2_{P^{med}_{ps,\phi}} + d^2.\sigma^2_{Q^{med}_{ps,\phi}} & c.d.(\sigma^2_{P^{med}_{ps,\phi}} - \sigma^2_{Q^{med}_{ps,\phi}}) \\ c.d.(\sigma^2_{P^{med}_{ps,\phi}} - \sigma^2_{Q^{med}_{ps,\phi}}) & d^2.\sigma^2_{P^{med}_{ps,\phi}} + c^2.\sigma^2_{Q^{med}_{ps,\phi}} \end{pmatrix} \tag{A.37}$$

where, the diagonal elements are the equivalent current measurements variances and the off-diagonal elements are the covariances due to the transformation, and $\sigma^2_{P^{med}_{ps,\phi}}$ and $\sigma^2_{Q^{med}_{ps,\phi}}$ are the original variances of the active and reactive power flow measurements, respectively. The values of c and d are given by:

$$c = \frac{e_p}{e_p^2 + f_p^2} \qquad d = \frac{f_p}{e_p^2 + f_p^2} \tag{A.38}$$

$$\sigma^2_{V^{eq}_{p,\phi}} = \left(\frac{e_p}{\sqrt{e_p^2 + f_p^2}} . \cos(\theta_p) + \frac{f_p}{\sqrt{e_p^2 + f_p^2}} . \sin(\theta_p) \right) . \sigma^2_{V^{med}_{p,\phi}} \tag{A.39}$$

where θ_p is the phase angle of $e_p + i.f_p$.

The resulting equivalent current measurements becomes correlated in the formulation of the AMB estimator, and in some cases are neglected (Almeida; Ochoa, 2017). Besides, the equivalent variances are a function of the state variables, and in this case the values are assumed from past values since using the calculated values the Gain matrix would not be constant any more. Such approximations directly affects the probabilistic formulation of the AMB state estimator, and may hamper its consistency, as shown in (Almeida; Ochoa, 2017). This way, although presenting good practical results, the approximations done in the measurement model presents fundamental changes on the estimation problem, thus solving an relaxed version of the original problem as pointed in (Feng; Yang; Peterson, 2012).

After obtaining the linear model for the AMB state estimator, the solution of state estimation problem consists of the Gauss-Newton method for the above measurement model, where the equivalent measurements are updated iteratively. The algorithm 3 presents the main steps for the AMB estimator.

Algorithm 3: AMB STATE ESTIMATOR

Input : Network parameter, Network topology, x_0, tol

Output : Estimated state vector x

Begin
- $k \leftarrow 0$ e $x_k \leftarrow x_0$
- Build measurement model: H e R^{-1}
- Calculate Gain matrix: $G = H.R^{-1}.H$
- **Do**
  - Caculate equivalent measurements: $z(x^k)$ and $h(x^k) = H.x^k$
  - Solve linear system: $G.\Delta x = H'.R^{-1}.[z(x^k) - (h(x^k)]$
  - Update state variables: $x^{k+1} = x^k + \Delta x$
  - k = k +1
- **while** $\|\Delta x^k\|_\infty \leq tol$;

end

Branch-Current Estimator

In another perspective towards improving computational efficiency, another estimator was proposed for distribution networks in (Baran; Kelley, 1995), that results in a decoupled estimator, and the uses concepts close to the backward/forward sweep power flow methods

for radial networks. This approach uses the complex currents in each phase of the branches of the electrical network as state variables in rectangular coordinates. This estimator is known as Branch-Current (BC) state estimator. Thus, the state vector is defined as follows for a system with total number of branches n_ramos:

$$x = \begin{bmatrix} V_{ref,\phi} & I^{real}_{ps,\phi} & I^{imag}_{ps,\phi} \end{bmatrix} \qquad \begin{matrix} ps \in \{1,...,n_{ramos}\} \\ \phi \in \{1,...,n_{fases}\} \end{matrix} \tag{A.40}$$

where ps denotes the branches of the electrical network (between nodes p and s), I^{real} and I^{imag} are the real and imaginary parts of the complex current phasor at phase ϕ of this branch, and $V_{ref,\phi}$ denotes the voltage magnitude of the reference bus. The inclusion of the voltage magnitudes of the reference bus is necessary since are the only state variables that cannot be formulated as a function of the branch currents in the network. It also needs to be estimated since it affects the state of the entire distribution network (Pau; Pegoraro; Sulis, 2013). Also, similar to the previous estimator, a special treatment for the reference bus voltage phase angles is also required.

The measurement model is composed by equivalent current measurements, according to each type of measurement and relating them with the branch currents as state variables. That is, active and reactive power flows and injections measurements and voltage magnitudes are transformed in equivalent current measurements (Baran, 2012; Pau; Pegoraro; Sulis, 2013) similarly to the AMB estimator using (A.32). In case of the BC estimator, since the branch currents are used as state variables, these new equivalent measurements present a direct relation with the state variables, as shown in equation (A.41).

$$h(x) = \begin{pmatrix} Re\{I^{eq}_{ps,\phi}\} \\ Im\{I^{eq}_{ps,\phi}\} \end{pmatrix} = \begin{pmatrix} I^{real}_{ps,\phi} \\ I^{imag}_{ps,\phi} \end{pmatrix} \tag{A.41}$$

In this estimator, the Jacobian matrix H of the equivalent measurement model equation (A.41) presents unitary elements for the equivalent current measurements, improving the numerical conditioning of the state estimation problem (Baran, 2012).

$$H(x) = \begin{pmatrix} \dfrac{\partial Re\{I^{eq}_{ps,\phi}\}}{\partial I^{real}_{ps,\phi}} & \dfrac{\partial Re\{I^{eq}_{ps,\phi}\}}{\partial I^{imag}_{ps,\phi}} & \dfrac{\partial Re\{I^{eq}_{ps,\phi}\}}{\partial V_{ref,\phi}} \\ \dfrac{\partial Im\{I^{eq}_{ps,\phi}\}}{\partial I^{real}_{ps,\phi}} & \dfrac{\partial Im\{I^{eq}_{ps,\phi}\}}{\partial I^{imag}_{ps,\phi}} & \dfrac{\partial Im\{I^{eq}_{ps,\phi}\}}{\partial V_{ref,\phi}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{A.42}$$

Among the advantages of this approach of branch currents as state variables, the state estimation process can be decoupled among phases. In both AMB estimator and WLS estimator, the Jacobian matrix was obtained through based on the elements of the admittance matrices of the network components, including their mutual coupling, and whose diversity of parameters may aggravate ill conditioning in such estimators.

Regarding voltage magnitude measurements, an special treatment must be applied. In the case of the BC estimator, since the branch currents are used as state variables, the voltage magnitude model may be obtained by a method called "Path Matrix" (Baran; Jung; McDermott, 2009; Teng, 2002). This matrix includes the impedances from the network path beginning at the reference bus until the bus that presents a voltage magnitude measurement. An equation is obtained through these impedances that relates the voltage drops across the path, and consequently the branch currents.

$$|V_p| = |\dot{V}_{ref} - \sum_{ps \in \Lambda_p} Z_{ps}.\dot{I}_{ps}| \tag{A.43}$$

$$\begin{aligned} V_p^{real} &= V_{ref}^{real} - \sum_{ps \in \Lambda_p} R_{ps}.I_{ps}^{real} - X_{ps}.I_{ps}^{imag} \\ V_p^{imag} &= V_{ref}^{imag} - \sum_{ps \in \Lambda_p} X_{ps}.I_{ps}^{real} + R_{ps}.I_{ps}^{imag} \end{aligned} \tag{A.44}$$

where $\dot{V}_{ref}$ is the complex nodal voltage of the reference bus, usually the primary substation node, $Z_{ps} = R_{ps} + j.X_{ps}$ is the impedance matrix of the branch ps, Λ_p is the set of branches that belong to the path between the reference bus and the bus p where the measurement is located. The incorporation of voltage magnitude measurements turn the method coupled by phases, since it now represents the mutual coupling of the branch impedances (Baran; Jung; McDermott, 2009). To maintain a decoupled algorithm, the work in (Teng, 2002) proposes to neglect such mutual coupling for the distribution system branches, that usually is small, and thus keeping the BC estimator decoupled by phases. Besides, the voltage magnitude is a non-linear model that requires updates on the Jacobian matrix and the Gain matrix of the estimator.

Working with equation (A.43) it is possible to obtain the following expression for the voltage magnitude measurements (Pau; Pegoraro; Sulis, 2013):

$$\begin{aligned} h(x) = |V_p| = |\dot{V}_{ref}|.\cos(\delta_{V_p}^{k-1}) &- \sum_{ps \in \Lambda_p} (R_{ps}.I_{ps}^{real} - X_{ps}.I_{ps}^{imag}).\cos(\delta_{V_p}^{k-1}) \\ &- \sum_{ps \in \Lambda_p} (X_{ps}.I_{ps}^{real} - R_{ps}.I_{ps}^{imag}).\sin(\delta_{V_p}^{k-1}) \end{aligned} \tag{A.45}$$

where, $\delta_{V_p}^{k-1}$ is the voltage phase angle at bus p calculated in the previous iteration $k-1$. With this, the partial derivatives that build the Jacobian matrix for the voltage magnitude measurements can also be obtained.

$$H(x) = \begin{pmatrix} \frac{\partial |V_p|}{\partial I_{ps}^{real}} & \frac{\partial |V_p|}{\partial I_{ps}^{imag}} & \frac{\partial |V_p|}{\partial V_{ref}} \end{pmatrix} =$$

$$\begin{pmatrix} \left(-R_{ps}.\cos(\delta_{V_p}^{k-1}) - X_{ps}.\sin(\delta_{V_p}^{k-1})\right) & \left(X_{ps}.\cos(\delta_{V_p}^{k-1}) - R_{ps}.\sin(\delta_{V_p}^{k-1})\right) & \cos(\delta_{V_p}^{k-1}) \end{pmatrix} \tag{A.46}$$

This way, the following non-linear measurement model can be written for the BC estimator:

$$z(\dot{V}) = h(x) + \varepsilon \tag{A.47}$$

where $\dot{V}$ denotes the complex nodal voltages in all buses of the electric network. It is noteworthy that the above model is non-linear due to voltage magnitude measurements across the feeders, but the equivalent current magnitude measurements are still linear and represent a direct relation with the state variables in the BC estimator. To update the complex voltages in all buses of the network, a forward sweep is employed. Basically it consists of a topological sweep on the network beginning at the reference bus towards the last buses of the feeder. The sweep updates the complex nodal voltages of the next buses in the sequence according to the impedances and admittances of the network, and the estimated branch currents. Besides, the transformations in the measurements to obtain the equivalent current measurements also carries the need of updating the measurement error covariance matrix. These covariance transformations are performed similarly as previously presented for the AMB estimator.

Finally, the BC estimator also uses a Gauss-Newton method to obtain the state variables according to the measurement model (A.36), thus obtaining the branch currents in all the components of the distribution network. The algorithm 4 presents the main steps for the BC estimator.

Algorithm 4: BC STATE ESTIMATOR

Input : Network parameters, Network topology, x_0, tol
Output : Estimated state vector x

Begin
  $k \leftarrow 0$ e $x_k \leftarrow x_0$
  Build measurement model: H e R^{-1}
  Do
    Calculate equivalent measurements: $z(\dot{V}^k)$ e $h(x^k)$
    Update Gain matrix: $G = H.R^{-1}.H$
    Solve the linear system: $G.\Delta x = H'.R^{-1}.[z(\dot{V}^k) - (h(x^k)]$
    Update state variables: $x^{k+1} = x^k + \Delta x$
    Obtain the complex voltages: $\dot{V}^k$ through a Forward Sweep
    k = k+1
  while $\|\Delta x^k\|_\infty \leq tol$;
end

Orthogonal WLS Estimator

In another perspective towards improving computational efficiency, an orthogonal version of the WLS estimator is proposed in this work to deal with highly ill-conditioned distribution systems. The proposed approach uses the complex nodal voltages (in polar coordinates) in all

buses and phases of the system as state variables. Thus, for distribution network with the number of phases n_{fases} and the number of buses n_{barras}, the state vector is given by:

$$x = \begin{bmatrix} V_\phi^p & \theta_\phi^p \end{bmatrix} \qquad \begin{matrix} p \in \{1,...,n_{barras}\} \\ \phi \in \{1,...,n_{fases}\} \end{matrix} \tag{A.48}$$

where p denotes the bus from the electric network and V and θ are the voltage magnitude and voltage phase angle at each phase ϕ. This way the amount of state variables depends on the quantity of phases in each bus (single, two or three-phase, as an example).

Besides, one of the buses is taken as angular reference for the network. In general, the voltage phase angles are assumed for each phase of the reference bus, for instance as a three-phase *abc* balanced bus:

$$\theta_a = 0^o \quad \theta_b = -120^o \quad \theta_c = 120^o \tag{A.49}$$

As shown in the last section, the traditional WLS fomulation for DS state estimation is prone to result in a ill-conditioned system because said method requires the solution of the normal equation in which the factorization of the Gain matrix is necessary. In order to counter the associated ill-conditioning of DSSE using the traditional WLS, a new strategy is proposed in this paper based on the application of an orthogonal method for solving the state estimation problem. Using the QR factorization as follows:

$$W^{1/2}H(x^k) = Q^T R \tag{A.50}$$

It is possible to rewrite the normal equations as:

$$[R^T Q Q^T R]\triangle x^k = R^T Q W^{1/2}[z - h(x^k)] \tag{A.51}$$

Since $QQ^T = I$, where I is the identity matrix, we obtain:

$$[R^T R]\triangle x^k = R^T Q W^{1/2}[z - h(x^k)] \tag{A.52}$$

Multiplying both sides by $(R^T)^{-1}$:

$$R\triangle x^k = Q W^{1/2}[z - h(x^k)] \tag{A.53}$$

From (A.53), it is possible to obtain the vector $\triangle x^k$ using backwards substitution since R was obtained via QR factorization and it is, therefore, upper triangular. The resulting system is better conditioned in terms of numerical robustness than that obtained with the normal equation. The algorithm of the orthogonal method for state estimations is given below:

Algorithm 5: ORTHOGONAL WLS STATE ESTIMATOR

Input : Network parameters, network topology, x_0, tol

Output : Estimated state vector x

Begin
  $k \leftarrow 0$ e $x_k \leftarrow x_0$
  Do
    Build measurement model: $h(x^k)$ and $H(x^k)$
    Perform the Orthogonal Factorization: $QR = W^{1/2}.H(x^k)$
    Solve the triangular linear system: $R\Delta x = Q.W^{1/2}.[z - (h(x^k)]$
    Update state variables: $x^{k+1} = x^k + \Delta x$
    k = k +1
  while $\|\Delta x^k\|_\infty \leq tol$;
end

Despite the formulation of the WLS Estimator using orthogonal factorization is widely explored for transmission systems, both from research as well as industrial implementations (Pires; Costa; Mili, 1999; Simões Costa; Albuquerque; Bez, 2013), as far as the author knows this is the first work that employ such technique for distribution systems. This is treated in this thesis as a secondary contribution, but a very important one to deal with practical aspects of distribution systems. More details about the implementation of this algorithm, using sparsity treatments and with numerical stability analysis are provided in the Appendix B. Also, more details of this will be also addressed in details in the next section to support the applications of Bayesian Inference, as part of the primary contribution of this thesis.

APPENDIX B – SPARSE AND NUMERICALLY STABLE STATE ESTIMATION BASED ON ORTOGONAL METHODS

The content of this appendix have been published on the following Journal:

- G.M. Hebling, J.A.D. Massignan, J.B.A. London Junior, M.H.M Camillo ," Sparse and numerically stable implementation of a distribution system state estimation based on Multifrontal QR factorization". *Electric Power Systems Research*, v. 189, p. 106734, 2020.

Distribution System State Estimation

Despite the consolidated position of the Weighted Least Squares (WLS) state estimator for transmission systems, specialized algorithms have been developed to perform Distribution System State Estimation (DSSE). In the following sections we present the classic WLS formulation as well as its problems when applied to distribution systems as well as alternative solutions found in the literature.

The distribution system state estimation is based on the nonlinear model of measurements (Abur; Gómez-Expósito, 2004; Monticelli, 1999):

$$z = h(x) + e, \tag{B.1}$$

where z is the measurement vector ($mx1$), x is the state variables vector ($nx1$), $h(.)$ is the nonlinear state estimation function ($mx1$) that relates the measurements to the states variables, and e is the vector of measurement errors ($mx1$) usually considered as independent random Gaussian variables with zero mean and diagonal covariance matrix R_z ($R_z = diag\ \{\sigma_{z1}^2, \sigma_{z2}^2, ..., \sigma_{zm}^2\}$, where σ_{zi} is the standard deviation of measurement z_i).

Through the classical Weighted Least-Squares (WLS) estimator, the state estimate vector $\hat{x}$ is obtained by minimizing the index $J(x)$ given by:

$$J(x) = [z - h(x)]^T W [z - h(x)], \tag{B.2}$$

where $W = R_z^{-1}$.

The minimum value for J(x) is obtained when its derivative, $\frac{\partial J(x)}{\partial x} = 0$, is equal to zero which can be written as:

$$H^T(\hat{x}) W [z - h(\hat{x})] = 0 \tag{B.3}$$

The matrix $H(\hat{x})$ is known as the Jacobian matrix and it is the matrix of first derivatives of $h(x)$ evaluated at the $x = \hat{x}$ point. Given the nonlinear nature of the index $J(x)$, an iterative algorithm

is used to obtain a solution to a linear equation, which calculates a correction, $\triangle x^k$, to the state vector where k is the iteration index. The linear approximation at the point x^k is:

$$h(x^{k+1}) \approx h(x^k) + H(x^k)\triangle x^k \tag{B.4}$$

Rewriting (B.1) considering the linear approximations presented beforehand we obtain:

$$z = h(x^k) + H(x^k)\triangle x^k + e \tag{B.5}$$

Which is equivalent to:

$$\triangle z(x^k) = z - h(x^k) = H(x^k)\triangle x^k + e \tag{B.6}$$

From the linear approximations, it is possible to write a new objective function $J(\triangle x)$:

$$J(\triangle x) = [\triangle z(x^k) - H(x^k)\triangle x^k]^T W[\triangle z(x^k) - H(x^k)\triangle x^k] \tag{B.7}$$

The minimum of $J(\triangle x)$ is obtained when:

$$\frac{\partial J(\triangle x)}{\partial x} = H(x^k)^T W[\triangle z(x^k) - H(x^k)\triangle x^k] = 0 \tag{B.8}$$

The correction $\triangle x^k$ is then given by the solution of:

$$\triangle x^k = [H(x^k)^T W H(x^k)]^{-1} H(x^k)^T W \triangle z(x^k) \tag{B.9}$$

Equation (B.9) is known as the *Normal Equation* and to obtain its solution it is necessary to invert or factorize $H(x^k)^T W H(x^k)$, which is called the Gain matrix G. Since it is computationally inefficient to invert the Gain matrix, factorization methods are used to obtain the solution such as Gauss Elimination, Cholesky Factorization, LU Decomposition and others. The numerical robustness of the solution of (B.9) depends not only on the chosen method for factorization but also on the condition number of the Gain matrix (Ebrahimian; Baldick, 2001).

It is possible to show that the condition number of the Gain matrix $K(G)$ (defined as the ratio between the largest and smallest of its eigenvalues $\lambda(G)$) is close to the magnitude of the square of the condition number of the weighted Jacobian matrix, as follows:

$$K(G) = \frac{\lambda_{max}(G)}{\lambda_{min}(G)} \approx [K(W^{\frac{1}{2}}H)]^2 \tag{B.10}$$

DSs have a set of particularities that make the resulting system of the WLS even more prone to ill-conditioning. These particularities are related to the topology of the network, the number of buses and availability of measurements. A DS is usually radial, with single-phase as well as two and three-phase unbalanced and short branches, containing different transformers connections and has a large number of nodes. The number of virtual and pseudo measurements

with different precision values from the real-time measurements also makes the weighting process a source of ill-conditioning.

These aforementioned characteristics have prevented the use of the WLS state estimator based on the Normal Equation method for DSSE (Lefebvre; Prévost; Lenoir, 2014), (Baran, 2012). Instead, some alternative formulations have been proposed in the literature such as the Branch Current approach in (Baran; Kelley, 1994) and the Admittance-matrix approach in (Lu; Teng; Liu, 1995). These alternative formulations are based on simplifications either in the state variables or in the measurement set that may not be as accurate when applied in modern DSs.

The proposed Sparse and Numerically Stable Algorithm for DSSE

In order to counter the associated ill-conditioning of DSSE using the traditional WLS, a new strategy is proposed in this paper based on the application of an orthogonal method for solving the state estimation problem. Using the QR factorization as follows:

$$W^{1/2}H(x^k) = Q^T R \tag{B.11}$$

It is possible to rewrite the normal equation as:

$$[R^T Q Q^T R]\triangle x^k = R^T Q W^{1/2}[z - h(x^k)] \tag{B.12}$$

Since $QQ^T = I$, where I is the identity matrix, we obtain:

$$[R^T R]\triangle x^k = R^T Q W^{1/2}[z - h(x^k)] \tag{B.13}$$

Multiplying both sides by $(R^T)^{-1}$:

$$R\triangle x^k = Q W^{1/2}[z - h(x^k)] \tag{B.14}$$

From (B.14), it is possible to obtain the vector $\triangle x^k$ using backwards substitution since R was obtained via QR factorization and it is, therefore, upper triangular. The resulting system is better conditioned in terms of numerical robustness than that obtained with the normal equation. Consider the condition number of the upper triangular matrix R:

$$K(R) = \|R\|\|R\|^{-1} \tag{B.15}$$

$$K(R) = \|[(Q^T)^{-1}W^{1/2}H]\|\|[(Q^T)^{-1}W^{1/2}H]^{-1}\| \tag{B.16}$$

$$K(R) = \|[(Q^T)^{-1}W^{1/2}H]\|\|[(W^{1/2}H)^{-1}Q^T]\| \tag{B.17}$$

$$K(R) \approx \|(Q^T)^{-1}\|\|(W^{1/2}H)\|\|(W^{1/2}H)^{-1}\|\|Q^T\| \tag{B.18}$$

Since Q is orthogonal, the additional following properties are true: $Q^T = Q^{-1}$ and $\|Q\| = \|(Q^T)^{-1}\| = 1$. Therefore we can approximate the condition number of R as:

$$K(R) \approx \|(W^{1/2}H)\|\|(W^{1/2}H)^{-1}\| = K(W^{1/2}H) \tag{B.19}$$

Equation (B.19) shows that the resulting system, when applying the QR factorization to the WLS formulation, has a condition number of the same order of magnitude as the Jacobian matrix whereas the system obtained with the Normal Equation has a condition number approximately the square of that of the Jacobian matrix. In addition to being a better conditioned system, when using the QR factorization another benefit lies in the stability of the method when dealing with floating point arithmetic.

In order to improve readability, we define the matrix A as:

$$A = W^{1/2}H \tag{B.20}$$

The corresponding QR factorization satisfies:

$$\tilde{Q}\tilde{R} = A + \delta A \tag{B.21}$$

Where $\frac{\|\delta A\|}{\|A\|} = O(\varepsilon_{machine})$. The forward step is stable (**??**) and proving the backwards stability we obtain:

$$(A + \triangle A)\tilde{x} = b \tag{B.22}$$

where $\frac{\|\triangle A\|}{\|A\|} = O(\varepsilon_{machine})$. The first step in the proof is to write $A + \triangle A$ as:

$$(\tilde{Q} + \delta Q)(\tilde{R} + \delta R)\tilde{x} = b \tag{B.23}$$

$$[\tilde{Q}\tilde{R} + \tilde{Q}(\delta R) + \tilde{R}(\delta Q) + (\delta Q)(\delta R)]\tilde{x} = b \tag{B.24}$$

Since $\tilde{Q}\tilde{R} = A + \delta A$, and δA is small when compared to A, we can write:

$$\frac{\|\tilde{R}\|}{\|A\|} \leq \|\tilde{Q}^*\|\frac{\|A + \delta A\|}{\|A\|} = O(1) \tag{B.25}$$

Using equation (B.25) and adding the terms of equation (B.24) we can show that:

$$\begin{aligned}\frac{\|\triangle A\|}{\|A\|} \leq \frac{\|\delta A\|}{\|A\|} + \frac{\|(\delta Q)\tilde{R}\|}{\|A\|} + \\ + \frac{\|(\delta R)\tilde{Q}\|}{\|A\|} + \frac{\|(\delta Q)(\delta R)\|}{\|A\|} = O(\varepsilon_{machine})\end{aligned} \tag{B.26}$$

One consequence of the backwards stability of the QR algorithm is that the obtained solution $\tilde{x}$ follows:

$$\frac{\|\tilde{x} - x\|}{\|x\|} = O(K(A)\varepsilon_{machine}), \tag{B.27}$$

where $K(A)$ is the condition number of the matrix A (**??**). Overall, the proposed algorithm that uses the QR factorization of the matrix $W^{1/2}H$ results in a smaller condition number than that of the correspondent Gain matrix. This fact also has influence on the backwards stability of the method because of equation (B.27). Therefore, the proposed algorithm in addition to being better conditioned is also stable.

Bad Data Processing with Orthogonal Formulation for DSSE

One of the most essential features of state estimation is the possibility to detect and identify bad data (or gross errors), besides filtering white noise from the measurement process. The effects of bad data in the estimation process can be very severe, reducing accuracy and even compromising functionalities that rely on the real-time state of the network. The bad data processing step is performed after the estimated state $\hat{x}$ is obtained. It is necessary to calculate the normalised residue and, for that, the residue's covariance matrix Ω is used as follows:

$$\Omega = W^{-1} - H(\hat{x})(H(\hat{x})^T W H(\hat{x}))^{-1} H(\hat{x})^T \tag{B.28}$$

Using the orthogonal relation defined in (B.11), it is possible to rewrite the residue's covariance matrix as:

$$\Omega = W^{-1} - H(\hat{x})((R^T Q)(Q^T R))^{-1} H(\hat{x})^T \tag{B.29}$$

Since Q is an orthogonal matrix and using inverse matrices properties, it follows that:

$$\Omega = W^{-1} - H(\hat{x})(R^{-1})(R^T)^{-1} H(\hat{x})^T \tag{B.30}$$

With this formulation, the inverse of the Gain matrix does not need to be calculated and the matrix Ω is obtained via a series of *Forward* and *Backward* substitutions. In order to increase numerical stability, the proposed method in this work also avoids the multiplications by Jacobian and its transpose. This is achieve by introducing an auxiliary matrix $Y(mxn)$ so that:

$$R^T Y = H^T(\hat{x}) \tag{B.31}$$

To increase computational efficiency, this work also exploits the fact that only the diagonal elements are employed in the gross errors processing. Each diagonal element of Ω is then obtained using Y and the following scalar product:

$$\Omega_{ii} = W_{ii}^{-1} - \langle Y_{:i}^T, Y_{:i} \rangle = W_{ii}^{-1} - \|Y_{:i}\|_2 \tag{B.32}$$

With the diagonal elements of the matrix Ω, three different indexes are calculated to guide the bad data detection and identification. First, the normalised residue is calculated for each measurement:

$$r_i^N = \frac{|z_i - h_i(\hat{x})|}{\sqrt{\Omega_{ii}}}. \tag{B.33}$$

With the normalised residue, it's possible to calculate the index of the b-hat method, following:

$$\hat{b}_i = \frac{\sigma_{ii}}{\sqrt{\Omega_{ii}}} r_i^N, \tag{B.34}$$

where σ_{ii} is the standard deviation of the $i-th$ measurement, obtained from the diagonal of the matrix W.

The maximum normalized residue test enables the detection and identification of a measurement with GE and the identification of critical measurements, that is, if the removal of said measurement renders the system not observable. The b-hat test enables the correction of a measurement affected by GE.

In order to complement the analysis of GEs, (Massignan; London Jr; Miranda, 2020) proposes the usage of the Undetectability Index (UI) as a a third index, which was first proposed in (Benedito et al., 2014). This index enables mapping measurements in terms of the difficulty of detection of GEs and in this work it is also obtained using the previously shown orthogonal method, following equation B.35. The matrix S is called Residue Sensitivity Matrix.

$$UI_i = \frac{\sqrt{1-\Omega_{ii}W}}{\sqrt{\Omega_{ii}W}} = \frac{\sqrt{1-S_{ii}}}{\sqrt{S_{ii}}} \tag{B.35}$$

Sparse QR Multifrontal Implementation

Householder Reflections (Householder, 1958) and Givens Rotations (Givens, 1958) are the methods usually applied to obtain a QR factorization. In order to extract peak computational performance when factorizing large sparse matrices other methods have been proposed such as the Multifrontal QR (Davis, 2011b). Traditional orthogonal methods for QR factorization of sparse matrices operate on one row or column at a time (George; Heath, 1980; GEORGE; LIU; NG, 1988) and are not able to achieve said maximum performance because of irregular access of memory. The Multifrontal QR performs on groups of rows and columns in a sequence of dense matrices called frontal matrices that can be handled in parallel.

There are three phases in solving the resulting linear system in (B.14) when using a direct method such as the Multifrontal QR. First, a fill-in reducing permutation is obtained using some form of heuristics which will be discussed further in the next section. Then a symbolic analysis is performed in order to obtain the non-zero pattern of the matrix by performing a symbolic Cholesky factorization of $A^T A$. Next, there are the numerical factorization phase and the proper solving phase that uses backward substitution.

Figure 56 represents the underlying structure of the algorithm in which the matrix is reordered. Its elimination tree is obtained via the symbolic factorization, and the supernodes are found in order to form the dense frontal matrices that will be subject to the numerical factorization using the Householder QR method , whereas other proposals use Givens Rotations to obtain

a orthogonal factorization (Costa; Salgado; Haas, 2007). The figure presents the elimination tree with supernodes as well as the corresponding R factor.This algorithm is available as an open-source package named *SuiteSparse*.

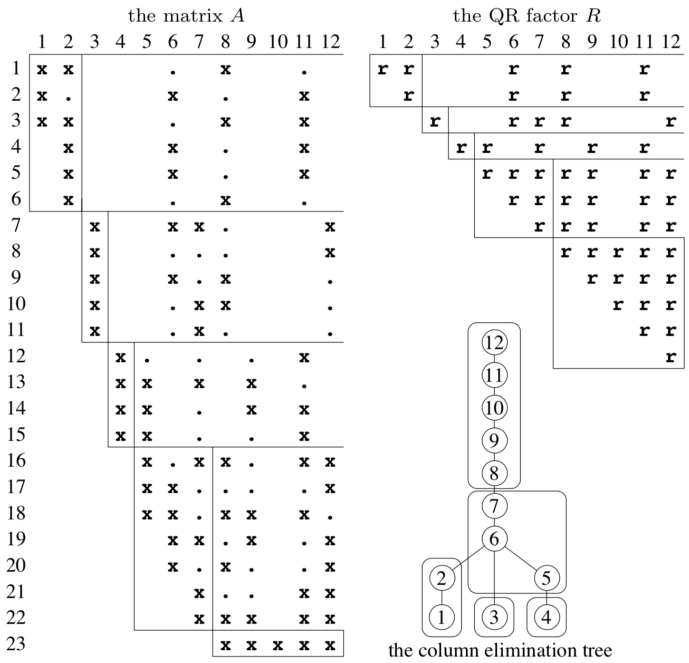


Figure 56 – Matrix A, its QR factor R and the corresponding elimination tree. Source: (Davis, 2011b)

The Thesis opted to use a Multifrontal QR method based on block Houselholder reflections, due to its good performance under very large scale linear systems. Alternative options exist, such as the Givens rotations, already successfully employed at transmission systems.

Efficient Ordering for Fill-In Reduction

A sparse matrix is a matrix with a high number of zero elements. In order to achieve higher computational efficiency, only the non-zero elements can be stored and specialized algorithms can be used to read and/or write data. Preserving the sparse characteristic is, therefore, beneficial in terms of memory usage and computational effort specially in application where real-time processing is expected, such as the State Estimator. The results in (Hebling et al., 2020) show that expressive performance gains can be expected when using dedicated sparse storage and algorithms.

Before the factorization step of the SE, an ordering step is performed so that the number of non-zero elements created, so called fill-in, is reduced. The ordering problem is formulated within an optimization framework in which the objective is to find a permutation matrix P so that the number of non-zero elements in PAP^T is minimized (Davis, 2011b; Amestoy; Davis;

DUFF, 2004). This problem is *NP-Complete* as shown in (Yannakakis, 1981) and the solution may be obtained by heuristics, as will be shown in a next section.

It is important to note, however, that in order to extract peak computational performance specialized algorithms must be used in addition to sparsity preserving ordering. These algorithms have the necessary sparse data structures and routines that take advantage of the high number of zero elements in the matrices. In this work, an algorithm called Multifrontal QR, based on block Householder reflections (Davis, 2011a), which is specialized in sparse matrices, is used in order to obtain the previously shown QR factorization of the matrix $W^{1/2}H$.

Some of the heuristics used to obtain the permutation matrix P are based on the undirected graph associated with the matrix being factorized since a factorization step is equivalent to a node elimination. Fill-ins are, therefore, new edges created in this graph after an elimination. This can be better visualized by figure 57 where, after node 2 is eliminated, three new edges, in red, are created.

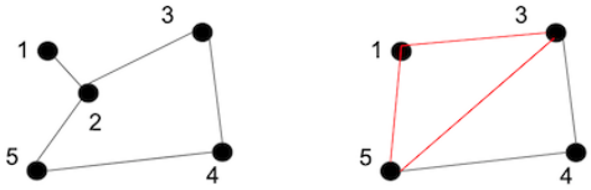


Figure 57 – Eliminating node 2 results in fill-in, shown in red.

One of these heuristics is called Approximate Minimum Degree. It chooses at each elimination step, the node with the minimum degree, that is, the least amount of adjacent nodes and builds the ordering matrix P. There can be ties and (Vempati; Slutsker; Tinney, 1991) proposes different criteria to choose the next node. In figure 57, the eliminated node 2 has a degree of three, if a node with a lower degree had been chosen, such as node 1, fewer elements would be created.

Another method that operates on the associated graph of the matrix being factorized is called Nested Dissection. This heuristic is based on the Divide and Conquer concept and, at each successive iteration, it chooses a separator so that each path between the resulting sub-graphs passes through this separator. After a sufficient number of iterations, any nodes that are not in a separator, are chosen to be first in order. The nodes in the first separators are set as last in the ordering.

The proposal in (Chan; George, 1980) features an implementation that aims to decrease the diagonal band of the matrix operating in the associated graph. It chooses, at first, the node with the minimum degree and orders the adjacent nodes' degree in ascending order. This process is repeated until there are no nodes left. (Liu; Sherman, 1976) noted that using the reverse order obtained with this process produced less fill-in naming, therefore, the method as *Reverse Cuthill-Mckee - RCM*.

The strategies to reduce the number of fill-ins are usually related to the undirected graph of the matrix being factorized because the factorization process is equivalent to a node elimination in the graph. Fill-in means, therefore, new edges created in the graph after the elimination process of a node and one strategy to reduce this number is to select nodes with the minimum degree, that is, the number of connections to be eliminated. Note that this number is updated after each elimination step because new edges are created. The resulting permutation matrix is that which orders the matrix being factorized in the sequence that the minimum degree nodes will be eliminated first.

APPENDIX C – LINEAR BAYESIAN INFORMATION FUSION

In this Appendix, a Linear Bayesian Information Fusion method is devised in the context of power system state estimation, to deal with with the diversity of measurements respecting their different sampling and updating rates. Bayesian Inference concepts guide the integration of such different measurements as detailed in the following. Initially, the proposition is devised for integration of SCADA and PMU measurements in two separate estimators, according to their sampling rates. In Chapter 6, the models are further detailed and generalized for multiple measurement sets and non-linear estimation.

Regarding only SCADA and PMU measurements, the method consists in two stages, the first to process SCADA measurements according to their updating rates (slower measurements), and the second for the PMUs (faster measurements). The first stage of the proposed Bayesian State Estimation treats the SCADA measurements by the conventional nonlinear measurement model described in (3.5), triggered accordingly to SCADA updates. The solution of this first SCADA stage consists of the traditional WLS state estimator presented in details in Appendix A, yielding an initial SCADA estimated state $\hat{\mathbf{x}}_0$ and its respective state covariance matrix $\mathbf{R}_{\hat{x}_0}$. Figure 58 illustrates the probabilistic model within the two stages of the method. It also illustrates the difference from the Kalman Filter.

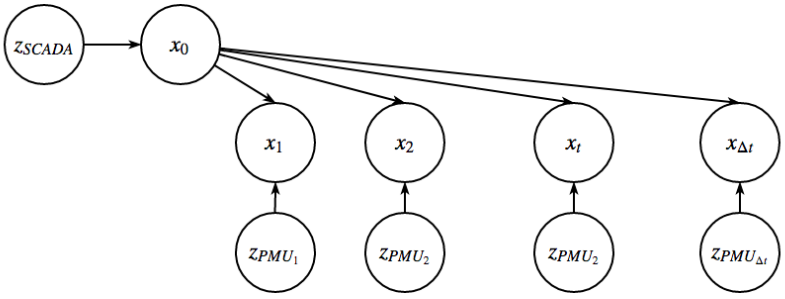
The Kalman filter is the recursive version of the above state space model, where the last posterior estimation updates the following prior belief (Barker; Brown; Martin, 1995) (x_0 as the prior for x_1, x_1 for x_2, and so on). The intention for breaking this recursive chain is to obtain a state estimator more sensitive to systems changes (non-stationary conditions). This difference shows how the Bayesian framework encompasses a probabilistic theory able to produce different state estimation strategies.

The second stage consists of employing the probabilistic model described above, which is inspired in the first-order state-space but breaking the recursive chain, in a stage composed only by PMUs as measured information. For metering systems containing only PMUs, using rectangular coordinates for the phasor measurements and state variables, the state estimation model (3.5) becomes linear (C.1).

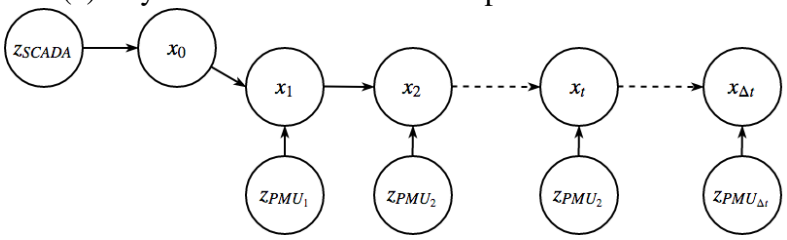
$$z = Hx + e_{PMU} \tag{C.1}$$

where H is a $(m_{PMU}$ x $n)$ matrix that relates the m_{PMU} PMU measurements with the state variables; e_{PMU} is the $(m_{PMU}$ x $1)$ noise vector also assumed as independent normally distributed random variables, with zero mean and known covariance matrix R_{PMU}.

By assuming the state follows a quasi-stationary process and that the last performed estimation provides the previous state information, we can write (C.2) and (C.3) for the PMUs



(a) Bayesian Information Fusion probabilistic model



(b) First-order state-space model for the Kalman Filter

Figure 58 – Probabilistic model of the proposed Bayesian Information Fusion in the context of power system state estimation. The first figure (a), the first stage is the static state estimation using SCADA measurements. The second stage consists in updating the state using the PMU observations, as they become available, by the posterior distribution obtained by Bayesian Inference principles. The second figure (b) illustrates the Kalman Filter perspective, as a recursive version of the proposed information fusion model.

as:

$$x = \hat{x}_{SCADA} + \omega \tag{C.2}$$

$$z_{PMU} = Hx + e_{PMU} \tag{C.3}$$

where, H is a $(m_{PMU}$ x $n)$ matrix that relates the m_{PMU} PMU measurements with the state variables; e_{PMU} is the $(m_{PMU}$ x $1)$ noise vector also assumed as independent normally distributed random variables, with zero mean and known covariance matrix R_{PMU}. These assumptions are relatively fair since state estimation is a steady state analysis and the aim of this study is to treat sampling rates of PMUs, much faster than SCADA measurements. Also by assuming the previously estimated state gives the forecasting uncertainty, the covariance of the forecasting error can be written as $Q_{SCADA} = R_{\hat{x}_{SCADA}} = (H_{SCADA}(\hat{x}_{SCADA})^T R_{SCADA} H_{SCADA}(\hat{x}_{SCADA}))^{-1}$ (the inverse of the SCADA gain matrix). It is noteworthy that to fully integrate the prior knowledge from a previous estimator, the state variables and state covariance matrix must also be transformed (or calculated) in rectangular coordinates, in this case to keep a linear model for the PMUs.

By using the probabilistic model in (C.2), x_t is modelled as a random variable, since it is associated with the previous state by a forecasting model with random error. Therefore, we can directly apply the concepts used in Bayesian Inference, where the estimation problem treats both measurements and states as random variables (Sivia; Skilling, 2006). Using the Bayes' Theorem, the posterior probability function for the state variables.

The estimation problem is then solved using the posterior probability distribution rather than only the measurement model. By assuming a Gaussian prior to the state, as assumed for the forecasting error, we can obtain the posterior probability:

$$Prior: \qquad x \sim \mathcal{N}(\hat{x}_{SCADA}, Q_{SCADA}) \tag{C.4}$$

$$Likelihood: \qquad z_{PMU} \mid x \sim \mathcal{N}(H\mathbf{x}, \mathbf{R}_{PMU}) \tag{C.5}$$

$$Posterior: \tag{C.6}$$
$$x \mid z_{PMU} \sim \mathcal{N}((H^T \mathbf{R}_{PMU}^{-1} H + Q_{SCADA}^{-1})^{-1}(H^T R_{PMU}^{-1} z_t + Q_{SCADA}^{-1} F \hat{x}_{SCADA}), (H^T R_{PMU}^{-1} H + Q_{SCADA}^{-1})^{-1})$$

The proposed Bayesian Information Fusion consists of setting the prior distribution as the estimation provided by the SCADA measurements, and keep it fixed while new PMU samples arrive. That is, both $\hat{x}_{SCADA}$ and $R_{\hat{x}_{SCADA}}$ fixed as the initially obtained by the SCADA measurements (with the state variables in rectangular coordinates).

In the case of a conjugate multivariate Gaussian prior model, the MAP estimate is the expected value of the posterior multivariate normal distribution shown in (C.6). Therefore, solving the following linear system results on the estimated state:

$$(H^T R_{PMU}^{-1} H + Q_{SCADA}^{-1})\hat{x} = (H^T R_{PMU}^{-1} z_{PMU} + Q_{SCADA}^{-1} \hat{x}_{SCADA}) \tag{C.7}$$

Demonstration in the Linear Case

By using the prior distribution (C.4) given by the state space model, and the likelihood (C.5) by the PMU linear measurement model, we can directly apply Bayes Theorem, which yields:

$$\begin{aligned} f_{X|Z}(x|z_{PMU}) \propto \exp\left(-\frac{1}{2}(z_{PMU} - Hx)' R_{PMU}^{-1}(z_{PMU} - Hx)\right) \\ \exp\left(-\frac{1}{2}(x - \hat{x}_{SCADA})' Q_{SCADA}^{-1}(x - \hat{x}_{SCADA})\right) \end{aligned} \tag{C.8}$$

Expanding the exponent:

$$\begin{aligned} z_{PMU}^T R_{PMU}^{-1} z_{PMU} + x^T H^T R_{PMU}^{-1} H x - x^T H^T R_{PMU}^{-1} z_{PMU} - z_{PMU}^T R_{PMU}^{-1} H x \\ + x^T Q_{SCADA}^{-1} x + \hat{x}_{SCADA}^T Q_{SCADA}^{-1} \hat{x}_{SCADA} - \hat{x}_{SCADA}^T Q_{SCADA}^{-1} x - x^T Q_{SCADA}^{-1} \hat{x}_{SCADA} \end{aligned} \tag{C.9}$$

In order to find the posterior probability distribution, we can search for its kernel to simplify the calculations. Thus we can neglect the constant values of the above equation, yielding:

$$\begin{aligned} x^T \left(H^T R_{PMU}^{-1} H + Q_{SCADA}^{-1}\right) x - x^T \left(H^T R_{PMU}^{-1} z_{PMU} + Q_{SCADA}^{-1} \hat{x}_{SCADA}\right) \\ - \left(z_{PMU}^T R_{PMU}^{-1} H + \hat{x}_{SCADA}^T Q_{SCADA}^{-1}\right) x \end{aligned} \tag{C.10}$$

The kernel of a Multivariate Normal distribution with expected value a and covariance matrix B has the following exponent structure:

$$(x-a)^T B(x-a) = x^T Bx - x^T Ba - a^T Bx - a^T Ba \tag{C.11}$$

Comparing (C.10) and (C.11) the expected value and covariance matrix of the posterior distribution kernel are given as:

$$B \equiv \left(H^T R_{PMU}^{-1} H + Q_{SCADA}^{-1}\right) \tag{C.12}$$

$$Ba \equiv \left(H^T R_{PMU}^{-1} z_{PMU} + Q_{SCADA}^{-1} \hat{x}_{SCADA}\right) \tag{C.13}$$

$$a \equiv \left(H^T R_{PMU}^{-1} H + Q_{SCADA}^{-1}\right)^{-1} \left(H^T R_{PMU}^{-1} z_{PMU} + Q_{SCADA}^{-1} \hat{x}_{SCADA}\right) \tag{C.14}$$

A first remark regards to the observability of the MAP estimation procedure. Note that the matrix of this linear system is a positive definite and full rank matrix since the previous observable SCADA stage Gain matrix provides Q_{SCADA}^{-1}. Therefore, the unique requirement for the method to provide an estimation is a SCADA observable system. Also a full covariance matrix is employed instead of diagonal, such as in the case of Thikonov regression. The correlation among the state variables arises from the electrical connections of the network, and the SCADA observable metering system captures them, embedded on the gain matrix (equal the inverse of the state covariance matrix). Once the SCADA metering system is observable, the gain matrix of the SCADA stage maps all the state correlations. By updating the posterior distribution with the few PMUs observations and the prior distribution from the SCADA, the MAP stage in (C.7) automatically updates all the variables.

A second important remark is the validity of the Gaussian assumption for the prior distribution. Indeed such assumption is an approximation of the true underlying state probability distribution function. Another practical goal is to maintain computational tractability, since the context of state estimation is a real-time tool, what makes impractical to employ nonparametric sampling methods with currently available computational methods. Such computational enhancements are a very exciting technological research direction, especially exploiting graphical processing units and novel numerical methods.

The above derivations was performed for linear measurement model, where the Conjugate-Prior yields a Gaussian posterior distribution. It is noteworthy that in case of nonlinear the posterior will not be Gaussian, and it may not possible to be obtained analytically. However, in the power system state estimation problem, only the MAP estimate is enough to evaluate system condition in real-time, which can be obtained without the requirement of computing the full posterior conditional, and employing nonlinear unconstrained optimization methods. The solution of the nonlinear version is presented in details in Chapter 6, under a concept of sampling layers to accommodate multiple types of sensors according to their sampling characteristics, that is, more types of sensors rather than only SCADA and PMUs.

Didatic Example

To validate the proposed Bayesian Information Fusion method, we performed simulations with the small distribution system IEEE 4 nodes test feeder using Monte Carlo simulation. A load flow calculation was used to create the reference values for the state variables (x^{lf}) and measurements (z^{lf}). The Monte Carlo simulation consists in including random noise in the reference load flow values to obtain the measured values used to perform state estimation (**??**). The i-th measurement value was calculated by adding a random noise with Normal distribution u_i $\mathcal{N}(0, \sigma_i)$ in the corresponding i-th measurement reference value, according to (C.15).

$$z_i = z_i^{lf} + u_i \tag{C.15}$$

The noise standard deviation σ_i was obtained using different precision levels for each measurement according to (C.16) (Singh; Pal; Jabr, 2009):

$$\sigma_i = \frac{\left| z_i^{lf} \right| pr_i}{3} \tag{C.16}$$

where, pr_i is the precision of the i-th meter (in this study, was assumed 2% for active and reactive power measurements from SCADA, 1% for voltage magnitude from SCADA, and 0.1% for voltage and current phasors measurements from PMUs).

This equation is also used to create the measurement covariance matrix used in the estimation process (R matrix for SCADA measurements and R_{PMU} for PMUs). However, in this case using the measured values instead of the measurement reference values obtained in the load flow calculation.

To simulate different sampling rates, a first sample from SCADA measurements was considered in the state estimation process, followed by 60 samples from PMUs. Each sample represents a different instant t and the respective measured values. For each instant, a load flow calculation was performed and the measured values are obtained according to the previously explained procedure. Thus, we obtained a sequence of quasi-stationary conditions being monitored by measurements with random noise in our simulations. When using such simulation strategy, based on load flow calculation, the system's dynamics is neglected, such as oscillations during transitory events (Zhao et al., 2016).

To evaluate the accuracy of the estimated state variables, we used the Mean Absolute Error (MAE) value:

$$\text{MAE}_t = \frac{1}{n_{trials}} \sum_{1}^{n_{trials}} \left| \hat{x}_t - x_t^{lf} \right|, \tag{C.17}$$

where $\hat{x}_t$ is the estimated state, x_t^{lf} is the reference state value in t, and n_{trials} is the number of trials during the simulation.

Performance with SCADA and PMU Observable Metering Systems and Fixed Loading Condition

The first test was performed using SCADA and PMU observable metering systems when the hypothesis of stationary condition is respected. Fig. 59 illustrates the IEEE 4 nodes feeder with the observable SCADA and PMU metering systems. The unbalanced load scenario with the Gr. Wye- Gr. Wye transformer is used in our simulation. The load, originally in node 4, was distributed among nodes 2, 3 and 4 to further represent the PMU non-observable scenario.

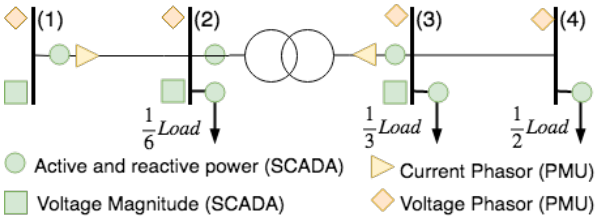


Figure 59 – IEEE 4 node test feeder used to evaluate the proposed Bayesian inference method for including PMUs in DS state estimation.

The simulations were performed with fixed nominal loading to represent a stationary condition and to demonstrate the effect of the PMU Posterior stage. Two different instants are considered: an initial instant with the acquisition of SCADA measurements; and a second instant with the arrival of a PMU sample. The simulations were executed with 300 repetitions, resulting in a calculation of MAE with $n_{trials} = 300$.

To illustrate the effect of the MAP estimation, Fig. 60 presents the error histogram for this simulation. As it can be seen, besides the reduction on the MAE indicator, a smaller error Variance is also obtained after the execution of MAP basing on the PMU sample. This because the higher precision assumed for such measurements are naturally considered in the proposed approach through the PMU measurements covariance matrix.

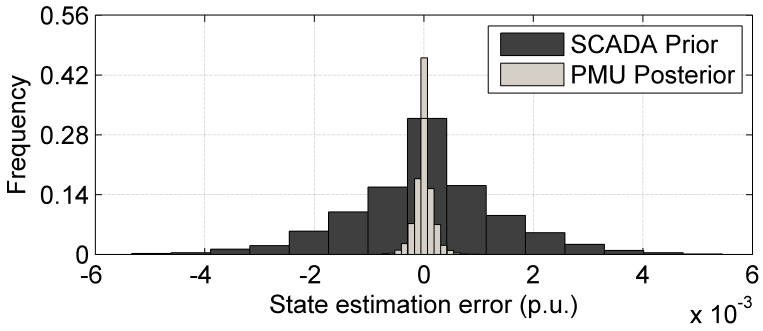


Figure 60 – Histogram of state estimation error among all state variables for the SCADA Prior and PMU Posterior.

Performance with SCADA and PMU Observable Metering Systems and Load Variation

A second test was performed with the inclusion of a 5% of load variation around the nominal loading. The IEEE 4 nodes feeder with the observable SCADA and PMU observable metering systems in Fig. 59 is used again. In this simulation, we considered all the 60 PMU samples. Despite the load variation, this situation is still considered as stationary since the load variation is within a 5% of its nominal value, keeping a constant mean value and variance along the time interval of the PMU samples. This scenario intends to show a more realistic load condition, since there are no guarantees that the loads will keep a fixed value while the PMU samples are arriving.

The proposed Bayesian Information Fusion was then compared with a hybrid static state estimator, which simultaneously processes both SCADA and PMU measurements in a single stage according to the nonlinear measurement model, and with a Kalman Filter tracking state estimator that uses recursively the estimated posterior to update the prior belief. Figure 61 presents the comparison using the MAE performance metric in both the fixed loading and the 5% of load variation scenarios.

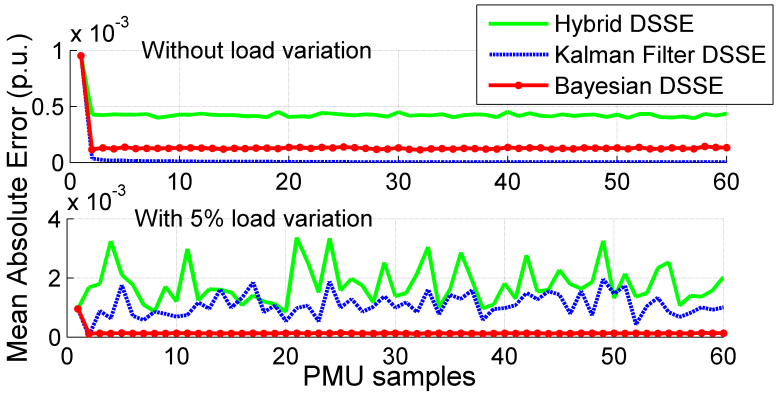


Figure 61 – Performance of state estiamtors in terms of the Mean Absolute Error in p.u. for each time sample (MAE_t) in the stationary scenarios.

The hybrid approach presented the worst precision, since it assumes both SCADA and PMUs samples represent the same instant. The KF approach presented a very good estimation for the fixed load scenario, but lost its accuracy for the load variation scenario. This result is expected since KF has good smoothing properties, and the fixed load value is the smoothest scenario possible. But when load variation was included this premise became false. Furthermore, the proposed Bayesian Information Fusion presented a similar precision in both scenarios, showing good results in both cases. This is related to the fact that the prior distribution only gives an initial estimate of the state vector and its covariance. Whenever a new PMU sample is gathered, the MAP estimation uses only the values of that sample to update the state vector.

Performance with SCADA observable and PMU Non-Observable Metering Systems and Non-Stationary Condition

In this third test, the observable SCADA metering system illustrated in Fig. 59 was considered. However, in terms of PMU metering system, only the PMU installed at the secondary winding of the transformer was considered (node 3). Therefore, the system is still SCADA observable, but no longer PMU observable, i.e., if only PMUs were considered it would not be possible to perform state estimation for the whole network. In this test, a large load variation at node 4 (60% increase), only in phase A, was included in the simulation from $t = 15$ to $t = 20$, in order to simulate a non-stationary event. In the other instants, the loads were kept within their nominal unbalanced loading with a 5% of variation. For instance, that large load variation could occur when a large load is connected, or due to intermittent distributed generation in the DS.

Fig. 62 presents the MAE performance index in such scenario. Both Hybrid and KF approaches had their performance significantly deteriorated during the load temporary event. While the proposed approach had a much smaller influence in its estimation accuracy. To illustrate the state estimation results, Fig. 63 presents the final estimate obtained by each tested DSSE for the voltage magnitude at node 4, phase A, in one of the Monte Carlo trials and the reference value.

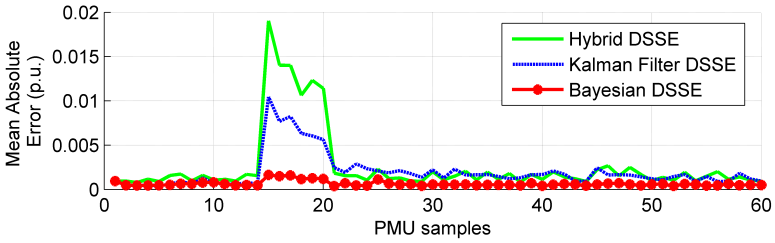


Figure 62 – Performance of state estimators in terms of the Mean Absolute Error in p.u. for each time sample (MAE_t) in the non-stationary scenario.

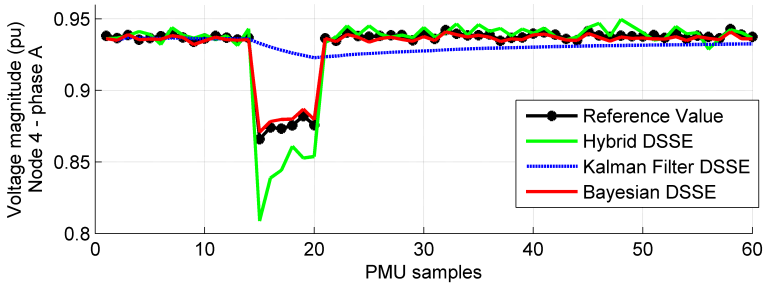


Figure 63 – Voltage magnitude at node 4 phase A in p.u. at one of the Monte Carlo trials.

The proposed approach obtained relatively good estimation even in this PMU non-observable scenario under a non-stationary condition, i.e. when the hypothesis that the previous state is a good approximation to current state is false. Fig. 64 shows a boxplot of the estimation

error in this simulation test obtained with the proposed Bayesian Information Fusion. The largest estimation error kept on node 4, where the event occurred. And the nodes that kept their loads near their initial nominal values did not present a significant change in estimation accuracy.

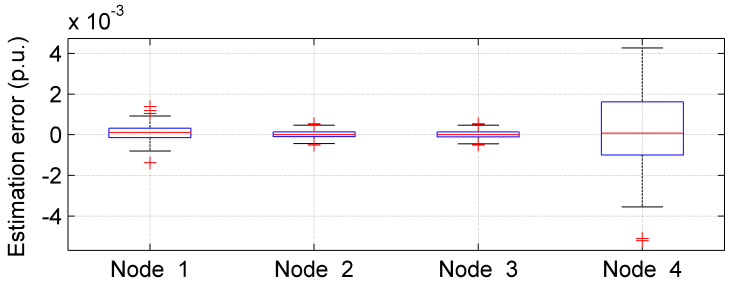


Figure 64 – State estimation error of the Bayesian Information Fusion state estimator ($\hat{x}_t - x_t^{lf}$) in p.u. for all time samples and all Monte Carlo repetitions according to each node.

This test was also used to show the credibility interval concept of the proposed Bayesian Information Fusion with PMUs. Fig. 65 shows the credibility intervals acquired with the prior and posterior distribution obtained by the proposed Bayesian DSSE for the real part of the complex voltage at node 4 (state variable x_{10}) at all samples.

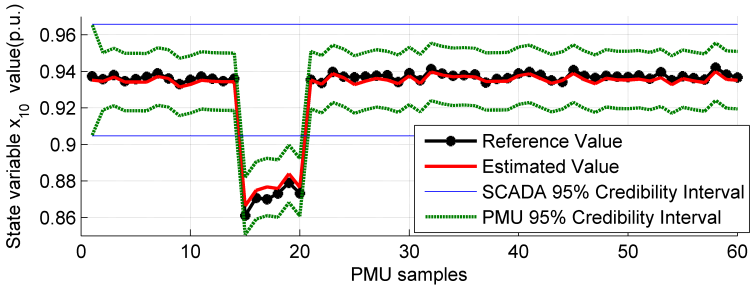


Figure 65 – Credibility intervals with SCADA prior and PMU posterior during an event occurrence, for state variable x_{10}.

Fig. 66 illustrates the concept of credibility interval in terms of estimated prior and posterior distributions for the state variable x_{10}. The prior distribution, and the posterior distribution at $t = 14$ (before the load event) and at $t = 15$ (after the load event) demonstrate the change in the state variable that occurred. While the DS is operating in a stationary condition, all the estimated state values with the PMUs fell inside the prior credibility interval, as it can be seen in the posterior distribution for $t = 14$. When the estimated value obtained in the MAP stage, at $t = 15$, fell outside the prior credibility interval, the previous state is not a good approximation to current state. That is, the system is no longer near the same state when the SCADA measurements were acquired. Since the event was detected with the prior distribution credibility interval, the test indicated a large event occurred.

The following chapter will present the detailed three-phase distribution systems models employed in this Thesis to accurately represent the unbalanced and asymmetrical nature of

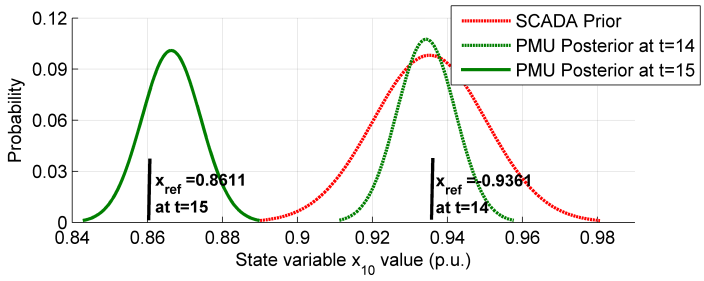


Figure 66 – SCADA prior distribution and PMUs posterior distributions, for state variable x_{10} at $t = 14$ and $t = 15$.

distribution systems. Chapter 6 presents the proposed Bayesian Information Fusion method to deal with different time scales of the diverse sources of information in distribution systems, such as load pseudo measurements, smart meters, SCADA and PMUs. Also, Chapter 7 a multiarea state estimation based on bayesian spatial fusion concepts is also presented to deal with the scalability of distribution networks along to demonstrate important computational requirements. Finally, Chapter 5 non-Gaussian characteristics for measurement noise and prior distributions under Information Theory concepts, a smoothed Bayesian MAP procedure.

APPENDIX D – MODELS OF THE PHYSICAL COMPONENTS OF A DISTRIBUTION SYSTEM

There is a variety of different components installed across the distribution systems, with different purposes, characteristics, connections, functionalities, and flexibility. Despite their intrinsic differences according to the type of component, fundamentally they can be individually described by their effects on the electrical quantities of the power system. This Appendix presents the main components and the respective characteristics that represent their operation inside the distribution networks. The model parameters are described and discussed, and the respective associated two-port model is derived.

Distribution Circuits

The distribution circuits represent the majority of the distribution networks, mainly composed of electric cables installed as overhead or underground circuits. The respective component model is obtained by electrical equations that capture the electromagnetic field relations for each phase and among different phases within the circuits. In general, they comprise a series impedance, to accommodate thermal losses and the effect of the magnetic fields surrounding the conductors, and a shunt admittance, to represent the potential electrical field between the conductors and the ground (Kersting, 2001; Sallam; Malik, 2018; Kagan, 2008; Kersting, 2008). Figure 67 illustrates a typical three-wire cable circuit geometric disposition for an overhead medium voltage distribution circuit. The respective electrical model and parameters are also presented. Such an electrical circuit comprises the two-port mathematical model and associated parameters.

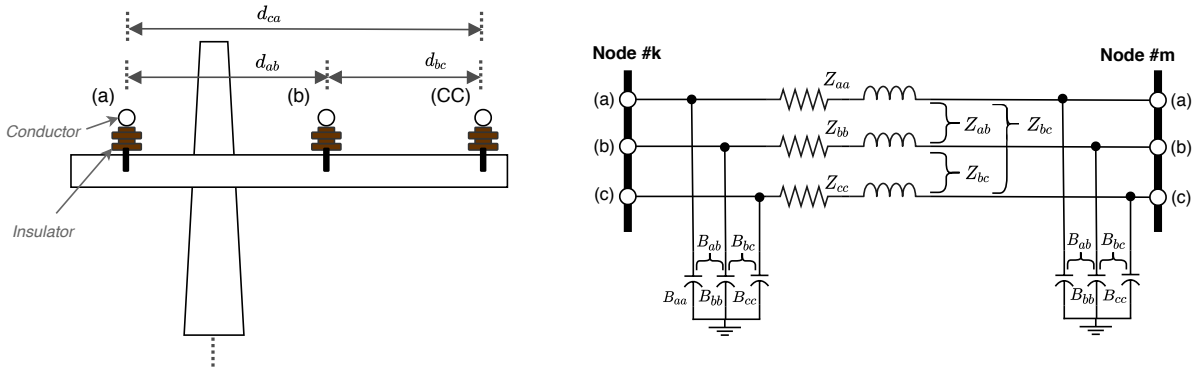


Figure 67 – Example of distribution circuit with an overhead configuration for a three-phase medium voltage circuit and its respective electrical model.

The series impedance (z) comprise a resistance (r) and inductive reactance (x), from the circuit inductance (l) and system frequency (f), for each phase of the circuit. Also, the model comprises mutual coupling among different phases. The series parameter may be expressed by

the general form described in (D.1) for the self impedance of each phase i, and in (D.2) for the mutual impedance among different phases i and j.

$$z^{ii} = r^{ii} + j2\pi f l^{ii} = r^{ii} + jx^{ii} \tag{D.1}$$

$$z^{ij} = r^{ij} + j2\pi f l^{ij} = r^{ij} + jx^{ij} \tag{D.2}$$

The series self resistance r^{ii} reflects the thermal losses in the circuit according to the current flowing through the cable and the Joule effect. The resistance of the cable is related to the type of material, cross-section of the cable, and the length of the circuit. Typically, manufacturers provide cable resistance values for direct current condition. For alternating current, the existence of a varying magnetic field inside the conductor results in a non-uniform distribution of current inside the cross-section of the cable, thickening the current density near the surface of the cables, known as *skin effect* (Kagan, 2008). The effect of the magnetic field of adjacent conductors, for instance, different phases near the conductor, also increases such effect, known as *proximity effect*. In practice, both effects increase the effective resistance of the cable and can be compensated by correction factors when calculating the cable resistance. Another important factor is temperature, which tends to increase the electrical resistivity of the materials, also compensated by a thermal coefficient related to the type of material.

Besides the series resistance, the series inductance comprises the magnetic field effects that arise inside and outside the conductors. It is a consequence of the current flowing in the conductors and among different phases, resulting in a self l^{ii} and a mutual l^{ij} inductance, respectively. Such magnetic fields are a direct result of the alternated electric current flowing through the conductors of the distribution circuits, derived from *Faraday's law of electromagnetic induction*. Besides the magnetic permeability physical constants, the geometric disposition of the cables is essential to the proper calculation of the distribution circuit inductance and magnetic coupling among phases (Kersting, 2001). The inner disposition of the cables comprises the first part of the inductance calculation. It encompasses the total magnetic flux inside each conductor of each phase. The geometric mean radius of the conductors is employed to calculate the self inductance parameters. The second part consists of the total magnetic flux outward the conductor among different conductors. This second part of the inductance calculation involves all different phases and neutral conductors in the vicinity of the distribution circuit, yielding the mutual inductance parameters. Typically, the inner geometric radius of the conductors is provided by manufacturers, and the distance among different conductors depends on the geometric disposition of the cables in the structures of the poles and cross arms, as illustrated in Figure 67, with the respective distances among the phase conductors (d_{ab}, d_{bc} and d_{ca}).

The effect of the return path of the current through the ground resistance can also be taken into account by performing a correction in the impedance values, known as *Carson's equations* (Kersting, 2001; Kagan, 2008; Carson, 1926). The Carson method consists basically

of reflecting the conductors in the soil and considering a uniform ground resistance and infinite extension. The method comprises introducing correction factors for the self and mutual series impedance to take into account the effect of grounding resistance. Indeed the mutual resistance r^{ij} term in (D.2) appears after this correction is performed. For practical purposes, a simplified version known as Modified Carson equations is typically used, by considering few terms for the correction factors.

The model may also comprise of a shunt admittance (y_{shunt}) from the circuit shunt capacitance (c) and system frequency (f), for each phase of the circuit. Such shunt parameter may be expressed by the general form described in (D.3) for the self admittance of each phase i and in (D.4) for the mutual admittance among different phases i and j. This shunt admittance captures the electric field among the conductors and the soil, also resulting in a self c^{ii} and a mutual c^{ij} capacitance. The electric field arises from the charged conductors and can be derived from *Coulomb's law* using the reflected image of the geometric disposition of the cables in the soil (Kersting, 2001). Besides the electric permittivity physical constants, the geometric disposition of the cables is essential to the proper calculation of the distribution circuit capacitance among phases. The calculation comprises the electric potentials difference among conductors, from which a capacitance matrix can be derived, as the inverse of the potential coefficient matrix. A shunt conductance may also be employed, for instance, to capture ionizing effects of the conductors due to the *Corona effect* or to insulator leakage, but is generally neglected in distribution systems (and in transmission systems), since are very small (Kersting, 2001; Glover; Sarma; Overbye, 2012).

$$y^{ii}_{shunt} = j2\pi f c^{ii} = jb^{ii} \tag{D.3}$$

$$y^{ij}_{shunt} = j2\pi f c^{ij} = jb^{ij} \tag{D.4}$$

In the case of underground cables, the same Carson method can be applied to model the series and shunt parameters, according to the geometric disposition of the cables and manufacturer parameters. There are, however, some additional considerations to be taken into account regarding different underground configurations and cables with concentric neutral conductors or with tape-shielded conductors. In this case, each phase will present a particular neutral conductor and associated variables and parameters for each phase. This will increase the dimension of the model to capture each neutral conductor for each cable of each phase, yielding in a $(6x6)$ impedance matrix with 3 dimensions for the phase conductors and the other 3 dimensions for each neutral conductor, all with its respective coupling terms.

Once distribution circuits consist of single, two, or three-phase untransposed lines, the Carson method provide an accurate model to represent them in any steady state analysis for distribution systems. There are some cases where sequence components (positive, negative, and zero sequence impedances and admittances) are employed to model three-phase circuits.

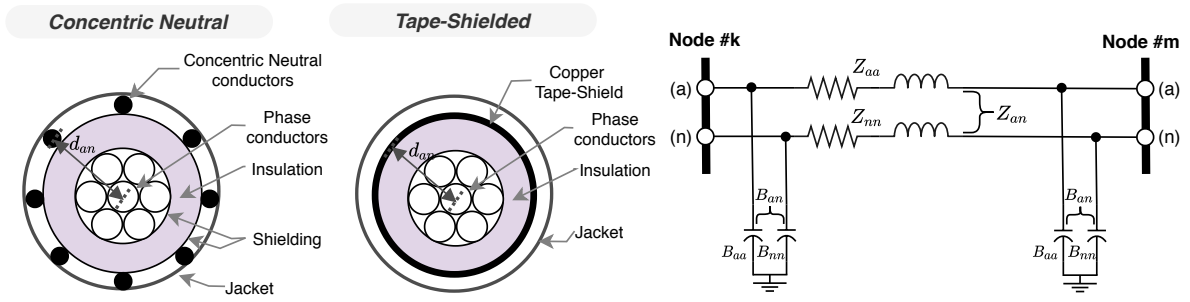


Figure 68 – Example of distribution circuit with an underground cable configuration, with concentric or with tape-shielded neutral, for a single-phase medium voltage circuit and its respective electrical model.

Such an approach, however, is only accurate for dealing with transposed circuits, where the mutual coupling terms of the model are equal (off-diagonal elements of the impedance and admittance matrices). It is not the case of distribution systems, where the majority of circuits are untransposed and asymmetrical, the main reason for using phase domain rather than symmetrical components in distribution systems analysis (Kersting, 2001). Details of the calculation of the resistance, inductance, and capacitance parameters for a distribution circuit, through Carson's method, can be found in (Kersting, 2001).

Finally, with the above parameters, it is possible to build the respective two-port admittance model described previously by the following equations. Single-phase or two-phase circuits are also represented by such model, by only considering the elements of the respectively connected phases in the two-port model.

$$\begin{aligned} Y_{kk} &= Y_{mm} = Z_{serie}^{-1} + Y_{shunt} \\ Y_{km} &= Y_{mk} = -Z_{serie}^{-1} \end{aligned} \tag{D.5}$$

where, Z_{serie} is the primitive series impedance matrix of the circuit composed by the self z^{ii} and mutual z^{ij} series impedances, and Y_{shunt} is the shunt admittance matrix of the circuit composed by the self y_{shunt}^{ii} and mutual z_{shunt}^{ij} shunt admittances. As an example, for a three-wire distribution circuit, those matrices would be:

$$Z_{serie} = \begin{pmatrix} r^{aa} + jx^{aa} & r^{ab} + jx^{ab} & r^{ac} + jx^{ac} \\ r^{ab} + jx^{ab} & r^{bb} + jx^{bb} & r^{bc} + jx^{bc} \\ r^{ac} + jx^{ac} & r^{bc} + jx^{bc} & r^{cc} + jx^{cc} \end{pmatrix} \tag{D.6}$$

$$Y_{shunt} = \begin{pmatrix} jb^{aa} & jb^{ab} & jb^{ac} \\ jb^{ab} & jb^{bb} & jb^{bc} \\ jb^{ac} & jb^{bc} & jb^{cc} \end{pmatrix} \tag{D.7}$$

The model can also be extended to represent the neutral conductors if available, which is the case of four-wire distribution circuits or underground cables with concentric neutral (Jesus;

Antunes, 2018; Kersting, 2008). In this case, an additional dimension is incorporated in the model to capture neutral conductor condition, and the associated neutral-ground voltages and neutral conductor current are considered in the two-port model as well, by the following matrices:

$$Z_{serie} = \begin{pmatrix} r^{aa} + jx^{aa} & r^{ab} + jx^{ab} & r^{ac} + jx^{ac} & r^{an} + jx^{an} \\ r^{ab} + jx^{ab} & r^{bb} + jx^{bb} & r^{bc} + jx^{bc} & r^{bn} + jx^{bn} \\ r^{ac} + jx^{ac} & r^{bc} + jx^{bc} & r^{cc} + jx^{cc} & r^{cn} + jx^{cn} \\ r^{an} + jx^{an} & r^{bn} + jx^{bn} & r^{cn} + jx^{cn} & r^{nn} + jx^{nn} \end{pmatrix} \tag{D.8}$$

$$Y_{shunt} = \begin{pmatrix} jb^{aa} & jb^{ab} & jb^{ac} & jb^{an} \\ jb^{ab} & jb^{bb} & jb^{bc} & jb^{bn} \\ jb^{ac} & jb^{bc} & jb^{cc} & jb^{cn} \\ jb^{an} & jb^{bn} & jb^{cn} & jb^{nn} \end{pmatrix} \tag{D.9}$$

Finally, a common practice for distribution circuits is to simplify the model by discarding the dimensions and variables associated with the neutral conductors through a technique known as *Kron reduction* (Kersting, 2001; Kron, 1951). The Kron reduction is based on the assumption of multi-grounded systems as if each node of the distribution circuit has the neutral terminals solidly grounded. Consequently, the neutral-to-ground voltage equals zero in all nodes of the distribution circuit, which enables discarding the associated variables and reducing the dimensionality of the distribution circuit only to the phase conductors. The Kron reduction technique then consists of a matrix elimination process applied to the series impedance and shunt admittance matrices obtained by the Carson method, which is described in (D.10) for the series impedance as an example.

$$Z^{abc} = Z^{ij} - Z^{in}(Z^{nn})^{-1}Z^{jn} \tag{D.10}$$

where the elements of the equation are the sub-matrices of the original full series impedance matrix: Z^{ij} is associated with all the phases conductors i and j; Z^{in} and Z^{jn} are associated with coupling elements between i and j phases and n neutral conductors; and, Z^{nn} is associated with the n neutral conductors. With this reduction, the final matrix Z^{abc} has the same size of phase conductors, for instance, a $(3x3)$ for three-phase *abc* circuits.

Power Transformers

Power transformers are equipment responsible for connecting different voltage levels of the distribution network, for instance, high voltage with medium voltage and medium voltage with low voltage circuits. They are mainly inductive elements constructed with magnetic coils in two separate windings, the primary side and the secondary side. Figure 69 illustrates the most common three-phase transformer connections, with Delta, Wye, and Grounded-Wye connections.

The transformers may also be represented by equation 4.1, in which the sub-matrices Y_{kk}, Y_{km}, Y_{mk} and Y_{mm} assume distinct values according to the connection of the three-phase transformer. The most common connections are presented in Table 14 for the step-down transformers, in case of a step-up transformer, columns regarding the mutual admittance have to be swapped (Chen et al., 1991; Peng Xiao; Yu; Wei Yan, 2006). Also, the sub-matrices Y_I, Y_{II} and Y_{III} in Table 14 are given as follows, in which y_t is the *per unit* transformer admittance (Arrillaga; Harker, 1978).

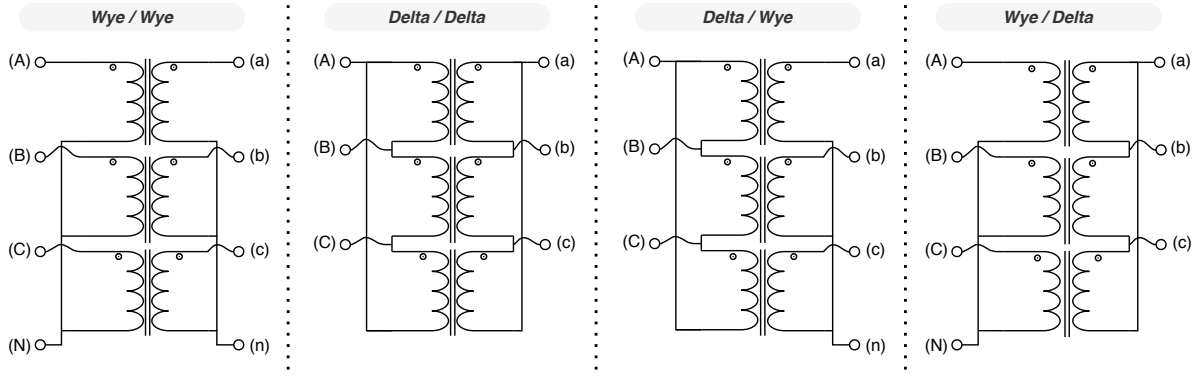


Figure 69 – Three-phase transformer connections of the primary and secondary windings. The Wye connection may present a solidly grounded connection (Wye-Grounded). Detailed grounding models may incorporate a reactor as well.

$$Y_I = \begin{pmatrix} y_t & 0 & 0 \\ 0 & y_t & 0 \\ 0 & 0 & y_t \end{pmatrix}; \qquad Y_{II} = \frac{1}{3}\begin{pmatrix} 2y_t & -y_t & -y_t \\ -y_t & 2y_t & -y_t \\ -y_t & -y_t & 2y_t \end{pmatrix}; \qquad \text{(D.11)}$$

$$Y_{III} = \frac{1}{\sqrt{3}}\begin{pmatrix} -y_t & y_t & 0 \\ 0 & -y_t & y_t \\ y_t & 0 & -y_t \end{pmatrix}$$

In case the transformer has an off-nominal tap, namely α for the primary and β for the secondary tap ratio, the self- and mutual-admittance matrices have the following adjustment (Arrillaga; Harker, 1978):

- divide the primary self-admittance matrix by α^2;
- divide the secondary self-admittance matrix by β^2;
- divide the mutual-admittance matrices by $\alpha\beta$.

Furthermore, all of the above models comprise three-wire or solidly grounded connections. In case of open grounded or in order to represent grounding systems, the above models can be extended for incorporating neutral-to-earth voltages as complementary state variables, or to

Table 14 – Sub-matrices for step down three-phase transformers according to its most common primary and secondary connections. *Source*: (Arrillaga; Harker, 1978)

Winding Connection		**Admittance sub-matrices**			
Primary	*Secondary*	Y_{kk}	Y_{mm}	Y_{km}	Y_{mk}
Yn	Yn	$\alpha^2.Y_I$	$\beta^2.Y_I$	$-\alpha.\beta.Y_I$	$-\alpha.\beta.Y_I$
Yn	Y	$\alpha^2.Y_{II}$	$\beta^2.Y_{II}$	$-\alpha.\beta.Y_{II}$	$-\alpha.\beta.Y_{II}$
Yn	D	$\alpha^2.Y_I$	$\beta^2.Y_{II}$	$\alpha.\beta.Y_{III}$	$\alpha.\beta.Y_{III}^t$
Y	Yn	$\alpha^2.Y_{II}$	$\beta^2.Y_{II}$	$-\alpha.\beta.Y_{II}$	$-\alpha.\beta.Y_{II}$
Y	Y	$\alpha^2.Y_{II}$	$\beta^2.Y_{II}$	$-\alpha.\beta.Y_{II}$	$-\alpha.\beta.Y_{II}$
Y	D	$\alpha^2.Y_{II}$	$\beta^2.Y_{II}$	$\alpha.\beta.Y_{III}$	$\alpha.\beta.Y_{III}^t$
D	Yn	$\alpha^2.Y_{II}$	$\beta^2.Y_I$	$\alpha.\beta.Y_{III}$	$\alpha.\beta.Y_{III}^t$
D	Y	$\alpha^2.Y_{II}$	$\beta^2.Y_{II}$	$\alpha.\beta.Y_{III}^t$	$\alpha.\beta.Y_{III}$
D	D	$\alpha^2.Y_{II}$	$\beta^2.Y_{II}$	$-\alpha.\beta.Y_{II}$	$-\alpha.\beta.Y_{II}$

complement the representation of solidly grounded connections where the sum of phase-voltages equals to zero. Similar to the four-wire distribution circuits, another variable is included in the two-port model that represents the effect of transformer grounding.

Besides those three-phase typical connections, there are also the Open-Wye and Open-Delta connections that may be employed in distribution systems. In this case, the two-port model must be adequate to represent the specific connections of each transformer, that is the proper input voltages and currents according to connected phases of the transformer, as presented by (Chen; Chang, 1992). Equation (D.12) presents the two-port model for an Open-Delta/Open-Delta (OD/OD) transformer, and equation (D.13), for an Open-Wye/Open-Delta (OY/OD) transformer.

$$Y_{kk} = \frac{y_t}{\alpha^2}.\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \qquad Y_{km} = Y_{mk} = \frac{y_t}{\alpha\beta}.\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \tag{D.12}$$

$$Y_{mm} = \frac{y_t}{\beta^2}.\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

$$Y_{kk} = \frac{y_t}{\alpha^2}.\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad Y_{km} = Y_{mk}^t = \frac{y_t}{\alpha\beta}.\begin{pmatrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0 \\ 0 & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \tag{D.13}$$

$$Y_{mm} = \frac{y_t}{\beta^2}.\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ 0 & -\frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

Another possibility, typically encountered for single-phase transformers, is the one with a center-tapped secondary winding. Such type of connection is common for distribution transformer in the United States, interfacing medium and low voltage networks. The center-tap is typically grounded, which enables phase to neutral and phase to phase voltage to be obtained

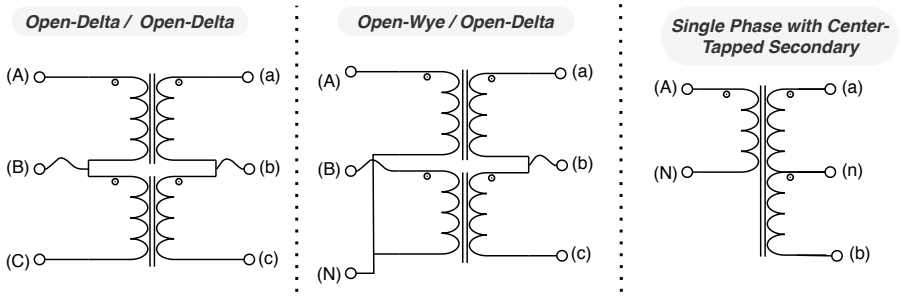


Figure 70 – Three-phase transformer open connections and single-phase with center-tapped secondary connection.

for the loads. The winding individual impedance values may be difficult to find in practice, and often empirical equations provide individual parameters for each winding and also their coupling (Kersting, 2001; Kersting, 2009). With such parameters, the transformer can be described by the following two-port admittance matrices:

$$Y_{kk} = \frac{y_{AN}}{\alpha^2} \qquad Y_{km} = Y_{mk}^t = \frac{1}{\alpha\beta} \cdot \begin{pmatrix} -y_{Aa} & -y_{Ab} \end{pmatrix} \qquad \text{(D.14)}$$

$$Y_{mm} = \frac{1}{\beta^2} \cdot \begin{pmatrix} y_{an} & -y_{ab} \\ -y_{ab} & y_{bn} \end{pmatrix}$$

where y_{AN} denotes the self admittance of the primary winding, and y_{Aa} and y_{Ab} the coupling admittance between the primary winding and each half of the secondary winding, y_{an} and y_{bn} the self admittance of the two half's of the secondary winding, and y_{ab} their mutual coupling admittance.

Besides the above connections, there are also the three-winding three-phase transformers, typically encountered at large distribution substations with two possibilities of the medium voltage supply. Another important application of this transformer is interfacing power electronic converters that may benefit from the phase displacement between the secondary and tertiary windings, with grounded-wye and delta connections, respectively. Three individual three-phase transformers represent the three-winding transformer according to the windings connections and manufacturer impedance data from the different terminals: from primary to secondary; from primary to tertiary; and from tertiary to secondary. Typically the secondary presents a Wye connection and the tertiary a Delta connection.

Finally, some analysis may even consider core losses and magnetizing reactances (Chen et al., 1991; Chen; Chang, 1992). These parameters were omitted in the above models but can be incorporated as shunt elements. A shunt resistance for the core loss and a shunt inductance for the magnetizing reactance, both represented at the primary terminal of the transformers. Typically such models may be necessary for assessing transient responses, evaluating harmonic content, or even for detailed losses calculation on distribution networks. Finally, Box 2 presents an example of a small distribution feeder to illustrate the distribution circuit and power transformer models.

Voltage Regulators

Voltage regulators are a special kind of transformer that controls the voltage at specific nodes of the network by increasing or decreasing the number of coils from a transformer, changing the transformation relation according to discrete tap switching operations. They perform an essential control task to keep voltage levels within operational limits. Typically they are built as a particular type of auto-transformers with tap switching mechanism, capable of switching tap positions while energized through on-load tap changers (OLTC). Constructively, the tap switching can be performed on the primary or in the secondary side of the voltage regulator, referred to as "Type A" and "Type B" voltage regulators (Kersting, 2001). The range for possible coil position, i.e. the taps of the voltage regulator, may vary for each manufacturer and regulator type. In distribution systems, it is often employed 16 taps that correspond to an increase/decrease of 10 % of the voltage on the secondary side (16 taps to increase, 16 taps to decrease, and a nominal position). Figure 71 illustrates in a single-phase schematic the main components of the voltage regulator and the associated instrumentation and control equipment.

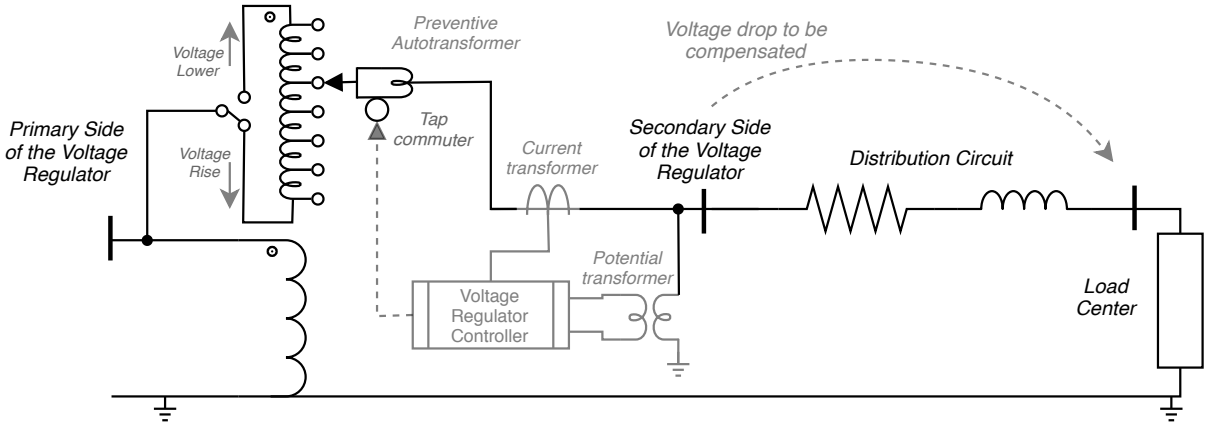


Figure 71 – Single-phase schematic of a "Type A" voltage regulator, its main components, and voltage control strategy illustration.

Regarding the voltage regulator mathematical model, a similar approach to the power transformer is employed to represent different connections of voltage regulators. The most common one is the three-phase Wye-grounded connection, but the open Delta connections are also encountered. In practice, due to constructive characteristics, their resistance and reactance are very small and are often represented as ideal transformers in the network model, neglecting its series impedance. Care must be taken in this approach, since it may severely aggravate numerical ill-conditioning (Cheng; Shirmohammadi, 1995). In this sense, the voltage regulator model may be associated with adjacent components to counter such effect, or by employing more stable numerical methods when performing analysis with voltage regulators. Alike the power transformer, the type of connection (Wye-Grounded, Wye, Delta, Open-Wye, Open-Delta, or Single Phase) will define the two-port model according to equations the admittance sub-matrices in (D.11). The difference relies on the off-nominal tap parameters that now comprise an

individual transformer relation for each phase i, denoted as α^i. Besides, typically only one side will present off-nominal tap relations. As an example, the Wye-Grounded voltage regulator, with tap switching represented on the secondary, two-port model is presented below in (D.15).

$$\begin{pmatrix} Y_{kk} & Y_{km} \\ Y_{mk} & Y_{mm} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} y_t & 0 & 0 \\ 0 & y_t & 0 \\ 0 & 0 & y_t \end{pmatrix} & \begin{pmatrix} -\frac{1}{\alpha^a} y_t & 0 & 0 \\ 0 & -\frac{1}{\alpha^b} y_t & 0 \\ 0 & 0 & -\frac{1}{\alpha^c} y_t \end{pmatrix} \\ \begin{pmatrix} -\frac{1}{\alpha^a} y_t & 0 & 0 \\ 0 & -\frac{1}{\alpha^b} y_t & 0 \\ 0 & 0 & -\frac{1}{\alpha^c} y_t \end{pmatrix} & \begin{pmatrix} \frac{1}{\alpha^{a2}} y_t & 0 & 0 \\ 0 & \frac{1}{\alpha^{b2}} y_t & 0 \\ 0 & 0 & \frac{1}{\alpha^{c2}} y_t \end{pmatrix} \end{pmatrix} \tag{D.15}$$

where, α^a, α^b and α^c are the off-nominal voltage regulator transformation relation according to the the tap position for the respective phase a, b and c.

Although an important detail for steady state analysis of distribution networks, in the context of distribution system state estimation, the tap values are typically assumed as monitored parameters, without the need of incorporating detailed controller models. Such an assumption is not always valid, and besides, the response time of such controllers may be faster than updates on the measurement set. It motivates some approaches to incorporate the tap positions as complementary variables to be estimated as in (Nanchian; Majumdar; Pal, 2017a; Massignan; Pereira; London, 2017).

Loads

In steady state analysis, loads of the system are often considered independent of the voltage frequency, leading to a representation with a constant load. However, in distribution system analysis, it may be required a more accurate representation of the loads, which leads to distinct models treating the load as voltage-dependent. One of the most traditional approaches is the exponential model (Kundur, 1994; Milanovic et al., 2013):

$$\begin{aligned} P_k^i &= P_k^{i_nom} \cdot \bar{V_k^i}^a \\ Q_k^i &= Q_k^{i_nom} \cdot \bar{V_k^i}^b \end{aligned} \tag{D.16}$$

In which the superscripts i denote the phase, the subscript k is the node, nom denotes the nominal load (in the beginning of the iterative process), and $\bar{V_k^i}$ corresponds to the relation between the updated voltage and its nominal value ($\bar{V_k} = V_k / V_k^{nom}$). The exponents a and b are used to represent a constant power, constant current, or constant impedance load characteristics, assuming the values of 0, 1 and 2, respectively, for each one of those. However, for composite loads these values usually range from 0.5 to 0.8 for a, and from 1.5 to 6 for b (Kundur, 1994).

Another model is the polynomial one, which is also known as the *ZIP* model (Kundur, 1994; Milanovic et al., 2013):

$$\begin{aligned} P_k^i &= P_k^{i-nom} \cdot \left(p_1 \bar{V}_k^{i\,2} + p_2 \bar{V}_k^i + p_3 \right) \\ Q_k^i &= Q_k^{i-nom} \cdot \left(q_1 \bar{V}_k^{i\,2} + q_2 \bar{V}_k^i + q_3 \right) \end{aligned} \tag{D.17}$$

This model is composed of three distinct parts, its constant impedance (Z), constant current (I), and constant power (P) portion (hence known as ZIP). Each one is multiplied by a coefficient that defines the proportion of each component in the load (p_1 to p_3 for the active load, and q_1 to q_3 for the reactive load). For example, a constant power load would have $p_3 = q_3 = 1$ while the others are set to zero. Any composite load may be represented as long as the sum of the distinct coefficients of the active or reactive load equals one (Kundur, 1994).

While the previous models efficiently represent most wye connected loads, delta connected loads need a few adjustments. Since most three-phase power system analyses are conducted using phase power and voltage, the consumed power of the delta connected load needs to be converted to its phase equivalent. To do so, it is assumed that the complex power and voltage of the delta-connected load are given by:

$$\dot{S}_k^{ab} = P_k^{ab} + jQ_k^{ab}; \quad \dot{S}_k^{bc} = P_k^{bc} + jQ_k^{bc}; \quad \dot{S}_k^{ca} = P_k^{ca} + jQ_k^{ca} \tag{D.18}$$

$$\dot{V}_k^{ab} = V_k^{ab} \angle \theta_k^{ab}; \quad \dot{V}_k^{bc} = V_k^{bc} \angle \theta_k^{bc}; \quad \dot{V}_k^{ca} = V_k^{ca} \angle \theta_k^{ca} \tag{D.19}$$

Therefore, the delta active and reactive power may be obtained using equations D.16 or D.17 according to the load model. After that, the line currents in the load are obtained with:

$$\dot{I}_k^{ab} = \left(\dot{S}_k^{ab} / \dot{V}_k^{ab} \right)^*; \quad \dot{I}_k^{bc} = \left(\dot{S}_k^{bc} / \dot{V}_k^{bc} \right)^*; \quad \dot{I}_k^{ca} = \left(\dot{S}_k^{ca} / \dot{V}_k^{ca} \right)^* \tag{D.20}$$

Furthermore, the wye equivalent complex phase load at bus k is obtained by:

$$\begin{aligned} \dot{S}_k^a &= \left(\dot{I}_k^{ab} - \dot{I}_k^{ca} \right)^* \dot{V}_k^a \\ \dot{S}_k^b &= \left(\dot{I}_k^{bc} - \dot{I}_k^{ab} \right)^* \dot{V}_k^b \\ \dot{S}_k^c &= \left(\dot{I}_k^{ca} - \dot{I}_k^{bc} \right)^* \dot{V}_k^c \end{aligned} \tag{D.21}$$

Consequently, the active and reactive powers (P_k^i and Q_k^i) of the wye equivalent specified complex phase load is obtained with the real and imaginary parts of the wye equivalent phase load, respectively. This same set of equations may also be applied to bi-phase loads.

The loads in distribution systems can also be associated with sensors, typically called smart meters, under an AMI concept (Huang; Lu; Lo, 2015). If the load values are effectively measured, they can be included in the state estimation measurement model, with some practical considerations that will be dealt with in future chapters. If such loads are not measured, as in the case of most of distribution systems, typically the load model comprises the use of typical load profiles (Jardini et al., 2000; Primadianto; Lu, 2017). Such load profiles are adopted as

an approximate estimate of the load values and are based mainly on consumer stratification and electricity monthly energy consumption. Although such profiles provide only a rough approximation of the real load values, in the absence of smart meters they are the only source of information regarding the load values for the steady state analysis. In the state estimation topic, such a technique is named as load pseudo measurement modelling.

Each consumer unit is associated with a typical load profile based on the consumer classification (e.g. residential, commercial or industrial), the monthly energy consumption in kWh, the number of phases of the consumer, and a respective typical load profile (Ghosh; Lubkeman; Jones, 1997; Jardini et al., 2000; Massignan et al., 2015). The load profiles are associated with the consumer's stratification and consist of daily curves with each point being associated with a normalized mean active load value and the respective standard deviation throughout the instants of a day and can be given for every 10 minutes up to each hour or more, often separated between a day of the week or weekends. The energy consumption is measured each month for billing purposes, or may be used previous day smart meter information, and is the typical choice as the scale factor for the typical load profile. With such information, each phase of the consumer can be associated with an active load curve with equation (D.22). The reactive power can be estimated using typical power factors or similar information, if available.

$$p_c^i(t) = \frac{kWh_c^i}{\Delta t_k} p_\omega(t), c \in \omega \tag{D.22}$$

where, $p_c^i(t)$ is the active load for each consumer c in the i-th phase at each instant t, kWh_c^i and Δt_k the respective energy consumption and measured time interval to be used as scaling factor, and $p_\omega(t)$ is the typical normalized load profile for the respective classification ω of the consumer unit among all possible classifications Ω (e.g. residential, commercial or industrial).

Typically, such pseudo measurements provide information for achieving observability at the medium voltage primary feeders, that is, they provide enough information to reach the minimum in order to perform the state estimation process in the primary feeder, in a complementary manner. Although they are associated with the consumer units at low voltage secondary systems, they are often represented in the primary feeders, for medium voltage system analysis. This way, the loads within a secondary low voltage system are aggregated at the respective power transformer that supplies energy from the primary feeder. The pseudo-measurement load consists of the expected value for the active load in each transformer k that connects the primary feeder to the secondary circuits ($E[P_k^i(t)]$), and the respective variance ($Var[P_k^i(t)]$), obtained by the aggregation process (Jardini et al., 2000; Massignan et al., 2015). It can be calculated as follow in equation (D.23) and (D.24).

$$E[P_k^i(t)] = E\left[\sum_{c \in \mathfrak{S}_k} p_c^i(t)\right] = \sum_{\omega \in \Omega} \left(\sum_{\substack{c \in \mathfrak{S}_k \\ c \in \omega}} \frac{kWh_c^i}{\Delta t_c} \right) E[p_\omega(t)] \tag{D.23}$$

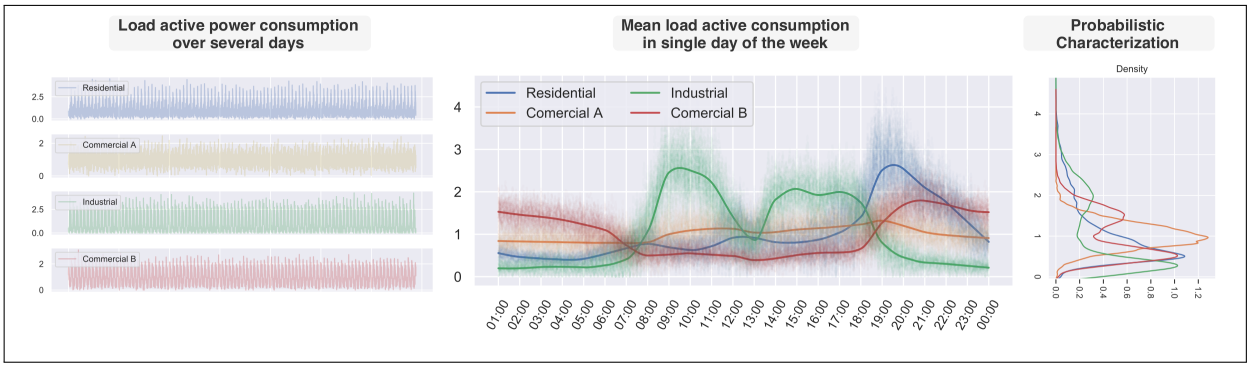


Figure 72 – Example of typical load curves and stratification of consumers to characterize distribution system loads.

$$Var[P_k^i(t)] = Var\left[\sum_{c\in\mathfrak{S}_k} p_c^i(t)\right] = \sum_{\omega\in\Omega}\left(\sum_{\substack{c\in\mathfrak{S}_k \\ c\in\omega}} \frac{kWh_c^i}{\Delta t_c}\right)^2 Var[p_\omega(t)] \tag{D.24}$$

where, $\mathfrak{S}_k$ is the set of consumer units c connected to the secondary circuits of transformer k, in each i-th phase. The expected value and variance, $E[p_\omega(t)]$ and $Var[p_\omega(t)]$, for each class of consumers are obtained by the typical load profiles associated with the respective class. Figure 72 illustrates the idea of typical load profiles and how they represent the normalized mean value of different consumers pertaining to the same class. The pseudo measurement model tries to capture this typical behavior of the load over a day. By performing the load aggregation, multiple consumers have their loads aggregated at the transformer that feeds their respective secondary circuits. The probabilistic interpretation associated with the typical load profiles can also provide additional information for more detailed methods (Ghosh; Lubkeman; Jones, 1997; Singh et al., 2011; Massignan et al., 2018).

The pseudo measurements are indeed approximations for the load values, based on the typical behavior of a class of consumers, and thus carry more uncertainty to steady state analysis. Thereby, the accuracy of the state estimation process is largely affected (Muscas et al., 2014). Nonetheless, in the absence of information from real time measurements, such typical profiles have been used in practice along with another step of refinement. Among the different approaches proposed in the literature for pseudo measurement modelling, the most common are based on: neural networks (Manitsas et al., 2012); associating probabilistic models based on Gaussian-mixtures (Singh et al., 2011); refined in real-time with load flow calculation (Massignan et al., 2018); using lookup tables and rescaling factors (Ghosh; Lubkeman; Jones, 1997); or also be included in special estimation formulations (Džafić et al., 2013). These approaches are very important for any practical and realistic distribution system state estimation implementation, since it can increase the quality of the pseudo measurements. Despite the current evolution towards the installation of smart meters, this kind of methodology still is very appealing, to compensate for the possible loss of communication or even to reduce traffic in the communication networks.

Shunt Capacitors and Reactors

Shunt capacitors are a type of equipment that offers reactive support for the grid, improving voltage quality aspects. They are in essence a shunt capacitive susceptance directly connected to the nodes of a distribution system. They can be installed as a fixed capacitor or as a variable capacitor bank, which consists of a group of capacitors and a local controller able to modify the total capacitance of the group by turning on or off some capacitors depending on the control strategy (Cheng; Shirmohammadi, 1995; Kersting, 2001). Since the shunt capacitors are directly connected to the nodes, their two-port model can be written as a single port:

$$\dot{I}_k = Y_{kk}\dot{V}_k \tag{D.25}$$

Which is, basically, the same equations that described the power flow, but, in this case, neglecting the terms of the second terminal of the model, denoted by the subscript m. The admittance matrix depends upon the type of connection of the shunt capacitor, which may be a Wye, a Wye-Grounded, a Delta, or a single phase connection. For three-phase modeling, the structure of the admittance matrix will be given by equation (D.26), similar to the different transformer connections, but now considering the capacitor susceptance b_c^{sh}, and considering the matrix Y_I for Wye-Grounded connection, Y_{II} for Wye connection and Y_{II} for Delta connection, respectively. For single-phase capacitors, they are modeled directly on the respectively connected phase.

$$Y_I = jb_c^{sh}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix};\quad Y_{II} = j\frac{b_c^{sh}}{3}\begin{pmatrix}2&-1&-1\\-1&2&-1\\-1&-1&2\end{pmatrix};Y_{III} = j\frac{b_c^{sh}}{\sqrt{3}}\begin{pmatrix}-1&1&0\\0&-1&1\\1&0&-1\end{pmatrix} \tag{D.26}$$

The capacitor's susceptance b_c^{sh} is obtained from the nominal parameters, following:

$$b_c^{sh} = \frac{MVAr_{nom}}{V_{nom}^2} \tag{D.27}$$

where $MVAr_{nom}$ is the nominal reactive power of the capacitor in MVAr, and V_{nom} is the capacitor nominal voltage in kV.

Another type of shunt equipment are reactors and grounding resistances, often used in grounding systems of distribution networks. The same idea as the above is employed for such elements to obtain detailed grounding representations; however, in this case, it expresses a relation between the voltage and current phasor of the neutral-to-ground variables:

$$\dot{I}_k^n = Y_k^{nn}\dot{V}_k^n \tag{D.28}$$

where the superscript n denotes the neutral-to-ground variables and parameters at node k. In the case of a shunt reactor, the parameter Y_k^{nn} is equal to $1/jX_{grounding}$, and $1/R_{grounding}$ for

resistance, which are the nominal parameters of the grounding reactor. Finally, Box 3 presents an example of a medium size radial distribution feeder to illustrate the shunt capacitors and voltage regulators models.

Annex

ANNEX A – DOCTORAL PROGRAM

In this section, the publications related to the doctoral program are presented. First the ones related to the presented Thesis, followed by the collaborations in other research topics.

Related with this Doctoral Research

Journals

1. **Published**: Bretas, N. G., Massignan, J. A. D.; London JR., J. B. A.; Bretas, A. S., "Hybrid Physics-Based Adaptive Kalman Filter State Estimation Framework". *Energies*, vol. 14, no. 20 2021;

2. **Published**: Massignan, J. A. D.; London JR., J. B. A.; Bessani, M.; Maciel, C. D.; Dantas, L. B., Miranda, V. "Bayesian Inference Approach for Information Fusion in Distribution System State Estimation," *IEEE Transactions on Smart Grid*, vol. 13, no. 1, pp. 526-540, Jan. 2022";

3. **Published**: G.M. Hebling, J.A.D. Massignan, J.B.A. London Junior, M.H.M Camillo ," Sparse and numerically stable implementation of a distribution system state estimation based on Multifrontal QR factorization". *Electric Power System Research*, v. 189, p. 106734, 2020.

4. **Published**: Massignan, J. A. D.; London JR., J. B. A. ; Miranda, V. . Tracking Power System State Evolution with Maximum-correntropy-based Extended Kalman Filter. Journal of Modern Power Systems and Clean Energy, v. 8, p. 616-626, 2020.

5. **Published**: Massignan, J. A. D., London JR., J. B. A.; Bessani, M.; Maciel, C. D.; Delbem, A. C. B.; Camillo, M. H. M.. "In-Field Validation of a Real Time Monitoring Tool for Distribution Feeders". *IEEE Transacations on Power Delivery*, vol. 33, no. 4, Aug. 2018;

Book Chapters

Four book chapters were accepted in the IET book "Power Distribution System State Estimation: Modelling and Alternative Approaches" that are currently under development and scheduled to be published in 2021.

1. **Accepted for publication**: João Bosco A. London Jr, Júlio Augusto Druzina Massignan, M. H.M. Camillo and Elizete Maria Lourenço , "Real-time monitoring of Distribution systems – COPEL experience"

2. **Accepted for publication**: João Bosco A. London Jr, Júlio Augusto Druzina Massignan, Gustavo Miranda Hebing, Odilon Luis Tortelli, Elizete Maria Lourenço and Renan Kovalczuk Portelinha, "Three-Phase Network Model for Steady State Analysis of Distribution Systems"

3. **Accepted for publication**: J. Bosco A. London Jr, Júlio Augusto Druzina Massignan, Paolo Attilio Pegoraro and Marco Pau, "Multiarea State Estimation for Distribution Systems"

4. **Accepted for publication**: Chan-Nan Lu, Mohammed Ansar Mohammed Manaz and Yu-Jen Lin, João Bosco A. London Jr, Júlio Augusto Druzina Massignan, "Including measurements with great difference in sampling rates in Distribution System State Estimation"

Conferences

1. Fernandes, J. P. R.; Massignan, J. A.D.; London Jr, J. B. A.; Fanucchi, R. Z., "Very Short-Term Current and Load Forecasting for Distribution Systems in Data Constrained Situations", in IEEE Power Tech 2021, Madrid, Jun 2021.

2. Melo, V. H. P.; Massignan, J. A.D.; London Jr, J. B. A.; Fanucchi, R. Z., "Estimation of Voltage Unbalance at the Reference Bus in Distribution System State Estimation", in IEEE Power Tech 2021, Madrid, Jun 2021.

3. Hebling, G. M.; Massignan, J. A.D.; London Jr, J. B. A.; Oliveira, R. "Sparse and Orthogonal Method for Fast Bad Data Processing in Distribution System State Estimation", in IEEE Power Tech 2021, Madrid, Jun 2021.

4. Massignan, J. A. D.; Melo, V. H. P. ; London JR, J. B. A. . Efeitos Não-Lineares na Redundância em Estimadores de Estado para Sistemas Elétricos de Potência. In: XXIII Congresso Brasileiro de Automática (aceito para publicação), 2020, Porto Alegre. Anais do XXIII Congresso Brasileiro de Automática, 2020.

5. Massignan, J. A. D.; Oliveira, R. ; Hebling, Gustavo M. ; London JR, J. B. A. . Aplicação de Métodos Ortogonais no Processamento de Erros Grosseiros em Estimadores de Estado para Sistemas de Distribuição. In: XXIII Congresso Brasileiro de Automática (aceito para publicação), 2020, Porto Alegre. Anais do XXIII Congresso Brasileiro de Automática, 2020.

6. Hebling, Gustavo M. ; Massignan, J. A. D. ; London JR, J. B. A. ; Dantas, L. B. . Análise de Padrões de Esparsidade e Métodos de Ordenação para Estimação de Estado em Sistemas de Distribuição. In: XXIII Congresso Brasileiro de Automática (aceito para publicação), 2020, Porto Alegre. Anais do XXIII Congresso Brasileiro de Automática, 2020.

7. Hebling, Gustavo M. ; Massignan, J. A. D. ; London, J. B. A. ; Camillo, M. H. M. . Aplicação de Equivalentes de Rede em Sistemas de Distribuição com Malhas entre Subestações. In: VIII Simpósio Brasileiro de Sistemas Elétricos, 2020, Santo André. Anais do VIII Simpósio Brasileiro de Sistemas Elétricos, 2020.

8. Massignan, J. A. D.; London, J. B. A. ; Miranda, V. . Vulnerability of Largest Normalized Residual Test and b-test to Gross Errors. In: IEEE Power & Energy Society General Meeting, 2020, Montreal. IEEE Power & Energy Society General Meeting, 2020.

9. Massignan, J. A. D.; London, J. B.A. ; Maciel, C. D. ; Bessani, M. ; Miranda, V. . PMUs and SCADA Measurements in Power System State Estimation through Bayesian Inference. In: 2019 IEEE Milan PowerTech, 2019, Milan. 2019 IEEE Milan PowerTech, 2019. p. 1.

10. Massignan, J. A. D.; London JR, J. B. A. ; Camillo, M. H. M. ; Miranda, V. . Estimação de Estado Multiárea para Monitoramento e Operação Descentralizada de Redes Inteligentes. In: XXV SNPTEE Seminário Nacional de Produção e Transmissão de Energia Elétrica, 2019, Belo Horizonte. (Aceito para publicação), 2019.

11. Massignan, J. A. D.; London JR, J. B. A. ; Cirino, R. B. Z. . Application and Comparison of Numerical Optimization Methods applied to Power System State Estimation. In: XIV Conferência Brasileira de DinÂmica, Controle e Aplicações, 2019, São C.. Anais do XIV Conferência Brasileira de DinÂmica, Controle e Aplicações, 2019.

12. Gonzaga Jr, R. M. ; Massignan, J. A. D. ; London JR, J. B. A. ; Maciel, C. D. ; ALMEIDA, R. M. A. ; Camillo, M. H. M. . An Embedded State Estimator for Reducing Data Volume and Processing in Smart Grids Monitoring. In: 3rd Workshop on Communication Networks and Power Systems, 2018, Brasília. 3rd Workshop on Communication Networks and Power Systems, 2018.

13. Avelino, L. L.; Massignan, J. A. D.; London JR., J. B. A.; Fernandes, J. P. R.. "Estimador de Demanda Trifásico em Tempo Real com Tratamento para Transformadores ΔY". No: *XXII Congresso Brasileiro de Automática (CBA)*, (Submetido);

14. Avelino, L. L.; Massignan, J. A. D.; London JR., J. B. A.; Oliveira, A. V.. "Application of a Three-Phase Load Estimation in a Real distribution Feeder". No: *IEEE URUCON 2017*, Outubro / 2017, Montevideu - Uruguai;

15. Camillo, M. H. M.; Fanucchi, R. Z.; Lima, T. W.; Soares, A. S.; Bessani, M.; Maciel, C. D.; Massignan, J. A. D.; Marques, L. T.;Delbem, A. B. C.; London JR., J. B. A.. "Detecção, Isolamento, Inspeção e Restabelecimento de Energia em Tempo Real em Sistemas de Distribuição de Larga Escala". No: *CITENEL 2017*, Agosto / 2017, João Pessoa - Paraíba;

16. Massignan, J. A. D.; Hebling, G. M.; Marques, L. T.; London JR., J. B. A.; Camillo, M. H. M.. "Modeling Issues on Load Flow Calculation for Meshed Distribution Systems". No: *IEEE Power Tech 2017*, Junho / 2017, Manchester - Inglaterra;

Collaborations

Journals

1. Bessani, M. ; Massignan, J. A.D. ; Santos, T. M.O. ; London, JOÃO B.A. ; Maciel, C. D. . "Multiple households very short-term load forecasting using bayesian networks". Electric Power System Research, v. 189, p. 106733, 2020.

2. Vigliassi, M. P.; Massignan, J. A. D.; Delbem, A. C. B.; London JR., J. B. A.; . "Multi-objective Evolutionary Algorithm in Tables for Placement of SCADA and PMU Considering the Concept of Pareto Frontier," *International Journal of Electrical Power and Energy Systems*, vol. 106, pp. 373-382 March 2019.

3. Bessani, M.; Massignan, J. A. D.; Fanucchi, R. Z.; Camillo, M. H. M.; Delbem, A. C. B.; London JR., J. B. A.; Maciel, C.D.. "Probabilistic Assessment of Power Distribution Systems Resilience under Extreme Weather," *IEEE Systems Journal*, vol. 13, no. 2, pp. 1747 - 1756, Jul. 2018.

Conferences

1. Agnoletto, E. J. ; Neves, R. V. A. ; Machado, R. Q. ; Oliveira, V. A. ; Massignan, J. A. D. . Gerenciamento De Energia Para Maximização Do Lucro Diário De Microrredes Com Dispositivos Armazenadores. In: XXII Congresso Brasileiro de Automática, 2018, João Pessoa. Anais do XXII Congresso Brasileiro de Automática, 2018.

2. Vigliassi, M. P. ; Santos, A. C. ; Massignan, J. A. D. ; London, J. B. A. ; DELBEM, A. C. B. . Estratégia Evolutiva Multiobjetivo Para Alocação De Medidores Scada E Umf Para Estimação De Estado. In: XXII Congresso Brasileiro de Automática, 2018, João Pessoa. Anais do XXII Congresso Brasileiro de Automática, 2018.

3. Gonzaga Jr, R. M. ; Massignan, J. A. D. ; London JR, J. B. A. . Analysis of the Hosting Capacity of a Real Distribution Feeder with Wind Generation. In: IEEE ISGT Innovative Smart Grid Technologies Latin America, 2019, Gramado. (Aceito para publicação), 2019.

4. Bessani, M. ; Massignan, J. A. D. ; London, J. B. A. ; Maciel, C. D. ; Fanucchi, R. Z. ; Camillo, M. H. M. . A hierarchical framework for complex networks robustness analysis to errors. In: 2017 Annual IEEE International Systems Conference (SysCon), 2017, Montreal. 2017 Annual IEEE International Systems Conference (SysCon), 2017. p. 1.

5. Camillo, M. H. M. ; Fanucchi, R. Z. ; Soares, T. W. L. ; Soares, A. S. ; Bessani, M. ; Maciel, C. D. ; Massignan, J. A. D. ; Marques, L. T. ; Delbem, A. C. B. ; London, J. B. A. . Detecção, Isolamento, Inspeção e Restabelecimento de Energia em Tempo Real em Sistemas de Distribuição de Larga Escala. In: IX Congresso de Inovação Tecnológica em Energia Elétrica CITENEL, 2017, João Pessoa. Anais do IX Congresso de Inovação Tecnológica em Energia Elétrica (IX CITENEL), 2017.