INTEGRITY AND ATTACK-RESILIENCE OF GPS-BASED POSITIONING AND TIMING: A BAYESIAN AND MEASUREMENT FUSION APPROACH

BY

SHUBHENDRA VIKRAM SINGH CHAUHAN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Aerospace Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2020

Urbana, Illinois

Doctoral Committee:

Assistant Professor Grace Gao, Chair and Director of Research
Associate Professor Timothy Bretl, Chair
Professor Peter Sauer
Research Assistant Professor Huy Tran

ABSTRACT

Robust Position, Velocity, and Timing (PVT) are essential for the safe operations of critical infrastructure sectors, such as transportation systems and power grids. Different transportation systems, both human-operated and autonomous vehicles, navigate using accurate position and velocity information. On the other hand, precise timing is crucial for various economic activities worldwide, such as banking, stock markets, and the power grid.

GPS serves as a backbone for many state-of-the-art applications related to these crucial infrastructures. GPS provides sub-microsecond accurate timing and meter level of accurate positioning. It has global coverage and is free for all users. The GPS positioning and timing service has some limitations. The positioning accuracy degrades in urban environments due to tall structures that block and reflect satellite signals. Degraded positioning is not safe for the operation of autonomously driving vehicles. Furthermore, GPS signals are susceptible to external attacks due to their low signal power and unencrypted signal structures. Researchers have shown that GPS Spoofing Attacks (GSAs) are feasible, and GSA for timing is able to alter timing without modifying the positioning solution. Such attacks create unsafe operating conditions for the modern power grid, which will use GPS timing for monitoring the wide-area network. The contribution of this work is to develop algorithms to mitigate the above limitations. We develop Bayesian algorithms that utilize multiple sensors and receivers.

For improving positioning, first, we design an adaptive filter based on Bayesian algorithms to augment GPS with the additional vision sensor. Second, we develop an integrity monitoring algorithm for Direct Positioning (DP), which is an advanced GPS receiver architecture that directly works on the position domain and is robust to signal blockage and multipath effects. To monitor integrity, we estimate vertical protection levels using a Bayesian approach. We further generate GPS datasets simulating open, semi-urban,

and urban environments for validating DP with multiple receivers.

For mitigating GSAs for timing, we design static and dynamic state estimators for the power grid. The static state estimator utilizes measurement residuals to correct power grid states. In the dynamic state estimator, we fuse GPS and power grid measurements to provide resiliency against GSAs. We create a virtual power grid testbed and generate datasets for a power grid network under different GSAs. These are the first datasets that contain both power grid and GPS measurements under GSAs, and we make them openly available. Our estimators are validated on various power grid networks and on the generated datasets.

To my parents, for their unconditional love and continual support.

ACKNOWLEDGMENTS

First, I wish to express my sincere gratitude to my advisor, Prof. Grace Xingxin Gao, for her continuous support, patience, and guidance throughout this dissertation. In her guidance, my technical and personal skills improved significantly. I am incredibly grateful to have her as my advisor and mentor throughout my Ph.D. life.

I want to extend my thanks to Cyber Resilient Energy Delivery Consortium (CREDC) team members for helping me in creating a virtual power grid testbed and generating datasets that are invaluable for this dissertation. Special thanks to Prosper Panumpabi for playing a significant role in creating the virtual power grid testbed, Prof. Peter Sauer, and Tim Yardley for allowing me to conduct experiments with RTDS, PMUs, and GPS clock.

I am fortunate to have worked with extraordinary members of our research group, past, and present. They were always available and never hesitated to help me. They are Akshay Shetty, Sriramya Bhamidipati, Arthur Chu, Craig Babiarz, Matthew Peretic, David Stier, Derek Chen, Enyu Luo, Ashwin Kanhere, Tara Mina, Siddharth Tanwar, Shubh Gupta, Pulkit Rustagi, and Adam Dai. I will always be grateful to have shared my time with them. I also wish to thank my friends Shamim Akhtar, Pratik Joshi, and Prakhar Gupta for making Urbana-Champaign memorable.

Finally, I would thank my parents Reena and Raghvendra Singh Chauhan, my brother Vishwendra Chauhan, and Sravya Aremanda for their unconditional love and continual support.

This material is based upon work supported by the Department of Energy (DOE) under Award Number DE-OE0000780. I would like to acknowledge the DOE for their financial support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AKF	Adaptive Kalman Filter
ALM	Altitude Likelihood Manifold
CSAC	Chip-Scale Atomic Clock
DP	Direct Positioning
ECEF	Earth Centered Earth Fixed
EKF	Extended Kalman Filter
ENU	East North Up
GSA	GPS Spoofing Attack
GSV	Google Street View
HIL	Hardware-In-the-Loop
IRIG-B	Inter-Range Instrumentation Group Timecodes-B
KF	Kalman Filter
LFTX	Low-Frequency Transmitter
LSE	Least Square Estimation
MC	Monte Carlo
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood
PDC	Phasor Data Concentrator
PEM	Pseudorange Error Models
PL	Protection Level

PMU	Phasor Measurement Unit
PVT	Position-Velocity-Time
RANSAC	RANdom SAmaple Consensus
RMSE	Root Mean Square Error
RTDS	Real-Time Digital Simulator
SCADA	Supervisory Control and Data Acquisition
SDR	Software Defined Radio
SE	State Estimator
SIFT	Scale Invariant Feature Transform
SR-SE	Spoofing-Resilient State Estimator
SR-SSE	Spoofing-Resilient Static State Estimator
SSE	Static State Estimator
USRP	Universal Software Radio Peripheral
VPL	Vertical Protection Level
WBX	Wide Bandwidth Transceiver

CHAPTER 1

INTRODUCTION

GPS provides sub-meter accurate positioning and sub-microsecond accurate timing service worldwide [1]. It is widely integrated into safety-critical infrastructure sectors such as power grid systems, transportation systems, banking, and communication systems, due to its free availability, reliability, and accuracy [2].

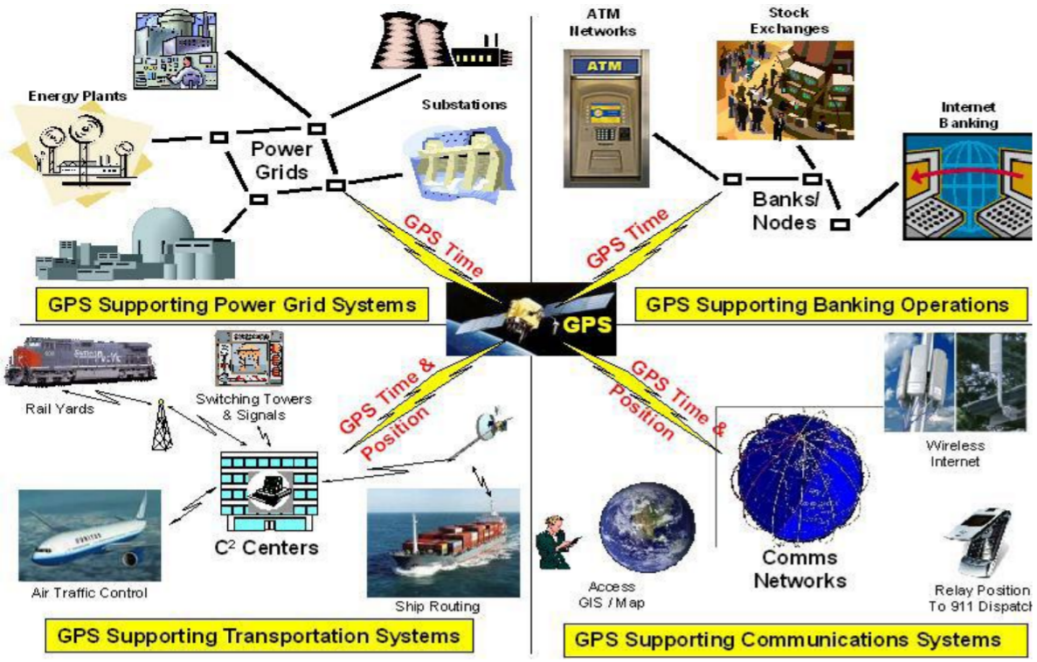


Figure 1.1: GPS positioning and timing service for many safety critical infrastructures [2]

Figure 1.1 shows the dependency of GPS positioning and timing services on various critical infrastructures. The modern power grid systems will utilize GPS timing to monitor a wide area network, which is necessary to perform essential tasks, including supervisory control and planning, bad data detection, optimizing power flows, security assessment for the grid, and detection of possible failures in power systems [3].

The transportation systems, including airplanes, railways, and ships, greatly use GPS positioning service to ensure safety. GPS plays an important role in the landing and takeoff phases of an aircraft flight. Its usage has resulted in reducing accidents, delays, and operating costs for railways. For maritime operations, it is being used for search and rescue, surveying, managing maritime port facilities, navigating to optimum fishing locations, and ensuring compliance with regulations. Furthermore, it has opened up a pathway for industry developing tools for autonomous navigation.

Apart from critical infrastructures, positioning and timing services are also used in public safety/ disaster relief. The precise location of police, fire, and rescue vehicles reduces the response time. Furthermore, ground and maritime vehicles equipped with GPS receivers can rapidly call for help and locate the crash site.

1.1 GPS Vulnerabilities and Limitations

The GPS positioning and timing service have some limitations, which can be broadly divided into two categories: positioning degradation in urban environments and vulnerability to external attacks. The presence of tall structures in the urban environment causes signal blockage and multipath, degrading the positioning service. Figure 1.2 illustrates a scenario consisting of signal blockage (SV_5) and multipath (SV_1 and SV_4).

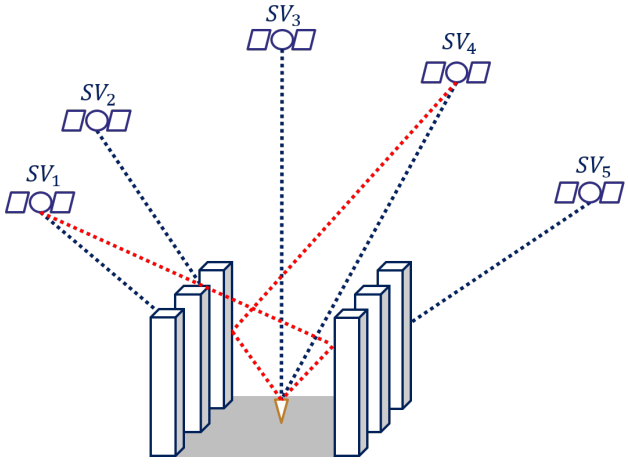


Figure 1.2: GPS positioning degrades in urban environments due to signal blockage and multipath.

The second vulnerability of susceptibility to external attacks arises from the low power and unencrypted structure of GPS signals [1]. A spoofing incident near the Russian port of Novorossiysk demonstrated that spoofing is a real threat to critical infrastructure [4]. More than 20 ships reported the same incorrect location that was off by 25 nm in this incident.

Ensuring the critical infrastructures' safe operations would require improved positioning service in GPS challenged environments and attack resilient timing service.

1.2 Related Work

In this work, we focus on addressing the aforementioned limitations of GPS positioning and timing service for transportation and power grid systems, respectively.

1.2.1 Improving Positioning in GPS Challenged Environments

Traditional GPS receiver architectures, such as scalar tracking loop [1] and vector tracking loop [5], use two steps to provide a Position-Velocity-Time (PVT) solution. First, the receiver estimates pseudoranges and then perform trilateration in the second step to obtain a PVT solution. In GPS challenging environments, such as the urban environment, the multipath's presence adds biases to pseudoranges and degrades GPS positioning service.

In the literature, the available approaches to improve positioning in GPS challenged environments can be put into two categories: fusing GPS measurements with a complementary sensor and developing advanced receiver architecture.

Urban environments are rich in features, such as corners, edges, planes, etc. Figure 1.3 shows the Scale Invariant Feature Transform (SIFT) [6] features in an urban environment. Utilizing these features in a vision-based localization approach will improve overall positioning. The accuracy of vision-based localization approaches is dependent on the number of unique features. It performs poorly in an open environment where such unique features are absent. However, GPS positioning is a more accurate open environment due to the absence of multipath or signal blockage. In this fashion, GPS and vision

are complementary to each other. Researchers utilize this complementary nature to fuse GPS and vision measurements to improve GPS positioning service in urban environments.

Figure 1.3: SIFT features are shown as red circles with a direction. The urban environment is abundant with unique features which improve localization using a vision-based approach.

The performance of a sensor fusion algorithm depends on process and measurement noise covariance matrices. Figure 1.4 shows the positioning estimate of a 1-D system using a Kalman Filter (KF) [7] in which the measurement noise changes in between the experiment. Figure 1.4 demonstrates that the estimator's performance also changes due to the measurement noise change.

The parameters of covariance matrices are, in general, manually tuned, which is a laborious task. In [8], researchers perform discriminative training using ground truth data to estimate these parameters to avoid manual tuning. However, this method and manual tuning suffer from overfitting, i.e., once tuned for a particular scenario, the parameters may not work with a different scenario. For instance, parameters tuned for an open environment for which the noise in GPS measurements is zero-mean Gaussian may not work for urban environments where the noise may no longer be Gaussian. The measurement noise in GPS measurements changes over time due to many factors, including the number of visible satellites, multipath, ionosphere and troposphere delays, and satellite geometry.

Covariance matching is a method that adaptively tunes the parameters of

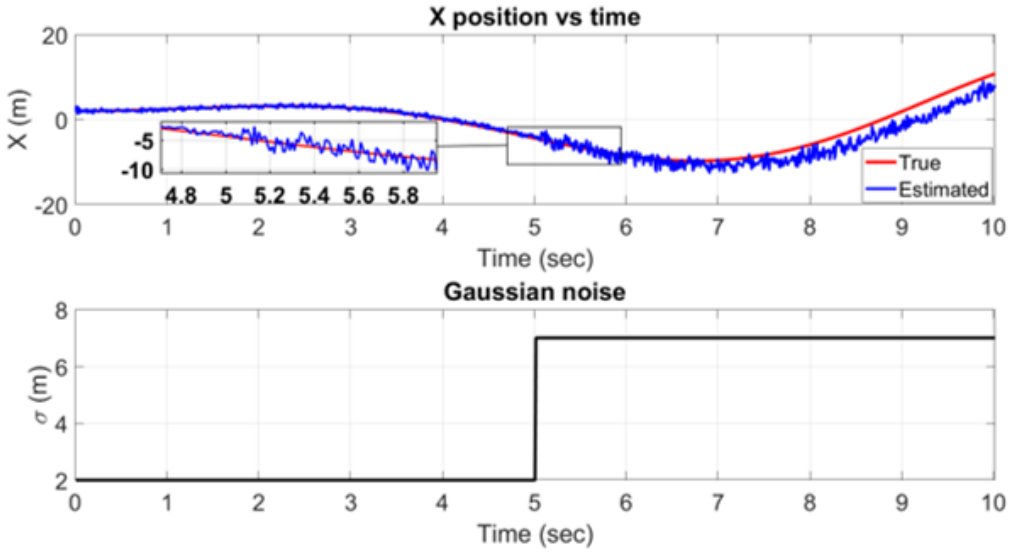


Figure 1.4: KF positioning estimate of a 1-D system in which measurement noise changes in between the experiment. The performance of the estimator depends on the measurement noise.

process and measurement noise covariance matrices using the received measurements [9]. The noise change due to different environmental conditions is captured in the received measurements, and thus, this method is robust to environmental conditions. In covariance matching, the size of measurements is assumed to be constant, which is an unrealistic assumption when we consider GPS measurements.

Another approach to improve GPS-positioning is to use an advance GPS receiver architecture. Direct Positioning (DP) is an unconventional GPS receiver architecture that directly operates in the PVT domain [10, 11, 12]. By the Maximum Likelihood (ML) estimation principle, DP estimates navigation solutions directly in the PVT domain in a single step [10, 13]. DP facilitates a deep coupling of the signals from different satellites, increases the effective signal power [11, 12], and utilizes a weak signal that would have otherwise been discarded [14, 15].

Existing works have shown the improved accuracy of DP in degraded signal environments using Cramer-Rao lower bound [16] to prove the higher achievable accuracy of DP than the two-step approach. Under various propagation models, software simulations have also demonstrated an improved accuracy performance of DP in noisy signal environments [16, 17, 18, 19, 20]. These improvements have been corroborated through live data experiments, includ-

ing stationary ground stations [12, 21], a hand-held device near a residential structure [22], and receivers mounted on automobiles [23].

Integrity measures the trustworthiness of a navigation solution [24]. It is one of the most critical requirements for safety-of-life applications. Protection levels (PLs) are used to assess the integrity requirement [25, 26, 27, 28]. PLs overbound positioning and timing errors by using error models.

A large amount of literature is available for overbounding positioning errors using Pseudorange Error Models (PEMs) [25, 26, 27, 28] for traditional receiver architecture, only a few are available for DP. With our best knowledge, there has not been a paper that empirically shows the error distribution for DP. Prior work [29] on DP-based integrity monitoring discusses the Solution-Separation Receiver Autonomous Integrity Monitoring framework. However, this framework is originally designed for PEMs. Another work [30] utilizes the correlation manifolds generated by DP to overbound the positioning errors. This approach is deeply coupled with the DP framework. However, the authors use many empirical parameters for overbounding vertical errors.

1.2.2 GPS Spoofing-Resilient Power Grid Monitoring

Wide-area monitoring of the power grid is necessary to perform essential tasks that ensure the power grid's safe operation, including supervisory control and planning, bad data detection, optimizing power flows, security assessment for the grid, and detection of possible failures in power systems [31, 32]. A State Estimator (SE) monitors the grid by estimating the substations' voltage phasors [33, 34]. Traditionally, SE monitors the power grid using Supervisory Control and Data Acquisition (SCADA) measurements, including injection power, power flow, and voltage magnitudes [32]. However, SCADA cannot provide real-time monitoring of the wide-area network as its measurements are asynchronous with a low update rate.

Compared to SCADA measurements, Phasor Measurement Units' (PMUs) measurements are 100 times faster [35] and are synchronized using GPS time. PMUs measure voltage magnitude, voltage phase angle, line current magnitude, line current phase angle, frequency, and rate of change of frequency [36]. Synchronized measurements with faster update rates make PMUs suitable for monitoring the power grid in real-time.

The Grid Modernization Initiative (GMI) is an effort to modernize the power grid network, ensuring greater resilience, improved reliability, enhanced security, additional affordability, superior flexibility, and increased sustainability [37]. Through this initiative, roughly 2500 PMUs have been installed throughout the North American power grid to establish wide-area situational awareness [3]. Figure 1.5 shows the PMU locations overs the North American power grid.

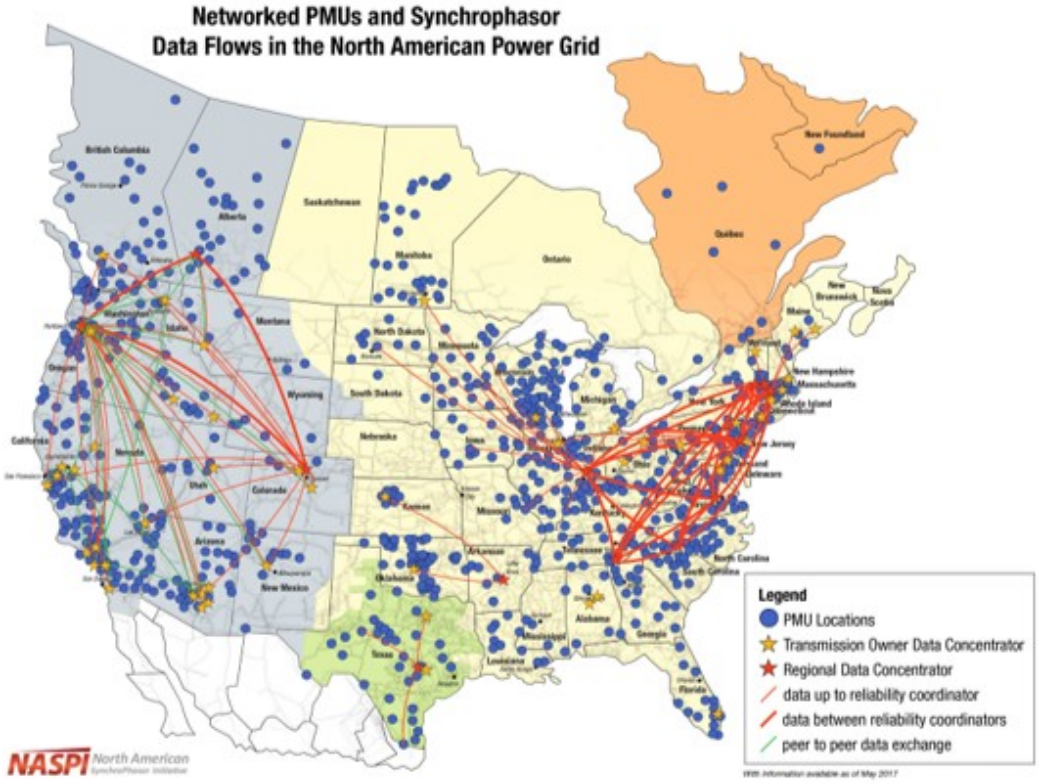


Figure 1.5: PMU locations over the North American power grid [3]

GPS provides sub-microsecond accurate timing [38] and plays a critical role in synchronizing PMU measurements. However, due to low signal power and the unencrypted structure of civilian GPS signals [1, 39, 40], PMUs are vulnerable to external attacks. GPS Spoofing Attack (GSA) is one such attack in which a counterfeit GPS signal is transmitted at a greater signal power than the civilian GPS signals, thereby inducing conventional GPS receivers [1, 41] to lock onto the counterfeit signal. These spoofing attacks can induce a time delay or modify the position of the GPS receiver. An induced time delay shifts the phase angle of the PMU measurements. The introduced

phase shift is referred to as the attack angle [42]. Various studies [43, 44, 45] show that GSAs are feasible. In [45], researchers design a portable spoofer that arbitrarily changes phase angles of PMU measurements by inducing time delays.

An incorrect phase angle degrades the performance of state estimation. The work in [46] demonstrates that a GSA induced time delay of 2.8 ms can mislead the fault line detection algorithm by 180 km. Furthermore, a state estimation algorithm can raise false warnings of power stability during a GSA and provide erroneous power flow estimates [44].

IEEE C37.118.1-2011 describes a standard for phase angle accuracy [47]. According to this standard, timing error must be less than 26.5 μs for a 60 Hz system to ensure the phase error is less than 0.01 radians. However, the IEC/IEEE 60255-118-1 standard highly recommends a time source to be at least ten times more accurate than 26.5 μs [36]. These standards are used to study power grid stability [48]. GSAs are capable of violating the IEEE C37.118.1-2011 standard [45]. GSA detection and mitigation are critical to ensure the safe operation of the power grid.

Similar to GSAs, bad data alter measurements by adding biases that degrade state estimation accuracy. For the power grid application, SEs use residuals to detect bad data [49, 50, 51, 52]. Although bad data usually induce an increased residual norm, some types of bad data can modify the states without increasing the residual norm [49]. A theoretical analysis of the impact of GSAs on residuals has not been provided in the literature. The theoretical analysis will identify if certain GSAs can alter the states without increasing the residual norm.

Related work is broadly divided into two parts: GSA detection and GSA mitigation. GSA detection algorithms utilize GPS signal properties to detect GSAs [53, 54, 55, 56, 57]. Various spoofing attacks and recommended countermeasures for commercial receivers are elucidated in [56]. The countermeasure strategies include drift monitoring and encryption-based defenses, as well as signal-geometry-based defenses. The presence of encrypted military P(Y) code in GPS raw signals is inspected in [53] for GSA detection. The inspection is performed by cross-correlating GPS raw signals between a wide network of power stations to find the spoofed node. Researchers in [54] design a novel GPS receiver architecture that uses multiple receivers to detect and locate a spoofer. The designed receiver architecture applies

the cross-correlation properties of GPS signals across multiple receivers for detecting GSAs. The work in [57] compares predicted and actual visible satellites to identify a GSA for a static receiver. A generalized likelihood ratio-based hypothesis test is devised in [58] to conduct GSA detection.

The GSA detection algorithm developed in [54] requires additional hardware that can provide GPS raw measurements, such as pseudoranges. The algorithms in [53, 57, 56, 58, 55] detect GSAs without providing a mechanism to mitigate the detected attacks, i.e., these algorithms do not estimate attack angles, which are essential for correcting PMU measurements.

After detecting a GSA, the SE must be able to mitigate the effects of the GSA in order to continue ensuring safe operations of the power grid network. The Spoofing-Matched (SpM) algorithm detects and mitigates the effect of a GSA on the power grid state estimation [59]. It is tested with simulated data and shown to be resilient against GSA. However, SpM is limited to mitigating a single GSA.

A joint state estimation and attack reconstruction algorithm is proposed in [42]. The joint estimation is formulated as an optimization problem. The proposed algorithm is capable of simultaneously estimating attack angles and states of a power grid network. This algorithm is computationally intensive, and it took 95 seconds to mitigate two GSAs for the IEEE 118 bus test case.

A PMU data correction algorithm is devised in [60]. This work utilizes the sparseness of attacked PMUs to correct PMU data. The devised method is applicable for correcting PMU data under multiple GSAs. However, state estimation is not performed in this work, which allowed the authors to relax the assumption of the power grid network's observability.

In all of these prior works, we observe that with a greater number of GSAs, the accuracy of these SEs decreases [61], and computation time increases. Furthermore, these SEs inherently assume a large GSA induced time delay (> 1 ms). Furthermore, these SEs are validated using low-fidelity simulations due to a lack of relevant datasets.

Experimental datasets containing both GPS and PMU measurements are essential to assess GSAs' impact on the power grid's SEs. It is challenging to validate GSA resilient SE through real-world experiments since it is illegal to broadcast signals at GPS frequency, and experimenting on the real power grid is costly. Available GSA datasets in the literature [62] are not relevant to the power grid applications due to the following reasons:

- Most of the GSAs in [62] alter the receiver position and are detectable since the power grid's substations are static. GSAs that modify the receiver time without changing its location pose a threat to the power grid. Such GSAs are known as timing GSAs [46].
- Timing GSAs in [62] induce a total time delay of 2 μs. According to the IEEE C37.118-2011 [48] standard, timing delay should be less than 26.5 μs to ensure the phase error is less than 0.01 radians. However, the literature does not cover the scenarios for which time delay lies between 26.5 μs and 1 ms.
- The current spoofing datasets in the literature are for GPS only. None of the available datasets contain both GPS and PMU measurements during a GSA.

1.3 Our Contributions

In this dissertation, we aim to improve GPS positioning service in urban environments and provide resilient GPS timing service. Our approach consists of fusing multi-sensor and multi-receivers measurements using a Bayesian approach. This dissertation presents algorithms for improving positioning in urban environments and providing resiliency against GSAs to power grid's SEs. The contributions of our works in [63, 64, 65, 66, 67] are as follows:

1. **Adaptive Sensor Fusion**: The performance of a sensor fusion algorithm depends on process and measurement noise covariance matrices, as illustrated in Figure 1.4. The parameters of these covariance matrices varies in size and with time. For instance, at a given position, the geometry of visible satellites changes with time as well as the number of the visible satellites. Tuning these parameters is a challenging and tedious task. Furthermore, the tuned parameters may result in poor performance if tested in different environmental scenarios.

   We develop an adaptive sensor fusion algorithm that estimates these parameters using the received measurements. Our algorithm implements the logic that noise implicitly affects measurements, and we can estimate the noise level by utilizing the received measurements.

The developed algorithm estimates the noise parameters that change in size and with time. Furthermore, we implemented a GPS-Vision fusion algorithm and compared the adaptive sensor fusion algorithm's performance with GPS only, vision only, and sensor fusion with fixed covariance matrices [63].

2. **Integrity Monitoring for DP**: The available integrity monitoring algorithms for DP either utilize integrity framework developed for conventional GPS receivers or use integrity framework for DP but with many environmental specific empirical parameters. Furthermore, the positioning error distribution for DP is not available in the literature.

   We devise a Bayesian algorithm to perform integrity monitoring for DP. The developed algorithm uses DP framework and does not require environmental specific empirical parameters. The developed algorithm is robust to the unknown positioning error distribution. We generate 24 hours of stationary GPS raw signal dataset using a high-fidelity GPS simulator.We obtain 4 million DP's positioning solution data points from the generated datasets. With our best knowledge, this is the

   first work that shows DP's positioning error distribution is multi-modal. The developed Bayesian algorithm is validated on the generated dataset, and we observe that the developed algorithm overbounds DP's positioning errors [64].

3. **GSA-Resilient Static State Estimation**: The modern power grid will utilize PMUs for monitoring a wide-area-network. These devices are vulnerable under GSAs due to their dependence on GPS signals for time synchronization. GSA-resilient SEs for the power grid in the literature utilize PMU measurements for mitigating GSAs. However, the literature does not provide theoretical analysis for identifying scenarios that might not be detectable using PMU measurements. In prior works, the GSA-resilient SEs mitigate at most one GSAs [59] or mitigate multiple GSAs but are computationally intensive to implement [42].

   We perform a theoretical analysis on the PMU measurement residuals and derive a necessary condition that demonstrates an increase in

residual norm during GSAs. We show that the distribution of residual changes during GSAs, which we exploit to detect GSAs. We propose a novel residual-based Spoofing-Resilient Static State Estimator (SR-SSE) for the power grid, which is resilient to multiple GSAs with different attack angles.

SR-SSE consists of two algorithms: Spoofing Detection and Measurement Correction. The Spoofing Detection algorithm monitors the residual norms to distinguish between a nominal and a spoofed scenario. In the nominal scenario, none of the PMU buses are spoofed, while in the spoofing scenario, one or more PMU buses are spoofed with different attack angles. The Measurement Correction algorithm corrects the false PMU measurements by iteratively minimizing the residual norms. We conduct MC simulations on IEEE 14, IEEE 39, IEEE 118, and Illinois 200-bus [68] test systems for different GSAs to test our derived necessary condition and validate SR-SSE [65].

4. **Experimental GPS and PMU Datasets Under GSAs**: Experimental datasets containing both GPS and PMU measurements under GSAs are unavailable in the literature. We generate the GPS and PMU datasets relevant to the power grid community for different GSAs.

   We devise a methodology for generating integrated datasets containing PMU measurements coupled with GPS measurements for nominal and spoof scenarios. The nominal scenario refers to an ideal environment in which GPS signals are authentic. In the spoof scenario, we simulate timing GSAs with time-walk that induce a linearly increasing timing delay greater than 26.5 μs.

   We generate openly available integrated datasets for nominal and spoof scenarios by performing Hardware-In-the-Loop (HIL) simulations with Real-Time Digital Simulator (RTDS), PMUs, and GPS clock. The integrated datasets have PMU datasets coupled with GPS datasets. The PMU datasets consist of GPS timestamped voltage phasor, current phasor, frequency, and frequency change rate. The GPS datasets contain GPS raw signals, satellite positions, and pseudoranges. With our best knowledge, these are the first openly available datasets that con-

tain both GPS and PMU measurements under various timing GSAs. These datasets will serve as an evaluation platform for testing power grid SE's performance under various timing GSAs [66].

5. **GSA-Resilient State Estimation Using an Extended Kalman Filter (EKF)**: The prior works [59, 42, 60, 65] minimize a complex objective function to estimate attack angles for different numbers of GSAs. The minimization step is computationally intensive and may not always reach the global minimum. Due to this, the computation time increases, and the accuracy decreases with an increase in the number of GSAs. Furthermore, these algorithms will fail to mitigate GSAs if the number of GSAs is higher than half of the number of installed PMUs [61].

   We propose a novel Spoofing-Resilient State Estimator (SR-SE) for the power grid that fuses time-varying GPS and PMU measurements using an EKF. The sequential estimator removes the minimization step by incorporating the time-varying GPS and PMU measurements in state estimation. With our best knowledge, this is the first estimator that utilizes both GPS and PMU measurements for mitigating GSAs.

   SR-SE jointly estimates power grid states and receiver clock biases. By keeping track of individual clock biases of the receivers, SR-SE is capable of mitigating GSAs even when the number of GSAs is higher than half of the number of installed PMUs. In SR-SE, we design a GPS-PMU coupled measurement model that relates GSA induced time delay to PMU measurements. We remove the minimization step by incorporating the time-varying GPS and PMU measurements in state estimation using a sequential EKF. The time-varying GPS measurements enable the SE to track the induced time delay for each PMU, thereby simultaneously tracking the attack angle during a GSA. This measurement model is essential to maintain estimates of attack angles, which is necessary to mitigate GSAs. We conduct HIL and MC simulations to test SR-SE on IEEE 14, IEEE 39, and Illinois 200-bus [68] test systems for different GSA scenarios.

1.4 Outline of the Dissertation

Chapter 2 describes the adaptive sensor fusion algorithm for improving positioning in urban environments. In this algorithm, we fuse pseudoranges from GPS receiver with vision measurements. The algorithm estimates the covariance matrices that adapt to environmental changes. We validate the algorithm through simulation and real-world experiments.

Chapter 3 provides an overview of DP and its implementation. The detailed Bayesian PL estimation algorithm, which is built on DP architecture, is presented afterward. We test the developed algorithm using high-fidelity simulations in which we generate GPS raw signals for a stationary receiver. We verify that DP's positioning error distribution is multi-modal and validate that the estimated PLs overbound the positioning errors.

In chapter 4, we present an overview of conventional SEs for the power grid. We describe our SR-SSE in detail and show theoretical analysis for detecting GSAs using PMU measurement residuals. The developed algorithm is validated on various power grid test systems for different GSA scenarios. We end the chapter with the limitations of SR-SSE.

Chapter 5 devises a methodology for generating GPS and PMU measurements under nominal and spoof scenarios. The nominal scenario represents an ideal environment in which GPS signals are authentic. In the spoof scenario, GPS signals are modified to mimic a timing GSA that modifies receiver time without altering the receiver location. Using the devised methodology, we generated openly available GPS, and PMU integrated datasets by performing HIL simulations with RTDS, physical PMUs, virtual PMUs, and GPS clock. This chapter provides details for the generated datasets. Furthermore, we demonstrate that the integrated datasets involve timing GSAs with time-walk. The integrated datasets will serve as an evaluation platform for testing the performance of SEs for the power grid.

Chapter 6 describes the SR-SE algorithm that addresses the limitations of the SR-SSE algorithm. In this algorithm, we remove the computation-intensive minimization step by incorporating GPS measurements in state estimation with PMU measurements using a sequential EKF. This is the first algorithm that utilizes both GPS and PMU measurements to estimate power grid states. SR-SE is tested with various power grid test systems through MC simulations for different GSAs. Furthermore, we validate SR-

SE on the generated integrated GPS and PMU datasets. Finally, Chapter 7 summarizes our contributions.

CHAPTER 2

ADAPTIVE SENSOR FUSION

Navigation in urban environments with standalone GPS is challenging due to tall structures that block and reflect GPS signals. The standalone GPS positioning accuracy degrades in urban environments because of signal blockage and multipath. Figure 2.1 shows GPS positioning in an urban environment which deviates from the true position.

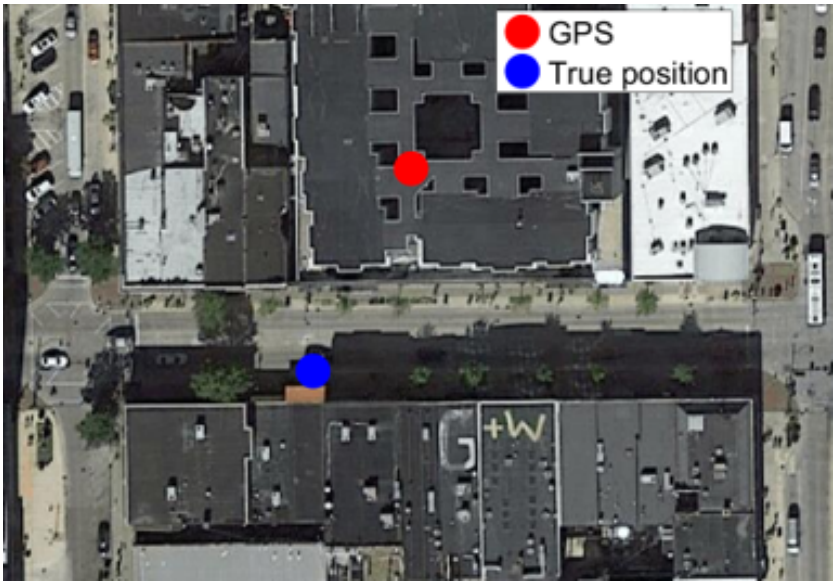


Figure 2.1: GPS positioning degrades in urban environment. The blue circle shows the true position and red circle shows the position estimated from GPS

One approach to improve positioning in an urban environment is to augment GPS with a complimentary sensor using a sensor fusion algorithm. The performance of a sensor fusion algorithm depends on covariance matrices. Tuning these matrices is time-consuming. Furthermore, once tuned for a particular scenario, these matrices may result in poor performance for a different scenario as noise may change with time. The noise in GPS measurements varies with time as the number of visible satellites and their geometry

change over time for a given position. Covariance matching [9] is not applicable due to the inherent assumption of the constant size of measurements.

This chapter presents our work on an adaptive sensor fusion algorithm that adapts to time and size varying noise. We fuse GPS and vision measurements to improve positioning in urban environments. The remainder of this chapter is organized as follows: Section 2.1 provides an overview of our approach. In Section 2.2, we describe the process of obtaining a position from camera images. The adaptive estimation of covariance matrices is elaborated in Section 2.3. Section 2.4 details the measurements used from GPS and vision. The developed algorithm is tested in simulation, and in the real-world, the details of simulation and real-world experiments are provided in Section 2.5 and 2.6, respectively.

2.1 Approach

The noise parameters for GPS and vision change with time and environmental scenarios. Figure 2.2 shows that the noise affects measurements, and in our approach, we utilize measurements to estimate noise parameters.

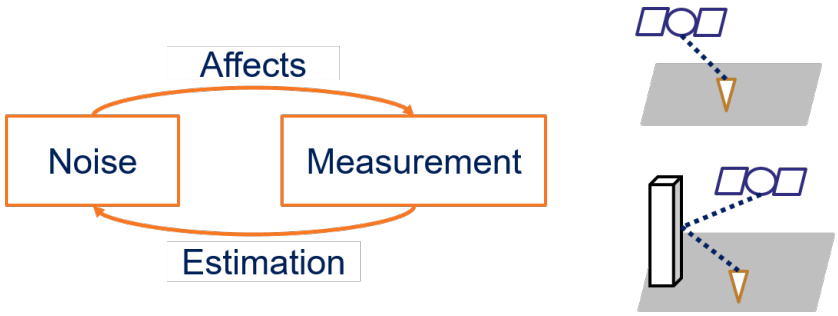


Figure 2.2: Noise affects measurements differently based on the scenario. The affected measurements can be utilized to estimate noise.

The overall architecture of our approach is shown in Figure 2.3. GPS and vision measurements are fused using an EKF. Pseudoranges from GPS receiver and position estimates from vision are input to EKF. In Image matching, we compare the raw image from the camera and Google Street View (GSV) to estimate the raw image's position. In Covariance estimation, innovation sequence and Kalman gain are used to estimate the covariance matrices that are used in EKF for fusing the measurements. The remaining

sections provide a detailed description of Image matching and Covariance estimation blocks.

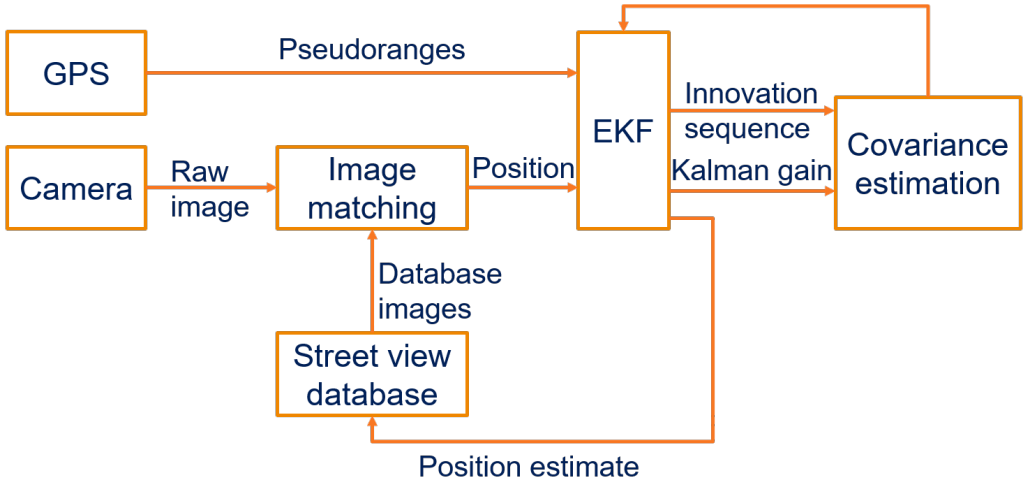


Figure 2.3: Overall architecture for adaptive sensor fusion algorithm.

2.2 Image matching

GSV contains a database of geo-referenced images. An image can be pulled from GSV by providing the latitude and longitude of the desired position. The resolution of this database is roughly 10 m, i.e., GSV gives the same image if the difference between the positions of two pull requests is less than 10 m.

It is challenging to match images that have a different scale, orientation, and lighting conditions. We use SIFT [6] features, which are invariant to scale, rotation, and lighting conditions, for image matching. Figure 2.4 shows the steps involved in matching the camera image with the GSV database.

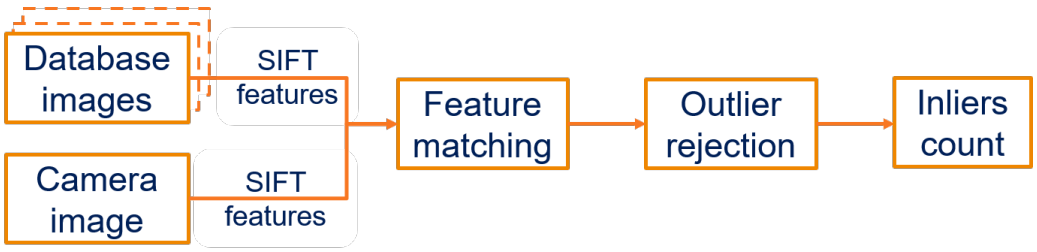


Figure 2.4: Image matching using SIFT features

First, we obtain SIFT features for all the images present in the database and the camera image. Then we find the feature correspondences by implementing Lowe's criteria, in which two features correspond to each other if the ratio of the distance between the two nearest neighbors is below a certain threshold. Feature matching is susceptible to outliers. After finding the correspondences, we utilize a RANdom SAmple Consensus (RANSAC) algorithm [69] on the matched features to estimate homography between the two images.

The homography is used to reject outliers by projecting features from the camera image to the database image and checking the difference between the projected features and their correspondences. Inliers are the features that lie within a certain threshold to their correspondences after applying the homography. We compute the number of inliers for each database image and consider the database image with the largest number of inliers as a match to the camera image. The position of the matched database image is vision measurement, i.e., the raw image's position.

2.3 Adaptive Covariance Estimation

Innovation sequence is the difference between received measurements and expected measurements based on the measurement model. Any time and size variation of noise will affect the measurements, and the effect will be observed in the innovation sequence. We apply innovation sequences to estimate process and measurement noise covariance matrices. Figure 2.5 outlines the steps involved in estimating the covariance matrices using the innovation sequence.

EKF is a sequential filter, to which the inputs are measurements and states from the previous time step. We utilize the innovation sequence and measurement model to get an expression of the measurement noise covariance matrix. Similarly, we use the process model, Kalman gain, and innovation sequence to express the process noise covariance matrix. The estimated covariance matrices are passed to EKF in the next time step. The subsequent subsections provide more details to filter equations and expressions of estimated covariance matrices.

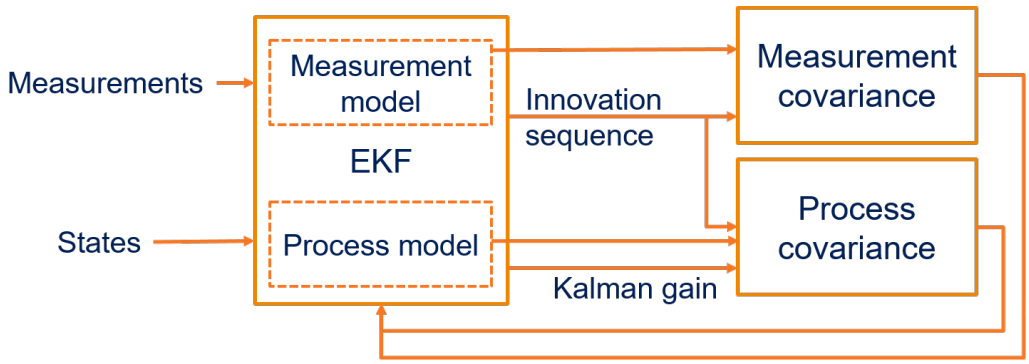


Figure 2.5: Adaptive estimation of covariance matrices using innovation sequence.

2.3.1 Filter Equations

An EKF consists of a prediction and a measurement update step, which utilizes the process and measurement models. A generic discrete process model is given by

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \boldsymbol{\omega}_k \tag{2.1}$$

where $\mathbf{x}$ is the state of a given system, k denotes the time instant, $\mathbf{f}$ is a state transition function that relates states from previous time step to current time step, $\boldsymbol{\omega}$ is zero-mean Gaussian noise. A generic discrete measurement model is given by

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \boldsymbol{\eta}_k \tag{2.2}$$

where $\mathbf{z}$ denotes the measurements, $\mathbf{h}$ is the measurement function that relates states to the measurements, and $\boldsymbol{\eta}$ is zero-mean Gaussian noise.

In the prediction step, EKF propagates states forward in time using the process model. The following equations are used in the prediction step

$$\mathbf{x}_{k|k-1} = f(\mathbf{x}_{k-1|k-1}) \tag{2.3}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^\top + \mathbf{Q}_k \tag{2.4}$$

where $\mathbf{x}_{p|q}$ denotes the state $\mathbf{x}$ at the p^{th} time instant given the measurements till the q^{th} time instant, $\mathbf{P}$ denotes the state covariance matrix, $\mathbf{F}_k$ is the Jacobian of $\mathbf{f}$ evaluated at $\mathbf{x}_{k-1|k-1}$ and $\mathbf{Q}_k = \mathbb{E}[\boldsymbol{\omega}_k \boldsymbol{\omega}_k^\top]$ is the process noise covariance matrix, which is a diagonal matrix.

New measurements allow EKF to update the states using the process and measurement models. The following equations are used in the measurement update step

$$\tilde{\mathbf{y}}_k = \mathbf{z}_k - \mathbf{H}_k \mathbf{x}_{k|k-1} \tag{2.5}$$

$$\mathbf{S}_k = \mathbf{R}_k + \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top \tag{2.6}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^\top \mathbf{S}_k^{-1} \tag{2.7}$$

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{K}_k \tilde{\mathbf{y}}_k \tag{2.8}$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} \tag{2.9}$$

where $\tilde{\mathbf{y}}$ denotes the innovation sequence, $\mathbf{H}_k$ is the Jacobian of $\mathbf{h}$ that is evaluated at $\mathbf{x}_{k|k-1}$, $\mathbf{S}$ is the theoretical covariance of innovation sequence, $\mathbf{R}_k = \mathbb{E}[\boldsymbol{\eta}_k \boldsymbol{\eta}_k^\top]$ is the measurement noise covariance matrix, $\mathbf{K}$ is the Kalman gain, and $\mathbf{I}$ is the identity matrix. EKF sequentially performs predict and update steps, and its performance depends on process and measurement noise covariance matrices, i.e., $\mathbf{Q}$ and $\mathbf{R}$ respectively.

2.3.2 Covariance Estimation

We assume the process and measurement noise to be independent and Gaussian in a small time interval of length W. For a given time instant k, we compute the covariance of innovation sequence using the received measurements as

$$\mathbb{E}[\tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_k^\top] = \frac{1}{W-1} \sum_{i=1}^{W} (\tilde{\mathbf{y}}_{k-i} - \bar{\tilde{\mathbf{y}}})(\tilde{\mathbf{y}}_{k-i} - \bar{\tilde{\mathbf{y}}})^\top \tag{2.10}$$

where $\bar{\tilde{\mathbf{y}}}$ is the mean of W samples of innovation sequence around the time instant k. Now, rearranging the terms in (2.6) will result in the following equation

$$\hat{\mathbf{R}}_k = \mathbf{S}_k^W - \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top \tag{2.11}$$

where $\hat{\mathbf{R}}$ is the estimate of measurement noise covariance matrix and $\mathbf{S}_k^W = \mathbb{E}[\tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_k^\top]$ is the computed covariance matrix of the innovation sequence. The $\hat{\mathbf{R}}$ is a diagonal matrix due to the assumption of noise being independent and Gaussian. The diagonal element of $\hat{\mathbf{R}}$ is given by

$$\hat{\mathbf{R}}_k(i,i) = \mathbf{S}_k^W(i,i) - \mathbf{H}_k(i,:)\mathbf{P}_{k|k-1}\mathbf{H}_k(i,:)^\top \tag{2.12}$$

where $\hat{\mathbf{R}}_k(i,i)$ is the i^{th} diagonal element of $\hat{\mathbf{R}}_k$, $\mathbf{S}_k^W(i,i)$ is the i^{th} diagonal element of $\mathbf{S}_k^W$, and $\mathbf{H}_k(i,:)$ denotes the i^th row of $\mathbf{H}_k$. Size variation in the measurements will modify the innovation covariance in (2.10) and the size of estimated covariance noise matrix will be adjusted accordingly in (2.12).

We use process model, estimated covariance of innovation sequence and Kalman gain to estimate elements of process noise covariance matrix. Equation (2.1) is rearranged to obtain an estimate of process noise, $\hat{\boldsymbol{\omega}}$, as

$$\begin{aligned}
\hat{\boldsymbol{\omega}}_k &= \mathbf{x}_{k|k} - f(\mathbf{x}_{k-1|k-1}) \\
&= \mathbf{x}_{k|k} - \mathbf{x}_{k|k-1} && \text{From rearranging (2.3)} \\
&= \mathbf{K}_k\tilde{\mathbf{y}}_k && \text{From rearranging (2.8)}
\end{aligned} \tag{2.13}$$

Taking expectation on both sides of (2.13) will result in following equation

$$\hat{\mathbf{Q}}_k = \mathbf{K}_k\mathbf{S}_k^W\mathbf{K}_k^\top \tag{2.14}$$

where $\hat{\mathbf{Q}}_k$ is the estimated process noise covariance matrix. We smooth the estimates of covariance matrices by weighing previous estimates, using the following equation

$$\hat{\mathbf{R}}_k(i,i) = \alpha\hat{\mathbf{R}}_k(i,i) + (1-\alpha)\hat{\mathbf{R}}_{k-1}(i,i) \tag{2.15}$$

$$\hat{\mathbf{Q}}_k = \alpha\hat{\mathbf{Q}}_k + (1-\alpha)\hat{\mathbf{Q}}_{k-1} \tag{2.16}$$

where $\alpha \in (0,1)$ denotes the parameter for weighing the estimates. Equations (2.15) and (2.16) show the expressions of estimated covariance matrices. To ensure that the estimated covariance matrices are positive definite, we take the absolute value of the right hand side in (2.15) and (2.16). This section completes the discussion of estimating covariance matrices for a generic sensor fusion algorithm which will be referred as Adaptive Kalman Filter (AKF). In the subsequent sections, we apply this algorithm for fusing GPS and vision measurements.

2.4 GPS-Vision Fusion

The states considered for GPS and vision are given by

$$\mathbf{x} = [x\ \dot{x}\ y\ \dot{y}\ z\ \dot{z}\ C\delta t\ C\dot{\delta t}]^\top \tag{2.17}$$

where (x, y, z) and $(\dot{x}, \dot{y}, \dot{z})$ denote position and velocity in Earth Centered Earth Fixed (ECEF) frame, respectively, C is the speed of light in vacuum, and $(\delta t, \dot{\delta t})$ are the receiver clock bias and drift, respectively.

We use a constant velocity as our process model which is linear. The state transition function in (2.1) becomes a constant matrix and the process model is given by

$$\dot{\mathbf{x}}_{k|k-1} = \mathbf{F}\mathbf{x}_{k-1|k-1} + \boldsymbol{\omega}_k \tag{2.18}$$

where $F = \begin{bmatrix} 1 & \Delta t & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \Delta t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$, Δt is a small time step and ω_k is process noise. The measurements used from GPS are pseudoranges and vision positions obtained by matching camera images with GSV database images. The overall measurement vector is given by

$$\mathbf{z} = \begin{bmatrix} \rho^{SV_1} & \cdots & \rho^{SV_i} & \cdots & \rho^{SV_N} & x_{vision} & y_{vision} & z_{vision} \end{bmatrix}^\top \tag{2.19}$$

where ρ^{SV_i} denotes pseudorange between the SV_i^{yh} satellite and the receiver, N is the total number of visible satellites and $(x_{vision}, y_{vision}, z_{vision})$ is the

position obtained from vision. The measurement model is given by

$$\mathbf{z} = \begin{bmatrix} \sqrt{(X^{SV_1}-x)^2+(Y^{SV_1}-y)^2+(Z^{SV_1}-z)^2}+C\delta t \\ \vdots \\ \sqrt{(X^{SV_i}-x)^2+(Y^{SV_i}-y)^2+(Z^{SV_i}-z)^2}+C\delta t \\ \vdots \\ \sqrt{(X^{SV_N}-x)^2+(Y^{SV_N}-y)^2+(Z^{SV_N}-z)^2}+C\delta t \\ x \\ y \\ z \end{bmatrix} + \boldsymbol{\eta} \qquad (2.20)$$

where $(X^{SV_i}, Y^{SV_i}, Z^{SV_i})$ denotes the satellite position in ECEF frame. Equations (2.18) and (2.20) describe the process and measurement model for fusing GPS and vision measurements. We implement the adaptive sensor fusion algorithm with the above process and measurement models.

2.5 Simulation Environment and Results

We simulate a scenario consisting of both open and urban environments. To simplify the calculations, we make following assumptions

- Motion is in 2-dimensional plane.
- Virtual satellites are stationary in 2-dimensional plane.
- Receiver and satellites' clock are synchronized.
- Noise in measurements have different variance in open environment compared to urban environment.

Figure 2.6 illustrates the 2-dimensional simulation scenario. There are four virtual stationary satellites located at (X^{SV_i}, Y^{SV_i}) with a pseudorange of ρ^{SV_i} from the receiver that is located at (x, y) and $1 \leq i \leq 4$ is a natural number. The receiver moves along the orange segment, i.e., from D-E-F-G. The segment EF represents an urban environment in which the signals from SV_2 and SV_3 are received after reflection from the blue wall. The remaining segments simulate the open environment in which the signals are received

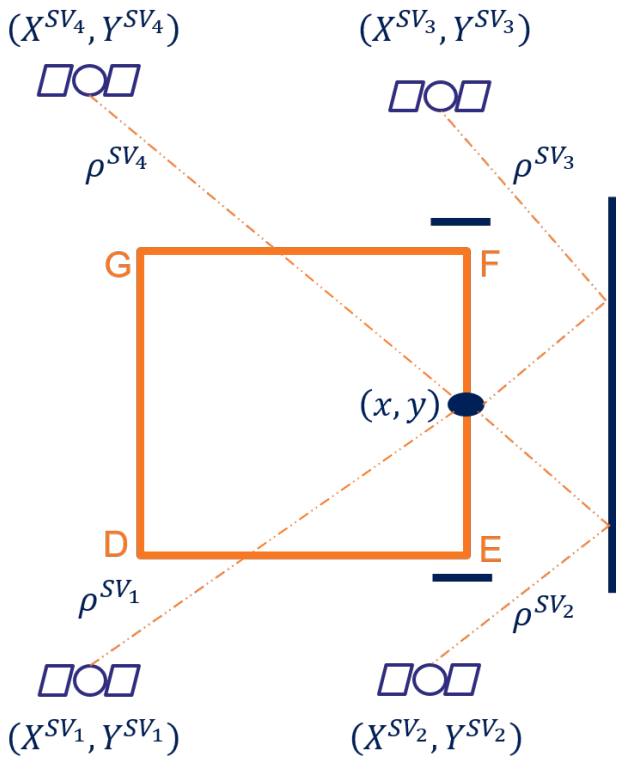


Figure 2.6: 2-dimensional simulation scenario with 4 stationary virtual satellites. The vehicle moves from point D-E-F-G-D. Segment EF simulates an urban environment, where the signals from SV_2 and SV_3 are received after reflection from the blue wall. The rest of the segments simulate an open environment, where the vehicle receives signals from all the satellites without any reflection.

without reflection. We add Gaussian noise to vision measurements with a larger variance in open environment than urban environment. The variable variance in vision measurements incorporates the hypothesis that the position obtained from vision is more reliable in urban environments.

Figure 2.7 shows the measurements generated for the simulation scenario, shown in Figure 2.6. The receiver enters segment EF at time instant 1000 and leaves at time instant 2000. From Figure 2.7, we observe sudden jumps in pseudoranges for SV_2 and SV_3 when the receiver enters and leaves the segment EF, simulating reflection from the wall. Similarly, the vision measurements have lower noise variance in segment EF than other segments, demonstrating that vision measurements are more reliable in the urban environment.

Conventional GPS receiver solves pseudorange equations using Least Squares

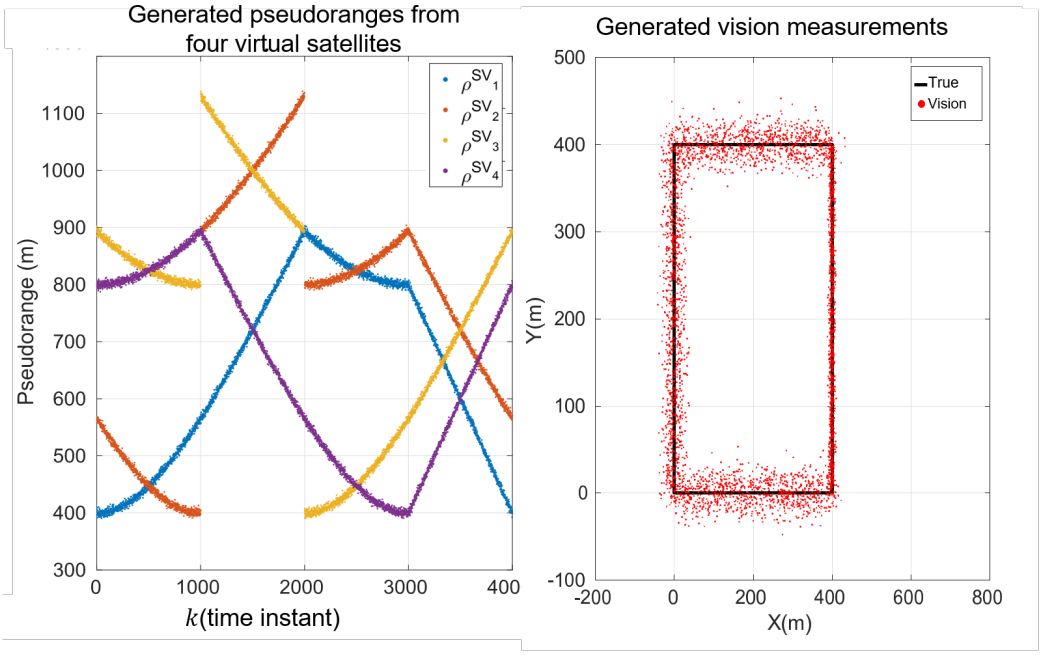


Figure 2.7: Generated pseudoranges and vision measurements for the simulation scenario

Estimation (LSE) to obtain a positioning solution. We apply LSE on the generated pseudoranges to obtain positioning solution which is shown in Figure 2.8. The reflection in the segment EF results in creating a bias in positioning solution which is illustrated in Figure 2.8

We utilize the generated pseudoranges and vision measurements in our AKF algorithm to obtain a positioning solution. In AKF, we also estimate the noise standard deviation for vision measurements. Figure 2.9 shows the estimated standard deviation of vision measurements' noise along X and Y directions. As mentioned in the simulation scenario, we added noise with less variance in segment EF to vision measurements compared to other segments. Figure 2.9 illustrates that AKF keeps track time-varying noise.

The estimated trajectory of the receiver is shown in Figure 2.10. Compared to pseudorange only solution, shown in Figure 2.8, and vision measurements, shown in Figure 2.7, the AKF positioning estimate is closer to the ground truth for all the segments. The positioning estimate has less variance in the open environment than vision measurements and does not have a bias in the urban environment. Therefore, AKF adapts to time-varying noise present in the sensor measurements. This section shows proof of concept via simulations. In the next section, we test the algorithm in real-world where

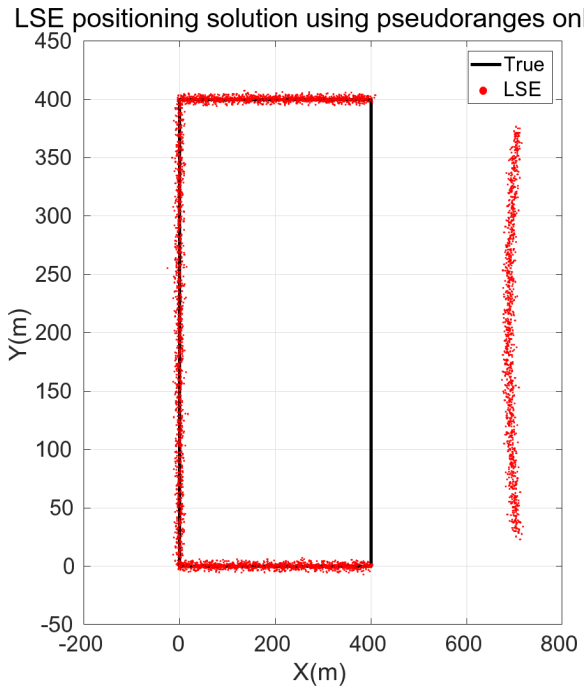


Figure 2.8: LSE positioning solution obtained by solving the generated pseudoranges

we encounter size variation in GPS measurements.

2.6 Real-World Testing

2.6.1 Experimental Scenario

We conducted a real-world experiment on the UIUC campus that consists of both urban and open environments. The route taken during the experiment is shown in Figure 2.11 with a red line. We recorded pseudoranges from a commercial GPS receiver and video from a handheld mobile camera in the shown route. The initial path along the west direction consisted of tall buildings, and we expect a few satellites in the north-south direction to be blocked. The rest of the path is considered an open environment due to the presence of shorter buildings.

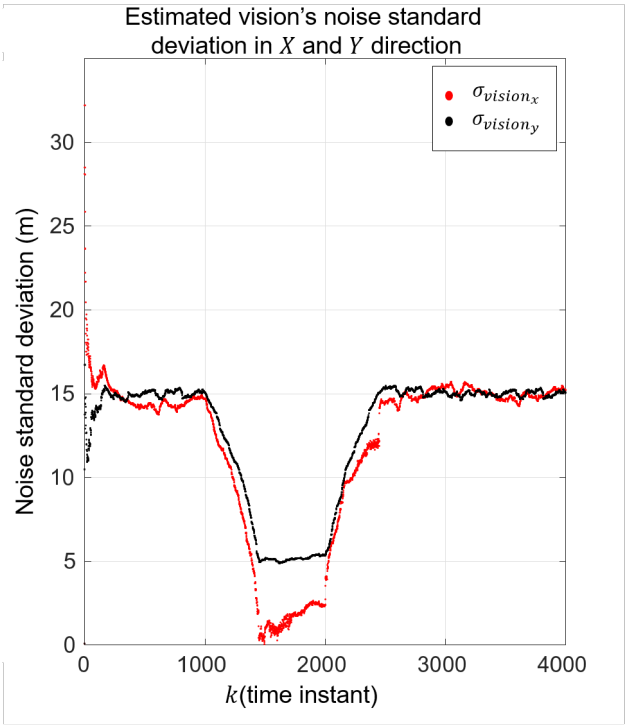


Figure 2.9: Estimated standard deviation of vision measurements' noise along X and Y directions.

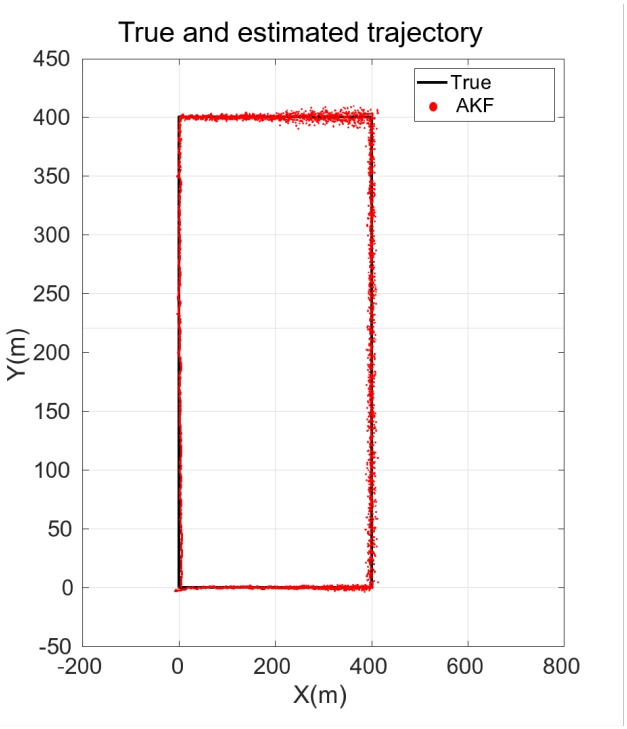


Figure 2.10: Estimated and true trajectory of the receiver

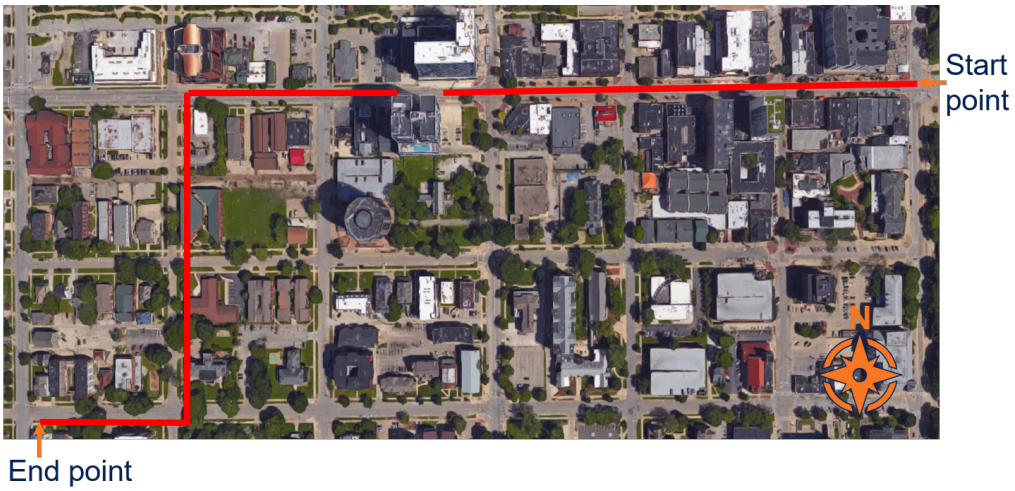


Figure 2.11: Route taken during the experiment on UIUC campus.

2.6.2 Results

Figure 2.12 shows the intermediate results of the image matching algorithm. The top image is obtained from GSV, and the bottom image is collected while experimenting. The subsequent images in each row, from left to right, correspond to the block diagram's steps. For a given database image and camera image, first SIFT features are computed, shown with green circles. Next, features are matched using Lowe's criteria, which are shown in the third column. This step is susceptible to outliers. For instance, the red car's features in the database image should not be matched with any camera image features as the car is not present in the camera image. The homography is obtained by applying the RANSAC algorithm to the matched features, and then it is used to remove the outliers. The last column of images shows the matched features after the outliers are rejected. From the last column of images, it can be visually verified that the features shown in the database image correspond to the features shown in the camera image.

Measurements used in our filter consist of pseudoranges from GPS receiver and position from vision. The number of visible satellites varies depending on the position and time of the day. Due to this, the size of the measurements changes with time. Figure 2.13 shows the variation of the length of the received measurements with time. It illustrates that the size of the measurements is not constant. Furthermore, the noise varies with time and environment. Figure 2.14 shows the estimated noise variance in vision mea-

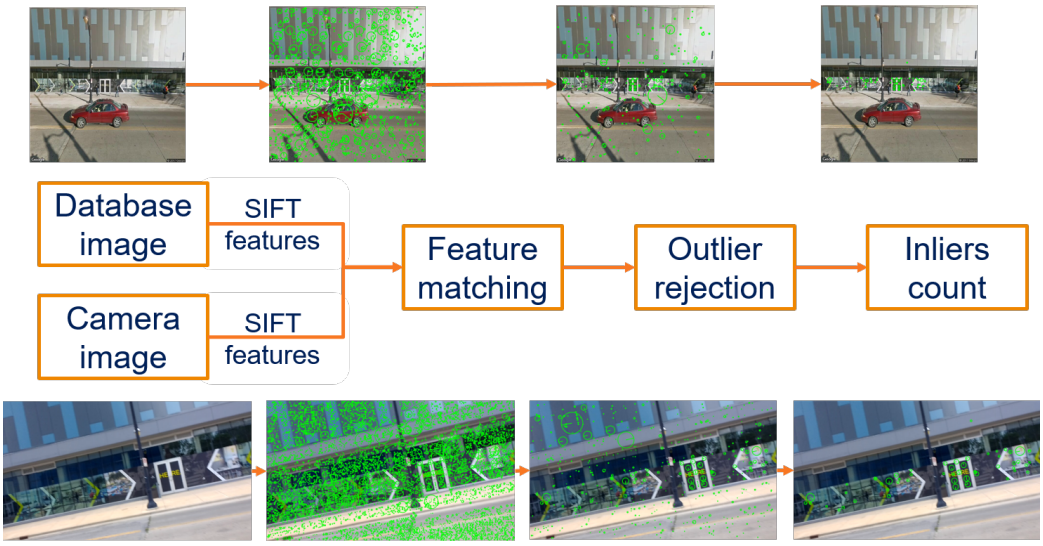


Figure 2.12: Demonstrating image matching algorithm for a single camera image.

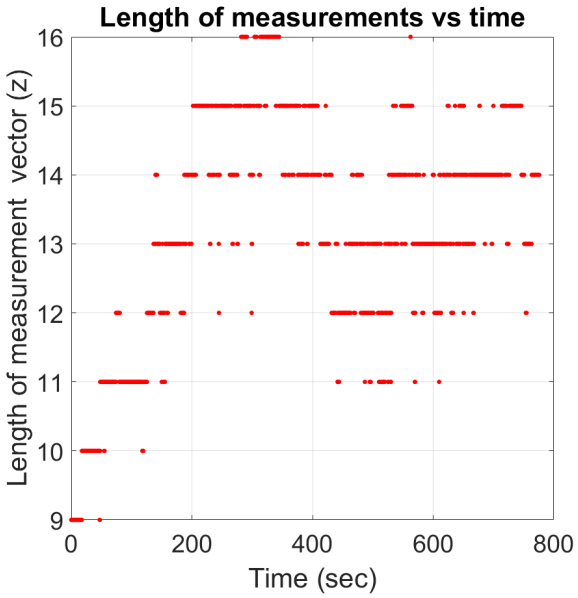


Figure 2.13: Variation in the size of received measurements with time.

surements. It demonstrates that noise in vision measurements varies with time, depending on the surrounding environment.

We implemented our AKF that adapts to time variation of noise and size variation of measurements. Figure 2.15 shows the variation of east and north position obtained from vision measurements and AKF. The noise in vision measurements changes around 450 seconds, and AKF positioning estimates are smooth compared to the vision measurements, showing AKF adapts to

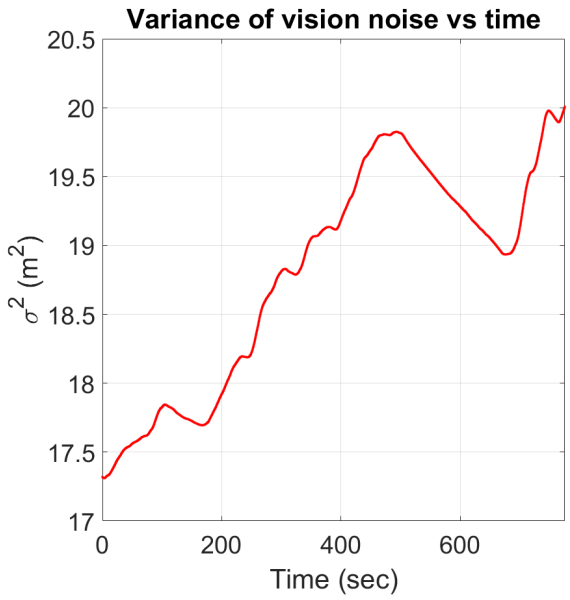


Figure 2.14: Variance of vision noise with time.

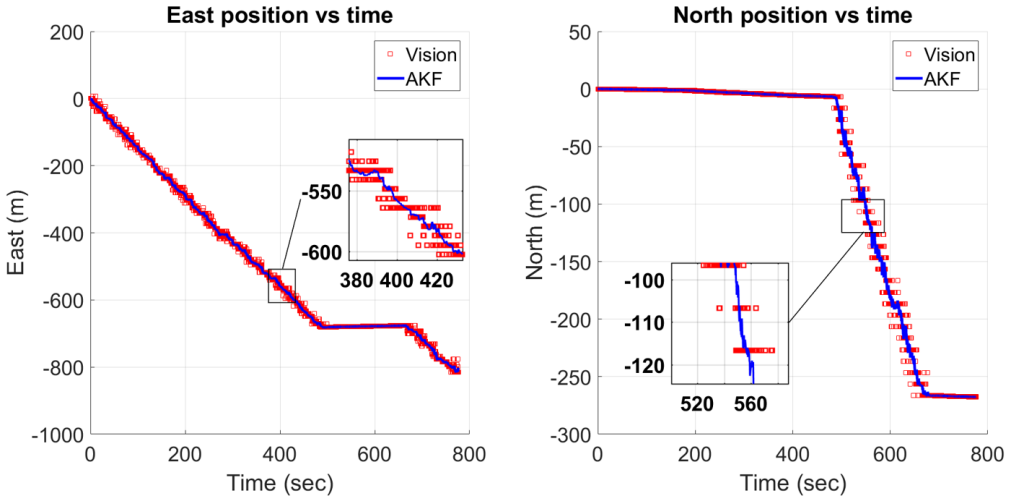


Figure 2.15: Variation of east and north position obtained from vision and AKF with time.

time variation of noise.

We compare the pseudorange only positioning estimate, obtained using LSE, EKF positioning estimate with constant covariance matrices, and AKF positioning estimate in Figure 2.16. It shows that AKF positioning estimates are closer to the ground truth than positioning estimates obtained from pseudorange only and EKF with constant covariance matrices..

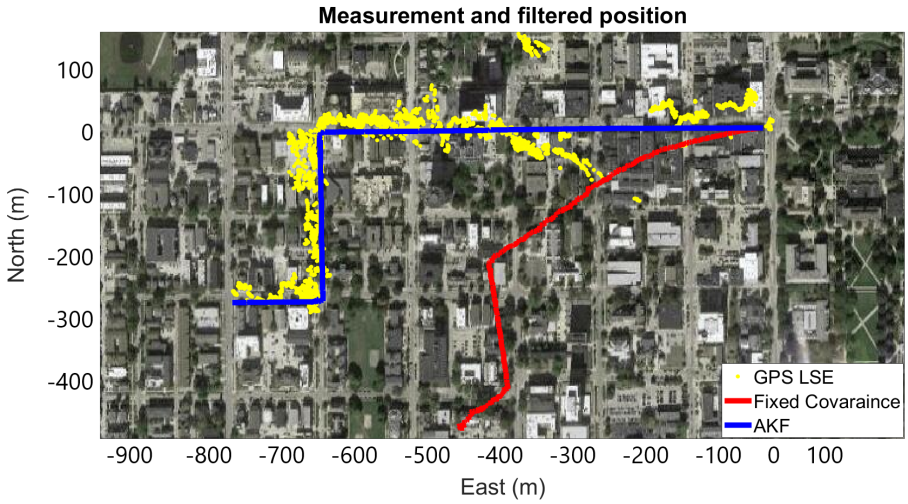


Figure 2.16: Positioning estimates obtained from LSE, EKF with constant covariance matrices, and AKF

2.7 Summary

In this chapter, we presented an AKF algorithm that adapts to time and size variation of noise. We discussed the implementation of a sensor fusion algorithm that fuses GPS and vision measurements. An image matching algorithm was devised, which provides a positioning estimate by matching a given image with GSV images. The developed AKF algorithm is tested in simulation as well as in the real world. We demonstrated that AKF positioning estimates are closer to ground truth, compared to the positioning estimates obtained from GPS only, vision only, and EKF with constant covariance matrices.

CHAPTER 3

INTEGRITY MONITORING FOR DIRECT POSITIONING

Critical infrastructures, such as the power grid, banking, and transportation system, use GPS timing, and positioning service to ensure safety [70]. Integrity [24]measures the trustworthiness of a navigation solution. It is one of the most critical requirements for safety-of-life applications. PLs are used to assess the integrity requirement for a system [25, 26, 27, 71]. PLs overbound positioning and timing errors by using error models.

Traditional receiver architectures, such as scalar tracking loop [1] and vector tracking loop [41], use two steps to provide a PVT solution. Pseudoranges are estimated in the first step, and trilateration is performed in the second step to obtain a PVT solution. PEMs are used to derive PLs. These models assume that each error component in PEMs is completely characterized by uni-modal symmetric Gaussian distribution. However, the error distribution is multi-modal due to changing environmental conditions [71]. Overbounding the multi-modal error distribution tails under such a scheme may result in the underbounding of errors due to asymmetry or bias in the distribution. This may cause a loss in integrity [27, 71].

One way to overbound multi-modal error distribution is to use multi-modal Gaussian distribution. In [71], the authors show that the error distribution is multi-modal and overbound the positioning errors using a bi-modal Gaussian distribution. Using this approach, the authors improve the overall availability of the system by 50%. This approach implicitly assumes that the number of modes present in the multi-modal error distribution is known. However, the number of modes in the error distribution is dependent on environmental conditions and is unknown.

DP [10, 11, 12, 17, 72, 20, 15, 73, 74, 75] is an unconventional GPS receiver architecture that directly operates in PVT domain. Compared to traditional two-steps methods, DP estimates PVT solution in a single step without estimating pseudoranges. Thus, DP removes errors arising from the

convolutions of PEMs.

A large amount of literature is available for overbounding positioning errors using PEMs [25, 26, 27, 71]; only a few are available for DP. With our best knowledge, there has not been a paper that empirically shows the error distribution for DP. We expect the error distribution to be multi-modal due to changing environmental conditions.

Prior work [29] on DP-based integrity monitoring discusses Solution-Separation Receiver Autonomous Integrity Monitoring framework. However, this framework is originally designed for PEMs. Another work [30] utilizes the correlation manifolds generated by DP to overbound the positioning errors. This approach is deeply coupled with the DP framework. However, the authors use many empirical parameters to overbound the vertical errors. These parameters are dependent on environmental conditions and may change with different environmental conditions.

As a starting step, we focus on estimating PLs for vertical errors. Compared to prior works, we develop a Bayesian algorithm for estimating Vertical Protection Levels (VPLs) using the DP framework and does not require empirical parameters to overbound vertical errors. The developed algorithm accounts for the unknown number of modes present in the vertical errors' multi-modal distribution. In this chapter, we describe our Bayesian algorithm for estimating VPLs. The remainder of this chapter is organized as follows: Section 3.1 provides an overview of a generic DP. In Section 3.2, we describe our Bayesian VPL estimation algorithm that is built on DP receiver architecture. Section 3.3 details the simulation environment and results for validating the developed Bayesian algorithm. Finally, Section 3.4 summarizes this chapter.

3.1 Overview of Direct Positioning

This section provides an overview of a generic DP receiver. The first subsection describes the mathematical formulation, and the second subsection provides details for DP's implementation.

3.1.1 Mathematical Formulation

The objective of DP [10] is to estimate PVT coordinates of a receiver $\mathbf{X}$ given the received signal Y, where:

$$\mathbf{X} = \begin{bmatrix} x & y & z & C\delta t & \dot{x} & \dot{y} & \dot{z} & C\dot{\delta t} \end{bmatrix}^\top = \begin{bmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \end{bmatrix} \tag{3.1}$$

$(C\delta t, C\dot{\delta t})$ denotes receiver specific clock bias and drift multiplied by speed of light. Also, $\mathbf{x} = \begin{bmatrix} x & y & z & C\delta t \end{bmatrix}^\top$ $\dot{\mathbf{x}} = \begin{bmatrix} \dot{x} & \dot{y} & \dot{z} & C\dot{\delta t} \end{bmatrix}^\top$. The PVT coordinates are in ECEF coordinate frame. The received signal at time t and at coordinate $\mathbf{X}$, after carrier wipe off is given by

$$Y(\mathbf{a}, \mathbf{X}, t) = \sum_{i}^{M} a^{(i)} g^{(i)}(t - \tau^i) \exp\left(j2\pi\Delta f^i t\right) + n(t) \tag{3.2}$$

where:

- $\mathbf{a} = \begin{bmatrix} a^{(1)} & a^{(2)} & \dots & a^{(M)} \end{bmatrix}^\top \in \mathbb{C}^M$ are the complex amplitudes of the visible satellites.
- $M \in \mathbb{N}$ is the number of visible satellites.
- $g^{(i)}$ is the L1 coarse acquisition (C/A) code of the i^{th} visible satellite.
- $\tau^{(i)}$ is the code delay of the i^{th} visible satellite:

$$\tau^{(i)} = \frac{||\mathbf{d}^{(i)}||}{C} + \left(\delta t - \delta t^{(i)}\right) \tag{3.3}$$

- $\Delta f^{(i)}$ is the carrier Doppler shift of the i^{th} visible satellite:

$$\Delta f^{(i)} = \frac{-f_{L1}}{C} \left\{ \frac{\mathbf{r}^{(i)} \dot{\mathbf{r}}^{(i)}}{||\mathbf{r}^{(i)}||} + C\left(\dot{\delta t} - \dot{\delta t}^{(i)}\right) \right\} \tag{3.4}$$

- $\mathbf{d}^{(i)} = \begin{bmatrix} x - x^{(i)} & y - y^{(i)} & z - z^{(i)} \end{bmatrix}^\top$ is the relative vector to the i^{th} visible satellite.
- $(\delta t^{(i)}, \dot{\delta t}^{(i)})$ are satellite specific clock bias and clock drift rate.

- $n(t) \in \mathcal{N}(0, \sigma^2) \in \mathbb{C}$ is an independent and identically distributed (i.i.d.) Gaussian process, same as the complex additive Gaussian nose (AWGN).

In DP, receiver's coordinates are obtained by maximizing the following likelihood

$$p(\mathbf{y}|\mathbf{a}, \mathbf{X}, \sigma^2) = \left(\frac{1}{\pi\sigma^2}\right)^N \exp\left\{-\frac{||\mathbf{y} - D\mathbf{a}||^2}{\sigma^2}\right\} \tag{3.5}$$

where:

- $\mathbf{y} = \begin{bmatrix} Y(\mathbf{a}, \mathbf{X}, t_1) & Y(\mathbf{a}, \mathbf{X}, t_2) & ... & Y(\mathbf{a}, \mathbf{X}, t_N) \end{bmatrix}^\top \in \mathbb{C}^N$ is a signal snapshot obtained over $\mathbf{t} = \{t_n\}_{n=1}^N$.
- $D(\mathbf{X}, \mathbf{t}) \in C^{\mathbb{N}\times M}$ is a matrix containing signal replicas of visible satellites for a given $\mathbf{X}$ and t.
- $\sigma^2 \in \mathbb{R}$ is the noise level of the receiver.
- $||\mathbf{b}||$ denotes $\mathbb{L}_2$ norm of a generic vector $\mathbf{b}$.

Under nice properties of D [75] and using the orthogonality principle, the likelihood [10] is simplified as

$$p(\mathbf{y}|\mathbf{X}, \sigma^2) = \left(\frac{1}{\pi\sigma^2}\right)^N \exp\left\{-\frac{||\mathbf{y}||^2 - \frac{1}{N}||D^*\mathbf{y}||^2}{\sigma^2}\right\} \tag{3.6}$$

The ML estimation is then obtained by

$$\hat{\mathbf{X}}_{ML} \approx \underset{\mathbf{X}}{\operatorname{argmax}} \frac{1}{N}\mathbf{y}^* DD^*\mathbf{y} = \underset{\mathbf{X}}{\operatorname{argmax}}\, \mathcal{R}(\mathbf{X}, \mathbf{t}) \tag{3.7}$$

where:

- $D^*, \mathbf{y}^*$ are conjugate transpose of D and $\mathbf{y}$ respectively.
- $\mathcal{R}(\mathbf{X}, \mathbf{t})$ denotes the correlation manifold obtained at coordinate $\mathbf{X}$ and time samples $\mathbf{t}$.

3.1.2 Implementation of Direct Positioning Receiver

Equations (3.3) and (3.4) show that time delay and Doppler shift are implicit functions of the receiver's PVT coordinates. Instead of first estimating time

delay and Doppler shift, the DP receiver directly estimates PVT coordinates by maximizing the correlation manifold. The implementation of a generic DP [76, 73, 74, 20] receiver is described below:

1. In the first step, the DP receiver generates a grid of candidates $\hat{\mathbf{X}}_m$. Each candidate represents a potential navigation solution and corresponds to a unique time delay and Doppler frequency shift. The grid is initialized in such a way to ensure $\mathbf{X}$ remains within the range of candidates.

2. DP generates an expected signal reception, $\hat{Y}_m$, for each candidate based on the PVT coordinates of the candidate.

$$\begin{aligned}\hat{Y}_m(\hat{\mathbf{X}}_m, t) = \sum_{i=1}^{M} g^{(i)}(t - \tau^{(i)}) \\ \exp\left\{j2\pi(f_{L1}t + \Delta f^{(i)}t + \phi^{(i)})\right\}\end{aligned} \tag{3.8}$$

   where f_{L1} is L1 carrier frequency (1575.42 MHz) and $\phi^{(i)}$ is the carrier phase of the i^{th} visible satellite. The signal synchronization parameters are derived from receiver coordinates $\hat{\mathbf{X}}_m$ using (3.3) and (3.4). Expected time sampled signal is given by time sampling of $\hat{Y}_m$ and is given below

$$\hat{\mathbf{y}}_m = \left[\hat{Y}_m(\hat{\mathbf{X}}_m, t_1) \quad \hat{Y}_m(\hat{\mathbf{X}}_m, t_2) \quad ... \quad \hat{Y}_m(\hat{\mathbf{X}}_m, t_N)\right]^\top \tag{3.9}$$

3. For each candidate, the receiver computes the cross-correlation between the expected reception $\hat{\mathbf{y}}_m$ and the received signal $\mathbf{y}$

$$\mathcal{R}(\hat{\mathbf{X}}_m, \mathbf{t}) = \text{corr}(\mathbf{y}, \hat{\mathbf{y}}_m) \tag{3.10}$$

   Correlation manifold is obtained by collectively obtaining correlation values over the grid. This manifold is typically unimodal, where the peak is at the candidate closest to the navigation solution. For illustration purpose, a typical correlation manifold is shown in Figure 3.1.

4. The navigation solution is obtained by selecting the candidate that has

the highest correlation value

$$\hat{\mathbf{X}}_{DP} = \underset{m}{\operatorname{argmax}}\, \mathcal{R}(\hat{\mathbf{X}}_m, \mathbf{t}) \tag{3.11}$$

where $\hat{\mathbf{X}}_{DP}$ denotes the estimated navigation solution provided by DP. This estimate is utilized in the next time step for populating the grid candidates, that is performed in step 1.

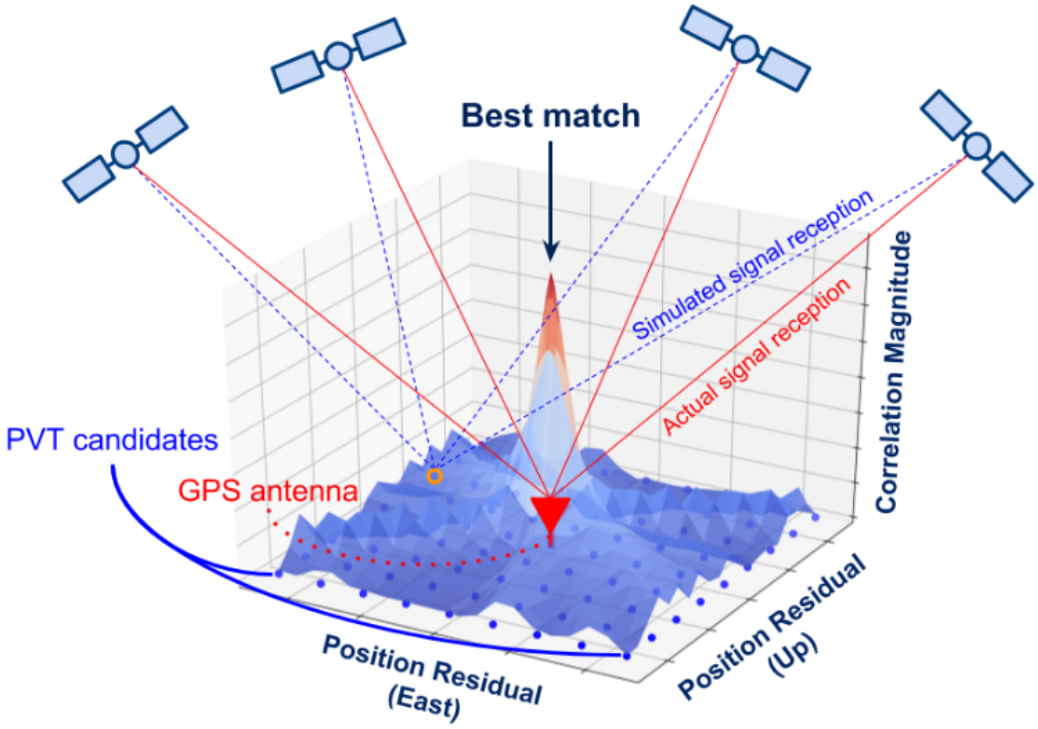


Figure 3.1: 2D example showing the correlation manifold $\mathcal{R}$ on the local East-Up plane [30]. The best match denotes the candidate closest to the receiver.

The high dimensional search space for $\mathbf{X}$ is decoupled into two subspaces: position and clock bias, $\mathbf{x}$ and velocity and clock drift, $\dot{\mathbf{x}}$. The decoupling is similar to Space Alternating Generalized Expectation (SAGE) algorithms, as discussed in [20, 77, 78]. This decoupling is performed to reduce the computation cost.

3.2 Bayesian Approach to Estimate Protection Levels

Figure 3.2 shows the overall architecture of the Bayesian VPL estimation algorithm in which we perform three steps for VPL estimation. In the first step, we estimate PVT coordinates and noise variance from the received raw signal snapshots, $\mathbf{y}$, using DP receiver. Next, we obtain Altitude Likelihood Manifold (ALM) by creating additional candidates in local up direction around the PVT coordinate and calculating the correlation manifold at the created candidates. Finally, we utilize ALM in Bayesian VPL Estimation to estimate VPL. The subsequent subsections details each of the above-mentioned steps.

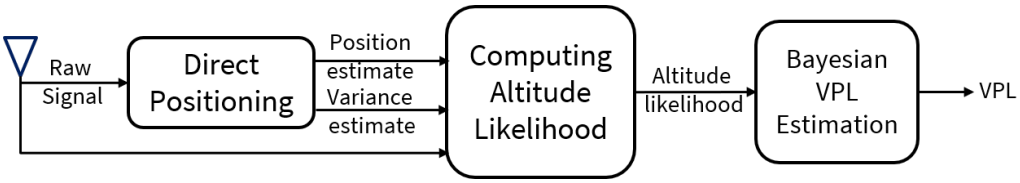


Figure 3.2: Overall architecture of the Bayesian VPL estimation algorithm

3.2.1 Direct Positioning

In the first step, we use a generic DP receiver architecture to estimate the receiver's PVT coordinates, $\hat{\mathbf{X}}_{DP}$. We model the vertical errors with time-varying Gaussian distribution and perform ML estimation on (3.6) to estimate the variance of the time-varying Gaussian distribution, $\hat{\sigma}^2$. The variance estimate is given by [75]

$$\hat{\sigma}^2 = \frac{||y||^2 - \mathcal{R}(\hat{\mathbf{X}}_{DP}, \mathbf{t})}{N} \tag{3.12}$$

Conventional DP integrity monitoring methods use PVT estimates only; however, we utilize both PVT and noise variance estimate in the developed algorithm. In the next step, we apply these estimates for computing ALMs.

3.2.2 Compute Altitude Likelihood

In Section 3.1, we described the implementation of a generic DP that provides two decoupled 4-dimensional correlation manifolds: one in position and clock bias subspace and another in velocity and clock drift subspace. Our primary

focus in this chapter is to obtain VPL for a given PVT estimate. We achieve this by converting the 4-dimensional correlation manifold, in position and clock bias subspace, into a 1-dimensional correlation manifold in the local up direction, i.e., ALM. The following steps are involved in obtaining ALM from the 4-dimensional correlation manifold:

1. We transform the DP's PVT estimate, $\hat{\mathbf{X}}_{DP}$, to local East-North-UP (ENU) frame using a reference coordinate. Then, we generate candidates z_j in local up direction around the reference coordinate.

2. Signal replicas are generated for these candidates. Note that for these replicas, PVT coordinates are same as $\hat{\mathbf{X}}_{DP}$ except for the local z coordinate.

$$\hat{\mathbf{y}}_{z_j} = \begin{bmatrix} \hat{Y}_{z_j}(z_j, t_1) & \hat{Y}_{z_j}(z_j, t_2) & \dots & \hat{Y}_{z_j}(z_j, t_N) \end{bmatrix}^\top \tag{3.13}$$

   where $\hat{\mathbf{y}}_{z_j}$ denotes the time sampled expected signal at local coordinate z_j.

3. Next, we perform cross-correlation between the received signal and the expected signal to obtain altitude correlation manifold.

$$\mathcal{R}(z_j, \mathbf{t}) = \text{corr}(\mathbf{y}, \hat{\mathbf{y}}_{z_j}) \tag{3.14}$$

4. Finally, we use altitude correlation manifold along with the estimate of noise variance to compute ALM.

$$p(\mathbf{y}|z_j) = \left(\frac{1}{\pi\hat{\sigma}^2}\right)^N \exp\left\{-\frac{||\mathbf{y}||^2 - \frac{1}{N}\mathcal{R}(z_j, \mathbf{t})}{\hat{\sigma}^2}\right\} \tag{3.15}$$

For illustration purpose, a sample 2-dimensional correlation manifold and the corresponding 1-dimensional correlation manifold in local up direction are shown in Figures 3.3 and 3.4 respectively. In the final step, we utilize ALM to estimate VPL.

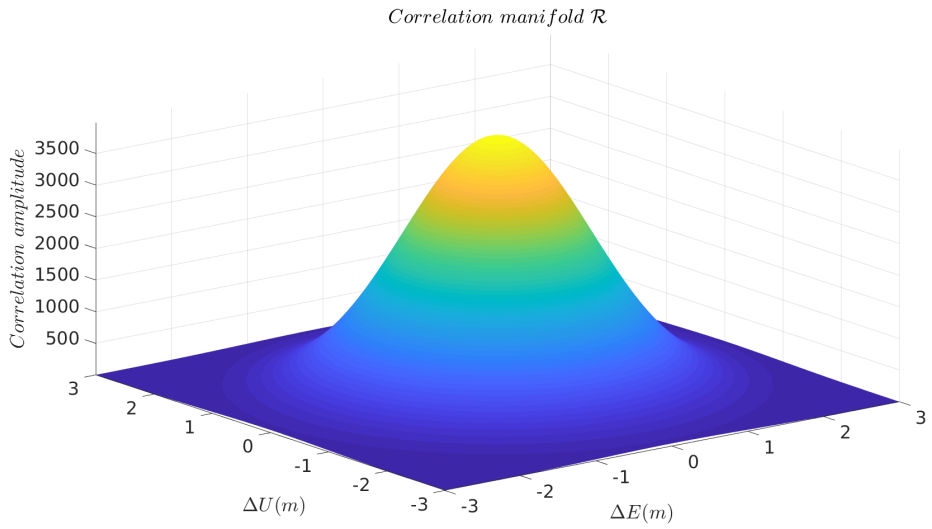


Figure 3.3: Sample 2-dimensional correlation for the candidates in local East and Up direction

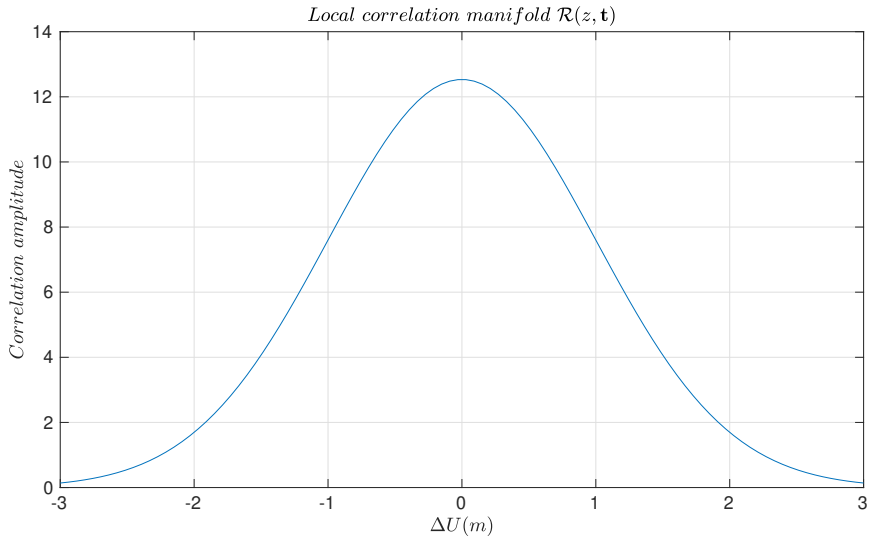


Figure 3.4: Sample 1 dimensional correlation for the candidates in local Up direction

3.2.3 Bayesian VPL Estimation

In the literature, VPL is defined using posterior probability and integrity risk requirements. Therefore, in the final step, we obtain the posterior probability from likelihood by applying Bayes theorem and then estimate VPL.

Estimation of Posterior Probability

According to Bayes theorem, the posterior probability is given by

$$p(z|\mathbf{y}) = \frac{p(\mathbf{y}|z)p(z)}{p(\mathbf{y})} \tag{3.16}$$

To simplify computations, we assume the prior probability, $p(z)$, to be uniformly distributed over the generated local z_j candidates. The normalizing probability, $p(\mathbf{y})$, is obtained by applying the law of total probability

$$p(\mathbf{y}) = \int p(\mathbf{y}|z)p(z)dz \tag{3.17}$$

The expression of $p(\mathbf{y}|z)$ contains non-linear terms and the closed-form solution of (3.17) is unavailable. We numerically integrate (3.17) using the Euler method, over the set of candidates. The posterior probability under such scheme of integration is given by

$$p(z|\mathbf{y}) = \frac{\exp\left\{\frac{\mathcal{R}(z,\mathbf{t})}{\hat{\sigma}^2 N}\right\}}{\sum_j \left(\exp\left\{\frac{\mathcal{R}(z_j,\mathbf{t})}{\hat{\sigma}^2 N}\right\}\right)\Delta z_j} \tag{3.18}$$

where:

- $\mathcal{R}(z, \mathbf{t})$ denotes altitude correlation manifold at z and time samples $\mathbf{t}$.
- Δz_j denotes the candidates' separation.

The summation in the denominator is taken over all the generated candidates z_j. The log-sum-exp trick is used to avoid numerical overflow errors. This trick is widely used by deep learning researchers for training the neural network [79]. The following equation explains this trick:

$$\log\left(\sum_k \exp(b_k)\right) = \log\left(\exp(b_{max})\right) + \log\left(\sum_k \exp(b_k - b_{max})\right) \tag{3.19}$$

where b_k is the k^{th} element of some vector $\mathbf{b}$ and $b_{max} = \max_k(\mathbf{b})$, i.e., the maximum element of $\mathbf{b}$. This trick is used to avoid numerical overflow errors that may occur in numerical computation.

Estimation of VPL

For a given integrity risk probability ϵ, VPL is defined by [71]

$$p(|z - \hat{z}| > VPL|\mathbf{y}) < \epsilon \tag{3.20}$$

The task is to estimate $\hat{z}$ and an interval I that satisfy (3.20). Equation (3.20) is simplified to [71]

$$\int_{z \in I} p(z|\mathbf{y})dz \leq \epsilon \tag{3.21}$$

There is no unique way to find the interval I. For simplicity, we assume that the probability of being on each side of the interval is $\epsilon/2$. Under this assumption, we numerically solve the inequality in (3.21) to obtain the interval I. Then, we compute upper and lower bounds for a given PVT estimate and thus obtain VPL. Figure 3.5 illustrates the steps used to obtain VPL by solving (3.21).

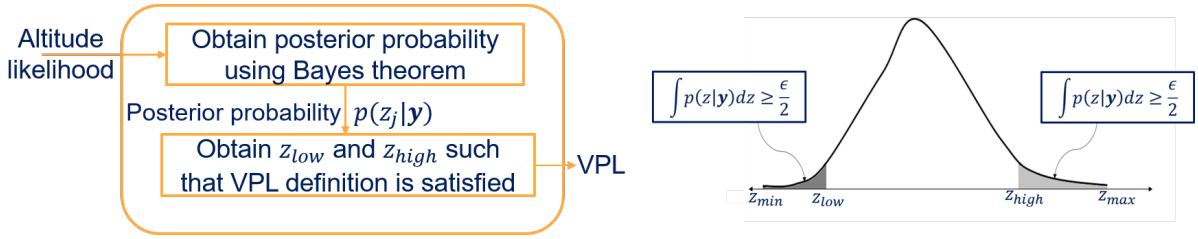


Figure 3.5: VPL estimation illustration using posterior probability

In Figure 3.5, z_{min}, z_{max} denotes the minimum and maximum value of the generated local candidates, respectively. We incrementally integrate the posterior probability from both ends of the distribution. At each incremental integration, we check if the integrated probability exceeds $\epsilon/2$. The first local z_j candidate for which condition specified in (3.21) is not satisfied, corresponds to upper and lower bounds. In Figure 3.5, these extremes are denoted by z_{low} and z_{high}. The estimated VPL is given by

$$VPL = |z_{high} - z_{low}| \tag{3.22}$$

where || denotes absolute value. Under similar Euler numerical integration scheme, the integration of posterior probability is given by

$$\int_{z\in I} p(z|\mathbf{y})dz \approx \frac{\sum_i \left(\exp\left\{\frac{\mathcal{R}(z_l,\mathbf{t})}{\hat{\sigma}^2 N}\right\}\right)\Delta z_l}{\sum_j \left(\exp\left\{\frac{\mathcal{R}(z_j,\mathbf{t})}{\hat{\sigma}^2 N}\right\}\right)\Delta z_j} \tag{3.23}$$

where l is such that $z_l \in I$, Δz_l denotes the candidates' separation and j is such that it covers all the generated candidates, i.e., the sum in the denominator is taken over all candidates.

3.3 Simulation Environment and Results

Real-life outdoor experiments involve uncontrolled conditions. These conditions often are undetectable and difficult to analyze. Obtaining ground-truth in such situations becomes a challenging task. In order to avoid uncontrolled situations and the problem of obtaining ground truth, we decide to work with a simulated dataset. We use a high-fidelity GPS simulator to generate GPS raw signals. The simulator provides the functionality of creating custom motion trajectories with adjustable individual satellite power levels and transmitting it directly to the receiver.

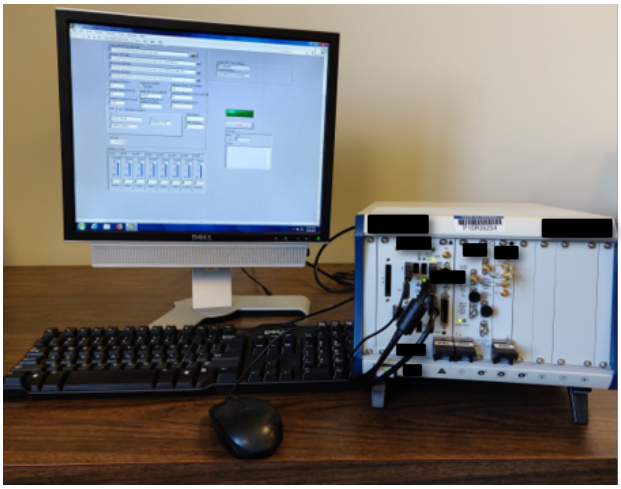


Figure 3.6: High fidelity GPS simulator

The simulator is capable of simulating up to 12 GPS satellites. The number GPS satellites are selected based on the almanac and ephemeris files, GPS time, and receiver position that a user specifies. The simulator validates several satellite parameters before simulating the satellites. If the satellite parameters are valid at the specified GPS time, the simulator selects the

satellite and uses it to simulate a GPS signal. The simulator continuously updates the Doppler shifts between the satellites and simulated receiver so that the simulated signal is as close to the specified position as possible.

We generate 24 hours of stationary GPS data using this simulator. At each time step, we use 20×10^{-3} seconds of data to get PVT coordinates using DP. The sampling frequency for the generated data is 2.5 MHz, providing us with four million positioning data points. These positioning data points are used to obtain the distribution of vertical positioning errors.

We implement our developed algorithm on *pyGNSS*, Python-based Software-Defined-Radio (SDR) research suite [76, 80]. The software suite provides the flexibility to test and verify new GPS receiver algorithms. It is also capable of analyzing raw GPS signal samples. Table 3.1 lists the parameters used in DP's implementation. Additional candidates generated to obtain ALM are spaced 1 cm apart if they are within 5 m of the estimated position. Otherwise, the spacing between the candidates is 3.5 m.

Table 3.1: Grid samples for DP's implementation

Domain	**Axis**	**Span**	**Spacing**	**Dim**
Position	East,North,Up	$[-351.6, 351.6]$ (m)	46.88 (m)	15
Velocity	East,North,Up	$[-17.58, 17.59]$ (m/s)	2.34 (m/s)	15
Time	$C\delta t$	$[-351.6, 351.6]$ (m)	46.88 (m)	15
Time Drift	$C\delta\dot{t}$	$[-0.22, 0.22]$ (m/s)	0.03 (m/s)	15

3.3.1 Validating Multi-modal Distribution

We obtain positioning estimates from DP on the generated GPS dataset. We compute the vertical positioning error for each DP's positioning estimate to get the vertical positioning error distribution. Figure 3.7 shows the histogram for vertical positioning errors.

Figure 3.8 displays a quantile-quantile (QQ) plot. The plot's x-axis denotes quantiles taken from a standard normal distribution, and the y-axis of the plot denotes quantiles taken from vertical positioning error distribution. The deviation of vertical positioning error distribution from the red dotted line shows that the distribution is non-Gaussian. Figure 3.7 verifies this observation, where the vertical positioning error distribution is multi-modal. There are roughly four million data points present in Figures 3.7

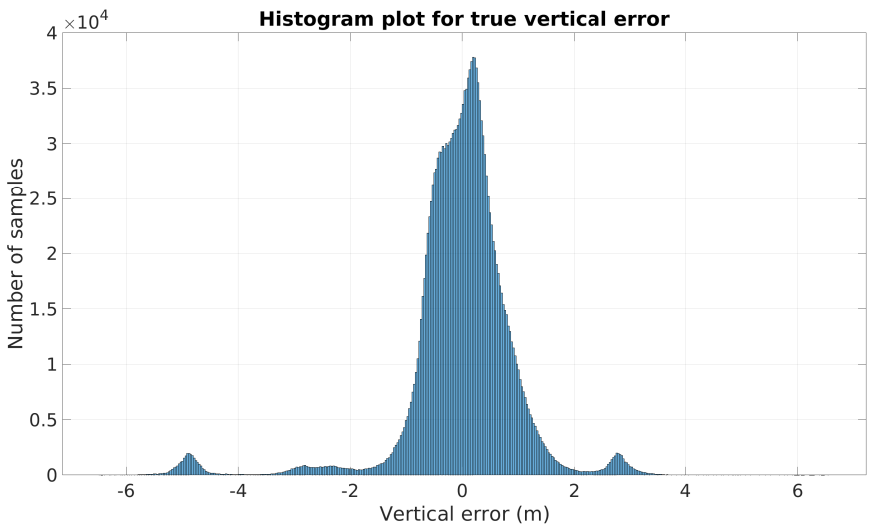


Figure 3.7: DP's vertical positioning error has a multi-modal distribution.

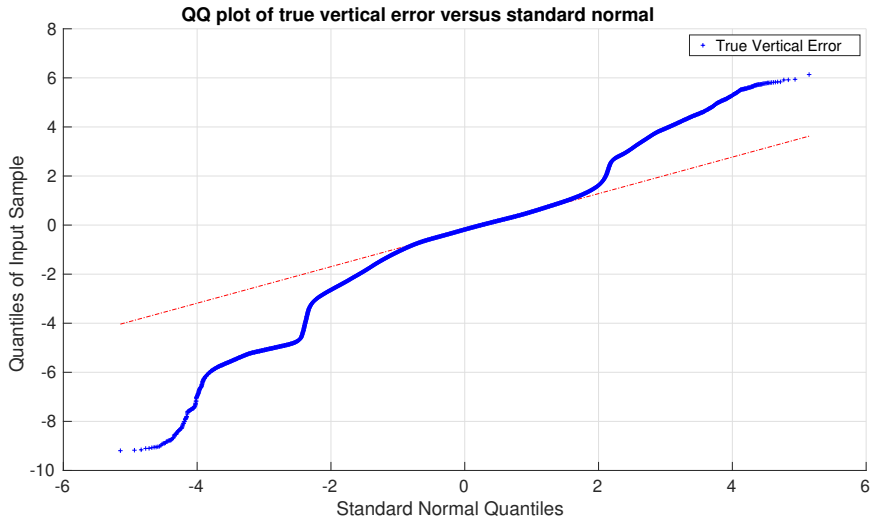


Figure 3.8: DP's vertical positioning error distribution is non-Gaussian as the blue crosses lie outside the red line in the QQ plot.

and 3.8. Overbounding of multi-modal distribution using a uni-modal Gaussian distribution may reduce performance or loss in integrity for a navigation system.

3.3.2 Estimated VPL

The developed Bayesian VPL estimation algorithm, described in Section 3.2, is implemented and tested on the generated dataset. Figure 3.9 shows the estimated VPL with red circles and vertical errors with black circles. The bottom subplot in Figure 3.9 shows a zoomed-in version of the top subplot. We compare the distribution of the estimated VPL and vertical errors in Figure 3.10 and 3.11.

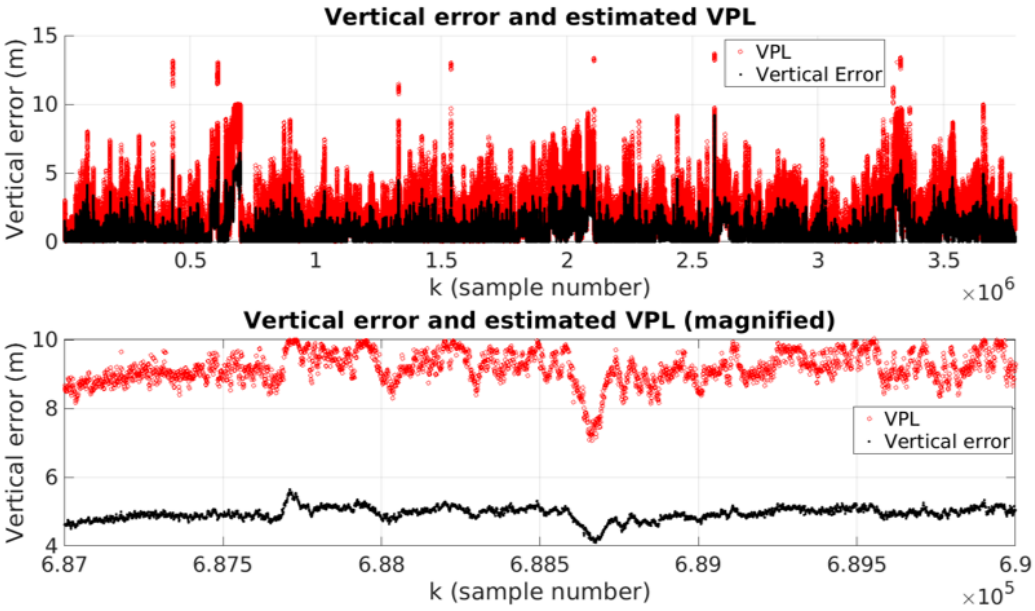


Figure 3.9: Estimated VPL overbounds the vertical positioning errors

Figure 3.11 shows that the distributions of estimated VPL and vertical positioning errors are similar for errors smaller than $6m$. Therefore, the derived algorithm is able to capture the multi-modal behavior. Even for larger error, the estimated VPL overbounds the vertical errors. This is evident from the histogram plot shown in Figure 3.12.

3.4 Summary

This chapter briefly gave an overview of DP and described a Bayesian algorithm for estimating VPL for DP, whose vertical error distribution has not been shown in the literature. The developed algorithm is tested on 24 hours of GPS dataset, which is generated using a high-

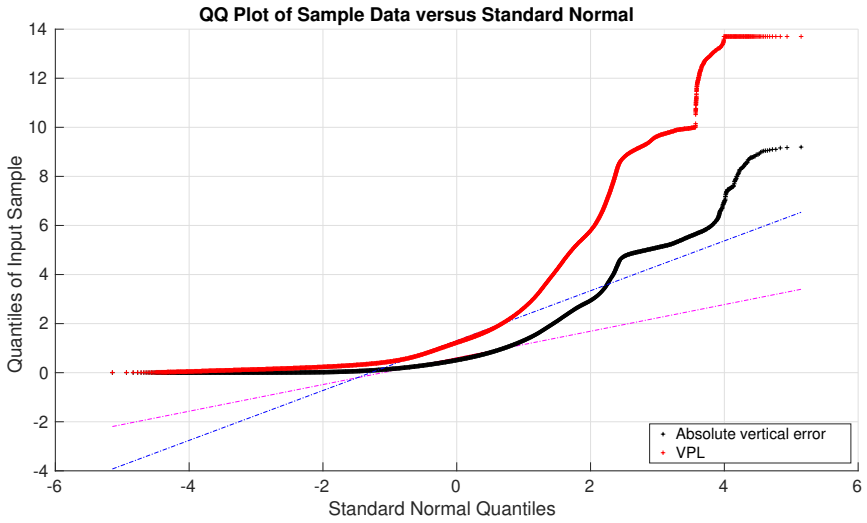


Figure 3.10: QQ plot of vertical positioning errors and estimated VPL vs standard normal

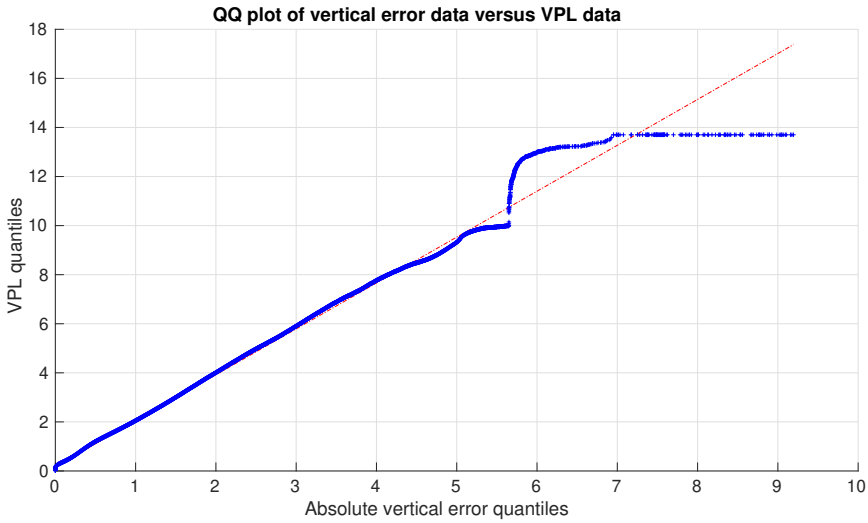


Figure 3.11: QQ plot of vertical positioning errors vs estimated VPL shows that vertical errors and estimated VPL have similar distribution for small vertical errors.

fidelity GPS simulator. We showed that the DP's vertical positioning error has a multi-modal distribution by obtaining 4 million positioning error data points from the generated dataset. The developed algorithm bounds the vertical errors and is insensitive to the unknown number of modes present in the distribution of vertical errors.

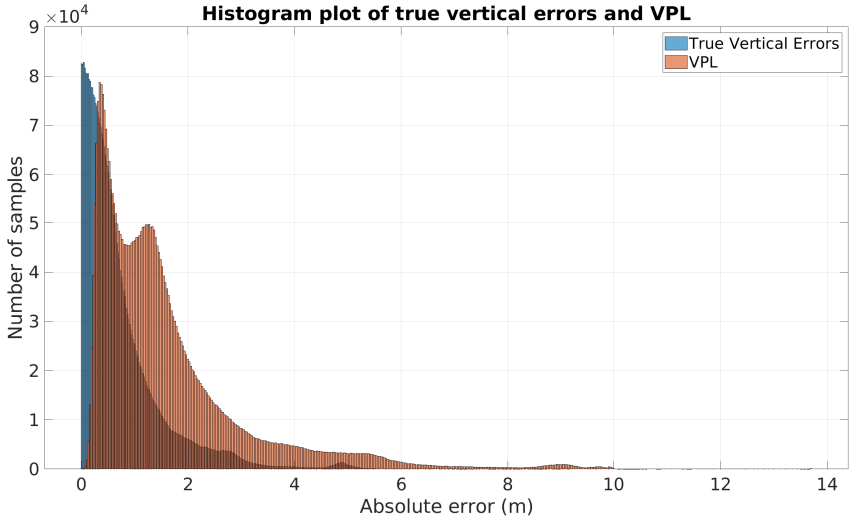


Figure 3.12: Histogram of vertical positioning errors and estimated VPL shows that estimated VPL bounds the vertical errors.

CHAPTER 4

GPS SPOOFING-RESILIENT STATIC STATE ESTIMATION FOR THE POWER GRID

The installation of PMUs is a step towards achieving wide-area situational awareness for the power grid. PMUs utilize GPS signals for time synchronization and are vulnerable to GSAs as civilian GPS signals are unencrypted and have low signal power. Studies [43, 44, 45] show that GSAs are feasible. GSAs shift the phase angle of the PMU measurement, and the phase shift is referred as attack angle [42]. The work in [44] and [46] demonstrate that an SE raises false warnings of power stability and provide erroneous flow estimates. Furthermore, GSAs are capable of violating the IEEE C37.118.1-2011 standard [45]. GSA detection and mitigation are critical to ensure the safe operation of the power grid.

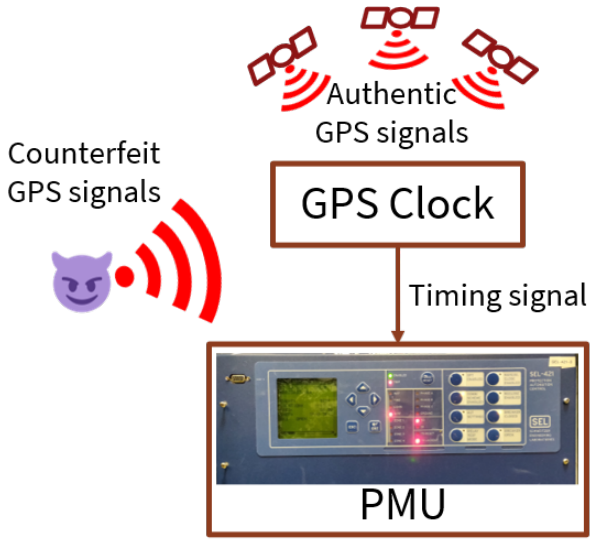


Figure 4.1: Conventional GPS receiver locks on counterfeit GPS signals under GSAs, providing incorrect timing to PMUs and making them vulnerable during GSAs.

In the literature, a theoretical analysis of GSAs' impact on PMU measurement residuals has not been provided. This analysis will identify if certain

GSAs can alter the power grid states without increasing the residual norm. The key objectives of this chapter are as follows:

1. Perform theoretical analysis to examine the impact of GSAs on PMU measurement residuals.

2. Provide an algorithm for correcting PMU measurements under multiple GSAs with different attack angles.

For the power grid, PMUs are stationary with a known position. GSAs that alter the receiver position would be detectable compared to timing GSAs that alter the receiver time without changing the receiver position. Timing GSAs pose a threat to the safe operations of the power grid. In this chapter, we consider timing GSAs and refer to them as GSAs only. This chapter's overall goal is to develop an SE that is resilient to multiple GSAs with different attack angles. We propose a novel residual-based SR-SSE for the power grid that is resilient to multiple GSAs. SR-SSE estimates voltage phasors using PMU measurements. During one or multiple GSAs, SR-SSE iteratively minimizes the residual norm to provide resilient voltage phasors. We derive a necessary condition for the proposed estimator to show that the residual norm increases during GSAs. We perform MC simulations to verify the derived necessary condition. We further simulate the IEEE 14, IEEE 39, IEEE 118, and Illinois 200-bus test systems to validate SR-SSE against multiple GSAs. The remainder of this chapter is organized as follows: Section 4.1 presents a conventional PMU-based Static State Estimator (SSE) for the power grid along with the spoofing attack measurement model. In Section 4.2, we provide details of our theoretical analysis and the derived necessary condition. Section 4.3 describes our proposed estimator and explains the simulation environment and implementation. Section 4.4 contains simulation results, and the summary of the chapter is presented in Section 4.5.

4.1 Background

The SE estimates the complex voltages of the buses present in a power grid network using SCADA or PMU measurements. Conventionally, an SSE is

used for estimating power grid states. The voltages in the power grid vary slowly, and the measurements at a given time instance are utilized in SSE. Depending on the type of measurements, different SSEs are available in the literature [34], such as SCADA-based SSEs [32], PMU-based SSEs [81], or SSEs which incorporate both SCADA and PMU measurements [82]. In this chapter, we focus on PMU-based SSE.

4.1.1 PMU-based SSE

Apart from having a faster update rate and synchronized measurements, PMUs directly measure voltage and current phasors. This results in a linear relationship between states of the grid and the PMU measurements. Consider a power grid network of N buses/nodes with M PMUs installed to ensure the network is observable. The system state $\mathbf{x} \in \mathbb{R}^{2N\times 1}$ can be written as

$$\begin{aligned}\mathbf{x} =&[Re(U_1), \cdots, Re(U_i), \cdots, Re(U_N),\\ &Im(U_1), \cdots, Im(U_i), \cdots, Im(U_N)]^\top\end{aligned} \tag{4.1}$$

where $Re(\cdot)$ denotes the real part, $Im(\cdot)$ is the imaginary part, and U_i is the complex voltage of the i^{th} bus. The PMU measurements at bus i, which is connected to k different buses, are given by

$$\begin{aligned}\mathbf{z}_i =&[Re(U_i), Im(U_i), Re(I_{i1}), \cdots,\\ &Re(I_{ik}), Im(I_{i1}), \cdots, Im(I_{ik})]^\top\end{aligned} \tag{4.2}$$

where $Re(I_{ik}), Im(I_{ik})$ are the real and imaginary parts of the complex current injected into line (i, k). The measurement model for the PMU at bus i is written as

$$\mathbf{z}_i = \mathbf{H}_i\mathbf{x} + \boldsymbol{\eta}_i \tag{4.3}$$

where $\mathbf{z}_i$ denotes the PMU measurements at bus i, $\mathbf{H}_i$ is the regression matrix associated with bus i, and $\boldsymbol{\eta}_i$ is assumed to be zero-mean Gaussian noise. In MATPOWER [83], a branch line is approximated using a π model. The regression matrix relates the complex current flowing in a line with the complex voltages at the buses of the π model. The construction of the regression matrix is given in [84, 83].

In order to have a concise representation, PMU measurements are verti-

cally stacked to create a total measurement vector $\mathbf{z}$. The overall PMUs' measurements are given by

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta} \tag{4.4}$$

where $\mathbf{z} = [\mathbf{z}_1, \cdots, \mathbf{z}_M]^\top$, $\mathbf{H} = [\mathbf{H}_1, \cdots, \mathbf{H}_M]^\top$, and $\boldsymbol{\eta} = [\boldsymbol{\eta}_1, \cdots, \boldsymbol{\eta}_M]^\top$.

In state estimation, the problem is to determine the unknown states, $\mathbf{x}$, using the available measurements, $\mathbf{z}$. Under the assumption of full observability, the least squares solution is given by

$$\hat{\mathbf{x}} = \left(\mathbf{H}^\top \mathbf{H}\right)^{-1} \mathbf{H}^\top \mathbf{z} \tag{4.5}$$

where $\hat{\mathbf{x}}$ is the estimated state of the network.

4.1.2 Spoofing Attack Model

Timing GSAs induce a time delay that shifts the phase angle of PMU measurements by an attack angle. Without loss of generality, assume the PMU at bus i is spoofed. During a timing GSA, the PMU measurements at bus i are modified as

$$\begin{aligned} \mathbf{z}_i^{spf} =& [|U_i| \cos(\theta_i + \Delta\theta_i), |U_i| \sin(\theta_i + \Delta\theta_i), \\ & |I_{i1}| \cos(\theta_{i1} + \Delta\theta_i), |I_{i1}| \sin(\theta_{i1} + \Delta\theta_i), \cdots, \\ & |I_{ik}| \cos(\theta_{ik} + \Delta\theta_i), |I_{ik}| \sin(\theta_{ik} + \Delta\theta_i)]^\top \end{aligned} \tag{4.6}$$

where $\mathbf{z}_i^{spf}$ denotes the spoofed measurements, θ_i is the phase angle at bus i, θ_{ik} is the phase angle for the line (i, k), and $\Delta\theta_i$ is the attack angle. The work in [46] demonstrates that timing GSAs introduce a constant attack angle to all PMU measurements as shown in (4.6). The attack angle at bus i is related to the induced time delay by

$$\Delta\theta_i = 2\pi f \Delta t_i \tag{4.7}$$

where f denotes the frequency of the system and Δt_i is the induced time delay. Using cosine identities, a linear relationship is obtained between spoofed and authentic measurements [42]

$$\mathbf{z}_i^{spf} = \boldsymbol{\gamma}_i \mathbf{H}_i \mathbf{x} + \boldsymbol{\eta}_i \tag{4.8}$$

where $\boldsymbol{\gamma}_i$ is a block diagonal matrix with the following sub-matrix

$$G = \begin{bmatrix} \cos(\Delta\theta_i) & -\sin(\Delta\theta_i) \\ \sin(\Delta\theta_i) & \cos(\Delta\theta_i) \end{bmatrix} \tag{4.9}$$

The PMU measurements are vertically stacked and (4.8) is re-written as

$$\mathbf{z}^{spf} = \boldsymbol{\Gamma}\mathbf{H}\mathbf{x} + \boldsymbol{\eta} \tag{4.10}$$

where $\mathbf{z}^{spf} = [\mathbf{z}_1^{spf}, \cdots, \mathbf{z}_M^{spf}]^\top$ and $\boldsymbol{\Gamma}$ is given by

$$\boldsymbol{\Gamma} = \begin{bmatrix} \mathbf{I}_1 & & & & 0 \\ & \ddots & & & \\ & & \boldsymbol{\gamma}_i & & \\ & & & \ddots & \\ 0 & & & & \mathbf{I}_M \end{bmatrix} \tag{4.11}$$

where $\mathbf{I}$ denotes an identity matrix. Without knowledge of $\boldsymbol{\Gamma}$, the SSE under timing GSAs will produce the following state estimates

$$\begin{aligned} \hat{\mathbf{x}}^{spf} &= \left(\mathbf{H}^\top\mathbf{H}\right)^{-1}\mathbf{H}^\top\mathbf{z}^{spf} \\ &= \left(\mathbf{H}^\top\mathbf{H}\right)^{-1}\mathbf{H}^\top\boldsymbol{\Gamma}\mathbf{z} = \mathbf{H}^\dagger\boldsymbol{\Gamma}\mathbf{z} \end{aligned} \tag{4.12}$$

where $\mathbf{z}$ denotes the PMU measurements under the nominal scenario and $\mathbf{H}^\dagger$ represents $\left(\mathbf{H}^\top\mathbf{H}\right)^{-1}\mathbf{H}^\top$. During the nominal scenario, none of the PMU buses are spoofed and the SSE state estimates match with power flow analysis. Under the spoofing scenario, one or more PMU buses are spoofed with different attack angles. The PMU measurements under the nominal and spoofed scenarios are related by following equation

$$\mathbf{z}^{spf} = \boldsymbol{\Gamma}\mathbf{z} \tag{4.13}$$

In the next section, we analyze the impact of GSAs on residuals and derive a necessary condition that ensures an increase in residual norm during GSAs.

4.2 Residual Statistics Under GSAs

In the power grid, bad data are detected using measurement residuals [49, 50, 51, 52]. The majority of the literature assumes the bad data to be of additive nature. However, the GSA introduces bad data of multiplicative nature as shown in Section 4.1.2. Some additive bad data do not increase the residual norm and are not detectable using residual-based detection algorithms [49]. Given that certain additive attacks are not detectable, we analyze the impact of GSAs on the residuals and derive a necessary condition showing that the residual norm increases under GSAs. We analyze residual distributions for two scenarios:

1. Nominal scenario: In this scenario, none of the PMU buses are spoofed and the residual is given by

$$\begin{aligned} \mathbf{r} &= \mathbf{z} - \mathbf{H}\hat{\mathbf{x}} \\ &= \mathbf{z} - \mathbf{H}\mathbf{H}^{\dagger}\mathbf{z} \end{aligned} \tag{4.14}$$

   where $\mathbf{r}$ denotes the PMU measurement residual vector. From (4.4), the expectation of $\mathbf{z}$ is $\mathbf{Hx}$. Using this and taking the expectation of the above equation yields

$$\begin{aligned} \mathbb{E}[\mathbf{r}] &= \mathbb{E}[\mathbf{z}] - \mathbf{H}\mathbf{H}^{\dagger}\mathbb{E}[\mathbf{z}] \\ &= \mathbf{Hx} - \mathbf{H}\mathbf{H}^{\dagger}\mathbf{Hx} \\ &= \mathbf{Hx} - \mathbf{H}\left(\mathbf{H}^{\top}\mathbf{H}\right)^{-1}\mathbf{H}^{\top}\mathbf{Hx} = 0 \end{aligned} \tag{4.15}$$

   where $\mathbb{E}[\cdot]$ is the expectation operator. The above equation shows that under the nominal scenario, the expectation of the residual vector is zero.

2. Spoofing scenario: In this scenario, one or more PMU buses are spoofed and the residual is given by

$$\begin{aligned} \mathbf{r}^{spf} &= \mathbf{z}^{spf} - \mathbf{H}\hat{\mathbf{x}}^{spf} \\ &= \mathbf{\Gamma z} - \mathbf{H}\mathbf{H}^{\dagger}\mathbf{\Gamma z} \end{aligned} \tag{4.16}$$

   where $\mathbf{r}^{spf}$ is the PMU measurement residual vector under GSA. To simplify notations, let m_{us} denote the set of PMUs that are not spoofed

and let m_s represent the set of PMUs that are spoofed. Now, the attacked measurements can be re-written as

$$\mathbf{z}^{spf} = \begin{bmatrix} \mathbf{I}_{m_{us}} & 0 \\ 0 & \boldsymbol{\gamma}_{m_s} \end{bmatrix} \mathbf{z} = \begin{bmatrix} \mathbf{z}_{m_{us}} \\ \boldsymbol{\gamma}_{m_s} \mathbf{z}_{m_s} \end{bmatrix} \tag{4.17}$$

where $\mathbf{z}_{m_{us}}$ denotes original measurements corresponding to unspoofed PMUs and $\mathbf{z}_{m_s}$ represents original measurements corresponding to spoofed PMUs. Note that $\mathbf{z} = [\mathbf{z}_{m_{us}}, \mathbf{z}_{m_s}]^\top$, which is a nominal case measurement. Using this, (4.16) is simplified as

$$\begin{aligned}
\mathbf{r}^{spf} &= \begin{bmatrix} \mathbf{z}_{m_{us}} \\ \boldsymbol{\gamma}_{m_s} \mathbf{z}_{m_s} \end{bmatrix} - \mathbf{H}\mathbf{H}^\dagger \begin{bmatrix} \mathbf{z}_{m_{us}} \\ \boldsymbol{\gamma}_{m_s} \mathbf{z}_{m_s} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{z}_{m_{us}} \\ \boldsymbol{\gamma}_{m_s} \mathbf{z}_{m_s} + \mathbf{z}_{m_s} - \mathbf{z}_{m_s} \end{bmatrix} \\
&\quad - \mathbf{H}\mathbf{H}^\dagger \begin{bmatrix} \mathbf{z}_{m_{us}} \\ \boldsymbol{\gamma}_{m_s} \mathbf{z}_{m_s} + \mathbf{z}_{m_s} - \mathbf{z}_{m_s} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{z}_{m_{us}} \\ \mathbf{z}_{m_s} \end{bmatrix} - \mathbf{H}\mathbf{H}^\dagger \begin{bmatrix} \mathbf{z}_{m_{us}} \\ \mathbf{z}_{m_s} \end{bmatrix} + \begin{bmatrix} 0 \\ (\boldsymbol{\gamma}_{m_s} - \mathbf{I}_{m_s}) \mathbf{z}_{m_s} \end{bmatrix} \\
&\quad - \mathbf{H}\mathbf{H}^\dagger \begin{bmatrix} 0 \\ (\boldsymbol{\gamma}_{m_s} - \mathbf{I}_{m_s}) \mathbf{z}_{m_s} \end{bmatrix} \\
&= \mathbf{r} + \begin{bmatrix} -\mathbf{H}_{m_{us}} \mathbf{H}_{m_s}^\dagger \left(\boldsymbol{\gamma}_{m_s} - \mathbf{I}_{m_s}\right) \mathbf{z}_{m_s} \\ \left(\mathbf{I}_{m_s} - \mathbf{H}_{m_s} \mathbf{H}_{m_s}^\dagger\right) \left(\boldsymbol{\gamma}_{m_s} - \mathbf{I}_{m_s}\right) \mathbf{z}_{m_s} \end{bmatrix}
\end{aligned} \tag{4.18}$$

where $\mathbf{H}_{m_{us}}$ and $\mathbf{H}_{m_s}$ are matrices which contain rows of $\mathbf{H}$ corresponding to unspoofed and spoofed PMUs, respectively, and $\mathbf{H}_{m_s}^\dagger$ is a matrix which contains columns of $\mathbf{H}^\dagger$ corresponding to spoofed PMUs. Taking the expectation of the last expression of (4.18) gives

$$\begin{aligned}
\mathbb{E}[\mathbf{r}^{spf}] &= \mathbb{E}[\mathbf{r}] \\
&\quad + \begin{bmatrix} -\mathbf{H}_{m_{us}} \mathbf{H}_{m_s}^\dagger \left(\boldsymbol{\gamma}_{m_s} - \mathbf{I}_{m_s}\right) \mathbb{E}[\mathbf{z}_{m_s}] \\ \left(\mathbf{I}_{m_s} - \mathbf{H}_{m_s} \mathbf{H}_{m_s}^\dagger\right) \left(\boldsymbol{\gamma}_{m_s} - \mathbf{I}_{m_s}\right) \mathbb{E}[\mathbf{z}_{m_s}] \end{bmatrix} \\
&= \begin{bmatrix} -\mathbf{H}_{m_{us}} \mathbf{H}_{m_s}^\dagger \left(\boldsymbol{\gamma}_{m_s} - \mathbf{I}_{m_s}\right) \mathbb{E}[\mathbf{z}_{m_s}] \\ \left(\mathbf{I}_{m_s} - \mathbf{H}_{m_s} \mathbf{H}_{m_s}^\dagger\right) \left(\boldsymbol{\gamma}_{m_s} - \mathbf{I}_{m_s}\right) \mathbb{E}[\mathbf{z}_{m_s}] \end{bmatrix} \\
&\neq 0
\end{aligned} \tag{4.19}$$

The expectation is non-zero since $\mathbb{E}[\mathbf{z}_{m_s}]$ is non-zero (assuming the states are non-zero). Therefore GSAs give rise to a bias in residuals. The change in the statistic is utilized in our approach to differentiate a nominal scenario from a spoofing scenario.

Comparing the residual norm under the nominal and spoofing scenarios, we demonstrated that the residual norm distribution changes under GSAs. We further show that the residual norm increases during a GSA, i.e., $\left\|\mathbf{r}^{spf}\right\|^2 \geq \|\mathbf{r}\|^2$ if $\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)$ is semi-positive definite. The proof is provided below:
To show: $\left\|\mathbf{r}^{spf}\right\|^2 \geq \|\mathbf{r}\|^2$ if $\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)$ is semi-positive definite.
Proof: Let $\mathbf{b} = \begin{bmatrix} 0 \\ (\boldsymbol{\gamma}_{m_s} - \mathbf{I}_{m_s})\mathbf{z}_{m_s} \end{bmatrix}$ to simplify calculations. From (4.18):

$$\begin{aligned} \mathbf{r}^{spf} &= \mathbf{r} + \left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)\mathbf{b} \\ &= \left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)\mathbf{z} + \left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)\mathbf{b} \end{aligned} \tag{4.20}$$

Taking square of norm on both sides:

$$\begin{aligned} \left\|\mathbf{r}^{spf}\right\|^2 &= \|\mathbf{r}\|^2 + \left\|\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)\mathbf{b}\right\|^2 \\ &\quad + 2\mathbf{z}^{\top}\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)^{\top}\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)\mathbf{b} \\ &= \left\|\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)\mathbf{b}\right\|^2 + 2\mathbf{z}^{\top}\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)^{\top}\mathbf{b} \\ &\quad - 2\mathbf{z}^{\top}\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)^{\top}\left(\mathbf{H}\mathbf{H}^{\dagger}\right)\mathbf{b} + \|\mathbf{r}\|^2 \\ &= \left\|\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)\mathbf{b}\right\|^2 + 2\mathbf{z}^{\top}\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)^{\top}\mathbf{b} \\ &\quad - 2\mathbf{z}^{\top}\left(\mathbf{H}\mathbf{H}^{\dagger}-\mathbf{H}\mathbf{H}^{\dagger}\mathbf{H}\mathbf{H}^{\dagger}\right)^{\top}\mathbf{b} + \|\mathbf{r}\|^2 \\ &= \|\mathbf{r}\|^2 + \left\|\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)\mathbf{b}\right\|^2 \\ &\quad + 2\mathbf{z}^{\top}\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)^{\top}\mathbf{b} \end{aligned} \tag{4.21}$$

as $\mathbf{H}\mathbf{H}^{\dagger} = \mathbf{H}\mathbf{H}^{\dagger}\mathbf{H}\mathbf{H}^{\dagger}$, due to the structure of pseudoinverse matrices. Now, note that the RHS contains a sum of positive numbers, the last term is positive since $\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)$ is semi-positive definite. This implies that

$$\left\|\mathbf{r}^{spf}\right\|^2 \geq \|\mathbf{r}\|^2 \tag{4.22}$$

The above proof shows that the residual norm under GSAs will be greater than that of the nominal residual norm. Semi-positive definiteness of the matrix $\left(\mathbf{I}-\mathbf{H}\mathbf{H}^{\dagger}\right)$ is the necessary condition that ensures an increase in residual

norm during GSAs. This observation is used to correct PMU measurements under GSAs. Note that the necessary condition implies that the residual norm will increase under GSAs. However, it is not a sufficient condition, i.e., an increase in residual norm does not always imply that there is a GSA.

4.3 Spoofing-Resilient Static State Estimation

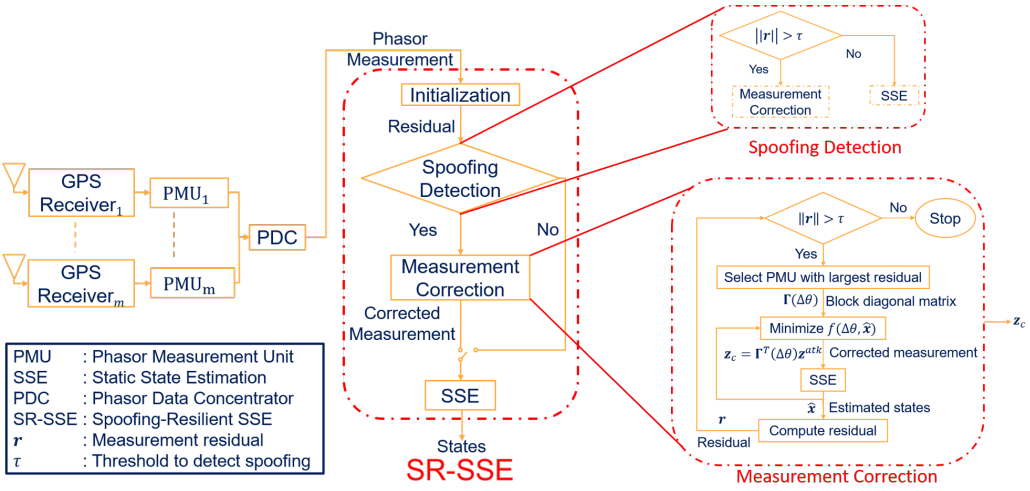


Figure 4.2: Flow chart of SR-SSE. First, we utilize residuals in the Spoofing Detection algorithm to detect GSAs. Later, if a GSA is detected, we correct PMU measurements in the Measurement Correction algorithm by iteratively minimizing the measurement residual norm. The corrected measurements are used in the SSE which provides GSA-resilient states.

The overall architecture of the proposed estimator is shown in Figure 4.2. In the proposed estimator, we perform the following steps:

1. **Initialization**: In this step, the power grid states are estimated using all of the PMU measurements and the measurement residuals are computed.

2. **Spoofing Detection**: The measurement residuals are compared with a predetermined threshold to detect spoofing. We consider measurements to be spoofed if the measurement residual norm is greater than a predetermined threshold.

3. **Measurement Correction**: Based on the output of the Spoofing Detection algorithm, the measurement correction step is carried out.

In this algorithm, we estimate the attack angles. Once the attack angles are estimated, we correct the PMU measurements.

4. **SSE**: The corrected measurements are passed to the SSE to estimate the power grid states.

The subsequent subsections provides more details on Spoofing Detection and Measurement Correction algorithms.

4.3.1 Spoofing Detection

In the previous section, we analyzed the measurement residuals and demonstrated that the residual norm increases under GSAs. We empirically compute a threshold to differentiate the nominal scenario from the spoofing scenario.

We perform MC simulations for the nominal scenario to obtain the maximum measurement residual norm. In MC simulations, we simulate a virtual power grid in steady state using MATPOWER [83]. For each simulation, we generate PMU measurements from (4.4) and estimate power grid states using SSE. The residual norm for each simulation is recorded. The maximum residual norm is selected as a threshold.

In this work, we primarily focus on GSAs and assume the PMU measurements are altered by GSAs only. The residuals obtained in the initialization step are passed to the Spoofing Detection algorithm, where we compare the measurement residual norm with the selected threshold. If the measurement residual norm is greater than the selected threshold we correct measurements in the Measurement Correction algorithm, otherwise we use the measurements as they are.

4.3.2 Measurement Correction

In this section, we will describe the developed algorithm to correct PMU measurements and thus provide GSA-resilient states. The flow chart of the algorithm is shown in the right side of Figure 4.2. The measurement correction is an iterative algorithm that estimates the attack angles for spoofed buses. The following steps are performed in the Measurement Correction algorithm:

1. The overall residual norm $\left(\left\|\mathbf{r}^{spf}\right\|\right)$ is compared with the selected threshold to identify whether the measurements are spoofed or not.

2. If the overall residual norm is less than the selected threshold, then the corrected measurements are used in the SSE to estimate states. Otherwise, we estimate the attack angles by iteratively minimizing the following objective function

$$f_{objective}(\Delta\theta_{m_s}, \hat{\mathbf{x}}) = \left\|\mathbf{r}^{spf}_{m_s}\right\| \tag{4.23}$$

   where m_s denotes the set of PMUs that are spoofed. In the developed algorithm, m_s is initialized as an empty set. We take the following steps to minimize the above objective function:

   (a) Select the PMU with the largest residual norm and add the measurements to the set m_s.

   (b) Given the previous state estimate ($\hat{\mathbf{x}}$), minimize the objective function with respect to $\Delta\theta_{m_s}$. We use gradient descent to minimize the objective function and estimate the attack angles.

   (c) Correct the PMU measurements by utilizing the estimated attack angles ($\Delta\theta_{m_s}$)

$$\mathbf{z}_c = \mathbf{\Gamma}^{\top}(\Delta\theta_{m_s})\mathbf{z}^{spf} \tag{4.24}$$

   (d) Update the power grid states

$$\hat{\mathbf{x}} = \mathbf{H}^{\dagger}\mathbf{z}_c \tag{4.25}$$

   (e) Repeat from Step 2.b until the estimated states have converged. The norm of the difference of the estimated states between two consecutive iterations is selected as the criterion for convergence.

3. Repeat from Step 1 until the residual norm falls below the predetermined threshold.

Equation (4.18) shows that the residuals are a function of the attack angle. This is the motivation for selecting residual norm as an objective function. The minimization of the residual norm is a non-convex problem. Our developed iterative algorithm is inspired from the alternating minimization algorithm. In order to estimate both the attack angles and states, we minimize

the objective function first with respect to the attack angles and then with respect to states.

4.4 Simulation Environment and Results

Real-life spoofing experiments cannot be performed without proper approval from the U.S. government. It is illegal to broadcast any signal at GPS frequency. Furthermore, it is costly to conduct real-life experiments with PMUs and the power grid. As a result, we perform simulations using MATPOWER [83], which generates steady-state PMU measurements for the IEEE 14, IEEE 39, IEEE 118, and Illinois 200-bus [68] test systems. The parameters required to create the regression matrix ($\mathbf{H}$) are provided in MATPOWER [83]. The noise covariance ($\mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^\top]$) is a diagonal matrix with standard deviation of 0.01 and 0.02 for bus voltage and line current measurements, respectively.

In our simulations, we have assumed the network to be observable. Table 4.1 presents the PMU buses for different test systems. For the nominal scenario, we perform power flow analysis to estimate the power grid states. We use these states as reference and calculate Root Mean Square Error (RMSE) relative to these reference states for different estimators. To simulate spoofing, we modify the nominal PMU measurements using (4.13).

The simulation results are divided into two parts: Residual Characteristics and State Estimation. In residual characteristics, we show the residual norm distribution under the nominal and spoofing scenarios. This distribution is empirically obtained by performing MC simulations. In state estimation, we compare the estimation results of SSE, SpM [59], and SR-SSE under different GSAs.

4.4.1 Residual Characteristics

For the considered systems, the minimum eigenvalue of $\left(\mathbf{I} - \mathbf{H}\mathbf{H}^\dagger\right)$ is found to be on the order of 10^{-17}, implying that $\left(\mathbf{I} - \mathbf{H}\mathbf{H}^\dagger\right)$ is semi-positive definite. Therefore, we expect the residual norm to increase under GSAs.

We conduct 1000 MC simulations for both nominal and spoofing scenarios. In the nominal scenario, none of the PMU measurements are modified. For the spoofing scenario, we modify the PMU measurements at bus 1 and 6

Table 4.1: PMU buses for different IEEE bus test systems

Test Case	Number of PMUs (M)	PMU buses
IEEE 14	8	[1,2,4-6,7,10,13]
IEEE 39	22	[1-3,5-10,12,14-17,19, 20,22,23,25,26,29,39]
IEEE 118	54	[1,3-6,8,9,11,12,15, 17,19,21,23,25,26,28,30, 34,35,37,40,43,45,46, 49,52,54,56,59,62,63,65, 68,70,71,75,76,77,78, 80,83,85,86,89,90,92,94, 96,100,105,108,110,114]
Illinois 200	136	[1,2,4,6,8-13,15-30,32,33,35, 37-41,43-45,47-53,55-63,65,67-73, 75-80,82,83,86,87,89-94,99,101, 103-105,107,108,110,113-115,117, 118,122,123,125-127,130,131, 135-138,145-149,151-155,157,161, 163-170,173,174,176,178,180-183, 185,186,189,190,195,196,197]

according to the spoofing attack model with an attack angle of 20 degrees. In each MC simulation, we calculate the residuals using SSE. Figure 4.3 plots the residual norms for each test system.

Figure 4.3 illustrates that spoofing introduces an increase in the residual norm, thus verifying (4.22). We use the maximum residual norm in the MC simulations for the nominal scenario as our threshold to differentiate the nominal and spoofing scenarios. We can effectively differentiate the nominal scenario from the spoofing using the threshold as long as the spoofed residual norm is greater than the selected threshold. We investigate if there are some cases that give rise to smaller residual norm compared to the threshold.

We perform 100 MC simulations in which the attack angle is varied between 1 to 20 degrees and the numbers of GSAs are varied from 1 to 3. In each GSA, we randomly select and spoof PMU buses with a given attack angle. For each GSA, we record the minimum residual norm among the 100 simulations and compare it with the selected threshold. Figure 4.4 shows the variation of the minimum residual norm with attack angle for different numbers of GSAs. This figure illustrates that GSAs with small attack angles (< 8 degrees) are not differentiable from the nominal scenario using the selected threshold.

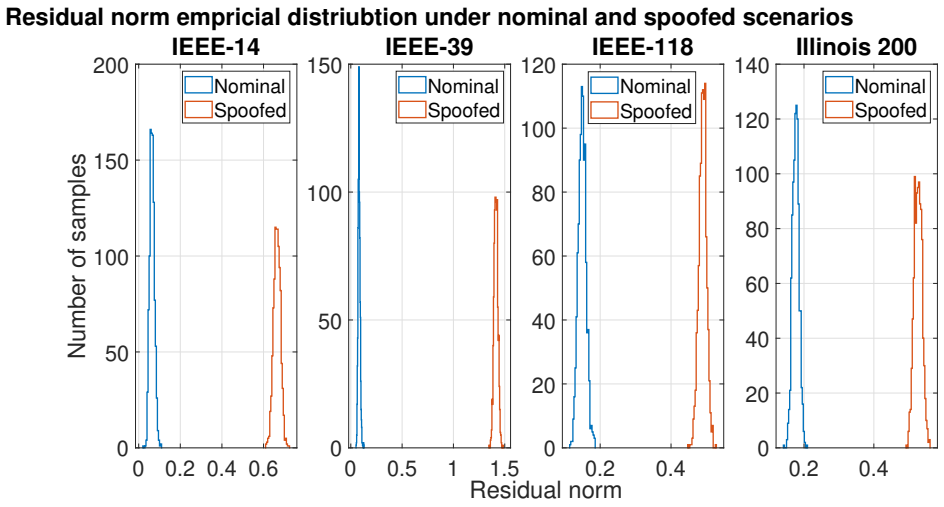


Figure 4.3: The histogram of measurement residual norms during the nominal, shown in blue, and during spoofing, shown in red, scenarios. The nominal and spoofing scenarios are clearly distinguishable due to the change in the distribution of the residual norms.

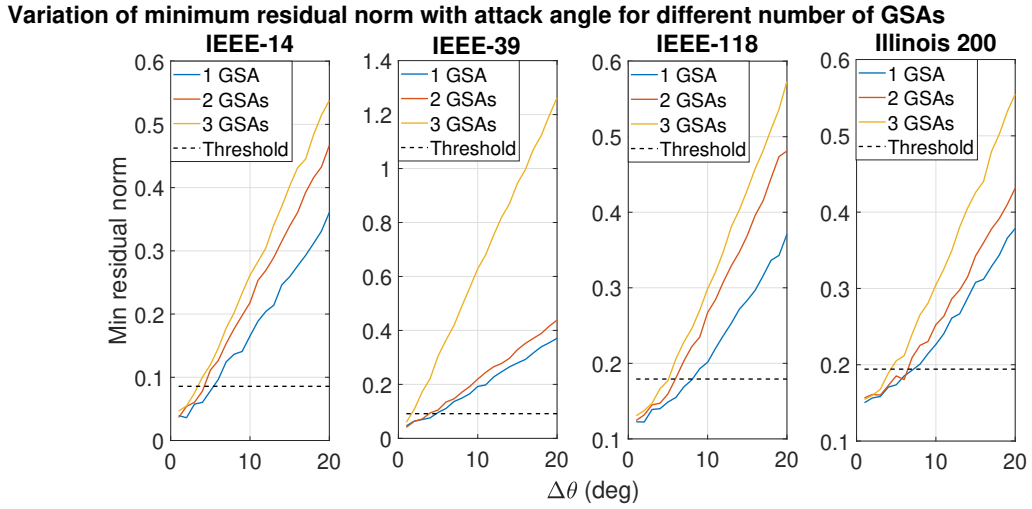


Figure 4.4: Variation of the minimum residual norm with attack angle for different numbers of GSAs. The nominal and spoofing scenarios are not distinguishable for small attack angles (< 8 degrees).

4.4.2 State Estimation

We test SR-SSE on the IEEE 14, IEEE 39, IEEE 118, and Illinois 200-bus test systems. We perform MC simulations in which number of GSAs are varied from 1 to 3. For each GSA, we perform 100 MC simulations in which we randomly spoof a given number of PMU buses with the attack angles.

The attack angles are randomly sampled from a normal distribution with a mean of 40 degrees and standard deviation of 5 degrees.

The RMSE and computation time of SR-SSE are compared with SpM and SSE. Figures 4.5, 4.6, 4.7, and 4.8 show RMSE box plots of voltage and phase estimates for different estimators. In these plots, the median value is illustrated with a red line, the blue box bounds the first and third quantile values, the black whiskers denote the 1.5× inter-quantile range, and the outliers are marked with red stars. Outliers are the points that lie outside 1.5× inter-quantile range.

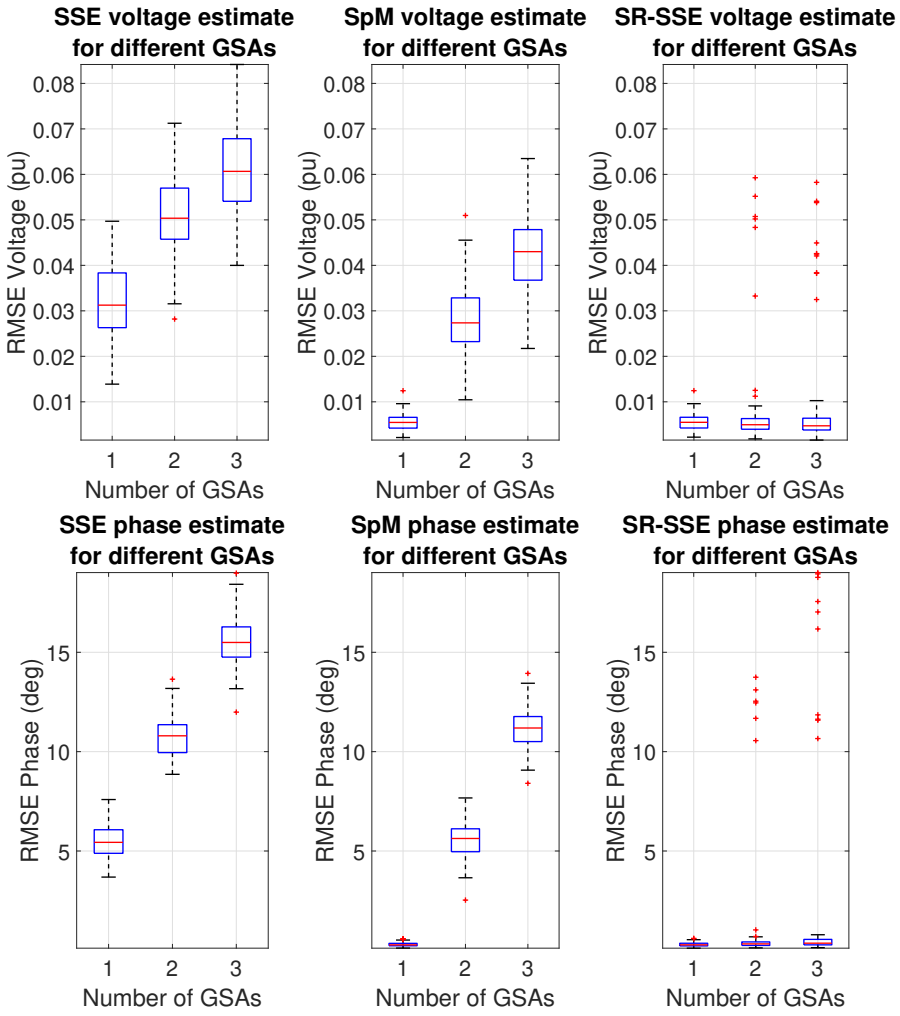


Figure 4.5: Comparison of the voltage and phase RMSE of SSE (first column), SpM (second column), and SR-SSE (third column) for the IEEE 14-bus test system. The SSE and SpM estimates degrade with the number of GSAs. The SR-SSE estimates are an order of magnitude more accurate than SSE and SpM estimates.

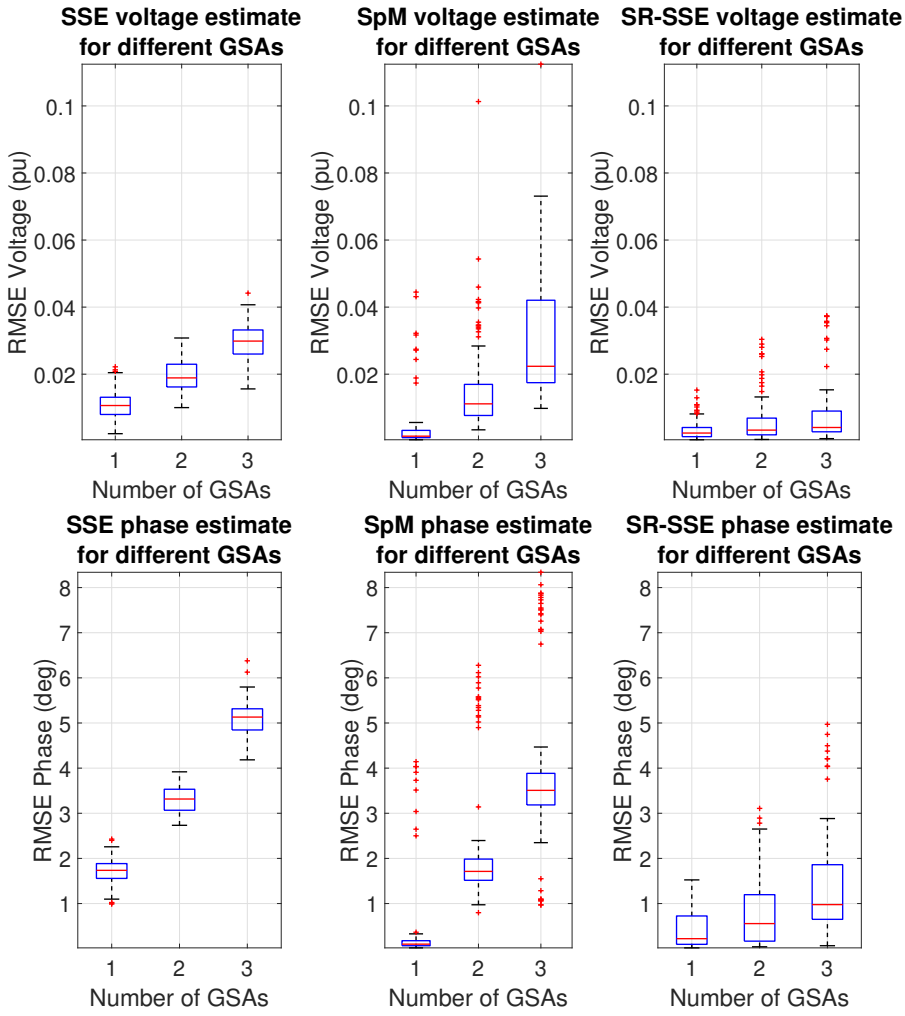


Figure 4.6: Comparison of the voltage and phase RMSE of the SSE (first column), SpM (second column), and SR-SSE (third column) for the IEEE 39-bus test system. The SSE and SpM estimates degrade with the number of GSAs. The SR-SSE voltage estimates are an order of magnitude more accurate than SSE and SpM estimates.

From Figures 4.5, 4.6, 4.7, and 4.8, we observe that the voltage RMSE of SR-SSE is an order magnitude smaller than SSE for all test systems. The performance of SR-SSE and SpM is similar for a single GSA; however, SR-SSE outperforms SpM for multiple GSAs.

Table 4.2 presents the median RMSE of 100 MC simulations and the computation time for all of the estimators. The computation time of SR-SSE is greater than SpM for the IEEE 14 and 39-bus test systems. However, the SR-SSE estimates have lower RMSE for all test systems. Computation time of SR-SSE increases with number of GSAs as it is an iterative estimator that

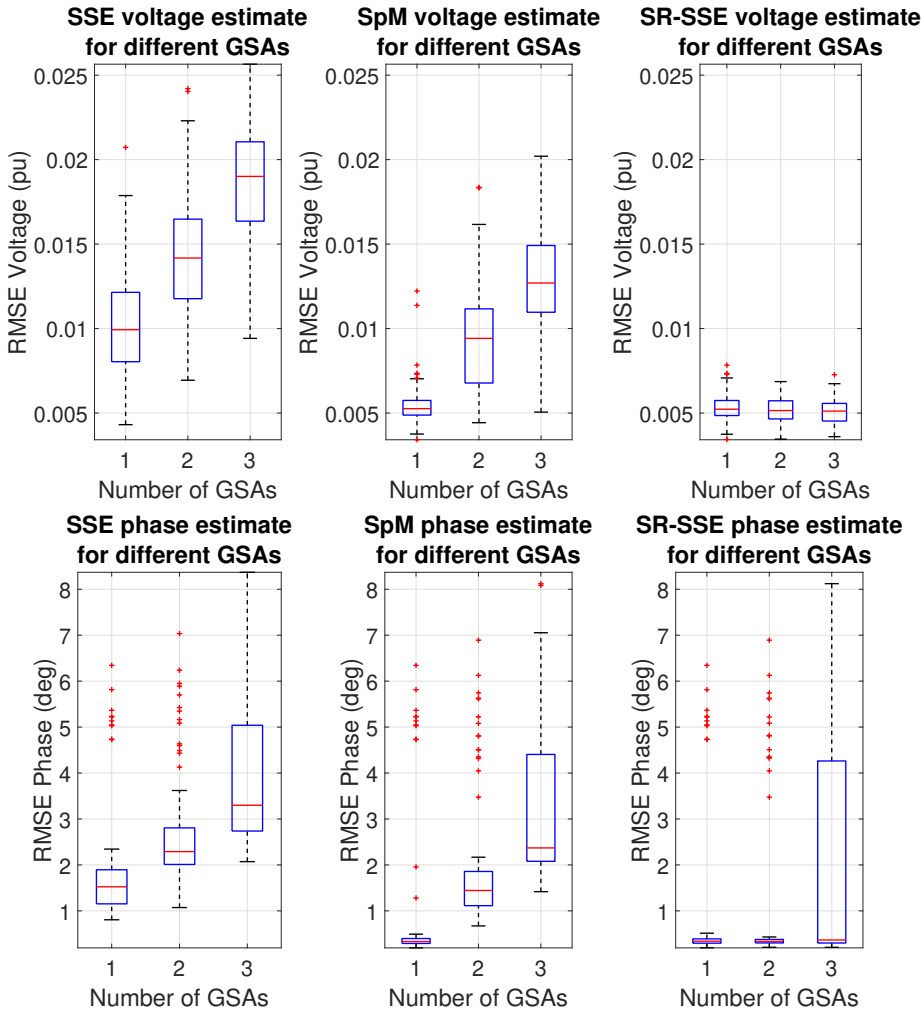


Figure 4.7: Comparison of the voltage and phase RMSE of the SSE (first column), SpM (second column), and SR-SSE (third column) for the IEEE 118-bus test system. The SSE and SpM estimates degrade with number of GSAs. The SR-SSE estimates are an order of magnitude more accurate than SSE and SpM estimates.

mitigates one GSA at a time. The largest computation time of 1.69 seconds is observed in the Illinois 200-bus test system for the scenario with 3 GSAs.

Depending on the time of day, the system time constant varies between 10 seconds and 10 minutes [33]. Presently, state estimation runs every 1-5 minutes for large power systems [85]. For the systems with time constant smaller than 1.69 seconds, SR-SSE can be used for real-time estimation.

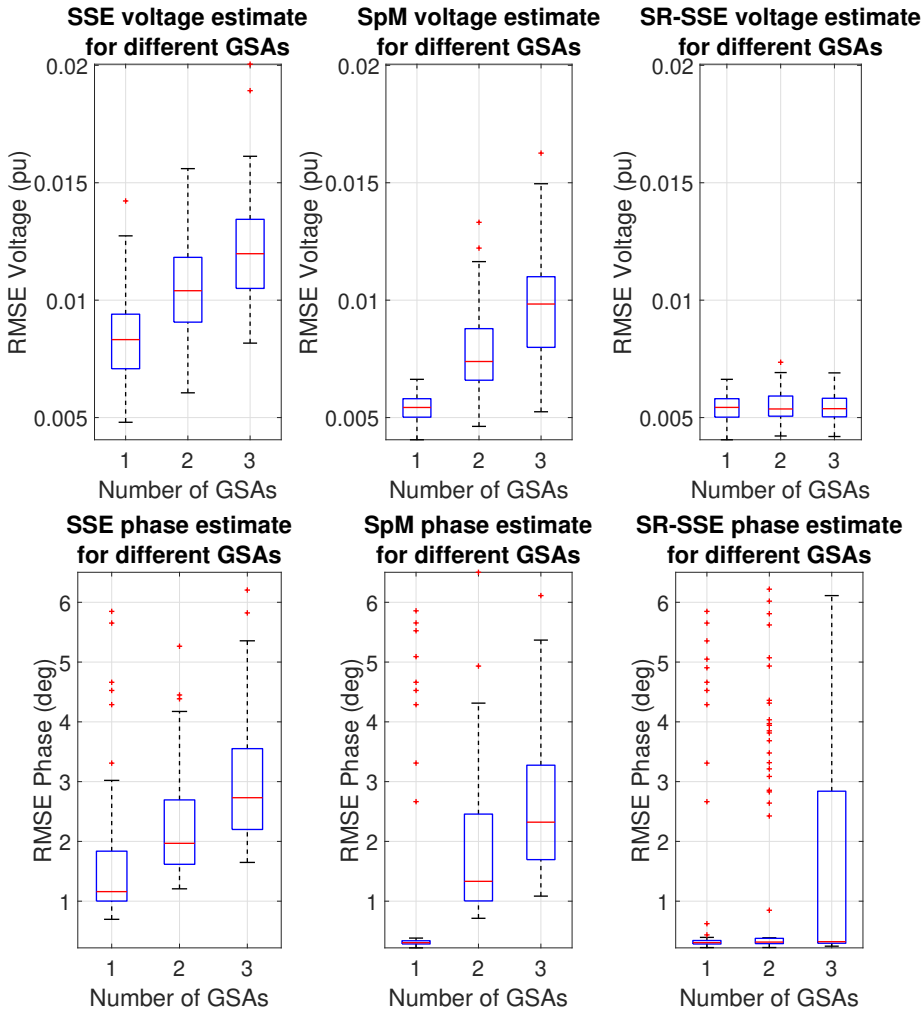


Figure 4.8: Comparison of the voltage and phase RMSE of the SSE (first column), SpM (second column), and SR-SSE (third column) for the Illinois 200-bus test system. The SSE and SpM estimates degrade with number of GSAs. The SR-SSE estimates are an order of magnitude more accurate than SSE and SpM estimates.

4.5 Limitations and Summary

This chapter described our proposed novel residual-based SR-SSE for the power grid that is resilient to multiple GSAs with different attack angles. The proposed estimator consists of two algorithms, one that detects a GSA and another that corrects the PMU measurements. The developed spoofing detection algorithm is based on measurement residuals. The measurement correction algorithm iteratively minimizes residual norms to correct PMU measurements under GSAs. Furthermore, we derived a necessary condition

to show that the residual norm increases under GSAs. The necessary condition was verified with MC simulations. When validating SR-SSE on the IEEE 14, IEEE 39, IEEE 118, and Illinois 200-bus test systems, we observed that SR-SSE estimates are an order of magnitude more accurate than SSE and SpM for multiple GSAs. We also observed that SR-SSE's computation time is lower than that of SpM for the IEEE 118 and Illinois 200-bus test systems, demonstrating that SR-SSE achieved a greater accuracy without compromising computation time.

We observed that the computation time of SR-SSE increases with the number of GSAs, and RMSE of phase estimates increases slightly with the number of GSAs. The increase in computation time is due to the iterative nature of the proposed estimator. SR-SSE's performance will degrade with the increase in the number of GSAs due to the reduction in the number of authentic measurements, as shown in [61]. Figure 4.4 showed that the GSAs with small attack angles (< 8 degrees) are not detectable. Furthermore, the derived necessary condition does not guarantee that increase in residual norm implies GSAs. In the next chapter 6, we improve the performance of SE by augmenting it with GPS measurements.

Table 4.2: Median RMSE and computation time of the SSE, SpM, and SR-SSE estimators for the IEEE 14, IEEE 39, IEEE 118, and Illinois 200-bus test systems under different GSAs. RMSE of SR-SSE estimates is smaller than that of SSE and SpM estimators for the multiple GSAs scenario. SR-SSE provides GSA-resilient states under multiple GSAs.

	Test Case	**Scenario**	Voltage Magnitude (pu)	Phase (deg)	Computation Time (sec)
SSE	**IEEE 14**	1 GSAs	0.0313	5.4354	0.0008
		2 GSAs	0.0504	10.7971	0.0008
		3 GSAs	0.0607	15.4949	0.0007
	IEEE 39	1 GSAs	0.0107	1.7355	0.0026
		2 GSAs	0.0189	3.3157	0.0025
		3 GSAs	0.0299	5.1311	0.0025
	IEEE 118	1 GSAs	0.0099	1.5235	0.0125
		2 GSAs	0.0142	2.2896	0.0123
		3 GSAs	0.0190	3.2987	0.0122
	Illinois 200	1 GSAs	0.0083	1.1598	0.0393
		2 GSAs	0.0104	1.9679	0.0414
		3 GSAs	0.0120	2.7309	0.0391
SpM	**IEEE 14**	1 GSAs	0.0055	0.2759	0.0074
		2 GSAs	0.0274	5.6298	0.0057
		3 GSAs	0.0430	11.1899	0.0054
	IEEE 39	1 GSAs	0.0015	0.1005	0.0711
		2 GSAs	0.0112	1.7123	0.0651
		3 GSAs	0.0224	3.5075	0.0622
	IEEE 118	1 GSAs	0.0053	0.3324	1.9515
		2 GSAs	0.0094	1.4440	1.7703
		3 GSAs	0.0127	2.3702	1.7530
	Illinois 200	1 GSAs	0.0054	0.3055	31.4554
		2 GSAs	0.0074	1.3320	30.0399
		3 GSAs	0.0098	2.3216	29.5627
SR-SSE	**IEEE 14**	1 GSAs	0.0055	0.2776	0.0282
		2 GSAs	0.0050	0.3254	0.1047
		3 GSAs	0.0047	0.3610	0.2629
	IEEE 39	1 GSAs	0.0024	0.2215	0.2510
		2 GSAs	0.0033	0.5558	0.6667
		3 GSAs	0.0041	0.9764	1.3421
	IEEE 118	1 GSAs	0.0052	0.3402	0.1287
		2 GSAs	0.0051	0.3345	0.3015
		3 GSAs	0.0051	0.3666	0.6991
	Illinois 200	1 GSAs	0.0054	0.3067	0.2842
		2 GSAs	0.0054	0.3160	0.8010
		3 GSAs	0.0054	0.3225	1.6912

CHAPTER 5

HARDWARE-IN-THE-LOOP GPS AND PMU INTEGRATED DATASETS FOR THE POWER GRID UNDER GSAS

Experimental datasets containing both GPS and PMU measurements are essential to assess GSAs' impact on power grid's SEs. The available datasets [86, 62] in the literature are not relevant to the power grid community due to the following reasons:

- Most of the GSAs in [86, 62] alter receiver position and are detectable since power grid sub-stations are static. GSAs that modify receiver time without changing its position pose a threat to the power grid. Such GSAs are known as timing GSAs [59].
- Timing GSAs in [86, 62] induce a total time delay of $2\mu s$. However, a timing delay greater than $26.5\mu s$ is required to violate IEEE C37.118-2011 standard.
- The current spoofing datasets in the literature are for GPS only. Integrated datasets, containing PMU measurements coupled with GPS measurements, are unavailable.

Due to lack of datasets, researchers in [59, 42, 60] evaluate SE's performance under GSAs using low-fidelity simulations [83] in which PMU measurements are obtained by performing software simulations without incorporating the PMU hardware. It is challenging to conduct real-world experiments involving GSAs as it is illegal to broadcast signals at GPS frequency. Additionally, experimenting on the real power grid under GSAs is costly as GSAs would implicitly affect the grid's power flow.

In this chapter, we address the limitations of the available datasets for power grid applications. We conduct HIL simulations with RTDS, PMUs, and GPS clock to simulate the IEEE 14-bus test system for different GSA scenarios. We generate PMU datasets coupled with GPS datasets for two scenarios: nominal, simulates an ideal environment in which GPS signals are

authentic and spoof, simulate timing GSAs with time-walk. The remainder of the chapter is organized as follows: Section 5.1 presents the developed methodology for generating integrated datasets. In this section, we provide information for generating relevant GPS datasets along with PMU datasets. Section 5.2 describes our integrated datasets for nominal and spoof scenarios. In Section 5.3, we show experimental results that validate our integrated datasets. Section 5.4 summarizes this chapter..

5.1 Methodology for Generating Integrated Datasets

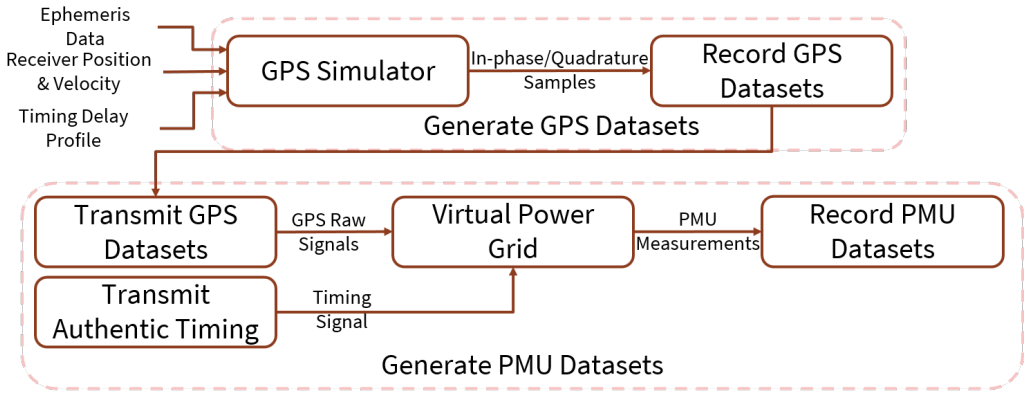
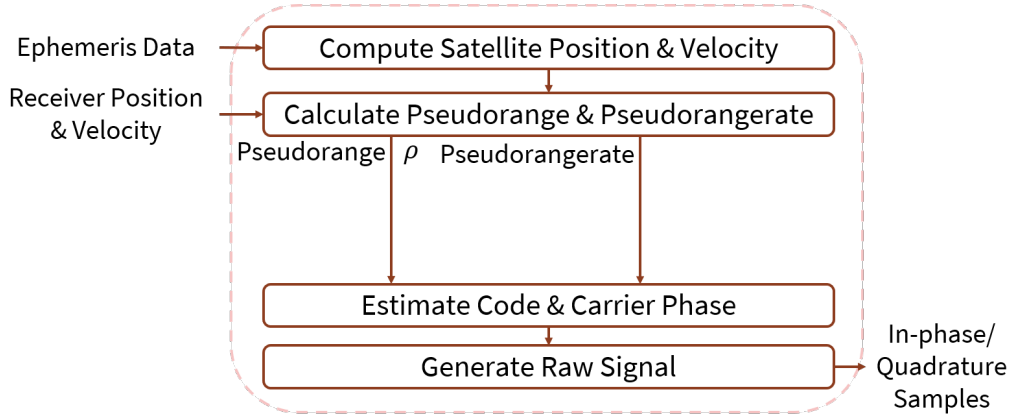


Figure 5.1: Overall architecture of our methodology.

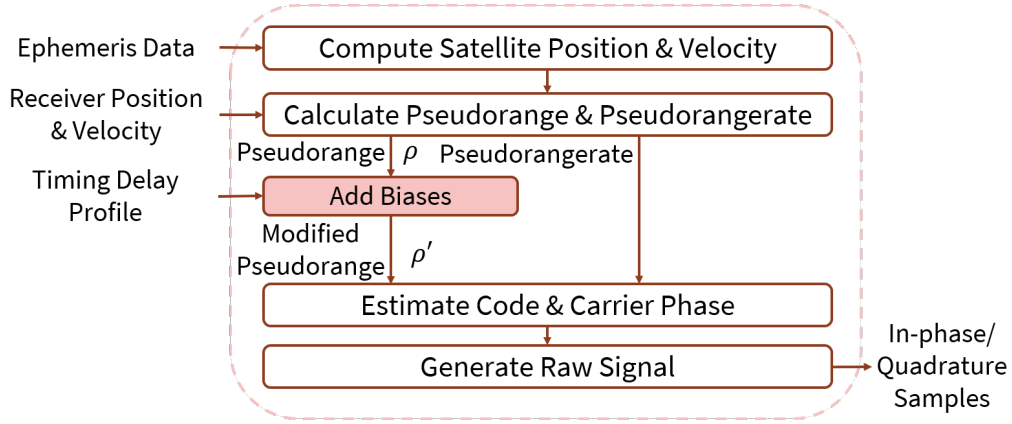
The developed methodology contains two coupled steps: generate GPS Datasets and generate PMU Datasets. In the first step, we generate GPS datasets using a modified GPS simulator [87], to which the inputs are ephemeris data, receiver position, velocity, and timing delay profile. This simulator generates In-phase/Quadrature (IQ) samples, which we record.

In the second step, we generate PMU datasets by conducting HIL simulations with a virtual power grid that receives GPS raw signals from the generated GPS datasets and authentic timing signals. Transmitting these two signals allows us to spoof some of the PMUs in the virtual power grid network. We record the PMU data in real-time from the virtual power grid for a given GPS dataset. Figure 5.1 shows the overall architecture of the developed methodology for generating integrated datasets. A detailed description of the two steps involved in our methodology is given in the following subsections.

5.1.1 Generating GPS Datasets



(a) Generic GPS Simulator



(b) Modified GPS simulator

Figure 5.2: High-level flow chart of a generic GPS simulator (a) and the modified GPS simulator (b).

The flow chart in Figure 5.2(a) shows the high-level steps involved in generating GPS raw signals. The first step is to compute the satellite positions and velocities using the ephemeris data. Next, pseudoranges and pseudorangerates are calculated by utilizing the receiver's position and velocity. Then the simulator estimates the code and carrier phase by applying the calculated pseudoranges and pseudorangerates [1]. Finally, the simulator generates IQ samples by employing the estimated code and carrier phase. We utilize GPS-SDR-SIM [87], an open-source GPS simulator that performs these steps and generates GPS raw signals.

We modify GPS-SDR-SIM by inserting the **Adding Biases** block to simulate timing GSAs. Figure 5.2(b) shows the steps involved in the modified

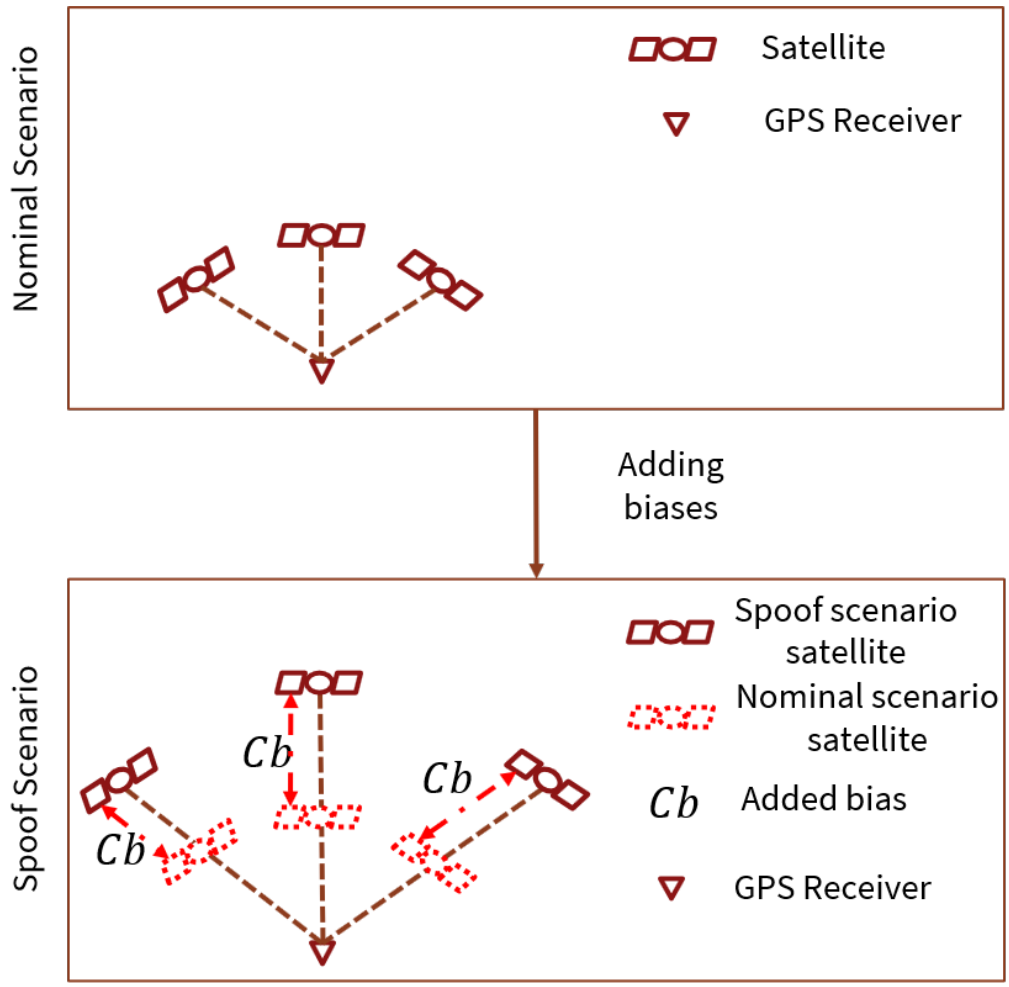


Figure 5.3: Adding equal biases to the pseudoranges for each of the visible satellites, modify the timing solution without changing the positioning solution.

GPS simulator. In **Adding Biases** block, we add equal biases to the pseudoranges for each of the visible satellites. The timing delay profile determines the magnitude of the biases.

Figure 5.3 illustrates the idea of simulating timing GSAs by adding equal biases to the pseudoranges for each of the visible satellites. In the nominal scenario, illustrated in the top plot of Figure 5.3, the receiver receives authentic GPS signals. In the spoof scenario, shown in the bottom plot of Figure 5.3, the added biases modify the timing solution without changing the positioning solution. To show this idea mathematically, consider the pseudorange equation in the nominal scenario which is given by

$$\rho^i = d^i + C\left(\delta t - \delta t^i\right) + \eta^i \tag{5.1}$$

where ρ^i denotes the pseudorange between the receiver and the i^{th} visible satellite, d^i is the true range between the receiver and the i^{th} visible satellite, C refers to speed of light, δt is the receiver clock bias, δt^i is the i^{th} satellite clock bias, and η^i is zero mean Gaussian noise. To simulate timing GSAs, we add Cb bias to the pseudorange that modifies the pseudorange equation as

$$\rho^{'i} = d^i + C\left(\delta t - \delta t^i\right) + Cb + \eta^i \tag{5.2}$$

where $\rho^{'i}$ denotes the modified pseudorange and Cb is the bias in meters corresponding to a timing delay of b seconds. Rearranging the terms in the last equation will result in the following equation

$$\begin{aligned} \rho^{'i} &= d^i + C\left((\delta t + b) - \delta t^i\right) + \eta^i \\ &= d^i + C\left(\delta t' - \delta t^i\right) + \eta^i \end{aligned} \tag{5.3}$$

where $\delta t' = \delta t + b$ is the modified receiver clock bias. From (5.1) and (5.3), we observe that by adding equal biases to the pseudoranges for each of the visible satellite we can achieve timing GSAs.

We use a timing delay profile that increases linearly i.e., we induce timing delay through time-walk to avoid sudden jumps in GPS timing and reduce the risk of detecting GSAs. The modified GPS simulator is utilized for generating GPS datasets for nominal and spoof scenarios.

5.1.2 Generating PMU Datasets

We simulate a virtual power grid network by performing HIL simulations with RTDS, physical PMUs, virtual PMUs, and GPS clock. Figure 5.4 elaborates on the intermediate blocks from our overall architecture, shown in Figure 5.1, that are utilized to generate PMU measurements. The virtual power grid testbed is adapted from [88], where author tested the performance of a PMU under meaconing and jamming attacks. We extended the testbed in [88] by incorporating multiple PMUs associated with a power grid network.

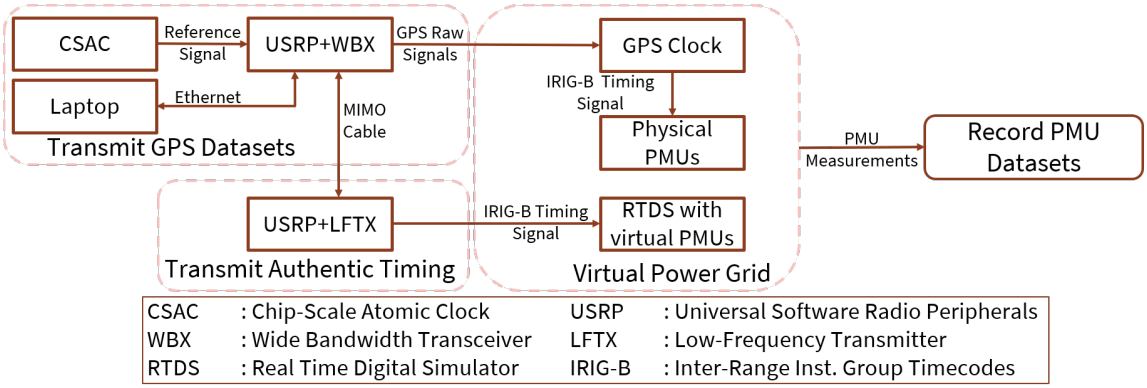


Figure 5.4: Generating PMU datasets by performing HIL simulations with RTDS, physical PMUs, virtual PMUs, and GPS clock.

The virtual power grid provides PMU measurements that are recorded in real-time. We transmit two signals to the virtual grid: GPS raw signals from the generated GPS datasets and authentic timing signals. The physical PMUs get the Inter-Range Instrumentation Group-B (IRIG-B) timing signals from the GPS clock, which receives the transmitted GPS raw signals. We also generate authentic IRIG-B timing signals that we transmit to RTDS. The virtual PMUs inside the RTDS always receive authentic timing signals.

In the Transmit GPS datasets block, we transmit GPS signals from the generated GPS datasets through a Universal Software Radio Peripheral (USRP) with a Wide Bandwidth Transceiver (WBX). In the Transmit Authentic Timing block, the second USRP, with Low-Frequency Transmitter (LFTX), transmits the authentic timing signals to RTDS. A Chip Scale Atomic Clock (CSAC) provides a reference signal to the USRP. Multiple Input Multiple Output (MIMO) cable syncs the two USRPs. We utilize openly available GNU-Radio Software to communicate between the two USRPs and the laptop for transmitting the two signals.

In the virtual power grid, we simulate the IEEE-14 bus system that is illustrated in Figure 5.5. There are 8 PMUs installed in this bus system to ensure observability, which is essential for SEs. Physical PMUs at bus 4 and 6 receive a timing signal from the GPS clock. The remaining PMUs at bus 1, 3, 5, 7, 10, 11, and 13 are virtual, which always receive authentic timing signals. For each PMU bus, PMU measurements, consisting of voltage phasors, line current phasors, frequency, and frequency change rate, are recorded in real-time using open-PDC software [89]. Figure 5.6 shows the hardware setup

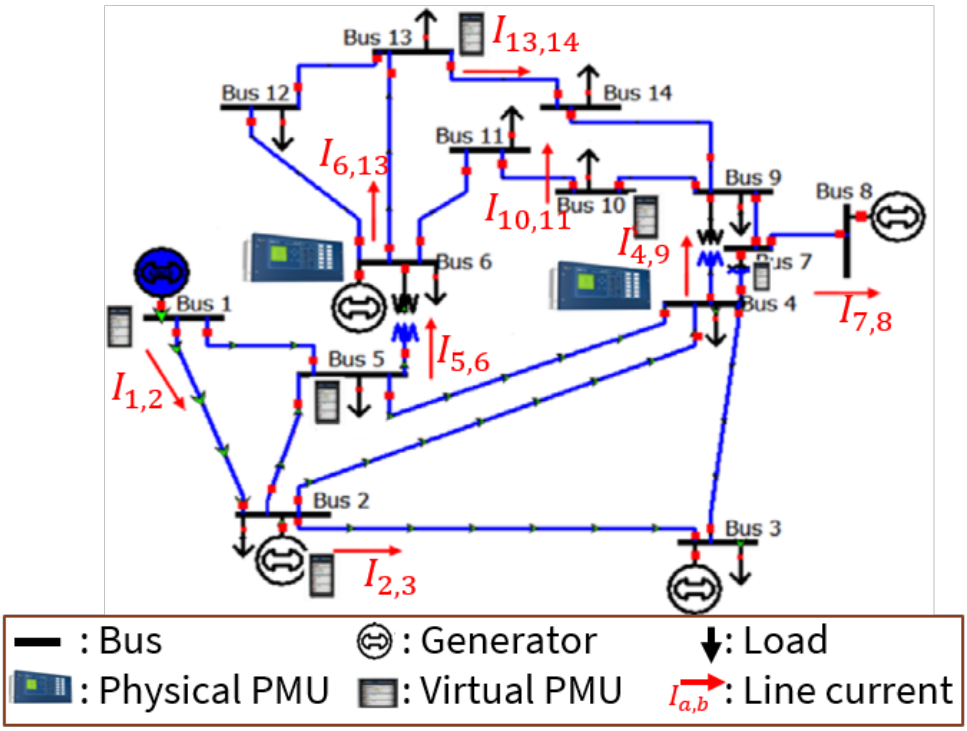


Figure 5.5: IEEE-14 bus system with physical and virtual PMUs.

for conducting HIL simulations and generating PMU datasets.

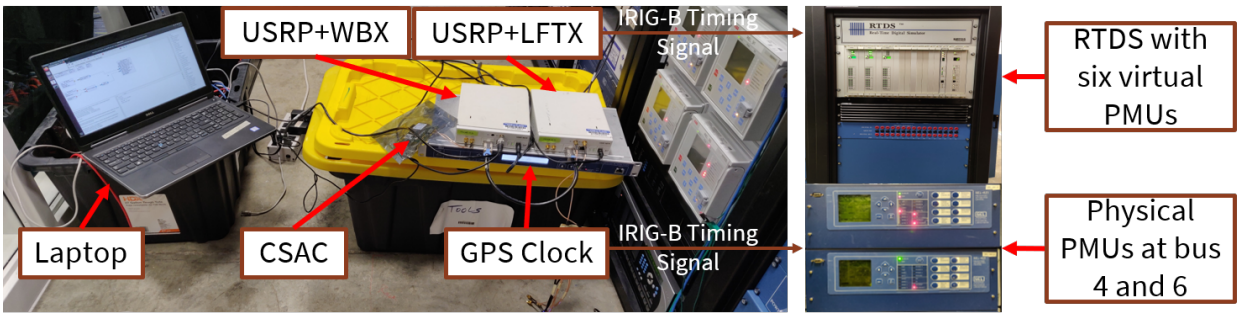


Figure 5.6: Hardware setup consisting of RTDS, physical PMUs, virtual PMUs, GPS clock and USRPs.

5.2 Integrated Datasets

We generated four GPS and PMU integrated datasets using the developed methodology for two scenarios: nominal and spoof. The nominal scenario refers to an ideal environment in which the transmitted GPS raw signals are

authentic. In the spoof scenario, we generated GPS raw signals by utilizing the modified GPS simulator that simulates timing GSAs with time-walk. Table 5.1 lists the scenarios for the integrated datasets.

Table 5.1: Scenarios for the integrated datasets.

S. No.	**Scenario**	**Total Induced Time Delay (ms)**	**Time Delay Profile**	**Description**
1	Nominal	0	-	Simulates ideal environment
2	Spoof	0.5	Time-walk	Simulates timing GSAs
3		2		
4		4		

Integrated datasets contain PMU datasets coupled with GPS datasets. We generated 30 minute long GPS datasets for each of the listed scenarios in Table 5.1. For each scenario, the simulated receiver position is static and the start time of the experiment is the same. Table 5.2 provides the ground truth for the receiver position and the start time. We obtained satellite positions from the ephemeris data and pseudoranges by performing scalar tracking using pyGNSS, a Python-based SDR research suite [76, 80]. The GPS datasets contain:

- GPS raw signals in binary files.
- GPS timestamped satellite positions and pseudoranges in CSV files.

Table 5.2: Ground truth of the receiver position and start time.

Date	**UTC start time (hh:mm:ss)**	**Receiver position in ECEF frame (m)**		
Dec 14, 2014	00:00:00	X	Y	Z
		151317.2428	−4882273.0498	4087975.6877

We generated PMU datasets by transmitting GPS datasets to the virtual power grid and recording PMU measurements. The PMU measurements under the nominal scenario are the ground truth measurements for the IEEE-14 bus system. The PMU datasets provide GPS timestamped voltage phasors, current phasors, frequency, and frequency change rate in CSV files for each PMU in the IEEE-14 bus system. The GPS and PMU integrated datasets are openly available at *https://navlab.stanford.edu/resources/datasets*.

5.3 Experimental Validation

We analyzed GPS datasets using *pyGNSS* software suite. We performed scalar tracking and obtained a navigation solution for each of the generated GPS datasets. Figure 5.7 shows the least-squares navigation solution and induced timing delays for different scenarios. The cyan, red, green and blue colors in Figure 5.7(a) show the navigation solution for Nominal, Spoof 0.5ms, Spoof 2ms, and Spoof 4ms, respectively. The Figure 5.7(b) shows induced timing delay for Spoof 0.5ms, Spoof 2ms, and Spoof 4ms with red, green, and blue colors, respectively. We observe that the navigation solution for Nominal and Spoof scenarios is indistinguishable. We further observe that the Spoof 0.5ms, Spoof 2ms, and Spoof 4ms scenarios induce a total timing delay of 0.5, 2, and 4 ms, respectively. Therefore, the generated GPS Spoof datasets induce timing GSAs with time-walk.

We tested the generated GPS datasets and our authentic timing signals with the GPS clock. Figure 5.8 shows the IRIG-B timing signal from the GPS clock, in blue color, and the authentic IRIG-B timing signal, in red color, at two-time instants during the experiment for nominal and spoof scenarios. In the nominal scenario, we notice that the authentic timing signal is in sync with the GPS clock's timing signal. In the spoof scenario, we observe that the GPS clock's timing signal lags behind the authentic timing signals. Figure 5.8 demonstrates that GPS Spoof datasets achieve the desired timing delays for different scenarios.

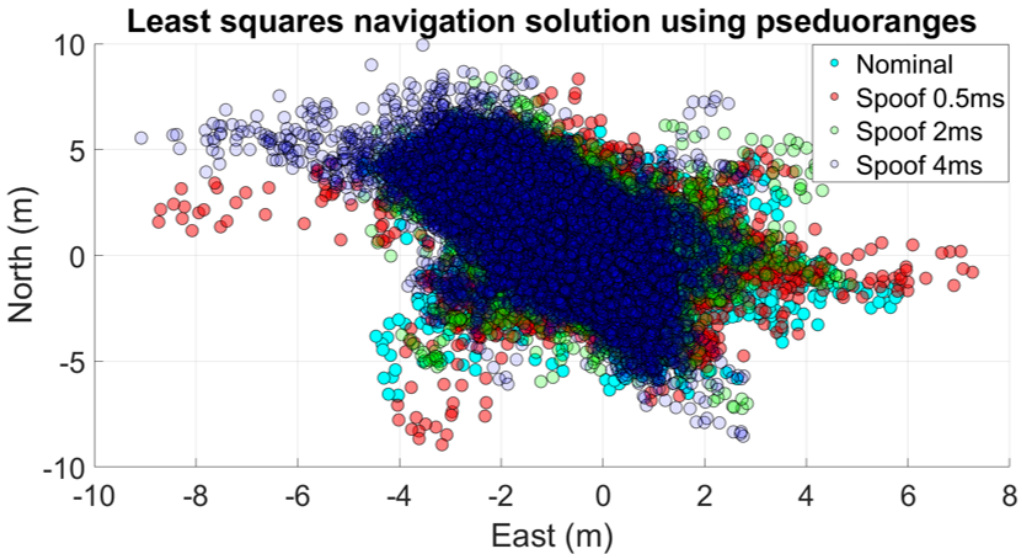
Timing GSAs alter phase angle of PMU measurements proportional to the induced timing delay [46], as shown in the following equation

$$\Delta\theta = 2\pi f \Delta t \tag{5.4}$$

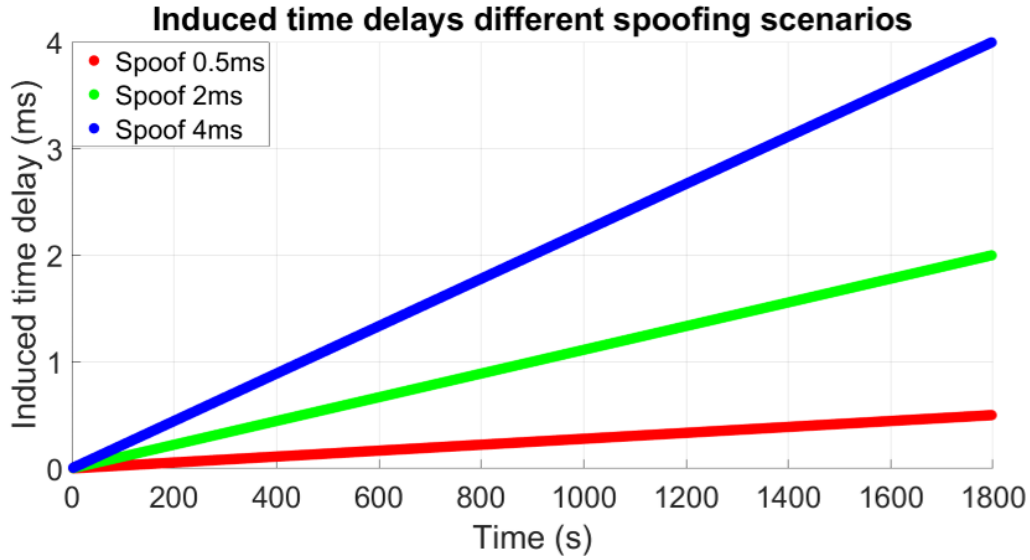
where $\Delta\theta$ is the phase delay, f is the frequency of the current and Δt is the timing delay induced by a timing GSA. Table 5.3 shows the expected phase delays, obtained from (5.4), for different spoof scenarios.

Table 5.3: Expected phase delays for different spoof scenarios.

Scenario	$\Delta\theta$ **(deg)**
Spoof 0.5ms	10
Spoof 2ms	43
Spoof 4ms	86



(a) Positioning Solution



(b) Induced Timing Delay

Figure 5.7: Positioning solution obtained using least squares (a) and induced timing delay (b) for different scenarios.

Figure 5.9 displays the voltage phase angle at spoofed bus 4 and 6 for different spoof scenarios. We observe that the voltage phase angle, under spoof scenarios, increases linearly and diverges from the nominal voltage phase angle. Furthermore, we notice a total phase delay of 10, 43, and 86 degrees is introduced for Spoof 0.5 ms, Spoof 2 ms, and Spoof 4 ms, respectively. Figure 5.9 validates that the GPS Spoof datasets induce phase delays proportional to the induced timing delays.

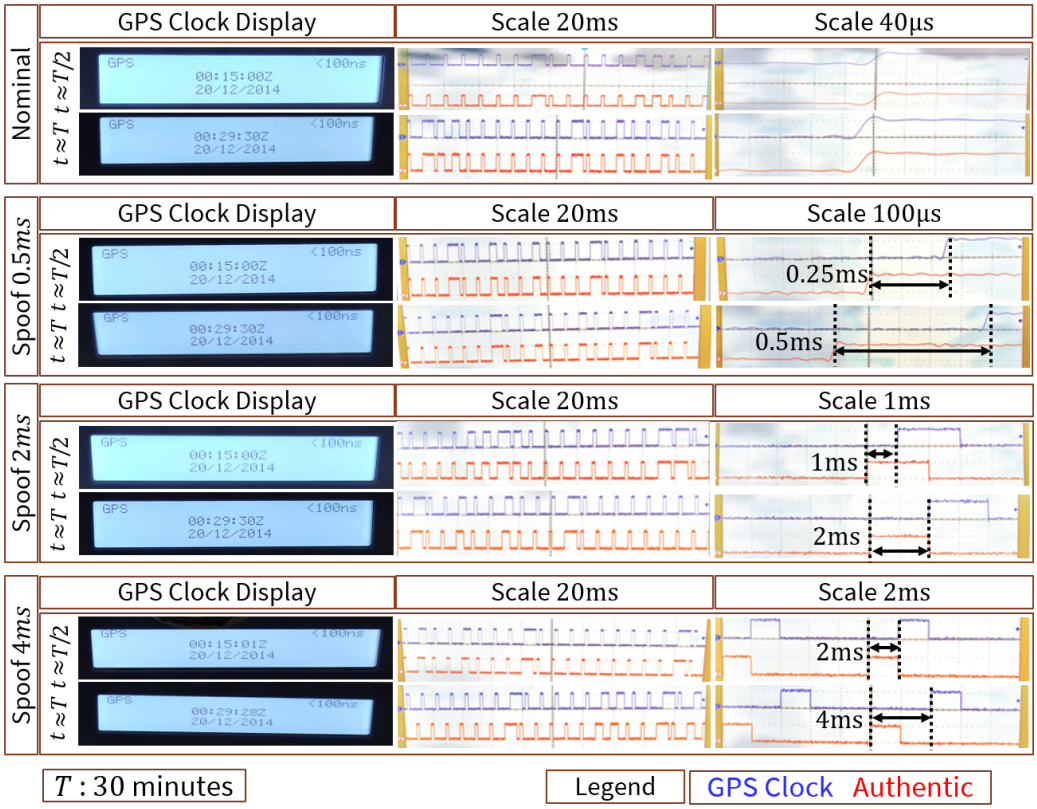


Figure 5.8: Timing signals at two-time instants during the experiment for Nominal and Spoof scenarios.

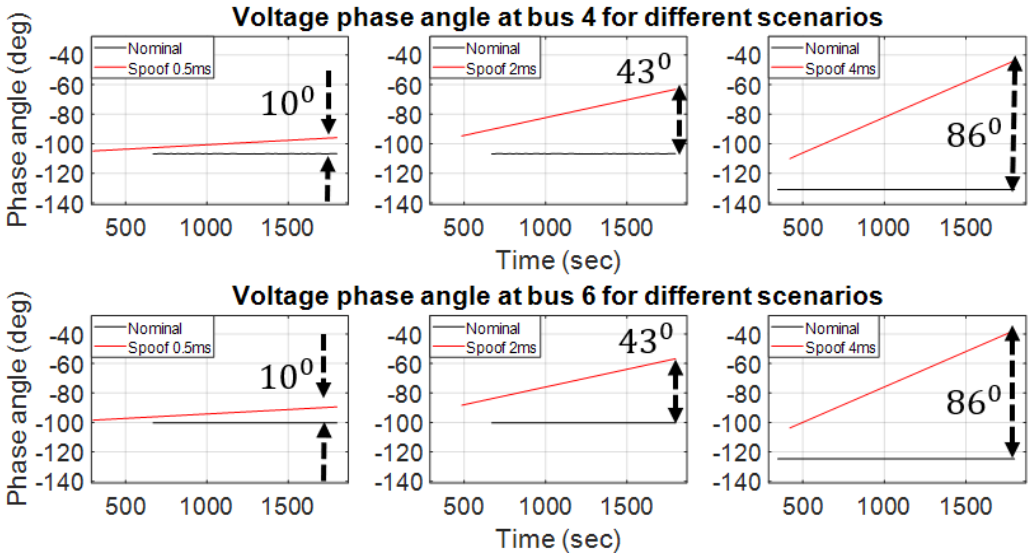


Figure 5.9: Voltage phase angle for spoofed bus 4 and 6 for different spoof scenarios.

5.4 Summary

In this chapter, a methodology was devised to generate GPS and PMU integrated datasets for a virtual power grid. Two scenarios were simulated: Nominal and Spoof. The Nominal scenario referred to an ideal environment in which GPS signals were authentic. In the Spoof scenario, the pseudoranges were modified to simulate timing GSAs with time-walk. Such GSAs are relevant to the power grid as its substations are static. We generated openly available GPS and PMU integrated datasets by performing HIL simulations with RTDS, physical PMUs, virtual PMUs, and GPS clock. The GPS datasets contained raw signals, satellite positions, and pseudoranges. The PMU datasets consisted of GPS timestamped voltage phasors, current phasors, frequency, and frequency change rate measurements for the IEEE-14 bus system. The integrated datasets were validated by demonstrating that the datasets involve timing GSAs with time-walk. The integrated datasets will serve as an evaluation platform for testing the performance of SEs for the power grid.

CHAPTER 6

GPS SPOOFING-RESILIENT STATE ESTIMATION FOR THE POWER GRID USING AN EXTENDED KALMAN FILTER

In the chapter 4, we presented SR-SSE that mitigates multiple GSAs using PMU measurements and also derived a necessary condition that showed GSAs increase PMU measurement residual norm. However, the computation time of SR-SSE increases with the number of GSAs. We observe similar trends with the prior works as the key part of the developed SEs involves minimizing a complex objective function to estimate attack angles for different GSAs [59, 42, 60, 65]. The minimization step is computationally intensive and may not always reach the global minimum. Due to this, the accuracy of SEs decreases with the increase in the number of GSAs. Another limitation of the prior works was the implicit assumption that GSAs introduce a large (> 10 degrees) attack angle. In this chapter, we present a novel SR-SE that mitigates the limitations of prior works. The developed SR-SE meets the following requirements:

- The order of computation time should remain the same for any number of GSAs.
- The order of RMSE of voltage magnitude and phase estimates should remain the same for any number of GSAs.

The remainder of this chapter is organized as follows: Section 6.1 provides a detailed description of SR-SE. This section presents our approach for mitigating multiple GSAs and process and measurement models for implementing an EKF. We validate SR-SE using HIL and MC simulations with different test systems under different GSA scenarios. Section 6.2 describes the simulation environment and experimental scenarios. In Section 6.3, we show the experimental results to validate SR-SE. Finally, Section 6.4 summarizes the chapter.

6.1 Spoofing-Resilient State Estimator

In our approach, we remove the minimization step by incorporating the time-varying GPS and PMU measurements in state estimation. The time-varying GPS measurements enable the SE to track the induced time delay for each PMU, thereby simultaneously tracking the attack angle during a GSA. Our proposed SR-SE jointly estimates power grid states and receiver clock biases using an EKF. The estimated power grid states are resilient to timing GSAs as SR-SE maintains an estimate of GSA induced time delay by tracking receiver clock biases. It is computationally efficient compared to prior works as it fuses measurements sequentially without requiring a computationally intensive minimization step.

6.1.1 Overview of SR-SE

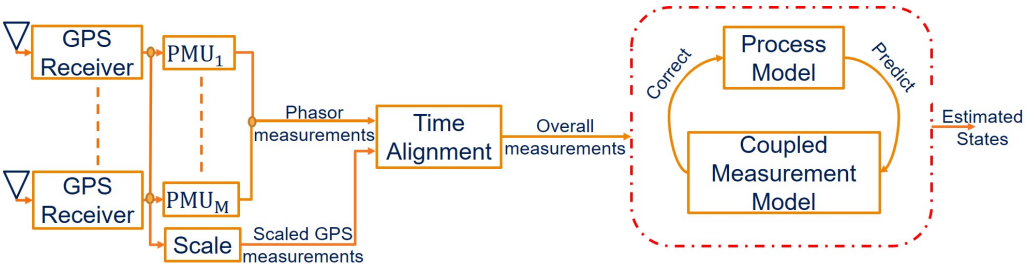


Figure 6.1: Overall architecture of SR-SE.

Figure 6.1 shows the overall architecture of SR-SE. The 'Scale' block in Figure 6.1 scales GPS measurements to avoid numerical instabilities since PMU measurements are on the order of 1 pu, and GPS measurements are on the order of 10^7 m. The time alignment block aligns PMU and GPS measurements by updating the overall measurements based on the time stamp of the latest received measurement. This is a necessary step as the update rates of GPS and PMU measurements are different. SR-SE sequentially performs prediction and measurement update steps using process and coupled measurement models.

Consider a power grid network of N buses with M PMUs installed to ensure observability. We refer to a bus with a PMU installed as a PMU bus. Each PMU uses a GPS receiver that provides GPS time. In SR-SE, we

augment the power grid states with the scaled GPS receiver clock bias state. The augmented state vector is given by

$$\mathbf{x} = \begin{bmatrix} \mathbf{x_v} & \mathbf{x_{clk}} \end{bmatrix}^\top \tag{6.1}$$

where $\mathbf{x} \in \mathbb{R}^{(2N+M)\times 1}$ is the state vector, $\mathbf{x_v} \in \mathbb{R}^{2N\times 1}$ denotes the power grid states which include the complex voltage phasors for each of the N buses in the network, and $\mathbf{x_{clk}} \in \mathbb{R}^{M\times 1}$ is the scaled clock bias of the GPS receivers at the PMU buses. We scale the receiver clock biases to avoid numerical instabilities. We perform a linear scaling and the scaled clock biases are given by

$$\mathbf{x_{clk}} = a(C\boldsymbol{\delta}\mathbf{t}) + b \tag{6.2}$$

where (a, b) are scaling constants that ensure the magnitude of each element in $\mathbf{x_{clk}}$ is less than or equal to one, and $\boldsymbol{\delta}\mathbf{t}$ is given by

$$\boldsymbol{\delta}\mathbf{t} = \begin{bmatrix} \delta t_1 & \cdots & \delta t_M \end{bmatrix}^\top \tag{6.3}$$

where δt_i denotes the receiver clock bias at the i^{th} PMU bus. The subsequent section outlines the details of process and measurement models for EKF implementation.

6.1.2 Process Model

We assume the power grid is operating in a quasi-steady state [90] in which voltage phasors change due to slow and smooth load generation changes. Under this assumption, voltage phasors become constant between two sets of PMU measurements. The following process model is used for the voltage phasors

$$\mathbf{x}_{\mathbf{v},k+1} = \mathbf{x}_{\mathbf{v},k} + \boldsymbol{\omega}_v \tag{6.4}$$

where k denotes the time instant and $\boldsymbol{\omega}_v$ is zero-mean Gaussian noise. In an unspoofed scenario, the clock bias of a stationary receiver varies slowly [1]. The following process model is used for scaled clock biases

$$\mathbf{x}_{\mathbf{clk},k+1} = \mathbf{x}_{\mathbf{clk},k} + \boldsymbol{\omega}_{clk} \tag{6.5}$$

where $\boldsymbol{\omega}_{clk}$ is zero-mean Gaussian noise. The overall process model is given by

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \boldsymbol{\omega} \tag{6.6}$$

where $\mathbf{x} = [\mathbf{x_v}\ \mathbf{x_{clk}}]^\top$ and $\boldsymbol{\omega} = [\boldsymbol{\omega}_v\ \boldsymbol{\omega}_{clk}]^\top$ is zero-mean Gaussian noise.

6.1.3 Measurement Model

We assume the positions of the PMUs are known and the power grid is observable. These assumptions reduce the number of unknowns in the pseudorange equations. The following subsections describe the coupled GPS and PMU measurement model that helps in mitigating timing GSAs.

GPS Measurement Model

A conventional GPS receiver [1, 41] executes two steps for estimating a PVT solution. In the first step, scalar acquisition and tracking estimate channel-specific parameters such as the code delay and Doppler shift for each visible satellite. The second step involves solving a set of pseudorange equations to obtain a positioning and timing solution. We use pseudoranges from GPS receiver as measurements in SR-SE. Without loss of generality, consider that $S_n \in \mathbb{N}$ is the number of visible satellites at the n^{th} PMU bus. The pseudorange equation is given by

$$\rho_n^j = d_n^j + C(\delta t_n - \delta t^j) + \eta_{clk_n}^j \tag{6.7}$$

where subscript n refers to the variables associated with the n^{th} PMU bus, ρ^j denotes pseudorange measurement between the receiver and the j^{th} visible satellite, d^j is the true range between the j^{th} visible satellite and the receiver, C is the speed of light, δt is the receiver clock bias, δt^j is the clock bias for the j^{th} satellite, and η_{clk}^j is zero-mean Gaussian noise. Because the PMU is stationary at a well-known position and because the satellite clock bias is provided to the PMU via external channels, the only unknown variable in (6.7) is the receiver clock bias. We scale (6.7) using the scaling constants (a, b) from (6.2) to obtain a relationship between the n^{th} scaled clock bias state (x_{clk_n}) and the pseudorange measurement.

$$
\begin{aligned}
a\rho_n^j &= ad_n^j + aC(\delta t_n - \delta t^j) + b - b + a\eta_{clk_n}^j \\
&= ad_n^j + (aC(\delta t_n) + b) - aC\delta t^j - b + a\eta_{clk_n}^j \\
&= ad_n^j + x_{clk_n} - aC\delta t^j - b + a\eta_{clk_n}^j \\
&= h_{GPS_n}^j(x_{clk_n}) + a\eta_{clk_n}^j
\end{aligned}
\tag{6.8}
$$

where x_{clk_n} denotes the scaled clock bias at the n^{th} PMU bus and $h_{GPS_n}^j(x_{clk_n})$ is a linear scalar function which is defined as

$$
h_{GPS_n}^j(x_{clk_n}) = x_{clk_n} + a(d_n^j - C\delta t^j) - b \tag{6.9}
$$

The overall GPS measurements at the n^{th} PMU bus is given by

$$
\begin{aligned}
\mathbf{z}_{GPS_n} =& a\begin{bmatrix}\rho_n^1 & \cdots & \rho_n^{S_n}\end{bmatrix}^\top \\
=& [h_{GPS_n}^1(x_{clk_n}) + a\eta_{clk_n}^1 \cdots \\
& h_{GPS_n}^{S_n}(x_{clk_n}) + a\eta_{clk_n}^{S_n}]^\top \\
=& \mathbf{h}_{GPS_n}(x_{clk_n}) + a\boldsymbol{\eta}_{clk_n}
\end{aligned}
\tag{6.10}
$$

where $\mathbf{z}_{GPS_n}$ denotes the GPS measurements at the n^{th} PMU bus, $\boldsymbol{\eta}_{clk_n}$ is zero-mean Gaussian noise, and $\mathbf{h}_{GPS_n}$ is a linear vector function given by

$$
\mathbf{h}_{GPS_n}(x_{clk_n}) = \begin{bmatrix}h_{GPS_n}^1(x_{clk_n}) & \cdots & h_{GPS_n}^{S_n}(x_{clk_n})\end{bmatrix}^\top \tag{6.11}
$$

The combined GPS measurements for the entire network is given by

$$
\begin{aligned}
\mathbf{z}_{GPS} &= \begin{bmatrix}\mathbf{z}_{GPS_1}^\top & \cdots & \mathbf{z}_{GPS_M}^\top\end{bmatrix}^\top \\
&= \mathbf{h}_{GPS}(\mathbf{x_{clk}}) + a\boldsymbol{\eta}_{clk}
\end{aligned}
\tag{6.12}
$$

where $\mathbf{z}_{GPS}$ denotes the GPS measurements of all receivers at M PMU buses, $\boldsymbol{\eta}_{clk}$ is zero-mean Gaussian noise, and $\mathbf{h}_{GPS}(\mathbf{x_{clk}}) = \begin{bmatrix}\mathbf{h}_{GPS_1}(x_{clk_1})^\top & \cdots & \mathbf{h}_{GPS_M}(x_{clk_M})^\top\end{bmatrix}^\top$. Equation (6.12) is the linear GPS measurement model.

PMU Measurement Model

The power grid state $\mathbf{x_v} \in \mathbb{R}^{2N\times 1}$ consists of

$$\begin{aligned}\mathbf{x_v} =&[Re(U_1), \cdots, Re(U_i), \cdots, Re(U_N),\\ &Im(U_1), \cdots, Im(U_i), \cdots, Im(U_N)]^\top\end{aligned} \tag{6.13}$$

where $Re(U_i)$ and $Im(U_i)$ denote the real and imaginary parts of the complex voltage U_i at the i^{th} bus, respectively. Each PMU measures voltage and current phasors, which are given by

$$\begin{aligned}\mathbf{z}_{PMU_i} =&[Re(U_i), Im(U_i), Re(I_{i1}), \cdots,\\ &Re(I_{ik}), Im(I_{i1}), \cdots, Im(I_{ik})]^\top\end{aligned} \tag{6.14}$$

where $\mathbf{z}_{PMU_i}$ denotes the PMU measurement at the i^{th} PMU bus that connects to k other buses, and $Re(I_{ik})$ and $Im(I_{ik})$ are the real and imaginary parts of the injected current phasors at the line connecting i^{th} and k^{th} buses, respectively. The PMU measurements at the i^{th} PMU bus are related to the power grid states by

$$\mathbf{z}_{PMU_i} = \mathbf{H}_{PMU_i}\mathbf{x_v} + \boldsymbol{\eta}_{PMU_i} \tag{6.15}$$

where $\mathbf{H}_{PMU_i}$ denotes the regression matrix associated with the i^{th} bus [84] and $\boldsymbol{\eta}_{PMU_i}$ is zero-mean Gaussian noise. The regression matrix for the i^{th} bus relates its voltage phasors and current phasors with voltage phasors of the entire network. The construction of the regression matrix is given in [84, 83].

The induced time delay under a timing GSA shifts the phase angle of the phasor measurements by an attack angle. Zhang *et al.* [46] showed that timing GSAs induce the same attack angle to all the PMU measurements. Assume the i^{th} PMU bus is attacked. Under a timing GSA, (6.15) is modified as

$$\begin{aligned}\mathbf{z}_{PMU_i} =&[|U_i|\cos(\theta_i+\Delta\theta_i), |U_i|\sin(\theta_i+\Delta\theta_i),\\ &|I_{i1}|\cos(\theta_{i1}+\Delta\theta_i), |I_{i1}|\sin(\theta_{i1}+\Delta\theta_i),\cdots,\\ &|I_{ik}|\cos(\theta_{ik}+\Delta\theta_i), |I_{ik}|\sin(\theta_{ik}+\Delta\theta_i)]^\top\end{aligned} \tag{6.16}$$

where θ_i is the phase angle of the i^{th} bus, θ_{ik} is the phase angle of the line connecting the i^{th} and the k^{th} buses, and $\Delta\theta_i$ is the attack angle. The following equation relates the attack angle at the i^{th} PMU bus to the GSA induced time delay

$$\begin{aligned}\Delta\theta_i &= 2\pi f\delta t_i\\ &= 2\pi f\left(\frac{x_{clk_i}-b}{aC}\right)\end{aligned} \tag{6.17}$$

where f denotes the frequency of the current, δt_i is the receiver's clock bias at the i^{th} PMU bus, (a, b) are the scaling constants from (6.2), and x_{clk_i} denotes the scaled clock bias at the i^{th} PMU bus. Effective timing GSAs modify the receiver time by adding bias to pseudorange measurements, thereby inducing errors in the receiver clock bias. Under timing GSAs, the receiver clock bias is approximately the same as the induced time delay. This approximation is valid when induced time delays are greater than a few microseconds, which is the case under timing GSAs. Risbud *et al.* [42] derived a linear relationship between spoofed and authentic measurements using cosine identities. This relationship is given by

$$\mathbf{z}_{PMU_i} = \boldsymbol{\gamma}_i(x_{clk_i})\mathbf{H}_{PMU_i}\mathbf{x_v} + \boldsymbol{\eta}_{PMU_i} \tag{6.18}$$

where $\boldsymbol{\gamma}_i(x_{clk_i})$ is a block diagonal matrix, given by the following equation:

$$\boldsymbol{\gamma}_i(x_{clk_i}) = \begin{bmatrix}\mathbf{G}_i(x_{clk_i}) & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{G}_i(x_{clk_i}) & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{G}_i(x_{clk_i})\end{bmatrix} \tag{6.19}$$

where $\mathbf{G}_i(x_{clk_i})$ is given by

$$\mathbf{G}_i(x_{clk_i}) = \begin{bmatrix} \cos\left(2\pi f\left(\frac{x_{clk_i}-b}{aC}\right)\right) & -\sin\left(2\pi f\left(\frac{x_{clk_i}-b}{aC}\right)\right) \\ \sin\left(2\pi f\left(\frac{x_{clk_i}-b}{aC}\right)\right) & \cos\left(2\pi f\left(\frac{x_{clk_i}-b}{aC}\right)\right) \end{bmatrix} \tag{6.20}$$

The PMU measurements for the entire power grid network are given by

$$\mathbf{z}_{PMU} = \mathbf{\Gamma}(\mathbf{x_{clk}})\mathbf{H}_{PMU}\mathbf{x_v} + \boldsymbol{\eta}_{PMU} \tag{6.21}$$

where $\mathbf{z}_{PMU} = [\mathbf{z}_{PMU_1}^\top \ \cdots \ \mathbf{z}_{PMU_M}^\top]^\top$ denotes the PMU measurements for the entire power grid network, $\mathbf{H}_{PMU} = [\mathbf{H}_{PMU_1}^\top \ \cdots \ \mathbf{H}_{PMU_M}^\top]^\top$ is the overall regression matrix, $\boldsymbol{\eta}_{PMU} = [\boldsymbol{\eta}_{PMU_1}^\top \ \cdots \ \boldsymbol{\eta}_{PMU_M}^\top]^\top$ is zero-mean Gaussian noise, and $\mathbf{\Gamma}(\mathbf{x_{clk}})$ is a block diagonal matrix, given by the following equation:

$$\mathbf{\Gamma}(\mathbf{x_{clk}}) = \begin{bmatrix} \boldsymbol{\gamma}_1(x_{clk_1}) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\gamma}_2(x_{clk_2}) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\gamma}_M(x_{clk_M}) \end{bmatrix} \tag{6.22}$$

Equation (6.21) shows the overall PMU measurement model, which is valid in the presence and absence of GSAs. Under GSAs, the induced time delay will be greater than a few microseconds, which will make the off-diagonal elements of $\mathbf{G}_i$ non-zero. However, in the absence of GSAs, the clock bias will be on the order of a few nanoseconds making $\mathbf{G}_i$ an identity matrix and correspondingly resulting in an identity $\mathbf{\Gamma}$ matrix.

The overall measurement of SR-SE consists of GPS and PMU measurements. The GPS-PMU coupled measurement model is given by

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \boldsymbol{\eta} \tag{6.23}$$

where $\mathbf{z} = \begin{bmatrix} \mathbf{z}_{GPS} \\ \mathbf{z}_{PMU} \end{bmatrix}$ denotes the overall measurements for the entire power grid network, $\mathbf{h}(\mathbf{x}) = \begin{bmatrix} \mathbf{h}_{GPS}(\mathbf{x_{clk}}) \\ \mathbf{\Gamma}(\mathbf{x_{clk}})\mathbf{H}_{PMU}\mathbf{x_v} \end{bmatrix}$ is the nonlinear vector function which relates states with measurements and $\boldsymbol{\eta} = \begin{bmatrix} a\boldsymbol{\eta}_{clk} \\ \boldsymbol{\eta}_{PMU} \end{bmatrix}$ is zero-mean Gaussian noise.

6.1.4 SR-SE Implementation

An EKF consists of a prediction and a measurement update step, which utilizes the process and measurement models. Since the process model is linear, the prediction step is given by

$$\mathbf{x}_{k|k-1} = \mathbf{x}_{k-1|k-1} \tag{6.24}$$

$$\mathbf{P}_{k|k-1} = \mathbf{P}_{k-1|k-1} + \mathbf{Q} \tag{6.25}$$

where $\mathbf{x}_{p|q}$ denotes the state $\mathbf{x}$ at the p^{th} time instant given the measurements till the q^{th} time instant, $\mathbf{P}$ denotes the state covariance matrix, and $\mathbf{Q} = \mathbb{E}[\boldsymbol{\omega}\boldsymbol{\omega}^\top]$ is the process noise covariance matrix, which is a diagonal matrix. New measurements allow the SR-SE to update the states using the process and measurement models. The following equations are used in the measurement update step

$$\tilde{\mathbf{y}}_k = \mathbf{z}_k - \mathbf{h}(\mathbf{x}_{k|k-1}) \tag{6.26}$$

where $\mathbf{z}$ is the overall measurements, $\mathbf{h}$ is the nonlinear measurement model from (6.23), and $\tilde{\mathbf{y}}$ is the innovation sequence. The Kalman gain matrix ($\mathbf{K}$) is calculated using the state and measurement covariance matrices as

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^\top \left(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^\top + \mathbf{R}\right)^{-1} \tag{6.27}$$

where $\mathbf{R} = \mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^\top]$ is the measurement noise covariance matrix, which is a diagonal matrix, and $\mathbf{H}$ is the Jacobian of the nonlinear measurement model $\mathbf{h}(\mathbf{x})$. The state vector is corrected using the following equation:

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{K}_k\tilde{\mathbf{y}}_k \tag{6.28}$$

where $\mathbf{x}_{k|k}$ is the posterior or the corrected state. The convergence of SR-SE depends on the $\mathbf{Q}$ and $\mathbf{R}$ matrices. We perform empirical tuning of these matrices using the expected noise levels of the GPS signals [1] and PMU measurements [42].

6.2 Simulation Environment and Experimental Setup

It is illegal to transmit signals at GPS frequency without approval from the U.S. government, making real-world spoofing experiments challenging. Furthermore, testing on a physical power grid network is costly and time-consuming. As a result, we perform both low- and high- fidelity simulations.

Low-fidelity simulations are mainly software-based and do not incorporate physical hardware in simulations. This type of simulations is ideal for performing MC simulations as it approximates the real-world, allowing it to perform simulations faster than real-time. High-fidelity simulations incorporate physical hardware in simulations and simulate an environment closer to the real-world.

In low-fidelity simulations, we simulate the steady-state power grid using the MATPOWER [83] tool. In high-fidelity simulations, we simulate the IEEE 14-bus test system using RTDS, physical PMUs, and GPS clock. The following subsections provide details for low- and high- fidelity simulations.

6.2.1 Low-Fidelity Simulations

In low-fidelity simulations, we generate 10 seconds of GPS and PMU measurements using measurement models described in Section 6.1.3. We generate these measurements for the IEEE 14, IEEE 39, and Illinois 200-bus [68] test systems.

Generating GPS Measurements

Generating pseudorange measurements from (6.7) requires satellite positions, PMU positions, and satellite clock biases. We simulate four stationary satellites with known clock biases in the same 2-dimensional plane as PMU devices for all the test systems. Table 6.1 specifies the positions of these satellites. The PMUs' positions for the IEEE 14 and 39-bus test systems are uniformly generated from a 10×10 km^2 area. The synthetic Illinois 200 bus [68] is based on the central Illinois power grid network. The synthetic buses' positions are specified in [68], which are spread all over central Illinois.

Table 6.1: Positions of simulated virtual satellites.

Satellite number	East (km)	North (km)
1	-26×10^3	30×10^3
2	26×10^3	-30×10^3
3	26×10^3	-30×10^3
4	-26×10^3	30×10^3

Generating PMU Measurements

We perform power flow analysis using MATPOWER [83] to obtain steady states of each test system from which we generate PMU measurements using (6.21). Table 6.2 lists the PMU buses for all the test systems. These PMU buses ensure the observability of each test system.

Table 6.2: PMU buses for IEEE 14, IEEE 39 and Illinois 200-bus test systems.

Test System	**Number of PMUs (M)**	**PMU buses**
IEEE 14	8	[1,2,4-6,7,10,13]
IEEE 39	20	[1-6,8,10,12,14-16,19, 20,21,22,23,25,26,29]
Illinois 200	136	[1,2,4,6,8-13,15-30,32,33,35,37-41, 43-45,47-53,55-63,65,67-73,75-80,82, 83,86,87,89-94,99,101,103-105,107, 108,110,113-115,117,118,122,123, 125-127,130,131,135-138,145-149, 151-155,157,161,163-170,173,174,176, 178,180-183,185,186,189,190,195,196,197]

6.2.2 High-Fidelity Simulations

In chapter 5, we simulated timing GSAs by shifting all of the visible satellites equally along the line-of-sight direction. We achieved timing GSAs by adding equal biases to all of the pseudoranges. The modified pseudoranges are utilized to generate GPS datasets using an openly available SDR receiver, GPS-SDR-SIM [87]. We generated a nominal GPS dataset, in which none of the pseudoranges were modified, as well as spoofing GPS datasets that induced total time delays of 0.5, 2, and 4 ms, respectively. In spoofed

GPS datasets, the time delay is linearly increased to avoid sudden jumps in the GPS timing to reduce the risk of detection. We analyze the generated datasets using our SDR, PyGNSS [80], with which we performed scalar tracking [1] to obtain positioning and timing solutions. Figures 5.7 and 5.8 showed the induced time delays and positioning solutions for each of the generated GPS datasets. From these figures, we observed that the different GPS datasets' positioning solutions coincided with each other. As a result, the generated GPS datasets mimic GSAs that induce time delays without altering the receiver's positioning solution.

In the high-fidelity simulation, we simulated the IEEE 14-bus test system in RTDS with virtual PMUs, physical PMUs, and GPS clock. Table 6.2 specifies the PMU buses for the IEEE 14-bus test system. Figure 5.6 illustrates the experimental setup for the HIL simulations. The setup consists of two physical PMUs, six virtual PMUs, a commercial GPS clock, and two Universal Software Radio Peripherals (USRPs). The physical PMUs are at buses 4 and 6, while the remaining PMUs are virtual.

6.3 Experimental Results

6.3.1 Monte Carlo Simulation Results

We test SR-SE for three GSA scenarios in which 25%, 50%, and 100% of the total number of PMU buses are spoofed, respectively. We perform 100 MC simulations for each GSA scenario. Each simulation generates 10 seconds of GPS and PMU measurements. In each MC simulation, for a given GSA scenario, we randomly select PMU buses from all PMU buses and spoof them by adding signed biases to the pseudoranges. The sign, positive or negative, is randomly chosen, and it determines whether the added bias is linearly increasing or decreasing. For each selected PMU bus, we start to add biases at random start times, which lie between 0 and 10 seconds.

We perform MC simulations with IEEE 14, IEEE 39, and Illinois 200-bus test systems. We compare the RMSE and computation time of our proposed SR-SE with that of SpM [59] and a conventional PMU-based SSE [91]. The RMSE is computed using the entire power grid state, i.e., for all N buses. Figures 6.2, 6.3, and 6.4 show the RMSE box plots of voltage and

phase estimates for the three GSA scenarios. In these box plots, the red lines indicate the median, the blue boxes bound the first and third quantiles, the black whiskers indicate the 1.5× inter-quantile range, and the red crosses denote the outliers. Outliers are the data points that lie outside the 1.5× inter-quantile range.

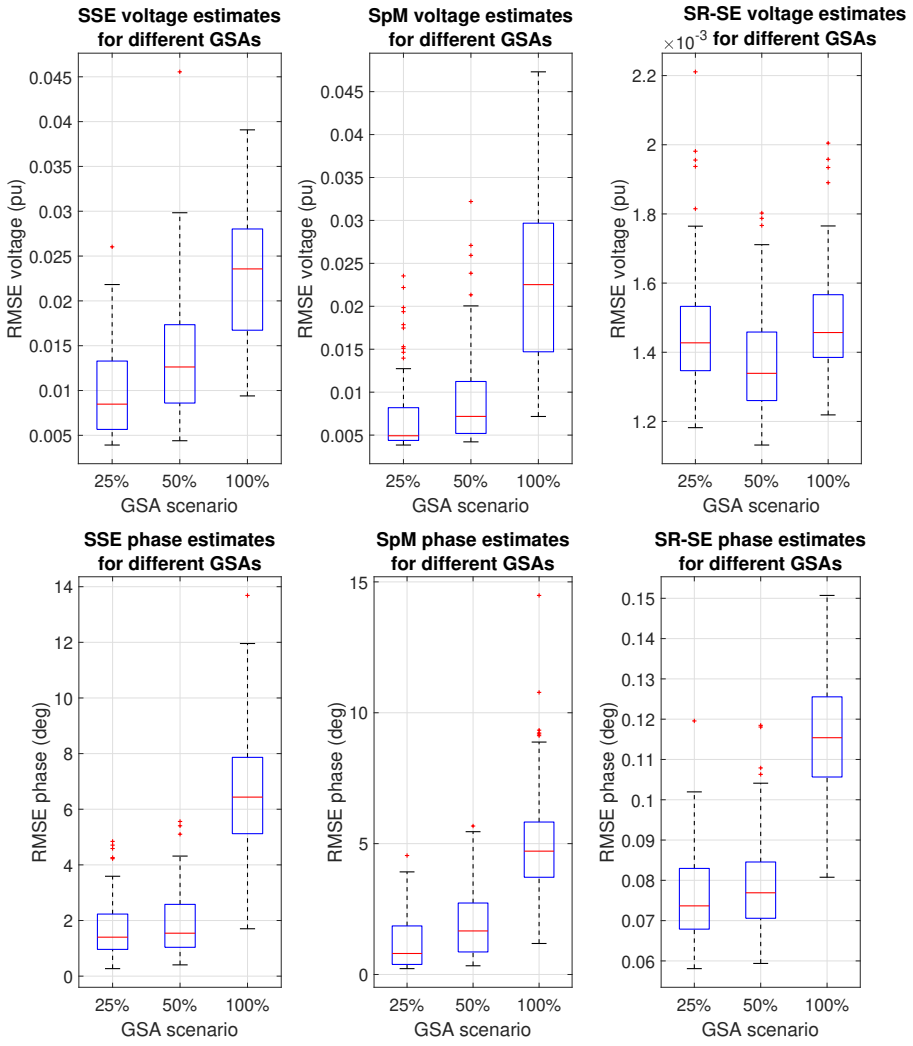


Figure 6.2: Voltage and phase RMSE of the SSE (first column), SpM (second column), and SR-SE (third column) for the IEEE 14-bus test system. SR-SE estimates are an order of magnitude more accurate than the SSE and SpM algorithm.

Figure 6.2 shows the RMSE box plot of voltage and phase estimates for the IEEE 14-bus test system. The SR-SE voltage and phase estimates, shown in Figure 6.2, are an order of magnitude more accurate than SSE and SpM estimates. We observe a similar trend for IEEE 39-bus test system from

Figure 6.3. Additionally, for the Illinois 200-bus test system, we observe that the SR-SE phase estimates are an order of magnitude more accurate than SSE and SpM.

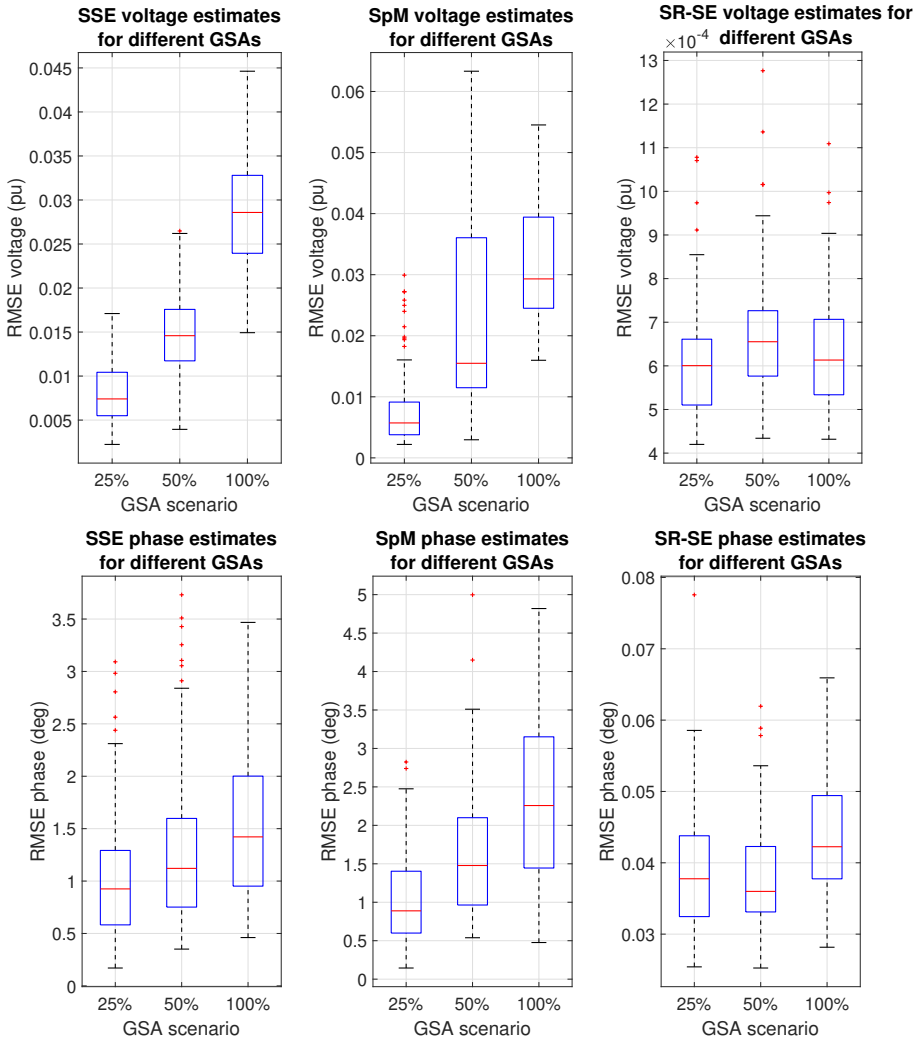


Figure 6.3: Voltage and phase RMSE of the SSE (first column), SpM (second column), and SR-SE (third column) for the IEEE 39-bus test system. SR-SE estimates are an order of magnitude more accurate than the SSE and SpM algorithm.

Table 6.3 provides the median RMSE and computation time of SSE, SpM, and SR-SE for all test systems under the three GSA scenarios. We observe in Table 6.3 that the SR-SE phase estimates are at least an order of magnitude more accurate than the SSE and SpM estimates for all test systems. The RMSE of the voltage estimate for SR-SE is smaller than the SSE and SpM voltage estimates for all test systems. SSE has the lowest computa-

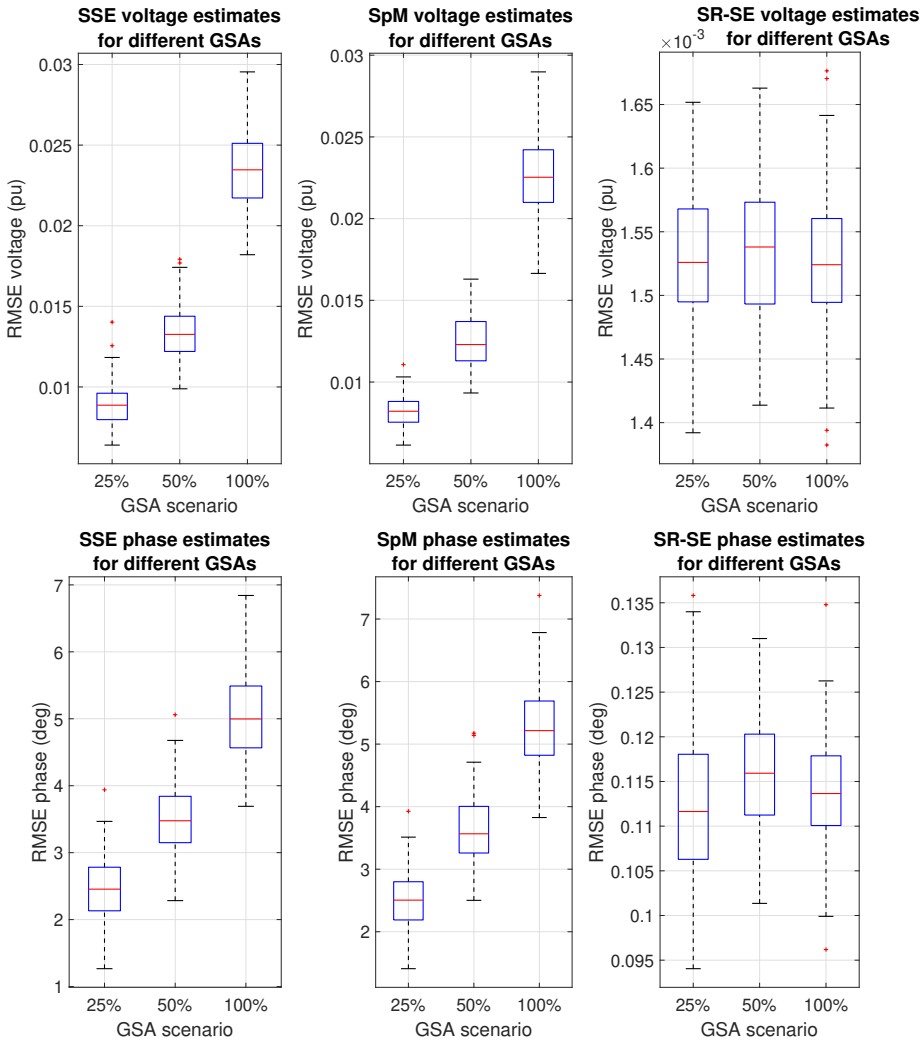


Figure 6.4: Voltage and phase RMSE of the SSE (first column), SpM (second column), and SR-SE (third column) for the Illinois 200-bus test system. SR-SE phase estimates are an order of magnitude more accurate than the SSE and SpM algorithm.

tion time, but its voltage and phase estimates degrade under an increasing number of GSAs. The sequential nature of SR-SE makes it computationally efficient compared to SpM, which minimizes a complex objective function. For each test system, the RMSE of the SR-SE estimates and computation time remains consistent for all GSA scenarios.

Due to the centralized nature of SR-SE, the computation time increases with the network's size, as observed in Table 6.3. A distributed state estimator would be more efficient for networks with a few thousand buses.

Table 6.3: Median RMSE and computation time of the SSE, SpM, and SR-SE for the IEEE 14, IEEE 39, and Illinois 200-bus test systems.

	Test System	**GSA scenario (% spoofed buses)**	Voltage Magnitude (pu)	Phase (deg)	Computation Time (sec)
SSE	**IEEE 14**	25	0.0085	1.4009	0.0001
		50	0.0126	1.5438	0.0001
		100	0.0236	6.4359	0.0001
	IEEE 39	25	0.0074	0.9252	0.0004
		50	0.0146	1.1210	0.0004
		100	0.0286	1.4216	0.0004
	Illinois 200	25	0.0089	2.4546	0.0020
		50	0.0133	3.4764	0.0020
		100	0.0235	4.9973	0.0019
SpM	**IEEE 14**	25	0.0049	0.8027	0.0421
		50	0.0072	1.6644	0.0432
		100	0.0225	4.7125	0.0404
	IEEE 39	25	0.0057	0.8883	0.2936
		50	0.0155	1.4779	0.2867
		100	0.0293	2.2575	0.2673
	Illinois 200	25	0.0082	2.5055	22.5612
		50	0.0123	3.5661	22.0030
		100	0.0225	5.2157	21.0423
SR-SE	**IEEE 14**	25	0.0014	0.0737	0.0032
		50	0.0013	0.0769	0.0032
		100	0.0015	0.1154	0.0032
	IEEE 39	25	0.0006	0.0378	0.0095
		50	0.0007	0.0360	0.0095
		100	0.0006	0.0423	0.0091
	Illinois 200	25	0.0015	0.1116	0.5982
		50	0.0015	0.1159	0.5993
		100	0.0015	0.1136	0.5996

6.3.2 Hardware Experimental Results

We test SR-SE on the three GPS-PMU integrated datasets that induce a total time delay of 0.5, 2, and 4 ms. We compare the RMSE of SR-SE estimates with SSE and SpM estimates. Table 6.4 provides the RMSE of voltage and phase estimates, and median computation times. From Table 6.4, we observe that for all three datasets

- The phase estimates provided by SR-SE are at least an order of magnitude more accurate than the SSE and SpM estimates.

- The RMSE of the SR-SE voltage estimates is smaller than the estimates from SSE and SpM.
- The computation time of SR-SE is an order of magnitude smaller than that of SpM.

Table 6.4: RMSE and median computation time of the SSE, SpM, and SR-SE for HIL simulations

	Total induced time delay	Voltage Magnitude (pu)	Phase (deg)	Computation Time (sec)
SSE	0.5 ms	0.0017	3.1299	0.0050
	2 ms	0.0157	13.2819	0.0027
	4 ms	0.0560	26.0038	0.0053
SpM	0.5 ms	0.0006	3.6902	0.0439
	2 ms	0.0006	14.9523	0.0154
	4 ms	0.0005	28.9547	0.0146
SR-SE	0.5 ms	0.0001	0.0734	0.0078
	2 ms	0.0002	0.2982	0.0073
	4 ms	0.0003	0.5925	0.0063

Figure 6.5 shows the estimated voltage magnitude and phase angle of the spoofed PMU bus 6. The first, second, and third columns of Figure 6.5 show the voltage and phase estimation results for GPS-PMU integrated datasets that induce a total time delay of 0.5, 2, and 4 ms. SR-SE phase estimates are closer to the truth. For each dataset, the SpM phase estimate oscillates due to its inability to reach the global minimum.

From Table 6.4 and Figure 6.5, we observe that SR-SE provides GSA resilient power grid states without compromising on the computation time for the IEEE 14-bus test system.

6.4 Summary

In this chapter, we presented a novel SE that addresses the limitations of the prior works. The proposed SR-SE fuses time-varying GPS and PMU measurements using an EKF. We designed a coupled measurement model that relates the GPS and PMU measurements. SR-SE jointly estimates the power grid states and receiver clock biases, thereby providing power grid state estimates that are resilient to timing GSAs.

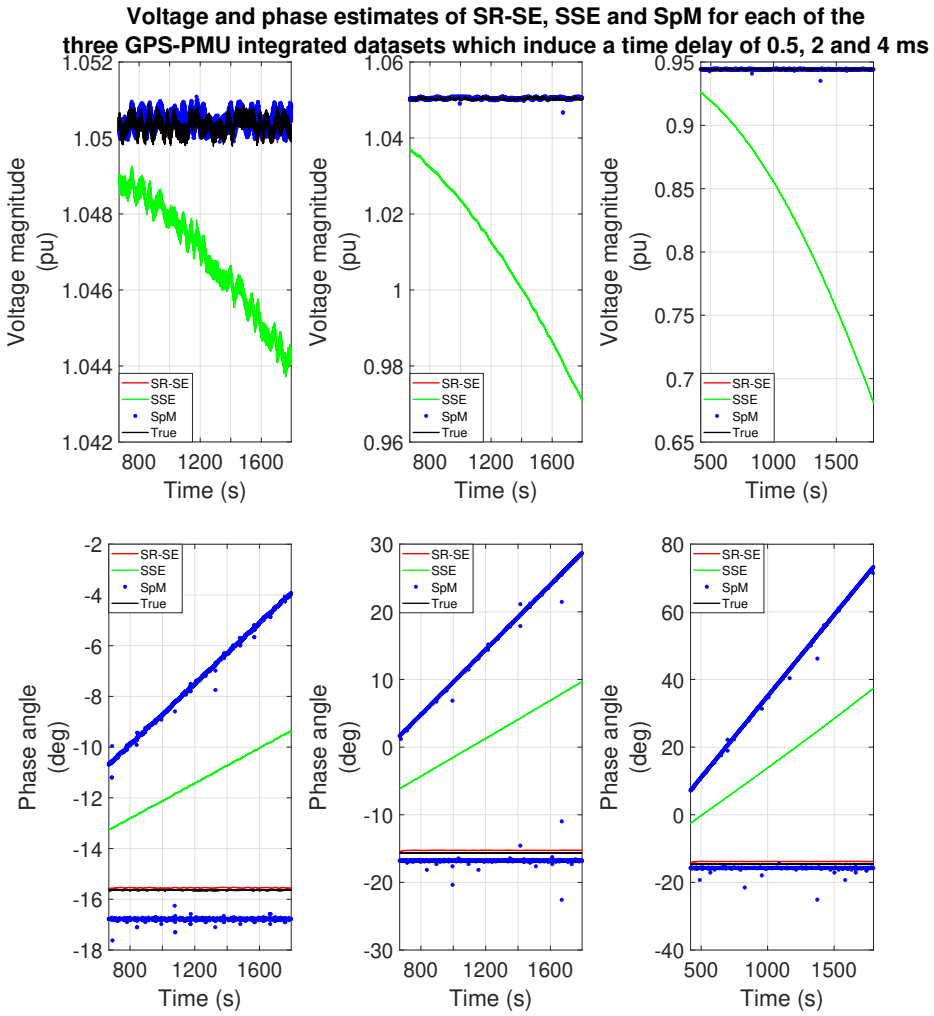


Figure 6.5: Estimated voltage magnitudes and phase angles of SR-SE, SSE, and SpM for each of the three GPS-PMU integrated datasets that induce a time delay of 0.5, 2, and 4 ms.

To validate SR-SE, we performed MC and HIL simulations. In MC simulations, we simulated IEEE 14, IEEE 39, and Illinois 200-bus test systems for different GSA scenarios and observed that SR-SE achieved greater accuracy than SSE and SpM for all test systems. In HIL simulations, we tested SR-SE on the generated GPS-PMU integrated datasets and observed at least an order of magnitude higher accuracy of phase estimates than SSE and SpM for all datasets.

The computation time of SR-SE is comparable with that of SSE for the IEEE 14 and 39-bus test systems. However, the computation time increases with the network's size due to the centralized nature of SR-SE. The next

step would be to explore distributed SEs, which will be efficient for grids with more than a few thousand buses.

CHAPTER 7

CONCLUSIONS

This dissertation presented algorithms that mitigate the limitations of GPS positioning and timing service. We developed algorithms that fuse multi-sensor and multi-receiver measurements using a Bayesian approach to improve GPS positioning in urban environments and provide resiliency against GSAs to the power grid's SEs.

The developed algorithms are tested in simulations as well as through real-world experiments. We created the first experimental datasets for GPS and PMU measurements under GSAs. The contributions of this dissertation are summarized below

- Chapter 2 presented an adaptive sensor fusion algorithm that adaptively estimates the time- and size-varying noise parameters for process and measurement noise covariance matrices. We described a sensor fusion algorithm that fuses GPS and vision measurements to improve positioning in urban environments. We tested the developed adaptive sensor fusion algorithm with simulated and real-world data. We showed that the developed algorithm improves positioning in urban environments compared to GPS only, vision only, and sensor fusion with fixed covariance matrices.

- Chapter 3 described a novel DP receiver that directly works in the position domain. We designed a novel Bayesian algorithm to estimate VPLs for DP that utilizes both PVT and variance estimates. The designed algorithm is robust to the unknown number of modes present in the vertical positioning error distribution. We validated the designed algorithm using a high-fidelity GPS simulator. We generated 24 hours of stationary GPS dataset and obtained 4 million vertical positioning error data points. This chapter showed that DP's vertical positioning error distribution is multi-modal and further validated that the

designed algorithm overbounds vertical positioning errors.

- Chapter 4 identified the unexplored areas for GSA-resilient SEs. It introduced PMU-based SSE and the negative impact of GSAs on SSE. We proposed a novel residual-based SR-SSE for the power grid, which is resilient to multiple GSAs with different attack angles. SR-SSE consisted of two algorithms: Spoofing Detection and Measurement Correction. We performed a theoretical analysis detailing GSAs' impact on residuals and derived a mathematically necessary condition that ensures an increase in residual norm during GSAs. We validated SR-SSE and verified our derived necessary condition by performing MC simulations on the IEEE 14, IEEE 39, IEEE 118, and Illinois 200-bus test systems for different GSAs.

- Chapter 5 described a methodology for generating GPS and PMU measurements under nominal and spoof scenarios. The nominal scenario represented an ideal environment in which GPS signals are authentic. In the spoof scenario, GPS signals were modified to mimic a timing GSA that modifies receiver time without altering the receiver location. Using the devised methodology, we generated openly available GPS and PMU integrated datasets by performing HIL simulations with RTDS, physical PMUs, virtual PMUs, and GPS clock. We demonstrated that the integrated datasets involved timing GSAs with time-walk. The integrated datasets will serve as an evaluation platform for testing the performance of SEs for the power grid.

- Chapter 6 outlined SR-SE that fuses GPS and PMU measurements using an EKF. Compared to prior works, we remove the minimization step for obtaining GSA-resilient states by incorporating time-varying GPS and PMU measurements in state estimation. The time-varying GPS measurements enabled the SE to track the induce time delay for each PMU, thereby simultaneously tracking the attack angle during a GSA. In SR-SE, we designed a GPS-PMU coupled measurement model that relates GSA induced time delay to PMU measurements. This measurement model is essential to maintain estimates of attack angles, which is necessary to mitigate GSAs.

SR-SE is validated using MC and HIL simulations. In MC simulations, we simulated IEEE 14, IEEE 39, and Illinois 200-bus test systems for different GSAs and observed that SR-SE achieved greater accuracy than SSE and SpM. In HIL simulations, we utilized the generated integrated datasets and observed at least an order of magnitude higher accuracy of phase estimates than SSE and SpM.

REFERENCES

[1] P. Misra and P. Enge, "Global positioning system: signals, measurements and performance second edition," *Global Positioning System: Signals, Measurements And Performance Second Editions,*, 2006.

[2] M. Graham, "GPS Use in U.S. Critical Infrastructure and Emergency Communications," (Accessed: 2020, Sep 23). [Online]. Available: https://www.gps.gov/multimedia/presentations/2012/10/USTTI/graham.pdf

[3] A. Silverstein, "Synchrophasors and the Grid," (Accessed: 2019, July 10). [Online]. Available: https://www.naspi.org/sites/default/files/reference_documents/naspi_naruc_silverstein_20170714.pdf

[4] M. Jones, "Spoofing in the black sea: What really happened," *GPS World*, vol. 11, 2017.

[5] J. J. Spilker Jr, "Vector delay lock loop processing of radiolocation transmitter signals," Mar. 14 1995, uS Patent 5,398,034.

[6] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov 2004. [Online]. Available: https://doi.org/10.1023/B:VISI.0000029664.99615.94

[7] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, 1960.

[8] P. Abbeel, A. Coates, M. Montemerlo, A. Y. Ng, and S. Thrun, "Discriminative training of kalman filters," in *Proceedings of Robotics: Science and Systems*, Cambridge, USA, June 2005.

[9] S. Akhlaghi, N. Zhou, and Z. Huang, "Adaptive Adjustment of Noise Covariance in Kalman Filter for Dynamic State Estimation," *ArXiv e-prints*, Feb. 2017.

[10] P. Closas, C. Fernández-Prades, and J. A. Fernández-Rubio, "Maximum likelihood estimation of position in gnss," *IEEE Signal Processing Letters*, vol. 14, no. 5, pp. 359–362, 2007.

[11] R. DiEsposti, "Gps prn code signal processing and receiver design for simultaneous all-in-view coherent signal acquisition and navigation solution determination," in *Proceedings of the 2007 National Technical Meeting of The Institute of Navigation*, 2001, pp. 91–103.

[12] P. Axelrad, B. K. Bradley, J. Donna, M. Mitchell, and S. Mohiuddin, "Collective detection and direct positioning using multiple gnss satellites," *Navigation*, vol. 58, no. 4, pp. 305–321, 2011.

[13] A. J. Weiss, "Direct position determination of narrowband radio frequency transmitters," *IEEE signal processing letters*, vol. 11, no. 5, pp. 513–516, 2004.

[14] G. X. Gao, M. Sgammini, M. Lu, and N. Kubo, "Protecting gnss receivers from jamming and interference," *Proceedings of the IEEE*, vol. 104, no. 6, pp. 1327–1338, 2016.

[15] P. Closas and A. Gusi-Amigo, "Direct position estimation of gnss receivers: Analyzing main results, architectures, enhancements, and challenges," *IEEE Signal Processing Magazine*, vol. 34, no. 5, pp. 72–84, 2017.

[16] P. Closas, C. Fernandez-Prades, and J. A. Fernandez-Rubio, "Cramér–rao bound analysis of positioning approaches in gnss receivers," *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3775–3786, 2009.

[17] J. Liu, H. Yin, X. Cui, M. Lu, and Z. Feng, "A direct position tracking loop for gnss receivers," *Proc. 24th ION GNSS, Portland, OR, USA*, pp. 3634–3643, 2011.

[18] J. Liu, M. Lu, Z. Feng, and X. Cui, "Direct position tracking loop based on linearised signal model for global navigation satellite system receivers," *IET Radar, Sonar & Navigation*, vol. 7, no. 7, Aug. 2013.

[19] P. Closas, C. Fernández-Prades, J. Fernández-Rubio et al., "Evaluation of GNSS direct position estimation in realistic multipath channels," in *Proceedings of the 28th International Technical Meeting of The Satellite Division of The Institute of Navigation (ION GNSS+ 2015)*, Tampa, FL, Sep. 2015, pp. 3693–3701.

[20] Y. Ng and G. X. Gao, "Mitigating jamming and meaconing attacks using direct GPS positioning," in *Position, Location and Navigation Symposium (PLANS), 2016 IEEE/ION*. IEEE, 2016, pp. 1021–1026.

[21] J. J. Brewer, "The differential vector phase-locked loop for Global Navigation Satellite System signal tracking," Ph.D. dissertation, Dept. of Elec. Eng., Air Force Institute of Technology, Wright-Patterson AFB, OH, June 2014.

[22] T. Lin, J. T. Curran, C. O'Driscoll, and G. Lachapelle, "Implementation of a navigation domain GNSS signal tracking loop," in *Proceedings of the 24th International Technical Meeting of The Satellite Division of The Institute of Navigation (ION GNSS+ 2011)*, Portland, OR, Sep. 2011.

[23] Y. Ng, "Improving the robustness of GPS direct position estimation," M.S. thesis, Dept. of Aero. Eng., Univ. of Illinois at Urbana-Champaign, Urbana, IL, Dec. 2016. [Online]. Available: http://hdl.handle.net/2142/95250

[24] J. ICAO, "International standards and recommended practices," in *Aeronautical Telecommunications, Annex 10 to the Convention of International Civil Aviation*, vol. 4, 73.

[25] T. Walter, P. Enge, J. Blanch, and B. Pervan, "Worldwide vertical guidance of aircraft based on modernized gps and new integrity augmentations," *Proceedings of the IEEE*, vol. 96, no. 12, pp. 1918–1935, 2008.

[26] J. Blanch, T. Walter, and P. Enge, "Satellite navigation for aviation in 2025," *Proceedings of the IEEE*, vol. 100, no. Special Centennial Issue, pp. 1821–1830, 2012.

[27] J. Rife, S. Pullen, B. Pervan, and P. Enge, "Paired overbounding and application to gps augmentation," in *PLANS 2004. Position Location and Navigation Symposium (IEEE Cat. No. 04CH37556)*. IEEE, 2004, pp. 439–446.

[28] J. Blanch, T. Walter, and P. Enge, "Position error bound calculation for GNSS using measurement residuals," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 3, pp. 977–984, 2008.

[29] P. Closas, A. Gusi-Amigó, and J. Blanch, "Integrity measures in direct-positioning," in *Proc. 30th Int. Technical Meeting of the Satellite Division of the Institute of Navigation*, 2017.

[30] A. H.-P. Chu and G. X. Gao, "Vertical integrity monitoring with direct positioning," in *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. IEEE, 2018, pp. 367–373.

[31] F. F. Wu, "Power system state estimation: a survey," *International Journal of Electrical Power & Energy Systems*, vol. 12, no. 2, pp. 80–87, 1990.

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

[33] F. C. Schweppe, "Power system static-state estimation, Part III: Implementation," *IEEE Transactions on Power Apparatus and systems*, no. 1, pp. 130–135, 1970.

[34] Y.-F. Huang, S. Werner, J. Huang, N. Kashyap, and V. Gupta, "State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 33–43, 2012.

[35] M. Asprou, S. Chakrabarti, and E. Kyriakides, "A two-stage state estimator for dynamic monitoring of power systems," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1767–1776, 2014.

[36] "IEEE/IEC international standard - measuring relays and protection equipment - part 118-1: Synchrophasor for power systems - measurements," *IEC/IEEE 60255-118-1:2018*, pp. 1–78, 2018.

[37] DOE, "Grid Modernization Initiative," (Accessed: 2019, July 10). [Online]. Available: https://www.energy.gov/grid-modernization-initiative

[38] P. Vyskocil and J. Sebesta, "Relative timing characteristics of GPS timing modules for time synchronization application," in *2009 International Workshop on Satellite and Space Communications*. IEEE, 2009, pp. 230–234.

[39] R. T. Ioannides, T. Pany, and G. Gibbons, "Known vulnerabilities of global navigation satellite systems, status, and potential mitigation techniques," *Proceedings of the IEEE*, vol. 104, no. 6, pp. 1174–1194, 2016.

[40] A. G. Dempster and E. Cetin, "Interference localization for satellite navigation systems," *Proceedings of the IEEE*, vol. 104, no. 6, pp. 1318–1326, 2016.

[41] J. J. Spilker Jr, "Vector delay lock loop processing of radiolocation transmitter signals," Mar. 14 1995, uS Patent 5,398,034.

[42] P. Risbud, N. Gatsis, and A. Taha, "Vulnerability analysis of smart grids to gps spoofing," *IEEE Transactions on Smart Grid*, 2018.

[43] T. E. Humphreys, B. M. Ledvina, M. L. Psiaki, B. W. O'Hanlon, and P. M. Kintner, "Assessing the spoofing threat: Development of a portable gps civilian spoofer," in *Radionavigation laboratory conference proceedings*, 2008.

[44] X. Jiang, J. Zhang, B. J. Harding, J. J. Makela, A. D. Domı et al., "Spoofing gps receiver clock offset of phasor measurement units," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3253–3262, 2013.

[45] D. P. Shepard, T. E. Humphreys, and A. A. Fansler, "Evaluation of the vulnerability of phasor measurement units to gps spoofing attacks," *International Journal of Critical Infrastructure Protection*, vol. 5, no. 3-4, pp. 146–153, 2012.

[46] Z. Zhang, S. Gong, A. D. Dimitrovski, and H. Li, "Time synchronization attack in smart grid: Impact and analysis," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 87–98, 2013.

[47] "IEEE standard for synchrophasor measurements for power systems," *IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005)*, pp. 1–61, 2011.

[48] K. Martin, D. Hamai, M. Adamiak, S. Anderson, M. Begovic, G. Benmouyal, G. Brunello, J. Burger, J. Cai, B. Dickerson et al., "Exploring the IEEE standard C37. 118–2005 synchrophasors for power systems," *IEEE transactions on power delivery*, vol. 23, no. 4, pp. 1805–1811, 2008.

[49] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1630–1638, 2016.

[50] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security (TISSEC)*, vol. 14, no. 1, p. 13, 2011.

[51] K. Mahapatra, N. R. Chaudhuri, and R. Kavasseri, "Bad data detection in PMU measurements using principal component analysis," in *2016 North American Power Symposium (NAPS)*. IEEE, 2016, pp. 1–6.

[52] S. Cui, Z. Han, S. Kar, T. T. Kim, H. V. Poor, and A. Tajer, "Coordinated data-injection attack and detection in the smart grid: A detailed look at enriching detection solutions," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 106–115, 2012.

[53] T. Y. Mina, S. Bhamidipati, and G. X. Gao, "Detecting gps spoofing via a multi-receiver hybrid communication network for power grid timing verification," in *31st International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS+ 2018*. Institute of Navigation, 2018, pp. 2963–2977.

[54] S. Bhamidipati and G. X. Gao, "Gps multi-receiver joint direct time estimation and spoofer localization," *IEEE Transactions on Aerospace and Electronic Systems*, 2018.

[55] D.-Y. Yu, A. Ranganathan, T. Locher, S. Capkun, and D. Basin, "Short paper: detection of gps spoofing attacks in power grids," in *Proceedings of the 2014 ACM conference on Security and privacy in wireless & mobile networks*. ACM, 2014, pp. 99–104.

[56] M. L. Psiaki and T. E. Humphreys, "Gnss spoofing and detection," *Proceedings of the IEEE*, vol. 104, no. 6, pp. 1258–1270, 2016.

[57] F. Zhu, A. Youssef, and W. Hamouda, "Detection techniques for data-level spoofing in gps-based phasor measurement units," in *2016 International Conference on Selected Topics in Mobile & Wireless Networking (MoWNeT)*. IEEE, 2016, pp. 1–8.

[58] P. Pradhan, K. Nagananda, P. Venkitasubramaniam, S. Kishore, and R. S. Blum, "Gps spoofing attack characterization and detection in smart grids," in *2016 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 2016, pp. 391–395.

[59] X. Fan, L. Du, and D. Duan, "Synchrophasor data correction under gps spoofing attack: A state estimation-based approach," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4538–4546, 2017.

[60] S. De Silva, T. Hagan, J. Kim, and E. Cotilla-Sanchez, "Sparse error correction for PMU data under GPS spoofing attacks," in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2018, pp. 902–906.

[61] S. D. Silva and T. Hagan, "Towards Attack Resilient Data Analytics for Power Grid Operations," (CREDC AHM Presented: 2019, March 8). [Online]. Available: https://cred-c.org/researchactivity/towards-attack-resilient-data-analytics-power-grid-operations

[62] A. Lemmenes, P. Corbell, and S. Gunawardena, "Detailed analysis of the texbat datasets using a high fidelity software gps receiver," in *Proceedings of the 29th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2016), Portland, OR, USA*, 2016, pp. 12–16.

[63] S. V. S. Chauhan and G. X. Gao, "Joint gps and vision estimation using an adaptive filter," in *30th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS 2017*. Institute of Navigation, 2017, pp. 808–812.

[64] S. V. S. Chauhan and G. X. Gao, "Vertical protection level estimation for direct positioning using a bayesian approach," in *Proceedings of the 32nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2019)*, 2019, pp. 2903–2915.

[65] S. V. S. Chauhan and G. X. Gao, "GPS spoofing-resilient static state estimation for the power grid using PMU measurements," *IEEE Transactions on Smart Grid*, 2020 [Second Revision Submitted].

[66] S. V. S. Chauhan and G. X. Gao, "Hardware-in-the-loop GPS and PMU integrated datasets for the power grid under GPS spoofing attacks," in *Proceedings of the 33rd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2020)*. Institute of Navigation, 2020.

[67] S. V. S. Chauhan and G. X. Gao, "Spoofing resilient state estimation for the power grid using an extended kalman filter," *IEEE Transactions on Smart Grid*, 2020 [Second Revision Submitted].

[68] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Grid structural characteristics as validation criteria for synthetic networks," *IEEE Transactions on power systems*, vol. 32, no. 4, pp. 3258–3265, 2016.

[69] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[70] S. M. Rinaldi, "Modeling and simulating critical infrastructures and their interdependencies," in *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the.* IEEE, 2004, pp. 8–pp.

[71] J. Blanch, T. Walter, and P. Enge, "Position error bound calculation for gnss using measurement residuals," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 3, 2008.

[72] O. Bialer, D. Raphaeli, and A. J. Weiss, "Maximum-likelihood direct position estimation in dense multipath." *IEEE Trans. Vehicular Technology*, vol. 62, no. 5, pp. 2069–2079, 2013.

[73] Y. Ng and G. X. Gao, "Computationally efficient direct position estimation via low duty-cycling," in *Proceedings of the 29th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2016), Portland, OR*, 2016, pp. 86–91.

[74] Y. Ng and G. X. Gao, "Direct position estimation utilizing non-line-of-sight (nlos) gps signals," *Proceedings of ION GNSS+*, 2016.

[75] A. H.-P. Chu, S. V. S. Chauhan, and G. X. Gao, "Gps multireceiver direct position estimation for aerial applications," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 1, pp. 249–262, 2019.

[76] Y. Ng and G. X. Gao, "Advanced multi-receiver vector tracking for positioning a land vehicle," in *Proceedings of the 26th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2014), Tampa, Florida*, 2015.

[77] P. Closas, C. Fernandez-Prades, and J. A. Fernkndez-Rubiot, "Ml estimation of position in a gnss receiver using the sage algorithm," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 3. IEEE, 2007, pp. III–1045.

[78] J. W. Cheong, J. Wu, A. G. Dempster, and C. Rizos, "Assisted-GPS based snap-shot GPS receiver with FFT-accelerated collective detection: Time synchronisation and search space analysis," in *Proceeding of the 2012 International Technical Meeting of the Satellite Division of the Institute of Navigation*, 2012.

[79] F. Nielsen and K. Sun, "Guaranteed bounds on information-theoretic measures of univariate mixtures using piecewise log-sum-exp inequalities," *Entropy*, vol. 18, no. 12, p. 442, 2016.

[80] E. Wycoff and G. X. Gao, "A python software platform for cooperatively tracking multiple gps receivers," *Proceedings of ION GNSS, Tampa, Florida, USA*, pp. 8–12, 2014.

[81] A. G. Phadke, J. S. Thorp, and K. Karimi, "State estimlatjon with phasor measurements," *IEEE Transactions on Power Systems*, vol. 1, no. 1, pp. 233–238, 1986.

[82] M. Zhou, V. A. Centeno, J. S. Thorp, and A. G. Phadke, "An alternative for including phasor measurements in state estimators," *IEEE transactions on power systems*, vol. 21, no. 4, pp. 1930–1937, 2006.

[83] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on power systems*, vol. 26, no. 1, pp. 12–19, 2010.

[84] V. Kekatos, G. B. Giannakis, and B. Wollenberg, "Optimal placement of phasor measurement units via convex relaxation," *IEEE Transactions on power systems*, vol. 27, no. 3, pp. 1521–1530, 2012.

[85] C. Yuan, Y. Zhou, G. Zhang, G. Liu, R. Dai, X. Chen, and Z. Wang, "Exploration of graph computing in power system state estimation," in *2018 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2018, pp. 1–5.

[86] T. E. Humphreys, J. A. Bhatti, D. Shepard, and K. Wesson, "The Texas spoofing test battery: Toward a standard for evaluating GPS signal authentication techniques," in *Radionavigation Laboratory Conference Proceedings*, 2012.

[87] T. Ebinuma, "GPS-SDR-SIM," (Accessed: 2019, October 1). [Online]. Available: https://github.com/osqzss/gps-sdr-sim

[88] S. Bhamidipati, "Attack resilient gps based timing for phasor measurement units using multi-receiver direct time estimation," M.S. thesis, Dept. of Aero. Eng., Univ. of Illinois at Urbana-Champaign, Urbana, IL, 2017. [Online]. Available: http://hdl.handle.net/2142/97515

[89] "Open Source Phasor Data Concentrator." [Online]. Available: https://github.com/GridProtectionAlliance/openPDC

[90] J. Zhao, A. Gómez-Expósito, M. Netto, L. Mili, A. Abur, V. Terzija, I. Kamwa, B. Pal, A. K. Singh, J. Qi et al., "Power system dynamic state estimation: Motivations, definitions, methodologies, and future work," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 3188–3198, 2019.

[91] W. Li, "PMU-based State Estimation for Hybrid AC and DC grids," Ph.D. dissertation, KTH Royal Institute of Technology, 2018.