ISBN 978-82-326-6806-9 (printed ver.)
ISBN 978-82-326-6299-9 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (electronic ver.)

NTNU
Norwegian University of
Science and Technology

Doctoral theses at NTNU, 2021:273

Mohammad Ali Abooshahab

Dynamic state estimation for electrical power grids

Doctoral thesis
Mohammad Ali Abooshahab

Doctoral theses at NTNU, 2021:273

NTNU
Norwegian University of
Science and Technology
Thesis for the degree of
Philosophiae Doctor
Faculty of Information Technology
and Electrical Engineering
Department of Engineering Cybernetics

NTNU
Norwegian University of
Science and Technology

Mohammad Ali Abooshahab

Dynamic state estimation for electrical power grids

Thesis for the degree of Philosophiae Doctor

Trondheim, August 2021

Norwegian University of Science and Technology
Faculty of Information Technology
and Electrical Engineering
Department of Engineering Cybernetics

NTNU
Norwegian University of Science and
Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Information Technology
and Electrical Engineering
Department of Engineering Cybernetics

ISBN 978-82-326-6806-9 (printed ver.)
ISBN 978-82-326-6299-9 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (electronic ver.)

ITK-report: 2021-6-W

Doctoral theses at NTNU, 2021:273

Printed by Skipnes Kommunikasjon AS

Contents

List of Tables

List of Figures

Abstract

Increased electrification of society, as well as increased incidence of distributed power generation, causes more dynamic operation of modern electrical power grids. Methods for monitoring of power grids have traditionally focused on static operation, but must in the future be expected to have to account for system dynamics. On the other hand, privacy concerns for individuals, commercial secrecy concerns for companies, and cost concerns for power system operators, may cause power systems to be only partly known and without sufficient sensors for traditional dynamic state estimation in power systems. This thesis contributes to the understanding of Simultaneous Input and State Estimation (SISE), in particular its stability properties, and develops methods for dynamic monitoring of partially known power systems based on SISE.

The other main focus of the thesis is decentralized state estimation. Both the sheer scale of power networks, as well as the fact that different sections of the system may be operated by different companies, make a single centralized estimation center impractical and/or impossible. The thesis therefore develops decentralized estimation approaches for power systems, in which the observability requirements of existing decentralized estimation methods are relaxed.

In addition, the thesis also addresses cybersecurity and fault detection as well as sensor placement for power systems, and shows how SISE can be exploited in power system stabilizer (PSS) design.

Preface

The dissertation contains my research results during the period of 2017-2020, and this is presented as part of Ph.D. requirements at the Norwegian University of Science and Technology (NTNU). My primary supervisor is Prof. Morten Hovd and the research has been conducted at the Department of Engineering Cybernetics (ITK), my co-supervisor is prof. Kjetil Uhlen from the department of Electric Power Engineering. My research is part of the Centre for Intelligent Electricity Distribution (CINELDI) project, directed by Gerd Kjølle, which is targeting digitalizing and modernizing the electricity distribution grid for higher efficiency, flexibility and resilience. The Norwegian Research Council funded this project (project number 90198147). Without their help, this Ph.D. dissertation would not have existed, their support in this dissertation is greatly appreciated, and I am grateful for their help- thanks are due to them all.

Acknowledgment

First of all, I would like to special thanks to my principal supervisor, Prof. Morten Hovd, for his generous support of my Ph.D. thesis and related research, for his patience, inspiration and exceptional knowledge. His guidance has supported me in the study and writing of this thesis. I couldn't have imagined my Ph.D. study getting a better supervisor.

My heartfelt thanks also go to Prof. Kjetil Uhlen and Prof. Robert R. Bitmead who gave me the chance to work with them. It would not be possible to carry out this work without their precious help.

Throughout my Ph.D., I want to thank Sheida for her love and her companionship.

I thank my colleague, Dr. Xianfeng Song, for making me feel better at work and providing me help with a range of concerns during my Ph.D.

I would like to thank my family: my parents, and my sisters for their spiritual support throughout my life.

Chapter 1

Introduction

This chapter first provides background and motivation for using dynamic state estimation for electrical power grids in order to describe the main objectives of this thesis. Then the main contributions are explained, and finally, the thesis outline is presented.

1.1 Background and Motivation

The traditional distribution grid is monitored using infrequent the supervisory control and data acquisition (SCADA) data and a global steady-state power system model. This situation is already starting to change due to the low-carbon requirements and introduction of distributed energy resources, including photovoltaic (PV) generators, small-scale hydropower and wind farms, which feed power into the electrical power grids. This change, which often is referred to as the modernization shift, is depicted in Figure 1.1.

The increased penetration of distributed energy resource (DER) has lead to important changes in distribution systems Eto et al. (2015):

- The first change occurs on direction of power flows which is no longer unidirectional. This change may have crucial consequences on planning and operations of distribution grids including protection and voltage management applications.
- Since DERs provide great quantities of generation in a grid, if a sudden load disturbance occurs, this sudden change can lead to activation of the protection systems. This can result in losing the power of many distributed generators. In this situation, control of surrounding dis-

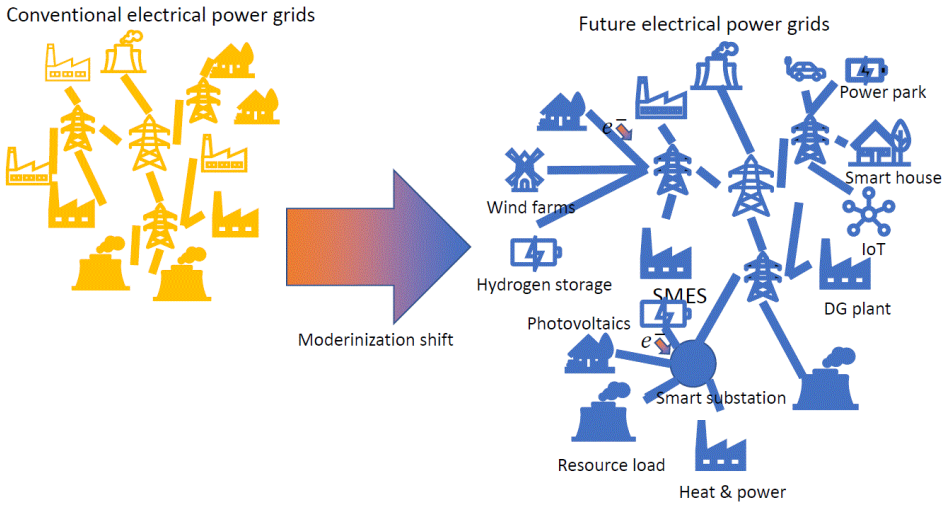


Figure 1.1: Power grids evolution

tribution systems' voltage and frequency would be more complicated than for conventional electrical grids.

- Distributed energy storage has unpredictable charge and discharge profiles. Hence, these profiles can cause unpredictable disturbances in the distribution system [1].

These effects, in addition to changes in the electrical load behavior and the requirement for reliable electricity, definitely increase the need for reliable dynamic monitoring of smart electric grids.

The distribution parts of the electrical grid typically have a low degree of instrumentation, so we need sophisticated tools to analyze and control them. To monitor and control such networks in an efficient way, the distributed SCADA (DSCADA) system and the information technology (IT) infrastructure in distribution grids have been improved by introducing new developments in sensor and communication technologies. The most advanced measurement units in distribution grids are smart meters and Phasor Measurement Units (PMUs).

[1] Although distributed energy consumers, such as Electric vehicles (EVs) technically can be exploited as grid storage when the electrical grid needs high demand and are therefore under high pressure, the mobility of EVs leads to variation in their availability unlike that of traditional, stationary storage.

GPS signals provide a globally synchronized time signal. This is being used in PMUs, developed in the 1980s, to offer dynamic measurements of voltage, current and phase angles. When a sufficient number of PMUs are placed in the electrical grids, then the system would be observable and as a result dynamic state estimation procedures such as Kalman filtering-based approaches can be utilized to estimate the system states Gomez-Exposito et al. (2011). However, replacing all the meters in an electric grid with PMUs is not feasible due to practical reasons, including their cost. Thus, proper monitoring and control of the electric grid requires the effective combination of information from PMUs and other sensors in the grid. Such information combination is commonly termed *sensor fusion*.

Furthermore, in order to handle the computational challenge of efficiently handling the significant volume of data generated by smart meters, PMUs, and SCADA measurements, distributed processing should be applied. This is done by dividing the monitoring and control tasks of the network into the different sub-systems. Each sub-system can estimate the state of the system locally. Globally accurate monitoring will then require communication and effective coordination of local estimates between sub-systems.

1.2 Background

Several methods have been proposed for state estimation of power grids, including hierarchical, as well as partially and fully distributed methods. In Van Cutsem and Ribbens-Pavella (1983) a hierarchical state estimation method for electrical power systems is presented. Early examples of research into distributed state estimation in electric power grids include Baran and Kelley (1994; 1995). A parallel and distributed state estimation was proposed in Falcao et al. (1995) for the case when the states of the entire system are observable locally, whereas a distributed approach for real-time state estimation of power grids is given in Khan et al. (2008) without the local observability assumption. An agent-based state estimation, as well as topology identification, observability analysis, and bad data detection, are considered in Nordman and Lehtonen (2005). Then, in Korres (2011), two-level state estimation for multi-area power system (MASE) has been studied in which distributed state estimates calculated at the primary level have coordination and communication with the top-level through PMUs. A detailed survey of MASE methods in power systems can be found in Gomez-Exposito et al. (2011). The robust distributed state estimation for discrete-time and time-varying nonlinear systems with stochastic parameters is investigated in Ding et al. (2012). The concept of distributed state estimation using sensor

networks and its convergence condition is studied in Wang and Ren (2017). In Sun et al. (2016), a real-time distributed state estimation approach for a network of linear systems using a distributed maximum a posteriori (MAP) estimation technique is investigated. Their approach coordinates local state estimates through communication with neighbors.

In accordance with the philosophy of design, these methods consider the power system operates in its steady-state or quasi-steady-state operating condition. However, as explained above, this contrasts with the new reality in modern power systems, which rarely operate near steady state for long periods, because variations in loads and distributed generators (DGs) are usually fast and unpredictable. Consequently, the assumption of a steady-state is dubious, and methods of static state estimation (SSE) can not capture dynamic effects of fast-changing operating conditions. Therefore, it seems vital to re-evaluate the SSE methods Schweppe and Wildes (1970), Abur and Exposito (2004) utilized in current power systems' energy management systems (EMS) and also to provide improved monitoring techniques based on dynamic state estimation (DSE) in modern EMS. DSE's ability to capture the dynamics of a fast-changing system is precisely what makes it suitable for power system control and protection Modir and Schlueter (1981), Meliopoulos et al. (2017), Liu et al. (2017), Fan et al. (2015), Cui et al. (2017).

In the existing literature on DSE for power systems, a number of techniques have been proposed: the extended Kalman filter (EKF) Ariff and Pal (2016), Albinali and Meliopoulos (2016), Farantatos et al. (2016), the iterated EKF (IEKF) Vittal et al. (2011), the unscented Kalman filter (UKF) NERC (2016), Huang et al. (2013), Zhao and Mili (2019b) and the particle filter (PF) Zhao and Mili (2018b;a), Chow (2013). Most of these techniques, however, suffer from some major drawbacks, preventing them from being implemented for real-time applications by power utilities.

1. They assume that a complete model of the system is available.
2. Unknown inputs are assumed to have immediate effects only locally.
3. They can not manage uncertainty and parameter errors in dynamic models.
4. They are not able to cope with cybersecurity problems.

There are a number of reasons for these shortcomings. First, in addition to the uncertainties caused by the implicit simplifications of the statistical

equations on which these methods are based, there may be many other noises due to unknown system inputs, including parametric uncertainties and actuator failures Chen and Patton (2012); unavailable inputs, such as unrecognized mechanical power, exciter field voltage, unknown fault location; and inaccuracies of the model variables of the synchronous generators, loads, lines, and transformers. Furthermore, due to privacy and protection concerns, fault occurrence, lack of sensors, and degradation of power equipment, certain parts of the network could be totally unknown to the local operator. A number of studies in this field have focused on cases where the unknown inputs are unknown local variables Ghahremani and Kamwa (2011a;c; 2016), Anagnostou and Pal (2018), Zhao et al. (2019c). In Abooshahab et al. (2019), it is stated that this unknown input/disturbance comes from a change in system-wide variables, and therefore cannot be considered a local variable; then, a new modeling strategy for a power system is derived to tackle the problem of facing system-wide unknown disturbance. A number of state estimation methods have focused on a condition that system model contain unknown inputs , including the extended state observers (ESOs) Chen et al. (2015), Chen et al. (2015), the augmented Kalman filter (AKF) Anderson and Moore (2012), Simon (2006), simultaneous input and state estimation (SISE) Gillijns and De Moor (2007a) and its Kalman filtering formulation (KF-SISE) which recently presented in Bitmead et al. (2019). Due to the common practice that the disturbance dynamics are selected as an integrator, the application of AKF is mostly limited to estimation of slow and constant signals e.g., parameters in power networks Plett (2004), Ritter et al. (2019), Bian et al. (2011).

Second, the current DSE methods assume that both the process and the observation noise of the nonlinear dynamic system models are Gaussian. Nevertheless, several studies conducted by Pacific Northwest National Laboratory (PNNL) Sauer et al. (2017), Crow (2015) have shown that PMU measurement errors of voltage and current magnitude obey non-Gaussian probability distributions Zhao (2018). The existence of non-Gaussian noise leads to a need for more development in the field of robust dynamic state estimation of power systems.

Third, as we stated before, there exist different sensors with different sampling rates in power networks, and it is essential to extract all the information in the power network from all the available sensors. Thus, a multi-rate framework for state estimation of power networks seems vital.

Additionally, cyber-attacks have become a significant concern due to the high dependence of smart grid functions on communications networks. They

are typically classified as *bias injection* attack, *denial of service* attack and *replay* attack Shivakumar and Jain (2008). Bias injection attack occurs when an adversary attempts to corrupt the content of either the measurement or the control signals; for example, the man-in-the-middle intercepts the PMU measurement signals and corrupts them with significant biases. Denial of service attack occurs when the actuator or sensor data can not reach their respective destinations, resulting lack of data for the DSE; replay attack happens when a hacker first conducts a disclosure attack over a certain period of time, collecting data sequences, and then starting to replace the data over a certain period of time. This type of attack may lead to misinterpreting the true state of the system, leading to inappropriate (and possibly dangerous) control and operational actions. Mitigating the impacts of cyber-attacks is obviously an important task that results in the need for developing fault detection methods Zhao (2018).

Lastly, PMUs are expensive devices, so it is crucial to know where to install them. Hence, developing methods for optimal PMU placement for state estimation purposes is also an essential issue for electrical grids.

1.3 Research Objectives and contributions

Improving the efficiency, stability, and resilience of electrical power systems depends on the availability of secure, reliable, and robust estimators of dynamic states. These estimators should handle unknown inputs and model uncertainties. Figure 1.2 shows the scope of covered topics in this thesis for power networks as further described below. In order to achieve the aims of this thesis, we have proposed a general theoretical framework for robust dynamic state estimation of a partially known system with unknown inputs and robust statistical theory and robust control theory. In particular, the principal contribution of this thesis is described in the following:

- A unified framework is developed for robust and stable simultaneous input and state estimation. These new estimators can handle both local and system-wide unknown inputs and can be used both for centralized and decentralized state estimation. To cope with system-wide unknown inputs, we develop and apply the stable SISE (S-SISE) algorithm Abooshahab et al. (2019; 2021), which is specifically targeted towards joint input and state estimation. Then we introduce a new modeling scheme for power systems containing unknown parts, called partially known power systems (PKPSs). Then, we apply SISE/S-SISE to estimate the states of the PKPS and its unknown inputs.

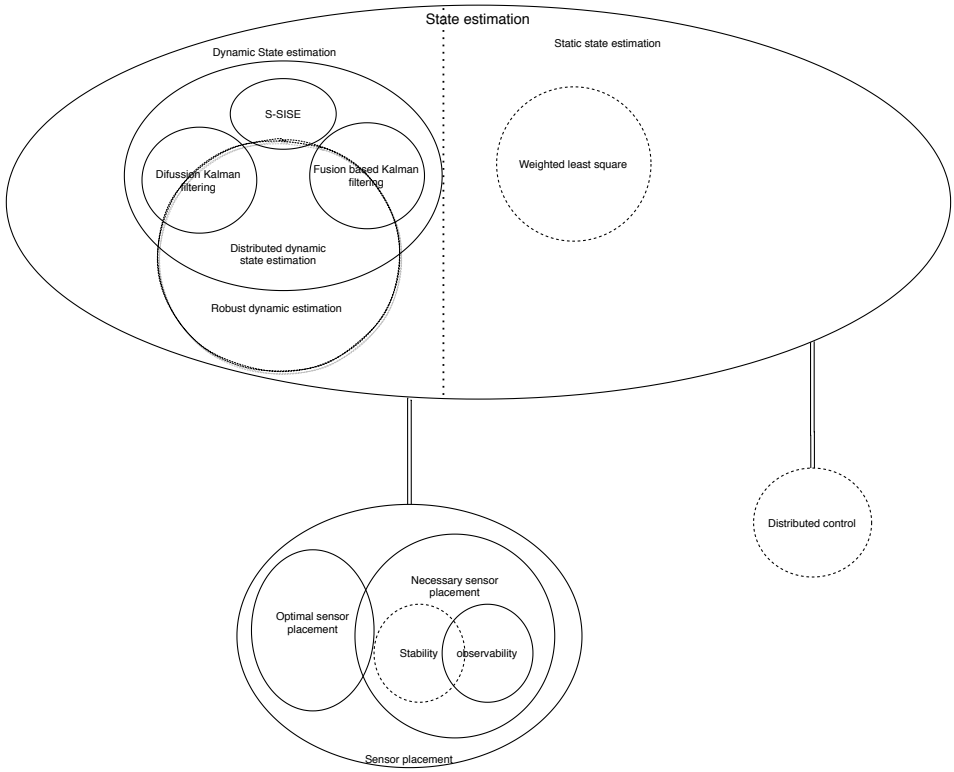


Figure 1.2: Scope diagram of covered topics

Our procedure for state estimation consists of the following steps:

1. Identify the known and the unknown parts of the system.
2. Model the unknown parts by disturbance inputs interacting with the known part.
3. Derive the PKPS model.
4. Examine the observability and recoverability conditions.
5. Apply SISE/S-SISE to estimate the states and the unknown inputs.

- We obtain the information matrix version of S-SISE and then propose 'best in' and 'worst out' greedy algorithms to optimize the placement of PMUs in a power network.

- Diffusion Kalman filtering is used to obtain a nonlinear distributed

state estimation algorithm for power networks and its stability is investigated.

- Observability assumptions are investigated and relaxed.
- Multirate state estimation is proposed using the covariance intersection that shows good consistency in widely different noise situations.
- As we will see, a precise local bus frequency is important for stabilizing the power grids. In this thesis, we also try to find the estimate of dynamical characteristics of bus frequency without having information about the models of the loads but having enough knowledge of swing equation parameters, admittance matrix and measurements.
- In order to obtain more robust distributed filtering method, we develop a nonlinear distributed/diffusion H-infinity filter using the unscented transformation (UT).

1.4 Thesis outline

This thesis is organized as follows. A general dynamic model for different components of a power network is given in in Chapter 2. A general introduction to various dynamic state estimation methods, including simultaneous input and state estimation, H_∞ filtering and diffusion Kalman filtering is provided in Chapter 3.

Chapters 4-11 presents the main contributions of the thesis. Each of these chapters is based on papers that are either published or submitted.

Chapter 12 describes additional topics on state estimation and its use as a basis for control design. Chapter 13 contains conclusions and proposals for further work.

1.5 List of Publications

Peer Reviewed Journal Papers

1. Robert R. Bitmead, Morten Hovd, and Mohammad Ali Abooshahab. "A Kalman-filtering derivation of simultaneous input and state estimation." *Automatica* 108 (2019): 108478.
2. Mohammad Ali Abooshahab, Mohammed M.J. Alyaseen, Robert R. Bitmead, Morten Hovd. "Simultaneous input & state estimation, singular filtering and stability." submitted to and provisionally accepted in *Automatica* (2020).

Peer Reviewed Conference Papers

1. Mohammad Ali Abooshahab, Morten Hovd, Edmund Brekke, Xianfeng Song. "A Covariance Consistent Data Fusion method for Power Networks with Multirate Sensors." 2020 IEEE Conference on Control Technology and Applications (CCTA).

2. Mohammad Ali Abooshahab, Morten Hovd, and Robert R. Bitmead. "Disturbance and State Estimation in Partially Known Power Networks." 2019 IEEE Conference on Control Technology and Applications (CCTA).

3. Mohammad Ali Abooshahab, Morten Hovd. "Distributed H_∞ Filtering for Linear and Nonlinear Systems." submitted to and accepted in the 2021 IEEE Conference on Control Technology and Applications (CCTA).

4. Mohammad Ali Abooshahab, Morten Hovd, Giorgio Valmorbida. "Optimal PMU placement for partially known power system dynamic state estimation" submitted to and accepted in the IEEE PES ISGT Europe 2021.

Journal Papers in Preparation

1. Mohammad Ali Abooshahab, Morten Hovd, and Robert R. Bitmead. "Monitoring Disturbances and States in Partially Known Power Systems." In preparation for submission to *IEEE Transactions on Power Systems* (2021).

2. Mohammad Ali Abooshahab, Morten Hovd. "Multi-Rate Distributed Unscented Kalman Filtering with Application to Power System Monitoring." In preparation for submission to *IEEE Transactions on Smart Grids* (2021).

Chapter 2

Model derivation

Various types of generators, loads, transmission networks, measuring devices, distributed generations (renewable energy sources), as well as monitoring and control centers are primary components of a modern power system Sadamoto et al. (2019). Different parts in a power system shares the power via transmission/distribution lines. Thus, electric power transmission/distribution systems are interconnected systems Sadamoto et al. (2019). Power systems are subject to continuous perturbation. Examples of typical disruptions in power systems are sudden changes in loads, losses of one or more transmission lines, system configuration modification, equipment outages, generator failures, and connection/disconnection of renewable energy sources.

Reliable and stable operation of electrical grids is of paramount importance for modern society. This requires effective monitoring of the state of the electrical grid. Any approach to monitoring a complex system requires a good understanding of system behaviour. Systematic monitoring approaches are generally based on a mathematical model of the system in question. This chapter will describe how the models used throughout this thesis for state estimation and monitoring are derived.

As we will discuss later, the modeling for different parts of a power system is not unique and is depending on the application of monitoring and control and the validation of simplifying assumptions. Throughout this Chapter, we provide general modelling details for different parts of a power system in order to derive a unified interconnected system model for a power system. However, as noted above, Chapters 4-11 are based on published or submitted

papers. Several of these papers describe the derivation of the models used therein. This causes some repetition, and readers may choose to skip the modelling sections of these chapters except for the simplifying assumptions and purpose of monitoring.

2.1 Dynamic distributed component-based modeling

In this section, we describe the dynamic models of the main components of a modern electrical grid including production, distribution, consumption, batteries and also their interactions. The energy production can be categorized into traditional generators such as synchronous or induction generators, and DERs, such as wind parks and PVs. Throughout the thesis, the terms *generators*, *wind park*, and PV apply for a combination of whole units forming the overall component. That is, we are concerned with how an entire wind park interacts with and affects the electrical grid, this thesis is not concerned with the behaviour of individual windmills. In the same manner, the term *load* refers to the combination of all energy consuming units connected to a bus in the system.

Note that an electrical power grid often has buses with no generators, DERs, loads, or energy storage devices. These buses are called non-unit buses Sadamoto et al. (2019) which means no energy unit is connected to these buses.

For clarity, we need to define the term 'component' which in this thesis means as either a non-unit bus, or a generator, DER, load or storage device connected to its bus (see Figure 2.1).

An overall dynamic model for the kth component of a power system (either for DERs, loads, energy consumption, non-unit bus and ...) is given as follows Sadamoto et al. (2019).

$$\mathfrak{C}_k : \begin{array}{c} \dot{x}_k = f_k(x_{\mathrm{k}}, (|\boldsymbol{V}|\angle\boldsymbol{\theta})_k, u_k, p_k)), \\ \mathbf{P}_k + j\mathbf{Q}_k = h_k(x_{\mathrm{k}}, (|\boldsymbol{V}|\angle\boldsymbol{\theta})_k, u_k, p_k)), \end{array} \tag{2.1}$$

where $\dot{x}_k$, $(|\boldsymbol{V}|\angle\boldsymbol{\theta})_k$, u_k and p_k represent the time derivative of the state, the voltage phasor, the input and the parameters of the of the kth component, respectively. Additionally, active and reactive power injected from the kth component is shown by P_k and Q_k Sadamoto et al. (2019).

For an interconnected power system, one may use the Krone reduction method Dorfler and Bullo (2013) to obtain a reduced model for the power

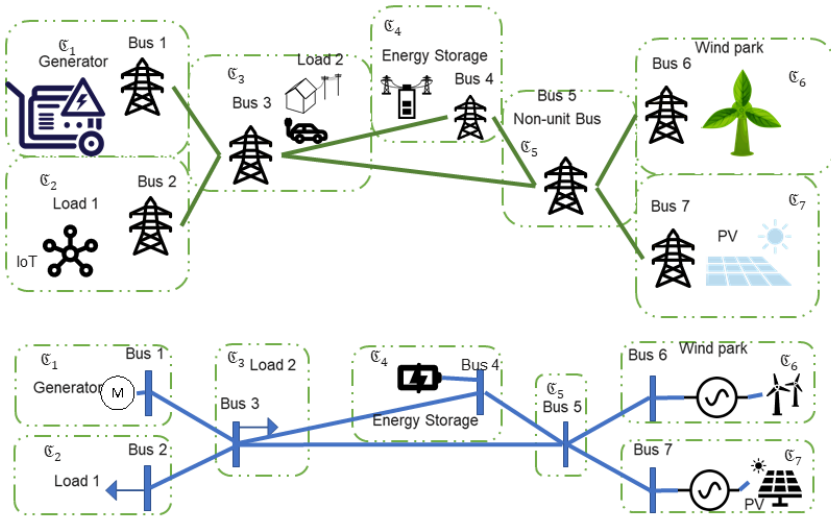


Figure 2.1: Different components of a modern electrical grid and its on-line digaram

network containing dynamic components only, resulting in the reduced admittance of the network $\boldsymbol{Y}^{red}$. Whence, the concepts of 'buses', 'nodes', and 'components' will be used interchangeably. The N components form an interconnected network, we can obtain the following equation using circuit theory Sadamoto et al. (2019):

$$0^N = (\boldsymbol{Y}^{red}(|\boldsymbol{V}|\angle\boldsymbol{\theta})_{1:N})^{\cdot *} \circ (|\boldsymbol{V}|\angle\boldsymbol{\theta})_{1:N} - (\mathbf{P}_{1:N} + j\mathbf{Q}_{1:N}), \tag{2.2}$$

where $\circ$ and $.*$ are the element-wise multiplication and conjugation operators, respectively. For $i \in \{1, \ldots, N\}$, the $(|\boldsymbol{V}|\angle\boldsymbol{\theta})_i, P_i$, and Q_i are accumulated denoting by $(|\boldsymbol{V}|\angle\boldsymbol{\theta})_{1:N}$, $P_{1:N}$, and $Q_{1:N}$ Sadamoto et al. (2019). For a measured or calculated $P_{1:N}$ and $Q_{1:N}$, $(|\boldsymbol{V}|\angle\boldsymbol{\theta})_{1:N}$ can be obtained by using (2.2). The model for the power system is the aggregation of (2.1) and (2.2).

When the system is operating without sudden changes, it reaches its equi-

librium point, which can be obtained as follows.

$$\begin{aligned} 0 &= f_k\left(x_{\mathrm{k}}^S, (|\boldsymbol{V}|\angle\boldsymbol{\theta})_k^S, u_k^S, p_k\right), \\ P_k^S + jQ_k^S &= \left(x_{\mathrm{k}}^S, (|\boldsymbol{V}|\angle\boldsymbol{\theta})_k^S, u_k^S, p_k\right) \end{aligned} \tag{2.3}$$

where superscript s denotes the steady state value.

For finding the equilibrium points for an interconnected power network, we need to consider the power flow equations. Power flow calculation is briefly the process of finding $(|\boldsymbol{V}|\angle\boldsymbol{\theta})_{1:N}^S$, $P_{1:N}^S$, and $Q_{1:N}^S$ satisfying (2.2), and then substitute them in (2.3) to find x_{k}^S and p_k. When sudden changes happen in the system, the power flow solutions can be used as initial conditions for the electrical grid (2.1)-(2.2). From now, we omit the subscript k for simplicity.

2.1.1 Generators

The generator can be assumed as an interconnected system containing different parts, including the synchronous machine, the power supply system (or the prime mover) and the excitation system Kundur et al. (1994). The excitation system produces currents in its winding to magnetize the rotor. In the generated magnetic field, the rotor starts to rotate as a result of mechanical power provided by the primary mover. The synchronous machine converts mechanical energy and transfers it to the grid. Because of its long time constant, the dynamics of the primary mover are typically ignored Dib et al. (2009), Tsolas et al. (1985), Machowski et al. (2020).

Synchronous Machine

For the synchronous machine, there exist several models to capture its dynamics Kundur et al. (1994). Here, we choose the flux-decay model which is widely used for synchronous machines. Then we will explain how and based on what assumptions it is possible to simplify this model for practical purposes. For obtaining this model, we need to consider the electromechanical swing and electromagnetic voltage dynamics Sadamoto et al. (2019).

Excitation System

Usually, the excitation system consists of various parts including an automatic voltage regulator (AVR) which controls the level of the generator voltage to the desired level, the exciter which provides excitation for the generator, and a power system stabilizer (PSS) whose function it is to maintain the stability of the power system through damping oscillations resulting from the generator swings.

The variables and parameters in the following sections are given in Table 2.1.

Table 2.1: Variables and parameters for the generator units.

State variable	External variables	Parameters
angular frequency ω	mechanical power P_m^S	governor gain κ_a
	AVR reference input u	PSS gain κ_{PS}
E EMF (q-axis)	field voltage setpoint V_f^S	
shaft angle δ	voltage setpoint $\|\mathbf{V}\|^S$	rotor damping R
V_f Field voltage		rotor inertia M
ψ PSS state		d-axis time constant $\mathfrak{T}_{\mathrm{d}}$,
		exciter time constant $\mathfrak{T}_e$,
		PSS time constant $\mathfrak{T}_{PS}$,
		$\mathfrak{T}_{J1}, \mathfrak{T}'_{J1}, \mathfrak{T}_{J2}, \mathfrak{T}'_{J2}$
		(PSS lead-lag controller for first (J_1) and second (J_2) stage).
		d- and q-axis reactance χ_{d}, χ_{q}

Swing model:

For obtaining the state-space representation of the overall generator model, we start with the swing equations.

$$\begin{aligned} \dot{\delta} &= \omega, \\ \dot{\omega} &= \tfrac{1}{M}\left(P_{\mathrm{m}}^S - R\omega - \tfrac{|\mathbf{V}|E}{\chi'_{\mathrm{d}}}\sin\left(\delta_{\mathbf{V}}\right) + \tfrac{|\mathbf{V}|^2}{2}\left(\tfrac{\chi_{\mathrm{q}}-\chi'_{\mathrm{d}}}{\chi'_{\mathrm{d}}\chi_{\mathrm{q}}}\right)\sin\left(2(\delta_{\mathbf{V}})\right)\right) \end{aligned} \tag{2.4}$$

where $\delta_{\mathbf{V}} := \delta - \angle\mathbf{V}$.

The electromagnetic dynamics are as follows Sadamoto et al. (2019):

$$\begin{aligned} \dot{E} &= \tfrac{1}{\mathfrak{T}_{\mathrm{d}}}\left(V_{\mathrm{f}} + \tfrac{\chi_{\mathrm{d}}}{\chi'_{\mathrm{d}}}(\,|\,\mathbf{V}\,|\,\cos\left(\delta_{\mathbf{V}}\right) - E) - |\,\mathbf{V}\,|\,\cos\left(\delta_{\mathbf{V}}\right)\right), \\ P + jQ &= \\ &\tfrac{E|\mathbf{V}|}{\chi'_{\mathrm{d}}}\sin(\delta_{\mathbf{V}}) - \tfrac{|\mathbf{V}|^2}{2}\left(\tfrac{\chi_{\mathrm{q}}-\chi'_{\mathrm{d}}}{\chi'_{\mathrm{d}}\chi_{\mathrm{q}}}\right)\sin\left(2(\delta_{\mathbf{V}})\right) + j\tfrac{E|\mathbf{V}|}{\chi'_{\mathrm{d}}}\cos\left(\delta_{\mathbf{V}}\right). \end{aligned} \tag{2.5}$$

The model for an exciter (with AVR) is given as follows:

$$\dot{V}_{\mathrm{f}} = \frac{1}{\mathfrak{T}_{\mathrm{e}}}\left(V_{\mathrm{f}}^S - V_{\mathrm{f}} + \kappa_{\mathrm{a}}\left(|\mathbf{V}| - |\mathbf{V}|^S - v + u\right)\right), \tag{2.6}$$

In (2.6), the input signal is u, and it provides the reference voltage for the AVR. In control theory, the traditional PSS can be considered as a feedback controller with the input as the rotor angular frequency and the output as the reference voltage for the AVR, v defined in (2.7) Sadamoto et al. (2019), Chow et al. (2004).

Note that there are different models for the PSS, especially in modern networks with the development in designing them. We consider the linear model for the PSS given in (2.7).

$$\dot{\psi} = A_{\mathrm{PS}}\psi + B_{\mathrm{PS}}\omega,\ v = C_{\mathrm{PS}}\psi + D_{\mathrm{PS}}\omega, \tag{2.7}$$

where

$$A_{\mathrm{PS}} = \begin{bmatrix} -\frac{1}{\mathfrak{T}_{\mathrm{PS}}} & 0 & 0 \\ -\frac{\kappa_{\mathrm{PS}}}{\mathfrak{T}_{\mathrm{PS}}\mathfrak{T}_{\mathrm{J1}}}\left(1-\frac{\mathfrak{T}'_{\mathrm{J1}}}{\mathfrak{T}_{\mathrm{J1}}}\right) & -\frac{1}{\mathfrak{T}_{\mathrm{J1}}} & 0 \\ -\frac{\kappa_{\mathrm{PS}}}{\mathfrak{T}'_{\mathrm{J1}}}\mathfrak{T}_{\mathrm{PS}}\mathfrak{T}_{\mathrm{J1}}\mathfrak{T}_{\mathrm{J2}}\left(1-\frac{\mathfrak{T}'_{\mathrm{J2}}}{\mathfrak{T}_{\mathrm{J2}}}\right) & \frac{1}{\mathfrak{T}_{\mathrm{J2}}}\left(1-\frac{\mathfrak{T}'_{\mathrm{J2}}}{\mathfrak{T}_{\mathrm{J2}}}\right) & -\frac{1}{\mathfrak{T}_{\mathrm{J2}}} \end{bmatrix}, \tag{2.8}$$

$$B_{\mathrm{PS}} = \begin{bmatrix} \frac{1}{\mathfrak{T}_{\mathrm{PS}}} \\ \frac{\kappa_{\mathrm{PS}}}{\mathfrak{T}_{\mathrm{PS}}\mathfrak{T}_{\mathrm{J1}}}\left(1-\frac{\mathfrak{T}'_{\mathrm{J1}}}{\mathfrak{T}_{\mathrm{J1}}}\right) \\ \frac{\kappa_{\mathrm{PS}}\mathfrak{T}'_{\mathrm{J1}}}{\mathfrak{T}_{\mathrm{PS}}\mathfrak{T}_{\mathrm{J1}}\mathfrak{T}_{\mathrm{J2}}}\left(1-\frac{\mathfrak{T}'_{\mathrm{J2}}}{\mathfrak{T}_{\mathrm{J2}}}\right) \end{bmatrix}, \tag{2.9}$$

$$C_{\mathrm{PS}} = \left[1 - \frac{\kappa_{\mathrm{PS}}\mathfrak{T}'_{\mathrm{J1}}\mathfrak{T}'_{\mathrm{J2}}}{\mathfrak{T}_{\mathrm{PS}}\mathfrak{T}_{\mathrm{J1}}\mathfrak{T}_{\mathrm{J2}}}\frac{\mathfrak{T}'_{\mathrm{J2}}}{\mathfrak{T}_{\mathrm{J2}}}\right], \tag{2.10}$$

$$D_{\mathrm{PS}} = \frac{\kappa_{\mathrm{PS}}\mathfrak{T}'_{\mathrm{J1}}\mathfrak{T}'_{\mathrm{J2}}}{\mathfrak{T}_{\mathrm{PS}}\mathfrak{T}_{\mathrm{J1}}\mathfrak{T}_{\mathrm{J2}}}. \tag{2.11}$$

Based on above dervations, the state space of a generator can be obtained by comparing (2.4) with (2.1).

Simplifications in Synchronous Machine Modeling

There exist several versions of models for synchronous generators, and selecting a proper model is based on the required accuracy Kundur et al. (1994). These models can be categorized as Park, sub-transient, one-axis, and classical versions. The relationship between these models is outlined next. The Park model is the famous synchronous machine model, which is an aggregation of electromagnetic and motion states showing the flux difference of circuits based on the d-axis and q-axis , the excitation and damper windings Kundur et al. (1994).

The approximation of the Park model can be obtained , when the flux states of the d-axis and q-axis circuit vary quickly. This model is called the sub-transient model that includes four coils on the rotor Pal and Chaudhuri (2006)The states of this model contain the motion and the excitation winding flux, d, and q-axis damper windings states.

When the damping effects can be ignored, and the connection line between the generator and its corresponding bus has insignificant resistance, then

the one-axis model can be obtained as in (2.4) and (2.5) Dib et al. (2009). This model only captures the flux changes of the excitation winding.

A further simplification assumption is that $\chi_q = \chi'_d$ in (2.4) and (2.5). Assuming that $\chi'_d = \chi_q$ (this is a reasonable assumption due to the symmetrical air gap in round rotor machines Machowski et al. (2020)), then $\mathfrak{T}_{do}\dot{E} = -E + V_f$. When the initial value of E is equal to E^S and $V_f(t) \equiv E^S$, then $E(t) \equiv E^S$. Thus, this model which is called the classical model can be derived as follows Kundur et al. (1994), Machowski et al. (2020):

$$\begin{aligned} \dot{\delta} &= \omega, \\ \dot{\omega} &= \tfrac{1}{M}\left(P_m - R\omega - \tfrac{|\mathbf{V}|E^S}{\chi}\sin(\delta - \angle\mathbf{V})\right). \end{aligned} \tag{2.12}$$

For the non-unit buses, we have

$$P + jQ = 0.$$

Algebraic power balance equations are commonly used as models for loads, although extensive literature also exists for dynamic loads Hiskens and Milanovic (1995), Hill (1993). The static loads can be modeled as Sadamoto et al. (2019):

$$\begin{aligned} \text{impedance model} \quad & P + jQ = \left(\bar{\mathbf{z}}^{-1}\mathbf{V}\right) * \mathbf{V}, \\ \text{current model} \quad & P + jQ = \bar{\mathbf{i}} * \mathbf{V}, \\ \text{power model} \quad & P + jQ = \bar{P} + j\bar{Q}, \end{aligned}$$

where $\bar{\mathbf{z}}$ is the given or obtained impedance parameter. Note that $\bar{\cdot}$ is used for constant parameters.

The synchronous motor is modeled in the same manner as a synchronous generator. The only difference is that, instead of a prime mover providing mechanical torque input to the generator, the motor drives a mechanical load Kundur et al. (1994).

2.1.2 Wind Turbine

Using wind as a source of energy source dates back to the 9th century, where the first windmill and wind pump were developed in Iran Glick et al. (2014). Contemporary wind turbines can be seen as a natural development of those early windmills, obviously with the difference that instead of pumping water they are equipped with electrical generators and are thus able to supply

green energy to the grid. Wind turbines convert wind power to mechanical power and then transfer the power to a doubly-fed induction generator (DFIG). The turbines are commonly designed with a two-inertia composition with a low-speed shaft, gearbox, and a high-speed shaft Sadamoto et al. (2019). Generally, changes in wind speed are not so fast compared with the grid frequency, so the aerodynamic power P_{ae} can be considered constant. The electrical components of a wind turbine contain different parts contributing to its dynamic including doubly-fed induction generator (DFIG), DC link, Back-to-Back (B2B) converter (the rotor-side converter (RSC) and the grid-side converter (GSC)) with its controllers, buck and boost converter and battery.

We will next describe the model for the wind turbine. The symbols used are explained in Table 2.2.

The wind turbine dynamics are given as follows:

$$\begin{aligned} J_{\mathrm{ls}}\dot{\omega}_{\mathrm{ls}} &= -(d_{\mathrm{c}} + B_{\mathrm{ls}})\omega_{\mathrm{ls}} + \tfrac{d_{\mathrm{c}}}{N_{\mathrm{ge}}}\omega_{\mathrm{hs}} - \kappa_{\mathrm{c}}\theta_{\mathrm{T}} + \tfrac{P_{\mathrm{ae}}}{\omega_{\mathrm{ls}}}, \\ J_{\mathrm{hs}}\dot{\omega}_{\mathrm{hs}} &= \tfrac{d_{\mathrm{c}}}{N_{\mathrm{ge}}}\omega_{\mathrm{ls}} - \left(\tfrac{d_{\mathrm{c}}}{N_{\mathrm{ge}}^2} + B_{\mathrm{hs}}\right)\omega_{\mathrm{hs}} + \tfrac{\kappa_{\mathrm{c}}}{N_{\mathrm{ge}}}\theta_{\mathrm{T}} - \chi_{\mathrm{t'}}\left(i_{\mathrm{ds}}i_{\mathrm{qr}} - i_{\mathrm{qs}}i_{\mathrm{dr}}\right), \\ \dot{\theta}_{\mathrm{T}} &= \bar{\omega}_{\mathrm{m}}\left(\omega_{\mathrm{ls}} - \tfrac{1}{N_{\mathrm{ge}}}\omega_{\mathrm{hs}}\right). \end{aligned} \tag{2.13}$$

.

The equations related to DFIG is given as follows:

$$\begin{aligned} \dot{i} &= A_{\mathrm{i}}^{\mathrm{DF}}(\omega_{\mathrm{hs}})i + G_{\mathrm{i}}^{\mathrm{DF}}[\mathbb{R}\mathrm{e}(\mathbf{V}), \mathbb{I}\mathrm{m}(\mathbf{V})]^\top + B_{\mathrm{i}}^{\mathrm{DF}}[v_{\mathrm{dr}}, v_{\mathrm{qr}}]^\top, \\ P_{\mathrm{s}} + jQ_{\mathrm{s}} &= \gamma_\Omega\left(\mathbb{R}\mathrm{e}(\mathbf{V})i_{\mathrm{ds}} + \mathbb{I}\mathrm{m}(\mathbf{V})i_{\mathrm{qs}}\right) + j\gamma_\Omega\left(\mathbb{I}\mathrm{m}(\mathbf{V})i_{\mathrm{ds}} - \mathbb{R}\mathrm{e}(\mathbf{V})i_{\mathrm{qs}}\right), \end{aligned}$$

$$i := [i_{\mathrm{dr}}, i_{\mathrm{qr}}, i_{\mathrm{ds}}, i_{\mathrm{qs}}]^\top, \tag{2.14}$$

where the definition of ω_{hs} is given in (2.13), and v_{dr} and v_{qr} are determined

in (2.23), and

$$A_{\mathrm{i}}^{\mathrm{DF}}(\omega_{\mathrm{hs}}) = \frac{1}{\eta}\begin{bmatrix} -R_{\mathrm{hs}}\chi_{\mathrm{t}} & \eta-\omega_{\mathrm{hs}}\chi_{\mathrm{t}}\chi_{\mathrm{hs}} & R_{\mathrm{t}}\chi_{\mathrm{t}'} & -\omega_{\mathrm{hs}}\chi_{\mathrm{t}}\chi_{\mathrm{t}'} \\ -\eta+\omega_{\mathrm{hs}}\chi_{\mathrm{t}}\chi_{\mathrm{hs}} & -R_{\mathrm{hs}}\chi_{\mathrm{t}} & \omega_{\mathrm{hs}}\chi_{\mathrm{t}}\chi_{\mathrm{t}'} & R_{\mathrm{t}}\chi_{\mathrm{t}'} \\ R_{\mathrm{hs}}\chi_{\mathrm{t}'} & \omega_{\mathrm{hs}}\chi_{\mathrm{hs}}\chi_{\mathrm{t}'} & -R_{\mathrm{t}}\chi_{\mathrm{hs}} & \eta+\omega_{\mathrm{hs}}\chi_{\mathrm{t}'}^2 \\ -\omega_{\mathrm{hs}}\chi_{\mathrm{hs}}\chi_{\mathrm{t}'} & R_{\mathrm{hs}}\chi_{\mathrm{t}'} & -\eta-\omega_{\mathrm{hs}}\chi_{\mathrm{t}'}^2 & -R_{\mathrm{t}}\chi_{\mathrm{hs}} \end{bmatrix}, \tag{2.15}$$

$$B_{\mathrm{i}}^{\mathrm{DF}} = \frac{1}{\eta}\begin{bmatrix} -\chi_{\mathrm{t}} & 0 \\ 0 & -\chi_{\mathrm{t}} \\ \chi_{\mathrm{t}'} & 0 \\ 0 & \chi_{\mathrm{t}'} \end{bmatrix}, \tag{2.16}$$

$$G_{\mathrm{i}}^{\mathrm{DF}} = \frac{1}{\eta}\begin{bmatrix} \chi_{\mathrm{t}'} & 0 & -\chi_{\mathrm{hs}} & 0 \\ 0 & \chi_{\mathrm{t}'} & 0 & -\chi_{\mathrm{hs}} \end{bmatrix}^{\top}, \tag{2.17}$$

$$\chi_{\mathrm{t}} := \chi_{\mathrm{t}'} + \chi_{\mathrm{ls}},\ \chi_{\mathrm{hs}} := \chi_{\mathrm{t}'} + \chi_{\mathrm{lr}}, \eta := \chi_{\mathrm{t}}\chi_{\mathrm{hs}} - \chi_{\mathrm{t}'}^2. \tag{2.18}$$

For the GSC, we have

$$\begin{aligned} &\tfrac{L_G}{\bar{\omega}}\dot{i}_{\mathrm{Gd}} = -R_G i_{\mathrm{Gd}} + L_G i_{\mathrm{Gq}} + \mathbb{R}\mathrm{e}\left(\mathbf{V}\right) - \tfrac{t_{\mathrm{Gd}}}{2} v_{\mathrm{bias}}, \\ &\tfrac{L_G}{\bar{\omega}}\dot{i}_{\mathrm{Gq}} = -R_G i_{\mathrm{Gq}} - L_G i_{\mathrm{Gd}} + \mathbb{I}\mathrm{m}\left(\mathbf{V}\right) - \tfrac{t_{\mathrm{Gq}}}{2} v_{\mathrm{bias}}, \\ &P_{\mathrm{hs}} + jQ_{\mathrm{hs}} = \gamma_\Omega\left(\mathbb{R}\mathrm{e}\left(\mathbf{V}\right) i_{\mathrm{Gd}} + \mathbb{I}\mathrm{m}\left(\mathbf{V}\right) i_{\mathrm{Gq}}\right) + j\gamma_\Omega\left(\mathbb{I}\mathrm{m}\left(\mathbf{V}\right) i_{\mathrm{Gd}} - \mathbb{R}\mathrm{e}\left(\mathbf{V}\right) i_{\mathrm{Gq}}\right), \end{aligned} \tag{2.19}$$

where definition of t_{Gd} and t_{Gq} are given in (2.23), and v_{bias} is defined in (2.27).

$$\begin{cases} \dot{\psi}_{\mathrm{Gd}} = \kappa_{\mathrm{I,dG}}(v_{\mathrm{bias}} - v_{\mathrm{bias}}^S), \\ i_{\mathrm{Gd}}^{\mathrm{ref}} = \kappa_{\mathrm{P,dG}}(v_{\mathrm{bias}} - v_{\mathrm{bias}}^S) + \psi_{\mathrm{Gd}}, \end{cases} \qquad \begin{cases} \dot{\psi}_{\mathrm{Gq}} = \kappa_{\mathrm{I,qG}}(Q_{\mathrm{hs}} - Q_{\mathrm{hs}}^S), \\ i_{\mathrm{Gq}}^{\mathrm{ref}} = \kappa_{\mathrm{P,qG}}(Q_{\mathrm{hs}} - Q_{\mathrm{hs}}^S) + \psi_{\mathrm{Gq}}, \end{cases} \tag{2.20}$$

where the definition of Q_{hs} and v_{bias} are given in (2.19) and (2.27).

$$\begin{aligned} &\mathfrak{T}_G\dot{\chi}_{\mathrm{Gd}} = i_{\mathrm{Gd}}^{\mathrm{ref}} - i_{\mathrm{Gd}}, \\ &t_{\mathrm{Gd}} = \mathrm{sat}\left(\tfrac{2}{v_{\mathrm{bias}}}\left(u_{\mathrm{Gd}} + \mathbb{R}\mathrm{e}(\mathbf{V}) + L_G i_{\mathrm{Gq}} - R_G\chi_{\mathrm{Gd}} - \tfrac{L_G}{\bar{\omega}\mathfrak{T}_G}(i_{\mathrm{Gd}}^{\mathrm{ref}} - i_{\mathrm{Gd}})\right)\right), \end{aligned} \tag{2.21}$$

$$\begin{aligned} &\mathfrak{T}_G\dot{\chi}_{\mathrm{Gq}} = i_{\mathrm{Gq}}^{\mathrm{ref}} - i_{\mathrm{Gq}}, \\ &t_{\mathrm{Gq}} = \mathrm{sat}\left(\tfrac{2}{v_{\mathrm{bias}}}\left(u_{\mathrm{Gq}} + \mathbb{I}\mathrm{m}(\mathbf{V}) - L_G i_{\mathrm{Gd}} - R_G\chi_{\mathrm{Gq}} - \tfrac{L_G}{\bar{\omega}\mathfrak{T}_G}(i_{\mathrm{Gq}}^{\mathrm{ref}} - i_{\mathrm{Gq}})\right)\right), \end{aligned} \tag{2.22}$$

where i_{Gd} and i_{Gq} are defined in (2.19), $i_{\mathrm{Gd}}^{\mathrm{ref}}$ and $i_{\mathrm{Gq}}^{\mathrm{ref}}$ in (2.20), and v_{bias} in (2.27).

v_{dr} and v_{qr} related to RSC are given bellow:

$$v_{\mathrm{dr}} = \frac{t_{\mathrm{Rd}}}{2} v_{\mathrm{bias}}, v_{\mathrm{qr}} = \frac{t_{\mathrm{Rq}}}{2} v_{\mathrm{bias}}, \tag{2.23}$$

where the definition of t_{Rd} and t_{Rq} are given in (2.25), and v_{bias} is determined in (2.27).

$$i_{\mathrm{dr}}^{\mathrm{ref}} = \kappa_{\mathrm{P,Rd}} \left(|\mathbf{V}| - |\mathbf{V}|^S \right), \; i_{\mathrm{qr}}^{\mathrm{ref}} = \kappa_{\mathrm{P,Rq}} \left(\omega_{\mathrm{hs}} - \omega_{\mathrm{hs}}^S \right), \tag{2.24}$$

where the definition of ω_{hs} is given in (2.13).

$$\begin{aligned} \dot{\chi}_{\mathrm{Rd}} &= \kappa_{\mathrm{I,Rd}} \left(i_{\mathrm{dr}} - i_{\mathrm{dr}}^{\mathrm{ref}} \right), \\ t_{\mathrm{Rd}} &= \mathrm{sat} \left(\tfrac{2}{v_{\mathrm{bias}}} \left(\kappa_{\mathrm{P,Rd}} \left(i_{\mathrm{dr}} - i_{\mathrm{dr}}^{\mathrm{ref}} \right) + \chi_{\mathrm{Rd}} + u_{\mathrm{Rd}} \right) \right), \end{aligned} \tag{2.25}$$

$$\begin{aligned} \dot{\chi}_{\mathrm{Rq}} &= \kappa_{\mathrm{I,Rq}} (i_{\mathrm{qr}} - i_{\mathrm{qr}}^{\mathrm{ref}}), \\ t_{\mathrm{Rq}} &= \mathrm{sat} \left(\tfrac{2}{v_{\mathrm{bias}}} \left(\kappa_{\mathrm{P,Rq}} (i_{\mathrm{qr}} - i_{\mathrm{qr}}^{\mathrm{ref}}) + \chi_{\mathrm{Rq}} + u_{\mathrm{Rq}} \right) \right), \end{aligned} \tag{2.26}$$

where i_{dr} and i_{qr} are defined in (2.14), $i_{\mathrm{dr}}^{\mathrm{ref}}$ and $i_{\mathrm{qr}}^{\mathrm{ref}}$ in (2.24), and v_{bias} in (2.27).

The voltage of the DC link can be obtained using (2.27).

$$\begin{aligned} \frac{C_{\mathrm{bias}}}{\bar{\omega}} \dot{v}_{\mathrm{bias}} = \frac{1}{2 v_{\mathrm{bias}}} &\left(\mathbb{R}\mathrm{e} \left(\mathbf{V} \right) i_{\mathrm{Gd}} + \mathbb{I}\mathrm{m} \left(\mathbf{V} \right) i_{\mathrm{Gq}} + v_{\mathrm{dr}} i_{\mathrm{dr}} + v_{\mathrm{qr}} i_{\mathrm{qr}} \right. \\ &\left. - R_G (i_{\mathrm{Gd}}^2 + i_{\mathrm{Gq}}^2) \right) - G_{\mathrm{sw}} v_{\mathrm{bias}} + \frac{1}{2} i_{\mathrm{bias}}, \end{aligned} \tag{2.27}$$

where i_{Gd} and i_{Gq} are defined in (2.19), v_{dr} and v_{qr} in (2.23), i_{dr} and i_{qr} in (2.14), and i_{bias} in (2.28). When the battery and dc/dc are not connected, $i_{\mathrm{bias}} = 0$.

The model for the buck-and-boost dc/dc converter is given as follows:

$$v'_{\mathrm{bias}} = e(Z + u_Z) v_{\mathrm{bias}}, \tag{2.28}$$

$$i_{\mathrm{bias}} = e(Z + u_Z) i'_{\mathrm{bias}}, \tag{2.29}$$

$$e\left(x\right) = \begin{cases} x & \text{if } x \geq 0, \\ 0 & \text{ow} \end{cases} \tag{2.30}$$

where the definition of v_{bias} and i'_{bias} are given in (2.27) and (2.31).

The battery's voltage v_{b} and the current injected from the battery i'_{bias} form the state space of a battery for the wind turbine.

$$\begin{aligned} \tfrac{C_{\mathrm{b}}}{\bar{\omega}} \dot{v}_{\mathrm{b}} &= -i'_{\mathrm{bias}} - G_{\mathrm{b}} v_{\mathrm{b}}, \\ \tfrac{L_{\mathrm{b}}}{\bar{\omega}} \dot{i}'_{\mathrm{bias}} &= v_{\mathrm{b}} - R_{\mathrm{b}} i'_{\mathrm{bias}} - v'_{\mathrm{bias}}, \end{aligned} \tag{2.31}$$

where the definition of v'_{bias} is given in (2.28).

The interconnection of a wind turbine to the grid can be obtained by using (2.32).

$$P + jQ = (P_{\mathrm{s}} - P_{\mathrm{hs}}) + j\,(Q_{\mathrm{s}} - Q_{\mathrm{hs}})\,, \tag{2.32}$$

where the definition of P_{s} and Q_{s} are given in (2.14) and P_{hs} and Q_{hs} in (2.19).

The model for a wind farm containing all of the above units can be concisely given as

$$x_{\mathrm{k}} := \left[\omega_{\mathrm{ls}}, \omega_{\mathrm{hs}}, \theta_{\mathrm{T}}, i^{\top}, i_{\mathrm{G}}^{\top}, \chi_{\mathrm{G}}^{\top}, \psi_{G}^{\top}, \chi_{\mathrm{R}}^{\top}, v_{\mathrm{bias}}, v_{\mathrm{b}}, i'_{\mathrm{bias}}\right]^{\top}, \tag{2.33}$$

$$\psi_{\mathrm{G}} := [\psi_{\mathrm{Gd}}, \psi_{\mathrm{Gq}}]^{\top}, \quad i_{\mathrm{G}} := [i_{\mathrm{Gd}}, i_{\mathrm{Gq}}]^{\top}, \tag{2.34}$$

$$u := [u_{\mathrm{Gd}}, u_{\mathrm{Gq}}, u_{\mathrm{Rd}}, u_{\mathrm{Rq}}, u_{\mathrm{S}}]^{\top}, \tag{2.35}$$

$$p := \left[v_{\mathrm{bias}}^{S}, Q_{\mathrm{hs}}^{S}, |\mathbf{V}|^{S}, \omega_{\mathrm{hs}}^{S}\right]^{\top}, \tag{2.36}$$

and $f(\cdot,\cdot,\cdot;\cdot)$ and $h(\cdot,\cdot;\cdot)$ follow from (2.13)-(2.32). Note that interconnection between different components are formed based on connecting lines including transmission/distribution lines. These lines transfer power between different units. Hence the intercation between different units can be obtained from the power flow equations given in (2.2).

2.1.3 Measurements in a power system

The measurement vector $\mathbf{y}$ for a power system is based on available sensors in that system. The main sensor for power systems that is fast enough to capture dynamics of a power network is the PMU. PMUs provide the phase and the magnitude of voltage and current at the installation bus, power and reactive power in all connected lines and also angular frequency at the installation bus. There also exist other measuring devices in a power system such as SCADA sensors which mainly measure the active and reactive power at the installation bus. Other units in a power network including windparks also have their own sensors and transducers providing information and measurement for the power system operators.

2.1.4 Overview of models used in different chapters

The detailed procedure to derive dynamic characteristics of different components in a power network is discussed above, which means each component is modeled based on its mechanical and electrical characteristics. Then each component's interconnection with other parts of the network is obtained through power flows in the power system, as it can be seen from equation (2.1). However, we should bear in mind that power system models are not unique, and the purpose of the monitoring/estimation can dictate what level of detail and what range of dynamics is included in the model.

In Chapters 4-7, a linear 4th order model for the synchronous generators are considered in order explain the proposed state estimation technique and to investigate its stability and observability properties. After that, we extend our framework with considering the PSS model for the synchronous generators (Chapter 12.5).

In Chapter 10, we use the classic nonlinear model for the synchronous generators (2.12). The measurement model in that section is based on PMUs, but with different sampling rates. The goal in Chapter 10 is to fuse the state estimates obtained from different sensors with different sampling rates. We also compare our method which is based on covariance intersection with the existing methods such as those proposed in Ghosal et al. (2017), Ghosal and Rao (2019).

In Chapter 9 (the diffusion and H_∞ filtering papers), we start with the classic nonlinear model for the synchronous generators (2.12) to develop and explain our method. Furthermore, we extend our framework for a distribution grid containing a wind park.

Table 2.2: Variables and parameters for the units of a wind farm.

State variable	External variables	Parameters
low speed frequency ω_l		inertia coefficients: J_{ls}, J_{hs},
high speed frequency, ω_r	wind power P_a	friction coefficients: B_{ls}, B_{hs} , damping coefficient d_c
angle of torsion θ_{T}	Steady state value of Q_r^S	Gear ratio N_g
rotor and stator currents: i_{dr}, i_{qr}, i_{ds}, i_{qs}	DC link voltage v_{bias}^S	torsional stiffness K_c
GSC currents i_{Gd}, i_{Gq},	reference signals $i_{\mathrm{Gd}}^{\mathrm{ref}}$, $i_{\mathrm{Gq}}^{\mathrm{ref}}$	number of turbines γ_{w}
GSC inner controller states: χ_{Gd}, χ_{Gq},		
RSC inner controller states: χ_{Rd}, χ_{Rq}	input signals u_{Gd}, u_{Gq}	Rotor leakage reactance X_{ls} , stator leakage reactance, X_{lr}
outer controller states ψ_{Gd}, ψ_{Gq}	reference signals $i_{\mathrm{dr}}^{\mathrm{ref}}$, $i_{\mathrm{qr}}^{\mathrm{ref}}$	stator resistance R_{s}
DC side voltage v_{bias}	input signals u_{Rd}, u_{Rq}	rotor resistance R_{hs} PI gains of outer: $K_{\mathrm{P,Gd}}$, $K_{\mathrm{P,Gq}}$, $K_{\mathrm{I,Gd}}$, $K_{\mathrm{I,Gq}}$,
battery voltage v_b		
battery current i'_{bias}		GSC time constant $\mathfrak{T}_G$
		PI gains of inner: $K_{\mathrm{P,Rd}}$, $K_{\mathrm{P,Rq}}$, $K_{\mathrm{I,Rd}}$, $K_{\mathrm{I,Rq}}$,
		DC link capacitance C_{bias}
		switching conductance G_{sw}
		battery capacity C_{b}
		battery conductance, G_{b}
		battery resistance R_{b}
		battery inductance L_{b}
		magnetizing reactance χ_{t},
		DFIG stator reactance χ_{t}
		DFIG rotor reactance χ_{hs}

Chapter 3

State estimation

3.1 Introduction

Whenever the dynamical states of a system cannot be directly measured, they have to be estimated based on measurements and a system model. This process is known as *state estimation*. Depending on the system model and the frequency of the measurements, state estimation can be performed dynamically or statically. A well-known static state estimation method is the weighted least square (WLS) method Gomez-Exposito and Abur (2004), Ule (1955), while the Kalman filter (KF) Kalman (1960) is widely used for dynamic state estimation.

The KF is constructed as a mean squared error (MSE) minimizer and is optimal (in the MSE sense) if the underlying assumptions are fulfilled Simon (2006):

- the system is linear, and the system model is accurate,
- the noises entering the system are Gaussian,
- the noise covariances are known.

A KF fulfilling these assumptions will have a Gaussian output prediction error. A non-Gaussian prediction error is an indication that the state estimate is not optimal. Although few systems fulfill the above assumptions perfectly, the KF has been found to perform well in a wide variety of applications Simon (2006).

Effective operation of many systems including power systems Zhao et al. (2019a;c), Sadamoto et al. (2019), chemical processes Yin et al. (2018), Guo and Huang (2015), complicated robots Kulic and Croft (2007), Lin et al. (2006) will typically require accurate information about the dynamic states of the system. Such accurate information is often not directly available, either because the required information is not directly measured, or because available measurements are too noisy. In these cases, the filtered information for operation and management of a system should be obtained by state estimation.

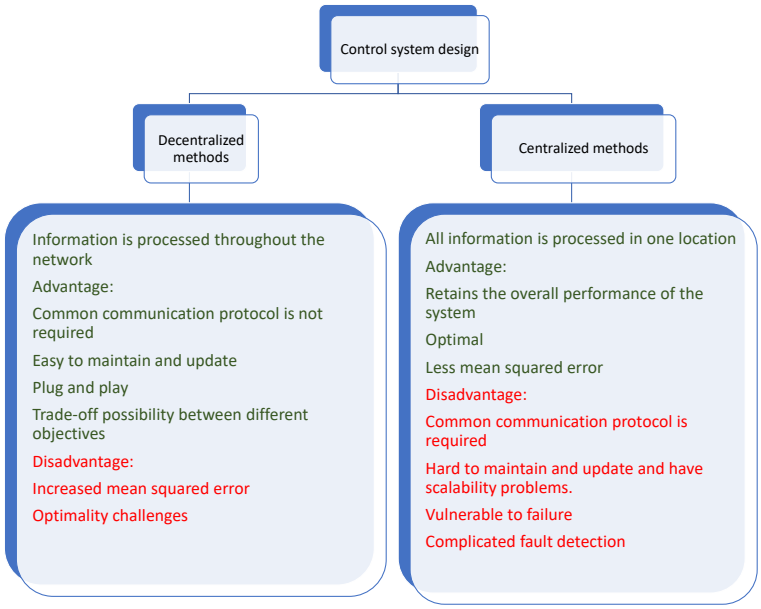


Figure 3.1: A comparison of centralized and decentralized estimation methods.

As illustrated in Fig. 3.1, dynamic state estimation algorithms can be divided into two subcategories: centralized and decentralized algorithms. In centralized algorithms, the measurements are gathered in a single measurement vector, and this vector is sent to a centralized estimator for computing an estimate of the global state vector. A high computational cost, vulnerability to failures and complicated communication protocols are the main challenges of these algorithms.

To address these problems, the decentralized filter variants can be used. Each local subsystem can calculate local estimates, so that the computational load decreases by splitting the centralized problem into smaller sub-

problems that can be solved in parallel. The obtained estimates based on local measurements in a local estimator can be sent to a centralized fusion center to obtain the global fused estimates. This strategy is known as centralized fusion method. Alternatively, local estimates can be sent between two or more local sub-systems using various strategies to estimate the local and global state estimates. In these methods, the local subsystems and their local estimators are known as *agents*, and their coordination will play a crucial role in estimating the states. Such methods are referred as distributed state estimation.

Different estimation algorithms such as moving horizon estimation Gao et al. (2018) and robust estimation algorithms Cao et al. (2010)-Lin et al. (2019) are presented in literature for decentralized state estimation. The strong properties of the centralized KF has lead to a significant volume of research into decentralized versions of different variants of the KF. This has led to decentralized versions of the extended Kalman filter (EKF) Salahshoor et al. (2008), unscented Kalman filter (UKF) Qing et al. (2015), cubature Kalman filter (CKF) Ge et al. (2014) and sigma-point Kalman filter (SPKF) Vercauteren and Wang (2005).

However, in modern industrial systems, some variables and parameters within a part of a system might be partially or totally unknown due to privacy and security considerations, fault occurrence, lack of sensors, and degradation of industrial devices.At the same time, the reliability, stability and functionality of industrial interconnected systems are highly dependent on monitoring of their dynamics Zhao et al. (2019a). Therefore, it is crucial to develop a method that performs state estimation of the known part while considering these knowledge limitations. There exist several state estimation methods with unknown inputs in the literature. The Extended State Observers (ESOs) is discussed in Chen et al. (2015), Chen et al. (2015) and the Augmented Kalman filter (AKF) in Simon (2006), Friedland (1969), Park et al. (2000). However, the major drawback of these methods is that they require a model for the unknown input. With the common practice of choosing the disturbance dynamics as integrators, the AKF is mainly used to estimate slow or even constant signals such as parameters in networked systems Plett (2004), Ritter et al. (2019), Bian et al. (2011). A two-stage state estimation method with unknown inputs is developed in Hou and Muller (1992), Hou and Patton (1998). In the first stage, an alternative formulation for a system decoupled from the unknown inputs is derived, and in the second stage, a minimum-variance unbiased (MVU) estimator for this decoupled system is designed.

Another approach is the simultaneous input and state estimation (SISE) Kitanidis (1987), Gillijns and De Moor (2007a). This method is an optimal and MVU filter which was developed firstly in Kitanidis (1987). In Gillijns and De Moor (2007a), this method was extended to simultaneous input and state estimation (SISE) for systems without direct feedthrough.

3.2 Simultaneous input and state estimation

Simultaneous Input and State Estimation (SISE) is an algorithm which takes a system with an unknown input, $\{d_t\}$,available output, $\{y_t\}$, and conceivably a known input signal, $\{u_t\}$, and tries to generate estimates of both the sequence of unknown input signal d_t and the state of the system x_t, using fixed-lag smoothing as it is shown in Bitmead et al. (2019).

3.2.1 The mathematical description of SISE

Although SISE formulation can be obtained for both linear Gillijns and De Moor (2007a), Yong et al. (2016) and nonlinear time-varying systems Fang et al. (2013), Kim et al. (2020), we consider the linear time-invariant system with zero known control input for clarity

$$x_{t+1} = Ax_t + Gd_t + w_t, \tag{3.1}$$
$$y_t = Cx_t + Hd_t + v_t, \tag{3.2}$$

where $x_t \in \mathbb{R}^n$, $d_t \in \mathbb{R}^m$, $y_t \in \mathbb{R}^p$. The zero-mean Gaussian noises $\{w_t\}$ and $\{v_t\}$ are uncorrelated. These are also independent from $\{d_t\}$ and x_0. The covariances of w_t and v_t are $Q \geq 0$ and the $R > 0$, respectively. The aggregated measurements up to and including time t are denoted by $\mathbf{Y}^t \triangleq \{y_t, y_{t-1}, \ldots, y_0\}$. The goal is to generate a state estimate, $\hat{x}_{t|t}$, and an unknown input filtered estimate, $\hat{d}_{t|t+1}$, by using $\mathbf{Y}^t$, and considering the properties of G and $\begin{bmatrix} C & H \end{bmatrix}$. The following assumption is required.

Assumption 3.1. *System (3.1-3.2) has* $[A, C]$ *observable,* $rank G = m$, $[A, Q]$ *reachable, and* $R > 0$.

SISE algorithm for full-rank direct feedthrough, namely $\text{rank}(H) = m$, is investigated in Gillijns and De Moor (2007a); SISE algorithm for zero direct feedthrough, i.e. $H = 0$, is treated in Gillijns and De Moor (2007a) with $\text{rank}(CG) = m$; furthermore, Yong et al. (2016) gives an extension, ULISE, with aggregated rank assumptions for H and CG.

3.2.2 SISE algorithm for zero direct feedthrough

When $H = 0$ in (3.2), SISE algorithm for zero direct feedthrough from Gillijns and De Moor (2007a) is the following recursion

$$X_t = AP_{t-1}A^T + Q, \tag{3.3}$$
$$K_t = X_t C^T (CX_t C^T + R)^{-1}, \tag{3.4}$$
$$M_t = [G^T C^T (CX_t C^T + R)^{-1} CG]^{-1} \times G^T C^T (CX_t C^T + R)^{-1}, \tag{3.5}$$
$$P_t = (I - K_t C)\left[(I - GM_t C)X_t \times (I - GM_t C)^T + GM_t R M_t^T G^T\right] + K_t R M_t^T G^T, \tag{3.6}$$
$$\hat{d}_{t-1|t} = M_t(y_t - CA\hat{x}_{t-1|t-1}), \tag{3.7}$$
$$\hat{x}_{t|t} = A\hat{x}_{t-1|t-1} + G\hat{d}_{t-1|t} + K_t \times (y_t - CA\hat{x}_{t-1|t-1} - CG\hat{d}_{t-1|t}). \tag{3.8}$$
$$\operatorname{cov}(x_t|\mathbf{Y}^t) = P_t, \tag{3.9}$$

considering the following condition.

Assumption 3.2.

$$rank\, CG = m. \tag{3.10}$$

An apparent finding would be that SISE does not utilize any description of the unmeasured disturbance model d_t. Therefore, it is often claimed that there is no model for the signal $\{d_t : t = 0, 1, \dots\}$. Obviously, Assumption 3.2 needs $p \geq m$ and $\operatorname{rank} C \geq \operatorname{rank} G = m$.

3.2.3 SISE algorithm for non-zero direct feedthrough

For $H \neq 0$ in (3.2), SISE algorithm for non-zero direct feedthrough can be obtained subject to the following, for the time-invariant case Gillijns and De Moor (2007a).

Assumption 3.3. *Rank* $H = m$.

Considering this assumption, the SISE algorithm for non-zero direct feed-

through, for the time-invariant system (3.1-3.2) is

$$\hat{x}_{t|t-1} = A\hat{x}_{t-1|t-1} + G\hat{d}_{t-1|t-1}, \tag{3.11}$$

$$P^x_{t|t-1} = \begin{bmatrix} A & G \end{bmatrix} \begin{bmatrix} P^x_{t-1|t-1} & P^{xd}_{t-1|t-1} \\ P^{dx}_{t-1|t-1} & P^d_{t-1|t-1} \end{bmatrix} \begin{bmatrix} A^T \\ G^T \end{bmatrix} + Q,$$

$$\tilde{R}_t = CP^x_{t|t-1}C^T + R,$$

$$M_t = (H^T\tilde{R}_tH)^{-1}H^T\tilde{R}_t^{-1},$$

$$\hat{d}_{t|t} = M_t(y_t - C\hat{x}_{t|t-1}), \tag{3.12}$$

$$P^d_{t|t} = (H^T\tilde{R}_tH)^{-1},$$

$$K_t = P^x_{k|k-1}C^T\tilde{R}_t^{-1},$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(y_t - C\hat{x}_{t|t-1} - H\hat{d}_{t|t}), \tag{3.13}$$

$$P^x_{t|t} = P^x_{t|t-1} - K_t(\tilde{R}_t - HP^dH^T)K_t^T,$$

$$P^{xd}_{t|t} = \left(P^{dx}_{t|t}\right)^T = -K_tHP^d_{t|t}.$$

For $\operatorname{rank} H < m$, in Yong et al. (2016), a version of SISE is developed based on the singular value decomposition (SVD). The resulting algorithm is called the Unified Linear Input & State Estimator (ULISE). In the following section, application of SISE for state estimation of interconnected industrial systems will be explained.

3.2.4 State estimation of interconnected systems using SISE

In geophysical exploration and environmental monitoring, SISE algorithms have generally been inspired by and applied to input estimation issues as a method for deconvolving input and state signals through a linear system. The primary methods given in Friedland (1969) focused on bias estimation. However, obtaining the state estimate is a direct consequence of this deconvolution. The emphasis of this algorithm can be both on linear system state estimation, interconnected with an unknown system, and on linear system input estimation, as shown in Figure 3.2 Kitanidis (1987). Note that there is no prerequisite statistical correlation between the unknown input and the other signals. The study of interconnected systems, however, reveals the potentially important application of SISE to estimate the state in isolation from the signals connecting the system to neighboring unknown parts, which are completely grasped from the model of unknown input. Irrespective of the size of the neighboring unknown parts, the complexity of this problem is given by the size of the state vector of the known system and the number of interconnecting signals. Further, the known system is required to be linear,

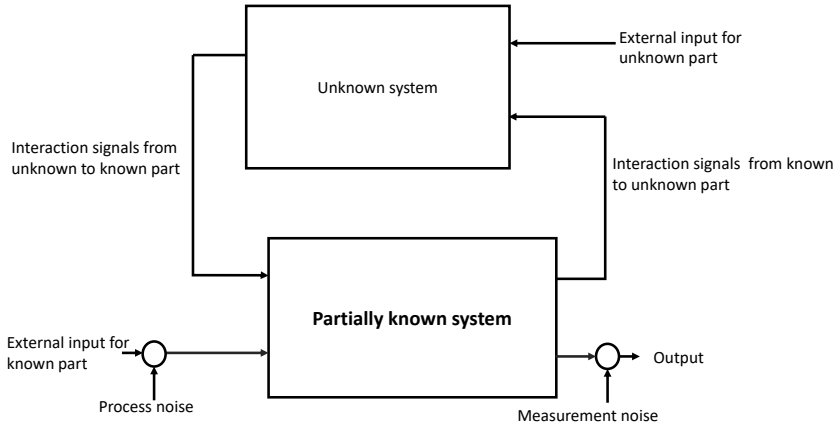


Figure 3.2: Depiction of interconnected known linear system with states need to be estimated, unknown system and unmeasured interconnection signals.

but this is not required for the unknown system. However, Assumption 3.2 has to hold in order to be able to estimate the disturbances, which depends on the topology and structure of the interconnection between the known and unknown parts. Note that we do not need knowledge of the unknown part of the system. Certainly the quality of state estimation is degraded compared to what would be achievable with access to a full model of the unknown system and sufficient measurements to make the unknown system observable[1]. A traditional approach might be to assume a model, typically lowpass or even static, for the unknown part. However, that introduces errors into the estimation of the states in the system.

As discussed above, the estimation accuracy is higher when more information is available. In many industrial systems, with improvement in information and communications technology (ICT), it is possible to share data between different regions or subsystems. In those cases, the adjoining parts of the system are no longer fully unknown. In this respect, it is important to take this communication between nodes into account and benefit from distributed algorithms. Our focus is on diffusion Kalman filtering proposed in Cattivelli and Sayed (2010).

[1] In which case the 'unknown system' would no longer be unkown.

3.3 Diffusion Kalman filtering

The class of distributed Kalman filtering algorithms proposed in Cattivelli and Sayed (2010) are based on diffusion adaptive algorithms Sayed (2014a), which enable both the measurement update and information fusion throughout the network to be applied in a single time-scale, and have become known as diffusion Kalman filters. An essential feature in the diffusion strategy is that only the state estimates, together with measurements, are shared in the network.

The diffusion Kalman filtering is chosen because each subsystem communicates only with its adjacent subsystems, without need for centralized fusion. Consequently, the diffusion Kalman filtering imposes no hierarchical structure on the estimation. In contrast, each subsystem applies the same procedure of processing data and communication (further in Chapter 9 we discuss how to synchronize local estimators if their sensors are not synchronized). This fully distributed architecture is robust to possible failures in subsystems, is scalable and can accommodate topology changes, and does not need complicated communication protocols. Importantly, it does not require performing consensus steps until convergence between measurements, which makes it appropriate for online and dynamic estimation. Last but not least, there is no limitation on the network's topology except for being fully connected Cattivelli and Sayed (2010).

3.3.1 Network Model

In keeping with the classical distributed processing approach, we consider a set of interconnected sensors, deployed for monitoring a dynamic system. Each node has access to one or more sensors, and each sensor is assigned to a single node. The nodes and their connections are modeled as the graph $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$, where the node set $\mathcal{N}$ representing the nodes and their communication links represented by the edge set $\mathcal{E}$. The neighbourhood of node l is defined as the set of nodes that node l can communicate with, including self communication, and is denoted as $\mathcal{N}_l = \{\ell_1, \ldots, \ell_{|\mathcal{N}_l|}\}$, (see Fig. 3.3) where ℓ_1 corresponds to the node itself. Finally, the cardinality of the set $\mathcal{N}_l$ is denoted by $|\mathcal{N}_l|$ with $|\mathcal{N}|$ denoting the total number of nodes in the network. In addition, all communication links are assumed to be bi-directional.

Based on these concepts, a diffusion estimation can be concisely described in two steps:

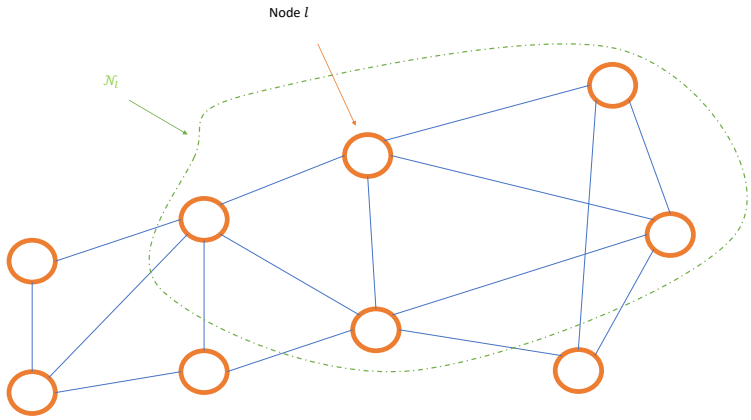


Figure 3.3: A simple network for clarifying the concept of the neighborhood, the green dotted line shows the neighbourhood of node l.

1. The nodes (as we will discuss in Chapter 9 it can be also a subsystem) share their measurements, and they use the received data with their own measurement to update their local estimates utilizing the Kalman filter's (or other state estimation methods such as recursive least square method Cattivelli et al. (2008)) measurement update equations resulting in local estimates, referred to *preestimates*.

2. Each node/subsystem sends its *preestimate*s to its neighborhood and then calculates a weighted average on all received *preestimate*s to obtain the diffused estimate as it is shown in Fig. 3.4.

Generally, in the diffusion approach, every agent should have access to the entire system state-space model and sufficient measurements to ensure observability. This clearly limits the scalability of the method. We address this problem, such that the diffusion filtering algorithm is modified to relax observability requirement to that of the centralized approach in Chapters 8 and 9.

The other drawback with diffusion Kalman filtering is that model and measurement uncertainties at some nodes can degrade the estimator's performance. In such cases, the use of robust estimators such as the H_∞ filter can enhance the performance of the estimation method by imposing a threshold for the uncertainties in the system and measurement models.

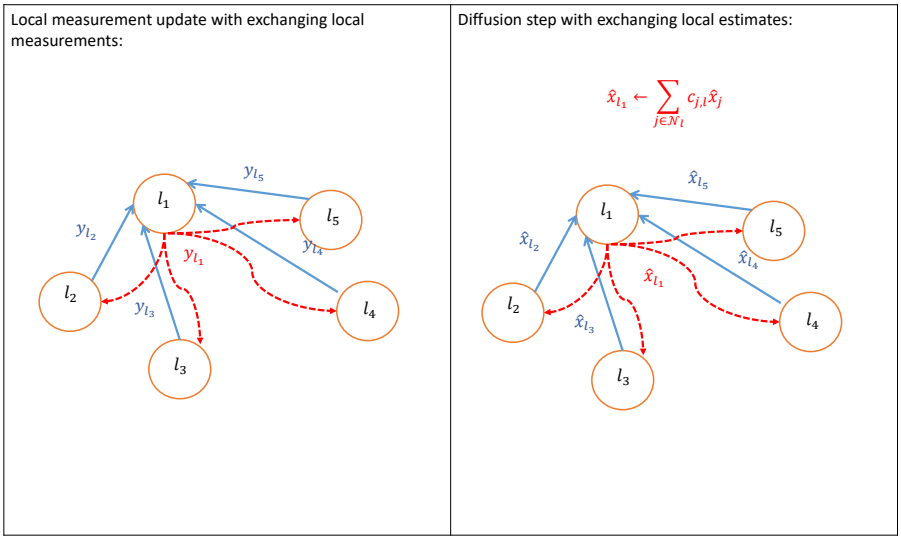


Figure 3.4: Diffusion algorithm at node l in network shown in Fig. 3.3.

3.4 Robust filtering

Consider the linear discrete-time dynamic model given in (3.14),

$$x_{t+1} = Ax_t + w_i \tag{3.14}$$

$$y_t = Cx_t + v_t \tag{3.15}$$

where at time instant t, the system state and measurement are denoted by $x_t \in \mathbb{R}^n$ and $y_t \in \mathbb{R}^m$. The system dynamics and the observation matrix are denoted by A and C, with the process noise, w_t, and the measurement noise, v_t, being energy bounded $l_2[0, +\infty)$ signals which have unknown statistical properties, so that

$$\sum_{i=0}^{\infty} w_t^{\mathsf{T}} w_t < \infty \quad \text{and} \quad \sum_{i=0}^{\infty} v_t^{\mathsf{T}} v_t < \infty. \tag{3.16}$$

The performance of the state estimation methods based on the the Kalman filtering techniques is strongly based on certain conditions Simon (2006):

1. At each moment, the mean and correlation of the process noise w_t and the measurement noise v_t and their covariances Q_t and R_t need to be available. Actually, Q_t and R_t are very important design variables for Kalman filtering approaches, and without having enough information

about them, it can be difficult to design a Kalman filter that works properly a practical application.

2. The Kalman filter is a MVU estimator which means the Kalman filter generates the estimates with the smallest deviation and variance of the estimation error (based on the Gaussian assumption for noises). For different objective functions, including minimizing the worst-case estimation error, we may with advantage use other strategies other than the Kalman filter to fulfill our goals.

3. The system model A and the measurement model C should be provided and be accurate to some extent.

In order to address above limitations, the H_∞ filter, also called the minimax filter, can be implemented. The aim is to estimate x_t on the basis of measurements prior to and including time t. Applying a game theoretic approach to the problem, nature takes the role of an adversarial player, which can in theory select any value for $\{w_t, v_t, x_0\}$ given the objectives. In contrast, in the Kalman filtering arena, the probability density functions (pdf)s of noises in the system are assumed to be known Simon (2006). Subsequently, this pdf knowledge is used to obtain a minimum variance state estimate. In this case, nature's possible actions to degrade the state estimate are constrained by the prescribed pdf's of the process and measurement noises Banavar (1992), Simon (2006).

In the H_∞ filtering problem, the following cost function is considered:

$$J_1 = \frac{\sum_{j=0}^{t} \|x_j - \hat{x}_j\|_2^2}{\|x_0 - \hat{x}_0\|_{\mathcal{P}_0^{-1}}^2 + \sum_{j=0}^{t}(\|w_j\|_{Q_j^{-1}}^2 + \|v_j\|_{R_j^{-1}}^2)} \tag{3.17}$$

where the estimate of x_t is denoted by $\hat{x}_t$, with Q_t and R_t being weighting matrices that are analogous to the covariance estimates of w_t and v_t in classical state-space filtering approaches, while $\mathcal{P}_0^{-1} > 0$ is a weighting matrix that can be used to integrate *a priori* information about the accuracy of the initial state estimate $\hat{x}_0$. The solution is to find $\hat{x}_t$ that minimizes the objective function in (3.16), while the opponent's goal is to find $\{w_t, v_t, x_o\}$ that maximize the error term $(x_t - \hat{x}_t)$ Simon (2006).

The solution of J_1 is not attainable in a straightforward fashion; therefore, an approximation method that meets a performance threshold is sought Simon (2006). In this setting, we have

$$J_1 < \lambda^2 \tag{3.18}$$

where λ is a user-specified performance bound. Substituting (3.17) into (3.18) results in

$$\begin{aligned} J = &- \lambda^2 \|x_0 - \hat{x}_0\|^2_{\mathcal{P}_0^{-1}} \\ &+ \sum_{j=0}^{t} [\|x_j - \hat{x}_j\|_2^2 - \lambda^2 (\|w_j\|^2_{Q_j^{-1}} + \|v_t\|^2_{R_j^{-1}})] < 0. \end{aligned} \tag{3.19}$$

A sub-optimal solution to the H_∞ filtering problem given in (3.19) can be reached through the following iterations Li and Jia (2010), Simon (2006)

At each time instant t:

$$\hat{x}_{t|t-1} = A\hat{x}_{t-1|t-1} \tag{3.20}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \mathcal{K}_t(y_t - C(\hat{x}_{t|t-1})) \tag{3.21}$$

$$\mathcal{P}_{t|t-1} = A\mathcal{P}_{t-1|t-1}A^\mathsf{T} + Q_t \tag{3.22}$$

$$\mathcal{K}_t = \mathcal{P}_{t|t-1}C^\mathsf{T}(R_t + C\mathcal{P}_{t|t-1}C^\mathsf{T})^{-1} \tag{3.23}$$

$$\mathcal{P}_{t|t} = \mathcal{P}_{t|t-1} - \mathcal{P}_{t|t-1} \begin{bmatrix} C^\mathsf{T} & I \end{bmatrix} R_{e,t}^{-1} \begin{bmatrix} C \\ I \end{bmatrix} \mathcal{P}_{t|t-1} \tag{3.24}$$

where $\hat{x}_{t|t-1}$ and $\hat{x}_{t|t}$ denote the *a posteriori* and *a priori* estimaties of x_t, respectively, while $R_{e,t}$ is given by

$$R_{e,t} = \begin{bmatrix} R_t & 0 \\ 0 & -\lambda^2 I \end{bmatrix} + \begin{bmatrix} C \\ I \end{bmatrix} \mathcal{P}_{t|t-1} \begin{bmatrix} C^\mathsf{T} & I \end{bmatrix}.$$

In addition, applying the matrix inversion lemma to (3.24), it can be shown that

$$\mathcal{P}_{t|t}^{-1} = \mathcal{P}_{t|t-1}^{-1} + C^\mathsf{T} R_t^{-1} C - \lambda_t^{-2} I. \tag{3.25}$$

The overall structure of the linear H_∞ filter which is very close to that of the Kalman filter are given in Fig. 3.5. We merge the idea of H_∞ filtering with the diffusion Kalman filtering to derive a robust diffusion filter for nonlinear systems. The details of our proposed algorithms are given in Chapters 8 and 9 .

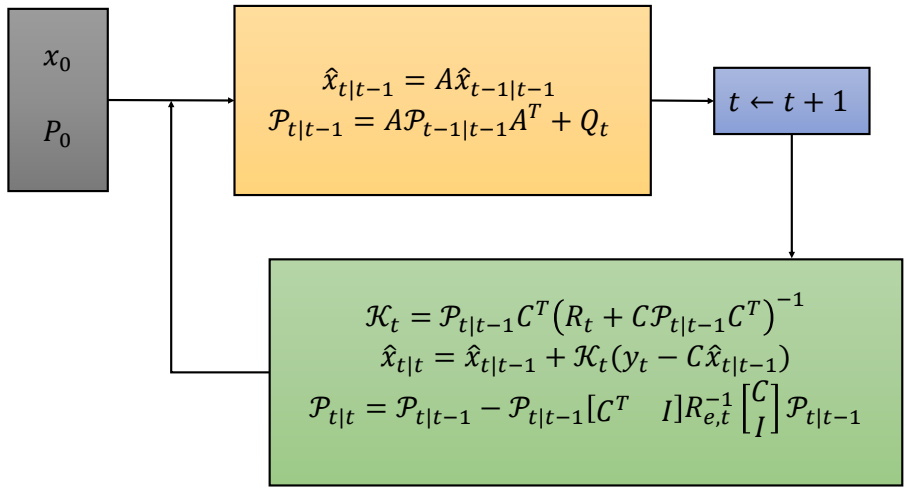


Figure 3.5: The overall structure of a H_∞ filter.

Chapter 4

Paper I: A Kalman-filtering deconstruction of simultaneous input and state estimation

This chapter is an extended version of Bitmead et al. (2019)

Robert R. Bitmead
Morten Hovd
Mohammad Ali Abooshahab
Automatica, **Vol. 108** (2019)

Abstract

Simultaneous input and state estimation algorithms are studied as particular limits of Kalman filtering problems. This admits interpretation of the algorithm properties and critical analysis of their claims to being partly model-free and to providing unbiased estimates. A disturbance model, white noise of unbounded variance, is provided and the bias feature is shown to be a geometric projection property rather than probabilistic in nature. As a consequence of this analysis, the algorithm is connected, in the stationary case, to Algebraic Riccati equation computations for the gains, estimate covariances and filter frequency response. Lastly, by focusing on the state estimation aspects, as opposed to the usual deconvolution input signal estimation features, we are able to present a powerful and intriguing property

of the algorithms in networked system state estimation.

4.1 Introduction

The on-line Cambridge English Dictionary defines

> *deconstruction*: the act of breaking something down into its separate parts in order to understand its meaning, especially when this is different from how it was previously understood.

Our objective in this chapter is to deconstruct the Simultaneous Input and State Estimation (SISE) algorithm to permit its interpretation as Kalman filtering. By doing this, we improve understanding of the algorithm and provide a path forward to its ready application, design and extension.

The SISE algorithm has been the subject of considerable research interest since its inception Glover (1969), Sanyal and Shen (1974), Kitanidis (1987), Mendel (1977) as an input reconstruction method suited to signal recovery in environmental and geophysical linear array analysis. Kitanidis Kitanidis (1987) is generally credited with the formulation which seeks also to generate reliable state estimates. More recent works Gillijns and De Moor (2007a;b), Fang et al. (2013), Yong et al. (2016) have developed the algorithm *per se* for systems with direct feedthrough and for nonlinear problems, again with the emphasis on environmental estimation when an application is developed. The genesis of the algorithm is clearly based on least-squares linear estimation but invokes a number of properties to motivate and guide its derivation. These focus on the absence of two features: any statistical signal model for the input signal and any 'bias' in the state or input estimates. Part of our aim in this chapter is re-derive and then extend the SISE algorithm by providing a specific input signal model (curiously suggested and then dismissed by both Glover (1969) and Kitanidis (1987)) and applying standard Kalman filtering ideas. Elucidating the algorithm in this way adds to clarity in application and design and opens the door to Riccati equation calculations and informed design. Our techniques, residing in the proof of Theorem 4.1 in the Appendix, are technically more demanding and distinct from the approaches proposed in these early references.

We take the SISE algorithm class and tie it to a basis in Kalman filtering as the limit of standard problems with unbounded noise. The pursuant contributions of the chapter are as follows.

- Since the Kalman filtering background is widely understood, the algorithms

and their derivation are more easily adopted and acquired. The algorithms are not fully replaced but they are extended and provided with context and genealogy and computational tools. This admits standard design approaches and gentler variants.

- The logically problematic concept of the absence of a model for the disturbance process is addressed by deriving the algorithms with a very explicit disturbance model.
- The misappropriated estimate property *unbiased* is clarified and the algorithm interpreted via projections and constrained optimization.
- For the stationary case, we provide a Riccati equation approach to the direct computation of filter gains and estimation performance and to assist in design. This was previously absent.
- The disjunction between consideration of systems with and without direct feedthrough is avoided and the methods extended to higher degrees of smoothing and more complicated observability structure.
- There is no need to impose additional stable invariant zero or minimum-phase conditions found in Yong et al. (2016).
- The algorithm's treatment of the inputs signals is exploited to derive state estimation for networked systems in which the models for the connected systems are not required since they may be unknown.

We note that the algebraic demonstration that SISE and Kalman filter algorithms coincide proves to be more intricate, difficult and revealing than had been suggested by Glover (1969), Kitanidis (1987) in their comments. Downstream consequences follow too.

The roadmap here is to present in Section 4.2 the SISE algorithm using the clear formulation of Gillijns and De Moor (2007a) and then to follow in Section 4.3 with the parallel Kalman filtering formulation with a white-noise input, making the link to Mendel's early work Mendel (1977) in geophysical signal recovery. A sequence of algebraic identities is presented in Section 4.4 for input variance, D, finite. This establishes the strong heritage of the SISE algorithm. There follows in Section 4.5 the limiting operation, suggested in Glover (1969), Kitanidis (1987), as $D^{-1} \to 0$ to arrive at the SISE algorithm of Section 4.2. The use of algebraic Riccati equation (ARE) methods for stationary performance analysis and design are presented in Section 4.6 followed by a numerical example with large but finite D. Section 4.7 presents

a new treatment of direct feedthrough and smoothing engendered by the new connections established in this chapter. Section 4.8 provides commentary and interpretations. All of the proofs are confined to the Appendix.

4.2 The SISE algorithm

Consider the linear time-invariant system without direct feedthrough and with zero known control input,

$$x_{t+1} = Ax_t + Gd_t + w_t, \tag{4.1}$$
$$y_t = Cx_t + v_t, \tag{4.2}$$

[We take the time-invariant and zero-control system (4.1-4.2) solely for clarity in exposition. The time-varying and control-inclusive versions are direct and available in the cited references.] Make the following assumptions.

Assumption 4.1. *1.* $x_t, w_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^q$, $d_t \in \mathbb{R}^m$, $v_t, y_t \in \mathbb{R}^p$.

2. these signals are mutually independent Gaussian white noises, $w_t \sim \mathcal{N}(0, Q)$, $v_t \sim \mathcal{N}(0, R)$ *and initial condition* $x_0 \sim \mathcal{N}(\hat{x}_{0|0}, P_0)$,

3. $R_t > 0$,

4. $rank\, CG = rank\, G = m$.

Then the simultaneous input and state estimation (SISE) algorithm, developed by Kitanidis (1987) and summarized, refined and analyzed by Gillijns and De Moor (2007a), is as follows. At time t with current state estimate $\hat{x}^{\text{SISE}}_{t|t}$ with covariance P_t and measurement y_{t+1},

$$X_{t+1} = AP_tA^T + Q, \tag{4.3}$$
$$K_{t+1} = X_{t+1}C^T(CX_{t+1}C^T + R)^{-1}, \tag{4.4}$$
$$M_{t+1} = [G^TC^T(CX_{t+1}C^T + R)^{-1}CG]^{-1}G^TC^T(CX_{t+1}C^T + R)^{-1}, \tag{4.5}$$
$$\begin{aligned} P_{t+1} = {} & (I - K_{t+1}C) \\ & \left[(I - GM_{t+1}C)X_{t+1}(I - GM_{t+1}C)^T + GM_{t+1}RM_{t+1}^TG^T\right] \\ & + K_{t+1}RM_{t+1}^TG^T, \end{aligned} \tag{4.6}$$
$$\hat{d}^{\text{SISE}}_{t|t+1} = M_{t+1}(y_{t+1} - CA\hat{x}^{\text{SISE}}_{t|t}), \tag{4.7}$$
$$\hat{x}^{\text{SISE}}_{t+1|t+1} = A\hat{x}^{\text{SISE}}_{t|t} + G\hat{d}^{\text{SISE}}_{t|t+1} + K_{t+1}(y_{t+1} - CA\hat{x}^{\text{SISE}}_{t|t} - CG\hat{d}^{\text{SISE}}_{t|t+1}). \tag{4.8}$$
$$\text{cov}(x_{t+1}|\mathbf{Y}^{t+1}) = P_{t+1}, \tag{4.9}$$

The purpose of this algorithm is to take the measurement sequence, $\mathbf{Y}^{t+1} \triangleq \{y_{t+1}, y_t, \ldots, y_1\}$, and the current state estimate, $\hat{x}^{\text{SISE}}_{t|t}$, and covariance, P_t, and to produce estimates, $\hat{d}^{\text{SISE}}_{t|t+1}$ and $\hat{x}^{\text{SISE}}_{t+1|t+1}$ respectively, of the input and the state signals. The properties claimed of these estimates are as follows.

1. No model whatsoever is provided for the evolution of the disturbance sequence $\{d_t\}$, including presumably that it might depend on $x_{t+\tau}$.

2. The estimates $\hat{d}^{\text{SISE}}_{t-1|t}$ and $\hat{x}^{\text{SISE}}_{t|t}$ are 'unbiased,' viz. $\mathrm{E}(\hat{d}^{\text{SISE}}_{t-1|t}|\mathbf{Y}^t) = d_{t-1}$ and $\mathrm{E}(\hat{x}^{\text{SISE}}_{t|t}|\mathbf{Y}^t) = x_t$, regardless of the values taken by $\{d_t\}$.

3. Subject to possession of the above properties, the estimates are least mean squares Kerwin and Prince (2000), minimizing the criterion

$$J = \text{trace cov}\left(x_t|\mathbf{Y}^t\right). \tag{4.10}$$

Our aim is to demonstrate that the SISE algorithm can be derived from a standard Kalman filtering problem and the non-properties of *no model* and *unbiasedness* can be linked to assumed signal properties. To achieve this, we provide a model for the $\{d_t\}$ sequence and for its relationship with the $\{x_t\}$ sequence; we assume that d_t is a Gaussian white noise sequence independent from other signals, with finite mean, $\mathfrak{d}$, but variance, D, tending to infinity. By doing so, we are able to provide a genealogy for the SISE algorithm and to show: that the algorithm's properties of convergence and stability in the time-invariant case, established by Fang et al. (2013), follow naturally; that there are aspects of the algorithm preserved for finite D; and that the algorithm might be derived in a standard way by selecting a specific augmenting disturbance model.

4.3 Kalman filtering formulation

Make the following assumptions regarding the signals

Assumption 4.2. *The disturbance signal $d_t \sim \mathcal{N}(\mathfrak{d}, D)$ and is independent from x_0, w_τ, v_τ for all t and τ.*

Theorem 4.1. *Given system (4.1-4.2) subject to Assumptions 4.1 and 4.2, the Kalman filtering solution to input and state estimation is given as fol-*

lows, from $\hat{x}_{t|t}$ *and* P_t*.*

$$\mathcal{X}_{t+1} = AP_tA^T + GDG^T + Q, \tag{4.11}$$

$$\mathcal{K}_{t+1} = \mathcal{X}_{t+1}C^T(C\mathcal{X}_{t+1}C^T + R)^{-1}, \tag{4.12}$$

$$\mathcal{M}_{t+1} = DG^TC^T(C\mathcal{X}_{t+1}C^T + R)^{-1}. \tag{4.13}$$

$$\mathcal{P}_{t+1} = \mathcal{X}_{t+1} - \mathcal{X}_{t+1}C^T(C\mathcal{X}_{t+1}C^T + R)^{-1}C\mathcal{X}_{t+1} = (I - \mathcal{K}_{t+1}C)\mathcal{X}_{t+1}, \tag{4.14}$$

$$\begin{aligned}\hat{x}_{t+1|t+1} &= \mathrm{E}\left[x_{t+1}|\mathbf{Y}^{t+1}\right] \\ &= A\hat{x}_{t|t} + G\mathfrak{d} + \mathcal{K}_{t+1}(y_{t+1} - CA\hat{x}_{t|t} - CG\mathfrak{d}),\end{aligned} \tag{4.15}$$

$$\begin{aligned}\hat{d}_{t|t+1} &= \mathrm{E}\left[d_t|\mathbf{Y}^{t+1}\right] \\ &= \mathfrak{d} + \mathcal{M}_{t+1}(y_{t+1} - CA\hat{x}_{t|t} - CG\mathfrak{d}).\end{aligned} \tag{4.16}$$

The criterion minimized is altered from (4.10)*, which deals with* $d_t - \hat{d}_{t|t+1}$ *via the 'unbiasedness' condition, to*

$$J = \text{trace cov}\left(d_{t-1}|\mathbf{Y}^t\right) + \text{trace cov}\left(x_t|\mathbf{Y}^t\right), \tag{4.17}$$

and these covariances are given by

$$\text{cov}\left(x_{t+1}|\mathbf{Y}^{t+1}\right) = \mathcal{P}_{t+1},$$

$$\text{cov}\left(d_t|\mathbf{Y}^{t+1}\right) = D - DG^TC^T(C\mathcal{X}_{t+1}C^T + R)^{-1}CGD, \tag{4.18}$$

$$= (I - \mathcal{M}_{t+1}CG)D \triangleq \mathcal{D}_t. \tag{4.19}$$

The proof, included in the Appendix, differs from those sketched by Glover (1969) and alluded to by Kitanidis (1987). Part of our aim is to establish, in Theorem 4.2 below, that as $D^{-1} \to 0$ the two algorithms coincide. This is more algebraic in nature than probabilistic.

4.4 Identities for finite D

From the earlier definitions of matrices: X_{t+1}, $\mathcal{M}_{t+1}$, $\mathcal{X}_{t+1}$, $\mathcal{K}_{t+1}$, K_{t+1}, for finite values of D, we have the following set of sequential identities linking quantities in the Kalman filtering formulation to SISE.

Identity 4.1 (divisors)**.**

$$CX_{t+1}C^T + R = (I_p - CG\mathcal{M}_{t+1})(C\mathcal{X}_{t+1}C^T + R).$$

Identity 4.2 (innovations)**.**

$$y_{t+1} - CA\hat{x}_{t|t} - CG\hat{d}_{t|t+1} = (I_p - CG\mathcal{M}_{t+1})(y_{t+1} - CA\hat{x}_{t|t} - CG\mathfrak{d}).$$

Identity 4.3 (state update gains)**.**

$$\mathcal{K}_{t+1} = G\mathcal{M}_{t+1} + K_{t+1}(I_p - CG\mathcal{M}_{t+1}),$$
$$I - \mathcal{K}_{t+1}C = (I - K_{t+1}C)(I - G\mathcal{M}_{t+1}C).$$

Identity 4.4 (state updates)**.**

$$\begin{aligned}\hat{x}_{t+1|t+1} &= A\hat{x}_{t|t} + G\mathfrak{d} + \mathcal{K}_{t+1}(y_{t+1} - CA\hat{x}_{t|t} - CG\mathfrak{d}),\\ &= A\hat{x}_{t|t} + G\hat{d}_{t|t+1} + K_{t+1}(y_{t+1} - CA\hat{x}_{t|t} - G\hat{d}_{t|t+1}).\end{aligned}$$

Identity 4.4 establishes that the filtered state estimate updates for the finite-D Kalman filter and for SISE starting from the same values of $\hat{x}_{t|t}$ and P_t coincide when $\hat{d}_{t|t+1}$ is the same. Since the matrices $\mathcal{M}_{t+1}$ and M_{t+1} are not identical for finite D, the algorithms will differ in the $\hat{d}_{t-1|t}$ update, which is addressed by the next finite-D identity.

Identity 4.5 (disturbance update gain)**.**

$$\mathcal{M}_{t+1} = \left[D^{-1} + G^TC^T(CX_{t+1}C^T + R)^{-1}CG\right]^{-1} G^TC^T(CX_{t+1}C^T + R)^{-1}.$$

Identity 4.6 (disturbance update)**.**

$$\hat{d}_{t|t+1} = (I_m - \mathcal{M}_{t+1}CG)\mathfrak{d} + \mathcal{M}_{t+1}(y_{t+1} - CA\hat{x}_{t|t}).$$

Identity 4.7 (disturbance estimation error covariance)**.** *The covariance of* $\hat{d}_{t|t+1}$, $\mathcal{D}_t$, *satisfies*

$$\mathcal{D}_t = \left[D^{-1} + G^TC^T(CX_{t+1}C^T + R)^{-1}CG\right]^{-1}.$$

Identity 4.8 (covariances)**.**

$$\mathcal{P}_{t+1} = (I - K_{t+1}C)\left\{(I - G\mathcal{M}_{t+1}C)X_{t+1} + G\mathcal{D}_tG^T\right\}.$$

4.5 Properties when $D^{-1} \to 0$: KF→SISE

Identity 4.9 (disturbance and filter update gains)**.**

As $D^{-1} \to 0$,

$$\mathcal{M}_t \to M_t, \quad \mathcal{K}_t \to GM_t + K_t(I_p - CGM_t).$$

Lemma 4.1. *For M_{t+1} given by (4.5),*

$$M_{t+1}CG = I_m. \tag{4.20}$$

Whence, the matrices

$$M_{t+1}CG \in \mathbb{R}^{m\times m},\ CGM_{t+1} \in \mathbb{R}^{p\times p},\ GM_{t+1}C \in \mathbb{R}^{n\times n},$$

are rank m projections on $\mathbb{R}^m$, $\mathbb{R}^p$, $\mathbb{R}^n$ respectively. The range spaces are given by

$$\begin{aligned}
\mathcal{R}a\left(M_{t+1}CG\right) &= \mathbb{R}^m,\\
\mathcal{R}a\left(CGM_{t+1}\right) &= \mathcal{R}a\left(CG\right) \subseteq \mathbb{R}^p,\\
\mathcal{R}a\left(GM_{t+1}C\right) &= \mathcal{R}a(G) \subseteq \mathbb{R}^n.
\end{aligned}$$

Theorem 4.2. *In the limit that $D^{-1} \to 0$, the Kalman filtering algorithm (4.11-4.16) coincides with the SISE algorithm (4.4-4.8).*

$$\hat{d}_{t|t+1} = \hat{d}^{\text{SISE}}_{t|t+1}, \quad \hat{x}_{t+1|t+1} = \hat{x}^{\text{SISE}}_{t+1|t+1},$$

with $\operatorname{cov}\left(x_{t+1}|\mathbf{Y}^{t+1}\right) = P_{t+1} = \mathcal{P}_{t+1}$ *and*

$$\operatorname{cov}\left(d_t|\mathbf{Y}^{t+1}\right) = \left[G^TC^T(CAP_{t+1}A^TC^T + CQC^T + R)^{-1}CG\right]^{-1}.$$

We may next combine: Theorem 4.2, Lemma 4.1, (4.14) and Identity 4.8, to yield an interpretation of the SISE algorithm.

Corollary 4.1. *For the SISE algorithm, define the signals*

$$\begin{aligned}
\hat{x}^{\text{SISE}}_{t+1|t} &\triangleq (I - GM_{t+1}C)\left(A\hat{x}^{\text{SISE}}_{t|t} + G\mathfrak{d}\right),\\
\hat{x}^{\text{SISE}}_{t+1|t+\frac{1}{2}} &\triangleq \hat{x}^{\text{SISE}}_{t+1|t} + GM_{t+1}y_{t+1}.
\end{aligned}$$

Then

$$\hat{x}^{\text{SISE}}_{t+1|t+1} = (I - K_{t+1}C)\,\hat{x}^{\text{SISE}}_{t+1|t+\frac{1}{2}} + K_{t+1}y_{t+1}.$$

Corollary 4.1 deconstructs SISE into three steps.

1. A time update projected onto the null space of G.
2. An update in the range space of G.
3. A Kalman-filter-like measurement update.

This sequence decodes the SISE covariance formula (4.6). Further, but consistent, reinterpretation of SISE unbiasedness as prioritizing the input signal estimate over state estimation is examined in Section 4.8 following.

4.6 Riccati-based steady-state SISE gains, performance and design

An evident and troubling absence from SISE is the Riccati difference equation associated with recursive linear least-squares optimal estimation. While, for the stationary case, the existence of and convergence to stationary values for the SISE gains and covariances has been established by Fang et al. (2013) and others, the computation of these values is problematic without an algebraic Riccati equation connection — our numerical example of SISE iterates the gain calculations (4.3-4.6) for 500 steps, i.e. effectively to convergence, before initiating stationary-gain SISE. The algorithm performance, and hence design in terms of Q, R and perhaps D, is evaluated from these error covariances. By the same token, appreciation of the noise amplification properties of the algorithm is wanting in earlier works. We remedy that here.

For the steady-state Kalman filter version of SISE, KF-SISE, we solve the following ARE in MATLAB.

```
Sig = dare(A',C',Q+G*D*G',R);
KF = Sig*C'/(C*Sig*C'+R);
MKF = D*G'*C'/(C*Sig*C'+R);
P = (eye(n)-KF*C)*Sig;
Dd = (eye(m)-MKF*C*G)*D;
```

The ARE solution, `Sig` here, is the steady-state prediction error covariance $\mathcal{X}_\infty$ from (4.11). Variables `KF`, `MKF`, `P`, `Dd` are the Kalman filter gain $\mathcal{K}_\infty$ from (4.12), the disturbance gain $\mathcal{M}_\infty$ from (4.13), the filtered state error covariance $\mathcal{P}_\infty$ from (4.14) and the smoothed disturbance error covariance $\mathcal{D}_\infty$ from the proof of Identity 4.7, which follows, in turn, from the proof of Theorem 4.1.

The systems from $d_t \to y_t$ and from $y_t \to \hat{d}_{t-1|t}$ may be computed as follows. We show this shortly for the example.

```
fwdsysd = ss(A,G,C,0,1)
deconsys = ss((eye(n)-KF*C)*A,KF,-MKF*C*A,MKF,1)
```

Thus, their frequency responses can be plotted.

As is apparent from (4.1), allowing $D \to \infty$ immediately implies that prediction error covariance $\mathcal{X}_\infty \to \infty$, which could be a numerical problem even

though SISE gains and covariances remain finite. In the example to follow, we used Ds as large as 10^{17} without issue but settled on 10^6.

An alternative strategy would be to use the Information Filter variant of the ARE Anderson and Moore (2012) yielding directly $\mathcal{P}_\infty^{-1}$, the inverse of the filter covariance. This, however, is an ARE for large D which exhibits singular optimal control issues, likely coupled to the non-minimum phase condition appearing in Yong et al. (2016).

4.6.1 Computational example from power systems

We apply three state estimators to a power system problem, taken from Blood (2011), for a 3-bus network with controlled generators connected to buses 1 and 2 and, at bus 3, a motor with a variable consumer power load. The phase differences $\delta_{2,t}$ between buses 1 and 2 and $\delta_{3,t}$ between buses 1 and 3, determine the power flows. Only the rotational speed, $\omega_{1,t}$, of generator 1 is measured. The bus 3 relative phase, $\delta_{3,t}$ is regarded as an external input with the rest of the network being modeled as a known linear system with seven states. The complete system is simulated with a realistic non-stationary load signal applied to the motor at bus 3. This yields the true state values. We apply three state estimators:

- a Kalman filter with 7-state model for the known 3-bus network driven by $\omega_{1,t}$ and external input $\delta_{3,t}$, which is made available to this estimator alone to provide a reference estimate quality,
- the SISE state and input estimator driven by $\omega_{1.t}$ to estimate the seven states and the external input $\delta_{3,t}$,
- the Kalman filter as above, denoted as KF-SISE, for various values of D driven by $\omega_{1,t}$ to yield state and $\delta_{3,t}$ estimates.

Even though the external load signal, and therefore $\delta_{3,t}$, is highly non-stationary, the model of the network is time-invariant and we compare the performance of steady-state estimators. For the Kalman filter and KF-SISE algorithms, this is done by solving the respective AREs. For the SISE algorithm, we iterate the algorithm for 500 steps before constructing the filter with constant coefficients.

Figure 4.1 displays state and state estimates for $\delta_{2,t}$, the phase angle between buses 2 and 1. The true signal is closely tracked by the full-information Kalman filter. The state estimates from SISE and KF-SISE with $D = 10^6$ are

indistinguishable and track around the changing state. However, both estimates exhibit a strong periodicity at 1.62 radians per sample corresponding to a pole of $(I - \mathcal{K}C)A$ at this angle with magnitude 0.92.

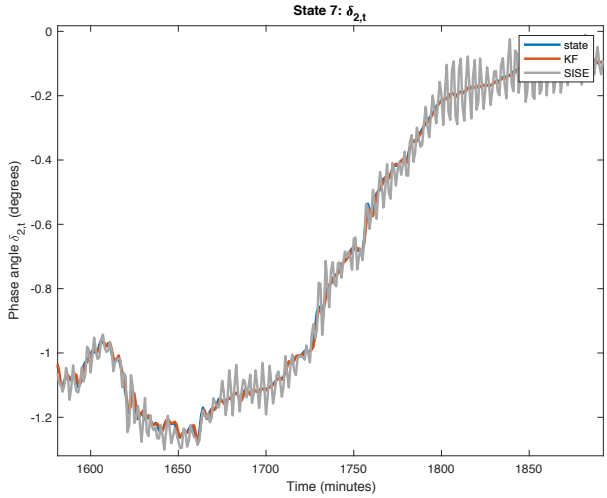


Figure 4.1: State $\delta_{2,t}$, phase difference between buses 2 and 1, in blue and its estimates with full-information Kalman filter, in red, and estimates from SISE and KF-SISE with $D = 10^6$ in gray.

Disturbance signal $\delta_{3,t}$ and effectively coincident estimates from SISE and KF-SISE for $D = 10^6$ are shown in Figure 4.2.

Figure 4.3 shows the logarithms of the sample signal norms of the error between SISE and KF-SISE estimates for states 1 to 7 and for the disturbance $\delta_{3,t}$ versus the logarithm of disturbance variance, D, for values between 1 and 10^{16}. Since SISE is not dependent on D, the glitch evident around 10^{14} is due to a numerical problem with the solution of the ARE for these choices of $Q\,(3\times10^{-2}\,I_7)$, $R\,(10^{-5})$. The subsequent recovery from this issue is a mystery. With different values of Q and R the problem disappears.

The availability of the ARE approach to the problem allows the examination of the frequency responses of the systems involved. Figure 4.4 shows the transfer functions and clearly exhibits the all-pass, pure-delay nature of the $d_t \to \hat{d}_{t-1|t}$ cascade system, which underpins the deconvolution operation. Also evident from the KF-SISE frequency response in this figure is the high-frequency noise amplification and resonant frequency around 1.6 rad/sec. These aspects appear in the time responses of Figures 4.1 and 4.2. Figure 4.5 shows the disturbance estimate in the absence of measurement noise. The

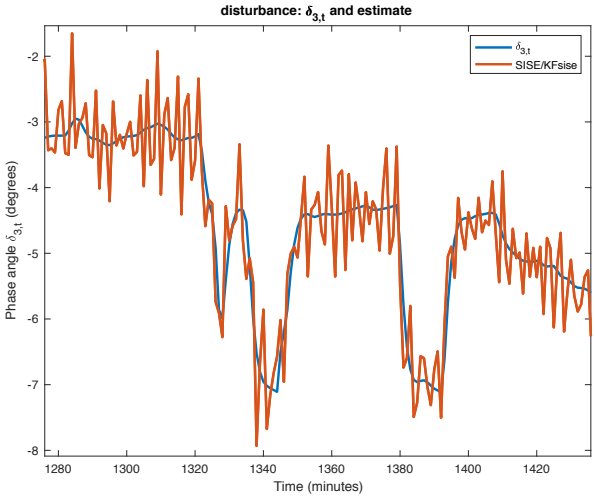


Figure 4.2: External input/disturbance $\delta_{3,t}$ in blue and its estimates with SISE/KF-SISE with $D = 10^6$ in red.

noise-excited ringing is absent.

The point of this numerical example if to highlight the accessibility of standard linear systems design tools for the KF-SISE algorithm via the application of the ARE to derive the steady-state gain and covariance values.

4.7 Direct feedthrough, delays and smoothing

Direct feedthrough modifies measurement equation (4.2) to

$$y_t = Cx_t + Hd_t + v_t. \tag{4.21}$$

The role of direct feedthrough in SISE algorithms has been a complication. Reflection on the problem also indicates that system delays will throw up related issues, since they affect the timing of appearance of d_t in $y_{t+\ell}$ for various values of ℓ. The reconstruction condition, $\operatorname{rank} CG = m$, from Gillijns and De Moor (2007a) and appearing as Assumption 4.1.4, for strictly proper ($H = 0$) systems, ensures that a full estimate of d_t is constructible from y_{t+1} without further delay. This follows since the rank condition admits a gain M satisfying $MCG = I_m$. The rank condition prohibits part of d_t being delayed in the system equation. The disturbance estimate satisfies

$$\hat{d}_{t|t+1} = d_t + CA(x_t - \hat{x}_{t|t}) + Cw_t + v_{t+1}.$$

The condition $\operatorname{rank} H = m$ from Gillijns and De Moor (2007b) for systems with direct feedthrough similarly ensures that y_t suffices for construction of

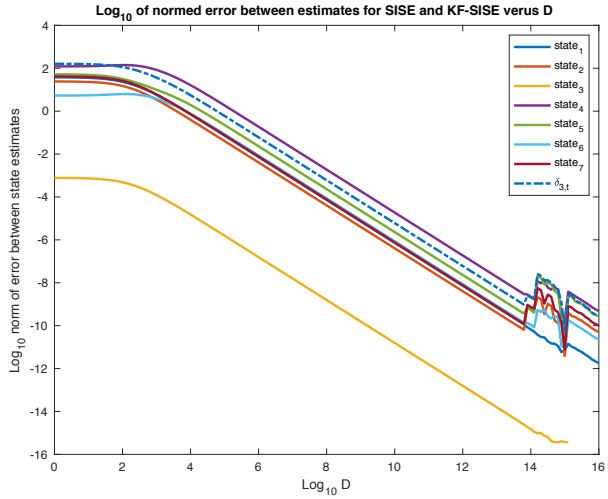


Figure 4.3: Log-log plot of normed errors between SISE and KF-SISE for each signal estimate as a function of $D \in [1, 10^{16}]$.

a full estimate of d_t via $MH = I_m$ in a modified algorithm.

Yang *et al.* Yong et al. (2016) relax this condition to rank $\begin{bmatrix} H^T & G^T C^T \end{bmatrix}^T = m$ and their ULISE algorithm permits a more complicated structure for $\begin{bmatrix} H^T & G^T C^T \end{bmatrix}$ which effectively constrains the Kronecker observability indices of the transfer function from $\{d_t\}$ to $\{y_t\}$ to be 0 or 1. This known observability structure is then used to present ULISE, which is unusually intricate and requires a non-minimum-phase assumption on the $\{d_t\}$ to $\{y_t\}$ transfer function. The signal d_t is estimated in two parts, one directly from y_t and the other via y_{t+1} and y_t. The twenty-three-step algorithm has many points of tangency with other SISE algorithms and with the Kalman filter.

The standard Gaussian pdf methods applied in the proof of Theorem 4.1 carry over to this set up, but with less machinery. Suppose that we are given at time t: $\mathbf{Y}^t$, conditional mean estimates $\hat{x}_{t|t}$ and $\hat{d}_{t|t}$, and corresponding joint conditional covariance matrix

$$\mathcal{S}_t = \begin{bmatrix} S_{x_t x_t^T} & S_{x_t d_t^T} \\ S_{d_t x_t^T} & S_{d_t d_t^T} \end{bmatrix},$$

Consider the conditional probability density function, subject to (4.1) and

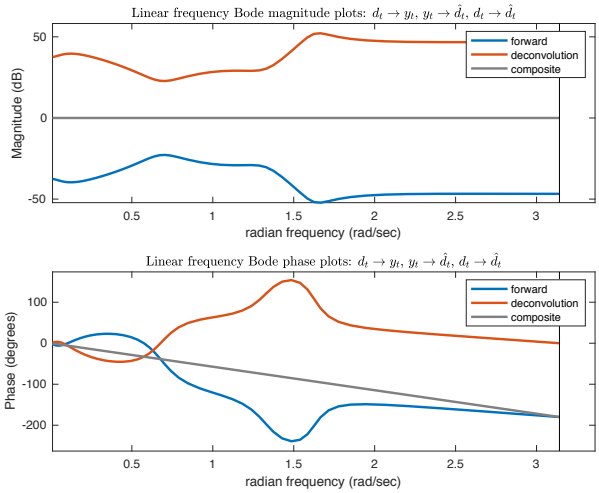


Figure 4.4: Linear-frequency Bode diagrams of discrete transfer functions: $d_t \to y_t$ (blue), $y_t \to \hat{d}_{t-1|t}$ (red) and cascade $d_t \to \hat{d}_{t-1|t}$ (gray) for SISE/KF-SISE with $D = 10^6$ in red.

(4.21),

$$\text{pdf}\left(\left.\begin{bmatrix} x_{t+1} \\ d_t \\ d_{t+1} \\ y_{t+1} \end{bmatrix}\right| \mathbf{Y}^t\right)$$
$$= \mathcal{N}\left(\begin{bmatrix} A\hat{x}_{t|t} + G\hat{d}_{t|t} \\ \hat{d}_{t|t} \\ \mathfrak{d} \\ CA\hat{x}_{t|t} + CG\hat{d}_{t|t} + H\mathfrak{d} \end{bmatrix}, \mathcal{A}S_t\mathcal{A}^T + \mathcal{H}D\mathcal{H}^T + \mathcal{C}Q\mathcal{C}^T + \mathcal{B}R\mathcal{B}^T\right),$$

where

$$\mathcal{A} = \begin{bmatrix} A & G \\ 0 & I_m \\ 0 & 0 \\ CA & CG \end{bmatrix}, \mathcal{H} = \begin{bmatrix} 0 \\ 0 \\ I_m \\ H \end{bmatrix}, \mathcal{C} = \begin{bmatrix} I_n \\ 0 \\ 0 \\ C \end{bmatrix}, \mathcal{B} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ I_p \end{bmatrix}.$$

Then the regular construction of conditional densities given $\mathbf{Y}^{t+1}$ allows us to recreate the SISE algorithm for this problem by taking $D_{t+1} \to \infty$ appropriately. The gains $\mathcal{K}_{t+1}$, $\mathcal{M}_{1,t+1}$ and $\mathcal{M}_{2,t+1}$, for $\hat{x}_{t+1|t+1}$, $\hat{d}_{t|t+1}$ and $\hat{d}_{t+1|t+1}$ respectively, follow naturally and converge to finite limits. Without the assumed constraints on observability indices as in Yong et al. (2016)

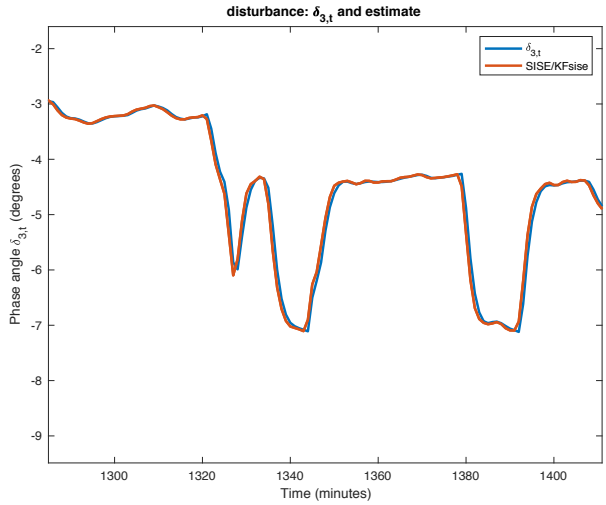


Figure 4.5: Version of Figure 4.2 with noise removed from measurements: external input/disturbance $\delta_{3,t}$ in blue and its estimate with SISE/KF-SISE with $D = 10^6$ in red.

there is no guarantee that the computed covariances also remain bounded. However, the extension of the SISE algorithms to accommodate direct feedthrough, delays and more complicated Kronecker observability indices is clear.

4.8 Non-models, unbiasedness and input estimation

It is usually attributed to John von Neumann or to Stanislaw Ulam that the study of non-equilibrium thermodynamics in Physics is akin to the study of non-elephants in Zoology. By the same token, the study of *model-free* estimation is an unhelpful even meaningless description in this domain. Theorem 4.2 establishes that the SISE algorithm does indeed correspond to a particular model for the disturbance input process $\{d_t\}$ and thereby admits access to standard tools of linear least-squares estimation. The SISE concept that estimates are independent is replaced by the sounder hypothesis that in the signal model the disturbance input is independent from early values of the state.

Unbiasedness of the estimates, used in a probabilistically non-standard (but at least consistent) fashion in SISE since Kitanidis (1987), refers to the property that, no matter the specific value taken by the disturbance, d_t,

the conditional expected value, $\hat{x}_{t|t} = \mathrm{E}\left(x_t|\mathbf{Y}^t\right) = x_t$. This is not so much a statistical property as a geometric one captured by the projection operations of Corollary 4.1. The juxtaposition of probabilistic signal properties with non-models and absence of assumptions concerning the disturbance leads to fundamental questions regarding the nature of filtrations over which one is meant to take the expected values. By assuming a model, albeit a singular one, we are able to clarify these statements and to prove that they disguise a deterministic projection property. This might better be interpreted as a prioritization of the estimation of the input signal over that of the state, with the constraint $\mathrm{E}(\hat{d}_{t|t+1}) = d_t$, trumping the subsequent optimization of (4.10). The formulation developed in this chapter can then be seen as a penalty function approach this same constrained optimization.

Likewise, the strong detectability properties in Yong et al. (2016) appear to be related to the projective features of the relevant algorithm rather than to statistical signal properties. These definitions of unbiasedness are definitions of convenience masking the projections forced by taking unbounded variances. The mathematical question lies in the nature fo the expectations when one assumes no model for d_t.

The computational example of Subsection 4.6.1 provides a good illustration of the state estimation features. The 3-bus network is captured by a known 7-state linear time-invariant model with a single measurement $\omega_{1,t}$. The connection to the customer network is via bus 3 and whatever draws power from or provides power to that bus. In our case, this is a motor with a highly time-varying consumer load. Yet, the (KF-)SISE methods admit the estimation of both the 7-state vector and the input signal $\delta_{3,t}$ using a seventh-order state estimator. Knowledge of the customer side of the network is not required. Certainly the quality of state estimation is degraded compared to having either access to the customer signals or an accurate model of the load signal. This is evident in the performance of the fully informed Kalman filter in, say, Figure 4.1 versus SISE. A traditional approach might be to assume a model, typically lowpass or even static, for the consumer load. But that introduces errors into the estimation of the states in the 3-bus system.

For a network state estimation problem, one can divide the network into known and unknown parts with constraints on the nature of the interconnection to admit SISE estimation of the known part in isolation. Certainly, as seen in Figure 4.1, the performance of a SISE estimator is degraded versus a more completely modeled coupled system or a more complete set of measurements. But the SISE estimator compartmentalizes the problem.

Further, armed with long term estimates of the disturbance input signals, one could explore fitting statistical models to describe their behavior and thereby to be incorporated into an augmented system model, provide one were certain of the time-invariance. These aspects of (KF-)SISE modeling in power system will be filled out elsewhere by the authors. Our aim here is to establish the appropriate genealogy, tools and assessment of the algorithms.

4.9 Conclusion

We have derived from an algebraic perspective the SISE algorithms as Kalman filters of a specific type, suggested by Mendel (1977) in his seismic deconvolution work. The input signal model is white noise, which if its variance tends to infinity, yields a Kalman filter coinciding with SISE. As we mention, this was hinted at earlier but not carried through. Equipped with a fuller understanding of the connections, we were able to present new interpretations and to connect the approach to algebraic Riccati equation computational methods. Further, we were able to clarify – the uncharitable might say *debunk* – the ideas of model-free state estimation and estimate unbiasedness, showing that the methods necessarily involve projections induced by the large variances.

The numerical example demonstrates the computational tools applied to a problem from power systems and illustrates the newfound accessibility of the algorithms via Kalman filtering. For systems with direct feedthrough, these methods also admit much simpler derivation and interpretation of the algorithms. Frequency responses were examined and connections to system inversion Moylan (1977) shown.

Finally, by moving the focus of SISE from deconvolution and input estimation to state estimation, we have been able to present, and illustrate via the numerical example, an approach to state estimation in networked systems, including those with poorly known network neighbors. We shall report on more details power system applications elsewhere.

Appendix – proofs

Proof of Theorem 4.1

From Assumption 4.1 and $\hat{x}_{t|t}$, P_t, write the joint conditional density

$$\text{pdf}\left(\begin{bmatrix} x_{t+1} \\ d_t \\ y_{t+1} \end{bmatrix} \middle| \mathbf{Y}^t\right) = \mathcal{N}\left(\begin{bmatrix} A\hat{x}_{t|t} + G\mathfrak{d} \\ \mathfrak{d} \\ CA\hat{x}_{t|t} + CG\mathfrak{d} \end{bmatrix}, \begin{bmatrix} \mathcal{X}_{t+1} & GD & \mathcal{X}_{t+1}C^T \\ DG^T & D & DG^TC^T \\ C\mathcal{X}_{t+1} & CGD & C\mathcal{X}_{t+1}C^T + R \end{bmatrix}\right).$$

Then appeal to the standard Gaussian conditional density calculation to yield

$$\mathrm{E}\left(\begin{bmatrix} x_{t+1} \\ d_t \end{bmatrix} \middle| \mathbf{Y}^{t+1}\right) = \begin{bmatrix} A\hat{x}_{t|t} + G\mathfrak{d} \\ \mathfrak{d} \end{bmatrix} + \begin{bmatrix} \mathcal{X}_{t+1}C^T \\ DG^TC^T \end{bmatrix} \times (C\mathcal{X}_{t+1}C^T + R)^{-1}(y_{t+1} - CA\hat{x}_{t|t} - CG\mathfrak{d}),$$

or,

$$\begin{aligned}\hat{x}_{t+1|t+1} &= \mathrm{E}\left(x_{t+1}|\mathbf{Y}^{t+1}\right), \\ &= A\hat{x}_{t|t} + G\mathfrak{d} + \mathcal{X}_{t+1}C^T(C\mathcal{X}_{t+1}C^T + R)^{-1}(y_{t+1} - CA\hat{x}_{t|t} - CG\mathfrak{d}),\end{aligned}$$

and

$$\begin{aligned}\hat{d}_{t|t+1} &= \mathrm{E}\left(d_t|\mathbf{Y}^{t+1}\right), \\ &= \mathfrak{d} + DG^TC^T(C\mathcal{X}_{t+1}C^T + R)^{-1}(y_{t+1} - CA\hat{x}_{t|t} - CG\mathfrak{d}).\end{aligned}$$

These are (4.15) and (4.16), respectively in Theorem 4.1.

The covariance calculation follows similarly.

$$\mathrm{cov}\left(\begin{bmatrix} x_{t+1} \\ d_t \end{bmatrix} \middle| \mathbf{Y}^{t+1}\right) = \begin{bmatrix} \mathcal{X}_{t+1} & GD \\ DG^T & D \end{bmatrix} - \begin{bmatrix} \mathcal{X}_{t+1}C^T \\ DG^TC^T \end{bmatrix}(C\mathcal{X}_{t+1}C^T + R)^{-1}\begin{bmatrix} C\mathcal{X}_{t+1} & CGD \end{bmatrix}.$$

The (1,1)-block-element gives (4.14) for $\mathcal{P}_{t+1}$ and the (2,2)-block-element gives (4.18) for $\mathcal{D}_t$.. □

Proof of Identity 4.5

Using Identity 4.1 and denoting $Y_{t+1} = CX_{t+1}C^T + R$,

$$\begin{aligned}
\mathcal{M}_{t+1} &= DG^TC^T(C\mathcal{X}_{t+1}C^T + R)^{-1}, \\
&= DG^TC^T(CX_{t+1}C^T + R)^{-1}(I - CG\mathcal{M}_{t+1}), \\
&= DG^TC^TY_{t+1}^{-1} - DG^TC^TY_{t+1}^{-1}CG\mathcal{M}_{t+1}, \\
\left[I + DG^TC^TY_{t+1}^{-1}CG\right]\mathcal{M}_{t+1} &= DG^TC^TY_{t+1}^{-1}, \\
\left[D^{-1} + G^TC^TY_{t+1}^{-1}CG\right]\mathcal{M}_{t+1} &= G^TC^TY_{t+1}^{-1}, \\
\mathcal{M}_{t+1} &= \left[D^{-1} + G^TC^TY_{t+1}^{-1}CG\right]^{-1}G^TC^TY_{t+1}^{-1}.
\end{aligned}$$

□

Proof of Identity 4.7

From (4.18),

$$\begin{aligned}
\mathcal{D}_t &= D - DG^TC^T(C\mathcal{X}_{t+1}C^T + R)^{-1}CGD, \\
&= (I - \mathcal{M}_{t+1}CG)D,
\end{aligned}$$

Now, using Identity 4.5 and continuing the notation $Y_{t+1} = CX_{t+1}C^T + R$,

$$\begin{aligned}
\mathcal{M}_{t+1}CG &= \left[D^{-1} + G^TC^TY_{t+1}^{-1}CG\right]^{-1}G^TC^TY_{t+1}^{-1}CG, \\
&= \left[D^{-1} + G^TC^TY_{t+1}^{-1}CG\right]^{-1}\left[-D^{-1} + D^{-1} + G^TC^TY_{t+1}^{-1}CG\right], \\
&= I - \left[D^{-1} + G^TC^TY_{t+1}^{-1}CG\right]^{-1}D^{-1}.
\end{aligned}$$

So,

$$\begin{aligned}
\mathcal{D}_t &= \left[D^{-1} + G^TC^TY_{t+1}^{-1}CG\right]^{-1}, \\
&= \left[D^{-1} + G^TC^T(CX_{t+1}C^T + R)^{-1}CG\right]^{-1}.
\end{aligned}$$

□

Proof of Identity 4.8

Substitute for $\mathcal{K}_{t+1}$ from Identity 4.3 into (4.14), drop the time indices, and pay attention to the dimensions and typefaces,

$$\begin{aligned}
\mathcal{P}_{t+1} &= (I_p - \mathcal{K}_{t+1}C)\,\mathcal{X}_{t+1}, \\
&= (I - KC)(I - G\mathcal{M}C)\mathcal{X}, \\
&= (I - KC)(I - G\mathcal{M}C)(X + GDG^T), \\
&= (I - KC)\left[(I - G\mathcal{M}C)X + (I - G\mathcal{M}C)GDG^T\right], \\
&= (I - KC)\left[(I - G\mathcal{M}C)X + G(I - \mathcal{M}CG)DG^T\right].
\end{aligned}$$

Now, denoting (as above) $Y_{t+1} = CX_{t+1}C^T + R$, and using Identity 4.5,

$$\begin{aligned} I - \mathcal{M}CG &= (D^{-1} + G^T C^T Y^{-1} CG)^{-1} \left\{ D^{-1} + G^T C^T Y^{-1} CG - G^T C^T Y^{-1} CG \right\}, \\ &= (D^{-1} + G^T C^T Y^{-1} CG)^{-1} D^{-1}, \\ &= \mathcal{D}_t D^{-1}. \end{aligned}$$

Whence, substituting this above,

$$\mathcal{P}_{t+1} = (I - K_{t+1}C) \left\{ (I - G\mathcal{M}C) X_{t+1} + G\mathcal{D}_t G^T \right\}. \qquad \square$$

Proof of Lemma 4.1

Property (4.20) is proven by Kitanidis (1987) and Gillijns and De Moor (2007a) but also is immediate from (4.5). The projection property follows directly from (4.20).

$$\begin{aligned} (M_{t+1}CG)^2 &= M_{t+1}CG \times M_{t+1}CG \\ &= I_m \times M_{t+1}CG = M_{t+1}CG, \\ (CGM_{t+1})^2 &= CGM_{t+1} \times CGM_{t+1} \\ &= CG \times M_{t+1}CG \times M_{t+1} = CGM_{t+1}, \\ (GM_{t+1}C)^2 &= GM_{t+1}C \times GM_{t+1}C \\ &= G \times M_{t+1}CG \times M_{t+1}C = GM_{t+1}C. \end{aligned}$$

From Assumption 4.1 the ranks of $M_{t+1}CG$, CGM_{t+1} and GCM_{t+1} are all less than or equal to m. Since $MCG = I_m$, rank $\mathcal{M}_{t+1}CG = m$ and $\mathcal{R}a\,(MCG) = \mathbb{R}^m$. Also, $CGM_{t+1} \times CG = CG$. So rank $CGM_{t+1} = m$ and $\mathcal{R}a\,(CM_{t+1}G) = \mathcal{R}a\,(CG)$. Similarly, since $GM_{t+1}CG = G$, rank $GM_{t+1}C = m$ and $\mathcal{R}a\,(GM_{t+1}C) = \mathcal{R}a\,(G)$. $\square$

Proof of Theorem 4.2

Identity 4.9 establishes the convergence of $\mathcal{M}_{t+1}$ and $\mathcal{K}_{t+1}$. With the convergence of $\mathcal{M}_{t+1}$ and $M_{t+1}CG = I$, from Lemma 4.1, we see that (4.7) and (4.16) are identical updates regardless of the value of $\mathfrak{d}$. Substituting the limiting value for $\mathcal{K}_{t+1}$ in (4.15) shows that this update and (4.8) also are identical. Thus, SISE and KF-SISE yield identical estimates at this t from the same starting data and, hence, the estimates remain identical. Since $\mathcal{P}_{t+1}$ is the conditional error covariance of x_{t+1} for KF-SISE and P_{t+1} is shown in Gillijns and De Moor (2007a) to be the conditional covariance of the SISE estimate, these covariances must also be identical. $\square$

Proof of Corollary 4.1

Applying Identity 4.4 to (4.8) and then Identities 4.3 and 4.9,

$$\begin{aligned}
\hat{x}^{\text{SISE}}_{t+1|t+1} &= A\hat{x}^{\text{SISE}}_{t|t} + G\mathfrak{d} + \mathcal{K}_{t+1}(y_{t+1} - CA\hat{x}^{\text{SISE}}_{t|t} - CG\mathfrak{d}),\\
&= (I - \mathcal{K}_{t+1}C)\, A\hat{x}^{\text{SISE}}_{t|t} + (I - \mathcal{K}_{t+1}C)\, G\mathfrak{d} + \mathcal{K}_{t+1}y_{t+1},\\
&= (I - K_{t+1}C)\,(I - GM_{t+1}C)\left(A\hat{x}^{\text{SISE}}_{t|t} + G\mathfrak{d}\right) + \mathcal{K}_{t+1}y_{t+1},\\
&= (I - K_{t+1}C)\,\hat{x}^{\text{SISE}}_{t+1|t} + \mathcal{K}_{t+1}y_{t+1},\\
&= (I - K_{t+1}C)\,\hat{x}^{\text{SISE}}_{t+1|t} + GM_{t+1}y_{t+1} + K_{t+1}(I - CGM_{t+1})y_{t+1},\\
&= (I - K_{t+1}C)\,\hat{x}^{\text{SISE}}_{t+1|t} + (I - K_{t+1}C)GM_{t+1}y_{t+1} + K_{t+1}y_{t+1},\\
&= (I - K_{t+1}C)\left(\hat{x}^{\text{SISE}}_{t+1|t} + GM_{t+1}y_{t+1}\right) + K_{t+1}y_{t+1},\\
&= (I - K_{t+1}C)\,\hat{x}^{\text{SISE}}_{t+1|t+\frac{1}{2}} + K_{t+1}y_{t+1}. \qquad \square
\end{aligned}$$

$$\text{cov}(d_t - \hat{d}_{t|t+1}|\mathbf{Y}^{t+1}) = (G^TC^T\tilde{R}^{-1}_{t+1}CG)^{-1}.$$

where,

$$\mathcal{X}_t = AP_tA^T + GD_tG^T + Q, \tag{4.22}$$

$$\mathcal{K}_{t+1} = \mathcal{X}_tC^T(C\mathcal{X}_tC^T + R)^{-1}, \tag{4.23}$$

$$\mathcal{M}_{t+1} = D_tG^TC^T(C\mathcal{X}_tC^T + R)^{-1}. \tag{4.24}$$

The conditional covariances of these estimates are given by

$$\text{cov}(\hat{x}_{t+1|t+1}) = \mathcal{X}_t - \mathcal{X}_tC^T(C\mathcal{X}_tC^T + R)^{-1}C\mathcal{X}_t, \tag{4.25}$$

$$\text{cov}(\hat{d}_{t|t+1}) = D_t - D_tG^TC^T(C\mathcal{X}_tC^T + R)^{-1}CGD_t. \tag{4.26}$$

Let $\{U, \Sigma, V\}$ be the singular-value decomposition of G. That is,

$$\begin{aligned}
G &= U\Sigma V^T,\\
&= \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_m \\ 0_{(n-m)\times m} \end{bmatrix} V^T,
\end{aligned}$$

where U_1 is an orthonormal basis for $\mathcal{R}a(G)$, U_2 is an orthonormal basis for $\mathcal{R}a(G)^\perp$, an Σ_m is a full-rank $m \times m$ diagonal matrix of positive singular values. Using the state space transformation

$$\bar{x}_t = \begin{bmatrix} \bar{x}_{1,t} \\ \bar{x}_{2,t} \end{bmatrix} = U^Tx_t = \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} x_t, \quad \begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix} = \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} w_t,$$

rewrite the state equation (4.1)

$$\begin{bmatrix} \bar{x}_{1,t+1} \\ \bar{x}_{2.t+1} \end{bmatrix} = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix} \begin{bmatrix} \bar{x}_{1,t} \\ \bar{x}_{2,t} \end{bmatrix} + \begin{bmatrix} \Sigma_m \\ 0 \end{bmatrix} d_t + \begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix}.$$

Chapter 5

Paper II: Simultaneous input & state estimation, singular filtering and stability

This chapter is taken from Abooshahab et al. (2020b).

Simultaneous input & state estimation, singular filtering and stability

Mohammad Ali Abooshahab
Mohammed M.J. Alyaseen
Robert R. Bitmead
Morten Hovd
submitted to and provisionally accepted in *Automatica*, (2020)

Abstract

Input estimation is a signal processing technique associated with deconvolution of measured signals after filtering through a known dynamic system. Kitanidis and others extended this to the simultaneous estimation of the input signal and the state of the intervening system. This is normally posed as a special least-squares estimation problem with unbiasedness. The approach has application in signal analysis and in control. Despite the connection to optimal estimation, the standard algorithms are not necessarily stable, leading to a number of recent papers which present sufficient conditions for stability. In this chapter we complete these stability results in two

ways in the time-invariant case: for the square case, where the number of measurements equals the number of unknown inputs, we establish exactly the location of the algorithm poles; for the non-square case, we show that the best sufficient conditions are also necessary. We then draw on our previous results interpreting these algorithms, when stable, as singular Kalman filters to advocate a direct, guaranteed stable implementation via Kalman filtering. This has the advantage of clarity and flexibility in addition to stability. En route, we decipher the existing algorithms in terms of system inversion and successive singular filtering. The stability results are extended to the time-varying case directly to recover the earlier sufficient conditions for stability via the Riccati difference equation.

5.1 Introduction

Simultaneous Input and State Estimation (SISE) algorithms take a system with an unknown disturbance input sequence, $\{d_t\}$, and measured output signal, $\{y_t\}$, plus possibly also a known input sequence, $\{u_t\}$, and produce estimates of both d_t and the system state x_t, based on fixed lag smoothing. There have been a great number of recent papers on these algorithms, with a recent subset Marro and Zattoni (2010), Fang et al. (2011), Fang and De Callafon (2012), Yong et al. (2016) focusing on conditions for stability. A feature of SISE algorithms is that the disturbance signals have uncertain provenance. So SISE presumes that no model is available for this signal and the algorithm proceeds without an explicit description of the disturbance signals' statistical properties. Thus techniques such as extended state observers Guo and Zhao (2016) and augmented Kalman filters Anderson and Moore (2012), Simon (2006) are inapplicable, as both rely on disturbance models. A recent paper by the authors Bitmead et al. (2019) establishes that, when stable, the linear system SISE algorithm of Gillijns and De Moor (2007a) coincides with the Kalman filter with $\{d_t\}$ modeled by white Gaussian noise of unbounded variance. Various approaches consider first estimating the state Marro and Zattoni (2010) and then using the state recursion to reconstruct d_t, or estimating the disturbance first and then reconstructing the state Gillijns and De Moor (2007a;b), Yong et al. (2016). These methods rely on geometric approaches and system inversion, although there is a strong overlap with least-squares state estimation concepts of unbiasedness and optimality.

SISE algorithms go back to least to Kitanidis Kitanidis (1987) with antecedents Glover (1969), Mendel (1977), Sanyal and Shen (1974) concentrating on input signal reconstruction. Here, we follow the formulation from Yong,

Zhu and Frazzoli Yong et al. (2016), which in turn builds on Gillijns and De Moor (2007a;b). We consider linear time-invariant systems to add clarity and to explore the connection to optimal estimation before extending to uniformly time-varying systems.

Input estimation is a signal processing technique associated with deconvolution of measured signals after filtering through a known dynamic system. Examples include the estimation of rainfall given river flow and the calculation of salinity in the ocean accommodating for sensor dynamics Fang et al. (2013). Here, the central objective is to estimate the driving disturbance signal d_t and there is little interest in the sensor state. The algorithm should be stable, however. Our particular driving problem, on the other hand, is the estimation of generator states in part of a power grid when the interconnection signals are unknown Abooshahab et al. (2019). Here the priority is to estimate network generator states in the face of unmodeled and unmeasured consumption, which is treated as the disturbance signal. In spite of these distinct objectives, the same algorithms have been used.

Contributions & organization

Our objective in this chapter is to attempt to bring some clarity and unity to this picture by establishing precisely the connection to system inversion and optimal estimation by deriving necessary and sufficient conditions for stability using explicit system inverse formulæ and algebraic Riccati equations, starting with the time-invariant case. Earlier stability conditions were sufficient only but derived in the time-varying situation. We recover these. Further, when stability is not achieved by these SISE algorithms, we propose a modification based on inner-outer factorization, which maintains state estimation performance at the expense of simple disturbance recovery. This can be compared with the techniques advanced in Marro and Zattoni (2010) for approximate system inversion with delay. Beyond this work, we know of no other which addresses estimation when the stability conditions fail.

Section 5.2 presents the SISE problem for a linear time-invariant system. Section 5.3 studies the zero direct feedthrough case and the corresponding SISE of Gillijns and De Moor (2007a) and shows that, in the square case where the number of measurements equals the number of disturbance channels, the input estimator is the inverse of the d_t-to-y_t system and the state estimator is a plant simulation. Stability depends on the transmission zeros of the former system. These necessary and sufficient stability conditions then are extended to the non-square case with more measurements. This involves the Riccati difference equation and a detectability condition.

Section 5.4 expands this analysis to the full-rank direct feedthrough case and comments on the non-full-rank case of Yong et al. (2016). Section 5.5 draws connections to earlier works of singular filtering and introduces an accommodation to circumvent stability issues using the inner-outer factorization. It also contains the extension to time-varying systems via the Riccati equation. Section 5.6 reinforces the connections to system inversion and concludes. The Appendix contains the proofs.

5.2 Problem statement

SISE algorithms have been formulated for linear time-varying systems Gillijns and De Moor (2007a;b), Yong et al. (2016) and for nonlinear time-varying systems Fang et al. (2013), Kim et al. (2020). However for clarity in development, we consider the linear, time-invariant system with zero known control input,

$$x_{t+1} = Ax_t + Gd_t + w_t, \tag{5.1}$$
$$y_t = Cx_t + Hd_t + v_t, \tag{5.2}$$

with $x_t \in \mathbb{R}^n$, $d_t \in \mathbb{R}^m$, $y_t \in \mathbb{R}^p$. Zero-mean white noises $\{w_t\}$ and $\{v_t\}$ are independent and independent from $\{d_t\}$ and x_0. The covariance of w_t is $Q \geq 0$ and the covariance of v_t is $R > 0$. Denote the signal measurements $\mathbf{Y}^t \triangleq \{y_t, y_{t-1}, \ldots, y_0\}$. The aim is to produce from $\mathbf{Y}^t$, a recursive filtered state estimate, $\hat{x}_{t|t}$, and filtered and/or smoothed estimates, $\hat{d}_{t|t+1}$ or $\hat{d}_{t|t}$, depending on the properties of G and $\begin{bmatrix} C & H \end{bmatrix}$. We make the following assumption.

Assumption 5.1. *System (5.1-5.2) has* $[A, C]$ *observable,* $rank G = m$, $[A, Q]$ *reachable, and* $R > 0$.

Full-rank direct feedthrough, i.e. $\text{rank}(H) = m$, is treated in Gillijns and De Moor (2007b); zero direct feedthrough, $H = 0$, in Gillijns and De Moor (2007a) with $\text{rank}(CG) = m$; and, Yong et al. (2016) provides a generalization, ULISE, with mixed rank properties between H and CG. A noise-free variant is treated in Marro and Zattoni (2010).

5.3 Zero direct feedthrough

For $H = 0$ in (5.2), SISE from Gillijns and De Moor (2007a) is the recursion.

$$X_t = AP_{t-1}A^T + Q, \tag{5.3}$$
$$K_t = X_t C^T (CX_t C^T + R)^{-1}, \tag{5.4}$$
$$M_t = [G^T C^T (CX_t C^T + R)^{-1} CG]^{-1} \times G^T C^T (CX_t C^T + R)^{-1}, \tag{5.5}$$
$$P_t = (I - K_t C)\left[(I - GM_t C)X_t \times (I - GM_t C)^T + GM_t R M_t^T G^T\right] + K_t R M_t^T G^T, \tag{5.6}$$
$$\hat{d}_{t-1|t} = M_t(y_t - CA\hat{x}_{t-1|t-1}), \tag{5.7}$$
$$\hat{x}_{t|t} = A\hat{x}_{t-1|t-1} + G\hat{d}_{t-1|t} + K_t \times (y_t - CA\hat{x}_{t-1|t-1} - CG\hat{d}_{t-1|t}). \tag{5.8}$$
$$\operatorname{cov}(x_t|\mathbf{Y}^t) = P_t, \tag{5.9}$$

under the following structural condition.

Assumption 5.2.

$$rank\, CG = m. \tag{5.10}$$

An immediate observation is that SISE contains no specific information related to a model for the unmeasured disturbance d_t. Indeed, it is frequently claimed that signal $\{d_t : t = 0, 1, \dots\}$ possesses no model whatsoever. Although, for bounded covariance X_t, i.e. when the algorithm is stable, the authors derived this version of SISE in Bitmead et al. (2019) as a Kalman filter with $\{d_t\}$ modeled as a white noise process of unbounded variance. We shall return to this point later. Evidently, Assumption 5.2 requires $p \geq m$ and $\operatorname{rank} C \geq \operatorname{rank} G = m$. Firstly, we treat the square case, $p = m$, where the number of measurements equals the dimension of the disturbance input. Then we shall derive more general results.

5.3.1 Square zero-feedthrough case

From Assumption 5.2 when $p = m$, CG is invertible. Since, from (5.5), $M_tCG = I$ or $M_t = (CG)^{-1}$, we have

$$\hat{d}_{t-1|t} = (CG)^{-1}(y_t - CA\hat{x}_{t-1|t-1}), \tag{5.11}$$
$$0 = y_t - CA\hat{x}_{t-1|t-1} - CG\hat{d}_{t-1|t}, \tag{5.12}$$
$$\hat{x}_{t|t} = A\hat{x}_{t-1|t-1} + G\hat{d}_{t-1|t}, \tag{5.13}$$
$$= [I - G(CG)^{-1}C]A\hat{x}_{t-1|t-1} + G(CG)^{-1}y_t. \tag{5.14}$$

This estimation algorithm:

- is time-invariant;
- does not depend on Q or R, the noise variances;
- is independent from the covariance calculations.,
- has zero $\hat{x}_{t|t}$ innovations (5.12), (5.8).

SISE reduces to (5.11-5.14).

$$\hat{x}_{t|t} = [I - G(CG)^{-1}C]A\hat{x}_{t-1|t-1} + G(CG)^{-1}y_t,$$
$$\hat{d}_{t-1|t} = -(CG)^{-1}CA\hat{x}_{t-1|t-1} + (CG)^{-1}y_t.$$

Note that, using the matrix inversion lemma, we may rewrite the SISE y_t-to-$\hat{d}_{t-1|t}$ transfer function as

$$\begin{aligned}&(CG)^{-1}\\ &\quad - (CG)^{-1}CA(zI - A + G(CG)^{-1}CA)^{-1}G(CG)^{-1}\\ &\qquad = \left[CG + CA(zI - A)^{-1}G\right]^{-1},\\ &\qquad = \left[zC(zI - A)^{-1}G\right]^{-1}.\end{aligned} \tag{5.15}$$

The filtered state estimate error satisfies

$$\begin{aligned}\tilde{x}_{t|t} &\triangleq x_t - \hat{x}_{t|t},\\ &= [I - G(CG)^{-1}C]A\tilde{x}_{t-1|t-1} + [I - G(CG)^{-1}C]w_{t-1} - G(CG)^{-1}v_t.\end{aligned}$$

The stability of SISE, i.e. the boundedness of the covariance of $\tilde{x}_{t|t}$, depends on the eigenvalues of $[I - G(CG)^{-1}C]A$.

Theorem 5.1. *For system (5.1-5.2) with $p = m$ and subject to Assumption 5.2, the eigenvalues of the SISE estimator system matrix, $\left[I - G(CG)^{-1}C\right]A$, lie at the transmission zeros of the square transfer function $zC(zI-A)^{-1}G$. Accordingly, the SISE estimator is asymptotically stable if and only if these transmission zeros all lie inside the unit circle.*

The proof of this theorem follows immediately from (5.15). An alternate is given in the Appendix for completeness and to establish connections to singular filtering. We note that condition (5.10) in Assumption 5.2 implies that $zC(zI-A)^{-1}G$ possesses exactly n finite transmission zeros with exactly m at zero.

We see that, in the square case, the poles of SISE can be located precisely at the transmission zeros of the d_t-to-y_t transfer function. SISE therefore is performing system inversion to recover $\hat{d}_{t-1|t}$ from $\mathbf{Y}^t$. The dependent recursion (5.13) for $\hat{x}_{t|t}$ is a simulation of the state equation (5.1) driven by $\hat{d}_{t-1|t}$. Effectively all the information in $\mathbf{Y}^t$ is used in generating the disturbance estimate, leaving simulation (5.13) to generate the state estimate.

When SISE is stable, it was shown in Bitmead et al. (2019) that the state estimation algorithm implements a Kalman filter with a model for $\{d_t\}$ as a white noise of unbounded variance, D. In this case, the state estimation problem has driving noise variance $Q+GDG^T$ and measurement noise variance R. The identical filter, but not the covariances, will be achieved by taking driving noise GDG^T for finite D and $R \to 0$. That is, SISE is a singular filter. The connection to Maciejowski (1985) in the proof is to the equivalent result in Loop Transfer Recovery for LQG control. When one selects $R = 0$, as opposed to $R \to 0$ from above, then the poles are placed at the transmission zeros. The limiting operation, on the other hand places the poles at the stable transmission zeros and the inverses of the unstable transmission zeros Shaked (1985).

5.3.2 Non-square zero-feedthrough case

From Assumption 5.2, we take $p \geq m$ and make a transformation of the output signal as follows. This is a variation on the technique of Yong et al. (2016). Take the singular value decomposition of $p \times m$ CG.

$$\begin{aligned} \texttt{svd}(CG) &= U\Sigma V^T, \\ &= \begin{bmatrix} U_m & U_{p-m} \end{bmatrix} \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^T. \end{aligned}$$

Define the $p \times p$ transformation

$$\mathcal{T} = \begin{bmatrix} U_m^T - U_m^T R U_{p-m} (U_{p-m}^T R U_{p-m})^{-1} U_{p-m}^T \\ U_{p-m}^T \end{bmatrix}, \tag{5.16}$$

and transform the original output signal, call it $\bar{y}_t$,

$$y_t = \mathcal{T} \bar{y}_t = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} x_t + \begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix}, \tag{5.17}$$

yielding

$$\det C_1 G \neq 0, \quad C_2 G = 0, \quad \operatorname{cov} \begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix} = \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}.$$

Theorem 5.2. *For system (5.1-5.2) with $p \geq m$ and subject to Assumptions 5.1 and 5.2, if and only if the pair $[A(I - G(C_1G)^{-1}C_1), C_2]$ is detectable then the filtered state covariance, P_t, is bounded and converges to a limit P_∞ as $t \to \infty$.*

The corresponding gain matrices, K_∞ and M_∞, yield the limiting SISE system matrix, $(I - K_\infty C)(I - GM_\infty C)A$, with all its eigenvalues strictly inside the unit circle.

The proof of this result appears in the Appendix and is based on proving that the state covariance satisfies a Riccati Difference Equation. Although this condition is not strictly the same as the condition in Yong et al. (2016), the theorem condition implies theirs. Hence, their condition is also necessary. Theorem 5.2 similarly extends the condition in Fang et al. (2011). The sufficient stability result in Kim et al. (2020) is predicated on P_t being bounded a priori. We already know from Theorem 5.1 the eigenvalues of $A(I - G(C_1G)^{-1}C_1)$ are stable if and only if $zC(zI - A)^{-1}G$ is minimum-phase.

We see that, when $p > m$, the surfeit of measurements beyond those strictly needed to produce $\hat{d}_{t-1|t}$ are brought to bear on estimating x_t. The stability of SISE depends on either the square case yielding stability via Theorem 5.1, i.e. via stable transmission zeros, or there being sufficient information in the additional measurements to stabilize the estimator.

When SISE is stable, then the algorithm implements a singular filter, as explained above. However, now the corresponding singular filter is *partially singular*, a term introduced in Priel and Shaked (1986). That is, the process noise variance is finite but the measurement noise variance, R, is less

than full rank rather than zero. The approach of Priel and Shaked (1986), under the banner of stable optimal filtering, in this case involves precisely a succession of a singular estimator and followed by a regular estimator, as in SISE. The result in Bitmead et al. (2019) derives this stable (partially) singular filter when the plant satisfies the conditions of Theorem 5.2.

5.4 Nonzero direct feedthrough

When $H \neq 0$ in (5.2), SISE alters. Gillijns and De Moor Gillijns and De Moor (2007b) provide a SISE algorithm, subject to the following, for the time-invariant case.

Assumption 5.3. *Rank* $H = m$.

Subject to this assumption, the SISE formulation for time-invariant system (5.1-5.2) is

$$\hat{x}_{t|t-1} = A\hat{x}_{t-1|t-1} + G\hat{d}_{t-1|t-1}, \tag{5.18}$$

$$\begin{aligned}
P^x_{t|t} &= \begin{bmatrix} A & G \end{bmatrix} \begin{bmatrix} P^x_{t-1|t-1} & P^{xd}_{t-1|t-1} \\ P^{dx}_{t-1|t-1} & P^d_{t-1|t-1} \end{bmatrix} \begin{bmatrix} A^T \\ G^T \end{bmatrix} + Q, \\
\tilde{R}_t &= CP^x_{t|t-1}C^T + R, \\
M_t &= (H^T\tilde{R}_tH)^{-1}H^T\tilde{R}_t^{-1}, \\
\hat{d}_{t|t} &= M_t(y_t - C\hat{x}_{t|t-1}), \qquad (5.19) \\
P^d_{t|t} &= (H^T\tilde{R}_tH)^{-1}, \\
K_t &= P^x_{k|k-1}C^T\tilde{R}_t^{-1}, \\
\hat{x}_{t|t} &= \hat{x}_{t|t-1} + K_t(y_t - C\hat{x}_{t|t-1} - H\hat{d}_{t|t}), \qquad (5.20) \\
P^x_{t|t} &= P^x_{t|t-1} - K_t(\tilde{R}_t - HP^dH^T)K_t^T, \\
P^{xd}_{t|t} &= \left(P^{dx}_{t|t}\right)^T = -K_tHP^d_{t|t}.
\end{aligned}$$

When rank $H < m$, Yong et al. (2016) provide ULISE, a carefully developed SISE algorithm which uses the singular value decomposition as in Subsection 5.3.2 but more widely to handle the more complicated interaction between filtered and smoothed estimates for d_t.

5.4.1 Square full-rank case

As with the $H = 0$ case, we consider first rank $H = m$ and $m = p$. That is H is invertible and, since $M_tH = I$, $M_t = H^{-1}$. Then SISE reduces to the

recursion

$$\begin{aligned}
\hat{d}_{t|t} &= H^{-1}(y_t - C\hat{x}_{t|t-1}) \\
0 &= y_t - C\hat{x}_{t|t-1} - H\hat{d}_{t|t}, \qquad (5.21)\\
\hat{x}_{t+1|t} &= A\hat{x}_{t-1|t-1} + G\hat{d}_{t-1|t-1}, \qquad (5.22)\\
&= (A - GH^{-1}C)\hat{x}_{t|t-1} + GH^{-1}y_t, \\
\tilde{x}_{t+1|t} &= (A - GH^{-1}C)\tilde{x}_{t|t-1} + w_t - GH^{-1}v_t.
\end{aligned}$$

Theorem 5.3. *For system (5.1-5.2) subject to Assumption 5.3, the eigenvalues of the SISE estimator system matrix, $A-GH^{-1}C$, lie at the transmission zeros of the square transfer function $H + C(zI - A)^{-1}G$. Accordingly, the SISE estimator is asymptotically stable if and only if these transmission zeros all lie inside the unit circle.*

*Proof:*Applying the matrix inversion lemma to the square transfer function between d_t and y_t,

$$\begin{aligned}
&\left[H + C(zI - A)^{-1}G\right]^{-1} \\
&\quad = H^{-1} - H^{-1}C(zI - A + GH^{-1}C)^{-1}GH^{-1}.
\end{aligned}$$

The poles of the square direct feedthrough SISE lie at the transmission zeros of the d_t to y_t transfer function. □

Again, this result adds necessity to that of Yong et al. (2016) in this case. Further, the result does not rely on optimality arguments. As in the square zero feedthrough case, the SISE estimator is time-invariant and independent from Q and R, and the state estimate filter innovations is zero. The condition $\operatorname{rank} H = m$ ensures that all n transmission zeros are finite. We note again that the state innovations sequence (5.21) is zero and the filter (5.22) simulates $\hat{x}_{t+1|t}$ from $\hat{d}_{t|t}$.

5.4.2 Non-square full-rank case

The careful derivation of ULISE to accommodate $\operatorname{rank} H \leq m$ is a central contribution of Yong et al. (2016) and involves separation into subspaces. Take the singular value decomposition of $p \times m$ H possessing rank r.

$$\begin{aligned}
\texttt{svd}(H) &= U\Sigma V^T, \\
&= \begin{bmatrix} U_r & U_{p-r} \end{bmatrix} \begin{bmatrix} \bar{H} & 0 \\ 0 & 0 \end{bmatrix} V^T.
\end{aligned}$$

Matrices take on the $(r, p-r)$ structure.

$$H = \begin{bmatrix} \bar{H} & 0 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, \quad G = \begin{bmatrix} G_1 & G_2 \end{bmatrix},$$
$$K_t = \begin{bmatrix} K_{1,t} & K_{2,t} \end{bmatrix}, \quad M_t = \begin{bmatrix} M_{1,t} & M_{2,t} \end{bmatrix}.$$

As earlier, define the $p \times p$ transformation

$$\mathcal{T} = \begin{bmatrix} U_r^T - U_r^T R U_{p-r} (U_{p-r}^T R U_{p-r})^{-1} U_{p-r}^T \\ U_{p-r}^T \end{bmatrix}, \tag{5.23}$$

and transform the original output signal, call it $\bar{y}_t$,

$$y_t = \mathcal{T}\bar{y}_t = \begin{bmatrix} \bar{C}_1 \\ \bar{C}_2 \end{bmatrix} x_t + \begin{bmatrix} \bar{H} \\ 0 \end{bmatrix} d_t + \begin{bmatrix} \bar{v}_{1,t} \\ \bar{v}_{2,t} \end{bmatrix}. \tag{5.24}$$

yielding $\det \bar{H} \neq 0$ and $\operatorname{cov} \begin{bmatrix} \bar{v}_{1,t} \\ \bar{v}_{2,t} \end{bmatrix} = \begin{bmatrix} \bar{R}_1 & 0 \\ 0 & \bar{R}_2 \end{bmatrix}$. When $\operatorname{rank} H = m$, $\bar{H}$ is $m \times m$ and we have the following result stemming from $M_{1,t}\bar{H} = I_m$.

Theorem 5.4. *Subject to Assumptions 5.1 and 5.3, $p \geq m$, SISE with feedthrough is stable if only if $[A - G\bar{H}^{-1}\bar{C}_1, \bar{C}_2]$ is detectable.*

The proof of this result is in the Appendix. This extends the detectability condition[1] of Theorem 5 of Yong et al. (2016) to a necessary and sufficient condition for stability of SISE in this case. It also is the analog of Theorem 5.2 for the full-rank feedthrough case.

An alternative way to view necessity is to write the system matrix of SISE as

$$\begin{aligned} &A - [AK_t + (G - AK_tH)M_t]C \\ &\quad = A - G\bar{H}^{-1}\bar{C}_1 - (AK_{2,t} + GM_{2,,t} - AK_{1,t}\bar{H}M_{2,t})\bar{C}_2. \end{aligned}$$

For this matrix to be stable, a multiple of $\bar{C}_2$ must stabilize $A - G\bar{H}^{-1}\bar{C}_1$.

5.4.3 Less than full rank feedthrough

We build again on the decomposition above of Yong et al. (2016) and make the following assumption.

Assumption 5.4. *$\operatorname{rank} \bar{C}_2 G_2 = m - \operatorname{rank} \bar{H}$.*

[1] Note that Yong et al. (2016) uses p to denote our m.

In Yong et al. (2016), the authors derive a sufficient condition for stability which we now extend to necessity.

Theorem 5.5. *Subject to Assumptions 5.1 and 5.4, general feedthrough SISE is stable if only if* $[A - G_1\bar{H}^{-1}\bar{C}_1, \bar{C}_2]$ *is detectable.*

This detectability condition is shown in Yong et al. (2016) to be sufficient for stability by using the filter recursion for $\hat{x}_{t|t}$. If one calculates the alternative recursive prediction, $\hat{x}_{t|t-1}$, then it is evident that the ULISE system matrix is again of the form

$$\begin{aligned}(A - G_1\bar{H}^{-1}\bar{C}_1)(I - \tilde{L}_t\bar{C}_2)(I - G_2M_{2,t}\bar{C}_2)\\ = A - G_1\bar{H}^{-1}\bar{C}_1 + W_t\bar{C}_2,\end{aligned}$$

for appropriate W_t. Evidently, this can be stable only if the detectability condition holds.

5.5 Upshots

5.5.1 Stability and singular filtering

The preceding analysis provides necessary and sufficient conditions for the stability of linear SISE algorithms. Further, for the square cases, it yields the precise locations on the algorithm poles and demonstrates that the emphasis is on d_t-to-y_t system inversion to recover d_t followed by best efforts to estimate the state. We have pointed out the successive estimation nature of SISE, as have others. The question remains as to actions to be taken when SISE proves to be unstable, noting these central properties:

1. SISE is stable when the d_t-to-y_t system is stably invertible.

2. When SISE is stable, it corresponds (at least in the zero feedthrough case) to a singular Kalman filter.

3. Subject to: detectability of $[A, C]$, stabilizability of $[A, Q^{\frac{1}{2}}]$, and $R \overset{+}{\to} 0$; the singular Kalman filter is a stable estimator by construction.

4. The stability conditions for the singular Kalman filter are more relaxed than Assumptions 5.2, 5.3 or 5.4.

5. When SISE proves to be unstable, it differs from the Kalman filter.

5.5.2 Kalman filtering for input and state estimation

Denote the transfer function from d_t to y_t by $T(z)$. As derived in earlier sections, when $zT(z)$ has all its transmission zeros inside the unit circle, then SISE is guaranteed stable and is equivalent to a specific stable singular Kalman filter. Compute a discrete-time inner-outer factorization[2] Green (1988), Ionescu and Oara (1996).

$$T(z) = T_o(z)T_i(z),$$

with $m \times m$ $T_i(z)$ an inner function, i.e. stable and all-pass, and $p \times m$ $T_o(z)$ an outer function, i.e. all transmission zeros inside the unit circle.

We note the following from the construction of the inner-outer factors.

Lemma 5.1. *If $T(z)$ has realization*

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + Gd_t + w_t, \\ y_t &= Cx_t + Du_t + Hd_t + v_t, \end{aligned}$$

then $T_o(z)$ has realization

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + \check{G}\check{d}_t + w_t, \\ y_t &= Cx_t + Du_t + \check{H}\check{d}_t + v_t, \end{aligned}$$

where $\check{d}_t$ is the output of $T_i(z)$.

That is, for the same initial conditions, $T_o(z)$ and $T(z)$ have the same states and have realizations which differ only in the G and H matrices. This factorization is depicted in Figure 5.1.

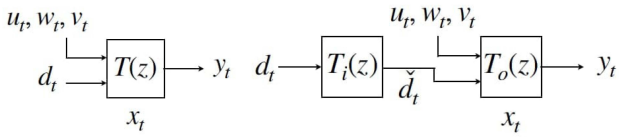


Figure 5.1: System inner-outer factorization

Applying SISE or the singular Kalman filter for the outer function $T_o(z)$ to the signal $\{y_t\}$ yields estimates of $\{\check{d}_t\}$ and $\{x_t\}$ via a stable algorithm

[2] Strictly speaking, this is a co-inner-co-outer factorization because of the ordering of T_i and T_o Francis (1987). It can be obtained from the inner-outer factorization of T^T.

with guaranteed statistical and optimality properties. This follows since T_o is stably invertible by construction. Further, this stability depends on the standard assumptions above for Kalman filter stability. If one uses SISE, then depending on the delay properties of $T_o(z)$, i.e. its behavior as $z \to \infty$, a modified variation of the algorithm and Assumption 5.2, 5.3, or 5.4 might be needed to accommodate d_t-to-y_t invertibility. This is discussed further in Section 5.6.

To recover estimates for original system inputs $\{d_t\}$ from those for $\{\check{d}_t\}$ requires deconvolution (input estimation) without state estimation for the maximum-phase but stable system T_i. If delay is not an issue, then this can proceed stably via a fixed-interval smoother or reverse-time input estimation.

If the state estimates of x_t themselves are the objective, then the reconstruction of $\check{d}_t$ versus d_t is immaterial. This is the nature of the problem addressed in partially-known power system state estimation Abooshahab et al. (2019).

The singular filters derived by Shaked and co-authors Shaked (1985), Priel and Shaked (1986), Shaked and Soroka (1987) rely on the Return Difference Equality and spectral factorization for their calculation. In the case where the transmission zeros are unstable, the filter solution replaces them by their inverses akin to the inner-outer factorization.

The Kalman filter of Bitmead et al. (2019) for $T_o(z)$ may be derived from the state-space model below with appropriate covariances,

$$\begin{bmatrix} x_{t+1} \\ \check{d}_{t+1} \end{bmatrix} = \begin{bmatrix} A & \check{G} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ \check{d}_t \end{bmatrix} + \begin{bmatrix} I_n & 0 \\ 0 & I_m \end{bmatrix} \begin{bmatrix} w_t \\ \delta_t \end{bmatrix},$$
$$y_t = \begin{bmatrix} C & \check{H} \end{bmatrix} \begin{bmatrix} x_t \\ \check{d}_t \end{bmatrix} + v_t,$$

or using the direct construction as in Bitmead et al. (2019), which avoids an explicit model for d_t but yields the same filter.

Marro and Zattoni Marro and Zattoni (2010) provide guidance on the recovery of the disturbance input signal when the d_t-to-y_t system is non-minimum-phase. Their approach involves the approximate inversion of this system using a long delay to accommodate the nominal instability of this inverse. Such techniques are reminiscent of those advanced in B. Widrow, E. Walach (1995). While the approach in Marro and Zattoni (2010) centers on state-estimation first, their development is geometric and noise free and so,

it is unclear how this affects performance. Of course, the geometric analysis throws up the same initial reliance on minimum-phase zeros for stability and exact inversion.

5.5.3 Extension to time-varying systems

Developments so far have been limited to the time-invariant case and have availed themselves of concepts of transmission zeros, stable invertibility and inner-outer factorization, each of which is problematic to extend to time-varying systems. However, since alternative results have been phrased for the time-varying case, we consider this extension now, relying on examination of SISE recursions via Riccati difference equations in the proofs of Theorems 5.2 and 5.4.

Appealing to Gillijns and De Moor (2007a;b) for the time-varying SISE algorithms in the case of Theorem 5.2 and zero direct feedthrough, Riccati equation (5.29) becomes

$$X_{t+1} = \bar{A}_t X_t \bar{A}_t^T - \bar{A}_t X_t C_{2,t}^T (C_{2,t} X_t C_{2,t}^T + R_{2,t})^{-1} \times (\bar{A}_t X_t C_{2,t}^T)^T + \bar{Q}_t,$$

where,

$$\bar{A}_t = A_t(I - G_{t-1}(C_{1,t}G_{t-1})^{-1}C_{1,t}),$$
$$\bar{Q}_t = A_t G_{t-1}(C_{1,t}G_{t-1})^{-1}R_{1,t}(G_{t-1}(C_{1,t}G_{t-1})^{-1})^T A_t^T + Q_t,$$

and, in the case of full-rank feedthrough, (5.31) becomes

$$X_{t+1} = \hat{A}_t X_t \hat{A}_t^T - (\hat{A}_t X_t \bar{C}_{2,t}^T)(\bar{C}_{2,t} X_t \bar{C}_{2,t} + \bar{R}_{2,t})^{-1} \times (\hat{A}_t X_t \bar{C}_{2,t}^T)^T + \hat{Q}_t,$$

where,

$$\hat{A}_t = A_t - G_t \bar{H}_t^{-1} \bar{C}_{1,t}, \quad \hat{Q}_t = Q_t + G_t \bar{H}_t^{-1} \bar{R}_{1,t} \bar{H}_t^{-T} G_t^T.$$

with now time-varying quantities $\{A_t, G_t, \dots, \}$. We may appeal to standard sufficient results, e.g. Jazwinski (1970), Anderson and Moore (2012) and Theorem 5.3 in Anderson and Moore (1981), on the exponential stability of the Kalman filter subject to uniform reachability and detectability. Subject to the uniform satisfaction of time-varying equivalents of Assumptions 5.1, 5.2 and/or 5.3 as appropriate, this extends these stability conditions to the uniformly time-varying case.

5.6 System Inversion and SISE

For the square cases of SISE satisfying Assumptions 5.2 or 5.3, we were able to demonstrate that the SISE d_t-estimator implements exactly the left inverse of the y_t-to-d_t system. The simultaneous x_t-estimate is the state of the inverse system the stability of which depends on the transmission zeros of the original system.

Conditions for left invertibility of a linear time-invariant system are provided by Sain and Massey Sain and Massey (1969) and for stable invertibility by Moylan Moylan (1977) via the Rosenbrock system matrix. Both papers construct the inverse system. Moreover, in Sain and Massey (1969), Sundaram (2020), left invertibility with delay, L, is studied, where stacked measurements $\begin{bmatrix} y_t^T & y_{t+1}^T & \dots & y_{t+L}^T \end{bmatrix}^T$ are used to estimate d_t. Marro and Zattoni blend into this picture stable approximate inversion with delay.

From Sain and Massey (1969), we see that, for any $p \geq m$:

- Assumption 5.2 is the left invertibility condition for $C(zI - A)^{-1}G$ with delay one.
- Assumption 5.3 is the left invertibility condition for $H + C(zI - A)^{-1}G$ with delay zero.
- Assumption 5.4 is the left invertibility condition for $H + C(zI - A)^{-1}G$ with delay one.

Given the input recovery objective of SISE, this is not surprising. But it is interesting to tie these ideas more closely.

It is worth remarking that many presentations of SISE algorithms make connections to 'unbiasedness' and 'optimality' of the state estimate. As Marro and Zattoni (2010), Bitmead et al. (2019) demonstrate, the probabilistic concept of unbiasedness is really tied to a geometric property of the algorithms and the nature of certain subspaces. The optimality of the state estimates is within the class of estimators already satisfying the geometric constraints. As is evident from, say, Theorems 5.1 and 5.2 and the Riccati equation proof, there is no degree of freedom left for the state estimator in the square case and limited degrees of freedom in the non-square case. Indeed, in the square cases, (5.13-5.22) show that $\hat{x}_{t|t}$ is computed by system simulation using the estimated input; the measurements play no further part. The detectability conditions on Theorems 5.2 and 5.4 show how the

remaining degrees of freedom are used in the Riccati difference equations (5.30) and (5.31).

In conclusion, the chapter attempts to unify the collection of SISE algorithms by revealing their explicit connections to system inversion to recover the otherwise unmodeled disturbance input d_t followed by their 'best efforts' subsequent estimation of the state x_t. The result has been to develop necessary and sufficient conditions for stability, at least in the linear time-invariant case, in terms of the transmission zeros of the d_t-to-y_t plant and then the detectability of the subsequent state estimator.

5.7 Appendix

Proof of Theorem 5.1

We loosely follow a calculation from Maciejowski Maciejowski (1985). From (5.14) and the system equations the (time-invariant) transfer function from d_t to $\hat{x}_{t|t}$ via y_t is given by

$$\begin{aligned}\Psi(z) &= \left\{zI - \left[I - G(CG)^{-1}C\right]A\right\}^{-1} zG(CG)^{-1} \times C(zI-A)^{-1}G, \quad (5.25)\\ &= \left\{zI - [I-\Pi]\,A\right\}^{-1} z\Pi(zI-A)^{-1}G,\end{aligned}$$

where we have used $\Pi \triangleq G(CG)^{-1}C$. Write

$$\begin{aligned}&[zI-(I-\Pi)A]^{-1}\, z\Pi\\ &\quad = [zI-(I-\Pi)A]^{-1}\,[z\Pi - zI + (I-\Pi)A] + I,\\ &\quad = -\,[zI-(I-\Pi)A]^{-1}\,(I-\Pi)(zI-A) + I.\end{aligned}$$

Then, since $(I-\Pi)G = 0$,

$$\begin{aligned}\Psi(z) = &\\ &-\,[zI-(I-\Pi)A]^{-1}\,(I-\Pi)(zI-A)(zI-A)^{-1}G\\ &\qquad\qquad + (zI-A)^{-1}G,\\ &= -\,[zI-(I-\Pi)A]^{-1}\,(I-\Pi)G + (zI-A)^{-1}G,\\ &= (zI-A)^{-1}G. \qquad (5.26)\end{aligned}$$

From (5.25), $\Psi(z)$ is the product of two transfer functions and nominally should have $2n$ poles; those at the eigenvalues of A and those at the eigenvalues of $(I-\Pi)A$. The transfer function $zC(zI-A)^{-1}G$ has McMillan degree n with n finite transmission zeros. We see from (5.26) that only poles at the eigenvalues of A are present in Ψ. This implies that the poles due to the eigenvalues of $(I-\Pi)A$ cancel the transmission zeros of $zC(zI-A)^{-1}G$.

Proof of Theorem 5.2

Define the following quantities.

$$\mathcal{H}_t = GM_t - K_tCGM_t + K_t, \quad Z = C_1G,$$
$$Y = (CX_tC^T + R)^{-1} = \begin{bmatrix} Y_1 & Y_2 \\ Y_2^T & Y_3 \end{bmatrix},$$

where Y is divided conformably with C and v_t in (5.24).

From (5.7) and (5.8) the filtered prediction error satisfies

$$\begin{aligned} \tilde{x}_t &\triangleq x_t - x_{t|t} \\ &= (I - \mathcal{H}_tC)A\tilde{x}_{t-1} + (I - \mathcal{H}_tC)w_{t-1} - \mathcal{H}_tv_t. \end{aligned}$$

Whence,

$$\begin{aligned} P_{t|t} &= \operatorname{cov}(x_t|\mathbf{Y_t}) \\ &= (I - \mathcal{H}_tC)(AP_{t-1|t-1}A^T + Q)(I - \mathcal{H}_tC)^T \qquad + \mathcal{H}_tR\mathcal{H}_t^T. \end{aligned}$$

Using (4.3) yields

$$X_{t+1} = A\left((I - \mathcal{H}_tC)X_t(I - \mathcal{H}_tC)^T + \mathcal{H}_tR\mathcal{H}_t^T\right)A^T + Q. \tag{5.27}$$

We show that this discrete Lyapunov equation is also a Riccati difference equation by substituting for $\mathcal{H}_t$ using $CG = \begin{bmatrix} Z^T & 0 \end{bmatrix}^T$.

$$\begin{aligned} \mathcal{H}_t = G\Bigg(&\begin{bmatrix} Z^T & 0 \end{bmatrix} Y \begin{bmatrix} Z \\ 0 \end{bmatrix}\Bigg)^{-1} \begin{bmatrix} Z^T & 0 \end{bmatrix} Y - X_t \begin{bmatrix} C_1^T & C_2^T \end{bmatrix} \\ &Y \begin{bmatrix} Z \\ 0 \end{bmatrix}\Bigg(\begin{bmatrix} Z^T & 0 \end{bmatrix} Y \begin{bmatrix} Z \\ 0 \end{bmatrix}\Bigg)^{-1} \begin{bmatrix} Z^T & 0 \end{bmatrix} Y + X_t \begin{bmatrix} C_1^T & C_2^T \end{bmatrix} Y \\ = &\begin{bmatrix} GZ^{-1} & GZ^{-1}Y_1^{-1}Y_2 \end{bmatrix} + X_t \begin{bmatrix} 0 & C_2(Y_3 - Y_2^TY_1^{-1}Y_2) \end{bmatrix}, \end{aligned} \tag{5.28}$$

Using partitioned matrix inversion with Y gives

$$\begin{aligned} (Y_3 - Y_2^TY_1^{-1}Y_2) &= (C_2X_tC_2^T + R_2)^{-1} \\ Y_1^{-1}Y_2 &= -(C_1X_tC_2^T)(C_2X_tC_2^T + R_2)^{-1}. \end{aligned}$$

Substituting this into (5.28) and (5.27) gives the following Riccati difference equation.

$$\begin{aligned} X_{t+1} = \bar{A}X_t\bar{A}^T - (\bar{A}X_tC_2^T)(C_2X_tC_2^T + R_2)^{-1} \\ \times (\bar{A}X_tC_2^T)^T + \bar{Q}, \end{aligned} \tag{5.29}$$

where,

$$\begin{aligned}\bar{A} &= A(I - G(C_1G)^{-1}C_1),\\ \bar{Q} &= AG(C_1G)^{-1}R_1(G(C_1G)^{-1})^TA^T + Q.\end{aligned} \tag{5.30}$$

Appealing to Theorem 14.3.1 Kailath et al. (2000) (p. 510), provided $[\bar{A}, \bar{Q}^{\frac{1}{2}}]$ is stabilizable and $[\bar{A}, C_2]$ is detectable, then X_t converges to the maximal solution of the algebraic Riccati equation, which is stabilizing.

Now, since by assumption $[A, Q^{\frac{1}{2}}]$ is stabilizable, there exists a $\mathcal{K}$ such that $A - Q^{\frac{1}{2}}\mathcal{K}$ is stable. Taking,

$$\bar{Q}^{\frac{1}{2}} = \begin{bmatrix} Q^{\frac{1}{2}} & AG(C_1G)^{-1}R^{\frac{1}{2}} \end{bmatrix},$$

and $\bar{\mathcal{K}} = \begin{bmatrix} \mathcal{K}^T & R^{\frac{T}{2}} \end{bmatrix}^T$, $\bar{A} - \bar{Q}^{\frac{1}{2}}\bar{\mathcal{K}}$ also is stable. So stabilizability of $[A, Q^{\frac{1}{2}}]$ implies stabilizability of $[\bar{A}, \bar{Q}^{\frac{1}{2}}]$.

Proof of Theorem 5.4

The proof parallels that of Theorem 5.2. Substitute (5.19) and (5.20) into (5.18) to yield

$$\hat{x}_{t+1|t} = (A - \mathcal{L}_tC)\hat{x}_{t|t-1} + \mathcal{L}_ty_t,$$

where $\mathcal{L}_t = AK_t - AK_tHM_t + GM_t$. Then

$$\begin{aligned}\tilde{x}_{t+1|t} &\triangleq x_t - \hat{x}_{t+1|t},\\ &= (A - \mathcal{L}_tC)\tilde{x}_{t|t-1} + w_t - \mathcal{L}_tv_t,\\ X_{t+1} &= (A - \mathcal{L}_tC)X_t(A - \mathcal{L}_tC)^T + \mathcal{L}_tR\mathcal{L}_t^T + Q,\end{aligned}$$

with $X_{t+1} \triangleq \text{cov}(x_{t+1}|\mathbf{Y}^t)$. Dividing K_t and M_t conformably with C^T: $K_t = \begin{bmatrix} K_{1,t} & K_{2,t} \end{bmatrix}$, $M_t = \begin{bmatrix} M_{1,t} & M_{2,t} \end{bmatrix}$, one arrives directly at the following Riccati difference equation.

$$X_{t+1} = \hat{A}X_t\hat{A}^T - \hat{A}X_t\bar{C}_2^T(\bar{C}_2X_t\bar{C}_2 + \bar{R}_2)^{-1}\bar{C}_2X_t\hat{A}^T + \hat{Q}, \tag{5.31}$$

where,

$$\hat{A} = A - G\bar{H}^{-1}\bar{C}_1, \quad \hat{Q} = Q + G\bar{H}^{-1}\bar{R}_1\bar{H}^{-T}G^T.$$

The proof follows as that for Theorem 5.2 using Kailath et al. (2000).

Chapter 6

Paper III: Disturbance and State Estimation in Partially Known Power Networks

This chapter is taken from Abooshahab et al. (2019).

Disturbance and State Estimation in Partially Known Power Networks

Mohammad Ali Abooshahab
Morten Hovd
Robert R. Bitmead
2019 IEEE Conference on Control Technology and Applications (CCTA), **Hong Kong, China** (2019)

Abstract

Due to privacy considerations, fault occurrence, geographical difficulties, lack of sensors and so forth, some parts of a power network are not precisely known. Thus, we name these systems 'partially known power networks'. To perform dynamic state estimation for partially known power networks, we study the application of the simultaneous input and state estimation algorithm to solve the problem. This algorithm jointly estimates the state of the system from a model and, through smoothing, the unmodeled disturbance signals. Although traditional Kalman filtering approaches for state

estimation of a power grid have achieved satisfactory results, they require that all parts of the system including disturbance models be provided, even if imprecisely known, which is problematic especially for the distribution part of power grids.We model the power grid as a system with known and unknown parts and derive the state estimation based solely on the model of the known part of the system with the connected unknown part captured by its disturbance signals. The specific nature of power grid models admits the application of this estimation approach more widely than is suggested by the disturbance reconstruction condition. Simulation results show the effectiveness and the accuracy of the proposed method.

6.1 Introduction

The state estimation problem in power systems has been studied both statically and dynamically Gomez-Exposito and Abur (2004), Tebianian and Jeyasurya (2015). Dynamic state estimation provides a fast response to disturbances and other system changes. The weakness of many previous works, including Aminifar et al. (2014), is that the estimation needs a model of the entire system, including statistical descriptions of external disturbances. However, a typical power system contains several interconnected elements or layers including generation, transmission, distribution and energy consumption. Among them, some variables might be known and modeled accurately, while others might be partially or totally unknown, and external disturbances may have unknown and/or possess time-varying statistical descriptions. Additionally, measurements from the distribution part of the power grid are generally few and restricted due to personal privacy considerations of the consumers and the lack of precise measuring devices. In this case, it is important to develop a method that performs state estimation considering these limitations.

We limit our attention to linear or linearized power systems and use the term *known* to describe a subsystem whose dynamic model is available and accurate in dimension and parameters. *Unknown* indicates the absence of such a precise model. In a power grid such as that illustrated in Fig. 6.1 showing a circuit cut dividing the grid into two parts, the left side is known and the right is unknown. The interaction between these parts if captured entirely by the two current signals flowing between them. We treat these currents as disturbances to the known part of the grid. We shall apply the Simultaneous Input and State Estimation algorithm (SISE) Bitmead et al. (2019) to the available measurements to yield estimates of the state of the known part and of the disturbance signals. This method is indifferent to

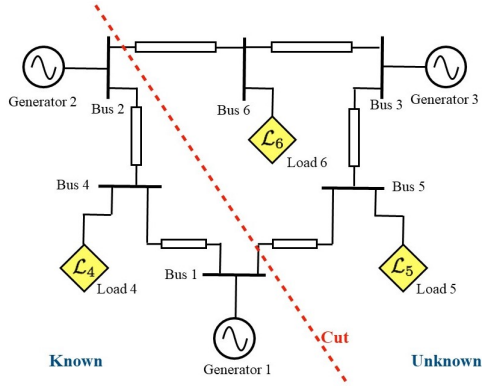


Figure 6.1: Transformerless dynamic power grid model of the Western System Coordinating Council 9-Bus System, WSCC-9, with circuit cut dividing known and unknown parts. Interactions between the parts are via two current signals.

the modeling of the unknown part and, provided an algebraic condition is satisfied, the state estimates are minimum-variance in the absence of a statistical description of the unknown signals.

In Ghahremani and Kamwa (2011a;c), a dynamic state estimator method based on extended Kalman filtering with unknown input is proposed, whereby the exciter output voltage measurement is not available. This method is studied in Ghahremani and Kamwa (2016) in a decentralized fashion. Recently, a derivative-free Kalman filtering based approach of dynamic state estimation for power systems with unknown inputs was proposed in Anagnostou and Pal (2018) that uses an augmented unscented Kalman filter for state estimation. The attention of the above approaches is on the individual network components' external signals ("bottom-up perspective"). On the other hand, an alternative perspective is proposed such that the unknown parts of a partially known power system are modeled as inputs flowing into the known part of the system ("top-down perspective"). Changing the perspective from the individual network components to a holistic view of the known part of the system can help operators monitor partially known power networks more simply.

The contribution of this chapter is to apply the viewpoint of the SISE state and (disturbance) input estimation algorithm to a power system which is split into known and unknown parts. This is the first stage of developing state estimation techniques for interconnected circuits with one (linear) circuit known and interacting with connected (possibly nonlinear) entit-

ies through port currents as external inputs Bitmead et al. (2019). The structural constraints on the SISE method for it to be able to estimate the unknown inputs are interpreted and discussed in a power system setting, where they prove to be more powerful than in the general setting.

This chapter is organized as follows. In Section 2, we introduce and derive the model for the Western System Coordinating Council (WSCC) 9-bus system Anderson and Fouad (2008), which is used to demonstrate the proposed state estimation method. Section 3 is devoted to the introduction of the state estimation algorithm, including the SISE algorithm in its Kalman filtering formulation. Moreover, the designed algorithms are developed for state estimation of the WSCC-9-bus network. The simulation results related to the WSCC-9-bus network are presented in Section 4 to verify the effectiveness of the proposed methods.

6.2 WSCC-9 Model Derivation

Because the example provides context for the methodology, we present it first. The WSCC-9 model in Anderson and Fouad (2008) is static. However, in order to have a dynamic model for state estimation, we make some modifications that are described below and in Table 6.1. For simplicity, we make the following assumptions:

Assumption 6.1. *1. We ignore transformers in our dynamic estimation procedure since their impedances are negligible compared to the impedance of the transmission lines.*

2. We ignore the conductances of transmission lines Hertem et al. (2006).

3. Bus 1 is a slack bus.

The 9-bus power network consists of three generators and three consumers or loads. The dynamics of generators 1, 2 and 3 – their moments of inertia, feedback control gains, mechanical and electrical properties – are available in the full model along with the admittances of the network. The consumer loads 4, 5 and 6 are also reliably modeled in the full model.

Ignoring the transformers, the 9-bus system can be reduced to a 6-bus system with three dynamic loads and three generators shown in Fig. 7.7. Generators 1, 2 and 3 are rotating machines that are described by their mechanical dynamics. The models presented in most works in this area, such as those investigating quasi-steady state estimation, are linearized incremental models describing the (slow) time evolution of system variables in

a periodically time-varying AC system. It is important to bear this in mind when interpreting the dynamic equations since they mask the (fast) variation of currents and voltages at mains frequency. Generators 1, 2 and 3 are rotating machines that are described by closed-loop mechanical dynamics Blood (2011):

$$\frac{d}{dt}\begin{bmatrix} a \\ P_m \\ \omega_r \\ \delta \end{bmatrix}_i = \begin{bmatrix} -kR & 0 & k & 0 \\ 1/T_{CH} & -1/T_{CH} & 0 & 0 \\ 0 & -1/M & -D/M & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}_i \begin{bmatrix} a \\ P_m \\ \omega_r \\ \delta \end{bmatrix}_i + \begin{bmatrix} -k & -k \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}_i \begin{bmatrix} L_{ref} \\ \omega_0 \end{bmatrix}_i + \begin{bmatrix} 0 \\ 0 \\ -1/M \\ 0 \end{bmatrix}_i P_{Ei}, \quad i = 1, 2, 3, \quad (6.1)$$

where variables and parameters are:

state variables: valve position a, angular frequency ω_r, mechanical power P_m, shaft angle δ;
parameters: governor feedback gain k, droop characteristic R, rotor damping D, rotor inertia M, machine time constant T_{CH};
external variables: external load P_E, frequency set point ω_0, load set point L_{ref}.

Similarly, loads containing rotating machines with known power, P_L, can be described as

$$\frac{d}{dt}\begin{bmatrix} \omega_r \\ \delta \end{bmatrix}_i = \begin{bmatrix} -D/M & 0 \\ 1 & 0 \end{bmatrix}_i \begin{bmatrix} \omega_r \\ \delta \end{bmatrix}_i + \begin{bmatrix} -1/M \\ 0 \end{bmatrix}_i [P_{Li}] + \begin{bmatrix} -1/M \\ 0 \end{bmatrix}_i P_{Ei}. \quad (6.2)$$

The component dynamic models of each generator or load in bus i can be rewritten concisely as

$$\dot{x}_i = A_i x_i + B_i u_i + B_i^{(P)} P_{Ei}. \quad (6.3)$$

where $B_i^{(P)}$ is external load matrix. Note that ω_0 and L_{ref} do not affect the state estimation procedure and can be neglected. The combination of the generator and load models is as follows

$$\begin{aligned} \mathbf{x} &= [x_1^T, \cdots, x_6^T]^T \\ \mathbf{P} &= [P_{E1}^T, \cdots, P_{E6}^T]^T \\ \boldsymbol{\delta} &= [\delta_1^T, \cdots, \delta_6^T]^T \\ \mathbf{A} &= \text{blockdiag}(A_1, \cdots, A_6) \\ \mathbf{B^{(P)}} &= \text{blockdiag}(B_1^{(P)}, \cdots, B_6^{(P)}). \end{aligned}$$

The susceptance matrix of the reduced model of the 9-bus network in Fig.7.7 relates the currents out of the six nodes to the voltages at the nodes. Considering the reduced model of the WSCC-9-bus system, the susceptance matrix for the network is

$$\mathrm{B} = \begin{bmatrix} -0.334 & 0 & 0 & 0.176 & 0.158 & 0 \\ 0 & -0.455 & 0 & 0.306 & 0 & 0.149 \\ 0 & 0 & -0.567 & 0 & 0.358 & 0.209 \\ 0.176 & 0.306 & 0 & -0.482 & 0 & 0 \\ 0.158 & 0 & 0.358 & 0 & -0.518 & 0 \\ 0 & 0.149 & 0.209 & 0 & 0 & -0.358 \end{bmatrix},$$

In fact, the power injections are nonlinear (sinusoidal) functions of the bus voltage angles Blood (2011). However, by considering Assumption 6.1, the power injections can be approximated as:

$$\mathbf{P} = \mathrm{B}\boldsymbol{\delta}.$$

Since the inputs are known and the models are linear, we may omit the inputs from the description,

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}^{(\mathbf{P})}\mathrm{B}\boldsymbol{\delta},$$

then,

$$\dot{\mathbf{x}} = (\mathbf{A} + \mathbf{B}^{(\mathbf{P})}\mathrm{B}\mathbf{S}_\delta)\mathbf{x},$$

where $\mathbf{S}_\delta$ is a δ selection matrix $\boldsymbol{\delta} = \mathbf{S}_\delta\mathbf{x}$. We re-organize the state vector and the system matrix for each component as follows,

$$x_i = \begin{bmatrix} \hat{x}_i \\ \delta_i \end{bmatrix}, \quad i = 1, \cdots, 6$$

$$\begin{aligned} A_i &= \left[\begin{array}{cc:cc} & G_i & & 0 \\ 0 & 0 & 1 & 0 \end{array}\right], \quad i = 1, \cdots, 3 \\ A_i &= \left[\begin{array}{cc:cc} & L_i & & 0 \\ 1 & & 0 & \end{array}\right], \quad i = 4, \cdots, 6 \end{aligned} \tag{6.4}$$

where G_i and L_i are generators and loads dynamic states except for their power angle.

For the re-organized system state vector, $\bar{\mathbf{x}} = [\hat{x}_1^T, \cdots, \hat{x}_6^T, \delta_2^T, \cdots, \delta_6^T]^T$, the

interconnected network has the system matrix as in (6.5),

$$\mathcal{A} = \left[\begin{array}{cccccccccccc:c}
 & & & & & & & & & & & & \mathcal{I}_1 \\
 & & & & & & & & & & & & \mathcal{I}_2 \\
 & & & & & & & & & & & & \mathcal{I}_3 \\
 & & & & & & & & & & & & \mathcal{I}_4 \\
 & & & & & & & & & & & & \mathcal{I}_5 \\
 & & & & & & & & & & & & \mathcal{I}_6 \\
\hdashline
0 & 0 & -1 & 0 & 0 & 1 & & & & & & & \\
0 & 0 & -1 & & & & 0 & 0 & 1 & & & & \\
0 & 0 & -1 & & & & & & & 1 & & & \\
0 & 0 & -1 & & & & & & & & 1 & & \\
0 & 0 & -1 & & & & & & & & & 1 &
\end{array}\right], \tag{6.5}$$

where $G_i,\ i = 1, 2, 3$ and $L_i,\ i = 4, 5, 6$ are defined in (6.4) and from (6.1)

$$\mathcal{I}_i = \begin{bmatrix} 0 \\ 0 \\ B(i, 2:6)/M_i, \end{bmatrix}, \quad i = 1, 2, 3,$$

$$\mathcal{I}_i = \begin{bmatrix} B(i, 2:6)/M_i, \end{bmatrix}, \quad i = 4, 5, 6.$$

Throughout the chapter, the simulations are performed with the full-order linear model for the entire grid driven by load disturbance signals taken from measured consumer data from Électricté de France Hebrail (2012). Thus, there are no exact or statistically accurate disturbance models. So the simulations provide a realistic test of the algorithms.

6.3 The SISE algorithm for interconnected systems

The aim of this section is to introduce the SISE algorithm and its formulation, and reveal the advantages of the SISE for our power network state estimation problem. We will use this algorithm in the next section in two computational state estimation examples associated with different compartmentalizations into known and unknown parts associated with power system circuit cuts. SISE is used to estimate the state of the known part and the disturbance signals impinging from the unknown parts of the power system.

6.3.1 The classic SISE algorithm

Consider the following linear time-invariant system:

$$x_{t+1} = \mathcal{A}x_t + Gd_t + w_t, \tag{6.6}$$

$$y_t = Cx_t + v_t, \tag{6.7}$$

where x_t is the state vector at time step t, d_t is the unknown part modeled as input demand estimation, and y_t is the measurement vector. We make the following basic assumptions.

Assumption 6.2. *1.* $x_t, w_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^q$, $d_t \in \mathbb{R}^m$, $v_t, y_t \in \mathbb{R}^p$.

2. *The initial condition* $x_0 \sim \mathcal{N}(\hat{x}_{0|0}, P_0)$, *and noise sequences* $w_t \sim \mathcal{N}(0, Q)$, $v_t \sim \mathcal{N}(0, R)$ *are Gaussian and independent with the noises being white.*

3. $R_t > 0$,

4. $[\mathcal{A}, C]$ *is completely observable,*

5. $rank CG = rank G = m$.

Then, the simultaneous input and state estimation (SISE) algorithm, Kitanidis (1987), Gillijns and De Moor (2007a), is as follows.

$$X_{t+1} = \mathcal{A} P_t \mathcal{A}^T + Q,$$

where X_{t+1} is the prior state covariance matrix

$$K_{t+1} = X_{t+1} C^T (C X_{t+1} C^T + R)^{-1},$$

where K_{t+1} is the Kalman gain for the state vector

$$M_{t+1} = [G^T C^T (C X_{t+1} C^T + R)^{-1} C G]^{-1} \times G^T C^T (C X_{t+1} C^T + R)^{-1},$$

where M_{t+1} is the Kalman gain for the unknown input vector

$$P_{t+1} = (I - K_{t+1} C) \left[(I - G M_{t+1} C) X_{t+1} (I - G M_{t+1} C)^T + G M_{t+1} R M_{t+1}^T G^T \right] + K_{t+1} R M_{t+1}^T G^T,$$

where P_{t+1} is the posterior state covariance matrix

$$\hat{d}^{\text{SISE}}_{t|t+1} = M_{t+1}(y_{t+1} - C\mathcal{A}\hat{x}^{\text{SISE}}_{t|t}),$$
$$\hat{x}^{\text{SISE}}_{t+1|t+1} = \mathcal{A}\hat{x}^{\text{SISE}}_{t|t} + G\hat{d}^{\text{SISE}}_{t|t+1} + K_{t+1} \times (y_{t+1} - C\mathcal{A}\hat{x}^{\text{SISE}}_{t|t} - CG\hat{d}^{\text{SISE}}_{t|t+1}).$$

where $\hat{x}^{\text{SISE}}_{t+1|t+1}$, $\hat{d}^{\text{SISE}}_{t|t+1}$ are posterior estimates for the state and the unknown input of the system.

The SISE algorithm subject to Assumption 6.2 yields asymptotically unbiased filtered state estimates and smoothed disturbance signal estimates. With zero mean noises, the state estimates are minimum-variance subject to projection into the null space of G Bitmead et al. (2019).

6.3.2 State estimation of interconnected systems

SISE algorithms have usually been motivated by and applied to input estimation problems in geophysical exploration and environmental monitoring, effectively as an approach to signal deconvolution through a known linear system. However, early approaches Friedland (1969) concentrated on bias estimation. As a byproduct of this deconvolution, a state estimate is produced. We focus on the state estimation of a linear system, depicted as $P_1(z)$ in Fig. 6.2, interconnected with an unknown system $\mathcal{S}_2$. We use

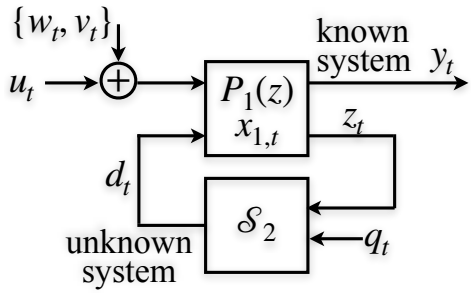


Figure 6.2: Depiction of interconnected known linear system $P_1(z)$ with state $x_{1,t}$ to be estimated, unknown system $\mathcal{S}_2$ and unmeasured interconnection signals d_t and z_t. Signals $\{w_t, v_t, u_t, q_t\}$ are the noises and known input of P_1, and other external driving signals, respectively.

the word *depicted* advisedly, since there is not necessarily any causal relationship implied between d_t and the other signals. However, the analysis of interconnected systems exposes a potentially powerful application of SISE in the capacity to estimate the state $x_{1,t}$ in isolation from its interconnections, which are entirely captured by d_t. This is an n-vector state estimation problem regardless of the complexity of the jointly interconnected system $\mathcal{S}_2$. Further, the linearity of the known system P_1, but not necessarily of $\mathcal{S}_2$, is all that required, along with observability of P_1 and Assumption 6.2.5 on the sufficiency of the measurements, which in turn relies on the topology and nature of interconnection between the systems, but not the internal dynamics of the known parts of the system.

6.4 Power network state estimation

In this section, we perform state estimation on the WSCC-9 test case with unknown parts. We make a cut at the line through the network diagram separating the known and the unknown parts of the system. The unknown part of the circuit interacts solely through the two current signals flowing

across the cut. We perform different cuts to see their effect on the state estimation feasibility. With each cut, the definitions of the matrices in the SISE description (6.6)-(6.7) alters to reflect the topology. Then, we apply SISE to estimate the states of generators 1 and 2 and the current signals from the unknown part. Our procedure for state estimation consists of the following steps:

1. Identify the known and the unknown parts of the system.
2. Model the unknown part by disturbance inputs into the known part.
3. Derive the subsystem model.
4. Examine the rank CG condition from Assumption 6.2.5.
5. Apply SISE to estimate the states and the unknown currents.

6.4.1 The first cut

The first cut, shown in Fig. 7.7, divides the network into the known and the unknown parts with the known part containing generators 1 and 2, dynamic load 4 and transmission lines between load 4 and generators 1 and 2. The unknown part contains generator 3 and loads 5 and 6 and their corresponding transmission lines. Note that we are using a complete model to simulate the full system, while the estimation is based on a partial model only.

The partial (known) system model is

$$\frac{d}{dt}\begin{bmatrix} a_1 \\ P_1 \\ \omega_1 \\ a_2 \\ P_2 \\ \omega_2 \\ \omega_4 \\ \delta_2 \\ \delta_4 \end{bmatrix} = \left[\begin{array}{ccccccccc} -5 & 0 & 100 & & & & & 0 & 0 \\ 0.2 & -0.2 & 0 & & & & & 0 & 0 \\ 0 & -0.1 & -0.15 & & & & & 0 & .02 \\ \hdashline & & & -5 & 0 & 125 & & 0 & 0 \\ & & & 0.3 & -0.3 & 0 & & 0 & 0 \\ & & & 0 & -0.2 & -0.3 & & -.1 & .06 \\ \hdashline & & & & & & -1.5 & .3 & -.5 \\ \hdashline & & -1 & & & 1 & & & \\ & & -1 & & & & 1 & & \end{array}\right] \times \begin{bmatrix} a_1 \\ P_1 \\ \omega_1 \\ a_2 \\ P_2 \\ \omega_2 \\ \omega_4 \\ \delta_2 \\ \delta_4 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1/M_1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1/M_2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{P}_{1,5} \\ \hat{P}_{2,6} \end{bmatrix}. \quad (6.8)$$

where $\hat{P}_{1,5}$ and $\hat{P}_{2,6}$ the unknown power disturbances entering bus 1 from bus 5, and entering bus 2 from bus 6, respectively. Then, a linear time-invariant system for the first cut can be rewritten as

$$\frac{dx_1}{dt} = \mathcal{A}x_1 + G_1 d_1, \quad y_1 = \begin{bmatrix}\omega_1\\ \omega_2\end{bmatrix} = C_1 x_1 = \begin{bmatrix}0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0\end{bmatrix} x_1. \tag{6.9}$$

Figures 6.3 and 6.4 show the performance of SISE's state estimation and disturbance estimation, respectively. The first six state estimates from SISE closely track the true states, the states of generators 1 and 2. Fig. 6.3(a) shows P_2 and its corresponding estimation. Load 4's speed and its estimate are shown in Fig. 6.3(b). The phase difference between bus 4 and 1 and its estimates is shown in Fig. 6.3(d). According to Fig. 6.4, we can deduce that while the estimated disturbances are somewhat less accurate than the estimates of the states of the known system, the disturbance estimates still reliably show the trends in the disturbances.

6.4.2 The second cut

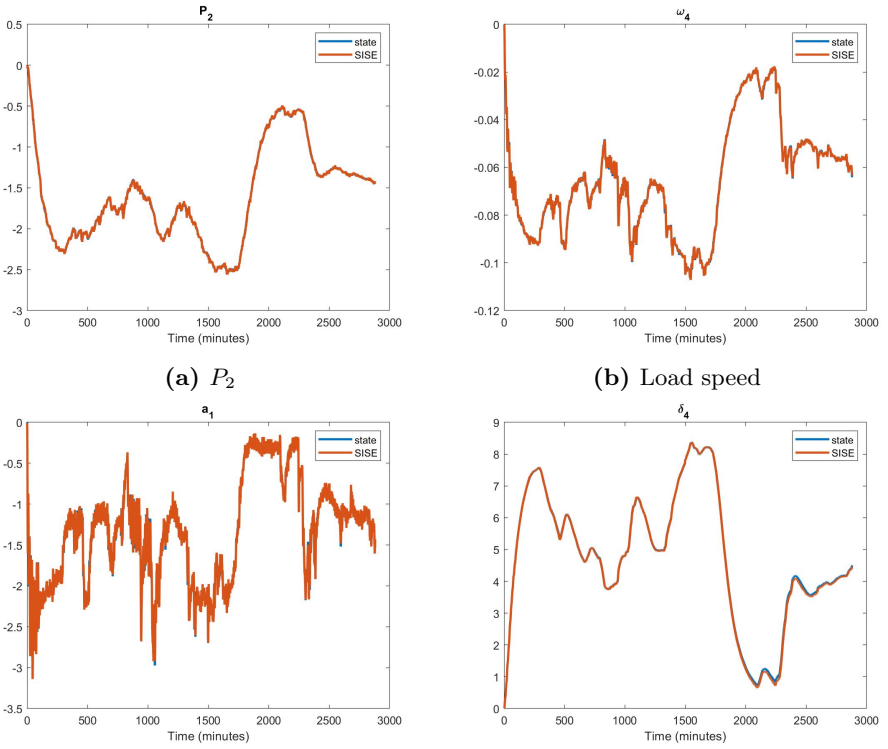
The network division is amended by replacing the first cut with that is shown in Fig. 6.5. The load at bus 4 is now part of the unknown part of the system and provides a third disturbance current signal to the network. The same two measurements, ω_1 and ω_2, are retained. But now, with three disturbance signals, Assumption 6.2.5 can no longer be satisfied. The system model for the known part associated with the second cut is given by:

$$\frac{d}{dt}\begin{bmatrix}a_1\\ P_1\\ \omega_1\\ a_2\\ P_2\\ \omega_2\\ \delta_2\end{bmatrix} = \left[\begin{array}{ccccccc} -5 & 0 & 100 & & & & 0\\ 0.2 & -0.2 & 0 & & & & 0\\ 0 & -0.1 & -0.15 & & & & 0\\ \hdashline & & & -5 & 0 & 125 & 0\\ & & & 0.3 & -0.3 & 0 & 0\\ & & & 0 & -0.2 & -0.3 & -0.09\\ \hdashline & & -1 & & & 1 & 0 \end{array}\right] \begin{bmatrix}a_1\\ P_1\\ \omega_1\\ a_2\\ P_2\\ \omega_2\\ \delta_2\end{bmatrix}$$

$$+ \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0.0176 & 1/M_1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\\ 0.06120 & 0 & 1/M_2\\ 0 & 0 & 0\end{bmatrix} \begin{bmatrix}\hat{\delta}_L\\ \hat{P}_{1,5}\\ \hat{P}_{2,6}\end{bmatrix},$$

$$\frac{dx_2}{dt} = \mathcal{A}x_2 + G_2 d_2,$$

$$y_2 = \begin{bmatrix}\omega_1\\ \omega_2\end{bmatrix} = C_2 x_2 = \begin{bmatrix}0 & 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0\end{bmatrix} x_2. \tag{6.10}$$



(a) P_2 **(b)** Load speed

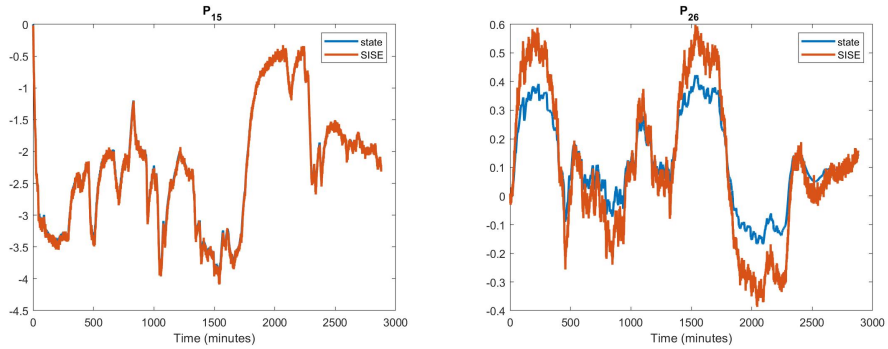
(c) a_1. **(d)** δ_4

Figure 6.3: System states estimated for the first cut: P_2 (mechanical power at bus 2), ω_4 (load speed at buss 4), a_1 (valve position of the generator at bus 1) and δ_4 (phase difference between buses 4 and 1). The actual value of the state is shown in blue, while the SISE estimate is shown in red.

where $\hat{P}_{1,5}$ and $\hat{P}_{2,6}$ are the same as the first cut, while $\hat{\delta}_L$ is the unknown power angle of the load at bus 4 (which determines the power drawn by the load).

The second cut defines three disturbances and, as remarked above, the two measurements do not suffice for precise reconstruction of the state and the disturbance signals, because Assumption 6.2.5 does not hold. Fig. 6.6 shows the degradation of estimation quality for $\hat{P}_{26}$, and the good estimation quality of a_1. Although Assumption 6.2.5 is not satisfied, we can still run SISE as our state estimator which leads to Lemma 6.1 below.

According to Assumption 6.2.5, one remedy to this collapse of disturbance signal reconstruction could be to include an additional measurement – say



(a) $\hat{P}_{15}$ **(b)** $\hat{P}_{26}$

Figure 6.4: External inputs/disturbances, $\hat{P}_{15}$ and $\hat{P}_{26}$, in blue and its estimates with SISE in red for the first cut.

a PMU at Bus 4 – yielding a new 3×7 measurement matrix C_2' satisfying $\operatorname{rank} C_2' G_2 = 3$. Indeed, this is a strength of SISE that the estimation conditions are simply checkable and easily related to the measurement location.

However, the surprising aspect of the SISE applied to the second cut is that the state estimates for the known part do not degrade in spite of the compromised disturbance reconstruction. This will be explained below.

6.4.3 Recoverable disturbance directions

In the previous sub-section, we saw that $rank(CG) = 2$, but we still were able to reconstruct all the signals in the known part of the system. In the second cut, there are three unknown disturbances to the known part of the system denoting with $[\hat{\delta}_L, \hat{P}_{1,5}, \hat{P}_{2,6}]$. However, it is apparent from (6.10) that the span of the three disturbance signals lies in a two-dimensional subspace of the state space $\mathbb{R}^n$. That is, no matter the set of disturbance signals, the vector Gd_t in (6.6) above belongs to this subspace. Indeed, this observation holds for any cut which leaves the known system observable.

A consequence of this property for the known part of the system for cut 2 is that all possible disturbances manifest as an equivalent set of two disturbances as in cut 1. Accordingly, the state estimation proceeds without issue and the Gd_t term is recovered but the separation of Gd_t into its constituent components can fail. We have the following result derived from Bitmead et al. (2019).

Lemma 6.1. *If* $\operatorname{rank} CG = \operatorname{rank} G = k$, *then the dimension-*$k$ *subspace*

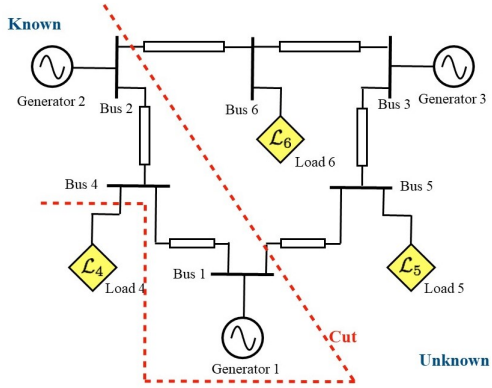


Figure 6.5: Performing the second cut. In this cut, load 4 is also assumed to be unknown, and the operator has only information and measurements from generator 1 and 2.

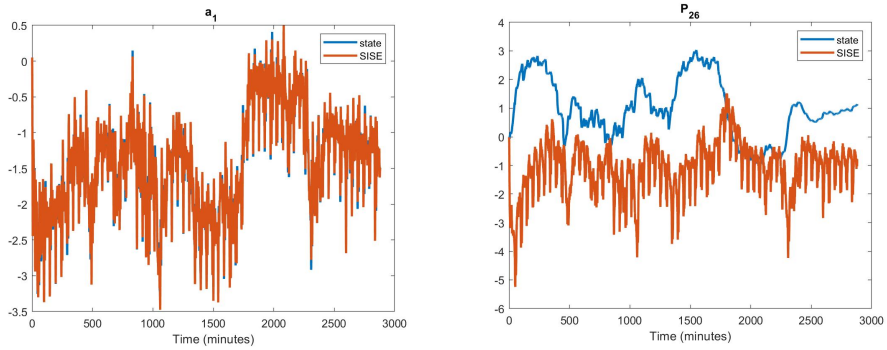
spanned by Gd_t *is recoverable by SISE with one-step-smoothed covariance*

$$G\left[G^TC^T(C\mathcal{A}P_{t+1}\mathcal{A}^TC^T+CQC^T+R)^{-1}CG\right]^{-1}G^T,$$

which is independent of the power of d_t*.*

6.4.4 Topology analysis

A central issue is whether the solution of our proposed method is feasible in the practical and physical power network. In the second cut, we see that we had only two paths of disturbances into the system. Analytically, it means $rank(CG) = 2$. Hence, the rank of G represents the maximum number of independent disturbance paths into the known system, and if the number of the unknown disturbances into the system is more than the rank of G, then we lose the ability to estimate the unknown disturbances affecting the system. Generally, the unknown disturbances in the electrical networks are power type signals, so they depend on the power angles of other network buses. Hence, for each rotating machine in the power network, there is only one input path into the system. For example, if we have two rotating machines in the known part of the network, then only two disturbance paths are identifiable. Thus, the important result is that if we have n power disturbance paths in the power system needing to be estimated, beside Assumption 6.2, we require at least n states connected to those power disturbances. This can be illustrated with a cut with unknown loads shown in Fig. 6.7. Since we have three disturbance inputs and three



(a) a_1 **(b)** $\hat{P}_{26}$

Figure 6.6: State, a_1 (valve position of the generator at bus 1), and external input, $\hat{P}_{26}$, in blue and their estimates for the second cut.

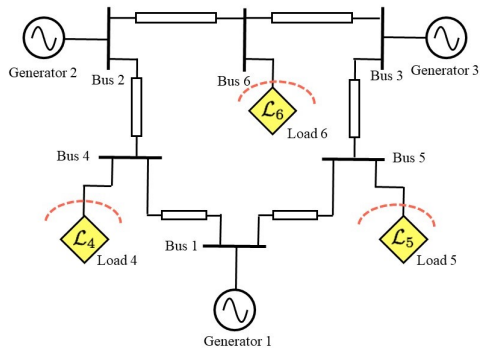


Figure 6.7: Load separation scheme. Loads 4, 5, and 6 are unknown parts of the network, and the known part contains generators 1, 2, 3 and the transmission lines.

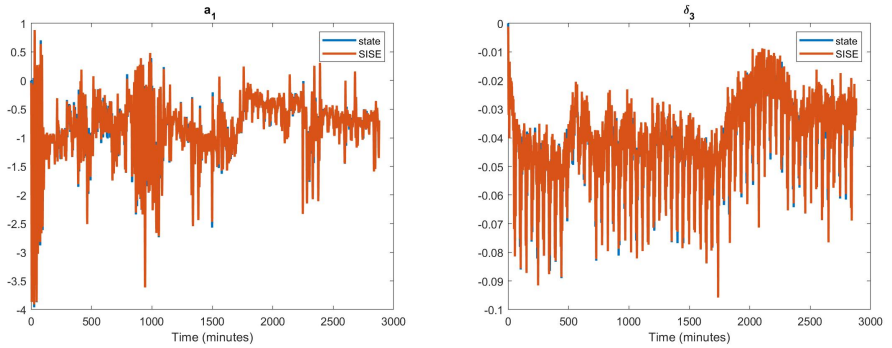
generators, the SISE would be able to estimate all three unknown disturbances. Thus, we can assume that the load parts are the unknown part of the power grid using the proposed cut. For this cut, the disturbance matrix is given as follows,

$$G_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.0176 & 0.0158 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.0612 & 0 & 0.0298 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0.0716 & 0.0418 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

so, by choosing

$$C_3 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix},$$

results in Fig. 6.8 and Fig. 6.9 can be obtained demonstrating good accuracy state estimation for the proposed cut.

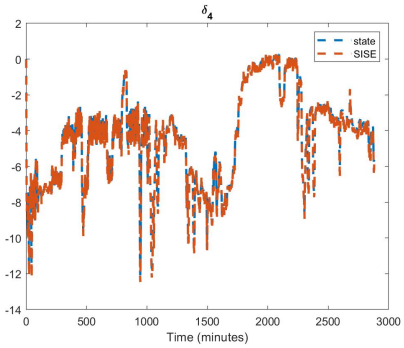


(a) a_1 **(b)** δ_3

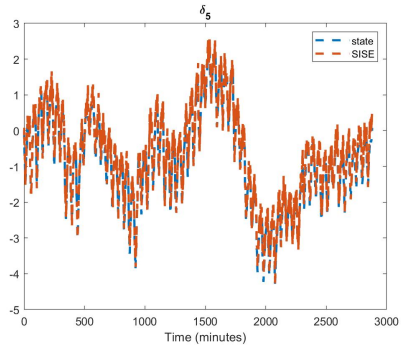
Figure 6.8: The estimates of the states of the known part of the system for the load separation. The blue line is the actual signal, and the red line is its estimate.

6.5 Conclusion

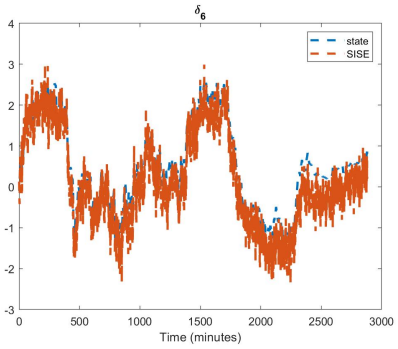
We apply the SISE method to dynamic state estimation in power grids with known and unknown parts. This is done by moving the focus of SISE from deconvolution and input estimation to state estimation. Thus we have been able to present, and illustrate an approach to state estimation in networked systems, including those with poorly known network neighbors. In addition, different cuts, their requirements, and their impacts have been investigated for a well-known power model. Simulation results reveal that the proposed approach can obtain accurate results.



(a) δ_4



(b) δ_5



(c) δ_6

Figure 6.9: The estimates of all unknown disturbances for the load separation. The blue line is the actual signal, and the red line is its estimate.

$D_i, i = 1, ..., 6$	1.5
$T_{CH,1}$	$5s$
$T_{CH,2}$	$3s$
$T_{CH,3}$	$4s$
R_1	0.05
R_2	0.04
R_3	0.033
k_1	100
k_2	125
k_3	150
M_1	10
M_2	5
M_3	8.33
M_4	1
M_5	3.33
M_6	3.33
B_{14}	0.176
B_{15}	0.158
B_{24}	0.306
B_{26}	0.149
B_{35}	0.358
B_{36}	0.209

Table 6.1: Power network parameters

Chapter 7

Paper IV: Monitoring Disturbances and States in Partially Known Power Systems

Monitoring Disturbances and States in Partially Known Power Systems

Mohammad Ali Abooshahab
Morten Hovd
Robert R. Bitmead
In preparation for submission to *IEEE Transactions on Power Systems*, (2021)

Chapter 8

Paper V: Distributed H_∞ Filtering for Linear and Nonlinear Systems

Distributed H_∞ Filtering for Linear and Nonlinear Systems

Mohammad Ali Abooshahab
Morten Hovd
submitted to and accepted in *2021 IEEE Conference on Control Technology and Applications (CCTA)*, **San Diego, USA** (2021)

Abstract

This chapter presents a fully distributed framework for implementing H_∞ filtering over multi-agent networked systems. This is obtained through using diffusion techniques for fusion of local filtering operations in order to enforce cooperation between agents and achieve a network-wide cohesive filtering operation. More importantly, we propose a diffusion-based algorithm which uses only locally observable states. Furthermore, the work includes the information formulation of the derived filtering framework. This information formulation not only provides the basis of the performance analysis and establishment of observability conditions, but also allows for the extension of the derived filtering framework to nonlinear systems via the use of the

unscented transform. Finally, the effectiveness of the derived estimation framework is demonstrated in two simulation examples.

8.1 Introduction

Since its introduction in the 1960s, the Kalman filter has seen use for a variety of purposes Kalman (1960), Kailath et al. (2000). However, the Kalman filter assumes accurate knowledge of the system model characteristics, including noise processes Kalman (1960). Unfortunately, this information is not readily available in most industrial applications Kumar (2015). In this setting, the class of H_∞ filters, that require no prior knowledge of noise characteristics, have attracted significant attention within the signal processing and control communities Li and Jia (2010), Wang et al. (2012). Unlike the Kalman filter which implements a minimum variance estimator based on a Gaussian assumption on the noise processes, the class of H_∞ filters aim to reduce the impact of the extreme case for the disturbance on the estimation error. H_∞ filters are therefore more robust against disturbances Simon (2006). The H_∞ filter is described as a filter/predictor minimizing the power density of the peak error while the classic Kalman filter reduces the average power density of error Simon (2006), Li and Jia (2010), Grimble and El Sayed (1990). Due to their natural ability to deal with model uncertainty, the class of H_∞ filters have found numerous applications, including signal processing Hu and Yang (2011), Ding and Guo (2015) power network monitoring Zhao and Mili (2018), Zhao et al. (2019b), Wang et al. (2019), and robotics Loo et al. (2019), Havangi (2015).

Another area that has seen increased interest is that of networked multi-agent state estimation applications Yin et al. (2018), Boem et al. (2019). Although optimal solutions are available through the framework of centralized state estimators Rego et al. (2019), in large-scale systems such as power distribution networks, centralized techniques impose high computational loads on the central processing unit and require complex communication protocols. This has led to the introduction of a wide range of distributed Kalman filtering techniques Sayed (2014b), Talebi and Werner (2019). In this setting, agents, i.e., buses in the power network, are able to communicate with their neighbors, while no centralized estimation strategy is present. In addition, there is no hierarchy for the agents in the network Cattivelli and Sayed (2010), Talebi and Werner (2019), Olfati-Saber (2007). Distributed Kalman filtering techniques for networked systems are shown to be robust against agent failure, more reliable when facing unforeseen topology changes, and do not require complicated communication protocols Cattivelli and Sayed

(2010), Talebi and Werner (2019), Olfati-Saber (2007). There are two major classes of distributed Kalman filtering techniques, consensus-based Olfati-Saber (2009) and diffusion-based Sayed (2014b), Cattivelli and Sayed (2009; 2010).

Akin to their single agent counterparts, distributed Kalman filtering techniques also assume knowledge of system model characteristics. In order to present a more robust distributed state estimation solutions, a consensus-based distributed H_∞ filter has been proposed in Ugrinovskii (2013). However, consensus-based distributed filtering schemes burden the network with timing constraints; furthermore, these approaches require the estimator implement the consensus step at faster rates than the local Kalman filtering operation Talebi and Werner (2018). A so-called decentralized H_∞ filter has been introduced in Li and Jia (2010); however, no comprehensive diffusion step is integrated into the algorithm. Thus, the estimation accuracy using local filters are lower than the truly distributed approaches.

In this work, a class of fully distributed H_∞ filtering techniques for state estimation in large-scale multi-agent networks is derived. Distributed filtering is achieved through exploiting the diffusion technique in order to coordinate the flow of information within the multi-agent network and achieve fusion of local filtering operations, thereby accomplishing a cohesive filtering operation. Importantly, the effectiveness of the derived framework is investigated and conditions for observability are provided. Furthermore, in order to loosen observability conditions to practical levels achievable in most real-world applications, a novel diffusion filtering method that allows each agent to track locally observable dynamic modes, while ascertaining sufficient information regarding its unobservable dynamic modes through the diffusion process is derived. For completeness, the framework is also extended to deal with nonlinear system equations using the unscented transform. Finally, an implementation of the framework is applied to target tracking problem and the IEEE-14-bus network, demonstrating the effectiveness of the proposed framework.

8.2 Preliminaries & Background

Consider the linear discrete-time dynamic model

$$x_{i+1} = F_i x_i + w_i \tag{8.1}$$

$$y_i = H_i x_i + v_i \tag{8.2}$$

where at time instant i, the system state and measurement are denoted by $x_i \in \mathbb{R}^n$ and $y_i \in \mathbb{R}^m$, while the system dynamics and the observa-

tion matrix are denoted by F_i and H_i, with the process noise, w_i, and the measurement noise, v_i, being energy bounded $l_2[0, +\infty)$ signals which have unknown statistical properties, so that

$$\sum_{i=0}^{\infty} w_i^{\mathsf{T}} w_i < \infty \text{ and } \sum_{i=0}^{\infty} v_i^{\mathsf{T}} v_i < \infty. \tag{8.3}$$

The aim is to estimate x_i on the basis of measurements prior to and counting time i. According to the methodology of game theory, nature takes the role of an adversarial player, which can, in theory select any value for $\{w_i, v_i, x_0\}$ given the objectives. In contrast, in the Kalman filtering arena, the probability density functions (pdf)s of noises in the system are assumed to be known Simon (2006). Subsequently, this pdf knowledge is used to obtain a minimum variance state estimate. In this case, nature's possible actions to degrade the state estimate are constrained by the prescribed pdf's of the process and measurement noises Banavar (1992), Simon (2006).

In the H_∞ filtering problem, , the following cost function is considered:

$$J_1 = \frac{\sum_{j=0}^{i} \|x_j - \hat{x}_j\|_2^2}{\|x_0 - \hat{x}_0\|_{\mathcal{P}_0^{-1}}^2 + \sum_{j=0}^{i} (\|w_j\|_{Q_j^{-1}}^2 + \|v_j\|_{R_j^{-1}}^2)} \tag{8.4}$$

where the estimate of x_i is denoted by $\hat{x}_i$, with Q_i and R_i being weighting matrices that are analogous to the covariance estimates of w_i and v_i in classical state-space filtering approaches, while $\mathcal{P}_0^{-1} > 0$ is a weighting matrix that can be used to integrate *a priori* information about the accuracy of the initial state estimate $\hat{x}_0$. The solution is to find $\hat{x}_i$ that minimizes the objective function in (8.3), while the opponent's, henceforth referred to as "nature", goal is to find $\{w_i, v_i, x_o\}$ that maximize the error term $(x_i - \hat{x}_i)$ Simon (2006).

The solution of J_1 is not attainable in a straightforward fashion; therefore, an approximation method that meets a performance threshold is sought Simon (2006). In this setting, we have

$$J_1 < \lambda^2 \tag{8.5}$$

where λ is a user-specified performance bound. Substituting (8.4) into (8.5) results in

$$\begin{aligned} J = &- \lambda^2 \|x_0 - \hat{x}_0\|_{\mathcal{P}_0^{-1}}^2 \\ &+ \sum_{j=0}^{i} [\|x_j - \hat{x}_j\|_2^2 - \lambda^2 (\|w_j\|_{Q_j^{-1}}^2 + \|v_i\|_{R_j^{-1}}^2)] < 0. \end{aligned} \tag{8.6}$$

A sub-optimal solution to the H_∞ filtering problem given in (8.6) can be reached through the following iterations Einicke and White (1999), Li and Jia (2010), Simon (2006)

At each time instant i:

$$\hat{x}_{i|i-1} = F_i \hat{x}_{i-1|i-1} \tag{8.7}$$

$$\hat{x}_{i|i} = \hat{x}_{i|i-1} + \mathcal{K}_i(y_i - H(\hat{x}_{i|i-1})) \tag{8.8}$$

$$\mathcal{P}_{i|i-1} = F_{i-1}\mathcal{P}_{i-1|i-1}F_{i-1}^{\mathsf{T}} + Q_i \tag{8.9}$$

$$\mathcal{K}_i = \mathcal{P}_{i|i-1}H_i^{\mathsf{T}}(R_i + H_i\mathcal{P}_{i|i-1}H_i^{\mathsf{T}})^{-1} \tag{8.10}$$

$$\mathcal{P}_{i|i} = \mathcal{P}_{i|i-1} - \mathcal{P}_{i|i-1}\begin{bmatrix} H_i^{\mathsf{T}} & I \end{bmatrix} R_{e,i}^{-1} \begin{bmatrix} H_i \\ I \end{bmatrix} \mathcal{P}_{i|i-1} \tag{8.11}$$

where $\hat{x}_{i|i-1}$ and $\hat{x}_{i|i}$ denote the *a posteriori* and *a priori* estimaties of x_i, respectively, while $R_{e,i}$ is given by

$$R_{e,i} = \begin{bmatrix} R_i & 0 \\ 0 & -\lambda^2 I \end{bmatrix} + \begin{bmatrix} H_i \\ I \end{bmatrix} \mathcal{P}_{i|i-1} \begin{bmatrix} H_i^{\mathsf{T}} & I \end{bmatrix}.$$

In addition, applying the matrix inversion to (8.11), it can be shown that

$$\mathcal{P}_{i|i}^{-1} = \mathcal{P}_{i|i-1}^{-1} + H_i^{\mathsf{T}} R_i^{-1} H_i - \lambda_i^{-2} I. \tag{8.12}$$

8.3 Diffusion H_∞ filtering

Hereafter, the neighborhood of agent k is defined as the set of agents connected to agent k, including agent k itself. The neighborhood of agent k can be represented by the index set $\mathcal{N}_k = \{\ell_1, \ldots, \ell_{|\mathcal{N}_k|}\}$. The number of neighbors of an agent is referred to as the degree of agent k and is denoted by $|\mathcal{N}_k|$ Cattivelli and Sayed (2009).

8.3.1 Local sequential H_∞ filtering

The observation of agent k at time instant i is modeled as

$$y_{k,i} = H_{k,i}x_i + v_{k,i} \tag{8.13}$$

where $y_{k,i}$, $H_{k,i}$, and $v_{k,i}$ denote the observation, observation function, and observation noise of at node k at time instant i. For agent k, we define

$$H_{k,i}^{col} = \mathrm{col}_{l\in\mathcal{N}_k}\{H_{l,i}\} \tag{8.14}$$

$$R_{k,i}^{di} = \mathrm{diag}_{l\in\mathcal{N}_k}\{R_{l,i}\} \tag{8.15}$$

where "$\text{col}_{l\in\mathcal{N}_k}\{H_{l,i}\}$" represents a block column matrix of the set $\{H_{l,i}, l \in \mathcal{N}_k\}$ and "$\text{diag}_{l\in\mathcal{N}_k}\{R_{l,i}\}$" is a block diagonal matrix of the set $\{R_{l,i}, l \in \mathcal{N}_k\}$.

A local state estimator gain and estimation error covariance matrix at each node l based on its own observation can be computed, so that at each time instant i and node l we have

$$\mathcal{K}_{l,i} = \mathcal{P}_{l,i|i-1} H_{l,i}^{\mathsf{T}} (R_{l,i} + H_{l,i} \mathcal{P}_{l,i|i-1} H_{l,i}^{\mathsf{T}})^{-1} \tag{8.16}$$

$$\mathcal{P}_{l,i|i} = \mathcal{P}_{l,i|i-1} - \mathcal{P}_{l,i|i-1} \begin{bmatrix} H_{l,i}^{\mathsf{T}} & I \end{bmatrix} R_{e,l,i}^{-1} \begin{bmatrix} H_{l,i} \\ I \end{bmatrix} \mathcal{P}_{l,i|i-1} \tag{8.17}$$

$$\mathcal{P}_{l,i|i+1} = F_i \mathcal{P}_{l,i|i} F_i^{\mathsf{T}} + Q_i \tag{8.18}$$

$$\tag{8.19}$$

where

$$R_{e,l,i} = \begin{bmatrix} R_{l,i} & 0 \\ 0 & -\lambda_l^2 I \end{bmatrix} + \begin{bmatrix} H_{l,i} \\ I \end{bmatrix} \mathcal{P}_{l,i|i-1} \begin{bmatrix} H_{l,i}^{\mathsf{T}} & I \end{bmatrix}$$

By applying the matrix inversion lemma to (8.17), one can obtain

$$(\mathcal{P}_{l,i|i})^{-1} = (\mathcal{P}_{l,i|i-1})^{-1} + H_{l,i}^{\mathsf{T}} R_{l,i}^{-1} H_{l,i} - \lambda_l^{-2} I$$

where it has been assumed that node k has access to the measurement matrices $\{H_{l,i}, l \in \mathcal{N}_k\}$ and $\lambda^{-2} = \sum_{l\in\mathcal{N}_k} \lambda_l^{-2}$. From (8.18) and considering that

$$\mathcal{P}_{l-1,i|i-1} = F_i \mathcal{P}_{l_{|\mathcal{N}_k|},i-1|i-1} F_i^T + Q_i$$

successive substitution for $|\mathcal{N}_k|$ state updates yields,

$$(\mathcal{P}_{l_{|\mathcal{N}_k|},i|i})^{-1} = (\mathcal{P}_{l-1,i|i-1})^{-1} + H_{k,i}^{col\,\mathsf{T}} {R_{k,i}^{di}}^{-1} H_{k,i}^{col} - \lambda^{-2} I \tag{8.20}$$

As far as the local estimator is concerned, the sequential incorporation of the new information from neighboring nodes, $y_{l,i} \in \mathcal{N}_l$ is implemented by sequentially applying the following operations constitute a local filtering operation:

For agent k, we perform $|\mathcal{N}_k|$ state updates on a local copy of the global state vector for $l \in \{l_1, \cdots l_{|\mathcal{N}_k|}\}$, where l_1 corresponds to the node itself and $l_{|\mathcal{N}_k|}$ corresponds to the last neighboring node:

$$\hat{x}_{l,i|i} = \hat{x}_{l-1,i|i} + \mathcal{K}_{l,i}[y_{l,i} - H_{l,i}\hat{x}_{l-1,i|i}], \tag{8.21}$$

By using the matrix inversion lemma, an alternative scheme for (8.20) can

be derived as follows:

$$(\mathcal{P}_{k,i|i})^{-1} = (\mathcal{P}_{k,i|i-1})^{-1} + \sum_{l \in \mathcal{N}_k} (H_{l,i}^{\mathsf{T}} R_{,l,i}^{-1} H_{l,i}) - \lambda^{-2} I. \tag{8.22}$$

Additionally,

$$\begin{aligned}
\mathcal{K}_{k,i} =& \mathcal{P}_{k,i|i-1} H_{k,i}^{col\,\mathsf{T}} (R_{k,i}^{di} + H_{k,i}^{col} \mathcal{P}_{k,i|i-1} H_{k,i}^{col\,\mathsf{T}})^{-1} \\
=& \mathcal{P}_{k,i|i-1} H_{k,i}^{col\,\mathsf{T}} \\
& [R_{k,i}^{di\,-1} + R_{k,i}^{di\,-1} H_{k,i}^{col} (\mathcal{P}_{k,i|i-1} + H_{k,i}^{col\,\mathsf{T}} R_{k,i}^{di\,-1} H_{k,i}^{col})^{-1} H_{k,i}^{col\,\mathsf{T}} R_{k,i}^{di\,-1}] \qquad (8.23) \\
=& \mathcal{P}_{k,i|i-1} [I - H_{k,i}^{col\,\mathsf{T}} R_{k,i}^{di\,-1} H_{k,i}^{col} (\mathcal{P}_{k,i|i}^{-1} + \lambda^{-2} I)^{-1}] H_{k,i}^{col\,\mathsf{T}} R_{k,i}^{di\,-1} \\
=& (\mathcal{P}_{k,i|i}^{-1} + \lambda^{-2} I)^{-1} H_{k,i}^{col\,\mathsf{T}} R_{k,i}^{di\,-1}. \qquad (8.24)
\end{aligned}$$

Based on these two formulations for decentralized H_∞ filtering and using a diffusion step, we propose the distributed H_∞ filtering in the following subsection.

8.3.2 Distributed H_∞ filtering

The purpose of distributed filtering is to enable each agent to estimate the state vector x_i, via collaboration with its neighborhood. The goal here is perform a state estimation approach of comparable accuracy with those obtained by centralized approaches. Using H_∞ filtering derivation in (8.13)-(8.23), results in the filtering technique detailed in Algorithm 8.1. Note that different strategies exist for selecting to select weight matrices $c_{l,k}$ for the diffusion step Talebi and Werner (2019), Cattivelli and Sayed (2010), Xiao et al. (2005); however, we simply calculate the averaged estimate of neighbors.

An alternative form of diffusion H_∞ filter is derived in Algorithm 8.2 that can be used to simplify the interpretation and implementation of distributed H_∞ filtering. Note that we assume $\mathcal{P}_i > 0$ for all i. Algorithm 8.2 is the information formulation of the proposed filter, which is derived using (8.22).

8.3.3 Convergence Analysis

Assumption 8.1 (Sayed (2014b), Kailath et al. (2000))**.** *: Matrices in (8.1)-(8.2), F, H, R and Q are time-invariant i.e., $F_i \to F, H_i \to H, R_i \to R$ and $Q_i \to Q$. It is further assumed that the solution of the filtering recursions, excluding the diffusion step, is convergent for each agent Sayed (2014b), Cattivelli and Sayed (2009; 2010).*

Algorithm 8.1: Distributed H_∞ filter using diffusion step

For the system, (8.1) do for all k:
Initialization:

$$\hat{x}_{k,0|-1} = \mathbb{E}\{x\} = x_0$$
$$\mathcal{P}_{k,0|-1} = \mathcal{P}_0$$

1. Update through measurement equation:

$$\psi_{k,i} \leftarrow \hat{x}_{k,i|i-1}$$

$$\mathcal{P}_{k,i} \leftarrow \mathcal{P}_{k,i|i-1}$$

   for $\{l \in \mathcal{N}_k\}$,

   (a) Update $\mathcal{K}_{l,i}$ through (8.16).

   (b) Update $\mathcal{P}_{l,i}$ through (8.17).

   (c)

$$\psi_{k,i} \leftarrow \psi_{k,i} + \mathcal{K}_{l,i}(y_{l,i} - H_{l,i}\psi_{k,i}).$$

2. Diffusion step: for $\sum_{l\in\mathcal{N}_k} c_{l,k} = 1$,

$$\hat{x}_{k,i|i} \leftarrow \sum_{l\in\mathcal{N}_k} c_{l,k}\psi_{l,i}$$

$$\mathcal{P}_{k,i|i} \leftarrow \mathcal{P}_{k,i}$$

$$\hat{x}_{k,i+1|i} \leftarrow F_i\hat{x}_{k,i|i}$$

$$\mathcal{P}_{k,i+1|i} \leftarrow F_i\mathcal{P}_{k,i|i}F_i^\mathsf{T} + Q_{k,i}.$$

Algorithm 8.2: Information formulation of the distributed H_∞ filter using diffusion step

Information step:

$$\mathcal{I}_{k,i} = \sum_{l\in\mathcal{N}_k} H_{l,i}^{\mathsf{T}} R_{l,i}^{-1} H_{l,i} \tag{8.25}$$

$$\mathcal{T}_{k,i} = \sum_{l\in\mathcal{N}_k} H_{l,i}^{\mathsf{T}} R_{l,i}^{-1} y_{l,i} \tag{8.26}$$

$$\mathcal{P}_{k,i|i}^{-1} = \mathcal{P}_{k,i|i-1}^{-1} - \lambda_k^{-2} I + \mathcal{I}_{k,i} \tag{8.27}$$

$$\psi_{k,i} = \hat{x}_{k,i|i-1} + (\mathcal{P}_{k,i|i}^{-1} + \lambda_k^{-2} I)^{-1} [\mathcal{T}_{k,i} - \mathcal{I}_{k,i}\hat{x}_{k,i|i-1}] \tag{8.28}$$

Diffusion step: for $\sum_{l\in\mathcal{N}_k} c_{l,k} = 1$,

$$\hat{x}_{k,i|i} = \sum_{l\in\mathcal{N}_k} c_{l,k}\psi_{l,i} \tag{8.29}$$

$$\hat{x}_{k,i+1|i} = F_i \hat{x}_{k,i|i} \tag{8.30}$$

$$\mathcal{P}_{k,i+1|i} = F_i \mathcal{P}_{k,i|i} F_i^{\mathsf{T}} + Q_{k,i}. \tag{8.31}$$

Recall the definition of H_k^{col} and R_k^{di} from (8.14) and (8.15), now we are able to rearrange the measurement steps in Algorithm 8.1 as follows:

$$\mathcal{P}_{k,i|i} = \mathcal{P}_{k,i|i-1} - \mathcal{P}_{k,i|i-1}\begin{bmatrix}\left(H_k^{\text{col}}\right)^{\mathsf{T}} & I\end{bmatrix} R_{e,k}^{-1}\begin{bmatrix}H_k^{\text{col}} \\ I\end{bmatrix}\mathcal{P}_{k,i|i-1} \tag{8.32}$$

where

$$R_{e,k} = \begin{bmatrix}R_k^{\text{di}} & 0 \\ 0 & -\lambda^2 I\end{bmatrix} + \begin{bmatrix}H_k^{\text{col}} \\ I\end{bmatrix}\mathcal{P}_k\begin{bmatrix}\left(H_k^{\text{col}}\right)^{\mathsf{T}} & I\end{bmatrix}$$

We derive the Riccati recursive equation by substituting (8.32) in (8.31) as follows:

$$\mathcal{P}_{k,i+1|i} = F\mathcal{P}_{k,i|i-1}F^{\mathsf{T}} + Q - \mathcal{K}_{p,k,i}R_{e,k,i}\mathcal{K}_{p,k,i}^{\mathsf{T}}$$

where $\mathcal{K}_{p,k,i} = FP_{k,i|i-1}\begin{bmatrix}\left(H_k^{\text{col}}\right)^{\mathsf{T}} & I\end{bmatrix} R_{e,k,i}^{-1}$. Now, denoting by $\mathcal{P}_k^s$ the steady state solution of the algebraic Riccati recursion obtained above, we have,

$$\mathcal{P}_k^s = F\mathcal{P}_k^s F^{\mathsf{T}} + Q - \mathcal{K}_{p,k}R_{e,k}\mathcal{K}_{p,k}^{\mathsf{T}}$$

Assumption 8.2 (Kailath et al. (2000), Karvonen et al. (2014), Cattivelli and Sayed (2010))**.** *: The pairs $\{F, H_k^{\text{col}}\}$ and $\{F, Q^{1/2}\}$ are detectable and stabilizable for every k, respectively.*

Assumption 8.3. *for each $i \in \mathcal{N}_k$, choose λ_i^2 as Labarre et al. (2007):*

$$\lambda_i^2 = \eta \max\{eig(\mathcal{P}_{i|i-1}^{-1} + H_i^{\mathsf{T}}R_i^{-1}H_i)^{-1}\} \tag{8.33}$$

where η is a scalar larger than one.

Assumption 8.2 guarantees the existence and the convergence of $\mathcal{P}_k^s$. Assumption 8.2 and Assumption 8.3 guarantee $R_{e,k,i}^{-1} > 0$.

Lemma 8.1. *The norm-2 expectation of estimates using distributed-state estimators proposed in Algorithm 8.1 is bounded by the largest norm-2 expectation of estimates in the network, i.e*

$$\mathbb{E}\{||\hat{x}_{k,i|i}||^2\} \leq \mathbb{E}\{||\psi_{max,i}||^2\}$$

where

$$\mathbb{E}\{||\psi_{max,i}||^2\} \stackrel{\text{def}}{=\!=} max(\mathbb{E}\{||\psi_{1,i}||^2\}, \cdots, \mathbb{E}\{||\psi_{N,i}||^2\}),$$

Proof. As a reminder, we have the following diffusion step:

$$\hat{x}_{k,i|i} = \sum_{l \in N_k} c_{l,k} \psi_{l,i}$$

Then,

$$\mathbb{E}\{||\psi_{k,i}||^2\} = \sum_{l \in N_k} c_{l,k}^2 \mathbb{E}\{\psi_{l,i}\} + \sum_{m \in N_k} \sum_{l \neq m \in N_k} c_{l,k} c_{m,k} \mathbb{E}\{\psi_{m,i}^{\mathsf{T}} \psi_{l,i}\}$$

Using the definition of, $\mathbb{E}\{||\psi_{\max,i}||^2\}$, and the fact that $\sum_{l \in N_k} c_{l,k} = 1$, concludes the proof. □

Generally, in the diffusion approach, every agent should have access to the system state-space model and sufficient measurements to ensure the observability. This clearly limits the scalability of the method. To address this problem, the diffusion filtering algorithm is modified to loosen observabiity condition to that of the centralized approach.

We start by determining the observable space for each neighborhood. Then, for the measurements y_k at agent k, there exist the corresponding observable state space $\mathcal{O}_{ok}$ and the unobservable state space $\mathcal{O}_{uk}$. The globally observable subspace is the union of locally observable subspaces, while the global unobservable subspace is the intersection of the locally unobservable subspaces Ghosal et al. (2017), Kotta (2005). To relax Assumption 8.2, we change the algorithm as follows: For every time instant i, perform the measurement update only for $\mathcal{O}_{ok}$. In this case, the local H_∞ filter will be stable on the observable space of $\mathcal{O}_{ok}$. Furthermore, it follows that the estimator will be stable. Note that for the unobservable states of the local agent, the measurement update is not performed, and the state estimates are kept unchanged. The existence of this state transition and the method to obtain it are given in Chapter 13 of Rugh (1996). The proposed approach is summarized briefly in Algorithm 8.3, where $F_{k,i}$ is the state transition matrix for the locally observable space, and Ω_k is the decomposing transformation. Note that superscript *ob* corresponds to observable spaces of the system, while *uob* corresponds to the unobservable subspace. Note that all nodes in the neighborhood must update their estimates within their individual neighborhoods before the diffusion step is performed.

8.4 Decentralized Unscented H_∞ filter

Here, we formulate the H_∞ filtering problem in accordance with the unscented transform (UT) methodology. Thereafter, we reformulate the obtained

Algorithm 8.3: Diffusion H_∞ filtering only on locally observable state vectors

Do for all neighborhood of node k, $\{l \in \mathcal{N}_k\}$:

1. Obtain the local observable dynamics and measurement equation for neighborhood of node k by using the following decomposition definition given in Kailath et al. (2000):

$$\begin{bmatrix} F_{k,i} & 0 \\ F_{k,i}^{uob1} & F_{k,i}^{uob2} \end{bmatrix} \stackrel{\text{def}}{=\!=} \Omega_k^{-1} F_i \Omega_k$$
$$\begin{bmatrix} H_{l,i}^{ob} & 0 \end{bmatrix} \stackrel{\text{def}}{=\!=} H_{l,i} \Omega_k \tag{8.34}$$

2. Update through measurement equation:

$$\psi_{k,i} \leftarrow \hat{x}_{k,i|i-1}$$
$$\mathcal{P}_{k,i} \leftarrow \mathcal{P}_{k,i|i-1}$$
$$\begin{bmatrix} \psi_{k,i}^{ob} \\ \psi_{k,i}^{uob} \end{bmatrix} \stackrel{\text{def}}{=\!=} \Omega_k^{-1} \psi_{k,i}$$

   (a) By substituting $H_{l,i}^{ob}$ instead of $H_{l,i}$, update $\mathcal{K}_{l,i}$ through (8.16).

   (b) By substituting $H_{l,i}^{ob}$ instead of $H_{l,i}$, update $\mathcal{P}_{k,i}$ through (8.17).

   (c)
$$\psi_{k,i}^{ob} \leftarrow \psi_{k,i}^{ob} + \mathcal{K}_{l,i}(y_{l,i} - H_{l,i}^{ob} \psi_{k,i}^{ob}).$$
$$\psi_{k,i} \leftarrow \Omega_k \begin{bmatrix} \psi_{k,i}^{ob} \\ \psi_{k,i}^{uob} \end{bmatrix}$$

3. Diffusion step: considering that agent l uses the state vector basis for agent k when it communicates its state estimate, then for $\sum_{l \in \mathcal{N}_k} c_{l,k} = 1$,

$$\hat{x}_{k,i|i} \leftarrow \sum_{l \in \mathcal{N}_k} c_{l,k} \psi_{l,i}$$
$$\mathcal{P}_{k,i|i} \leftarrow \mathcal{P}_{k,i}$$
$$\hat{x}_{k,i+1|i} \leftarrow F_i \hat{x}_{k,i|i}$$
$$\mathcal{P}_{k,i+1|i} \leftarrow F_{k,i} \mathcal{P}_{k,i|i} F_{k,i}^{\mathsf{T}} + Q_{k,i}.$$

H_∞ filter into an information filtering format, so that it can be extended to the distributed setting in a more straightforward manner.

8.4.1 Unscented H_∞ filter

To extend our method for the nonlinear case, we assume the following nonlinear discrete-time model

$$x_{i+1} = f_i(x_i) + w_i \tag{8.35}$$

$$y_i = h_i(x_i) + v_i \tag{8.36}$$

where $f_i(\cdot)$ is the state evolution function and $h_i(\cdot)$ is the measurement function. Consider the estimate $\hat{x}_{i-1|i-1}$ and the estimation covariance $\mathcal{P}_{i-1|i-1}$ at time $i-1$. Applying the unscented transform Julier et al. (2000), H_∞ filtering can be obtained as described next Li and Jia (2010).

The prediction step is the propagation of $2n$ sigma points, $\tilde{\mathbf{x}}^s_{i-1|i-1}, s = 1, \ldots, 2n$, which are the representatives of whole distribution, via system equation at time $i-1$. The generation of sigma points is as follows:

$$\begin{aligned}
\tilde{\mathbf{x}}^0_{i-1|i-1} &= \hat{x}_{i-1|i-1}, \\
\tilde{\mathbf{x}}^s_{i-1|i-1} &= \hat{x}_{i-1|i-1} + (\sqrt{n\mathcal{P}_{i-1|i-1}})_s, s \in \{1, \cdots, n\} \\
\tilde{\mathbf{x}}^s_{i-1|i-1} &= \hat{x}_{i-1|i-1} - (\sqrt{n\mathcal{P}_{i-1|i-1}})_s, s \in \{n+1, \cdots, 2n\}
\end{aligned} \tag{8.37}$$

Next, we apply the known system model to each sigma points, and then, obtain the average value of the propagated sigma points. Thus, the implementation of the unscented transform Julier et al. (2000), Li and Jia (2010) yields:

$$\tilde{\mathbf{x}}^s_{i|i-1} = f(\tilde{\mathbf{x}}^s_{i-1|i-1}) \tag{8.38}$$

$$\hat{x}_{i|i-1} = \frac{1}{2n}\sum_{s=1}^{2n} \tilde{\mathbf{x}}^s_{i|i-1} \tag{8.39}$$

$$\mathcal{P}_{i|i-1} = \frac{1}{2n}\sum_{s=1}^{2n} [\tilde{\mathbf{x}}^s_{i|i-1} - \hat{x}_{i|i-1}][\tilde{\mathbf{x}}^s_{i|i-1} - \hat{x}_{i|i-1}]^\mathsf{T} + Q_i. \tag{8.40}$$

The corresponding predicted measurement would be [1]:

$$\hat{y}_{i|i-1} = \frac{1}{2n}\sum_{s=1}^{2n} h_i(\tilde{\mathbf{x}}^s_{i|i-1})$$

[1] It is recommended to re-sample sigma points around $\hat{x}_{i|i-1}$ using $\mathcal{P}_{i|i-1}$, but this step is omitted here for brevity.

The measurement covariance, $\mathcal{P}^{yy}$, and cross-correlation covariance between sigma points in state space and sigma points in the measurement space, $\mathcal{P}^{xy}$, can be obtained applying the sigma points prediction as follows

$$\mathcal{P}^{yy}_{i|i-1} = E[y_i - \hat{y}_{i|i-1}][y_i - \hat{y}_{i|i-1}]^\mathsf{T} \tag{8.41}$$

$$= \frac{1}{2n}\sum_{s=1}^{2n}[h_i(\tilde{\mathbf{x}}^s_{i|i-1}) - \hat{y}_{i|i-1}][h_i(\tilde{\mathbf{x}}^s_{i|i-1}) - \hat{y}_{i|i-1}]^\mathsf{T} \tag{8.42}$$

$$\mathcal{P}^{xy}_{i|i-1} = E[x_i - \hat{x}_{i|i-1}][y_i - \hat{y}_{i|i-1}]^\mathsf{T} \tag{8.43}$$

$$= \frac{1}{2n}\sum_{s=1}^{2n}[\tilde{\mathbf{x}}^s_{i|i-1} - \hat{x}_{i|i-1}][h_i(\tilde{\mathbf{x}}^s_{i|i-1}) - \hat{y}_{i|i-1}]^\mathsf{T} \tag{8.44}$$

In this case, we rearrange the unscented transform framework equations by taking advantage of the statistical linear error propagation method Julier et al. (2000), Li and Jia (2010), which helps us to derive equations for distributed H_∞ filtering. Namely,

$$\mathcal{P}^{yy}_{i|i-1} \approx H_i \mathcal{P}_{i|i-1} H_i^\mathsf{T} \tag{8.45}$$

$$\mathcal{P}^{xy}_{i|i-1} \approx \mathcal{P}_{i|i-1} H_i^\mathsf{T} \tag{8.46}$$

Consequently, based on (8.11) and (8.46), we obtain the filtered estimates as:

$$\mathcal{P}_{i|i} = \mathcal{P}_{i|i-1} - \begin{bmatrix}\mathcal{P}^{xy}_{i|i-1} & \mathcal{P}_{i|i-1}\end{bmatrix} R^{-1}_{e,i} \begin{bmatrix}[\mathcal{P}^{xy}_{i|i-1}]^\mathsf{T} \\ \mathcal{P}^\mathsf{T}_{i|i-1}\end{bmatrix} \tag{8.47}$$

where

$$R_{e,i} = \begin{bmatrix} R_i + \mathcal{P}^{yy}_{i|i-1} & [\mathcal{P}^{xy}_{i|i-1}]^\mathsf{T} \\ \mathcal{P}^{xy}_{i|i-1} & -\lambda^2 I + \mathcal{P}_{i|i-1}\end{bmatrix}$$

Choosing the parameter λ is an important step in H_∞ filtering problems, since it affects the positiveness of $\mathcal{P}_{i|i}$. By using Assumption 8.3 , (8.33), (8.45) and (8.46), we can obtain λ_i as follows:

$$\lambda_i^2 = \eta \max\{\text{eig}(\mathcal{P}^{-1}_{i|i-1} + \mathcal{P}^{-1}_{i|i-1}\mathcal{P}^{xy}_{i|i-1}R_i^{-1}[\mathcal{P}^{-1}_{i|i-1}\mathcal{P}^{xy}_{i|i-1}]^\mathsf{T})^{-1}\}$$

With these developments, the unscented H_∞ filtering method is obtained using only calculated covariances, without the matrix H_i from (8.45) and (8.46). This will be shown next.

8.4.2 Information form of the Unscented H_∞ filter

In order to obtain the decentralized diffusion procedure, the unscented H_∞ information filter is derived. For this purpose, we can apply the matrix inversion lemma, and rewrite the state estimator gain $\mathcal{K}_i$ as follows

$$\begin{aligned}\mathcal{K}_i =&\mathcal{P}_{i|i-1}H_i^{\mathsf{T}}(I + R_i^{-1}H_i\mathcal{P}_{i|i-1}H_i^{\mathsf{T}})^{-1}R_i^{-1}\\ =&\mathcal{P}_{i|i-1}(I + H_i^{\mathsf{T}}R_i^{-1}H_i\mathcal{P}_{i|i-1})^{-1}H_i^{\mathsf{T}}R_i^{-1}\\ =&(\mathcal{P}_{i|i-1}^{-1} + H_i^{\mathsf{T}}R_i^{-1}H_i)^{-1}H_i^{\mathsf{T}}R_i^{-1}\end{aligned}$$

As a reminder, $\mathcal{P}_{i|i}$ can be rewritten by taking the advantage of the matrix inversion lemma,

$$\begin{aligned}\mathcal{P}_{i|i}^{-1} =&\mathcal{P}_{i|i-1}^{-1} + \begin{bmatrix}H_i^{\mathsf{T}} & I\end{bmatrix}\begin{bmatrix}R_i & 0\\ 0 & -\lambda_i^{-2}I\end{bmatrix}^{-1}\begin{bmatrix}H_i\\ I\end{bmatrix}\\ =&\mathcal{P}_{i|i-1}^{-1} - \lambda_i^{-2}I + H_i^{\mathsf{T}}R_i^{-1}H_i\end{aligned} \tag{8.48}$$

Then, we can obtain

$$\begin{aligned}\mathcal{K}_i =&(\mathcal{P}_{i|i}^{-1} + \lambda_i^{-2}I)^{-1}H_i^{\mathsf{T}}R_i^{-1}\\ =&(\mathcal{P}_{i|i}^{-1} + \lambda_i^{-2}I)^{-1}\mathcal{P}_{i|i-1}^{-1}\mathcal{P}_{i|i-1}^{xy}R_i^{-1}\end{aligned}$$

Thus, the filtered estimates can be obtained as

$$\begin{aligned}\hat{x}_{i|i} =&\hat{x}_{i|i-1} + \mathcal{K}_i(y_i - h_i(\hat{x}_{i|i-1}))\\ =&\hat{x}_{i|i-1} + (\mathcal{P}_{i|i} + \lambda_i^{-2}I)^{-1}\mathcal{P}_{i|i-1}^{-1}\mathcal{P}_{i|i-1}^{xy}R_i^{-1}(y_i - h_i(\hat{x}_{i|i-1}))\end{aligned}$$

8.4.3 Distributed nonlinear filtering

Based on the obtained information matrix (8.48) and substituting the (8.45)-(8.46) into (8.48), the diffusion H_∞ filtering can be obtained as in Algorithm 8.4.

8.5 Simulation results

We give two examples to verify our investigations. The first example is a well known target tracking problem Talebi and Werner (2019), Talebi and Werner (2018), Li and Jia (2010). The dynamics of target tracking problem in xy-plane at time instant i is given by

$$\mathbf{x}_i = \begin{bmatrix}1 & 0 & \frac{1}{f_s} & 0\\ 0 & 1 & 0 & \frac{1}{f_s}\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\mathbf{x}_{i-1} + \underbrace{\begin{bmatrix}\frac{1}{2}(\frac{1}{f_s})^2 & 0\\ 0 & \frac{1}{2}(\frac{1}{f_s})^2\\ \frac{1}{f_s} & 0\\ 0 & \frac{1}{f_s}\end{bmatrix}}_{\mathbf{w}_i}\mathbf{a}_i + \delta_{\mathrm{x}}$$

Algorithm 8.4: Distributed nonlinear H_∞ information filter using diffusion step

For the system model (8.35), starting with $\hat{x}_{k,0|-1} = 0$ and $\mathcal{P}_{k,0|-1} = \Pi_0$, we have
Information update:

$$\mathcal{I}_{k,i} = \sum_{l \in \mathcal{N}_k} \mathcal{P}^{-1}_{l,i|i-1} \mathcal{P}^{xy}_{l,i|i-1} R^{-1}_{l,i} \mathcal{P}^{xy}_{l,i|i-1}{}^{\mathsf{T}} \mathcal{P}^{-1}_{l,i|i-1} \tag{8.49}$$

$$\mathcal{T}_{k,i} = \sum_{l \in N_k} \mathcal{P}^{-1}_{l,i|i-1} \mathcal{P}^{xy}_{l,i|i-1} R^{-1}_{l,i} (y_{l,i} - h_{l,i}(\hat{x}_{k,i|i-1})) \tag{8.50}$$

$$\mathcal{P}^{-1}_{k,i|i} = \mathcal{P}^{-1}_{k,i|i-1} + \mathcal{I}_{k,i} - \lambda^{-2} I \tag{8.51}$$

$$\psi_{k,i} = \hat{x}_{k,i|i-1} + (\mathcal{P}^{-1}_{k,i|i} + \lambda^{-2} I)^{-1} \mathcal{T}_{k,i}. \tag{8.52}$$

Diffusion update:

1. $$\hat{x}_{k,i|i} = \sum_{l \in N_k} c_{l,k} \psi_{l,i} \tag{8.53}$$

2. Update $\hat{x}_{k,i+1|i}$ through (8.37)-(8.39).

3. Update $\mathcal{P}_{k,i+1|i}$ through (8.40).

$$\text{if } \; l \in \mathcal{N}_k : \; \mathbf{y}_{l,i} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \mathbf{x}_i + \mathbf{v}_{l,i}$$

where $\mathbf{x}_i = [p_{x_i}, p_{y_i}, \dot{p}_{x_i}, \dot{p}_{y_i}]$ for positions states p_{x_i}, p_{y_i}. The simulation results are performed using sampling rate $f_s = 25$Hz. $\mathbf{w}_i$ is the process noise assumed to be zero-mean Gaussian, and $\mathbf{a}_i$ can be interpreted as accelerations in two orthogonal horizontal directions Talebi and Werner (2018). The measurement noise $v_{l,i}$ is also assumed to be zero-mean Gaussian. The topology of the network used for simulations is shown in Fig. 8.1. Furthermore, Fig. 8.2 shows the mean square estimation error of first and the second state of the system, p_{x_i}, p_{y_i} using a diffusion based Kalman and H_∞ filtering. A Monte Carlo simulation with 100 runs has been performed which has similar statistic's characteristics for both distributed Kalman and H_∞ filtering methods. The process noise covariance is assumed to be $Q_s = 0.0023\,\mathrm{diag}\,\{0,\ 0,\ 1,\ 1\}$, and δ_{x} is additional unknown process noise

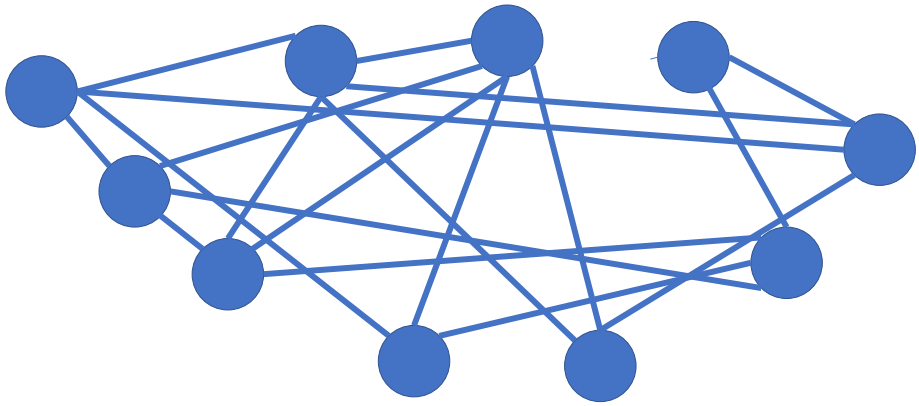


Figure 8.1: A network with 10 agents and 20 edges

which we assume is modeled as uniformly distributed noise within ± 0.002 and the measurement covariance is $R_s = 0.16I$. Even though, the diffusion H_∞ filtering requires higher computational cost because of the extra computation of the parameter η, it achieves a lower mean square error, as shown in Fig. 8.2. The superior performance observed for H_∞ filtering clearly comes from the Kalman filter being unable to represent accurately the uniformly distributed noise source δ_{x}. This illustrates the proposed H_∞ filter is more robust to errors and uncertainties in process noise and consequently in process model. It should be mentioned that the mean square periodically goes to zero for both filters. This is because the disturbance/acceleration is zero for these times in all runs. Note that the diffusion step for this example serves only as a means for propagating the global state estimate throughout the system, as the local measurements for agent l are the only ones that contain information about the states of agent l.

8.5.1 IEEE-14-bus network

The second example is the distributed dynamic state estimation for IEEE-14-bus network, Fig. 8.3. The derivation of the model for this network is as in Ghosal and Rao (2019), Abooshahab et al. (2019). The classical synchronous generator model is given in Kundur et al. (1994), Machowski

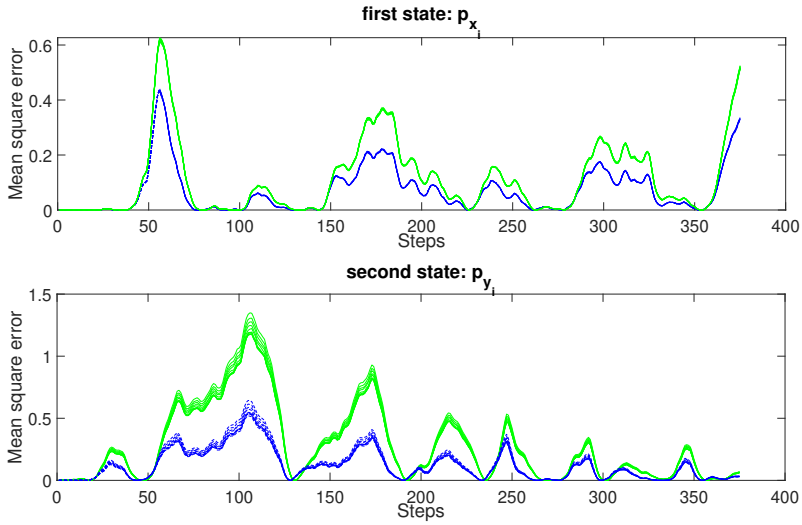


Figure 8.2: Mean square estimation error of distributed H_∞ filter (blue dashed lines) and Kalman filter (green solid line) using diffusion approach for 100 runs of Monte Carlo simulation with different error covariances.

et al. (2020) as follows:

$$\dot{\delta} = \omega$$
$$\dot{\omega} = \frac{1}{M}(P_m - P_e - D\omega) \tag{8.54}$$

where P_e is the air-gap electrical power. The other variables and parameters are defined as follows:

state variables: ω denotes the difference between rotor rotational speed and the synchronous speed, and δ refers to the rotor angle;
parameters for rotor: D and M represent the damping and inertia coefficients, respectively.
external variables: P_m is the mechanical turbine power.

For this example, by measuring the power angles of buses one, two, three, six, and eight, we get the result given in Fig. 8.4. In this example, the covariances of the process noise assumed to be known, but the measurement error $\mathrm{v_i}$ with covariance matrix $0.0004I$ assumed to be added by an unknown additive measurement noise δ_y, uniformly distributed within ± 0.0002. We plot power angles at buses three and five, and the rotational speeds at buses three and four in Fig. 8.4 which shows that the H_∞ filter also better deals with the uncertainties in the measurement. It is also shows that the

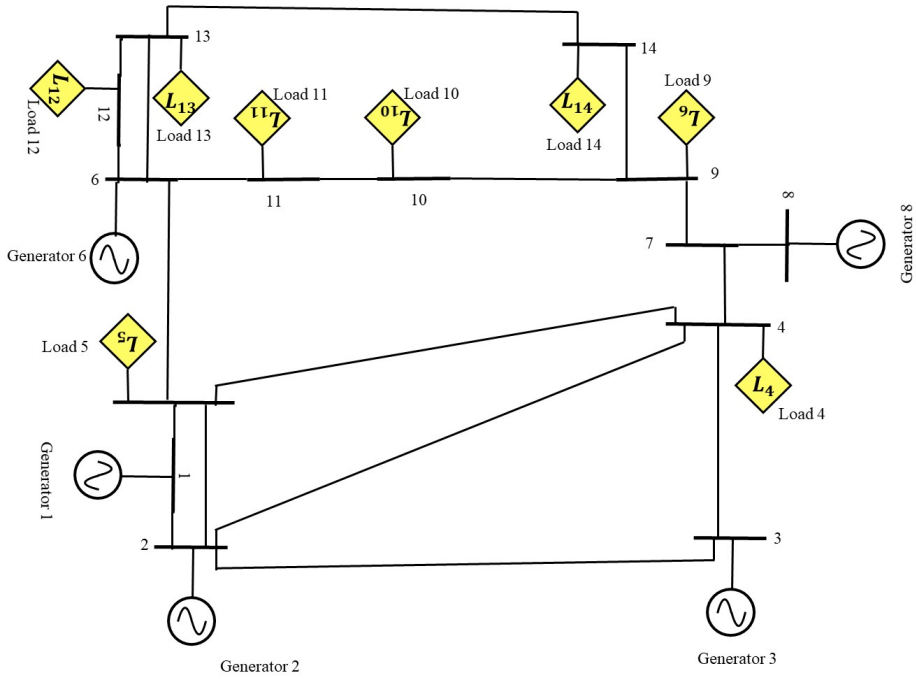


Figure 8.3: The IEEE-14-bus network Li et al. (2010).

estimates of the H_∞ filter fluctuate less than those from the Kalman filtering one. However, in this example, only the measurement error assumed to be unknown and changed in different Monte Carlo runs, that The results for the other states are similar and are therefore not shown.

8.6 Conclusion

In this chapter, we propose a fully distributed linear H_∞ filtering by using the diffusion strategy. Then we relax the assumption of global observability for each agent by proposing a diffusion based algorithm which uses only the locally (in the neighborhood) observable state vector. Then we extend our framework to nonlinear systems using the unscented transform. Finally, we verify the effectiveness of proposed method using two examples, a target tracking problem and the distributed state estimation for IEEE-14-bus system

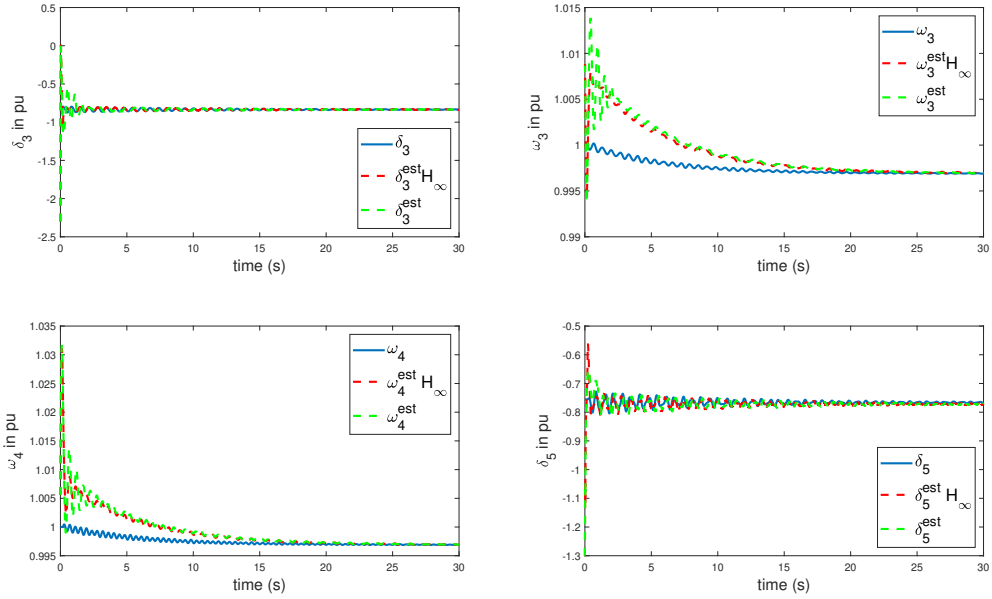


Figure 8.4: States of a power system with their estimation using distributed H_∞ (dashed red lines) and Kalman filter (dashed green line) for 100 runs of Monte Carlo simulation.

Chapter 9

Paper VI: Multi-Rate Distributed Unscented Kalman Filtering with Application to Power System Monitoring

Multi-Rate Distributed Unscented Kalman Filtering with Application to Power System Monitoring

Mohammad Ali Abooshahab
Morten Hovd
In preparation for submission to *IEEE Transactions on Smart Grids*, (2021)

Chapter 10

Paper VII: A Covariance Consistent Data Fusion method for Power Networks with Multirate Sensors

This chapter is taken from Abooshahab et al. (2020a).

A Covariance Consistent Data Fusion method for Power Networks with Multirate Sensors [1]

Mohammad Ali Abooshahab
Morten Hovd
Edmund Brekke
Xianfeng Song
2020 IEEE Conference on Control Technology and Applications (CCTA), **Montréal, Canada** (2020)

Abstract

The quasi-static assumption has been often employed in the analysis of a power system for state monitoring/estimation. Accordingly, only static state estimates can be obtained. However, the increased penetration of

[1] This chapter (paper) use the same algorithms 9.5 and 9.6 as in chapter 9, and to reduce repetition these algorithms have been removed from the present chapter.

renewable generation, especially photovoltaic generators and wind farms, introduces significant variability in the system and challenges this quasi-static assumption. Thus, it is crucial to extract information about the system dynamic states, which influences the stability of the system. Therefore, dynamic schemes for state estimation, particularly Kalman filtering, have been introduced for the power systems to perform dynamic state estimation. However, power systems usually have a low degree of instrumentation, which renders it necessary to exploit all the available information and measurements in a power network. There are different sources of measurements in distribution grids, such as SCADA and Phasor Measurement Units (PMUs). These sensors, however, provide different rates of data. Hence, multi-rate data fusion is required in a power system containing different types of sensors. Considering the demonstrated consistency with the covariance intersection method (CI), we propose an unscented Kalman filter (UKF)-based CI data fusion approach to fuse the estimates based on sensors with different data rates. This method is then compared to an existing multi-rate data fusion algorithm for power systems. The results show that the proposed dynamic approach is effective and provides robust state estimates for the power systems.

10.1 Introduction

In electric power systems, systematic, consistent, and precise state estimation is required for most energy management system (EMS) operations Ghosal et al. (2017). Any severe fault or sudden change can lead to major problems for the power system, such as instability and blackout. In addition, in control procedures for power systems, including linear quadratic regulator (LQR) Yang and Cimen (1996), estimation based model predictive control (MPC) Roshany-Yamchi et al. (2011), sliding mode He et al. (2019) and backstepping Cupelli et al. (2015), the controller needs to monitor the variables (known as states) of the system. In general, these states are estimated through processing noisy measured data from different sensors. The states estimated based on these noisy measurements are then forwarded to the control center, where the operator will respond accordingly. Historically, it has been common for the state estimation methods applied to power systems to assume the system to be static/quasi-static, and for the estimator to update its state estimates every few minutes. In contrast, electromagnetic faults, dynamic loads, and distributed generation can introduce much faster dynamics into the power system. Thus, we must engage dynamic estimation instead of static estimation to detect and identify sudden changes in a power system.

Although the Kalman filter was discovered around six decades ago Kalman (1960), the industrial application in power systems has been limited. Important reasons for this are the relatively slow measurements available in SCADA systems and the large physical distances in power systems - making it hard to synchronize measurements accurately relative to the operating frequency of AC systems Gomez-Exposito and Abur (2004). The first PMU was developed in the 1980s Phadke and Thorp (2006), and they have both provided fast measurements and taken advantage of a global clock from GPS signals. PMUs are becoming the most important sensors for modern power grids Ghosal and Rao (2019). Thus, there is both an opportunity and a need to perform dynamic state estimation in power grids. On the other hand, many power systems still have limited instrumentation Ghosal and Rao (2019), and it is therefore necessary to combine the information from slow/infrequent measurements from SCADA systems and smart meters with information from faster/more frequent measurements such as PMUs Ghosal and Rao (2019).

In the last decade, the power systems community has spent significant effort in studying filtering methods for electromechanical Abooshahab et al. (2019), Ghosal and Rao (2019), electromagnetic Anagnostou and Pal (2018), and quasi-static Xie et al. (2017) states of the power systems. In Blood (2011), the authors proposed a procedure for shifting from static to dynamic electric power network state estimation using a linear Kalman filter. Various dynamic estimation methods, including particle filtering, the unscented, extended, and ensemble Kalman filtering for a power system, were investigated in Zhou et al. (2015). In Ghahremani and Kamwa (2011a), a dynamic state estimation procedure using the extended Kalman filter (EKF) with unknown inputs was proposed. Such a method has been extended to monitor partially known power networks Abooshahab et al. (2019) based on a method introduced in Bitmead et al. (2019). In Ghahremani and Kamwa (2011b), the UKF is implemented to obtain the state estimates of the power system. In Huang et al. (2009), the effects of sampling rate on the performance of an estimator for a power grid were investigated.

Multi-rate data fusion methods have been successfully implemented in robotics and signal processing applications Mitchell (2007). By using the method introduced in Yan et al. (2010), multi-rate data fusion for power networks using the Bar-Shalom Campo (BSC) fusion method is presented in Ghosal and Rao (2019), Ghosal et al. (2017). Despite the good performance of the BSC fusion method in the proposed examples in Ghosal and Rao (2019), the BSC method is subject to the challenges of instability or signific-

ant error. Data fusion methods are performed based on some assumptions, especially for the correlation between two local estimates. Such assumptions, however, can result in the estimates that are too optimistic. This feature is called the trend of inconsistency, which might cause the filter to diverge in specific circumstances, as explained in Maybeck (1982). Clearly, consistent estimation is critical for power systems state estimation since monitoring and advanced control methods are based on such state estimates. Hence, large estimation errors can be hazardous. Besides, even though power networks are sparsely equipped with PMUs, the results proposed in Ghosal and Rao (2019) are based on strong observability assumptions for both fast and slow sensors. Furthermore, the propagation of sigma points exploited for the slow sensors to obtain the predicted states without measurement update is complicated and computationally expensive.

It is proven in Deng et al. (2012) that the CI Kalman fuser is consistent; therefore, we choose the covariance intersection as our fusion method to obtain a consistent state estimation, which is essential for control schemes of the power networks. Also, as a key step to exploit all data from sensors with various sampling rates, we employ an additional operation in the state prediction step, which simply extrapolates the state estimates and their associated covariances from the slower sensor. The simulation analysis shows that the performance of estimators using CI is more reliable compared to the BSC for the studied benchmark. The comparison shows applying the extrapolation of estimates instead of propagating sigma points through the equations for the system dynamics can reduce the execution time and the complexity of the estimator without a significant influence on the covariance of the estimation error. Furthermore, we perform the fusion only for the intersection of the observable subspaces. In this case, the observability condition is relaxed to: the system is observable if each state is in the observable subspace of at least one subsystem.

The outline of this paper is as follows. We present our methodology and obtain a model for the power system to validate the performance of the proposed data fusion methods in Section 10.2. In Section 10.3, the state estimation and data fusion algorithms, including the UKF and covariance intersection, are presented. The simulation results using the standard benchmark system WSCC-9-bus network L. Chen et al. (2005) are presented in Section 10.4 to verify the robustness of the proposed methods and compare them to methods proposed in the literature.

10.2 Methodology

In this section, we introduce the full-order nonlinear model for our power network.

10.2.1 Model Derivation

Power system models are not unique, and choosing an appropriate dynamical model, therefore, depends on the emphasis of components in the power system and purpose of the monitoring. The purpose of the monitoring/estimation will determine what dynamics will need to be included in the model. One approach to characterizing power system dynamics is to separate them into three types of states based on their time-scale. These three types are electromagnetic transients (few milliseconds), electro-mechanical transients (tens of milliseconds to a few seconds), and quasi-steady-state (more than ten seconds) Ghosal and Rao (2019). We know that fast dynamics (electromagnetic transients) has virtually no impact on the slower dynamics (electro-mechanical and quasi-steady-state). In this study, we consider the electro-mechanical transient dynamics to represent the power system. Hence, the swing equations of the generators, as well as the power flow equations, are used to describe the power system states. The classical synchronous generator model is given in Kundur et al. (1994), Ghosal et al. (2017), Machowski et al. (2020) as follows:

$$\begin{aligned} \dot{\delta} &= \omega \\ \dot{\omega} &= \frac{1}{M}(P_m - P_e - D\omega), \end{aligned} \tag{10.1}$$

where P_e is the air-gap electrical power. The other variables and parameters are defined as follows:

state variables: ω denotes the difference between rotor rotational speed and the synchronous speed, and δ refers to the rotor angle;
parameters for rotor: D and M represent the damping and inertia coefficients, respectively.
external variables: P_m is the mechanical turbine power.

For power systems with multiple machines, we distinguish between two types of nodes, which are dynamic component nodes, and bus nodes. The numbers of bus nodes and component nodes are denoted by $n_{\mathcal{B}}$ and $n_{\mathcal{C}}$, respectively. The $n_{\mathcal{C}}$ components determine the dynamics of the interconnected network. For $Y \in \mathbb{C}^{n_{\mathcal{B}} \times n_{\mathcal{B}}}$ as the admittance matrix of the network, we can obtain

the following equation using circuit theory Sadamoto et al. (2019):

$$0^{n_{\mathcal{B}}} = (Y(|V|\angle\theta)_{1:n_{\mathcal{B}}})^{\cdot *} \circ (|V|\angle\theta)_{1:n_{\mathcal{B}}} - (P_{1:n_{\mathcal{B}}} + jQ_{1:n_{\mathcal{B}}}), \tag{10.2}$$

where $\circ$ and $.*$ is the element-wise multiplication and conjugation operator, respectively Sadamoto et al. (2019). The active and reactive power injected from the ith bus are denoted by P_i and Q_i Sadamoto et al. (2019). The notations, $(|V|\angle\theta)_{1:n_{\mathcal{B}}}$, $P_{1:n_{\mathcal{B}}}$, and $Q_{1:n_{\mathcal{B}}}$ are the accumulated matrices of $(|V|\angle\theta)_i, P_i$, and Q_i for $i \in \{1, \ldots, n_{\mathcal{B}}\}$ Sadamoto et al. (2019). Note that we drop all time indices for simplicity of exposition.

For buses without generators or loads, $P_i + jQ_i = 0$; Hence we can rewrite (10.2) as:

$$\begin{bmatrix} P_{1:n_{\mathcal{C}}} + jQ_{1:n_{\mathcal{C}}} \\ 0 \end{bmatrix} = \begin{bmatrix} Y_{\mathcal{C}\mathcal{C}} & Y_{\mathcal{C}\mathcal{B}} \\ Y_{\mathcal{B}\mathcal{C}} & Y_{\mathcal{B}\mathcal{B}} \end{bmatrix} (|V|\angle\theta)_{1:n_{\mathcal{B}}})^{\cdot *} \circ (|V|\angle\theta)_{1:n_{\mathcal{B}}} \tag{10.3}$$

Thus, by applying the Krone reduction method Dorfler and Bullo (2013), we obtain the bus reconstruction matrix $R_v = -Y_{\mathcal{B}\mathcal{B}}^{-1} Y_{\mathcal{B}\mathcal{C}}$ and, the reduced admittance matrix $Y_r = Y_{\mathcal{C}\mathcal{C}} - Y_{\mathcal{C}\mathcal{B}} Y_{\mathcal{B}\mathcal{B}}^{-1} Y_{\mathcal{B}\mathcal{C}}$. Our test case only contains synchronous generators as dynamic components. Thus, for an $n_{\mathcal{C}}$-generator multi-machine power network, we have the following dynamics for the ith synchronous generator Kundur et al. (1994), Ghosal and Rao (2019)

$$\dot{\delta}_i = \omega_i, \quad i = 1, \cdots, n_{\mathcal{C}}$$

$$\dot{\omega}_i = \frac{(P_{m,i} - P_{e,i}(\bar{\delta}))}{M_i} - \frac{D_i}{M_i}\omega_i,$$

and

$$P_{e,i}(\bar{\delta}) = \sum_{j=1}^{n_{\mathcal{C}}} [\bar{G}_{ij} \cos(\delta_i - \delta_j) + \bar{B}_{ij} \sin(\delta_i - \delta_j)],$$

$$\bar{G}_{ij} = |E_i||E_j|G_{ij}, \quad \bar{B}_{ij} = |E_i||E_j|B_{ij},$$
$$\bar{\delta} = [\delta_1, \ \cdots, \ \delta_{n_{\mathcal{C}}}]^{\mathsf{T}},$$

where G_{ij} is the real part of Y_r, and B_{ij} is the imaginary part of Y_r. Consequently, we are able to define the state space as follows Ghosal and Rao (2019):

$$\mathrm{x} = [(\delta_2 - \delta_1), \cdots, (\delta_{n_{\mathcal{C}}} - \delta_1), (\omega_2 - \omega_1), \cdots, (\omega_{n_{\mathcal{C}}} - \omega_1)]^{\mathsf{T}}, \tag{10.4}$$

where δ_1 is the power angle at the reference bus. In this case, the state space contains $2(n_{\mathcal{C}}-1)$ state variables. By applying this state space definition, the multi-machinery dynamical equations can be re-written as Ghosal and Rao (2019):

$$\dot{x}_i = x_{j+n_{\mathcal{C}}-1},$$

$$\dot{x}_{i+n_{\mathcal{C}}-1} = \frac{P_{m,i}}{M_i} - \frac{P_{m,1}}{M_1} - \frac{D_i}{M_i}x_{i+n_{\mathcal{C}}-1} + \frac{\bar{G}_{11}}{M_1} - \frac{1}{M_i}[\bar{B}_{i1}\sin(x_i) + \bar{G}_{i1}\cos(x_i)]$$
$$-\frac{1}{M_i}\sum_{j=1}^{n_{\mathcal{C}}-1}[\bar{G}_{ij}\cos(x_i - x_j) + \bar{B}_{ij}\sin(x_i - x_j)]$$
$$-\frac{1}{M_1}\sum_{i=1}^{n_{\mathcal{C}}-1}[\bar{B}_{i1}\sin(x_i) + \bar{G}_{i1}\cos(x_j)].$$

Then the compact nonlinear form of the multi-machine power network can be expressed as follows Ghosal and Rao (2019),

$$\dot{\mathrm{x}} = \mathrm{f}(\mathrm{x}, u) + \mathrm{w},$$

where the process noise signal w is a Gaussian zero-mean noise with covariance matrix Q_x.

10.2.2 Measurement model

It is recognized in power networks that there are two major types of sensors: the fast and high-frequency sensors like PMUs, and the slow and low frequency sensors like smart meters, SCADA sensors and also slower PMUs. A set of PMU-based measurements usually contains the information of the frequency, the voltage, and the current phasors. On the other hand, a set of SCADA measurements can assess the real power, active power and voltage magnitudes. For simplicity, we assume that our measurement model only contains PMUs but involves different sampling intervals. The sensor selection set is thus $B = \{b_1, b_2, \cdots, b_{n_f}\} \subset \{1, 2, \cdots, n_{\mathcal{B}}\}$ where n_f is the number of fast PMUs. For the PMU-based measurements, the magnitude and phase angle of the corresponding voltage can be determined by Ghosal and Rao (2019)

$$|v_j| = \left| R_{v,i1}|e_1| + \sum_{i=1}^{n_{\mathcal{C}}-1} R_{v,ji}|E_i|\angle x_i \right|$$

$$\theta_j = \delta_1 + \arg(R_{v,j1}|e_1| + \sum_{i=1}^{n_{\mathcal{C}}} R_{v,ji}|E_i|\angle x_i)$$

Note that $R_{v,ji}$, $j \in \{1, 2, \cdots, n_{\mathcal{B}}\}$, $i \in \{1, \cdots, n_{\mathcal{C}}\}$ is an element of the reconstruction matrix introduced in section 10.2.A. Hence, the fast PMU-based measurement vector is

$$\mathrm{y}_f = [|v_{b1}|, |v_{b2}|, \cdots, |v_{b_{n_f}}|, |\bar{\theta}_{b2}|, \cdots, |\bar{\theta}_{b_{n_f}}|]^{\mathsf{T}},$$

where $\bar{\theta}_{b_i} = \theta_{b_i} - \theta_{b_1}$, and θ_{b_1} is the reference angle. The fast PMU measurement equation can be concisely rewritten as,

$$\mathrm{y}_f = \mathrm{h}_f(x) + \nu_f.$$

In this stage, the noise signal ν_f in the PMU-based measurement is assumed to follow a Gaussian zero-mean distribution, and the covariance matrix of this noise signal is R_f. Similarly, the slow PMU measurement equation vector would be as follows:

$$\mathrm{y}_s = \mathrm{h}_s(x) + \nu_s.$$

10.3 State estimation

The importance of state estimation for the monitoring of power systems has been explained in Section 10.1. However, power systems are highly nonlinear. Consequently, nonlinear approaches should be used in the case of dynamical state estimation Ghosal et al. (2017).

In standard Kalman filtering (as described in, e.g., Simon (2006)), all measurements are sampled synchronously and with the same sampling rate, and new measurements are immediately available to the Kalman filter Ghosal and Rao (2019). Accordingly, the prior state estimates in the prediction step should correspond to the concurrent measurements when updating the estimates. In this work, we assume that we have two different types of sensors with sampling rates r_f (fast PMU) and r_s (slow PMU), although the sensor fusion methodology is straightforward to apply when there are different sensor types in the system. Then we perform distributed state estimation on each set of measurements and subsequently introduce data fusion with CI.

10.3.1 Multi time-scale dynamical and measurement models

For the fast PMU with sampling rate of r_f, we can discretize the dynamical system at time-step $T_f = 1/r_f$ as

$$x_{t_f+1} = f_f(x_{t_f}) + w_{t_f}, \tag{10.5}$$

where w_{t_f} is a Gaussian white noise with variance $\mathbf{E}\{w_{t_f} w_{t_f}^{\mathsf{T}}\} = Q_f$ Ghosal and Rao (2019), Huang et al. (2007). The discrete measurement equation for the fast sensor can be expressed as,

$$y_{t_f} = h_f(x(t_f), u_{t_f}) + \nu_{t_f}. \tag{10.6}$$

Using the same procedure, we can obtain discrete dynamical system of slow PMU measurements at time-steps $T_s = 1/r_s$ as

$$x_{t_s+1} = f_s(x_{t_s}, u_{t_s}) + w_{t_s}, \tag{10.7}$$

where $Q_s = \mathbf{E}\{w_{t_s} w_{t_s}^{\mathsf{T}}\}$. The discrete measurement equation for the slow sensor can be expressed as,

$$y_{t_s} = h_s(x(t_s)) + \nu_{t_s}. \tag{10.8}$$

We aim to fuse the data from both types of sensors with different sampling rates for power systems state estimation.

10.3.2 Unscented Kalman filter

In this paper, we use the UKF because it is derivative-free and inherently has no need to calculate partial derivatives and the Jacobian matrix. Such derivative-free state estimation procedures can reduce the heavy computational load. Furthermore, instead of performing Taylor expansion, the UKF can give a better performance for the estimation of random variables in nonlinear functions, especially when the higher-order terms in Taylor's expansion cannot be ignored Julier and Uhlmann (1996), Ghosal and Rao (2019).

We choose arbitrary values for the initial conditions of the state estimates and error covariance matrix (the initial covariance matrix has to be positive definite). In the case of the initial condition selected, we can perform UKF state estimation by exploiting the algorithm depicted in Figure 10.1. The state estimation procedure using UKF shown in Figure 10.1 is summarized in Algorithm 9.5-9.6.

10.3.3 Multi-sensor data fusion

It is noted the sensing elements collect data at different sampling rates. For example, SCADA devices provide data with sampling rates around 0.5-2 samples per second, while PMUs have much higher sampling rates (30-120

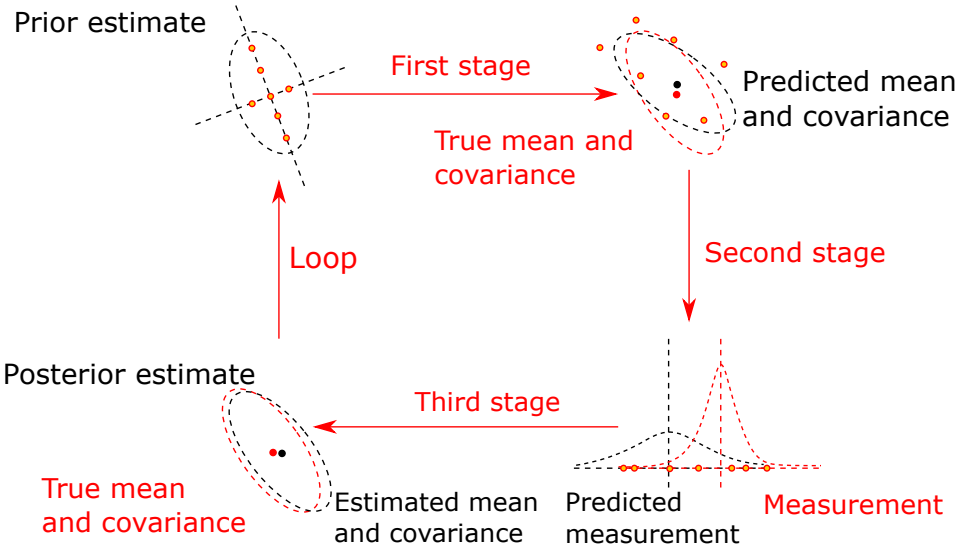


Figure 10.1: Summarized depiction of UKF in three stages

samples per second). In this subsection, we solve the multi-rate multi-sensor data fusion to exploit all possible and appropriate data.

Lemma 1 Ghosal et al. (2017) For a given observable system, the *a posteriori* estimation error covariance is $\mathcal{P}^{i}_{t|t}$. Provided an additional measurement is available, the corresponding *a posteriori* estimation error covariance is denoted $\mathcal{P}^{i+1}_{t|t}$. In this case,

$$trace\{\mathcal{P}^{i}_{t|t}\} \geq trace\{\mathcal{P}^{i+1}_{t|t}\}.$$

Thus, additional measurements with finite error covariance do not degrade the performance of state estimators like Kalman filter.

In Ghosal and Rao (2019), it is assumed that the system is observable when each of the sensor sets used. This assumption is very restrictive. The proposed method in Algorithm 1 is characterized by more flexibility. Consider the state estimation with slow sensor measurements at fast sensor sampling rates; each set of estimated values is supposed to synchronize at the same time instant for data fusion. Therefore, it is suggested to develop an approach with slow measurements and predict the states between the sampling points for the slow sensor, i.e., Ghosal and Rao (2019), Yan et al. (2010):

$$\hat{x}_{sf}(t_f + 1|t_f + 1) = \mathbb{E}\left[x_{t_f+1}|y_1, y_2, \cdots, y_{t_s+1}\right]$$

where $\hat{x}_{sf}$ is the state estimate of the system using the slow measurements with fast measurements sampling rate. We address this problem by applying two methods. The obtained results in Section 10.4 suggest that the two methods bring very similar results.

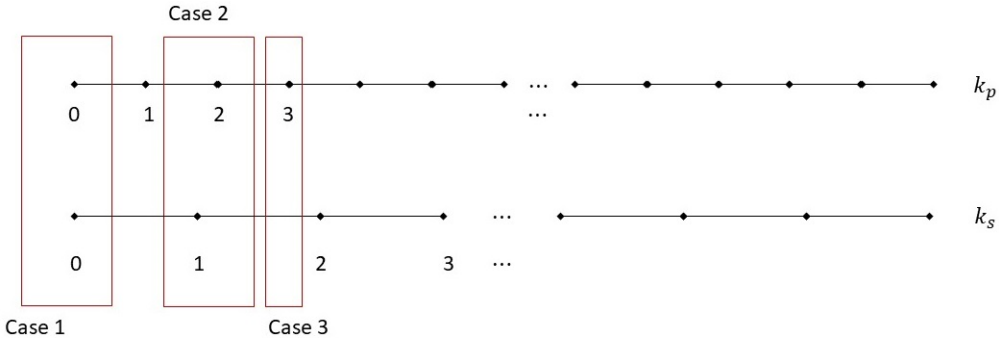


Figure 10.2: Three cases of measurement arrival Ghosal et al. (2017)

.

M1 The first approach was introduced in Ghosal and Rao (2019), and three cases are considered as it is depicted in Figure 10.2.

- The measurements from the sensors are gathered at the same time. In this case, data from the slower sensor would be used for the state estimation.
- The latest slow measurements arrive between the current and previous fast measurements. In this case, the UKF prediction step in Algorithm 10.1 would be used to estimate the states using the most recent slow measurement estimation, $\hat{x}_s$.
- The latest fast measurements arrive after fast measurements in the previous step, while no measurement from a slow sensor is available at this instant. The predicted x_s is obtained using the 'slow' model to propagate the estimate from its most recent estimate to the 'present' time.

For further clarification, please refer to Ghosal and Rao (2019).

M2 The second solution, which is more straightforward, is to extrapolate estimates from a slow sensor to the fuse estimates when slow measurements are not available. This extrapolation can be weighted averaging or regression-based extrapolations. The easiest version is to keep the

estimated from the slow measurements constant until observing a new measurement. This is the approach taken in this paper. The simulation results show that this approach works effectively for the conducted test case.

By assuming that the union of the observable subspaces of both sensors is the observable space of the entire system, we propose the Algorithm 10.1.

Algorithm 10.1: Algorithm for multi-rate sensor data fusion

- Determine the observable space for each set of measurements. For the fast measurements y_f there exist the corresponding observable state space $\mathcal{N}(\mathcal{O}_{of})$ and the unobservable state space $\mathcal{N}(\mathcal{O}_{uf})$. Similarly, sub-spaces for y_s are $\mathcal{N}(\mathcal{O}_{os})$ and $\mathcal{N}(\mathcal{O}_{us})$. The intersections of these sub-spaces lead to the state space decomposition Ghosal et al. (2017), Kotta (2005). Thus, we have

$$X_{oo} \in \mathcal{N}(\mathcal{O}_{of}) \cap \mathcal{N}(\mathcal{O}_{os}),$$

- For each set of sensors:
  - Find the lowest common multiple rates of sensors.
  - Calculate the sigma-points and propagate sigma-points through the system model (Algorithm 9.5) based on the lowest common multiple rates.
  - If a measurement is sensed, then update and propagate the sigma points through the measurement equation using Algorithm 9.6 and then perform a measurement-update using Algorithm 9.6 (M1), or
  - Propagate and update the slow model only at the sampling rate of the slow measurements, and keep the slow model estimates and covariances constant between each sample for the slow measurements (M2).
- For each time-step, apply the fusion step only for the states in the subspace X_{oo} only.

10.3.4 Covariance intersection

Covariance intersection is a method of combining two pieces of information in a way that the convex combination of the covariances characterizes the intersection Julier and Uhlmann (2009). The CI algorithm takes a convex combination of mean and covariance estimates that are represented information (inverse covariance) space Julier and Uhlmann (2009). The covariance intersection uses the following equations for obtaining the fused covariance and estimate,

$$\mathcal{P}_j^{-1} = \gamma \mathcal{P}_f^{-1} + (1-\gamma)\mathcal{P}_{sf}^{-1}$$

$$\mathcal{P}_j^{-1}\bar{x}_j = \gamma \mathcal{P}_f^{-1}\bar{x}_f + (1-\gamma)\mathcal{P}_{sf}^{-1}\bar{x}_{sf},$$

where the subscript f represents the fast sensor estimate. The subscript sf refers to the estimate based on the slow sensor but using fast sensor time steps, and the subscript j is the joint (fused) estimate. $\gamma \in [0, 1]$ is a coefficient which changes the weights assigned to $\bar{x}_f$ and $\bar{x}_{sf}$. The coefficient γ can be adjusted to optimize the update to meet different performance criteria, such as minimizing the trace or the determinant of $\mathcal{P}_f$. Note that the covariance size must be minimized at each time-update to guarantee non-divergence; otherwise, an updated estimate could be larger than the prior estimate. We use the optimization strategy in Julier and Uhlmann (2009) to optimize γ for fusing the estimates using outputs from the fast and slow filters.

Remark 10.1. *Note that the CI method automatically updates the states in the subspace X_{oo} only.*

Alternatively, the Bar Shalom-Campo (BSC) formula Bar-Shalom and Campo (1986) introduced in Ghosal and Rao (2019) can be applied. The fusion rule for BSC is given bellow:

Fusion rule. The fused state estimate for the system observed via multi-rate sensors can be defined as follows Ghosal and Rao (2019),

$$\hat{x}_j = [c_f \quad c_s][\hat{x}_f, \quad \hat{x}_{sf}]^{\mathsf{T}},$$

where $c_f + c_s = I$. The optimal Bar Shalom-Campo coefficients obtained in Bar-Shalom and Campo (1986) are as follows:

$$c_s^* = \mathcal{P}_{sf}^{-1}(\mathcal{P}_f^{-1} + \mathcal{P}_{sf}^{-1})^{-1}$$

and

$$c_f^* = \mathcal{P}_f^{-1}(\mathcal{P}_f^{-1} + \mathcal{P}_{sf}^{-1})^{-1}.$$

10.4 Simulation result

We use the WSCC-9 bus system Abooshahab et al. (2019), which is an IEEE benchmark Delavari et al. (2018), to assess the performance of the two different multi-rate data fusion approaches (sketched in Fig. 10.3). It is assumed that measurement with PMUs at buses six, seven, eight, and nine are sampled with higher rates, while measurement with PMUs at buses two, three, four, and five are sampled slowly. The standard deviation of the

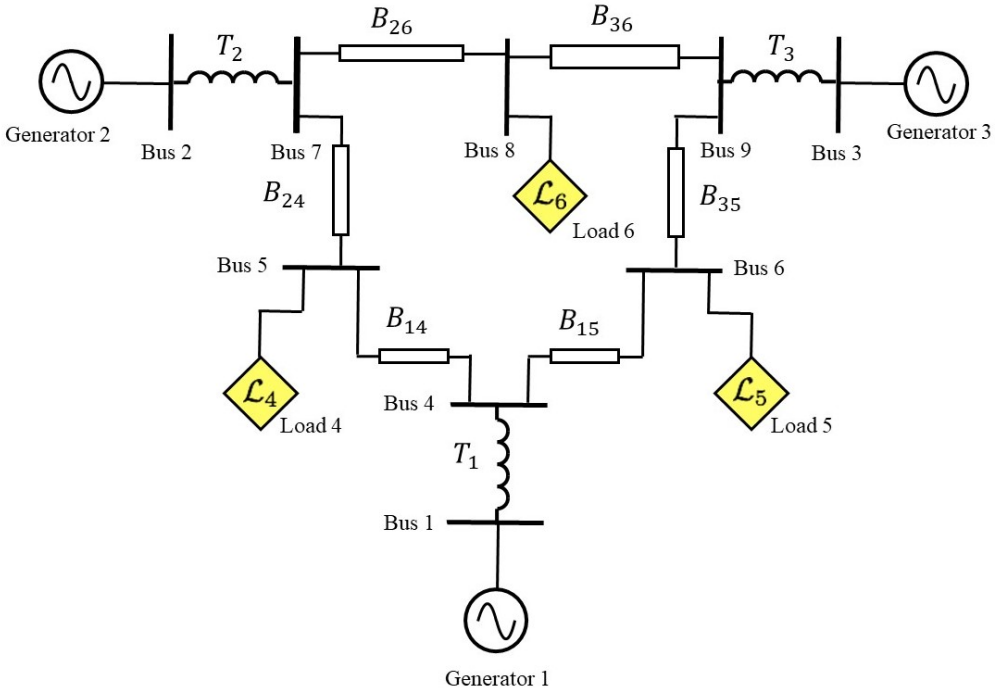


Figure 10.3: WSCC-9 bus system Abooshahab et al. (2019).

process noise is assumed to be 0.0003 p.u., and the measurement noise for the fast and slow sensors are chosen to be 0.01 and 0.02 p.u., respectively. It is assumed that a balanced three-phase fault is applied between the bus four and six during time $t = \{5s, 5.2s\}$. The sampling rate for the fast PMU is 0.03, which is four times faster than the slow PMU. According to the results given in Fig. 10.4 and Fig. 10.5, it can be seen that when the process and measurement noises are small, and enough information is available, all algorithms work well; besides, Fig. 10.4 shows that the BSC data fusion algorithm has the best performance in this case since it retains the smallest trace value for the error covariance. In addition, we can see the performance

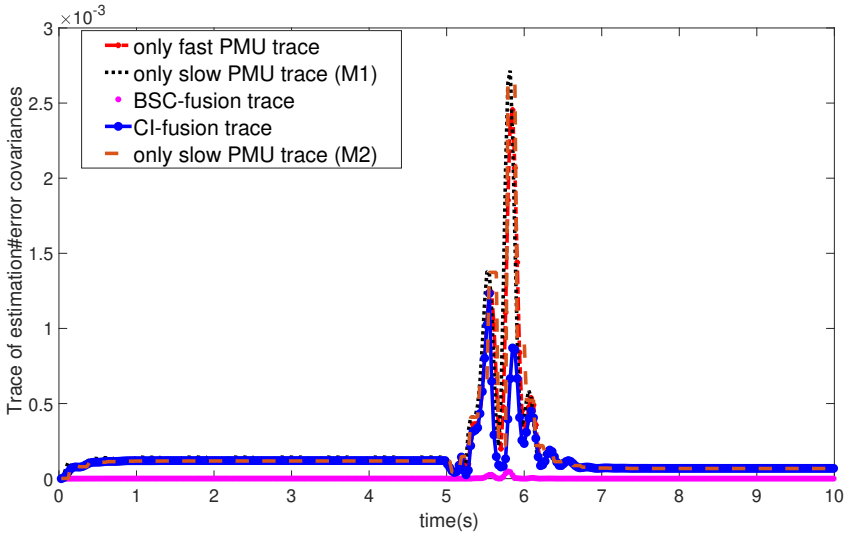


Figure 10.4: Trace of estimation error covariance for each algorithm when standard deviation of process noise is 0.0003 p.u.

of the method M2, which gives virtually the same results as M1. Hence, one may choose M1 proposed in Ghosal and Rao (2019), if the very accurate estimation is required while the computational effort is not important, and vice versa. The execution time for the total length of the simulation using M1 is 5.494 s, and for M2 is 3.687 s.

A series of 100 Monte-Carlo runs with different conditions were conducted processing the same data (to create an appropriate analogy), to investigate the trend of consistency for the BSC data fusion and the proposed method applying on our specific example. We use two methods to validate the performance of data fusion methods in terms of consistency. The first method is to find a condition that makes one of the filters unstable while the other one is stable. Thus, we repeat the procedure introduced above (only increase the value of the standard deviation of the process noise sampling at fast sensor time-steps to 0.3 p.u.) to investigate the trend of consistency. The obtained results illustrated in Fig. 10.6 shows that when we increase the standard deviation, state estimates' covariances using BSC starts to increase. In this situation, CI performs better than BSC. It can be observed that when the process noise's standard deviation is 0.8 or larger, then both state estimation based on slow sensor measurements and BSC fusion method tend to be unstable while the CI-fusion is stable, see Fig. 10.7. Hence, the method

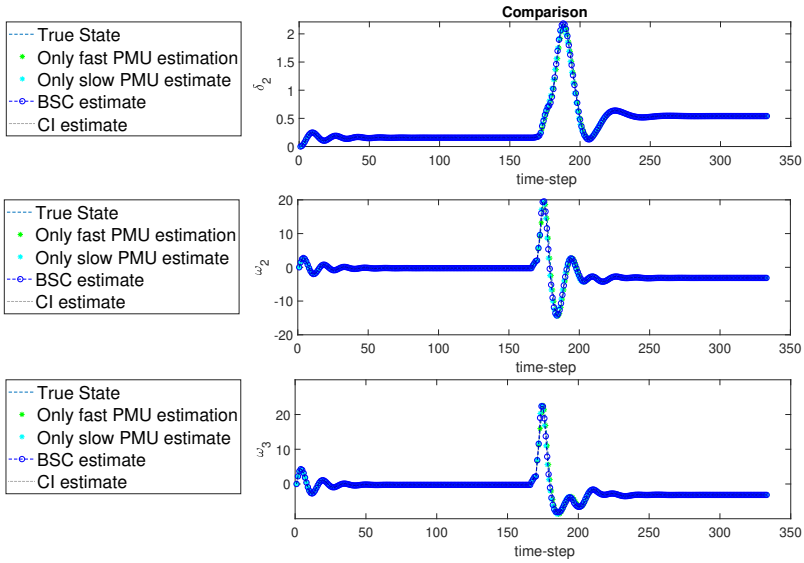


Figure 10.5: States of the system and its estimates for each algorithm when the standard deviation of process noise is 0.0003 p.u.

proposed in Ghosal and Rao (2019) is not applicable.

Secondly, we use the error metric; namely, the average normalized estimation error squared (ANEES) Bar-Shalom and Kirubarajan (2004). For all states, by averaging over all Monte Carlo runs and states, we can calculate ANEES for each time step, refer to Huang et al. (2008) for more details. Fig. 10.8 shows compared to that of BSC, the covariance intersection data-based fusion method exhibits better performance in the view of the ANEES. As a result, based on these two observations, we can conclude that the CI-based fusion method is more consistent in our nonlinear power system.

The substantial benefit of CI is that it is covariance consistent, which means if one of the estimates' covariance remains bounded, then the fused covariance and its corresponding estimate would be bounded. However, good consistency does not necessarily imply good Root Mean Square Error (RMSE). On the contrary, one may lose precision to get a good consistency, which is required for robust state estimation.

10.5 Conclusion

In this paper, we use CI as an alternative method for Bar Shalom-Campo in order to perform sensor fusion in power systems. It is observed that for

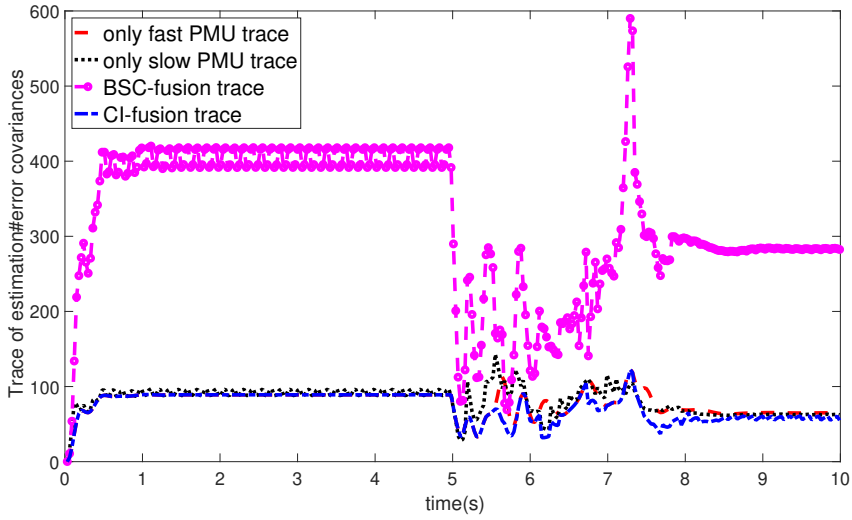


Figure 10.6: Trace of estimation error covariance for each algorithm when the standard deviation of process noise is 0.3 p.u. for the first and the second method handling the covariance of the slowly sampled measurements.

power networks, CI can deliver more consistent and robust results. Although BSC will, in some cases, provide more accurate results, state consistency is, however, found to be more important for power-related problems since the state estimation is used for operation critical purposes, including control, fault detection, and optimization of power grids. Hence, divergence in state estimation can lead to severe problems. In addition, we use extrapolation instead of propagation in the state prediction step when there is no new measurement. Finally, the fusion is only performed for the intersection of the observable subspaces, so each state only needs to be observable in (at least) one model/measurement for the overall system estimation to be observable.

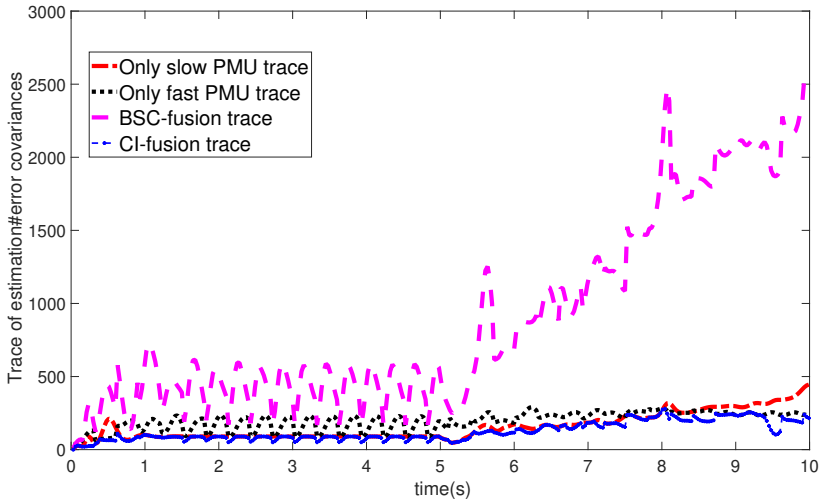


Figure 10.7: Unstable trace of estimation error covariance for the BSC algorithm when standard deviation of process noise is 0.8 p.u.

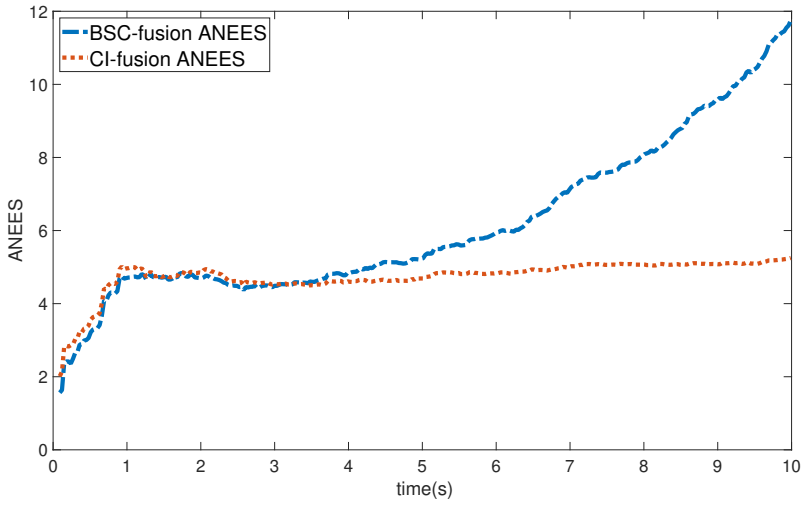


Figure 10.8: ANEES for a series of 100 Monte Carlo simulations for the WSCC-9-bus example.

Chapter 11

Paper VIII: Optimal PMU placement for partially known power system dynamic state estimation

Optimal PMU placement for partially known power system dynamic state estimation

Mohammad Ali Abooshahab
Morten Hovd
Giorgio Valmorbida
submitted to and accepted in *IEEE PES ISGT Europe 2021*, **Espoo, Finland** (2021)

Abstract

The synchronized phasor measurement unit provides fast, precise, and synchronized measurements, which is very important for the dynamic monitoring of the power systems. However, currently, very few nodes in the European power grid is covered by PMU measurements, since PMUs are expensive devices. Hence, with restricted budgets, the installation of these measuring devices should be selective. Previous work has been focused on steady-state estimation, and hence they focused on network topology to find

optimal configurations making the whole network observable at steady-state. However, this approach is not applicable for dynamic state estimation. Instead, we develop a greedy approach to obtain optimal PMU placement for dynamic estimation of partially known power grids.

Monitoring is a crucial part of the power grid's energy management system. Thanks to the introduction and development of synchronized phasor measurement units (PMUs), more frequent measurements are available in different parts of a power grid. This allows for the exploitation of dynamic approaches for power grid monitoring instead of steady-state or quasi-steady-state estimation. In this case, it is possible to track the sudden changes, including electromagnetic transients Zhao et al. (2019a), cyber attacks Ghosal and Rao (2019), Ghosal (2018), and dynamics of distributed generators Sadamoto et al. (2019). However, few nodes can be covered by PMUs in practice because they are expensive devices. Thus, it is necessary to place the PMUs selectively and optimally. Several methods have been conducted on PMU placement problems. A summary of different approaches to solving the PMU placement problem can be found in Manousakis et al. (2012). One of the most popular methods for solving the PMU placement problem is the greedy algorithm Tran and Zhang (2018), Li et al. (2012) due to its simplicity.

There are several interconnected elements or layers in a typical power system, including generation, transmission, distribution, and loads. Many parts of the system may be fully known with the information required for dynamic modeling available, while other parts may be entirely or partially unknown. Moreover, the availability of the measurements in some parts of the modern electrical grids, especially in the distribution part, is limited. This is due to the consumers' personal privacy and the lack of expensive and accurate measuring devices such as PMUs. Hence, the *partially known power system* concept has been recently introduced in Abooshahab et al. (2019). In this context, the Kalman Filtering - Simultaneous Input and State Estimation (KF-SISE), which has been proposed and investigated in Bitmead et al. (2019), Abooshahab et al. (2021), is used to estimate the states and unknown parts of the partially known power system Abooshahab et al. (2020b). This method requires fewer measurements in an extensive power network because it reduces the complete model to the partially known system, which needs only one measurement for each unknown signal; for further information, see Abooshahab et al. (2019; 2020b). In previous studies, however, optimal PMU placement for state estimation of partially known systems has not been investigated.

The aim of this chapter is, first, to formulate the optimization problem to solve optimal PMU placement for state estimation of partially known systems. To achieve this goal, we obtain the information version of KF-SISE and then propose 'best in' and 'worst out' greedy algorithms to optimize the placement of PMUs in a power network. The criteria we optimize is minimizing the covariance of estimation error. Finally, we present the results of numerical simulations on the IEEE-14-bus and WSCC-9-bus test systems, validating the performance of the proposed PMU placement method for KF-SISE based state estimation of partially known power networks.

This chapter is organized as follows. Section 11.1 introduces the KF-SISE and concept of partially known power systems. In Section 11.2, we formulate the sensor placement problem, and we solve this optimization problem using the greedy algorithm. The simulation results related to the WSCC-9-bus and IEEE-14-bus network are presented in Section 11.3 to show the proposed methods' performance.

11.1 Partially known power networks state estimation

We aim to minimize the overall covariance of the estimation error for the same number of PMUs. We first describe the filtering method. A summary of a KF-SISE, presented in Bitmead et al. (2019), is given in the following subsection.

11.1.1 Filtering algorithm

To simplify matters, we consider initially a linear time-invariant formulation without direct feedthrough for the known system:

$$x_{t+1} = Ax_t + Gd_t + w_t, \tag{11.1}$$

$$y_t = Cx_t + v_t, \tag{11.2}$$

where $x_t, w_t \in \mathbb{R}^n$ are the state vector and process noise at time step t, $d_t \in \mathbb{R}^m$ is the disturbance signal from the unknown part of the system, and $y_t, v_t \in \mathbb{R}^m$ are the measurements and the measurement noise. The following assumptions are required to use the Kalman filtering formulation of SISE Bitmead et al. (2019).

Assumption 11.1. *The assumptions are listed below:*

i $w_t \sim \mathcal{N}(0, Q)$, $v_t \sim \mathcal{N}(0, R)$ *and initial condition* $x_O \sim \mathcal{N}(\hat{x}_{0|0}, P_O)$ *are mutually independent Gaussian white noises,*

ii $R > 0$,

iii *Pair* (A, C) *is observable, and* $rank\,CG = rank\,G = m$.

iv *The disturbance signal* $d_t \sim \mathcal{N}(\mathfrak{d}, D)$ *is independent from* x_O, w_τ, v_τ *for all* t *and* τ.

By defining $\mathcal{X}_{t+1}$, $\mathcal{K}_{t+1}$, $\mathcal{M}_{t+1}$, $\mathcal{D}_t$ and $\mathcal{P}_{t+1}$ respectively as the prior state covariance matrix, the Kalman gain for the state vector, the Kalman gain for the unknown input vector, the posterior disturbance covariance matrix and the posterior state covariance matrix, and $\hat{x}_{t+1|t+1}$ as the posterior estimates for the state and $\hat{d}_{t|t+1}$ as the estimate for the unknown input of the system, the measurement sequence, $\mathbf{Y}^{t+1} \triangleq \{y_{t+1}, y_t, \ldots, y_1\}$, we can give the summarized algorithm for KF-SISE as in Algorithm 1.

Algorithm 11.1: State and disturbance estimation using KF-SISE

Prediction step: compute $\mathcal{X}_{t+1}$, $\mathcal{K}_{t+1}$, $\mathcal{M}_{t+1}$ as follows:

$$\mathcal{X}_{t+1} = A\mathcal{P}_t A^T + GDG^T + Q, \tag{11.3}$$

$$\mathcal{K}_{t+1} = \mathcal{X}_{t+1}C^T(C\mathcal{X}_{t+1}C^T + R)^{-1}, \tag{11.4}$$

$$\mathcal{M}_{t+1} = DG^TC^T(C\mathcal{X}_{t+1}C^T + R)^{-1}. \tag{11.5}$$

Update step: compute $x_{t+1|t+1}$, $d_{t|t+1}$, P_{t+1} and D_t, as follows:

$$\hat{x}_{t+1|t+1} = A\hat{x}_{t|t} + G\mathfrak{d} + \mathcal{K}_{t+1}(y_{t+1} - CA\hat{x}_{t|t} - CG\mathfrak{d}), \tag{11.6}$$

$$\hat{d}_{t|t+1} = \mathfrak{d} + \mathcal{M}_{t+1}(y_{t+1} - CA\hat{x}_{t|t} - CG\mathfrak{d}), \tag{11.7}$$

$$\mathcal{P}_{t+1} = (I - \mathcal{K}_{t+1}C)\mathcal{X}_{t+1}, \tag{11.8}$$

$$\mathcal{D}_t = (I - \mathcal{M}_{t+1}CG)D. \tag{11.9}$$

11.1.2 Information matrix derivation for KF-SISE

In this subsection, we derive the information matrix for KF-SISE Bitmead et al. (2019), which can simplify the analysis of estimation error and its covariance further in this study. The following lemma provides us with the information version of the KF-SISE.

Lemma 11.1. *Considering* (11.3)-(11.8)*, the Information matrix* $\mathfrak{I}_{t+1} = \mathcal{P}_{t+1}^{-1}$ *for the SISE can be obtained as follows:*

$$\mathfrak{I}_{t+1} = \mathcal{X}_{t+1}^{-1} + C^T R^{-1} C \tag{11.10}$$

Proof. Let us show that $\mathfrak{I}_{t+1} = \mathcal{P}_{t+1}^{-1}$. This will be done by showing that $\mathfrak{I}_{t+1}\mathcal{P}_{t+1} = I$, using the expression for $\mathcal{P}_{t+1}$ in (11.8).

Note that $\mathcal{P}_{t+1}^{-1}$ is full rank due to the observability condition in Assumption 11.1. We have, using (11.4):

$$\begin{aligned}
\mathfrak{I}_{t+1}\mathcal{P}_{t+1} &= (\mathcal{X}_{t+1}^{-1} + C^T R^{-1} C)(I - K_{t+1}C)\mathcal{X}_{t+1} \\
&= (\mathcal{X}_{t+1}^{-1} + C^T R^{-1} C) \\
&\times (\mathcal{X}_{t+1} - \mathcal{X}_{t+1}C^T(C\mathcal{X}_{t+1}C^T + R)^{-1}C\mathcal{X}_{t+1}) \\
&= I - C^T \Big[(C\mathcal{X}_{t+1}C^T + R)^{-1} - R^{-1} \\
&+ R^{-1}C\mathcal{X}_{t+1}C^T(C\mathcal{X}_{t+1}C^T + R)^{-1} \Big] C\mathcal{X}_{t+1} \\
&= I - C^T \big[(C\mathcal{X}_{t+1}C^T + R)^{-1}(I + R^{-1}C\mathcal{X}_{t+1}C^T) \\
&- R^{-1} \big] \times C\mathcal{X}_{t+1} \\
&= I - C^T [R^{-1} - R^{-1}] C\mathcal{X}_{t+1} \\
&= I
\end{aligned}$$

Since the inverse of a full rank matrix is unique, we conclude $\mathfrak{I}_{t+1} = \mathcal{P}_{t+1}^{-1}$. □

11.1.3 Steady state formulation of estimation

For the steady-state Kalman filter version of SISE, KF-SISE, we would have to solve an ARE (11.11) to obtain $\mathcal{X}_\infty$ - and then (11.12)-(11.15) follow Bitmead et al. (2019):

$$\mathcal{X}_\infty = \text{dare}(A^T, C^T, Q + GDG^T, R); \tag{11.11}$$

$$\mathcal{K}_\infty = \mathcal{X}_\infty C^T (C\mathcal{X}_\infty C^T + R)^{-1}; \tag{11.12}$$

$$\mathcal{M}_\infty = DG^T C^T (C\mathcal{X}_\infty C^T + R)^{-1}; \tag{11.13}$$

$$\mathcal{P}_\infty = (I_n - C\mathcal{K}_\infty)\mathcal{X}_\infty; \tag{11.14}$$

$$\mathcal{D}_\infty = (I_m - \mathcal{M}_\infty CG)D. \tag{11.15}$$

The ARE solution, here denoted $\mathcal{X}_\infty$, is the steady-state prediction error covariance from (11.3). $\mathcal{P}_\infty$ indicates the expected state estimation error corresponding to the given measurement matrix. For the same number of PMUs, we aim to minimize this overall error.

11.1.4 Model for the test system

This subsection is devoted to describe a dynamic model for the power network and to clarify the partially known power network concept. To derive a

State variables		External variables		Parameters	
angular frequency	ω_r	external load	P_G	rotor damping coeficient	D
mechanical power	P_m	frequency setpoint	f_O	time constant	T_S
rotor shaft angle	δ	load setpoint	L_{cre}	governor feedback gain	k
valve position	a			droop characteristic	R
				motor inertia	M

Table 11.1: Variables and parameters for components in the test cases.

dynamic model for our power network, the following assumptions are consideredAbooshahab et al. (2019):

Assumption 11.2. *The assumptions are listed below:*

i *Transformers are neglected because their impedances are negligible compared to the impedances of the transmission lines.*

ii *The ratio of the transmission lines reactance to their resistance is assumed to be significantly above unity Hertem et al. (2006).*

iii *The power angle at Bus 1 is assumed to be the network's reference power angle.*

In the test systems, synchronous generators and loads contribute to the power system dynamics. Generators dynamic model can be described as Blood (2011), Kundur et al. (1994):

$$\frac{\partial}{\partial t}\begin{bmatrix}\omega_r\\ P_m\\ a\\ \delta\end{bmatrix}_k = \begin{bmatrix}-D/M & -1/M & 0 & 0\\ 0 & -1/T_S & 1/T_S & 0\\ k & 0 & -kR & 0\\ 1 & 0 & 0 & 0\end{bmatrix}_k \begin{bmatrix}\omega_r\\ P_m\\ a\\ \delta\end{bmatrix}_k + \begin{bmatrix}0 & 0\\ 0 & 0\\ -k & -k\\ 0 & 0\end{bmatrix}_k \begin{bmatrix}L_{cre}\\ \omega_O\end{bmatrix}_k + \begin{bmatrix}-1/M\\ 0\\ 0\\ 0\end{bmatrix}_k P_{Gk}, \quad (11.16)$$

where variables and parameters are given in Table 11.1. Similarly, buses containing rotating loads with known P_L is Blood (2011),

$$\frac{\partial}{\partial t}\begin{bmatrix}\omega_r\\ \delta\end{bmatrix}_k = \begin{bmatrix}-D/M & 0\\ 1 & 0\end{bmatrix}_k \begin{bmatrix}\omega_r\\ \delta\end{bmatrix}_k + \begin{bmatrix}-1/M\\ 0\end{bmatrix}_k P_{Lk} + \begin{bmatrix}-1/M\\ 0\end{bmatrix}_k P_{Gk}. \quad (11.17)$$

The model derivation process for the WSCC-9-bus network as in Abooshahab et al. (2019) is given in the Appendix. The abridged version of the power

network model can be given as:

$$\dot{\bar{\mathbf{x}}} = \mathcal{A}\bar{\mathbf{x}} + \mathbf{w}$$

where $\mathbf{w}$ represents the noise originating from modeling uncertainties, and $\bar{\mathbf{u}}$ is the input corresponding to reordered state vector with input matrix $\mathcal{B}$. For this simplified model, the states associated with unknown parts will not appear in the system dynamics. The unknown inputs are denoted by d and the truncated state vector and its system matrix are denoted by $\bar{\mathbf{x}}_{tr}$ and $\mathcal{A}_{tr}$, respectively. Thus, the model for a partially known power network can be given as:

$$\dot{\bar{\mathbf{x}}}_{tr} = \mathcal{A}_{tr}\bar{\mathbf{x}}_{tr} + \mathbf{G}\mathbf{d} + \mathbf{w}_{tr}$$

where $\mathbf{w}_{tr}$ denotes the modeling noise for the partially known power network. Note that the truncated part of $\mathbf{B}^{(\mathbf{P})}\delta$ corresponds to the Gd here. We refer to Abooshahab et al. (2019; 2020b) for further details.

11.1.5 The measurement model

The modern PMUs can measure several variables in the power system, including bus voltage, bus current, valve position, and the output of the power system stabilizer Yang et al. (2007), Wang et al. (2012). Moreover, state variables for a bus with a PMU can be measured directly Wang et al. (2012). In this case, we can choose the valve position a_k, the mechanical power $P_{m,k}$, rotor velocity $\omega_{r,k}$, and power angle δ_k as available measurements at node i with PMU Wang et al. (2012). Thus, the PMU-based measurement equation can be modeled as a linear function of states.

$$\mathrm{y} = C\mathrm{x} + \mathrm{v}, \quad y, v \in \mathbb{R}^p$$

Next, we assume measurement noises are independent, and therefore that R is diagonal. Let $C_{[k]}$ denote the kth row of the measurement matrix C, and $R_{[k,k]}$ denote the kth diagonal element of R. Next, we introduce the following definition:

Definition 11.1. *Sensing precision matrix.*

For each measurement k, the sensing precision matrix S_k, and its assimilated version $\tilde{S}$ are defined in Yang et al. (2015; 2013) as follows:

$$S_k = C_{[k]}^T R_{[k,k]}^{-1} C_{[k]}, \tag{11.18}$$

$$\tilde{S} = \sum_{k=1}^{p} C_{[k]}^T R_{[k,k]}^{-1} C_{[k]} \tag{11.19}$$

Remark 11.1. *Note that the assimilated sensing precision matrix $\tilde{S}$ is the second term of the information matrix equation (11.10), thanks to the structure of R.*

This model fits the model (11.1)-(11.2), so filtering algorithms described in Section 11.1 are applicable for partially known systems.

11.2 Sensor placement problem formulation

This section aims to propose a systematic way for sensor placement for power network state estimation using S-SISE. The sensor selection vector is defined in Liu et al. (2016) as

$$\mathrm{w} = [w_1, \cdots, w_m]^T, \quad w_k \in \{0, 1\},$$

where w_k indicates whether the ith sensor is available or not. Hence, available measurement equation can be expresses as

$$y_t^w = C_w x_t + \Phi_w \nu_t$$
$$\text{and} \quad C_w = \Phi_w C, \tag{11.20}$$

where $\Phi_w \in \{0, 1\}^{||\mathrm{w}||_1 p}$ is a matrix constructed from $diag(\mathrm{w})$ such that all rows corresponding unselected sensors have been omitted from this matrix.

11.2.1 Problem statement

The problem we wish to solve can be stated as

$$\min_{w} \operatorname{trace}(\mathcal{P}_{\infty,\mathrm{w}}) \quad \text{subject to} \quad 1^T \mathrm{w} \leq n_s, \tag{11.21}$$

with $(\mathcal{P}_{\infty,\mathrm{w}})$ in (11.14) and defined above, where $n_s \leq p$ is the maximum number of sensors available. Note that Boolean variables in w in the constraints of the optimization problem and the matrix C_w the optimization problem make it non-convex. The two lemmas below will be used in our first result, which shows that adding a sensor can only improve the cost in (11.21).

Lemma 11.2 (Zhang et al. (2017), Yang et al. (2015))**.** *For two selections* w *and* $\tilde{\mathrm{w}}$*, if* $w_k = \tilde{w}_k$ *for* $i \in \{1, \cdots, m\} \backslash j$*,* $w_j = 0$ *and* $\tilde{w}_j = 1$ *then* $\mathcal{X}_{t+1,\tilde{\mathrm{w}}} \leq \mathcal{X}_{t+1,\mathrm{w}}$ *and* $\mathcal{X}_{\infty,\tilde{\mathrm{w}}} \leq \mathcal{X}_{\infty,\mathrm{w}}$*.*

Lemma 11.3 (Zhang et al. (2017), Horn and Johnson (2012))**.** *For two given matrices A, B if $A \geq B$ then $A^{-1} \leq B^{-1}$*

Theorem 11.1. *If* w *and* $\tilde{\mathrm{w}}$ *are two sensor selections such that* $w_k = \tilde{w}_k$ *for* $i \in \{1, \cdots, m\} \backslash j$, $w_j = 0$ *and* $\tilde{w}_j = 1$ *then* $\mathcal{P}_{\infty,\tilde{\mathrm{w}}} \leq \mathcal{P}_{\infty,\mathrm{w}}$.

Proof. Based on Lemma 3, we exploit information matrix for $t = 0, \cdots, k-1$, using (11.10) and (11.18) we obtain

$$\mathfrak{I}_{1,\tilde{\mathrm{w}}} = \mathcal{X}_{1,\tilde{\mathrm{w}}}^{-1} + \tilde{S}_{\mathrm{w}} + S_j$$

where $\tilde{S}_{\mathrm{w}}$ is the assimilated sensing precision matrix for the sensor selection set w, and S_j is the the sensing precision matrix defined in (11.18). From Lemma 2 and Lemma 3, we have that $\mathcal{X}_{1,\tilde{\mathrm{w}}}^{-1} \geq \mathcal{X}_{1,\mathrm{w}}^{-1}$; thus,

$$\mathfrak{I}_{1,\tilde{\mathrm{w}}} \geq (\mathcal{X}_{1,\mathrm{w}}^{-1} + \tilde{S}_{\mathrm{w}}) + S_j = \mathfrak{I}_{1,\mathrm{w}} + S_j$$

$S_j \geq 0$; hence,

$$\mathfrak{I}_{1,\tilde{\mathrm{w}}} \geq \mathfrak{I}_{1,\mathrm{w}}.$$

Repeating this procedure until $k-1$, yields

$$\mathfrak{I}_{k,\tilde{\mathrm{w}}} \geq \mathfrak{I}_{k,\mathrm{w}}.$$

Then, following Lemma 1, we have

$$\mathcal{P}_{k,\tilde{\mathrm{w}}} \leq \mathcal{P}_{k,\mathrm{w}},$$

which holds for all k, letting $k \to \infty$ gives

$$\mathcal{P}_{\infty,\tilde{\mathrm{w}}} \leq \mathcal{P}_{\infty,\mathrm{w}}.$$

□

As a result, activating a new sensor does not degrade the estimation performance. Therefore the inequality constraint in (11.21) can be changed to an equality constraint.

Note that, if we in KF-SISE also wish to take the input estimation into account, so the problem given in (11.21) can be extended to:

$$\min_{w} \operatorname{trace}(\mathcal{P}_{\infty,\mathrm{w}}) + \operatorname{trace}(\mathcal{D}_{\infty,\mathrm{w}}) \quad \text{subject to} \quad 1^T \mathrm{w} \leq n_s, \tag{11.22}$$

Theorem 11.2. *If* w *and* $\tilde{\mathrm{w}}$ *are two sensor selections such that* $w_k = \tilde{w}_k$ *for* $i \in \{1, \cdots, m\} \backslash j$, $w_j = 0$ *and* $\tilde{w}_j = 1$ *then* $\mathcal{D}_{\tilde{\mathrm{w}}} \leq \mathcal{D}_{\mathrm{w}}$.

Proof. Take

$$\mathcal{K}_{t+1} = \mathcal{X}_{t+1} C^T (C\mathcal{X}_{t+1} C^T + R)^{-1}$$

By performing matrix inversion, we can obtain

$$\mathcal{K}_{t+1} = \mathcal{P}_{t+1} C R^{-1} = \mathcal{X}_{t+1} C^T (C\mathcal{X} C^T + R)^{-1}.$$

Thus,

$$C^T (C\mathcal{X}_{t+1} C^T + R)^{-1} = \mathcal{X}_{t+1}^{-1} \mathcal{P}_{t+1} C R^{-1}.$$

Therefore,

$$\mathcal{M}_{t+1} = D G^T \mathcal{X}_{t+1}^{-1} \mathcal{P}_{t+1} C R^{-1}.$$

In addition,

$$\mathcal{D}_t = (I - \mathcal{M}_{t+1} C G) D. \tag{11.23}$$

Consider the second term of above equation

$$\mathcal{M}_{t+1} C G D = D G^T \mathcal{X}_{t+1}^{-1} \mathcal{P}_{t+1} C R^{-1} G D.$$

Now, we rewrite the above equation for $y_{t+1}^{\tilde{w}}$ as

$$\mathcal{M}_{t+1}^{\tilde{w}} C_{\tilde{w}} G D = D G^T {\mathcal{X}_{t+1}^{\tilde{w}}}^{-1} \mathcal{P}_{t+1}^{\tilde{w}} C_{\tilde{w}} R^{-1} C_{\tilde{w}} G D.$$

Using Theorem 11.1 leads to:

$$\mathcal{M}_{t+1}^{\tilde{w}} C_{\tilde{w}} G D \geq D G^T {\mathcal{X}_{t+1}^{w}}^{-1} ({\mathcal{X}_{t+1}^{w}}^{-1} + S_j)^{-1} (C_w R^{-1} C_w^T + S_j) G D.$$

Note that since ${\mathcal{P}_{t+1}^{w}}^{-1} \geq {\mathcal{X}_{t+1}^{w}}^{-1} \geq C_w^T R^{-1} C_w$, we have

$$\mathcal{M}_{t+1}^{\tilde{w}} C_{\tilde{w}} G D \geq D G^T {\mathcal{X}_{t+1}^{w}}^{-1} \mathcal{P}_{t+1}^{w} (C_w^T R^{-1} C_w) G D,$$

and then

$$\mathcal{M}_{t+1}^{\tilde{w}} C_{\tilde{w}} G D \geq \mathcal{M}_{t+1}^{w} C_w G D,$$

which means

$$\mathcal{D}_{t+1}^{\tilde{w}} \leq \mathcal{D}_{t+1}^{w}.$$

□

11.2.2 Greedy algorithm for sensor placement

This section proposes two methods based on a greedy algorithm to solve Problem (11.21). The solution is obtained such that the equality constraint in (11.21) holds. One algorithm starts with zero sensors, $n_w = 0$, and increments n_w until (11.21) holds; we call this the 'best in' greedy algorithm. The second algorithm starts with $n_w = p$ and decrease n_w until $n_w = n_s$, we call this the greedy algorithm 'worst out'.

Note that Assumption 11.1. iii should be satisfied to reconstruct all signals of our interest, The two greedy algorithms are given in Algorithm 11.2 and 11.3.

Algorithm 11.2: Greedy algorithm 'best in' Shamaiah et al. (2010)

1. Initialization:
$$k = 1, \quad \Phi_k = \Phi_w, \mathrm{w}_k = \{\}$$
where Φ_w is defined in (11.20).

   Necessary measurements selection:
   Identify measurement sets w_k of minimum cardinality satisfying Assumption 11.1. iii.

   (a) From sets of measurements satisfying Assumption 11.1. iii, select w_k such that:
$$w_k = \arg\min \operatorname{trace}(\mathcal{P}_\infty^w)$$

   (b) Update the measurement set:
$$\Phi_{k+1} = \Phi_k \backslash w_k, \mathrm{w}_{k+1} = \mathrm{w}_k \cup w_k$$

2. Next measurement selection:
$$w_k = \arg\min \operatorname{trace}(\mathcal{P}_\infty^w)$$

3. Update the measurement set (as in 1)b) above)

4. If $1^T w_k < n_s$ go to step 2) else Stop.

Algorithm 3 is called the Greedy descent algorithm or greedy 'worst out'

Algorithm 11.3: Greedy descent 'worst out'

1. Initialization:
$$k = 1, \quad \Phi_k = \Phi_w,$$
and w_k is initialized as vector of ones.

2. Next measurement selection for deletion:
$$w_k = \arg\min |\mathrm{trace}(\mathcal{P}_\infty^w(\mathrm{w}_k \backslash \{w_k\}))|$$

3. Update the measurement set and select the next measurement:
$$\Phi_{k+1} = \Phi_k \backslash w_k, \mathrm{w}_{k+1} = \mathrm{w}_k \backslash w_k$$

4. If $1^T w_k > n_s$ go to step 2) else Stop.

algorithm. In this algorithm, it is assumed that all measurements are available. Then the worst measurement from the measurement set would be omitted.

Remark 11.2. *Algorithm 11.3 is easier to implement, and it does not need to perform a recoverability check (Assumption 11.1. iii). Hence, it is more straightforward to exploit Algorithm 11.3 when using the KF-SISE method in general. However, for large systems with a low number of sensors, Algorithm 11.3 should be used because of its lower computational cost and faster speed.*

11.3 Numerical results

In this section, we test the proposed method on the WSCC-9-bus network Fig. 11.1 and evaluate its optimality. To illustrate the proposed method, we assume each sensor can measure one state at each bus. Hence, we need to solve Problem (11.21) with $n_s = 4$. After solving the Problem (11.21) with the Greedy algorithm given in Algorithm 11.2, we found that first, we need two measurements to assess the angular frequency at bus 1 and 2; in addition, measuring valve position first at bus one and then at bus two can give us the minimum value for our objective function, while Table 11.2 shows that this solution is the optimal solution for this problem.

We also implement Algorithm 11.3 for this problem. We start with the full set of available sensors (for simplicity, it is assumed the maximum num-

ber of sensors is nine and each sensor can measure only one state of the system), and then select the worst sensor concerning the optimality criteria and remove the worst sensor from the optimal sensor set. This is done until the optimal sensor set contains the maximum number of measurements that is $n_s = 4$ here. Using Algorithm 11.3 yields removing ω_{r4}, P_1, P_2, δ_2, δ_4 from the the optimal sensor set, respectively and results in $C_{\mathrm{w}^*} = \{\omega_{r,1},\ \omega_{r,2}, a_1,\ a_2\}$ as the the optimal sensor set. This sensors set is the same as the one we obtained with Algorithm 11.2.

The estimation results for the WSCC-9-bus system is given in Fig. 11.2. From Fig. 11.2-11.4, we see that the estimates are accurate for the system states.

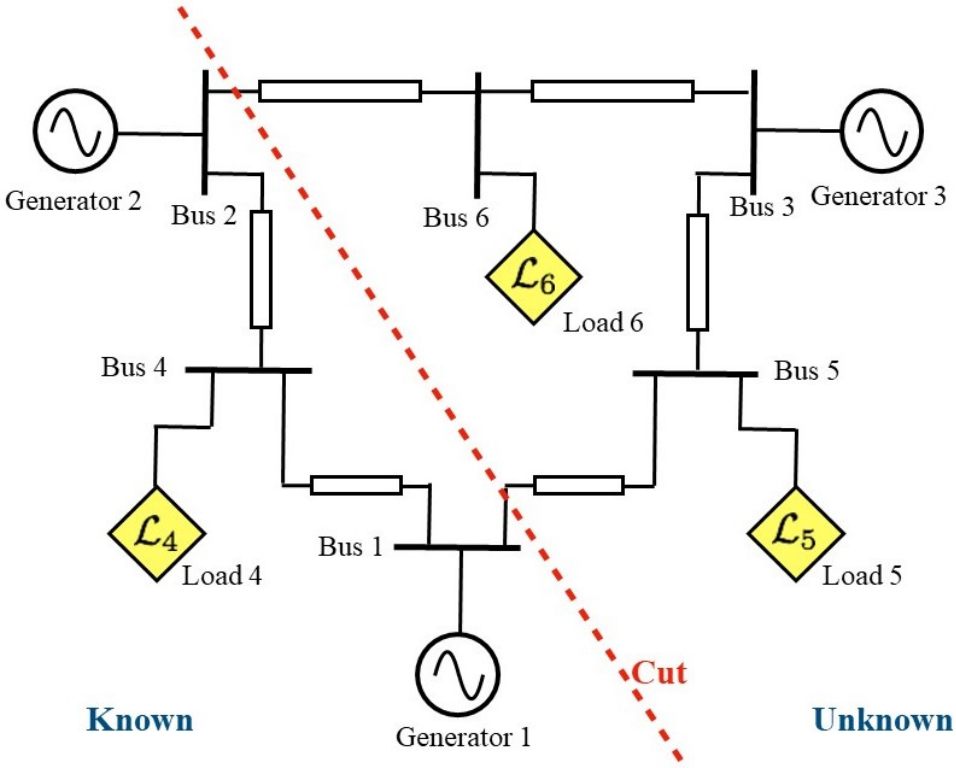


Figure 11.1: WSCC-9 bus system Abooshahab et al. (2019).

We further use the sensor selection method for the IEEE-14 bus system having only 3 measurements. The optimal sensor selection set is found as $\{\omega_{r,1},\ \omega_{r,2},\ \omega_{r,3}\}$. For this sensor selection, $\mathrm{trace}(\mathcal{P}_\infty + \mathcal{D}_\infty) = 86.75 + 17.61$.

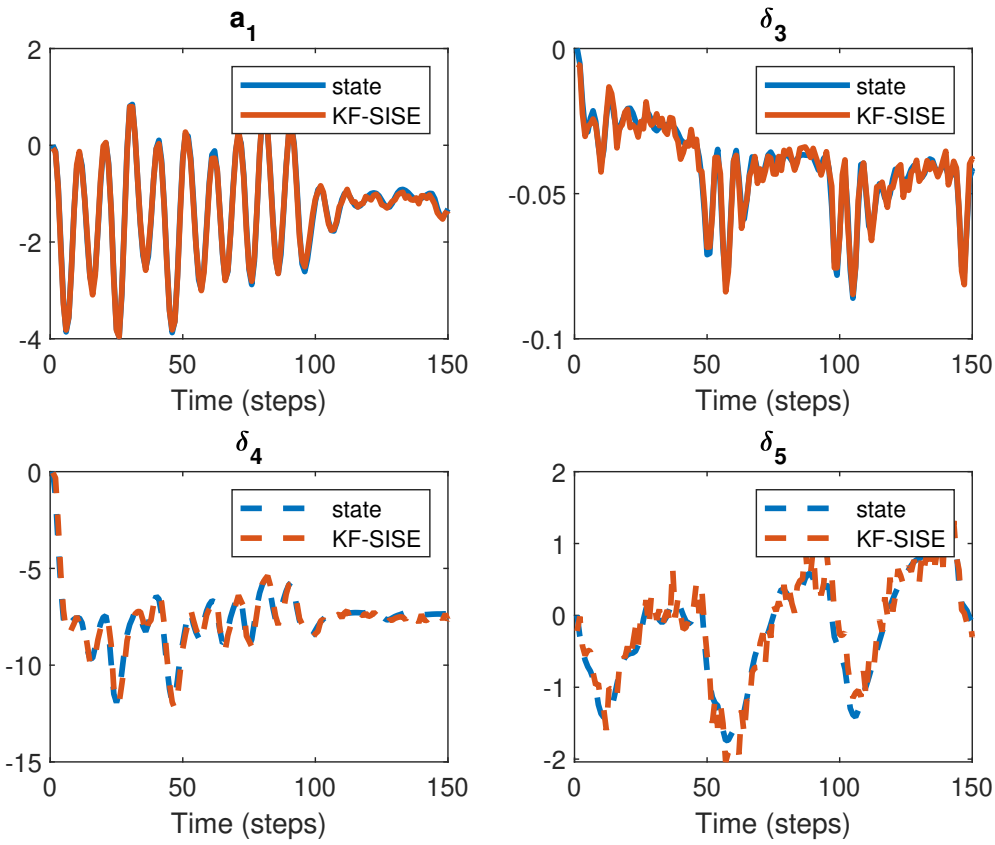


Figure 11.2: State estimation for the WSCC-9-bus system using KF-SISE and the proposed sensor placement algorithms.

Measurements	trace$\mathcal{P}_\infty$ + trace$\mathcal{D}_\infty$
ω_{r1}, ω_{r2}	$30.62 + 6.84$
$\omega_{r1}, \omega_{r2}, \delta_2, \delta_4$	$11.84 + 2.52$
$\omega_{r1}, \omega_{r2}, \delta_2, a_1$	$7.59 + 1.69$
$\omega_{r1}, \omega_{r2}, \delta_2, P_1$	$9.59 + 2.09$
$\omega_{r1}, \omega_{r2}, \delta_2, a_2$	$4.6 + 1.24$
$\omega_{r1}, \omega_{r2}, \delta_2, P_2$	$5.79 + 1.52$
$\omega_{r1}, \omega_{r2}, \delta_2, \omega_{r4}$	$12.61 + 3.12$
$\omega_{r1}, \omega_{r2}, \delta_4, a_1$	$10.14 + 2.24$
$\omega_{r1}, \omega_{r2}, \delta_4, P_1$	$11.97 + 2.82$
$\omega_{r1}, \omega_{r2}, \delta_4, a_2$	$5.1 + 1.46$
$\omega_{r1}, \omega_{r2}, \delta_4, P_2$	$6.19 + 1.62$
$\omega_{r1}, \omega_{r2}, \delta_4, \omega_{r4}$	$10.15 + 2.42$
$\omega_{r1}, \omega_{r2}, a_1, P_1$	$14.52 + 3.49$
$\omega_{r1}, \omega_{r2}, a_1, a_2$	$4.16 + 0.84$
$\omega_{r1}, \omega_{r2}, a_1, P_2$	$7.61 + 1.71$
$\omega_{r1}, \omega_{r2}, a_1, \omega_{r4}$	$15.41 + 3.76$
$\omega_{r1}, \omega_{r2}, P_1, a_2$	$6.59 + 1.65$
$\omega_{r1}, \omega_{r2}, P_1, P_2$	$10.15 + 2.34$
$\omega_{r1}, \omega_{r2}, P_1, \omega_{r4}$	$21.591 + 4.42$
$\omega_{r1}, \omega_{r2}, a_2, P_2$	$8.56 + 1.74$
$\omega_{r1}, \omega_{r2}, a_2, \omega_{r4}$	$9.90 + 2.11$
$\omega_{r1}, \omega_{r2}, P_2, \omega_{r4}$	$14.56 + 3.62$
$\omega_{r1}, \omega_{r2}, P_1, P_2, \delta_2$	$4.77 + 1.02$

Table 11.2: Trace of $\mathcal{P}_\infty$ + trace of $\mathcal{D}_\infty$

Several state estimates and their actual states for IEEE-14 bus system are depicted in Fig. 11.4.

Conclusion

This chapter presents two greedy algorithms for the optimal placement of PMUs for partially known power networks. Simulations are carried out on two power benchmarks, and results are presented. The information version of KF-SISE is derived, and it is used to solve the optimization problem with the greedy algorithm. Also, we investigate the optimality of the two proposed greedy algorithms. By considering the effects of disturbance on KF-SISE, we derive the new formulation for the optimization problem.

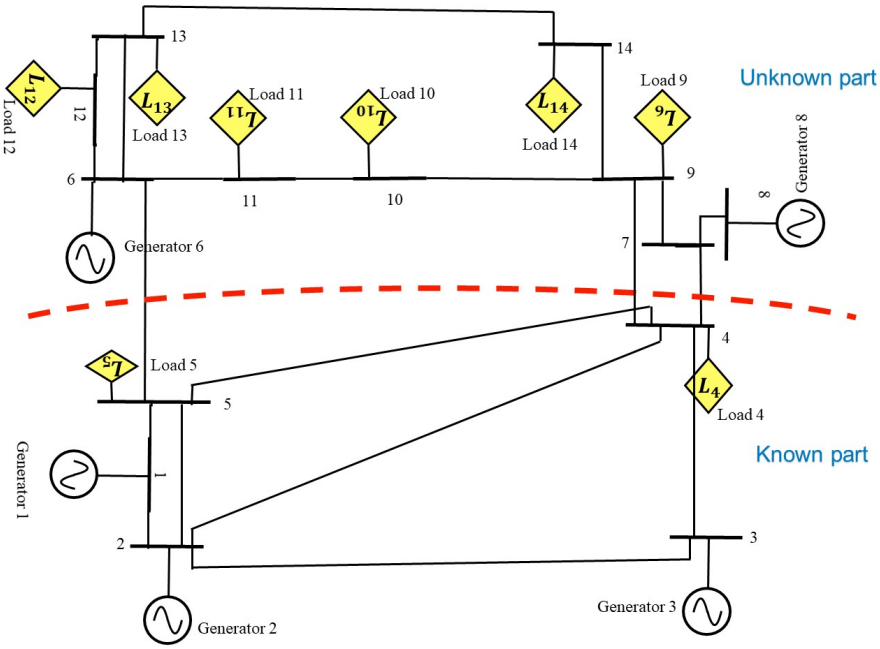


Figure 11.3: Transformer-less dynamic power grid model of IEEE-14-bus-system, with circuit cut dividing known and unknown parts Abooshahab et al. (2020b).

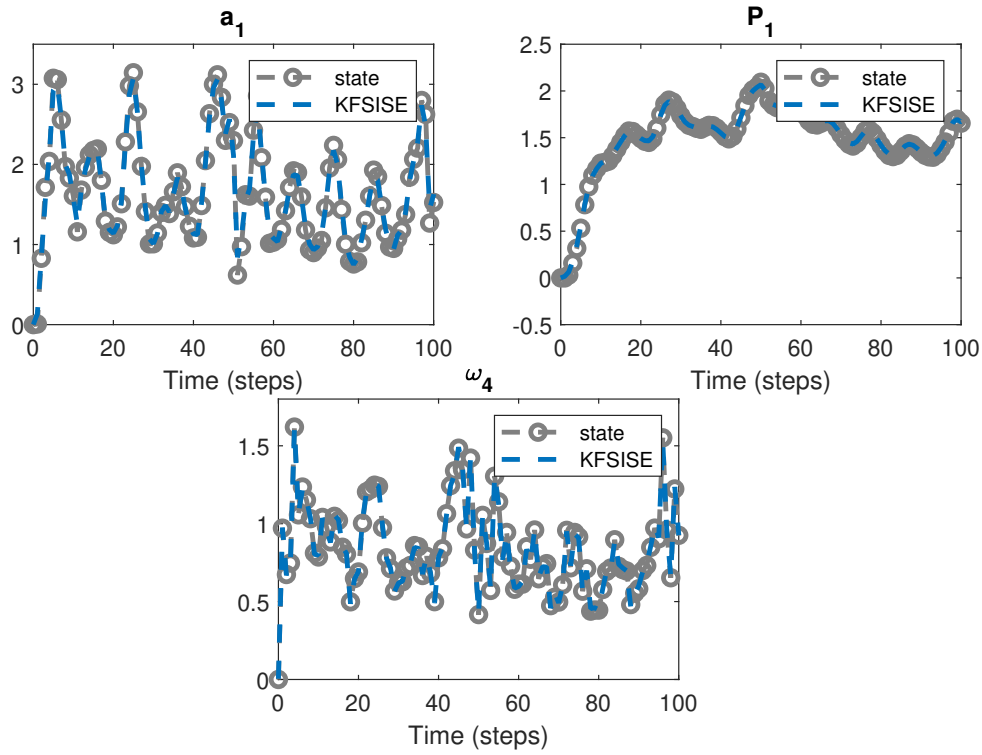


Figure 11.4: State estimation for the IEEE-14-bus system using KF-SISE and the proposed sensor placement algorithms.

Chapter 12

Additional topics on dynamic power system monitoring

The preceding chapters document that benefits of using the dynamic state estimation specifically S-SISE (KF-SISE) for the dynamic state monitoring of power systems include:

1. Although power systems are continuous-time systems, their state estimation process is mainly done discretely. The discrete-time model of a continuous-time system is more likely to contain a non-minimum phase zeros since the standard sampling process is known to introduce zeros, some of which may lie outside the unit circle Maciejowski (1985). The schemes given Ghahremani and Kamwa (2011a;c; 2016) for the DSE of power networks with unknown inputs are not stable when the system has NMP zeros. Hence, they do not seem to be appropriate approaches dealing with DSE problem for power networks with unknown inputs. However, our method is a good tool to handle unwanted non-minimum phase behavior of the discrete model of the process.

2. The topological description of the state estimation does not have to change for different operating conditions.

3. Our proposed method needs fewer measurements compared with conventional state estimation methods.

4. Another possible alternative to address the problem of monitoring partially known power networks is the augmented Kalman filter (AKF).

In the S-SISE model, we seek to construct an estimate of the disturbance first and then compute an estimate of the state next; and we do this without imposing a model structure on disturbance. With the AKF model, we assign a model to the disturbance and try to estimate both the plant state and the disturbance model state from the same output. Therefore, the observability properties of AKF for power systems applications are more stringent. Furthermore, if we have an exact model of the network, then it is completely" known" otherwise it is "unknown" at least in part. Choosing the disturbance dynamics as an integrator, as is commonly done with the AKF is otherwise hard to justify, whereas S-SISE is insensitive to the disturbance model. For power grids, the unknown parts are typically fast rather than slow due to DERs, responsive loads, and microgrids.

Several advantages of our proposed estimation methods will be described in this chapter.

1. For the small signal stability analysis, electromechanical states of a power system are needed. Subsequently, the estimates of these states can be exploited to adjust the PSSs' parameters, and so improving the system stability. The conventional PSS uses local measurements for stabilizing local states. The functionality of local PSS will be improved if it has information about the other unknown parts of the system. Thus, by applying our framework, we can estimate the unknown signals from other parts of the system without the need for additional measurement from other parts. That can be a big development for the PSS (this advantage will become clear after reading the details of point in Section 12.5).

2. The TSO/DSO interacting signals can be easily estimated using our proposed method (more details are provided in Section 12.4).

3. The reliability of the protection system can be improved Zhao et al. (2019a). Checking consistency between the measurements and the outputs of the dynamical model of the protection zone leads to detect anomalies stemming from internal and external faults. This results in more reliable protection systems compared with the traditional coordinated settings-based schemes (we will propose fault detection methods in Section 12.2.

We first need to introduce the decentralized version of S-SISE, since it is

not only useful for fault detection but also for scalability of the estimation and the handling of topology changes.

12.1 Decentralized filter using S-SISE

A decentralized scheme of estimation is a practical alternative for large and complex interconnected systems. We term a local assembly of interconnected components a *unit*. Now assume that we want to estimate the states of one of the units in the system, and we call this unit the *local unit* while all units connected directly to the local unit are termed *interconnecting units*. Interconnecting (coupling) signals between interconnecting units and the local unit are assumed to be unknown inputs for the local state estimator. This formulation, which is not unique, allows us to decouple the dynamic equations of the local unit from the dynamic equations of interconnecting units. Furthermore, just the knowledge of interconnecting signals for many networked systems is sufficient to estimate the dynamics of the local unit. These interconnecting signals can be obtained by measuring them with measuring devices or estimating them as unknown inputs using SISE. Note that the reference angle must be the same for all units, and this can be assumed as a coordinating signal between different units.

Consider the following linear time-invariant equation as a model for the jth unit satisfying Assumption 4.1.

$$x_{t+1}^j = A^j x_t^j + G^j d_t^j + w_t^j, \tag{12.1}$$

$$y_t^j = C^j x_t^j + v_t^j, \tag{12.2}$$

where w_t^j and v_t^j are mutually independent Gaussian white noises, $w_t^j \sim \mathcal{N}(0, Q^j)$, $v_t^j \sim \mathcal{N}(0, R^j)$. Algorithm 12.1 proposes the decentralized version of S-SISE.

12.2 Bad data and fault detection

There are two objectives in fault detection in the distribution system. The first is to identify the faulted line section or the faulted network area to make a fast system restoration by switching and changing some loads or switching topology to recover generation. The second objective is to identify the exact fault location to repair the faulty component. Here, we focus on the identification of the faulty section. The identification of the exact fault location to repair the component is usually made by visual inspection for overhead lines or with dedicated devices for underground cables Gerstner et al. (2013).

Algorithm 12.1: Decentralized DSE for the jth unit using S-SISE

Begin Determine $A^j,\ G^j,\ C^j,\ Q^j,\ R^j$. Initialization;

$$x_0^j \sim \mathcal{N}(\hat{x}_{0|0}^j, P_0^j)$$

while $t \geq 1$ **do**

Covariance prediction;

$$\mathcal{X}_{t+1}^j = A^j \mathcal{P}_t^j A^{j^T} + G^j D^j G^{j^T} + Q^j,$$

Kalman Update;

$$\mathcal{K}_{t+1}^j = \mathcal{X}_{t+1}^j C^{j^T} (C^j \mathcal{X}_{t+1}^j C^{j^T} + R^j)^{-1},$$

$$\mathcal{M}_{t+1}^j = D^j G^{j^T} C^{j^T} (C^j \mathcal{X}_{t+1}^j C^{j^T} + R^j)^{-1}.$$

$$\begin{aligned} \mathcal{P}_{t+1}^j &= \mathcal{X}_{t+1}^j - \mathcal{X}_{t+1}^j C^{j^T} (C^j \mathcal{X}_{t+1}^j C^{j^T} + R^j)^{-1} C^j \mathcal{X}_{t+1}^j, \\ &= (I - \mathcal{K}_{t+1}^j C^j) \mathcal{X}_{t+1}^j, \end{aligned}$$

Output and time update;

$$\begin{aligned} \hat{x}_{t+1|t+1}^j &= A^j \hat{x}_{t|t}^j + G^j \mathfrak{d}^j \\ &\quad + \mathcal{K}_{t+1}^j (y_{t+1}^j - C^j A^j \hat{x}_{t|t}^j - C^j G^j \mathfrak{d}^j), \\ \hat{d}_{t|t+1}^j &= \mathfrak{d}^j + \mathcal{M}_{t+1}^j (y_{t+1}^j - C^j A^j \hat{x}_{t|t}^j - C^j G^j \mathfrak{d}^j), \end{aligned}$$

$$t = t + 1$$

end

Many solutions have been suggested to identify the fault and its location in distribution systems such as using fault passage indicators. Fault passage indication can be used for short-circuit and/or earth-fault location in distribution networks Lehtonen et al. (1995). These devices trigger if they sense a certain amount of current, which indicates if the fault is upstream or downstream at the measurement point. This information can be used for the identification of the faulty zone of the network. In Lehtonen et al. (1995), it is found that such methods are prone to having not enough directional functions for networks with distributed generation (DG), while it becomes more complicated to find which way the current is flowing in closed loop systems.

The other method is to use the traveling wave algorithms in distribution systems Saha et al. (2009). However, the transmission of power from the generators to the distribution parts are often composed of different types of line sections: the cross sectional area decreases further away from the generators as less power is transmitted. This results in discontinuities of impedance in the distribution part, which generates many wave reflections. Generation of many wave reflections requires complicated signal processing and needs fast devices Saha et al. (2009), in order to reliably use traveling wave algorithms for fault detection. In isolated or islanded networks, the distribution system's charging transients should be taken into account for islanding fault location Janssen (2014). The reason is that the fault current value can be too small to be sensed in steady-state measurements. An alternative is to measure current magnitude. A short-circuit calculation may be derived for every part of the system such that a list of current magnitudes for every generation become available. Then during a fault occurrence, the computed current magnitude will be sent to the SCADA, and this value is compared with the list in order to determine the faulted part Lehtonen et al. (1995). An extension of such methods is to determine the reactance measured by distance sensors which is known as the impedance-based fault detection algorithm Janssen (2014). However, impedance-based fault detection in distribution systems has the following issues when comparing with its transmission counterpart:

- The distribution system topology is mostly radial.
- The substations are unbalanced.
- The connecting lines are shorter, which can complicate the detection of the fault location.

- The network length can range from a few kilometers for urban power networks to hundreds of kilometers for rural areas Lehtonen et al. (1995).
- The ratio of the system resistance to the system reactance, r/x is often large.

There exist several impedance-based fault detection methods taking advantage of substation voltage and current measurements for radial networks Girgis et al. (1993), Filomena et al. (2009). A problem with these methods is that they cannot handle introduction of DGs in a power system. Hence, they will not be applicable for modern distribution grids with several DGs.

In Bretas and Salim (2006), the DGs are modeled as their Thevenin equivalent models to simplify the system. In Kezunovic (2011), it is suggested to use the feeders' voltage measurements as well as other branches voltage magnitudes for fault detection. In Venkata et al. (2013), they extend such approach for distributed synchronized voltage phasor measurements. The main weakness with all aforementioned fault detection methods is that they cannot detect transient faults since they are designed for quasi-static operating points in power networks. Hence, we present a fault detection method based on decentralized SISE in the following section.

Principle of the proposed fault detection algorithm

The use of distributed measurements may enhance the robustness and accuracy of fault detection, especially in the context of networks with DGs. In this chapter, a fault detection algorithm using distributed measurements is proposed. The algorithm is based on the state estimation framework already developed in this thesis, S-SISE for partially known power systems, which is a novel strategy for fault detection. The main idea is based on the fact that faults change the system admittance; and after fault occurrence the system matrix will be changed; further, SISE performance for each local system is completely independent of unknown parts; hence, during fault occurrence in one local subsystem only its own estimator would have large estimation error, but estimation accuracy for other parts of the system would be unchanged. In addition, our fault detection method can distinguish between different types of bad data in a power network. Hence, we start with introducing different types of data.

12.2.1 Different types of bad data

Bad data can be determined by using the normalized measurement residuals.

Definition 12.1 (Normalized measurement residuals Gomez-Exposito and Abur (2004), Ghosal et al. (2017))**.** *The normalized measurement residuals is defined as*

$$r_{y_{t+1}} = \epsilon_{t+1}^{\mathsf{T}} \tilde{R}_{t+1}^{-1} \epsilon_{t+1} \tag{12.3}$$

where $\tilde{R}_{t+1}^{-1} = (CX_{t+1}C^{\mathsf{T}} + R)^{-1}$ *and* $\epsilon_{t+1} = y_{t+1} - CA\hat{x}_{t|t}$.

$r_{y_{t+1}}$ is a random variable with χ-squared distribution with m_i degrees of freedom, where m_i is the number of independent measurements Gomez-Exposito and Abur (2004), Ghosal et al. (2017), Caro et al. (2011). In a state estimator, bad data is detected if the following inequality is true:

$$r_{y_{t+1}} > \mathrm{r}_{bd},$$

at any time instant $t+1$, where r_{bd} is an appropriately chosen threshold value.

Definition 12.2 (Malicious measurement Ghosal et al. (2017))**.** *Malicious measurement,* $y^a \in \mathbb{R}^m$*, is represented by*

$$y_{t+1}^a = y_{t+1} + a_{t+1} = Cx_{t+1} + v_{t+1} + a_{t+1},$$

where $a \in \mathbb{R}^m$ *is termed the attack vector.*

The malicious data is a type of bad data and consequently is detected if the following inequality is true:

$$r_{y_{t+1}} > \mathrm{r}_{bd},$$

at any time instant $t+1$. Thus, there is a requirement for determination of different types of bad data, while each one requires a specific action for being handled. The value of r_{bd} could be determined by ensuring that under normal conditions (no attack) the value of the normalized measurement residuals is less than r_{bd} with certain high probability (referred to as the confidence level).

12.3 Methodology based on decentralized version of SISE

By computing the normalized measurement residuals, defined as $r_{y_{t+1}}$ in (12.3) , we can determine the suspicious measurements. If the majority of measurements in the system (especially measurements close to each other) are found to be suspicious, then it is possible that a system fault is happened. In this section, we propose a novel fault detection method to distinguish between the different faults. We start with the following definitions for various faults happening in a power system.

12.3.1 Stealth attack

During a stealth attack the original measurement y_{t+1} is modified to y^s_{t+1} but still satisfying the following inequality Miao et al. (2014),

$$r^s_{y_{t+1}} = \epsilon^{s\mathsf{T}}_{t+1} \tilde{R}^{-1}_{t+1} \epsilon^s_{t+1} \leq \mathrm{r}_{bd}$$

Definition 12.3. *A stealth attack is defined as the act of perturbing the measurements y between time-instants (t_0+1) to (t_0+t_f), for some $t_f \geq 1$, such that $r^s_{y_{t+1}} \leq \alpha \;\; \forall t \in \{t_0, \cdots, t_0+t_f-1\}$ but the resulting deviation in the state estimates due to the attack exceeds some pre-defined threshold $\beta > 0$ by the time instant (t_0+t_f).*

Here, we focus on a specific scenario for the stealth attack by assuming that the attacker wants to inject a stealth attack signal in such a way that the state estimation error is large at the same time instant, i.e. $t_f = 1$. In this work, we consider an additive attack model as presented below. Suppose the modified measurement vector is:

$$y^s_{t+1} = y_{t+1} + \Gamma a_{t+1}$$

Here, Γ is a normalized attack vector whose elements in each row signify how the attack is channelized into the measurements, while $a_{k+1} \geq 0$ is a non-negative scalar which quantifies the magnitude of the injected attack Kwon et al. (2013), Ghosal et al. (2017).

In our proposed method, we first try to detect anomalies in the system. When a suspicious measurement y^s is found such that $\forall\, k$, $|\frac{y^k_{t+1} - C^k A^k \hat{x}_{t|t}}{\tilde{R}^{k,k}_{t+1}}| > \mathrm{r}_{bd}$. Then we enable a number of secure measurements satisfying $rankCG = m$ for each unit. The remaining measurements are treated as (potentially) suspicious. After that each node shares the estimated interacting signals with their neighbours. From Corollary 4.1, we concluded S-SISE prioritize the estimation of the unknown input over that of the state. Hence, there is a possibility to use only reliable measurements for recovering unknown inputs and then use these estimated unknown inputs to estimate states of the system. In the presence of reliable measurements, the S-SISE can be described in four steps. In this setting, the secure measurements and estimated unknown inputs based on S-SISE approach have impact on the state estimates; however, the unreliable measurements would not affect the estimation of the input signals. By comparing different estimates of interacting signals from different units/nodes i.e. the interacting signals from neighbour d^{flow} and their estimates $\hat{d}^{\text{flow}}$, we can determine the faults type such that if

the unknown input estimation error, which is found from comparing input estimates with neighboring units i.e, $||\hat{d}^{\text{flow}}_{t|t+1} - \hat{d}_{t|t+1}||$, is large then we will have a system fault. For the small unknown input error, less than Th_d, the suspicious measurements are malicious data (note that the uncertainties due to the measurement noise are considered in the confidence interval). The threshold Th_d can be determined by selecting a proper confidence value based on the inverse cumulative probability distribution. For the cases that we have no suspicious data but we have large unknown input estimation error then the system may face stealth attack. The algorithm is summarized in Algorithm 12.2.

12.4 TSO/DSO state estimation

Monitoring in a traditional electric grid typically has been done in a tree structured way. A conventional hierarchical multilevel state estimation structure is shown in Fig. 12.1. At the lowest level, a local SE can be implemented in the distribution level to preliminarily deal with the information collected within a substation or small set of adjacent substations. A great majority of raw measurements will be processed at this distributed level, where a modest but sufficient computing power already exists. Distribution substations, delivering power to a large number of secondary transformers through a set of radial feeders, constitute a particular relevant case. In those substations, it is advantageous and makes sense to process each radial feeder in a decoupled manner, leading to a fourth level of information processing.

The results provided by the local SE have to be transmitted through existing RTUs and communication channels to the TSO-level SE (TSE). At this intermediate level, commercially available software can be adopted with minor modifications, the major difference with respect to a conventional SE being that prefiltered rather than raw measurements are handled.

At the uppermost level, a regional SE (RSE) will be needed to synchronize and refine the results separately provided by each TSO affiliated with the interconnected system, particularly near the border nodes. The RSE will be a customized tool, designed in such a way that the amount of information exchanged with subordinate TSEs is kept to a minimum. This SE level will significantly benefit from wide-area measurements provided by PMUs.

12.4.1 A flat structure TSO/DSO state estimation based on SISE

Here our aim is to construct a pedagogical approach to flatten the conventional estimation methods using SISE and synchronization methods for

Algorithm 12.2: A systematic way to find bad data and their types in a power network

Available information: measurements y such that $\text{rank}(C) > m$ and $\text{rank}(CG) = m$, and the input estimation error prior to time t, $||\hat{d}^{\text{flow}}_{t-1|t} - \hat{d}_{t-1|t}||$. For each time step $(t+1)$

1. Identify: suspicious measurements y^s, by $\forall\ k,\ |\frac{y^k_{t+1} - C^k A^k \hat{x}_{t|t}}{\tilde{R}^{k,k}_{t+1}}| > \text{r}_{bd}$.
2. Determine: the threshold Th_d for $||\hat{d}^{\text{flow}}_{t|t+1} - \hat{d}_{t|t+1}||$.
3. If suspicious measurement is found
   (a) Enable a set of sufficient number of measurements (rank $CG = m$) as secure ones.
   (b) Receive from neighboring units estimates of the flowing (interconnecting) signals into the system defined as $\hat{d}^{\text{flow}}_{t|t+1}$.
   (c) Estimate the unknown inputs (interconnecting signals) $\hat{d}_{t|t+1}$ using SISE.
   (d) Check: $||\hat{d}^{\text{flow}}_{t|t+1} - \hat{d}_{t|t+1}|| > Th_d$ then we have system fault, else we have malicious data.
4. If suspicious measurement is not found, but $||\hat{d}^{\text{flow}}_{t|t+1} - \hat{d}_{t|t+1}|| > Th_d$, then the system is exposed to a stealth attack. When a stealth attack is detected, one may find which suspicious measurement(s) contribute more to the input estimation error - and thereby identify the measurement(s) under attack.

When a suspicious measurement (bad data), a system fault or a stealth attack is found, then recognize the susbsystem as a possible fault location.

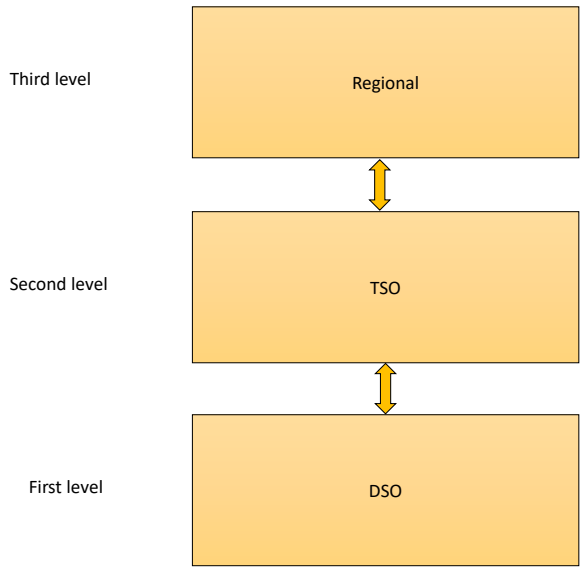


Figure 12.1: Hierarchical multilevel architecture of conventional state estimation. The adjacent levels may share information with each other that is represented by the double arrows.

sensors. As we stated in Chapter 1, the uncertainty and speed of changes in modern electrical grids are increasing, and many basic assumptions for power systems are therefore becoming doubtful. As a result, there should be fundamental changes in energy management systems (EMSs) of future power systems. One of those assumptions being doubtful is the quasi-static assumption. The hierarchical structure shown in Fig. 12.1 is designed based on quasi-static assumption. In this multi-level state estimation, information in each layer will be shared with upper layer estimator corresponding to the sensors' sampling rates.

In modern power systems, fast sensors such as PMUs will be placed on nodes both in the transmission and distribution levels of these systems, while only these types of sensors can capture the fast changes and their dynamics in modern grids. The other change in the future smart grids is that it will not be easy to partition power networks into high and low voltage parts, as we may have higher voltage parts in the distribution part of the system than that of the current distribution systems. In such situations, it is important to use all information provided in the system to obtain the best state estimates. In this case, we need to capture the fast dynamics of the system, while at the same time synchronize state estimators' frequencies of

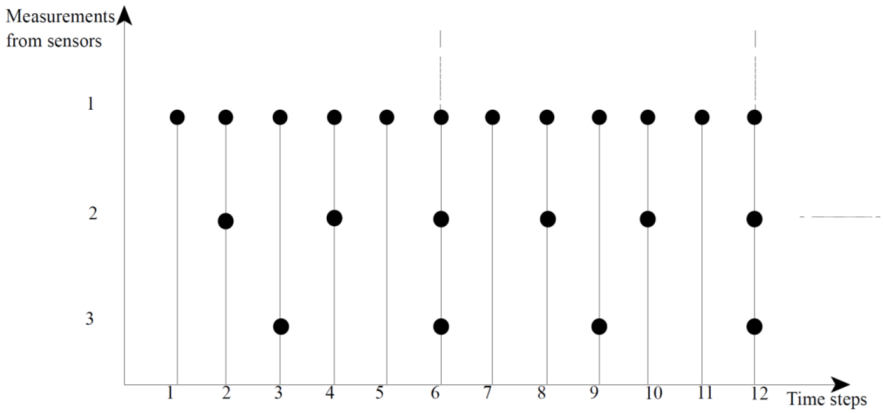


Figure 12.2: Sampling rate and time steps for three types of sensors.

different types of sensors.

12.4.2 Synchronization between estimates results from multirate sensors

In this section, we aim to establish a mathematical approach for synchronizing state estimators' frequencies of different types of sensors. The sampling case of multisensor multirate systems can be described in Fig. 12.2. Three types of sensors for example are considered in Fig. 12.2. The circles show the sampling time of different sensors. The sensors have the sampling rate $r_{T_l} = 1/T_l,\ l = 1, 2, 3$. The sampling period of the second sensor is $r_{T_1}/2$ and the third sensor is $r_{T_1}/3$. It is clear that the least common multiple of three sample rates and sampling time are r_{T_1} and $6T_1$, respectively. This means that the samplings of different sensors are asynchronous in each data block of the length $6T_1$. Individual sensors provide their local estimates based on their own observations. Then local estimators are sent into a centralized or local fusion center. with sampling rate r_{T_1}. However, from Fig.12.2, one can conclude that sensor 2 and 3 do not have measurements at many time steps. In addition, there exist some issues regarding discretization of a continuous system which are listed below Axelsson and Gustafsson (2015):

- For the nonlinear case, the first or second order Taylor series expansion is considered for the system approximation.
- The Euler or Runge-Kutta integration is used to discretize the con-

tinuous system function.

- The discrete-time noise is an aggregation of the total effect of the process noise between two time-steps. The problem with this approximation is that the process noise is based on the sampling time, while in practice it can have its own statistical characteristics.

To address these problems, the comminuted version of Kalman filtering based estimation is given.

12.4.3 The Comminuted Estimator (CE)

Consider the continuous-time system (4.1). The comminuted version of Kalman filtering based estimation for a sensor is implemented in a sampling time smaller than the sensor's sampling time i.e. the CE is implemented with sampling time $T_{CE} = T/\mathsf{n}$, where $\mathsf{n} \in \mathbb{N}$. For the missing measurement instances, the prediction step is performed. Doing more iterations in the Kalman filter's prediction step is referred to *oversampling* in Axelsson and Gustafsson (2015).

As mentioned, some parts of a power network can be unknown for some particular operators, for example, distribution system operators may have no information about the transmission part of the system. In this respect, the effects of transmission system on the distribution part can be modeled as unknown input signals, and then apply S-SISE to estimate the system dynamics and also TSO/DSO interaction signal. First we need to divide the DSO/TSO parts by performing cuts and obtain the decoupled DSO subsystems and TSO part (as shown in Fig. 12.3). S-SISE allows for the unknown power inputs to be estimated simultaneously with the state estimation, for each subsystem independently. The algorithm to perform a flat-structured state estimation for a power network can be summarized as in Algorithm 12.3.

12.5 Power system stabilizer

The power system stabilizer (PSS) control target is to damp the generator rotor angle swings, which are in a broad range of frequencies in the power system. These range from low-frequency modes (0.1 - 1.0 Hz), to local modes (typically 1 - 2Hz), to intra-plant modes (about 2 -3 Hz). The low-frequency modes, commonly called interarea modes, are caused by coherent groups of generators swinging against other groups in the interconnected system. These modes are present in all large-scale power systems, and the damping is a function of tie-line strength and unit loading factors. Weak ties due to

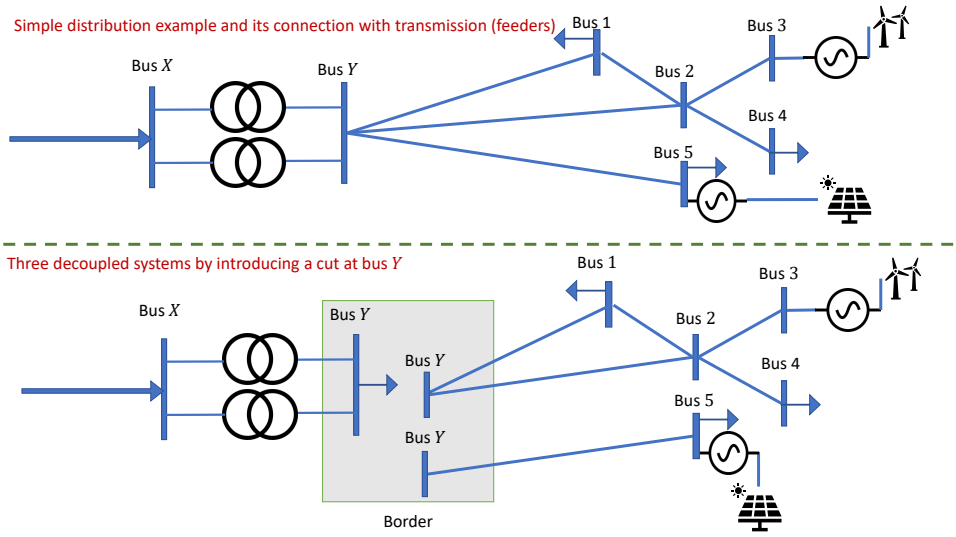


Figure 12.3: Decoupling a system with TSO/DSO parts into different subsystems (Bus X is part of transmission part and Bus Y is connection between the DSO and TSO.

Algorithm 12.3: a flat-structured state estimation for a power network

1. Synchronize all estimators in a DSO part. Synchronize all estimators in a TSO part.
2. Distinguish connecting points between TSO and DSO parts.
3. Decouple the system into TSO/DSO parts. For the DSO, the unknown disturbance is the signals coming from TSO and vice versa.
4. Apply S-SISE to jointly estimate the interacting signals and states of the DSO. Apply S-SISE to jointly estimate the interacting signals and states of the TSO.

line outages and heavy system loads can lead to poorly damped modes. PSS control can generally provide significant improvements in interarea mode damping by applying stabilizers to most units that participate in power swing modes.

The block diagram of a synchronous machine dynamics containing PSS and auto voltage regulator (AVR) is shown in Fig. 12.4. As we can see in Fig. 12.4, the input for the PSS is the rotor speed deviation $\Delta\omega_r$, which is used to generate a damping torque component.

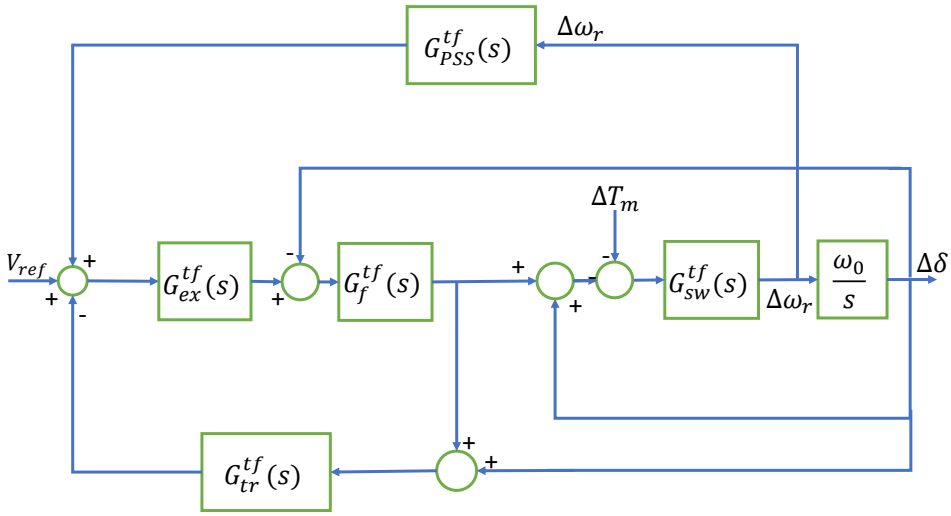


Figure 12.4: Block diagram of a synchronous machine dynamics containing PSS and automatic voltage regulator (AVR). G_{tr}^{tf}, G_{ex}^{tf}, G_f^{tf} and G_{sw}^{tf} are denoting the transfer functions of the voltage transducer, exciter system, field current dynamics and swing dynamics. ΔT_m is the external mechanical torque.

If the G_{ex}^{tf} and swing dynamics were pure gains, direct feedback of $\Delta\omega_r$ would generate the damping torque. However, both the generator and the exciter input contain frequency dependant gain and phase characteristics. Thus, the PSS transfer function, G_{PSS}^{tf} should have phase compensation, see Fig. 12.5.

12.5.1 Loop shaping regulator for power system stabilizers

The optimal linear quadratic Gaussian power system stabilizer (LQG-PSS) has been investigated before in Zolotas et al. (2006), Seo et al. (1996), Singh and Pal (2018), Singh and Pal (2016), Florescu et al. (2012). These references all assume that all parts of the system are known. Using SISE for the

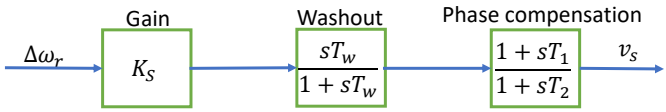


Figure 12.5: Block diagram of a PSS containing a phase compensation.

linear-quadratic Gaussian (LQG) problem can make it complicated because, in SISE, we have unknown inputs needing to be estimated, and control inputs needing to be properly determined at the same time. However, as we discussed before, the S-SISE algorithm is a regular Kalman filter for finite D. Thus, it is possible to employ Kalman filtering based control strategies including the LQG with S-SISE. In order to investigate the effect of LQG, we need to consider the control signal coming from the LQG-PSS in the state space model of a generator. This is modeled as follows:

$$\dot{x}_i = A_i x_i + B_i u_i + \mathcal{B}_{PSSi} u_{PSSi} + B_i^{(P)} P_{Ei}. \tag{12.4}$$

where

$$\mathcal{B}_{PSSi} = \begin{bmatrix} 0 & 0 & \frac{1}{M_i \tau_{Ci}} & 0 \end{bmatrix}^T$$

The procedure of designing the S-SISE-LQG regulator for the partially known power system (PKPS) contains two steps. First, S-SISE is designed to obtain the state estimates of the PKPS. Then based on the estimated states rather than actual states the control input for the system can be calculated (see Fig. 12.6). In this section, we investigate the application of S-SISE for PKPS from different aspects. We start with a simple example to explain the basics of our proposed method, and then we discuss some of the properties related to power systems applications.

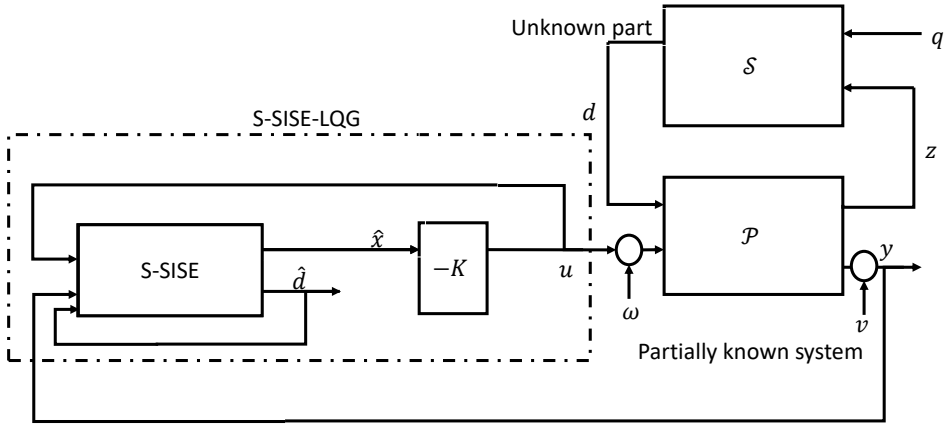


Figure 12.6: Block diagram of the S-SISE-LQG regulator

12.5.2 IEEE-3-bus test case

Fig. 12.7 (from Blood (2011)) depicts a three-bus power network consisting of two generators and one consumer. The link admittances are as displayed, and Generator 1 is managed by the network operator with Generator 2 being managed by an alternative supplier. Accordingly, the dynamics of generator 1 – its moment of inertia, feedback control gain, mechanical and electrical properties – are fully known to the operator along with the admittances of the network. A model is available for Generator 2, meaning that its precise mechanical and electrical properties are known. The consumer load is unreliably modeled and could exhibit sudden rapid changes.

Suppose that no model is available for the machine at bus 3. and we regard the power flow at that bus as an input signal. Then states associated with that machine no longer appear in the system model, and the model for the

known part of the system becomes

$$\frac{d}{dt}\begin{bmatrix} a_1 \\ P_1 \\ \omega_1 \\ a_2 \\ P_2 \\ \omega_2 \\ \delta_2 \end{bmatrix} = \left[\begin{array}{ccccccccc} -5 & 0 & 100 & & & & & & \\ 0.2 & -0.2 & 0 & & & & & & \\ 0 & -0.1 & -0.15 & & & & & & 0.08 \\ \hdashline & & & -5 & 0 & 125 & & & \\ & & & 0.3 & -0.3 & 0 & & & \\ & & & 0 & -0.2 & -0.3 & -0.1 & -0.3 & \\ \hdashline & & -1 & & & 1 & & & \end{array}\right] \times \begin{bmatrix} a_1 \\ P_1 \\ \omega_1 \\ a_2 \\ P_2 \\ \omega_2 \\ \delta_2 \end{bmatrix} + \left[\begin{array}{c} \\ 0.01 \\ \hdashline \\ \\ \\ 0.14 \\ \hdashline \end{array}\right] \left[\delta_3\right]. \quad (12.5)$$

With measurement vector

$$C_1 = [0 \;\; 0 \;\; 1 \;\; 0 \;\; 0 \;\; 0 \;\; 0],$$

we have observability – meaning that, given the external inputs and measurement of Generator 1 rotor frequency, we can estimate the system state for the known part of the network. Importantly, we also satisfy the single input estimation condition Assumption 4.1 that rank $C_1 G_1 = 1$, with $G_1 = [0 \;\; 0 \;\; 0.01 \;\; 0 \;\; 0 \;\; 0.14 \;\; 0]^T$. In this example, we have two generators located at bus one and two. At each generator bus, we have a PSS, but we measure only ω_1. Based on this measurement, the S-SISE-LQG regulators for the PSSs at bus 1 and 2 are implemented. For the external load input, we apply a square signal in order to observe the rotor angle swings. The measurement noise is chosen to be $R = 1e-8$ and the process noise is $Q = 3e-3$. The weights used in the design of the LQ controller are $\mathcal{W} = 3 \times \mathcal{B}_{PSS} \times \mathcal{B}_{PSS}^T$ and $\mathcal{V} = R$, then the dynamic response of the state variables a_1 and ω_1 are obtained as in Fig. 12.8. This contrasts with both previously published LQG-PSS design (which assumes a model of the complete system is available), and PSS design using local measurements only (which would require measuring both ω_1 and ω_2). It can be observed that the controller dampens the system response and reduces overshoot and oscillations; also, the low-frequency behaviour of the system is improved.

If the measurement noise is increased, the LQG is unstable which is related to the fact that the model of the known part of the system does not describe

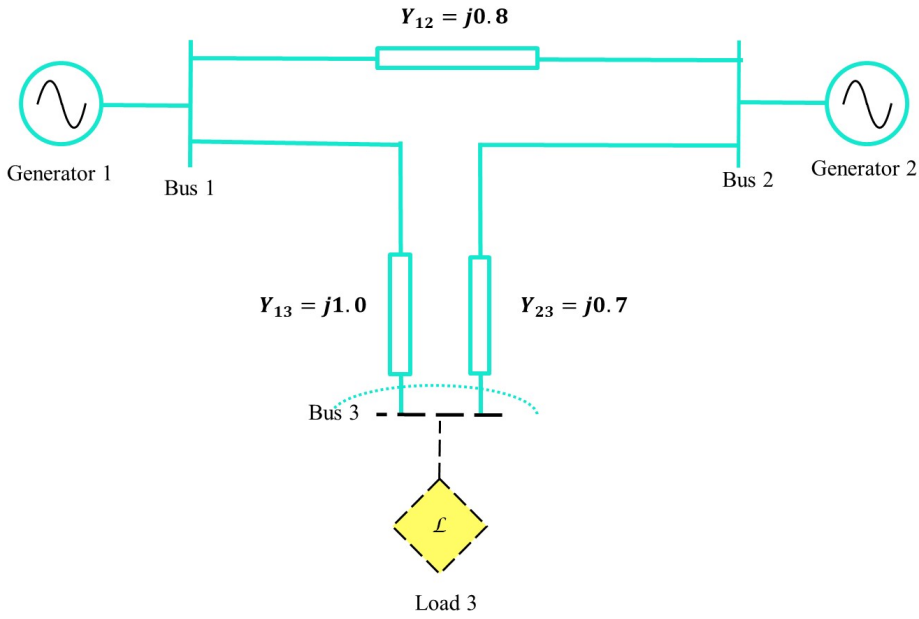


Figure 12.7: Transformer-less dynamic power grid model of IEEE-3-bus performing the first cut. The dashed line represents the virtual cut separating the known and unknown parts of the network.

the total dynamics, and it is illustrated in Fig. 12.9. It is known that LQG lacks robustness to system uncertainty Doyle (1978). One can decrease the controller sensitivity by artificially increasing the LQG weight $\mathcal{W}$; this method is known as LQG/LTR Stein and Athans (1987). However, this can degrade the the performance of the closed-loop system significantly George (2014). As a remedy to the lack of robustness of LQG, Glover-McFarlane loop-shaping method McFarlane and Glover (1992) that uses a combination of loop shaping and robust stabilization is applied.

12.5.3 McFarlane Glover loop shaping method

This method is chosen because it can obtain the closed-loop stability and improved stability margins without significant degradation in the large and small gain regions of the open-loop response. In this method, we obtain the H_∞ normalized coprime factor loop-shaping controller $\mathcal{K}_s$ for a plant $\mathcal{G}$ with pre-compensator and post-compensator weights Υ_1 and Υ_2. In our

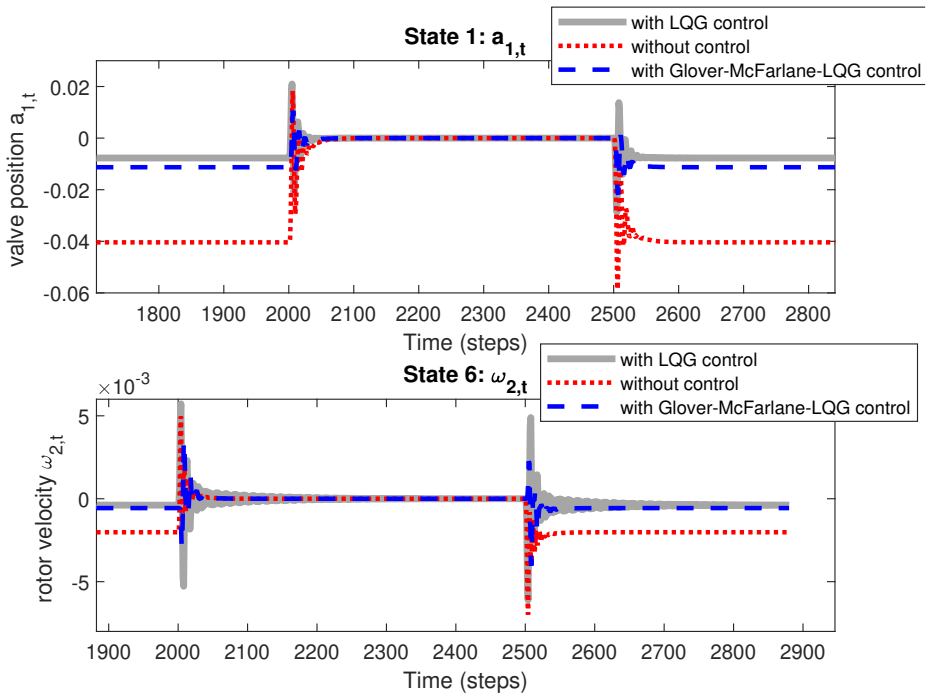


Figure 12.8: The rotor frequency at bus 2 and the valve position at bus 1 for the first cut when no control, the S-SISE-LQG and the S-SISE-LQG-MCFarlane methodologies are applied for the PSS.

example, $\Upsilon_1 = I$ and $\Upsilon_2 = K_{LQG}$. Then we perform a robust coprime factors stabilization to compute $\mathcal{K}_s$ by minimizing:

$$|| \begin{bmatrix} I \\ \mathcal{K}_s \end{bmatrix} (I + \Upsilon_2 \mathcal{G}_{\mathcal{P}} \Upsilon_1 \mathcal{K}_s)[I \quad \Upsilon_2 \mathcal{G}_{\mathcal{P}} \Upsilon_1]||_\infty = \gamma$$

where γ specifies the bound of uncertainty, and $\mathcal{G}_{\mathcal{P}}$ is the transfer function for PKPS. The control procedure as it is shown in Fig. 12.10 consists of following steps:

1. The S-SISE is designed to obtain the state estimates of PKPS.
2. The LQR controller is designed based on SISE estimates.
3. Based on the controller gain obtained from the LQG compensator, a Glover-McFarlane controller is designed.

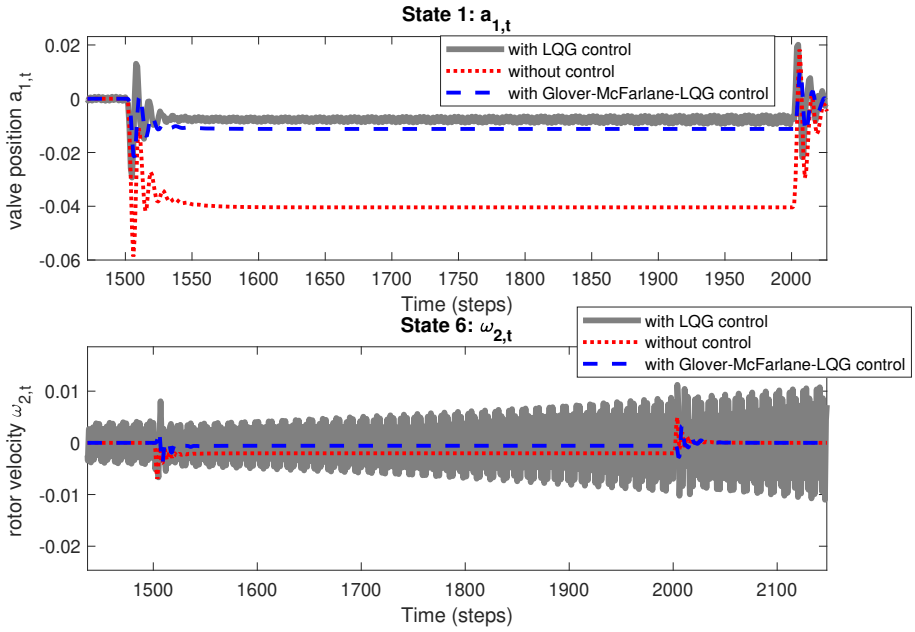


Figure 12.9: The rotor frequency at bus 2 and the valve position at bus 1 and their estimates for the first cut by using the S-SISE-LQG and S-SISE-LQG-MCFarlane based PSS, with $R = 1e-6$.

The transfer functions [1] in Fig. 12.11 are listed bellow:

- The S-SISE-LQG compensator:

$$((I - \mathcal{K}C)\mathcal{A} - \mathcal{B}_{PSS}K_{LQG}, \mathcal{K}, K_{LQG}, 0)$$

.

- The S-SISE-LQG-McFarlane compensator:

$$((I - \mathcal{K}C)\mathcal{A} - \mathcal{B}_{PSS}K_{LQG}, \mathcal{K}, K_{LQG}, 0)$$

$$\times(A_{Mc}, B_{Mc}, C_{Mc}, D_{Mc})$$

where

$$\mathcal{K}_s : (A_{Mc}, B_{Mc}, C_{Mc}, D_{Mc}).$$

[1] where each quadruple (A, B, C, D) shows a dynamic system with system matrix A, input matrix B, measurement matrix C, and feedthrough matrix D

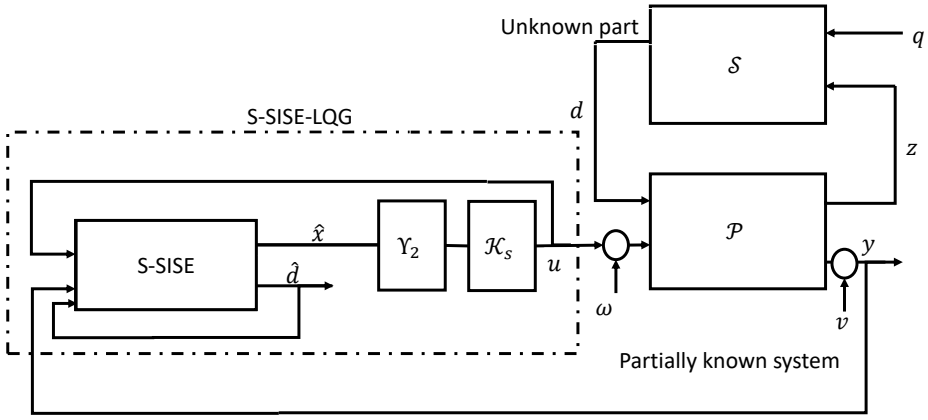


Figure 12.10: Block diagram of the S-SISE-LQG-McFarlane regulator, where Υ_2 is the original LQG-based controller, and $\mathcal{K}_s$ is the McFarlane-Glover robust loop-shaping controller.

- Full state uncontrolled system:

$$(\mathcal{A}, \mathcal{B}_{PSS}, C, 0)$$

To assess robustness, we check the maximum singular value of the sensitivity function at the plant input, while large singular values of the sensitivity function are indicators of robustness problems Skogestad and Postlethwaite (2007). Fig. 12.12 shows the robustness improvement using Glover-McFarlane method, as the peak value of the maximum singular value is significantly reduced.

Conclusion

Based on the frequency response of the closed loop system, one can observe the oscillatory behaviour of the system and the possible frequency that the oscillations are happening. Based on this information and by using robust control schemes, one can stabilize the system by damping the oscillatory swing modes. In addition, we have demonstrated how S-SISE can be used to design PSS, with a reduced number of measurements.

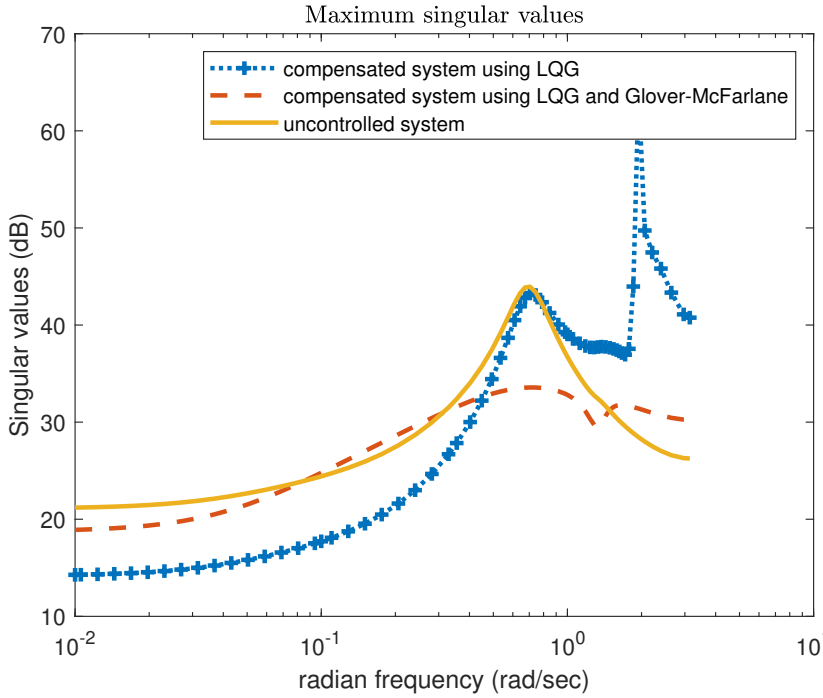


Figure 12.11: Singular values of the open loop transfer function matrix, at the plant's input of the compensated system using LQG, LQG-Glover-MCFarlane and the full-state system of the PKPS (without control).

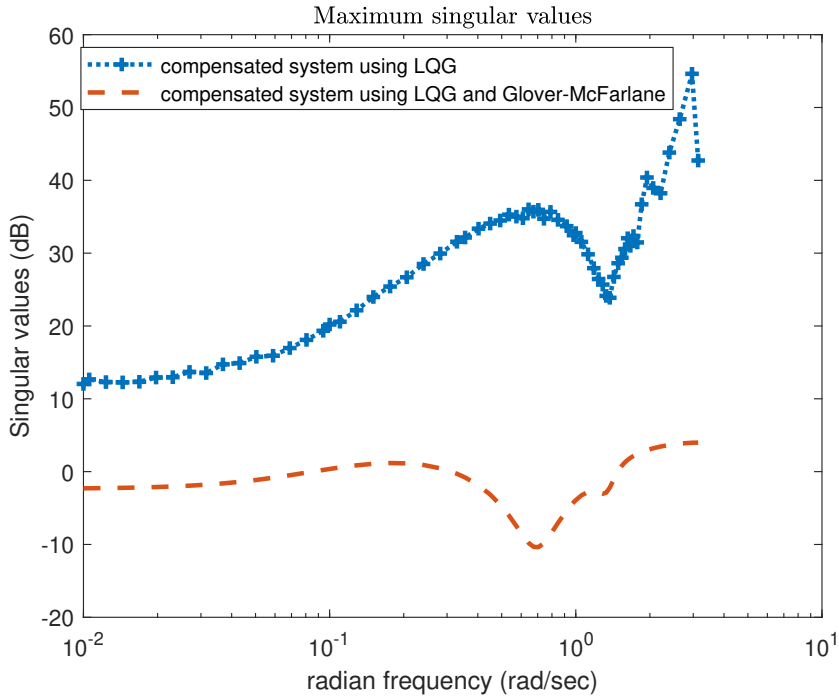


Figure 12.12: The maximum singular value of the sensitivity function at the plant input of the compensated system using LQG, LQG-Glover-McFarlane.

Chapter 13

Conclusion and future works

In this thesis, we clarify some "myths" regarding SISE that have existed for a while, by deconstructing SISE and deriving an equivalent Kalman filtering formulation in which unknown disturbances are modeled as noises with unbounded variance. Then we study the stability properties of (conventional) SISE in detail, and show that the Kalman filtering based version has improved stability properties. This Kalman filtering based version of SISE, called S-SISE, is used to develop an approach to dynamic state estimation for partially known power systems.

For the case where coordination between different parts of a power network is possible, we present a distributed algorithm using diffusion techniques to implement H_∞ and Kalman filtering in a multi agent setting. In order to relax observability conditions in the diffusion filtering algorithms, we use three strategies but the most novel one is using only the locally observable states, which avoids local covariance matrices growing without bounds in unobservable directions.

For a power system containing sensors with different data rates, a covariance intersection (CI) method is applied to perform fusion sensor in power systems measurement problems. We compare this method with that of Bar-Shalom-Campo and demonstrate that CI is better in terms of consistency but not necessarily concerning precision.

Finally, benefits and applications of using the proposed dynamic state estimation methods for the dynamic state monitoring of power systems are investigated. A fault detection methodology, a sensor placement strategy, an estimation approach that can account for the interactions between the

transmission and distribution levels of the power system, and a power system stabilizer have been proposed. A number of power systems examples of increasing complexity illustrate the approaches and their capabilities.

To clarify our approach, we have assumed that the transmission lines reactance are significantly larger than their resistance in Chapters 7 and 11. However, for distribution grids, the resistance can be significantly larger than the inductance. This can raise the question of scalability for using our approaches for distribution grids. In Section 7.5, however, the assumption of ignoring the resistance of the transmission lines has been relaxed, and a more general method has been proposed. Furthermore, using the S-SISE method results in reducing the order of the global state by allowing each PKPS to estimate only a subset of the global state space in addition to the number of unknown inputs. This solution consequently addresses the scalability problem of state estimation when the order of the system increases. This scalability property of using PKPS concept and S-SISE method has been investigated in Chapter 7.

On the basis of the findings presented in this thesis, a number of extensions are described that can be explored in the future. Below, they are addressed concisely:

- The design of PSS can be supported with a more solid theoretical framework: The disturbance model would make better control possible. However, the idea used here is to exploit loop shaping for robust stability. The link is the LQG/LTR to recover the closed loop transfer function of the estimator/controller and then use the obtained identified transfer function to design a robust controller. The theoretical framework with strong mathematical guarantees is needed to evaluate this loop-shaping approach.

- The design of proposed estimators can be supported with real-world experiments: Our proposed methods are based on simplifying assumptions, and we are using 3-bus, 9-bus, 14-bus and 118-bus power systems. Hence it seems vital to design laboratory experiments for more complex interconnected power systems to test and validate the applicability of S-SISE for practical application.

- More complicated and larger test cases for S-SISE based PSS: More complex simulation studies for larger systems should be tried out to validate the performance of LQG loop shaping robust dynamic PSS using S-SISE as estimation method. In addition, the number of PSSs

to use and their location would be relevant issues when dealing with large systems.

- Cyber attack detector: While the S-SISE is robust and recovering n interaction signals only need n measurement, then by using secure measurements cyber attacks can be detected in a power system. However, PMUs are exposed to GPS signal spoofing. Therefore, the secure measurements are likely to be SCADA measurements, with low sampling rates. The problem of cyber attack detection with low sampling rate sensor information needs further investigations.

- The fault detection method and the diffusion strategy need coordination between different subsystems, sending and receiving these coordinating signals need privacy preservation. This can be done by state decomposition and the application of coding strategies to share data securely. Using secure data sharing between different agents in a power network should get more attention in future.

- In modern power systems, dynamic models of ICT systems contain dynamic states and would contribute to the states of the model of the overall system. The concept of PKPS can be used to decouple a power system states from ICT ones. Developing a methodology for decoupling ICT dynamics from power system dynamics can be a topic for future work.

Bibliography

Abooshahab, M., Hovd, M., Bitmead, R., 2019. Disturbance and state estimation in partially known power networks, in: 2019 IEEE Conference on Control Technology and Applications (CCTA), Hong Kong, China. pp. 98–105.

Abooshahab, M., Hovd, M., Brekke, E., Song, X., 2020a. A covariance consistent data fusion method for power networks with multirate sensors, in: 2020 IEEE Conference on Control Technology and Applications (CCTA), Montreal, Canada.

Abooshahab, M.A., Alyaseen, M.M., Bitmead, R.R., Hovd, M., 2021. Simultaneous input and state estimation, singular filtering and stability. Provisionaly accepted in Automatica .

Abooshahab, M.A., Hovd, M., Bitmead, R.R., 2020b. Monitoring disturbances and states in partially known power systems. In preparation for submission to IEEE Transactions on Power Systems .

Abur, A., Exposito, A.G., 2004. Power system state estimation: theory and implementation. CRC press.

Albertini, F., D'Alessandro, D., 1996. Remarks on the observability of nonlinear discrete time systems. Springer US, Boston, MA. pp. 155–162.

Albinali, H.F., Meliopoulos, A.P., 2016. A centralized substation protection scheme that detects hidden failures, in: 2016 IEEE Power and Energy Society General Meeting (PESGM), pp. 1–5.

Aminifar, F., Shahidehpour, M., Fotuhi-Firuzabad, M., Kamalinia, S., 2014. Power system dynamic state estimation with synchronized phasor meas-

urements. IEEE Transactions on Instrumentation and Measurement 63, 352–363.

Anagnostou, G., Pal, B.C., 2018. Derivative-free Kalman filtering based approaches to dynamic state estimation for power systems with unknown inputs. IEEE Transactions on Power Systems 33, 116–130.

Anderson, B.D., Moore, J.B., 1981. Detectability and stabilizability of time-varying discrete-time linear systems. SIAM Journal on Control and Optimization 19, 20–32.

Anderson, B.D., Moore, J.B., 2012. Optimal filtering. Courier Corporation.

Anderson, P.M., Fouad, A.A., 2008. Power system control and stability. John Wiley & Sons.

Andersson, G., Donalek, P., Farmer, R., Hatziargyriou, N., Kamwa, I., Kundur, P., Martins, N., Paserba, J., Pourbeik, P., Sanchez-Gasca, J., et al., 2005. Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance. IEEE Transactions on Power Systems 20, 1922–1928.

Ariff, M.A.M., Pal, B.C., 2016. Adaptive protection and control in the power system for wide-area blackout prevention. IEEE Transactions on Power Delivery 31, 1815–1825.

Axelsson, P., Gustafsson, F., 2015. Discrete-time solutions to the continuous-time differential lyapunov equation with applications to Kalman filtering. IEEE Transactions on Automatic Control 60, 632–643.

B. Widrow, E. Walach, 1995. Adaptive inverse control. Prentice-Hall, Englewood Cliffs, NJ.

Banavar, R.N., 1992. A game theoretic approach to linear dynamic estimation. Ph.D. thesis. University of Texas Austin.

Bar-Shalom, Y.and Li, X.R., Kirubarajan, T., 2004. Estimation with applications to tracking and navigation: theory algorithms and software. John Wiley & Sons.

Bar-Shalom, Y., Campo, L., 1986. The effect of the common process noise on the two-sensor fused-track covariance. IEEE Transactions on Aerospace and Electronic Systems , 803–805.

Baran, M.E., Kelley, A.W., 1994. State estimation for real-time monitoring of distribution systems. IEEE Transactions on Power Systems 9, 1601–1609.

Baran, M.E., Kelley, A.W., 1995. A branch-current-based state estimation method for distribution systems. IEEE Transactions on Power Systems 10, 483–491.

Battistelli, G., Chisci, L., 2014. Kullback–Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability. Automatica 50, 707–718.

Bian, X., Li, X.R., Chen, H., Gan, D., Qiu, J., 2011. Joint estimation of state and parameter with synchrophasors—Part II: Parameter tracking. IEEE Transactions on Power Systems 26, 1209–1220.

Bitmead, R.R., Hovd, M., Abooshahab, M.A., 2019. A Kalman-filtering derivation of simultaneous input and state estimation. Automatica 108, 108478.

Blood, E.A., 2011. From static to dynamic electric power network state estimation: the role of bus component dynamics. Ph.D. thesis. CMU.

Boem, F., Sabattini, L., Secchi, C., 2019. Decentralized state estimation for the control of network systems. Journal of the Franklin Institute 356, 860–882.

Boutayeb, M., Rafaralahy, H., Darouach, M., 1997. Convergence analysis of the extended Kalman filter used as an observer for nonlinear deterministic discrete-time systems. IEEE Transactions on Automatic Control 42, 581–586.

Bretas, A.S., Salim, R.H., 2006. Fault location in unbalanced DG systems using the positive sequence apparent impedance, in: 2006 IEEE/PES Transmission & Distribution Conference and Exposition: Latin America, IEEE. pp. 1–6.

Cao, Y., Ren, W., Meng, Z., 2010. Decentralized finite-time sliding mode estimators and their applications in decentralized finite-time formation tracking. Systems & Control Letters 59, 522–529.

Caro, E., Conejo, A.J., Minguez, R., Zima, M., Andersson, G., 2011. Multiple bad data identification considering measurement dependencies. IEEE Transactions on Power Systems 26, 1953–1961.

Cattivelli, F.S., Lopes, C.G., Sayed, A.H., 2008. Diffusion recursive least-squares for distributed estimation over adaptive networks. IEEE Transactions on Signal Processing 56, 1865–1877.

Cattivelli, F.S., Sayed, A.H., 2009. Diffusion LMS strategies for distributed estimation. IEEE Transactions on Signal Processing 58, 1035–1048.

Cattivelli, F.S., Sayed, A.H., 2010. Diffusion strategies for distributed Kalman filtering and smoothing. IEEE Transactions on Automatic Control 55, 2069–2084.

Cattivelli, F.S., Sayed, A.H., 2010. Distributed nonlinear Kalman filtering with applications to wireless localization, in: 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 3522–3525.

Chen, J., Patton, R.J., 2012. Robust model-based fault diagnosis for dynamic systems. volume 3. Springer Science and Business Media.

Chen, W., Ohnishi, K., Guo, L., 2015. Advances in disturbance/uncertainty estimation and attenuation [guest editors' introduction]. IEEE Transactions on Industrial Electronics 62, 5758–5762.

Chen, W.H., Yang, J., Guo, L., Li, S., 2015. Disturbance-observer-based control and related methods—an overview. IEEE Transactions on Industrial Electronics 63, 1083–1095.

Chow, J.H., 2013. Power system coherency and model reduction. volume 84. Springer.

Chow, J.H., Boukarim, G.E., Murdoch, A., 2004. Power system stabilizers as undergraduate control design projects. IEEE Transactions on Power Systems 19, 144–151.

Ciccarella, G., Dalla Mora, M., Germani, A., 1993. Observers for discrete-time nonlinear systems. Systems & Control Letters 20, 373 – 382.

Conte, G., Moog, C.H., Perdon, A.M., 2007. Algebraic methods for nonlinear control systems. Springer Science & Business Media.

Crow, M.L., 2015. Computational methods for electric power systems. Crc Press.

Cui, Y., Kavasseri, R.G., Brahma, S.M., 2017. Dynamic state estimation assisted out-of-step detection for generators using angular difference. IEEE Transactions on Power Delivery 32, 1441–1449.

Cupelli, M., Mirz, M., Monti, A., 2015. A comparison of backstepping and LQG control for stabilizing mvdc microgrids with constant power loads, in: 2015 IEEE Eindhoven PowerTech, IEEE. pp. 1–6.

Dehghanpour, K., Wang, Z., Wang, J., Yuan, Y., Bu, F., 2019. A survey on state estimation techniques and challenges in smart distribution systems. IEEE Transactions on Smart Grid 10, 2312–2322.

Delavari, A., Kamwa, I., Brunelle, P., 2018. Simscape power systems benchmarks for education and research in power grid dynamics and control, in: 2018 IEEE Canadian Conference on Electrical Computer Engineering (CCECE), pp. 1–5.

Deng, Z., et al., 2012. Sequential covariance intersection fusion Kalman filter. Information Sciences 189, 293–309.

Desoer, C.A., Kuh, E.S., 1966. Basic circuit theory. McGraw-Hill.

Dib, W., Ortega, R., Barabanov, A., Lamnabhi-Lagarrigue, F., 2009. A globally convergent controller for multi-machine power systems using structure-preserving models. IEEE Transactions on Automatic Control 54, 2179–2185.

Ding, D., Wang, Z., Dong, H., Shu, H., 2012. Distributed H_∞ state estimation with stochastic parameters and nonlinearities through sensor networks: the finite-horizon case. Automatica 48, 1575–1585.

Ding, L., Guo, G., 2015. Distributed event-triggered H-infinity consensus filtering in sensor networks. Signal Processing 108, 365–375.

Dorfler, F., Bullo, F., 2013. Kron reduction of graphs with applications to electrical networks. IEEE Transactions on Circuits and Systems I: Regular Papers 60, 150–163.

Doyle, J.C., 1978. Guaranteed margins for LQG regulators. IEEE Transactions on Automatic Control 23, 756–757.

Einicke, G.A., White, L.B., 1999. Robust extended Kalman filtering. IEEE Transactions on Signal Processing 47, 2596–2599.

Eto, J.H., Stewart, E., Smith, T., Buckner, M., Kirkham, H., Tuffner, F., Schoenwald, D., 2015. Scoping study on research and development priorities for distribution-system phasor measurement units. Lawrence Berkeley Nat. Lab., Berkeley, CA, USA, Tech. Rep .

Falcao, D.M., Wu, F.F., Murphy, L., 1995. Parallel and distributed state estimation. IEEE Transactions on Power Systems 10, 724–730.

Fan, R., Meliopoulos, A.P.S., Cokkinides, G.J., Sun, L., Yu Liu, 2015. Dynamic state estimation-based protection of power transformers, in: 2015 IEEE Power Energy Society General Meeting, pp. 1–5.

Fang, H., De Callafon, R.A., 2012. On the asymptotic stability of minimum-variance unbiased input and state estimation. Automatica 48, 3183–3186.

Fang, H., De Callafon, R.A., Cortés, J., 2013. Simultaneous input and state estimation for nonlinear systems with applications to flow field estimation. Automatica 49, 2805–2812.

Fang, H., Shi, Y., Yi, J., 2011. On stable simultaneous input and state estimation for discrete-time linear systems. International Journal of Adaptive Control and Signal Processing 25, 671–686.

Farantatos, E., Huang, R., Cokkinides, G.J., Meliopoulos, A.P., 2016. A predictive generator out-of-step protection and transient stability monitoring scheme enabled by a distributed dynamic state estimator. IEEE Transactions on Power Delivery 31, 1826–1835.

Filomena, A.D., Resener, M., Salim, R.H., Bretas, A.S., 2009. Fault location for underground distribution feeders: An extended impedance-based formulation with capacitive current compensation. International Journal of Electrical Power & Energy Systems 31, 489–496.

Florescu, A., Bratcu, A.I., Munteanu, I., Bacha, S., 2012. Energy management system within electric vehicles using ultracapacitors: An LQG-optimal-control-based solution. IFAC Proceedings Volumes 45, 229–234.

Francis, B., 1987. A Course in H_∞ Control Theory. Springer-Verlag.

Friedland, B., 1969. Treatment of bias in recursive filtering. IEEE Transactions on Automatic Control 14, 359–367.

Gao, C., Zhao, G., Lu, J., Pan, S., 2018. Decentralized state estimation for networked spatial-navigation systems with mixed time-delays and quantized complementary measurements: The moving horizon case. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering 232, 2160–2177.

Ge, Q., Xu, D., Wen, C., 2014. Cubature information filters with correlated noises and their applications in decentralized fusion. Signal Processing 94, 434–444.

George, J., 2014. On adaptive loop transfer recovery using Kalman filter-based disturbance accommodating control. IET Control Theory Applications 8, 267–276.

Gerstner, A., L'eplattenier, R., Bawart, M., 2013. New enhancements for cable fault location in complex medium voltage distribution networks, in: 22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013), pp. 1–4.

Ghahremani, E., Kamwa, I., 2011a. Dynamic state estimation in power system by applying the extended Kalman filter with unknown inputs to phasor measurements. IEEE Transactions on Power Systems 26, 2556–2566.

Ghahremani, E., Kamwa, I., 2011b. Online state estimation of a synchronous generator using unscented Kalman filter from phasor measurements units. IEEE Transactions on Energy Conversion 26, 1099–1108.

Ghahremani, E., Kamwa, I., 2011c. Simultaneous state and input estimation of a synchronous machine using the extended Kalman filter with unknown inputs, in: Electric Machines & Drives Conference (IEMDC), 2011 IEEE International, IEEE. pp. 1468–1473.

Ghahremani, E., Kamwa, I., 2016. Local and wide-area PMU-based decentralized dynamic state estimation in multi-machine power systems. IEEE Transactions on Power Systems 31, 547–562.

Ghosal, M., 2018. Diagnosis of anomaly in the dynamic state estimator of a power system using system decomposition, in: 2018 European Control Conference (ECC), pp. 1–6.

Ghosal, M., Rao, V., 2019. Fusion of multirate measurements for nonlinear dynamic state estimation of the power systems. IEEE Transactions on Smart Grid 10, 216–226.

Ghosal, M., et al., 2017. Enhancing Smartgrid Situational Awareness and Resiliency to Data Attack using Dynamic State Estimation. Ph.D. thesis. Texas Tech University.

Gillijns, S., De Moor, B., 2007a. Unbiased minimum-variance input and state estimation for linear discrete-time systems. Automatica 43, 111–116.

Gillijns, S., De Moor, B., 2007b. Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough. Automatica 43, 934–937.

Girgis, A.A., Fallon, C.M., Lubkeman, D.L., 1993. A fault location technique for rural distribution feeders. IEEE Transactions on Industry Applications 29, 1170–1175.

Glick, T.F., Livesey, S., Wallis, F., 2014. Medieval science, technology, and medicine: an encyclopedia. volume 11. Routledge.

Glover, J., 1969. The linear estimation of completely unknown signals. IEEE Transactions on Automatic Control 14, 766–767.

Gomez-Exposito, A., Abur, A., 2004. Power system state estimation: theory and implementation. CRC press.

Gomez-Exposito, A., Abur, A., de la Villa Jaen, A., Gomez-Quiles, C., 2011. A multilevel state estimation paradigm for smart grids. Proceedings of the IEEE 99, 952–976.

Green, M., 1988. On inner-outer factorization. Systems & control letters 11, 93–97.

Grimble, M.J., El Sayed, A., 1990. Solution of the H-infinity sub-optimal linear filtering problem for discrete-time systems. IEEE Transactions on Acoustics, Speech, and Signal Processing 38, 1092–1104.

Grossman, W.D., 1999. Observers for discrete-time nonlinear systems. Ph.D. thesis. New Jersey Institute of Technology.

Guo, B.Z., Zhao, Z.L., 2016. Active disturbance rejection control for nonlinear systems: An introduction. John Wiley & Sons.

Guo, Y., Huang, B., 2015. State estimation incorporating infrequent, delayed and integral measurements. Automatica 58, 32–38.

Hashemipour, H.R., Roy, S., Laub, A.J., 1988. Decentralized structures for parallel Kalman filtering. IEEE Transactions on Automatic Control 33, 88–94.

Havangi, R., 2015. Unscented H-infinity filtering based simultaneous localization and mapping with evolutionary resampling. Journal of the Franklin Institute 352, 4801–4825.

He, J., et al., 2019. Research on multi-axis servo synergic control system based on sliding mode variable structure, in: IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society, IEEE. pp. 6908–6915.

Hebrail, G., 2012. Individual household electric power consumption data set. http://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption.

Hermann, R., Krener, A., 1977. Nonlinear controllability and observability. IEEE Transactions on Automatic Control 22, 728–740.

Hertem, D.V., Verboomen, J., Purchala, K., Belmans, R., Kling, W.L., 2006. Usefulness of dc power flow for active power flow analysis with flow controlling devices, in: The 8th IEE International Conference on AC and DC Power Transmission, pp. 58–62.

Hill, D.J., 1993. Nonlinear dynamic load models with recovery for voltage stability studies. IEEE Transactions on Power Systems 8, 166–176.

Hiskens, I., Milanovic, J., 1995. Load modelling in studies of power system damping. IEEE Transactions on Power Systems 10, 1781–1788.

Horn, R.A., Johnson, C.R., 2012. Matrix analysis. Cambridge university press.

Hou, M., Muller, P., 1992. Design of observers for linear systems with unknown inputs. IEEE Transactions on Automatic Control 37, 871–875.

Hou, M., Patton, R., 1998. Optimal filtering for systems with unknown inputs. IEEE Transactions on Automatic Control 43, 445–449.

Hu, J., Yang, C., 2011. Second-order extended H_∞ filter for nonlinear discrete-time systems using quadratic error matrix approximation. IEEE Transactions on Signal Processing 59, 3110–3119.

Huang, G.P., Mourikis, A.I., Roumeliotis, S.I., 2008. Analysis and improvement of the consistency of extended Kalman filter based SLAM, in: 2008 IEEE International Conference on Robotics and Automation, IEEE. pp. 473–479.

Huang, Z., Du, P., Kosterev, D., Yang, B., 2009. Application of extended Kalman filter techniques for dynamic model parameter calibration, in: Power & Energy Society General Meeting, 2009. PES'09. IEEE, IEEE. pp. 1–8.

Huang, Z., Du, P., Kosterev, D., Yang, S., 2013. Generator dynamic model validation and parameter calibration using phasor measurements at the point of connection. IEEE Transactions on Power Systems 28, 1939–1949.

Huang, Z., Schneider, K., Nieplocha, J., 2007. Feasibility studies of applying Kalman filter techniques to power system dynamic state estimation, in: Power Engineering Conference, 2007. IPEC 2007. International, IEEE. pp. 376–382.

Ionescu, V., Oara, C., 1996. Spectral and inner-outer factorizations for discrete-time systems. IEEE Transactions on Automatic Control 41, 1840–1845.

Janssen, P., 2014. Monitoring, protection and fault location in power distribution networks using system-wide measurements. Ph.D. thesis. Ph. D. dissertation, Ecole Polytechnique de Bruxelles, 2013-2014.

Jazwinski, A., 1970. Stochastic processes and filtering theory. Number 64 in Mathematics in science and engineering, Acad. Press, New York, NY.

Julier, S., Uhlmann, J., Durrant-Whyte, H.F., 2000. A new method for the nonlinear transformation of means and covariances in filters and estimators. IEEE Transactions on Automatic Control 45, 477–482.

Julier, S., Uhlmann, J.K., 2009. General decentralized data fusion with covariance intersection. Handbook of multisensor data fusion: theory and practice , 319–344.

Julier, S.J., Uhlmann, J.K., 1996. A general method for approximating nonlinear transformations of probability distributions. Technical Report. Technical report, Robotics Research Group, Department of Engineering Science, University of Oxford, UK.

Kailath, T., Sayed, A.H., Hassibi, B., 2000. Linear Estimation. Prentice Hall.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Journal of Basic Engineering 82, 35–45.

Karimipour, H., Dinavahi, V., 2015. Extended Kalman filter-based parallel dynamic state estimation. IEEE Transactions on Smart Grid 6, 1539–1549.

Karvonen, T., et al., 2014. Stability of linear and non-linear Kalman filters. Master's thesis. University of Helsinki.

Kawano, Y., u. Kotta, 2015. On integrability of observable space for discrete-time polynomial control systems. IEEE Transactions on Automatic Control 60, 1987–1991.

Kerwin, W.S., Prince, J.L., 2000. On the optimality of recursive unbiased state estimation with unknown inputs. Automatica 36, 1381–1383.

Kezunovic, M., 2011. Smart fault location for smart grids. IEEE Transactions on Smart Grid 2, 11–22.

Khan, U.A., Ilić, M.D., Moura, J.M., 2008. Cooperation for aggregating complex electric power networks to ensure system observability, in: Infrastructure Systems and Services: Building Networks for a Brighter Future (INFRA), IEEE. pp. 1–6.

Kim, H., Guo, P., Zhu, M., Liu, P., 2020. Simultaneous input and state estimation for stochastic nonlinear systems with additive unknown inputs. Automatica 111, 108588.

Kitanidis, P.K., 1987. Unbiased minimum-variance linear state estimation. Automatica 23, 775–778.

Korres, G.N., 2011. A distributed multiarea state estimation. IEEE Transactions on Power Systems 26, 73–84.

Kotta, U., 2005. Decomposition of discrete-time nonlinear control systems, in: Proceedings of the Estonian Academy of Sciences: Physics, Mathematics, Estonian Academy Publishers; 1999. p. 154.

Kulic, D., Croft, E.A., 2007. Affective state estimation for human–robot interaction. IEEE Transactions on Robotics 23, 991–1000.

Kumar, P. R.and Varaiya, P., 2015. Stochastic systems: Estimation, identification, and adaptive control. volume 75. SIAM.

Kundur, P., Balu, N.J., Lauby, M.G., 1994. Power system stability and control. volume 7. McGraw-hill New York.

Kwon, C., Liu, W., Hwang, I., 2013. Security analysis for cyber-physical systems against stealthy deception attacks, in: 2013 American control conference, IEEE. pp. 3344–3349.

L. Chen, Mercorelli, P., Liu, S., 2005. A Kalman estimator for detecting repetitive disturbances, in: Proceedings of the 2005, American Control Conference, 2005., pp. 1631–1636.

Labarre, D., Grivel, E., Najim, M., Christov, N., 2007. Dual H_∞ algorithms for signal processing- application to speech enhancement. IEEE Transactions on Signal Processing 55, 5195–5208.

Lee, D.J., 2008. Nonlinear estimation and multiple sensor fusion using unscented information filtering. IEEE Signal Processing Letters 15, 861–864.

Lefebvre, T., Bruyninckx, H., De Schuller, J., 2002. Comment on "a new method for the nonlinear transformation of means and covariances in filters and estimators" [with authors' reply]. IEEE Transactions on Automatic Control 47, 1406–1409.

Lehtonen, M., Harmand, Y., Huber, A., de Vylder, J., et al., 1995. Fault management in electrical distribution systems. Number 1678 in VTT Tiedotteita - Meddelanden - Research Notes, VTT Technical Research Centre of Finland, Finland.

Lewis, F.L., Zhang, H., Hengster-Movric, K., Das, A., 2014. Cooperative control of multi-agent systems: Optimal design and adaptive control. Springer-Verlag.

Li, F., Qiao, W., Sun, H., Wan, H., Wang, J., Xia, Y., Xu, Z., Zhang, P., 2010. Smart transmission grid: Vision and framework. IEEE Transactions on Smart Grid 1, 168–177.

Li, L., Xia, Y., 2012. Stochastic stability of the unscented Kalman filter with intermittent observations. Automatica 48, 978–981.

Li, Q., Cui, T., Weng, Y., Negi, R., Franchetti, F., Ilic, M.D., 2012. An information-theoretic approach to PMU placement in electric power systems. IEEE Transactions on Smart Grid 4, 446–456.

Li, W., Jia, Y., 2010. H-infinity filtering for a class of nonlinear discrete-time systems based on unscented transform. Signal Processing 90, 3301–3307.

Li, W., Wei, G., Han, F., Liu, Y., 2016. Weighted average consensus-based unscented Kalman filtering. IEEE Transactions on Cybernetics 46, 558–567.

Li, W., Wei, G., Ho, D.W., Ding, D., 2018. A weightedly uniform detectability for sensor networks. IEEE Transactions on Neural Networks and Learning Systems 29, 5790–5796.

Lin, C., Wu, W., Guo, Y., 2019. Decentralized robust state estimation of active distribution grids incorporating microgrids based on PMU measurements. IEEE Transactions on Smart Grid 11, 810–820.

Lin, P.C., Komsuoglu, H., Koditschek, D.E., 2006. Sensor data fusion for body state estimation in a hexapod robot with dynamical gaits. IEEE Transactions on Robotics 22, 932–943.

Liu, S., Chepuri, S.P., Fardad, M., Maşazade, E., Leus, G., Varshney, P.K., 2016. Sensor selection for estimation with correlated measurement noise. IEEE Transactions on Signal Processing 64, 3509–3522.

Liu, Y., Meliopoulos, A.P.S., Fan, R., Sun, L., Tan, Z., 2017. Dynamic state estimation based protection on series compensated transmission lines. IEEE Transactions on Power Delivery 32, 2199–2209.

Loo, J.Y., Tan, C.P., Nurzaman, S.G., 2019. H-infinity based extended Kalman filter for state estimation in highly non-linear soft robotic system, in: 2019 American Control Conference (ACC), pp. 5154–5160.

Lu, T.T., Shiou, S.H., 2002. Inverses of 2×2 block matrices. Computers & Mathematics with Applications 43, 119–129.

Machowski, J., Lubosny, Z., Bialek, J.W., Bumby, J.R., 2020. Power system dynamics: stability and control. John Wiley & Sons.

Maciejowski, J., 1985. Asymptotic recovery for discrete-time systems. IEEE Transactions on Automatic Control 30, 602–605.

Manousakis, N.M., Korres, G.N., Georgilakis, P.S., 2012. Taxonomy of PMU placement methodologies. IEEE Transactions on Power Systems 27, 1070–1077.

Marro, G., Zattoni, E., 2010. Unknown-state, unknown-input reconstruction in discrete-time nonminimum-phase systems: Geometric methods. Automatica 46, 815–822.

Maybeck, P.S., 1982. Stochastic models, estimation, and control. volume 3. Academic press.

McFarlane, D., Glover, K., 1992. A loop-shaping design procedure using H_∞ synthesis. IEEE Transactions on Automatic Control 37, 759–769.

Meliopoulos, A.P.S., Cokkinides, G.J., Myrda, P., Liu, Y., Fan, R., Sun, L., Huang, R., Tan, Z., 2017. Dynamic state estimation-based protection: Status and promise. IEEE Transactions on Power Delivery 32, 320–330.

Mendel, J., 1977. White-noise estimators for seismic data processing in oil exploration. IEEE Transactions on Automatic Control 22, 694–706.

Miao, F., Zhu, Q., Pajic, M., Pappas, G.J., 2014. Coding sensor outputs for injection attacks detection, in: 53rd IEEE Conference on Decision and Control, pp. 5776–5781.

Mitchell, H.B., 2007. Multi-sensor data fusion: an introduction. Springer Science and Business Media.

Modir, H., Schlueter, R.A., 1981. A dynamic state estimator for dynamic security assessment. IEEE Transactions on Power Apparatus and Systems PAS-100, 4644–4652.

Moylan, P., 1977. Stable inversion of linear systems. IEEE Transactions on Automatic Control 22, 74–78.

NERC, 2016. Reliability guideline-power plant dynamic model verification using PMUs. Technical Report. North American Electric Reliability Corporation Std.

Nijmeijer, H., Van der Schaft, A., 1990. Nonlinear dynamical control systems. volume 175. Springer.

Nordman, M.M., Lehtonen, M., 2005. Distributed agent-based state estimation for electrical distribution networks. IEEE Transactions on Power Systems 20, 652–658.

Olfati-Saber, R., 2007. Distributed Kalman filtering for sensor networks. In Proceedings of IEEE Conference on Decision and Control , 5492–5498.

Olfati-Saber, R., 2009. Kalman-consensus filter: Optimality, stability, and performance. In Proceedings of IEEE Conference on Decision and Control , 7036–7042.

Olfati-Saber, R., Shamma, J.S., 2005. Consensus filters for sensor networks and distributed sensor fusion. In Proceedings of IEEE Conference on Decision and Control .

Pal, B., Chaudhuri, B., 2006. Robust control in power systems. Springer Science & Business Media.

Park, S.H., Kim, P.S., Kwon, O.K., Kwon, W.H., 2000. Estimation and detection of unknown inputs using optimal fir filter. Automatica 36, 1481–1488.

Phadke, A., Thorp, J., 2006. History and applications of phasor measurements, in: 2006 IEEE PES Power Systems Conference and Exposition, IEEE. pp. 331–335.

Plett, G.L., 2004. Extended Kalman filtering for battery management systems of lipb-based hev battery packs: Part 3. state and parameter estimation. Journal of Power sources 134, 277–292.

Priel, B., Shaked, U., 1986. Sequential decomposition of the partially singular discrete time filtering problem, in: 1986 25th IEEE Conference on Decision and Control, pp. 1048–1053.

Qi, J., Sun, K., Wang, J., Liu, H., 2018. Dynamic state estimation for multi-machine power system by unscented Kalman filter with enhanced numerical stability. IEEE Transactions on Smart Grid 9, 1184–1196.

Qing, X., Karimi, H.R., Niu, Y., Wang, X., 2015. Decentralized unscented Kalman filter based on a consensus algorithm for multi-area dynamic state estimation in power systems. International Journal of Electrical Power & Energy Systems 65, 26–33.

Rego, F.F., Pascoal, A.M., Aguiar, A.P., Jones, C.N., 2019. Distributed state estimation for discrete-time linear time invariant systems: A survey. Annual Reviews in Control .

Reif, K., Gunther, S., Yaz, E., Unbehauen, R., 1999. Stochastic stability of the discrete-time extended Kalman filter. IEEE Transactions on Automatic Control 44, 714–728.

Ritter, B., Mora, E., Schild, A., Doekemeijer, B., Konigorski, U., 2019. Adaptive master-slave cubature Kalman filters subject to state inequality constraints for wind turbine state estimation, in: 2019 American Control Conference (ACC), IEEE. pp. 3482–3487.

Roshany-Yamchi, S., Cychowski, M., Negenborn, R.R., De Schutter, B., Delaney, K., Connell, J., 2013. Kalman filter-based distributed predictive control of large-scale multi-rate systems: Application to power networks. IEEE Transactions on Control Systems Technology 21, 27–39.

Roshany-Yamchi, S., et al., 2011. Kalman filter-based distributed predictive control of large-scale multi-rate systems: Application to power networks. IEEE Transactions on Control Systems Technology 21, 27–39.

Rugh, W.J., 1996. Linear systems theory. volume 2. Prentice Hall Upper Saddle River, NJ.

Sadamoto, T., Chakrabortty, A., Ishizaki, T., Imura, J., 2019. Dynamic modeling, stability, and control of power systems with distributed energy resources: Handling faults using two control methods in tandem. IEEE Control Systems Magazine 39, 34–65.

Sagosen, Ø., 2013. Analysis of large scale integration of electric vehicles in Nord-Trøndelag. Master's thesis. NTNU.

Saha, M.M., Izykowski, J.J., Rosolowski, E., 2009. Fault location on power networks. Springer Science & Business Media.

Sain, M., Massey, J., 1969. Invertibility of linear time-invariant dynamical systems. IEEE Transactions on Automatic Control 14, 141–149.

Salahshoor, K., Mosallaei, M., Bayat, M., 2008. Centralized and decentralized process and sensor fault monitoring using data fusion based on adaptive extended Kalman filter algorithm. Measurement 41, 1059–1076.

Sanyal, P., Shen, C., 1974. Bayes' decision rule for rapid detection and adaptive estimation scheme with space applications. IEEE Transactions on Automatic Control 19, 228–231.

Sauer, P.W., Pai, M.A., Chow, J.H., 2017. Power system dynamics and stability: with synchrophasor measurement and power system toolbox. John Wiley and Sons.

Sayed, A.H., 2014a. Adaptive networks. Proceedings of the IEEE 102, 460–497.

Sayed, A.H., 2014b. Diffusion adaptation over networks, in: Academic Press Library in Signal Processing. Elsevier. volume 3, pp. 323–453.

Sayed, A.H., Lopes, C.G., 2007. Adaptive processing over distributed networks. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences E90-A, 1504–1510.

Schweppe, F.C., Wildes, J., 1970. Power system static-state estimation, part i: Exact model. IEEE Transactions on Power Apparatus and Systems PAS-89, 120–125.

Seo, J., Kim, T., Park, J., Moon, S., 1996. An LQG based pss design for controlling the ssr in power systems with series-compensated lines. IEEE transactions on energy conversion 11, 423–428.

Shaked, U., 1985. Explicit solution to the singular discrete-time stationary linear filtering problem. IEEE Transactions on Automatic Control 30, 34–47.

Shaked, U., Soroka, E., 1987. A simple solution to the singular linear minimum-variance estimation problem. IEEE Transactions on Automatic Control 32, 81–84.

Shamaiah, M., Banerjee, S., Vikalo, H., 2010. Greedy sensor selection: Leveraging submodularity, in: 49th IEEE Conference on Decision and Control (CDC), pp. 2572–2577.

Shamma, J.S., 2007. Cooperative Control of Distributed Multi-Agent Systems. John Wiley & Sons.

Shivakumar, N.R., Jain, A., 2008. A review of power system dynamic state estimation techniques, in: 2008 Joint International Conference on Power System Technology and IEEE Power India Conference, pp. 1–6.

Simon, D., 2006. Optimal state estimation: Kalman, H infinity, and nonlinear approaches. John Wiley & Sons.

Singh, A.K., Pal, B., 2018. Dynamic estimation and control of power systems. Academic Press.

Singh, A.K., Pal, B.C., 2016. Decentralized control of oscillatory dynamics in power systems using an extended LQR. IEEE Transactions on Power Systems 31, 1715–1728.

Skogestad, S., Postlethwaite, I., 2007. Multivariable feedback control: analysis and design. John Wiley & Sons, New York.

Speyer, J., 1979. Computation and transmission requirements for a decentralized linear-quadratic-Gaussian control problem. IEEE Transactions on Automatic Control 24, 266–269.

Stein, G., Athans, M., 1987. The LQG/LTR procedure for multivariable feedback control design. IEEE Transactions on Automatic Control 32, 105–114.

Sun, Y., Fu, M., Wang, B., Zhang, H., Marelli, D., 2016. Dynamic state estimation for power networks using distributed map technique. Automatica 73, 27–37.

Sundaram, S., 2020. Inversion of linear systems. https://engineering.purdue.edu/~sundara2/talks/invert.pdf. [accessed June 2020].

Talebi, S.P., Werner, S., 2018. Distributed Kalman filtering: Consensus, diffusion, and mixed, in: 2018 IEEE Conference on Control Technology and Applications (CCTA), pp. 1126–1132.

Talebi, S.P., Werner, S., 2019. Distributed Kalman filtering and control through embedded average consensus information fusion. IEEE Transactions on Automatic Control 64, 4396–4403.

Tebianian, H., Jeyasurya, B., 2015. Dynamic state estimation in power systems: Modeling, and challenges. Electric Power Systems Research 121, 109–114.

Theodor, Y., Shaked, U., 1996. Robust discrete-time minimum-variance filtering. IEEE Transactions on Signal Processing 44, 181–189.

Tran, V., Zhang, H., 2018. Optimal PMU placement using modified Greedy Algorithm. Journal of Control, Automation and Electrical Systems 29, 99–109.

Tsolas, N., Arapostathis, A., Varaiya, P., 1985. A structure preserving energy function for power system transient stability analysis. IEEE Transactions on Circuits and Systems 32, 1041–1049.

Tzyh-Jong Tarn, Rasis, Y., 1976. Observers for nonlinear stochastic systems. IEEE Transactions on Automatic Control 21, 441–448.

Ugrinovskii, V., 2013. Distributed robust estimation over randomly switching networks using H-infinity consensus. Automatica 49, 160–168.

Ule, L., 1955. Weighted least-squares smoothing filters. IRE Transactions on Circuit Theory 2, 197–203.

Van Cutsem, T., Ribbens-Pavella, M., 1983. Critical survey of hierarchical methods for state estimation of electric power systems. IEEE Transactions on Power Apparatus and Systems , 3415–3424.

Venkata, S.S.M., Wilson, D., Ren, J., Miller, M., 2013. Advanced and adaptive protection for active distribution grid, in: 22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013), pp. 1–4.

Vercauteren, T., Wang, X., 2005. Decentralized sigma-point information filters for target tracking in collaborative sensor networks. IEEE Transactions on Signal Processing 53, 2997–3009.

Vittal, V., Ma, F., Hou, G., McCalley, J., Tang, L., 2011. Next generation on-line dynamic security assessment. Power Systems Engineering Research Center., Arizona, Rep. PSERC Publication , 11–09.

Wang, S., Gao, W., Meliopoulos, A.P.S., 2012. An alternative method for power system dynamic state estimation based on unscented transform. IEEE Transactions on Power Systems 27, 942–950.

Wang, S., Ren, W., 2017. On the convergence conditions of distributed dynamic state estimation using sensor networks: A unified framework. IEEE Transactions on Control Systems Technology .

Wang, Y., Sun, Y., Dinavahi, V., Wang, K., Nan, D., 2019. Robust dynamic state estimation of power systems with model uncertainties based on adaptive unscented H-infinity filter. IET Generation, Transmission & Distribution 13, 2455–2463.

Wang, Z., Shen, B., Liu, X., 2012. H-infinity filtering with randomly occurring sensor saturations and missing measurements. Automatica 48, 556–562.

Wei, G., Li, W., Ding, D., Liu, Y., 2018. Stability analysis of covariance intersection-based Kalman consensus filtering for time-varying systems. IEEE Transactions on Systems, Man, and Cybernetics: Systems , 1–12.

Xiao, L., Boyd, S., 2004. Fast linear iterations for distributed averaging. Systems & Control Letters 53, 65–78.

Xiao, L., Boyd, S., Lall, S., 2005. A scheme for robust distributed sensor fusion based on average consensus, in: IPSN 2005. Fourth International Symposium on Information Processing in Sensor Networks, 2005., IEEE. pp. 63–70.

Xie, B., et al., 2017. Distributed quasi-dynamic state estimation with both gps-synchronized and non-synchronized data, in: 2017 North American Power Symposium (NAPS), IEEE. pp. 1–6.

Xie, L., Choi, D., Kar, S., Poor, H.V., 2012. Fully distributed state estimation for wide-area monitoring systems. IEEE Transactions on Smart Grid 3, 1154–1169.

Xiong, K., Zhang, H., Chan, C., 2006. Performance evaluation of ukf-based nonlinear filtering. Automatica 42, 261–270.

Yan, L., Zhou, D., Fu, M., Xia, Y., 2010. State estimation for asynchronous multirate multisensor dynamic systems with missing measurements. IET Signal Processing 4, 728–739.

Yang, C., Wu, J., Ren, X., Yang, W., Shi, H., Shi, L., 2015. Deterministic sensor selection for centralized state estimation under limited communication resource. IEEE Transactions on Signal Processing 63, 2336–2348.

Yang, C., Wu, J., Zhang, W., Shi, L., 2013. Schedule communication for decentralized state estimation. IEEE Transactions on Signal Processing 61, 2525–2535.

Yang, Q., Bi, T., Wu, J., 2007. WAMS implementation in China and the challenges for bulk power system protection, in: 2007 IEEE Power Engineering Society General Meeting, pp. 1–6.

Yang, T.C., Cimen, H., 1996. Applying structured singular values and a new LQR design to robust decentralized power system load frequency control, in: Proceedings of the IEEE International Conference on Industrial Technology, pp. 880–884.

Yin, X., Li, Z., Zhang, L., Han, M., 2018. Distributed state estimation of sensor-network systems subject to markovian channel switching with application to a chemical process. IEEE Transactions on Systems, Man, and Cybernetics: Systems 48, 864–874.

Yong, S.Z., Zhu, M., Frazzoli, E., 2016. A unified filter for simultaneous input and state estimation of linear discrete-time stochastic systems. Automatica 63, 321–329.

Zhang, H., Ayoub, R., Sundaram, S., 2017. Sensor selection for Kalman filtering of linear dynamical systems: Complexity, limitations and greedy algorithms. Automatica 78, 202 – 210.

Zhao, J., 2018. A Robust Dynamic State and Parameter Estimation Framework for Smart Grid Monitoring and Control. Ph.D. thesis. Virginia Tech.

Zhao, J., Gómez-Expósito, A., Netto, M., Mili, L., Abur, A., Terzija, V., Kamwa, I., Pal, B., Singh, A.K., Qi, J., et al., 2019a. Power system dynamic state estimation: Motivations, definitions, methodologies, and future work. IEEE Transactions on Power Systems 34, 3188–3198.

Zhao, J., Mili, L., 2018. A decentralized H-infinity unscented Kalman filter for dynamic state estimation against uncertainties. IEEE Transactions on Smart Grid 10, 4870–4880.

Zhao, J., Mili, L., 2018a. Power system robust decentralized dynamic state estimation based on multiple hypothesis testing. IEEE Transactions on Power Systems 33, 4553–4562.

Zhao, J., Mili, L., 2018b. A robust generalized-maximum likelihood unscented Kalman filter for power system dynamic state estimation. IEEE Journal of Selected Topics in Signal Processing 12, 578–592.

Zhao, J., Mili, L., 2019a. A decentralized H-infinity unscented Kalman filter for dynamic state estimation against uncertainties. IEEE Transactions on Smart Grid 10, 4870–4880.

Zhao, J., Mili, L., 2019b. Robust unscented Kalman filter for power system dynamic state estimation with unknown noise statistics. IEEE Transactions on Smart Grid 10, 1215–1224.

Zhao, J., Mili, L., Gómez-Expósito, A., 2019b. Constrained robust unscented Kalman filter for generalized dynamic state estimation. IEEE Transactions on Power Systems 34, 3637–3646.

Zhao, J., Zheng, Z., Wang, S., Huang, R., Bi, T., Mili, L., Huang, Z., 2019c. Correlation-aided robust decentralized dynamic state estimation of power systems with unknown control inputs. IEEE Transactions on Power Systems .

Zhou, N., Meng, D., Huang, Z., Welch, G., 2015. Dynamic state estimation of a synchronous machine using PMU data: A comparative study. IEEE Transactions on Smart Grid 6, 450–460.

Zolotas, A.C., Chaudhuri, B., Jaimoukha, I.M., Korba, P., 2006. A study on LQG/LTR control for damping inter-area oscillations in power systems. IEEE Transactions on Control Systems Technology 15, 151–160.